Multi-Scale Network Analysis on Passenger Flight Data

Network Analysis of USA, UK, China, and Australia

s437451 – Julian Gerald Dcruz

Department: Applied Artificial Intelligence

Date: 30/12/23

Information categorisation:
Open

Contents

1. INTRODUCTION 4

1.2 DATASET OVERVIEW 4

2. DATA PREPROCESSING 5

2.1 PRELIMINARY ANALYSIS 5

2.2 DATA PREPARATION 5

3. METHODOLOGY 6

4. MULTI-SCALE NETWORK ANALYSIS FOR USA 8

4.1 NETWORK VISUALIZATION OF DOMESTIC FLIGHT NETWORK IN THE USA 8

4.2 MACRO-SCALE ANALYSIS OF DOMESTIC FLIGHT NETWORK IN THE USA 8

4.3 MICRO-SCALE ANALYSIS OF DOMESTIC FLIGHT NETWORK IN THE USA 9

4.4 MESO-SCALE ANALYSIS OF DOMESTIC FLIGHT NETWORK IN THE USA 11

4.5 ASSORTATIVITY ANALYSIS OF DOMESTIC FLIGHT NETWORK IN THE USA 11

5. MULTI-SCALE NETWORK ANALYSIS FOR UK 12

5.1 NETWORK VISUALIZATION OF DOMESTIC FLIGHT NETWORK IN THE UK 12

5.2 MACRO-SCALE ANALYSIS OF DOMESTIC FLIGHT NETWORK IN THE UK 12

5.3 MICRO-SCALE ANALYSIS OF DOMESTIC FLIGHT NETWORK IN THE UK 13

5.4 MESO-SCALE ANALYSIS OF DOMESTIC FLIGHT NETWORK IN THE UK 14

5.5 ASSORTATIVITY ANALYSIS OF DOMESTIC FLIGHT NETWORK IN THE UK 15

6. MULTI-SCALE NETWORK ANALYSIS FOR CHINA 16

6.1 NETWORK VISUALIZATION OF DOMESTIC FLIGHT NETWORK IN THE CHINA 16

6.2 MACRO-SCALE ANALYSIS OF DOMESTIC FLIGHT NETWORK IN THE CHINA 16

6.3 MICRO-SCALE ANALYSIS OF DOMESTIC FLIGHT NETWORK IN THE CHINA 17

6.4 MESO-SCALE ANALYSIS OF DOMESTIC FLIGHT NETWORK IN THE CHINA 18

6.5 ASSORTATIVITY ANALYSIS OF DOMESTIC FLIGHT NETWORK IN THE CHINA 19

7. MULTI-SCALE NETWORK ANALYSIS FOR AUSTRALIA 20

7.1 NETWORK VISUALIZATION OF DOMESTIC FLIGHT NETWORK IN THE AUSTRALIA 20

7.2 MACRO-SCALE ANALYSIS OF DOMESTIC FLIGHT NETWORK IN THE AUSTRALIA 20

7.3 MICRO-SCALE ANALYSIS OF DOMESTIC FLIGHT NETWORK IN THE AUSTRALIA 21

7.4 MESO-SCALE ANALYSIS OF DOMESTIC FLIGHT NETWORK IN THE AUSTRALIA 22

7.5 ASSORTATIVITY ANALYSIS OF DOMESTIC FLIGHT NETWORK IN THE AUSTRALIA 23

8. ANALYSIS AND DISCUSSIONS 23

9. ADAPTING AIRCRAFT DESIGN: ANALYSING THE INTERPLAY BETWEEN FUEL ECONOMICS AND NETWORK CONFIGURATIONS 25

9.1 RANDOM SPATIAL GRAPH ANALYSIS SIMULATING IMPACT OF FUEL PRICES ON CONNECTIVITY FOR USA 25

9.2 RANDOM SPATIAL GRAPH ANALYSIS SIMULATING IMPACT OF FUEL PRICES ON CONNECTIVITY FOR UK 26

9.3 RANDOM SPATIAL GRAPH ANALYSIS SIMULATING IMPACT OF FUEL PRICES ON CONNECTIVITY FOR CHINA 27

9.4 RANDOM SPATIAL GRAPH ANALYSIS SIMULATING IMPACT OF FUEL PRICES ON CONNECTIVITY FOR AUSTRALIA 28

9.5 DISCUSSIONS 29

10. CONCLUSION 30

11. REFERENCES 30

Multi-Scale Network Analysis on Passenger Flight Data

Network Analysis of USA, UK, China, and Australia

1. Introduction

The aviation industry is increasingly leveraging big data analytics to enhance operational efficiency and customer satisfaction. Flight data analysis, using predictive analytics, is crucial in managing delays, disruptions, and improving safety. This approach enabled airlines to anticipate and mitigate potential issues by analysing large volumes of data from various sources, such as booking systems and customer feedback.

Multi-scale network analysis is vital in the aviation industry as it offers a comprehensive view of network performance, resilience, and recovery, especially in response to unforeseen events like epidemics or policy changes. This type of analysis helps in understanding the structure and dynamic of complex transportation systems, enabling better management of disruptions and improving overall network robustness.

Studies have shown that analysing complex networks like aviation networks through this method can reveal insights into the impact of unexpected events on network performance. It allows for the assessment of network resilience and recovery, providing a clearer understanding of how various factors influence the network's ability to withstand and recover from disturbances. Such analysis is crucial in policy-making and strategic planning, helping to improve safety, efficiency, and customer satisfaction in air transport.

1.2 Dataset Overview

For this study, the 'Airports.csv' and 'Flights Data.xlsx' files are used for network analysis. The 'Flight Data.xlsx' file contains domestic flight data across different countries in the world, it has the following columns:

- *"Source"*: which represents the source airport id in IATA format.
- *"Source City"*: which represents the city where the source airport is located.
- *"Source Country"*: which represents the country where the source airport is located.
- *"Target"*: which represents the destination airport id in IATA format.
- *"Target City"*: which represents the city where the destination airport is located.
- *"Target Country"*: which represents the country where the destination airport is located.
- *"Weight"*: which represents the passenger volume for each flight.

The ‘Airports.csv’ file contains information about the location of the airports, it has the following columns:

- *“id”*: which represents the airport id in IATA format.
- *“label”*: which represents the city where the airport is located.
- *“country”*: which represents the country where the airport is located.
- *“Lat”*: which represents the Latitude coordinates for the airport.
- *“Lon”*: which represents the Longitude coordinates for the airport.

2. Data Preprocessing

2.1 Preliminary Analysis

A preliminary data analysis was done to analyse the data files, the analysis concluded the following:

- Found a corrupt file entry in ‘Airports.csv’ that was not in ‘utf-8’ format, this had to be corrected. It was for the Simón Bolívar International in Venezuela.
- The ‘Airports.csv’ contained zero records with NaN values for the Longitude and Latitude columns.
- The ‘Flights Data.xlsx’ contained zero records with Nan or 0 values for the passenger volume, ‘Weight’ column.
- Manual analysis of the data files found that some records in ‘Airports.csv’ and ‘Flight Data.xlsx’ had different naming conventions for the same country, like USA was also seen as United States, and UK was also seen as United Kingdom. To map this a dictionary key was used as referenced and the ‘Source Country’, ‘Target Country’ and ‘country’ columns in the data files were updated with the same naming convention for the country, this allows for easy filtering.

2.2 Data Preparation

For this study domestic flight data for USA, UK, China, and Australia is considered for network analysis. Using the ‘Source’ and ‘Target’ columns in ‘Flight Data.xlsx’ with the ‘id’ column from ‘Airports.csv’ a merged dataset with the flight routes now having the geographic coordinates of latitude and longitude for their source and destination airports was created.

Analysis of this merged data revealed 49 records with missing geographic latitude and longitude information. To fix the missing and incorrect latitude and longitude information in the merged data, the IATA/ICAO List data available from IP2Location was used along with data from other online sources (1).

Also 9620 duplicate flight routes were found, this usually signifies a bi-directed connection, however the weighting of these duplicate routes need to be updated with the sum of 2 weights and then treated as a single record (2)

The corrected data was then stored to ‘cleaned_routes.csv’ for further analysis. This csv contains the domestic flight routes data for the countries in the focus of this study, along with the geographic latitude and longitude information for each of the source and destination airports. This can now be used for network analysis.

3. Methodology

In this research, the application of a suite of Python libraries facilitated a nuanced examination of domestic air transportation networks in the USA, UK, China, and Australia. The study employed Matplotlib for data visualization, NetworkX(3) for network graph construction, Pandas for data processing, and MPL_toolkits Basemap for geographic mapping. These tools collectively enabled a multi-scale network analysis, encompassing macro-scale, meso-scale, and micro-scale assessments, along with an investigation into assortativity.

At the macro-scale, a statistical analysis, focusing on overarching patterns and trends within the data is visualized, particularly the degree distribution. This level of analysis is crucial for airline operators as it provides a broad overview of the air traffic network. Degree distribution is a fundamental metric in network science, representing the frequency of nodes having a certain number of connections (or 'degrees'). A skewed distribution might suggest a network dominated by a few highly connected hubs, while a more uniform distribution could indicate a decentralized network. The utilization of a logarithmic scale against the descending rank of each location based on these degrees enabled a clearer visualization of the connectivity disparities among different airports, crucial in understanding the network's structure. For airline operators, this degree distribution analysis is of paramount importance. It identifies the most influential nodes - major hubs that play a vital role in maintaining efficient operations. Recognizing these key locations assists in optimizing route planning, improving connectivity, and identifying potential network bottlenecks. Furthermore, this insight is instrumental in strategic decision-making, such as resource allocation, service expansion, and rerouting strategies during disruptions. Overall, this approach enhances the understanding of the hierarchical structure within the air transportation networks, providing valuable information for effective traffic flow and network management.

