|  |  |
| --- | --- |
|  |  |

Graph Analysis of Airline Networks

# Data Analytics and Visualization

#### Student\_ID: S398713

#### Department: Applied Artificial Intelligence

#### Date: 4/1/2023

|  |  |  |
| --- | --- | --- |
|  |  | Graph Analysis of Airline Transport Networks |

# Assessment Questions

1. **Construct separate airline networks for each country (internal flights only) and visualise them using geospatial coordinates provided. Comment how it is done**

The following steps were followed to construct internal flight network for the given countries:-

Firstly, Data was provided in the form of two csv files ‘*Airports.csv’ and ‘Flight\_Data.xlsx’.*

The first one consisted of data regarding Airport locations along with their IATA identifier and

Latitudeand longitude coordinates. The second file consisted of flight data between various

Airports along with the weight which states how frequent the flights are along with the IATA

Identifiers to correlate with the ‘*Airports.csv’.*

The next step was to merge both the Airports and Flight data. After merging the two csv files together a final data frame was created which consisted of the countries of interest United States of America, United Kingdom, China and Australia. Then separate csv files for the above-mentioned countries were created to construct their internal flight network.

With the help of Networkx and Geopandas in Python internal flight network was created

where airport’s geospatial coordinates act as the nodes whereas the flight data are the

edges connecting the nodes. Labels were added in the airline network because there were

some abnormalities observed in the USA Airline Network graph with some of the

airports(nodes) getting plotted outside the boundaries of USA. By adding labels, I was able

to track the wrong airport locations, and which needed rectification.

After tracking down the airports having wrong geospatial coordinates with the help of Google, exact latitude and longitude coordinates were found and respective changes were made in the csv file. After plotting the points, the labels were dropped as it made the graph look too packed and cluttered (only for US Flight network).

(*Code attached below in appendix)*

**Graphical user interface

Description automatically generated with medium confidence**

**Figure 1.1: Flight Network for United States of America**

Chart, diagram, radar chart

Description automatically generated

**Figure 1.2: Flight Network for United Kingdom**

**Diagram

Description automatically generated**

**Figure 1.3: Flight Network for China**

Chart, radar chart

Description automatically generated

**Figure 1.4: Flight Network for Australia**

1. **Plot the following:**

**2.1: Degree Distribution (x-axis- Descending Rank, y-axis- Logscale of Weighted Degree)**

Chart, line chart

Description automatically generated Chart, scatter chart

Description automatically generated

Chart

Description automatically generated Chart, line chart

Description automatically generated

**2.2 Degree vs. Betweenness Distribution**

Chart, scatter chart

Description automatically generatedChart, scatter chart

Description automatically generated

Chart, scatter chart

Description automatically generatedChart, scatter chart

Description automatically generated

**2.3 Assortativity (Degree Correlation)**

**Chart, scatter chart

Description automatically generatedChart, scatter chart

Description automatically generated**

Chart, scatter chart

Description automatically generatedChart, scatter chart

Description automatically generated

**2.4 Core Periphery**

**Chart

Description automatically generated with medium confidence**

**Chart, line chart

Description automatically generated**

Chart, line chart

Description automatically generated

Graphical user interface

Description automatically generated with medium confidence

1. **Analysis, discussion and conclusions that include selection and application of metrics for analysis in Q1-Q2 previously with performance considerations relevant to airline networks**

Figure 2.1 shows the weighted degree distribution for the four countries USA, UK, China and Australia. A weighted degree distribution graph is a graph that shows the distribution of the weighted degree of vertices in a graph. The weighted degree of a vertex is the sum of the weights of the edges incident to it. Weights indicating the number of passengers boarding the flight from one airport to another. The weighted degree graph for USA shows an exponential graph showcasing King-Pauper effect in the head and tail. It was observed (from Core periphery graph for USA) that all the core nodes are found in the King part of the model whereas the low connected nodes are found in the tail part of the model.

