18 MAT333 : GRAPH ANALYTICS & ALGORITHMS

CASE STUDY : FLIGHT ROUTES AND AIRPORTS NETWORK IN INDIA

Nandana N.S *CB.SC.I5DAS18023*

Uthara Koliyot *CB.SC.I5DAS18044*

Integrated Msc.Data Science, Semester 6

Amrita Vishwa Vidyapeetham

Abstract

The world's eminent airports are directly or indirectly connected to many other airports. Every airport is considered as a node and the route can be considered as edge connecting them. The work analyzes the Indian airport network using different centrality measures of network analysis. The centrality measures calculated on airport network help in identification of certain characteristics of the airports. Some of the characteristics are like the busiest airport and the airports which influence trade, alternate path, fastest route, nearest airports, etc. The characteristics helps to find the designated airports meant for improving the economy. The results of this paper say about the prominent communication and connections among the airports in India.

Table of Contents

Introduction

Airport network has an important role in improving the economy of the country. The analysis of such network results in designating an airport for better economy. There are various tools and ways to analyze the airport network like the other networks. In this project we try to analyze the Indian Airport Network using networkx. The project has implemented in Python Programming Language and conducted in JUPYTER platform.

AIM : To analyse the airport routes and flight networks in India.

Source of the Dataset : OpenFlights website
https://openflights.org/data.html

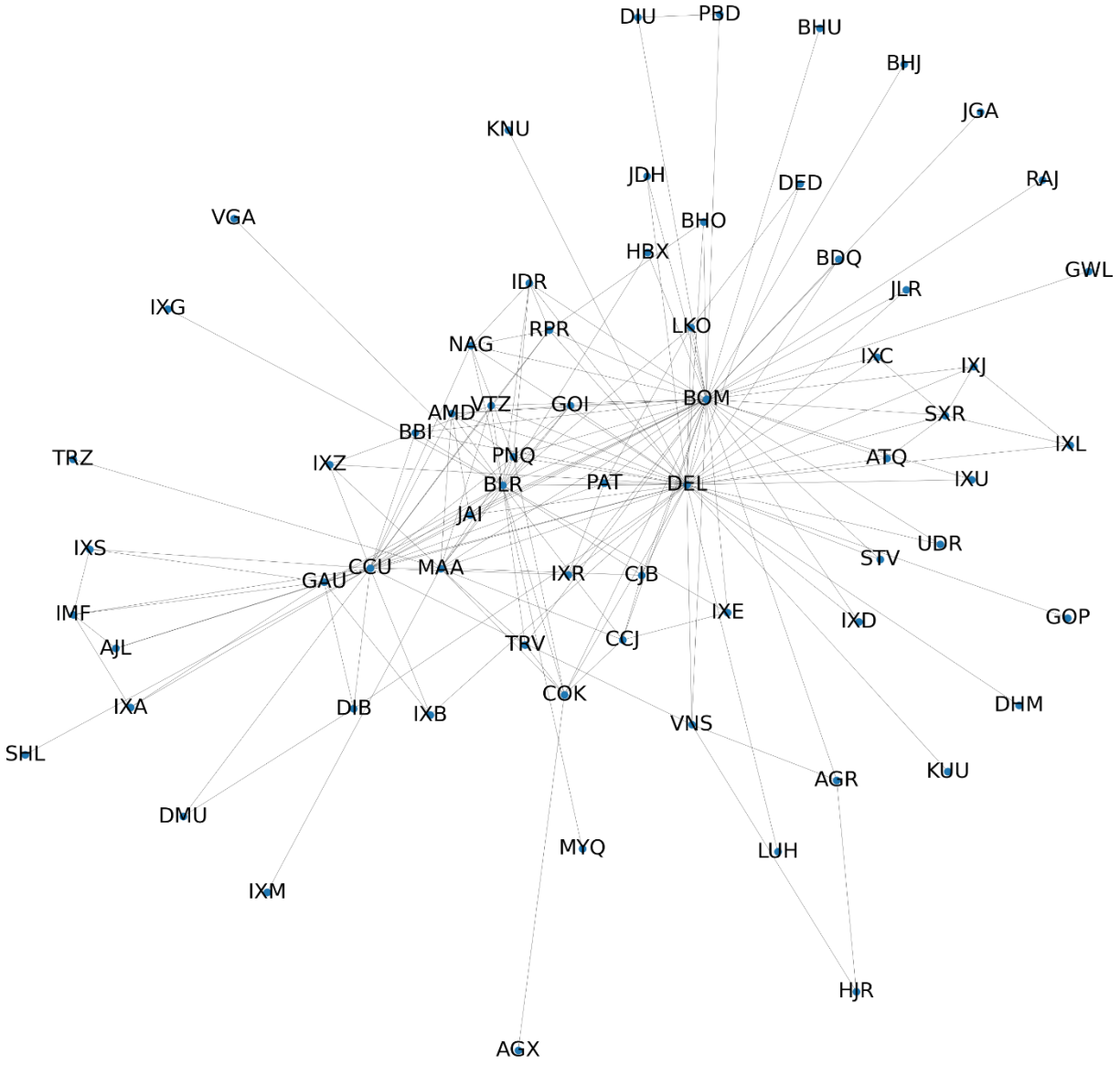
About the dataset :