The meso-scale analysis centred around core community analysis, diverging from traditional modularity-based approaches. This analysis is critical in network science for identifying closely-knit groups or 'communities' within a larger network, which in this context, refers to groups of airports that have more frequent flights amongst themselves than with other parts of the network. The methodology involved creating network graphs from the air transportation data and applying the k-core decomposition technique. The k-core of a network is a subset of nodes that are interconnected with at least 'k' connections. This core often represents the most robust and connected part of the network. Identifying the main k-core in the airport network revealed the most interconnected airports, which are pivotal in maintaining the cohesiveness and efficiency of the entire network. By analysing the largest core community and determining the 'k' value, the study provided valuable insights into the strength and stability of these core networks. For airline operators, understanding these core communities is invaluable for strategic planning and network optimization. It aids in identifying which airports form the backbone of the air traffic network and how these hubs are interconnected. This knowledge is essential for optimizing flight routes, improving service efficiency, and developing robustness against potential disruptions. Additionally, understanding the structure of these core communities can guide decisions related to expanding services, enhancing connectivity in under-served regions, and improving overall network resilience.

In the micro-scale analysis of the study, the focus shifted to individual nodes (airports) within the air transportation networks, emphasizing their specific characteristics and roles. This granular perspective is pivotal in network science, as it allows for the examination of individual elements within the broader system and their influence on the network's functionality and resilience. Two key metrics were central to this analysis: degree and betweenness centrality. The degree of a node signifies the number of direct connections it has, essential for identifying highly connected airports, which are crucial for network robustness and efficiency. Betweenness centrality measures the extent to which a node lies on the shortest path between other nodes, highlighting airports that serve as critical bridges within the network. The combination of these metrics offers a comprehensive understanding of the influence and importance of individual airports in the network. For airline operators, this detailed analysis at the node level is invaluable. It helps in pinpointing strategic airports based on their connectivity (degree) and their role in facilitating traffic flow across the network (betweenness centrality). Airports with high degrees are central to maintaining connectivity, while those with high betweenness centralities are vital for efficient navigation and flow within the network. This knowledge is crucial for optimizing routing, enhancing operational efficiency, and preparing contingency plans for potential disruptions. Furthermore, the correlation between these metrics was examined to understand the relationship between an airport's connectivity and its strategic position within the network. Such insights are instrumental for airlines in making informed decisions regarding route expansion, hub development, and resource allocation.

Additionally, the research incorporated an analysis of assortativity, a measure of the similarity of connections in the network. In the context of air traffic networks, assortativity can indicate whether large hubs tend to connect with other large hubs or if they predominantly connect to smaller, regional airports. This metric can inform airline operators about the structure of the network and potential biases in connectivity, enabling more informed strategic planning and network development.

The multi-scale approach, from macro to micro, can equip airline operators with critical information for optimizing their operations and understanding the intricate relationships within their networks.

4. Multi-Scale Network Analysis for USA

4.1 Network Visualization of Domestic Flight Network in the USA

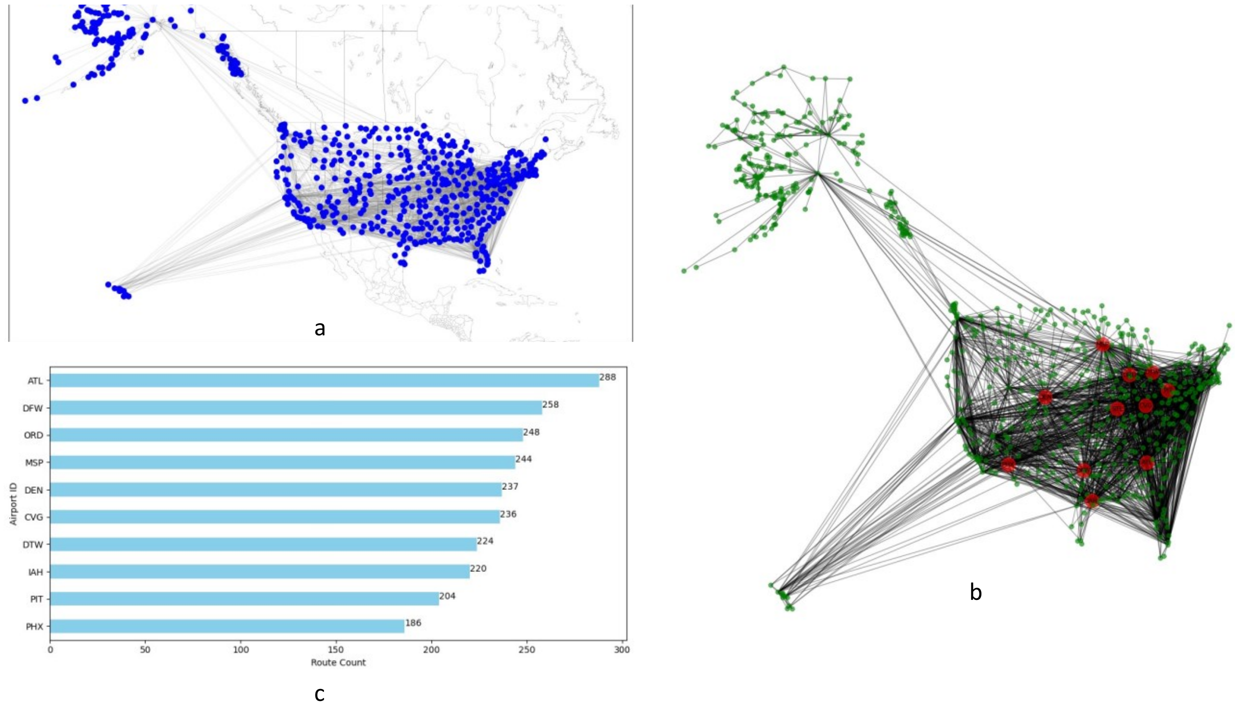


Figure 1: Comprehensive Network Visualization of the United States's Domestic Air Traffic — Part (a) illustrates a geospatial overlay of flight routes upon the UK map, part (b) delineates the intricate network architecture, spotlighting the top 10 airports by route volume in red with corresponding IATA codes, and part (c) provides a horizontal bar chart quantifying the route counts for these leading airports, each labelled with their respective IATA identifier.

4.2 Macro-Scale Analysis of Domestic Flight Network in the USA

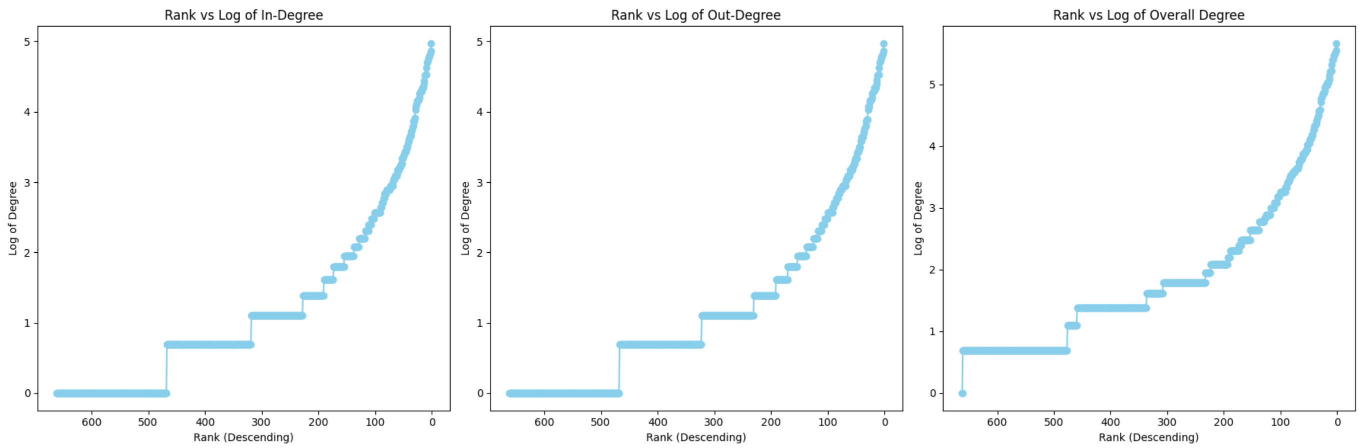


Figure 2: Logarithmic scale of in-degree, out-degree, and overall-degree against their descending rank for each location in USA

Information categorisation:
Open

In Figure 2, The first plot, labelled "Rank vs Log of In-Degree," shows the logarithm of the in-degree (the number of incoming routes) for each airport. The steep curve towards the 'top' end of the plot suggests a small number of airports with significantly higher in-degrees compared to the rest, a common characteristic of scale-free networks often found in air transportation systems.

The second plot, "Rank vs Log of Out-Degree," follows a similar pattern, displaying the logarithm of the out-degree (the number of outgoing routes) against the descending rank of each airport. Again, the plot shows a small number of airports with a high out-degree, decreasing sharply as the rank increases, indicating these few airports dominate outbound flights.

The third plot, "Rank vs Log of Overall Degree," combines both the in-degree and out-degree, giving the total number of connections each airport has. This plot also shows a skewed distribution, with the 'top-ranked' airports having a disproportionately high overall-degree, reinforcing their importance as central hubs in the flight network.