Figure 2.2 shows the degree vs betweenness distribution. A degree vs betweenness graph is a graph that shows the relationship between the degree and betweenness centrality of the nodes in a graph. This type of graph can provide insights into structure of the graph and relative importance of nodes within it. For example, a node with high degree and high betweenness centrality may indicate a hub in the graph while a node with a low degree and low betweenness centrality may indicate a peripheral node. It is observed in China and Australia that there is Hub-Spoke model in existence. We can also see it from figures 1.3,1.4 and 2.4 that there is a hub in existence and the core periphery structure follows a wider bell-shaped structure as compared to other countries.

Figure 2.3 shows the Assortativity or degree correlation of the countries. Assortativity is a measure of the extent to which nodes in a graph tend to connect with other nodes that are similar to them in some way. In other words, nodes with high degree tend to connect with other nodes with high degree and nodes with low degree tend to connect with other nodes with low degree. It can be quantified with the help of an Assortativity coefficient which is a value between -1 and 1. A positive Assortativity indicates that there is a tendency for nodes to connect with other nodes that are similar to them, while a negative Assortativity coefficient indicates that high degree nodes tend to connect with low degree nodes. It is observed that UK shows a positive Assortativity meaning that the nodes are connected with nodes that are similar in nature. Whereas US, Australia and China show a negative Assortativity value indicating that the nodes of higher degree are connected with nodes of lower degree.

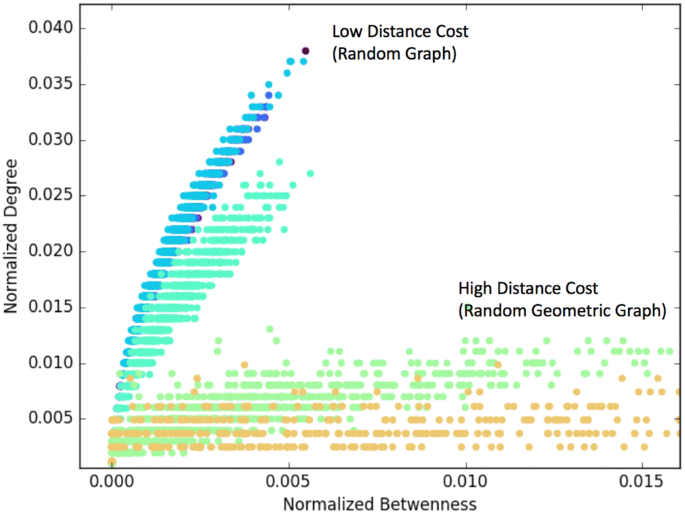
Figure 2.4 indicates Core Periphery structure of the countries. It shows the Core nodes and the Periphery nodes in the graph. It can provide insights into the organization and structure of a network. Core indicates hub of activity or influence while periphery may indicate less active group of nodes. The level of separation can also provide insights into the level of integration and cohesion within the network. A network with a large separation between the core and periphery may be less integrated and cohesive while a network with small separation may be more integrated and cohesive. The lower core nodes show that the network is effective but also shows that the network is less robust under random attacks. (Zhou et al., 2019)

1. **Hypothesize with evidence on the potential impact of results on the design of future aircrafts in these different countries. Hint: how will changing the fuel prize impact the distance penalty and the formation of random graphs that mimic the airline network?**

It is possible that changes in fuel price can have significant impacts on the design of future aircraft in different countries.

One potential impact of changing fuel prices is on the distance penalty, which is a measure of the cost associated with flying a certain distance. As fuel prices rise, the distance penalty is likely to increase, which could lead to the design of aircraft that are more fuel efficient in order to reduce operating costs. This could involve the development of new technologies such as more efficient engines or lighter materials, or the adoption of existing technologies such as winglets or blended wing bodies.

Another potential impact of changing fuel prices is on the formation of random graphs that mimic the airline network.