The original dataset containing many flight routes and airports around the world .The DataFrame index is the IATA code, a 3-characters code identifying the airports.

Our aim is to identify the airport routes in India ,hence we select only airports in India and routes connecting to them.

The final dataset consists of 68 airports and 179 routes connecting them.

The network looks like :



Results and analysis

Degree Centrality

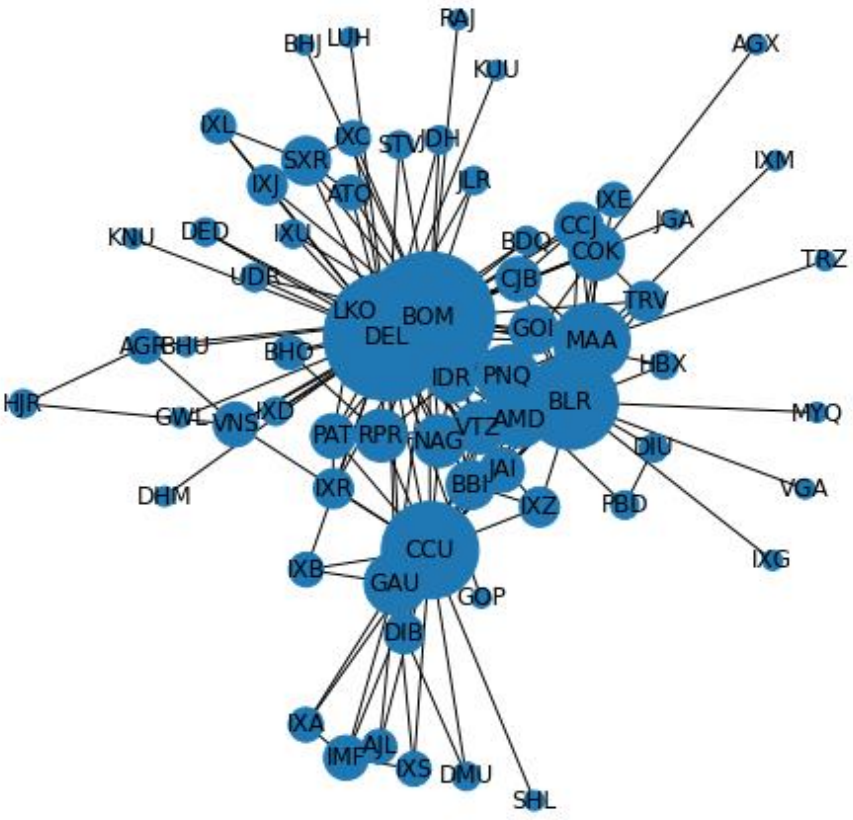
The degree centrality for a node v is the number of nodes it is connected to.

Degree centrality of this dataset shows the number of direct flights boarded or arrived from or to a particular airport. Here Bombay has the highest degree centrality of 44.

Which means total of 44 flights are boarded or departure from or arrived from Bombay Airport.

Top 10 nodes with Maximum Degree Centrality are:

BOM, DEL, CCU, BLR, MAA, PNQ, GAU, AMD, COK & VTZ.



Closeness centrality of the network:

Closeness centrality of a node is a measure of centrality in a network, calculated as the sum of the length of the shortest paths between the node and all other nodes in the graph. Thus the more central a node is, the closer it is to all other nodes.