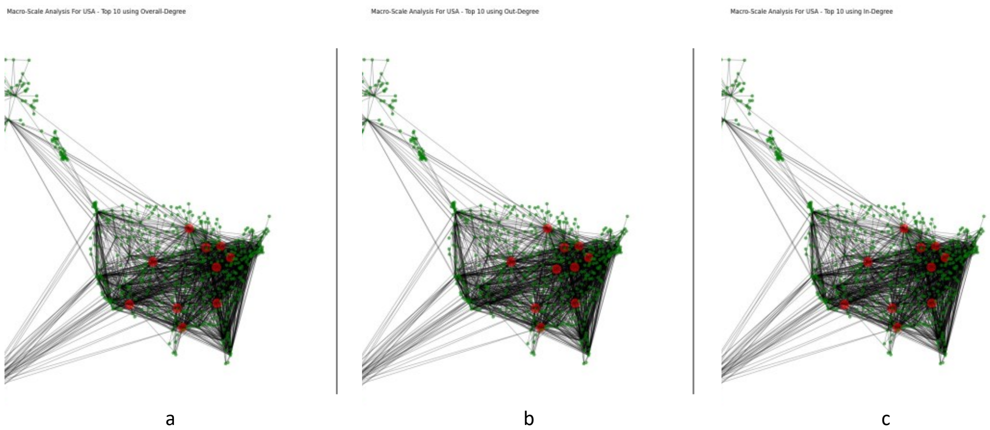


Figure 3: Visualization of Top 10 Airports in the USA based on their (a) overall-degree, (b) out-degree (c) in-degree distributions.

4.3 Micro-Scale Analysis of Domestic Flight Network in the USA

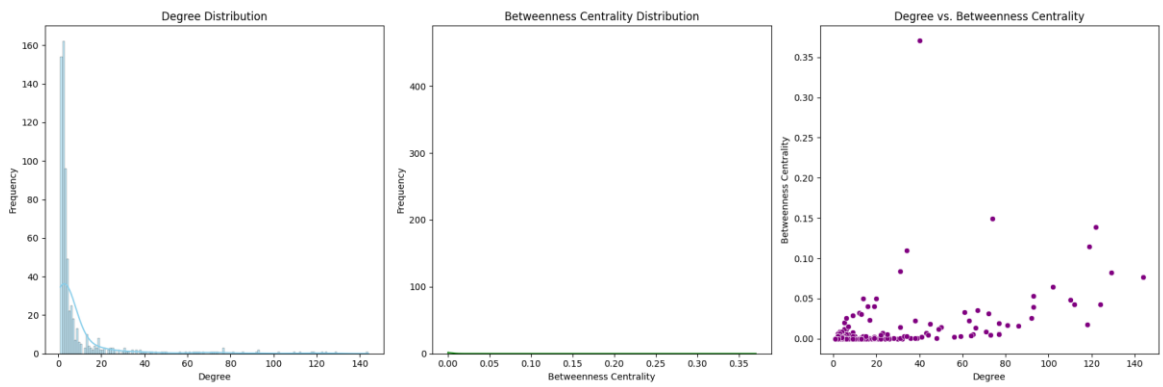
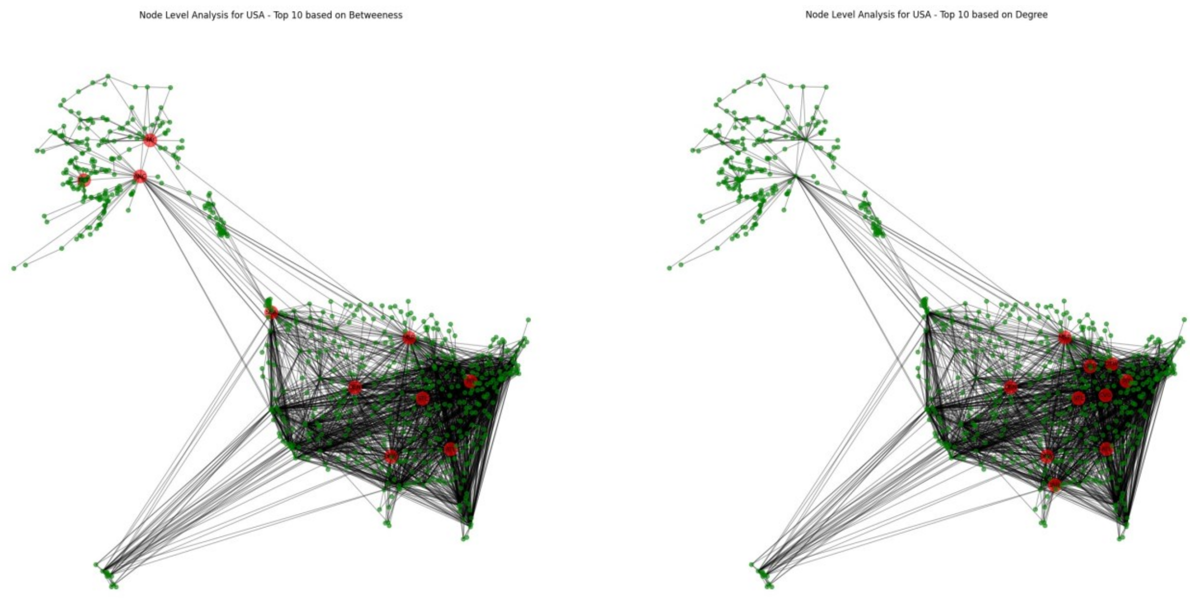


Figure 4: Node Level analysis using Degree Distribution and Betweenness Centrality for USA

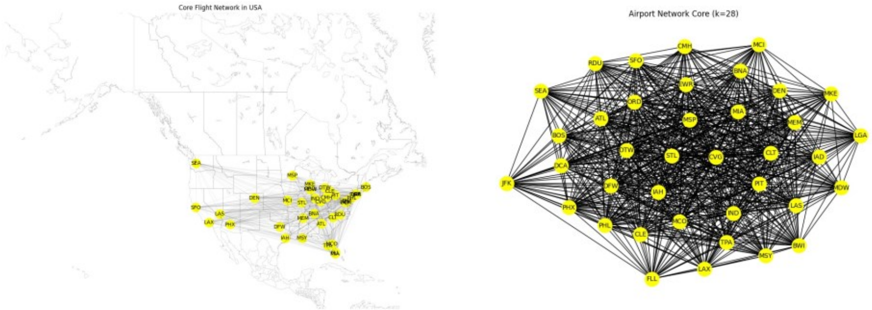


a b

Figure 5: Visualization of Top 10 Airports in the USA based on their (a) Betweenness Centrality and (b) Degree.

Figure 4 presents a node-level analysis of an air transportation network using two central concepts in network science: degree distribution and betweenness centrality. The first plot in Figure 4 depicts the degree distribution of the network's nodes. The plot shows a common characteristic of many real-world networks, where most nodes have a low degree, and a small number of nodes have a high degree, as evidenced by the steep drop-off in the curve. This indicates that the network might be a scale-free network, with a few highly connected nodes (hubs) and many nodes with fewer connections. The middle plot illustrates the betweenness centrality distribution across the network. This plot likely shows that most nodes have a low betweenness centrality, with only a few nodes acting as significant connectors within the network. The third plot in Figure 4 is a scatter plot correlating each node's degree with its betweenness centrality in an air transportation network.

4.4 Meso-Scale Analysis of Domestic Flight Network in the USA



a b

Figure 6: Core Community Analysis of the USA's Domestic Flight Network via K-Core Decomposition, Highlighting a Cohesive Cluster of Airports Each with a Minimum of 28 Connections.