In order to achieve this, we refer a 2-D random geometric graph (RGG) using a Poisson Point Process (random uniform)(Guo et al., 2019), whereby the probability of connection is given by Waxman(Waxman, 1988)

p(dij ) = β e−dij /Lα

Where dij is the distance between two nodes and L is the maximum distance in the graph

between two nodes. and are parameters in the range (0, 1] and in this case set to =

1 and = 1/α2, where α is the penalty due to price impact.

The spatial graph displays a weak to non-existent link between degree and betweenness for a high value of α (i.e., 2-3). This shows that flying point-to-point is preferable to not flying at all, and as a result, prominent transfer hubs are not well connected (high degree) (high betweenness). The non-spatial graph displays a significant link between degree and betweenness for a low-medium value of α (i.e., 0-2). Accordingly, the hub airports are the finest airports for short-haul transfers. As a result, we can come to the following conclusion. As fuel prices rise, it may become more costly for airlines to operate long-haul flights, which could lead to a shift towards shorter routes or the development of alternative modes of transportation. This could result in changes to the overall structure of the airline network such as increase in number of hub-and-spoke systems or shift towards more direct flights between major cities.(Guo et al., 2019)

On the other hand, the aircraft industry is making promising strides towards being less fuel dependant and shifting its focus on developing electric and hydrogen powered aircrafts. This will help in reduction of flight tickets in the future.(Mukhopadhaya & Rutherford, 2035)

# References

Guo, W., Toader, B., Feier, R., Mosquera, G., Ying, F., Oh, S. W., Price-Williams, M., & Krupp, A. (2019). Global air transport complex network: multi-scale analysis. *SN Applied Sciences*, *1*(7). <https://doi.org/10.1007/s42452-019-0702-2>

Mukhopadhaya, J., & Rutherford, D. (2035). *PERFORMANCE ANALYSIS OF EVOLUTIONARY HYDROGEN-POWERED AIRCRAFT*. [www.theicct.orgcommunications@theicct.org](http://www.theicct.orgcommunications@theicct.org)

Waxman, B. M. (1988). Routing of Multipoint Connections. *IEEE Journal on Selected Areas in Communications*, *6*(9), 1617–1622. <https://doi.org/10.1109/49.12889>

Zhou, Y., Wang, J., & Huang, G. Q. (2019). Efficiency and robustness of weighted air transport networks. *Transportation Research Part E: Logistics and Transportation Review*, *122*, 14–26. <https://doi.org/10.1016/J.TRE.2018.11.008>

Python Script

(Appendix)

Graphical user interface, text, application, email

Description automatically generated

Graphical user interface, text, application

Description automatically generated

Graphical user interface, application

Description automatically generated



Table

Description automatically generated

Text

Description automatically generated

Chart

Description automatically generated with medium confidence

Graphical user interface

Description automatically generated

Application

Description automatically generated with medium confidence

Graphical user interface, application

Description automatically generated

Text, application

Description automatically generated

Graphical user interface, text, application

Description automatically generated

Graphical user interface

Description automatically generated

Graphical user interface, text, application, Teams

Description automatically generated

Graphical user interface, text, application, Teams

Description automatically generated

Chart

Description automatically generated

Text

Description automatically generated

Graphical user interface

Description automatically generated

Graphical user interface, chart

Description automatically generated with medium confidence

A picture containing graphical user interface

Description automatically generated

Graphical user interface

Description automatically generated with medium confidence

Graphical user interface

Description automatically generated with medium confidence

Graphical user interface

Description automatically generated



Chart, line chart

Description automatically generated

A picture containing chart

Description automatically generated

Graphical user interface

Description automatically generated with medium confidence

Graphical user interface, application

Description automatically generated

Graphical user interface

Description automatically generated with medium confidence

Graphical user interface, application

Description automatically generated

Text

Description automatically generated

Chart, line chart

Description automatically generated

Graphical user interface, application

Description automatically generated

Background pattern

Description automatically generated with low confidence

Graphical user interface, application, Word

Description automatically generated

Graphical user interface, application

Description automatically generated

Graphical user interface, application

Description automatically generated

Text

Description automatically generated

Graphical user interface

Description automatically generated