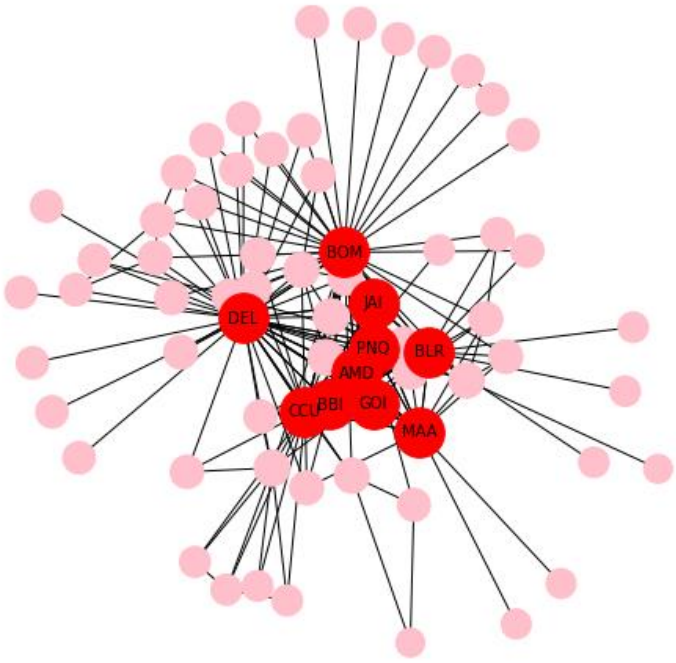
Formula :

$$CC(i) = \frac{N-1}{\sum_j d(i,j)}$$

From the perspective of the Indian air transport dataset, closeness centrality measures airport which is nearer to all other airports so as to measure how fast the flight can reach the destination and find which airport is closer to many airports and the flights which departures from this airport takes shorter time to any other airports.

Here Ahmedabad has the highest closeness centrality of 0.523438.

The 10 airports with maximum closeness centrality are : 'AMD', 'BLR', 'BOM', 'CCU', 'DEL', 'GOI', 'JAI', 'MAA', 'PNQ' & 'BBI'



Betweenness Centality

The betweenness of a vertex measures how much flow will pass through that particular vertex.

Formula :

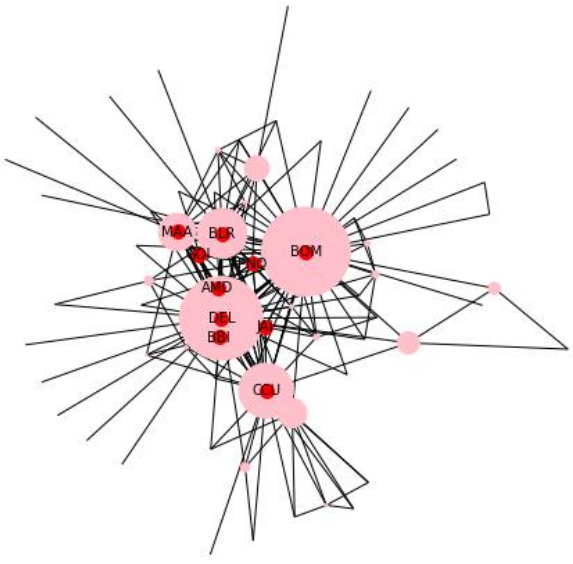
$$c_B(v) = \sum_{s,t \in V} \frac{\sigma(s,t|v)}{\sigma(s,t)}$$

Betweenness centrality which shows which airport has the highest influence within that network. The change or disorder of this airport leads to disrupt communication between other airports, which are considered in this dataset. It acts as an intermediate airport for many airports in India i.e., without these airports the flights which have connection to it leads to take long path to reach the destination or it won't have any path to the destination. Therefore, this airport plays a vital role to reach the destination airport with shortest path.

Here Ahmedabad has the highest betweenness centrality of 0.523438.

The top 10 airports with maximum betweenness centrality are :

'AMD', 'BLR', 'BOM', 'CCU', 'DEL', 'GOI', 'JAI', 'MAA', 'PNQ' & 'BBI'.