4.5 Assortativity Analysis of Domestic Flight Network in the USA

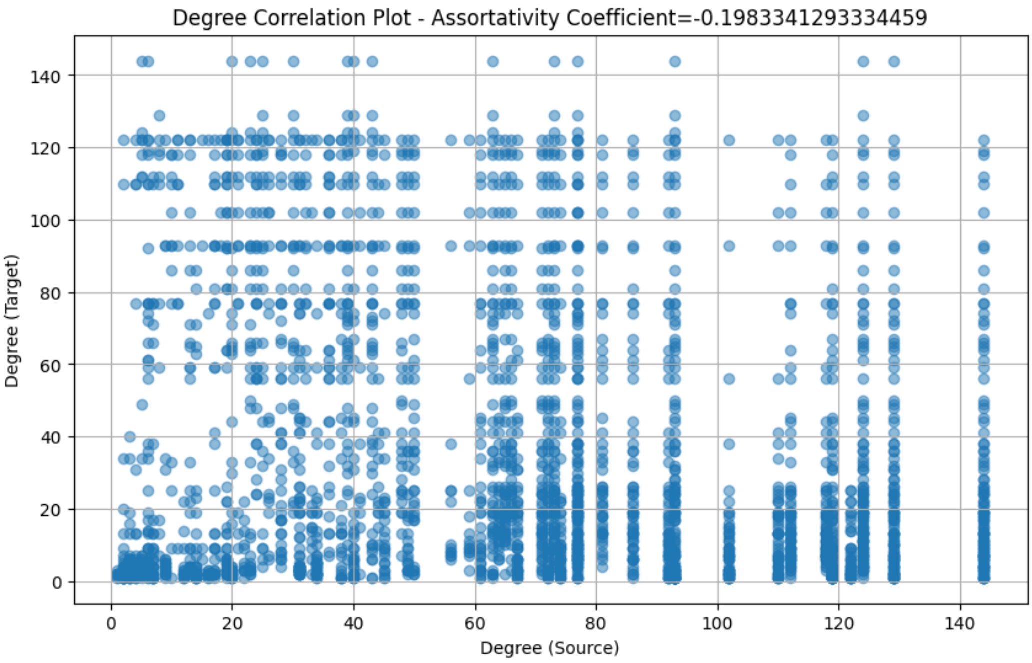
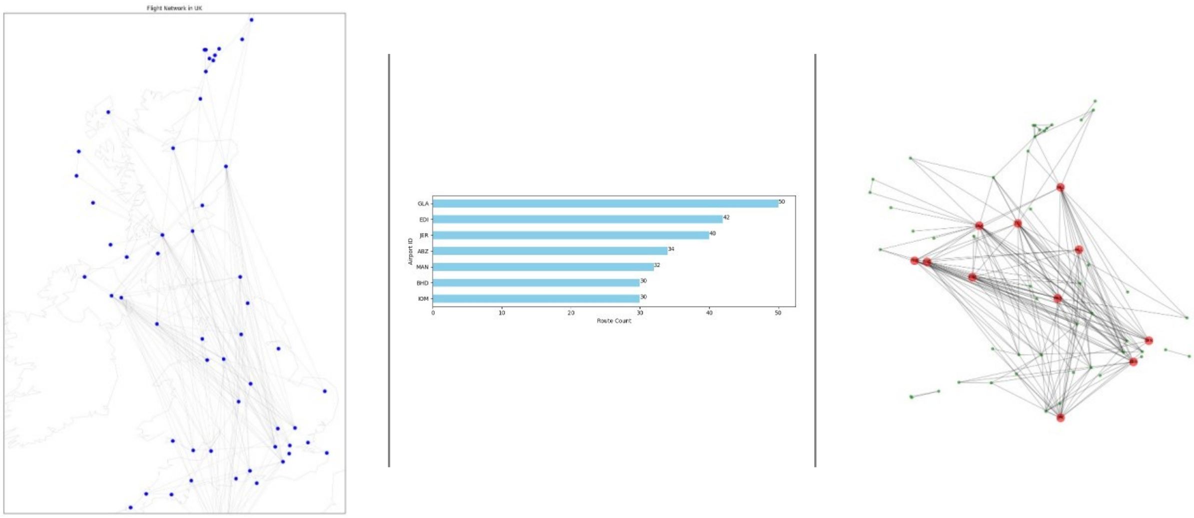


Figure 7: Scatter Plot Illustrating Assortativity in the USA's Domestic Flight Network, revealing a Slight Disassortative Tendency as Evidenced by the Negative Assortativity Coefficient, with High-Degree Airports Preferentially Connecting to Low-Degree Counterparts

5. Multi-Scale Network Analysis for UK

5.1 Network Visualization of Domestic Flight Network in the UK



a b c

Figure 8: Comprehensive Network Visualization of the United Kingdom's Domestic Air Traffic — Part (a) illustrates a geospatial overlay of flight routes upon the UK map, part (b) provides a horizontal bar chart quantifying the route counts for these leading airports, each labelled with their respective IATA identifier, part (c) delineates the intricate network architecture, spotlighting the top 10 airports by route volume in red with corresponding IATA codes.

5.2 Macro-Scale Analysis of Domestic Flight Network in the UK

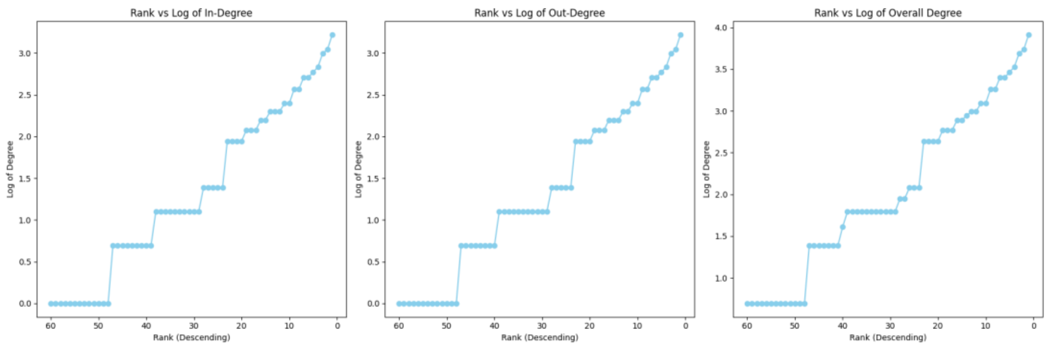
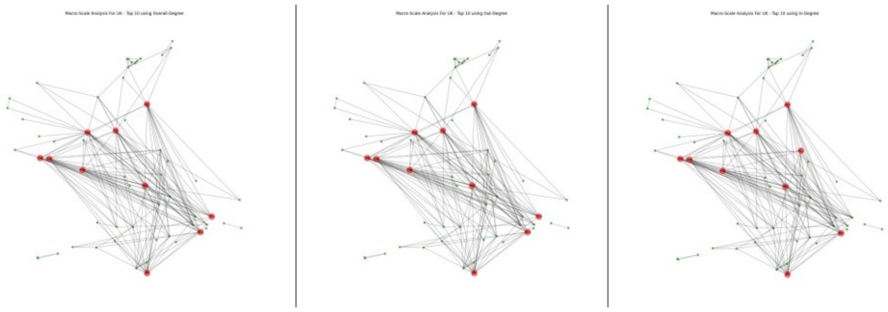


Figure 9: Logarithmic scale of in-degree, out-degree, and overall-degree against their descending rank for each location in UK



a b c

Figure 10: Visualization of Top 10 Airports in the UK based on their (a) overall-degree, (b) out-degree (c) in-degree distributions.

In Figure 9, the first plot shows a trend that a small number of airports have very high in-degree, as indicated by the steep curve at the higher ranks, as the rank decreases the in-degree decreases more gradually, indicating fewer incoming flights for lower-ranked airports. The middle plot again shows a steep initial curve which indicates that the highest-ranked airports have significantly more outgoing flights compared to the rest. The third plot is consistent with the previous plots, where those with the highest overall degree, are significantly more connected than those with lower ranks.

5.3 Micro-Scale Analysis of Domestic Flight Network in the UK

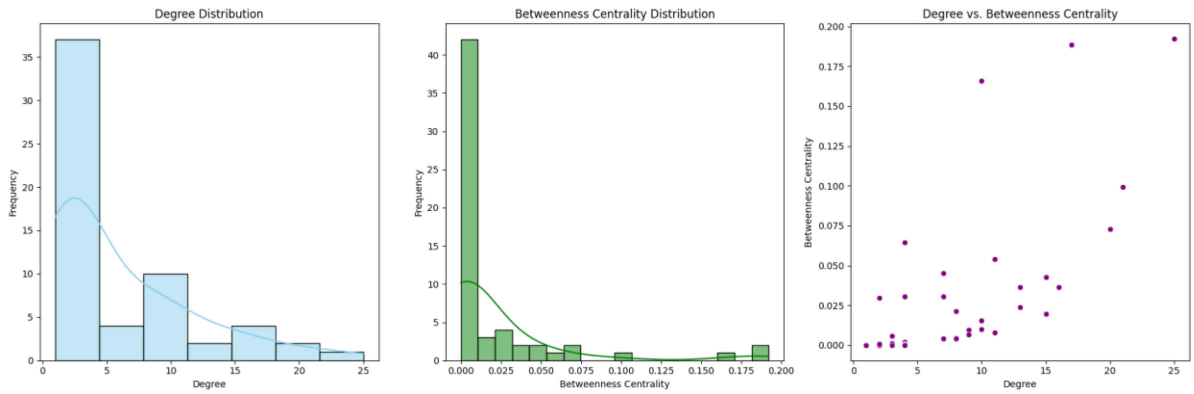
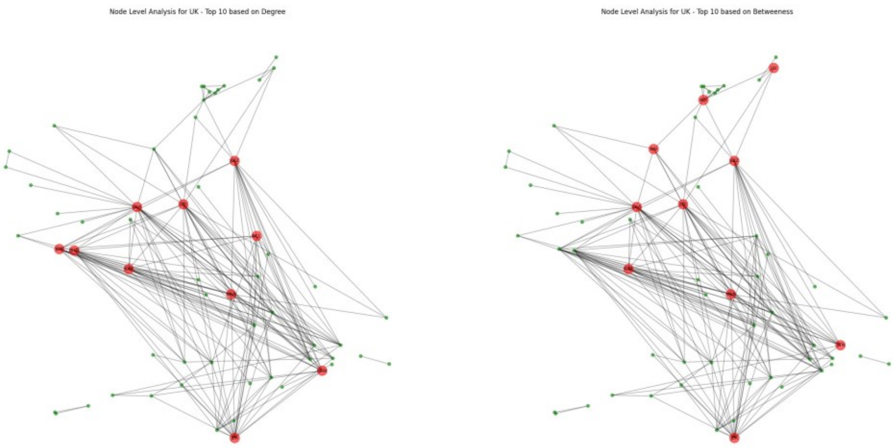


Figure 11: Node Level analysis using Degree Distribution and Betweenness Centrality for UK

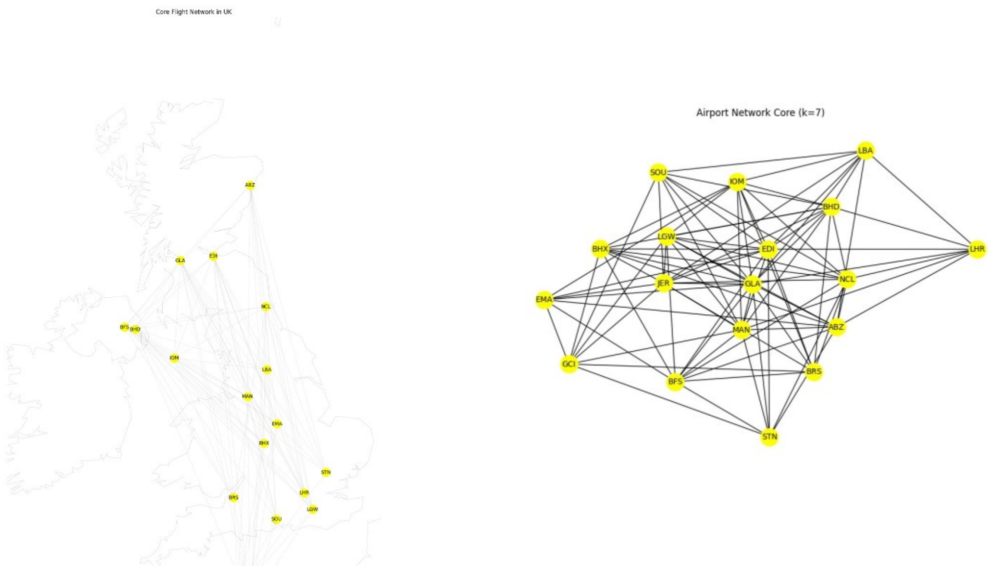
In Figure 11, the first plot it is seen that most nodes have a low degree, with the frequency quickly dropping off as the degree increases, this trend suggests that the network is not dominated by any single hub but rather has a dispersed connectivity pattern with possibly several smaller hubs. The second plot is also like the degree distribution plot, there is a steep decline in the frequency as betweenness centrality increases, indicating that few nodes act as critical intermediaries in the network's shortest path. In the third plot, the points do not show a clear upward trend instead the points are more dispersed, with a cluster at the lower end of both measures, suggesting that nodes with fewer connections are not generally those through which many shortest paths pass.



a b

Figure 12: Visualization of Top 10 Airports in the UK based on their (a) Degree and (b) Betweenness Centrality

5.4 Meso-Scale Analysis of Domestic Flight Network in the UK



a b

Figure 13: Core Community Analysis of the UK's Domestic Flight Network via K-Core Decomposition, Highlighting a Cohesive Cluster of Airports Each with a Minimum of 7 Connections.

5.5 Assortativity Analysis of Domestic Flight Network in the UK

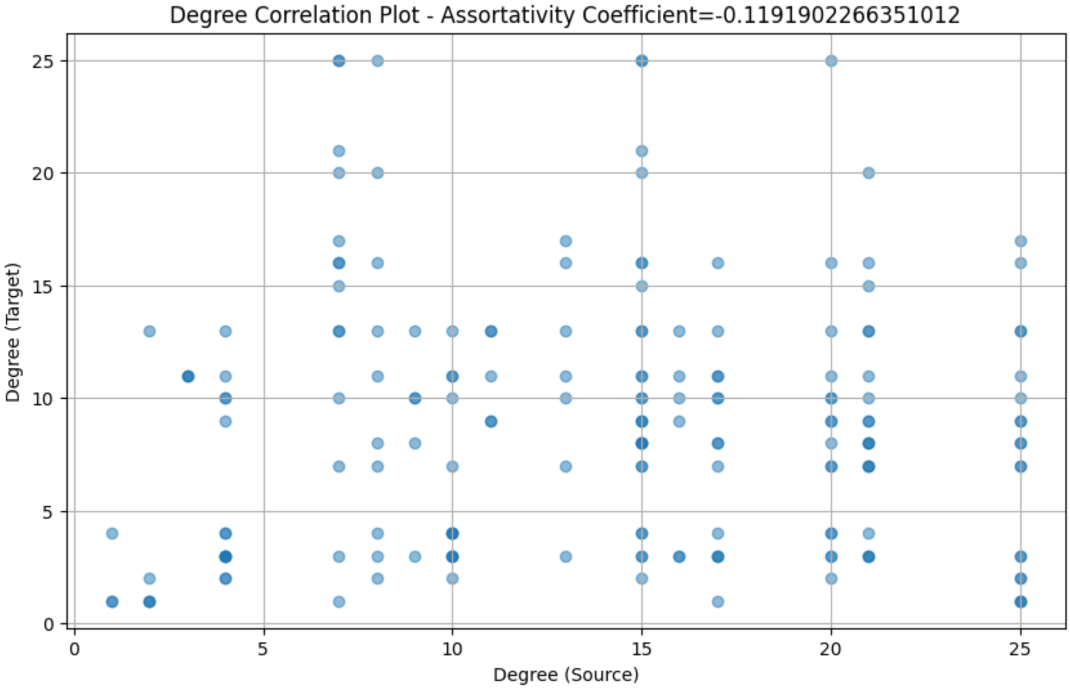


Figure 14: Scatter Plot Illustrating Assortativity in the UK's Domestic Flight Network, revealing a Slight Disassortative Tendency as Evidenced by the Negative Assortativity Coefficient, with High-Degree Airports Preferentially Connecting to Low-Degree Counterparts

6. Multi-Scale Network Analysis for China

6.1 Network Visualization of Domestic Flight Network in the China

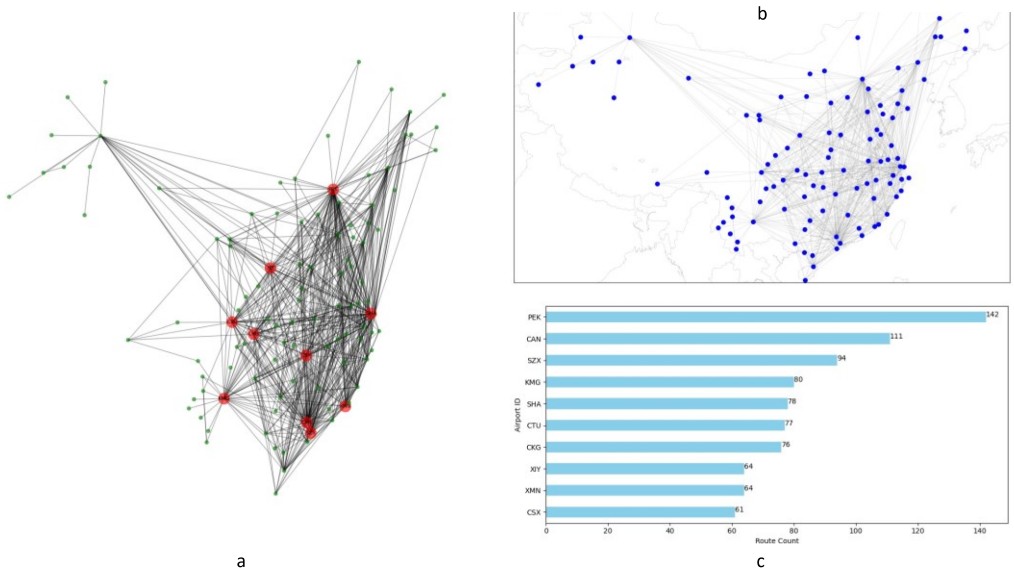


Figure 15: Comprehensive Network Visualization of the China's Domestic Air Traffic — Part (a) delineates the intricate network architecture, spotlighting the top 10 airports by route volume in red with corresponding IATA codes, part (b) illustrates a geospatial overlay of flight routes upon the China map, and part (c) provides a horizontal bar chart quantifying the route counts for these leading airports, each labelled with their respective IATA identifier.

6.2 Macro-Scale Analysis of Domestic Flight Network in the China

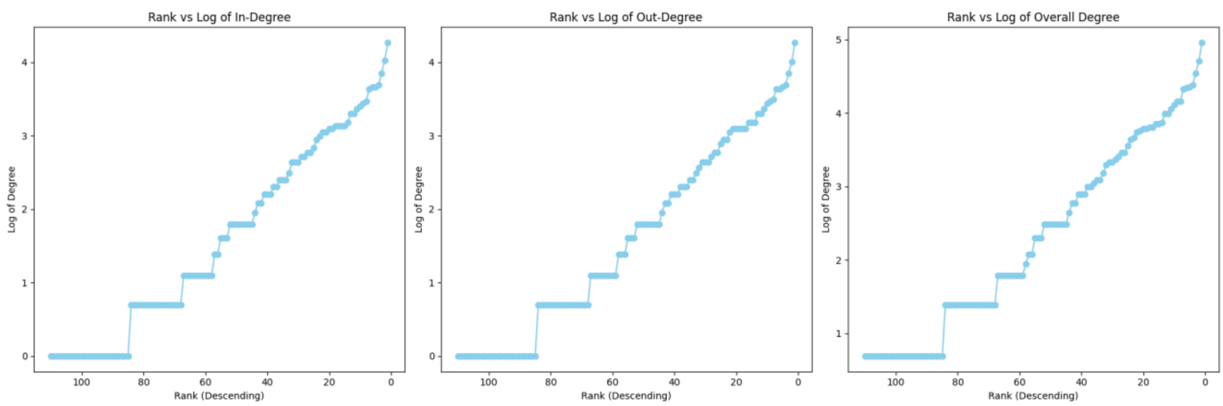


Figure 16: Logarithmic scale of in-degree, out-degree, and overall-degree against their descending rank for each location in China

In Figure 16, the first plot shows a steep curve at the highest ranks, indicating that a few locations have a very high in-degree, and as the rank decreases the in-degree decreases more gradually, suggesting that most locations have a relatively low number of incoming connections. The second plot also shows a similar pattern to the in-degree plot, here a few locations have a high out-degree, and many have low out-degree. The steepness in the initial part of the curve suggests that the locations with highest out-degree are significantly more connected than others. The third plot also follows a similar pattern with a steep drop-off for the most connected locations and then a more gradual decline.