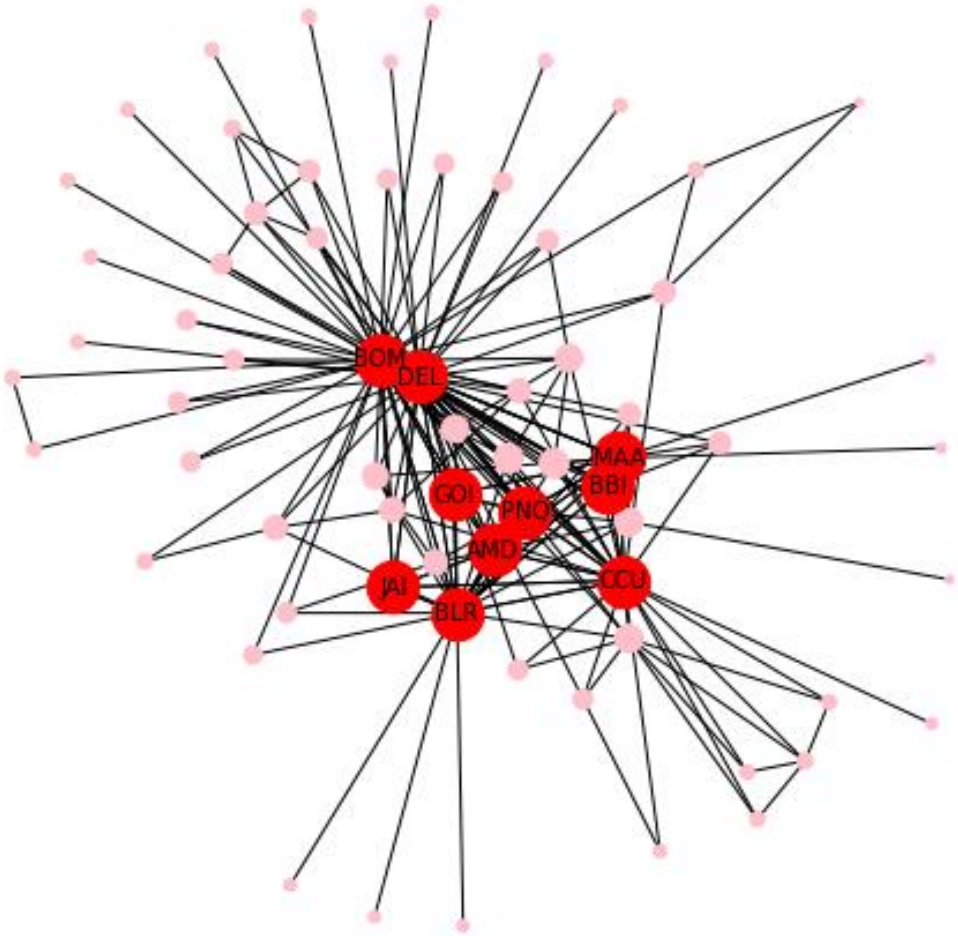
Eigen Vector Centrality

Eigenvector centrality is a measure of the influence of a node in a network. It indicates which of the nodes have influenced the maximum other nodes. The eigenvector centrality network metric takes into consideration not only how many connections a vertex has, but also the degree of the vertices that it is connected to.

The most influential node in our network is Ahmedabad with an eigen vector centralit y of 0.17533994885647974.

The top 10 airports with maximum betweenness centrality are :

'AMD', 'BLR', 'BOM', 'CCU', 'DEL', 'GOI', 'JAI', 'MAA', 'PNQ' & 'BBI'



Clustering Coefficient

Clustering coefficient is the overall probability for the network to have adjacent nodes interconnected, thus revealing the existence of tightly connected communities.

Formula

$$c_u = \frac{2T(u)}{deg(u)(deg(u)-1)},$$

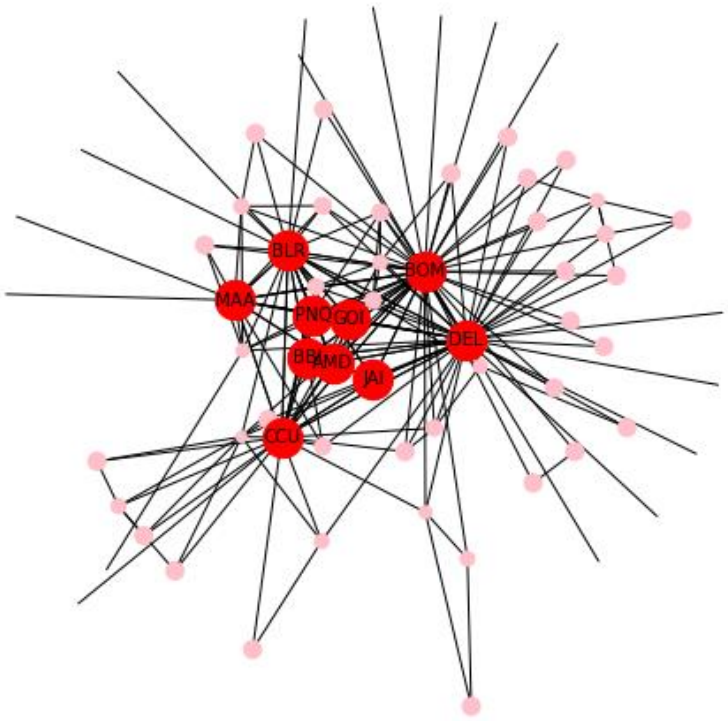
where $T(u)$ is the number of triangles through node u and $deg(u)$ is the degree of u.