Figure 17: Visualization of Top 10 Airports in China based on their (a) overall-degree, (b) out-degree (c) in-degree distributions.

6.3 Micro-Scale Analysis of Domestic Flight Network in the China

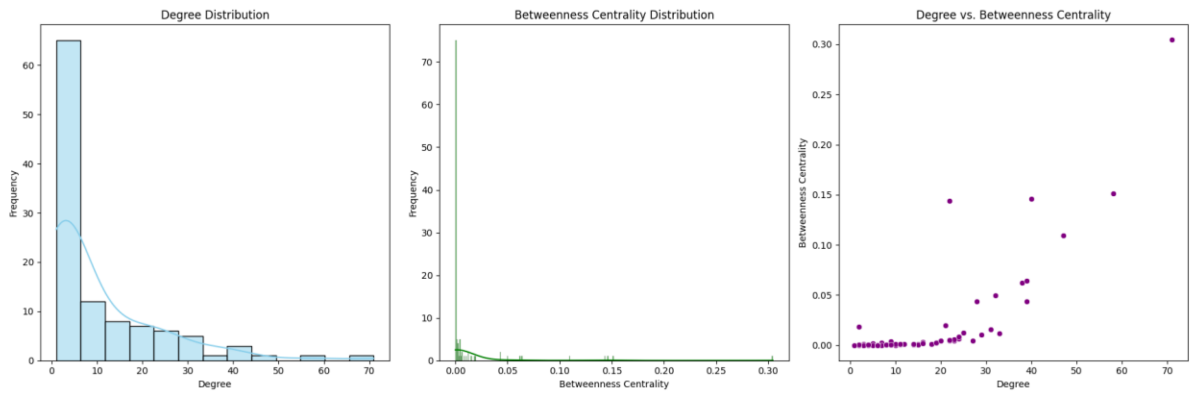
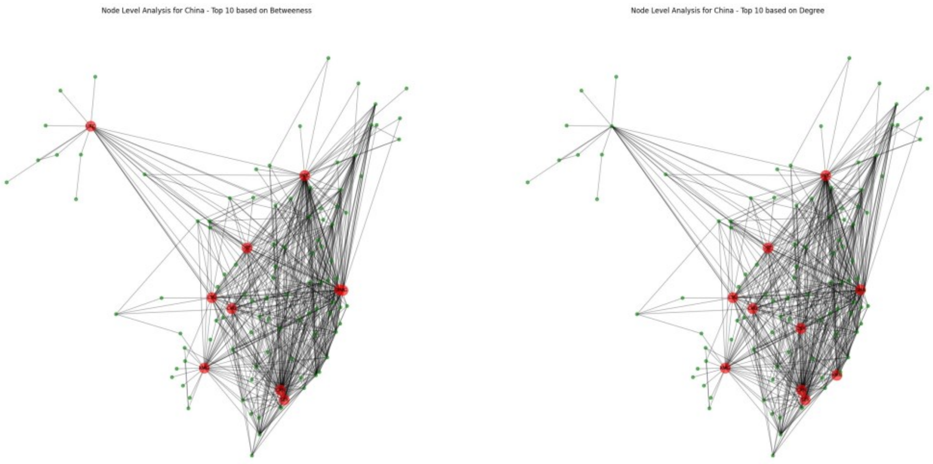


Figure 18: Node Level analysis using Degree Distribution and Betweenness Centrality for China

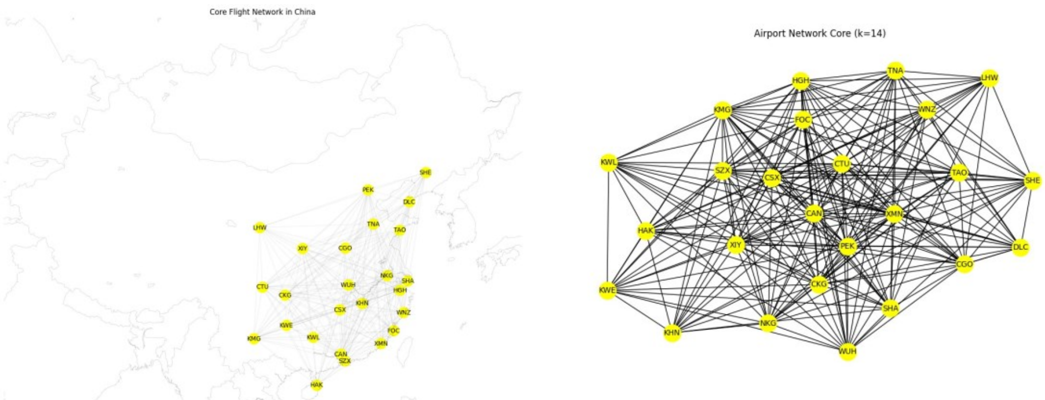
In Figure 18, the first plot shows most of the nodes having a low degree with very few nodes having a high degree, the distribution is heavily skewed to the left. The second plot shows that most nodes have a very low betweenness centrality, indicating that they are not often on the shortest path between other nodes. There are a few nodes with high betweenness centrality, suggesting that these few nodes could have significant control over the network’s flow or connectivity. In the third plot, there doesn’t appear to be a clear trend or correlation between degree and betweenness centrality, which indicates that nodes with many connections are not necessarily those through which most paths in the network pass.



a b

Figure 19: Visualization of Top 10 Airports in China based on their (a) Betweenness Centrality and (b) Degree.

6.4 Meso-Scale Analysis of Domestic Flight Network in the China



a b

Figure 20: Core Community Analysis of the China's Domestic Flight Network via K-Core Decomposition, Highlighting a Cohesive Cluster of Airports Each with a Minimum of 14 Connections.

6.5 Assortativity Analysis of Domestic Flight Network in the China

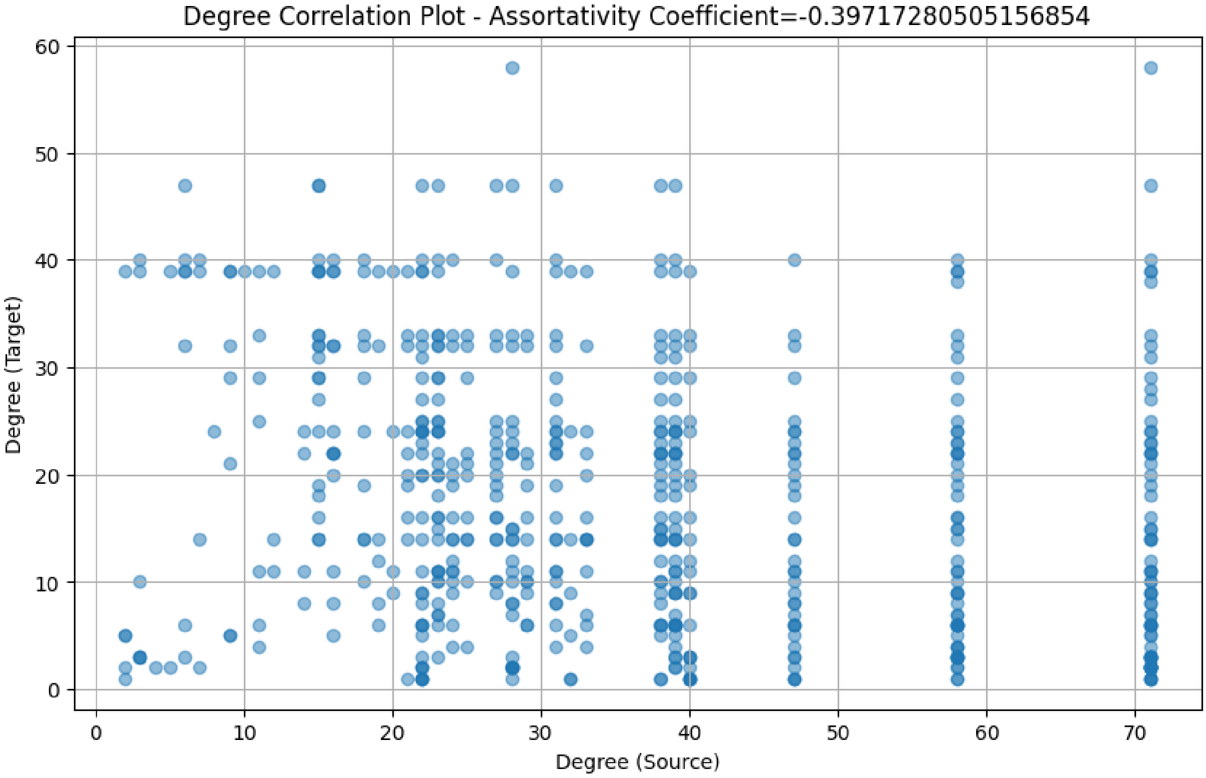
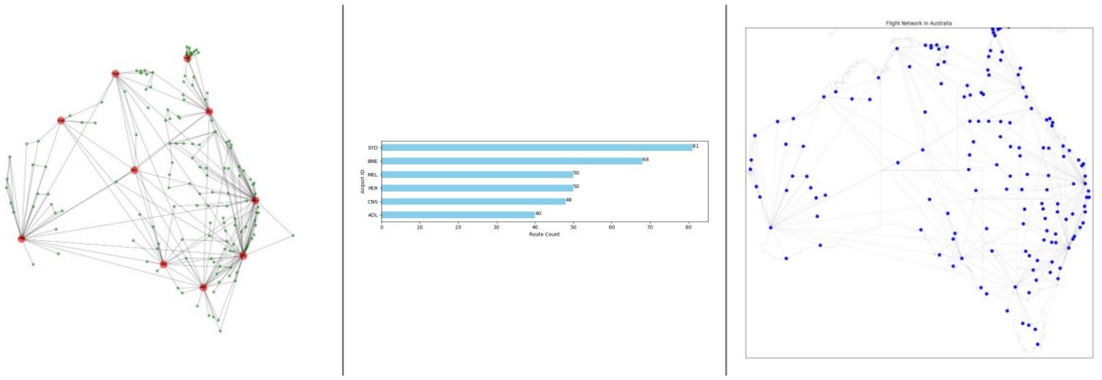


Figure 21: Scatter Plot Illustrating Assortativity in the China's Domestic Flight Network, revealing a Disassortative Tendency as Evidenced by the Negative Assortativity Coefficient, with High-Degree Airports Preferentially Connecting to Low-Degree Counterparts

7. Multi-Scale Network Analysis for Australia

7.1 Network Visualization of Domestic Flight Network in the Australia



a b c

Figure 22: Comprehensive Network Visualization of the Australia's Domestic Air Traffic — Part (a) delineates the intricate network architecture, spotlighting the top 10 airports by route volume in red with corresponding IATA codes, part (b) provides a horizontal bar chart quantifying the route counts for these leading airports, each labelled with their respective IATA identifier, and part (c) illustrates a geospatial overlay of flight routes upon the Australia map.

7.2 Macro-Scale Analysis of Domestic Flight Network in the Australia

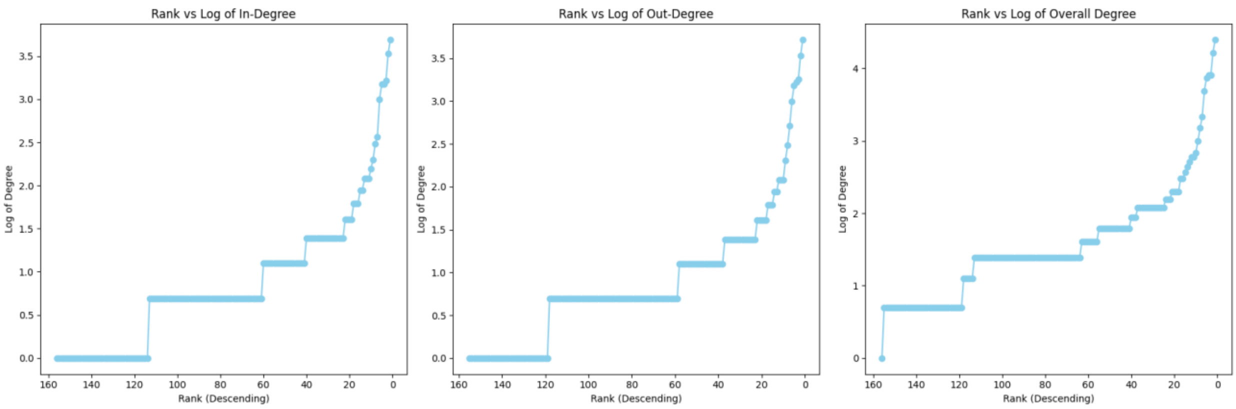


Figure 23: Logarithmic scale of in-degree, out-degree, and overall-degree against their descending rank for each location in Australia

In Figure 23, the first plot shows a sharp increase at the lower end of the x-axis with significantly higher in-degree values. The step-like nature of the plot on a log scale is indicative of a heavily tailed distribution, where only a few locations have a high number of incoming connections, while majority have significantly fewer. From the out-degree plot it is noticed that small number of locations have a high-out degree, and there's a steep decrease as the rank increases, signifying that most locations have a lower out-degree. In the third plot also, the trend is consistent with the in-degree and out-degree plots, a few locations have very high overall connectivity, while the rest have much lower connectivity.

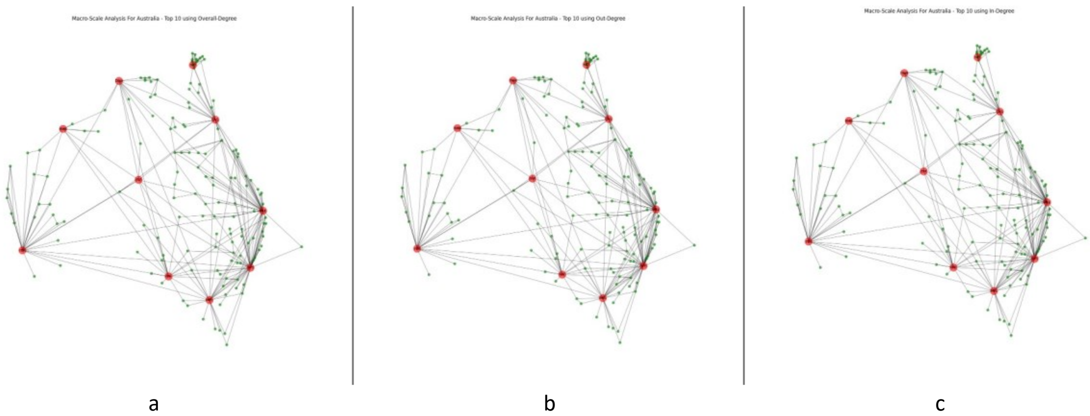


Figure 24: Visualization of Top 10 Airports in Australia based on their (a) overall-degree, (b) out-degree (c) in-degree distributions.

7.3 Micro-Scale Analysis of Domestic Flight Network in the Australia

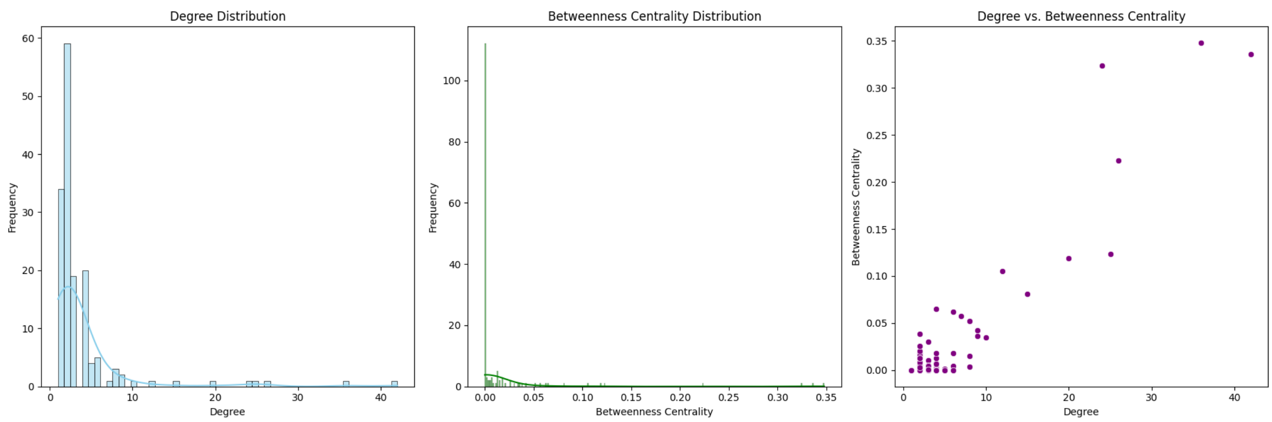
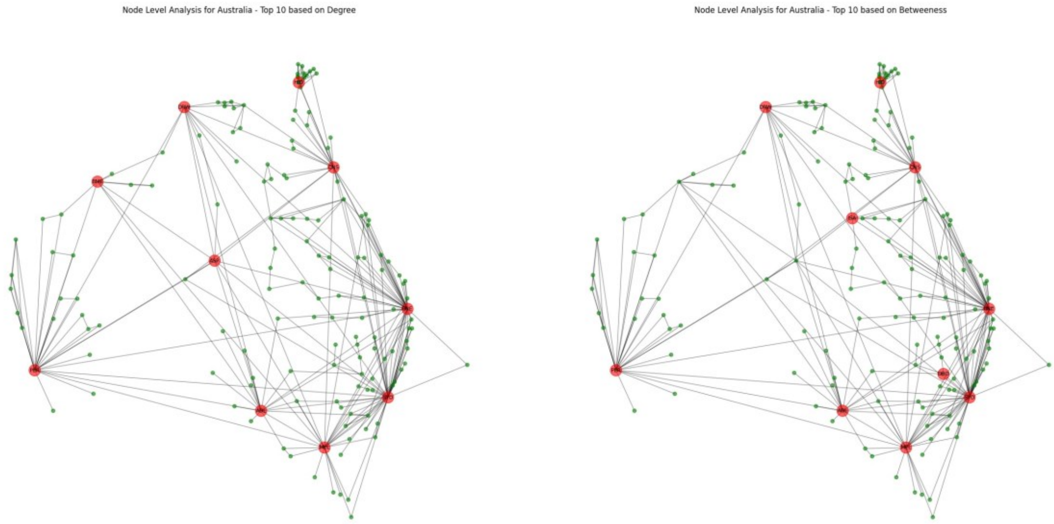