Here Ahmedabad has the highest clustering coefficient of 0.09830866807 610994.

This means the flight network of Ahmedabad is the densest compared to o ther cities due to the fact that most of its nodes have smaller values of clus tering coefficient.

The top 10 airports with maximum clustering coefficient centrality are :

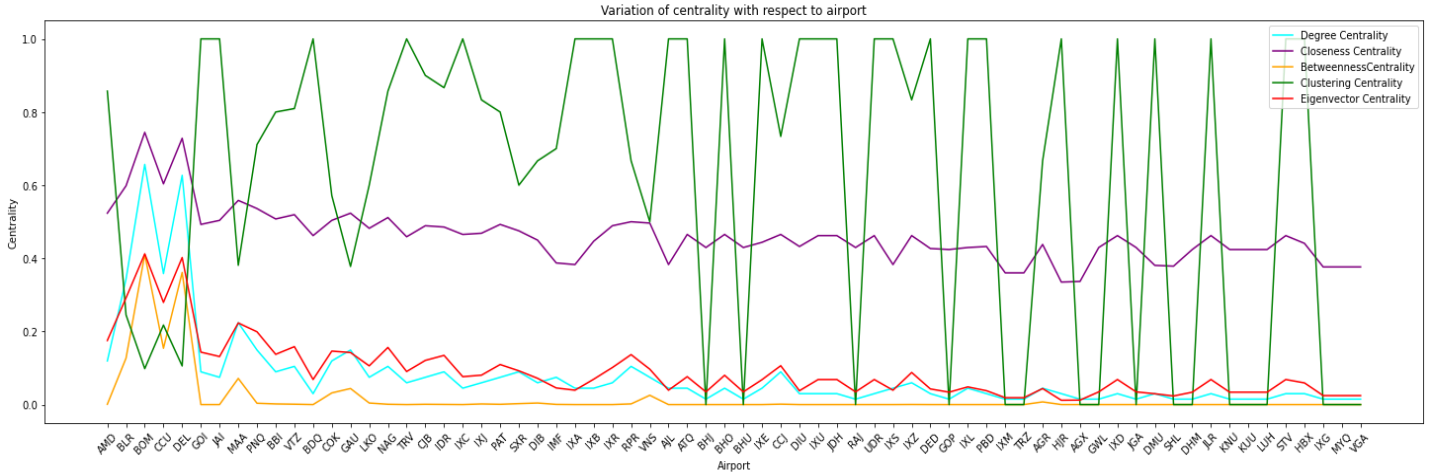
'AMD', 'BLR', 'BOM', 'CCU', 'DEL', 'GOI', 'JAI', 'MAA', 'PNQ' & 'BBI'



Average Clustering Coefficient of the Network :

The average clustering coefficient of the network is 0.6087970051414424. When these connections are dense, the clustering coefficient is high. So, we can say that our network forms a moderately dense cluster.

Variation of different centralities with respect to Airports :



Here we can see that major Airports like Ahmedabad, Bombay Bangalore and Kolkata CCU Airports show high peaks in all the centralities signifying their importance in the Airport Network of India.

Conclusion

The analysis of Indian airport network using centrality measures like degree, betweenness, closeness, and clustering coefficients identifies the important airports and its characteristics that help in revamping the economy of the country. The analysis of the Indian airport networks shows that Ahmedabad, Bangalore and Bombay are the designated cities in India and which acts as a connection airport for most of the airports within India and help in boosting the trade within and outside of the country. So, Ahmedabad, Bangalore and Bombay are some of the main cities that promotes the economic growth of the country. Analysis of such datasets on different centrality measures helps in identifying the importance of the airports in various aspects.

References:

1. *https://www.sciencedirect.com/science/article/pii/S209252121004 1X*
2. *https://www.researchgate.net/publication/326148275_Analysis_an d_visualization_of_airport_network_to_strengthen_the_economy*
3. *https://towardsdatascience.com/applied-network-analysis-using-python-25021633a702*
4. *https://www.analyticsvidhya.com/blog/2018/04/introduction-to-graph-theory-network-analysis-python-codes/*