Figure 25: Node Level analysis using Degree Distribution and Betweenness Centrality for Australia

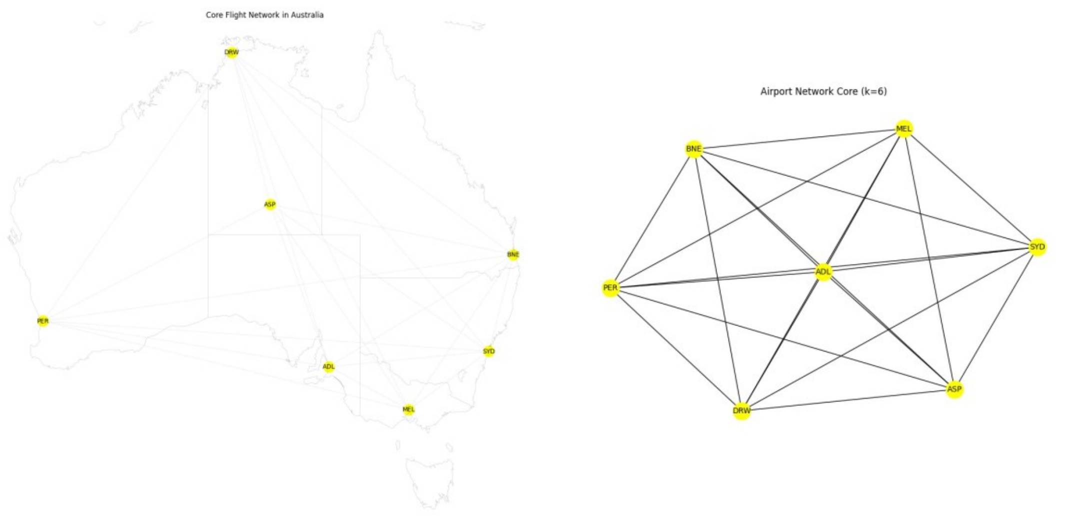
In Figure 25, the first plot has a long tail indicating that most nodes have a low degree, while a few have a high degree. The second plot indicates that most nodes have a very low betweenness centrality, suggesting that only a few nodes are critical in terms of controlling the flow of information or resources in the network. The third plot shows a trend where nodes with the highest degree also exhibit higher betweenness centrality, which is expected since nodes with more connections are more likely to be on the shortest path between other nodes.



a b

Figure 26: Visualization of Top 10 Airports in Australia based on their (a) Degree and (b) Betweenness Centrality.

7.4 Meso-Scale Analysis of Domestic Flight Network in the Australia



a b

Figure 27: Core Community Analysis of the Australia's Domestic Flight Network via K-Core Decomposition, Highlighting a Cohesive Cluster of Airports Each with a Minimum of 6 Connections.

7.5 Assortativity Analysis of Domestic Flight Network in the Australia

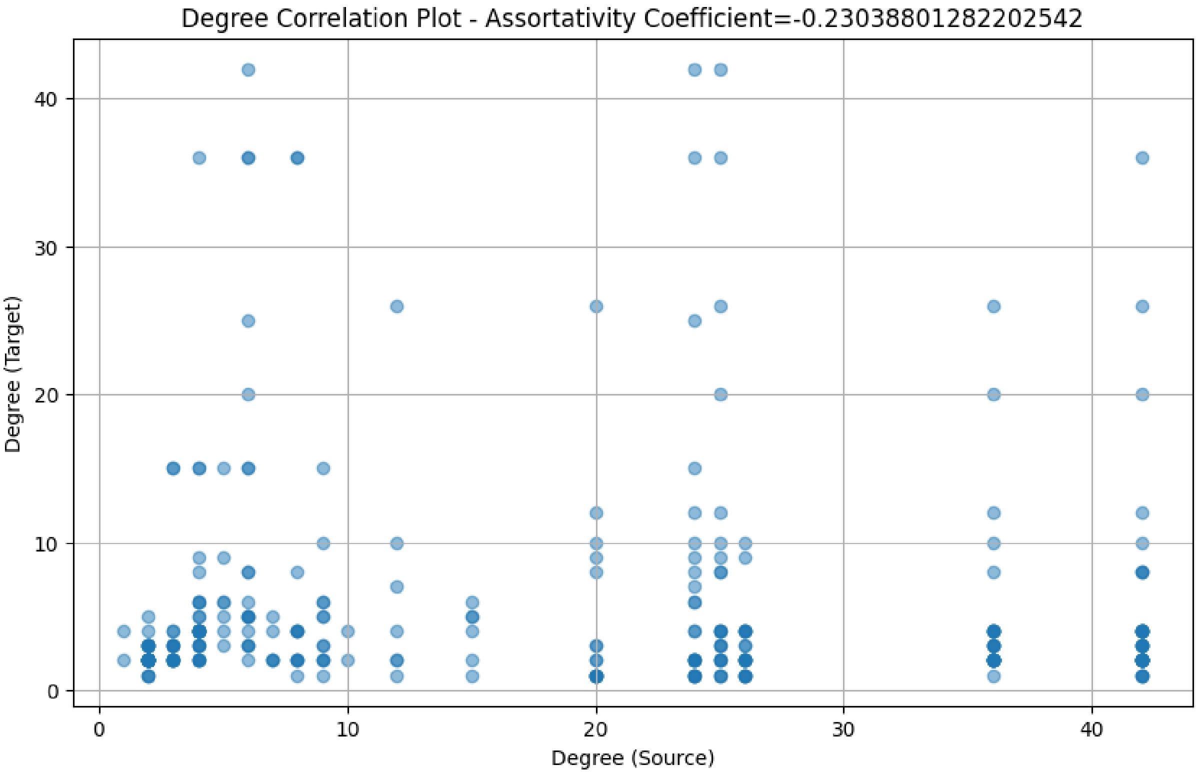


Figure 28: Scatter Plot Illustrating Assortativity in the Australia's Domestic Flight Network, revealing a Disassortative Tendency as Evidenced by the Negative Assortativity Coefficient, with High-Degree Airports Preferentially Connecting to Low-Degree Counterparts

8. Analysis and Discussions

In this comprehensive study of multi-scale network analysis on domestic flight data across the United States, the United Kingdom, China, and Australia, the selection of specific metrics for analysis was meticulously tailored to address the intricate dynamics of aviation networks. The decision to employ metrics such as degree distribution, betweenness centrality, and assortativity was rooted in their profound ability to unravel the complex interplay of nodes within the aviation network, thereby providing actionable insights for airline operators.

The degree distribution was primarily chosen for its efficacy in identifying the most connected airports within the network. These airports, characterized by their high degree of connections, are pivotal in the aviation network, serving as central hubs around which traffic is orchestrated. Understanding the distribution of these hubs is critical for airlines, as it directly influences routing, scheduling, and resource allocation decisions. The degree distribution metric provides a clear view of how traffic is dispersed across the network, highlighting potential areas of congestion, and identifying key nodes that warrant additional focus for maintaining efficient operations.

In conjunction with degree distribution, betweenness centrality was incorporated into the analysis to gauge the strategic importance of individual airports within the network. Airports with high betweenness centrality scores emerge as crucial nodes, not merely in terms of the volume of traffic they handle but also in their role as integral connectors in the network. These airports often form the most efficient paths in the network, making them indispensable in routing and especially crucial in formulating contingency plans during network disruptions.

The assortativity measure was another cornerstone of this analysis, providing insights into the nature of connections between different airports. Understanding whether larger, more connected hubs predominantly link with similar airports or with smaller regional ones is vital for discerning the network's structural tendencies. Such insights are paramount for airlines in strategizing network development and exploring new, potentially lucrative routes.

While these metrics were central to the study, the exclusion of other metrics such as clustering coefficients or network density was a considered choice. The primary rationale behind this was the direct relevance and applicability of the chosen metrics to the operational aspects of airline networks. The aim was to streamline the analysis to metrics that not only provided profound insights but were also pragmatic in application, ensuring that the findings could be directly transposed to operational strategies and decision-making processes within the aviation industry.

The alignment of the network analysis with population distribution and airline operations revealed critical correlations. Airports with high degrees of connectivity often coincided with major population centres and economic hubs, indicating a direct relationship between the demand for air travel and the structure of the flight network. This correlation underscores the importance of aligning flight operations with passenger demand, optimizing resource allocation, and ensuring the profitability of routes.

9. Adapting Aircraft Design: Analysing the Interplay Between Fuel Economics and Network Configurations

The study commences with the creation of spatial graphs based on actual airline route data. These graphs form the basis for a comparative analysis against spatial graphs generated under varying distance thresholds, using a distance penalty to limit connectivity among nodes. Adjusting these thresholds simulates different fuel price scenarios, providing insights into how network connectivity might respond to economic fluctuations. This method offers a detailed perspective on the potential adaptability of airline networks to changes in fuel costs.

To understand the impact of fuel prices with connectivity, the following formula was implemented:

$$\begin{aligned}&Adjusted\ Distance\ Threshold\\&\quad = Base\ Distance\ Threshold\ \times \left(1 - \frac{Percentage\ of\ Fuel\ Price\ Change}{100}\right)\end{aligned}$$

The Spatial Graph analysis for each country is visualized along with visualization for fuel price impact on network connectivity.

9.1 Random Spatial Graph Analysis Simulating Impact of Fuel Prices on Connectivity for USA

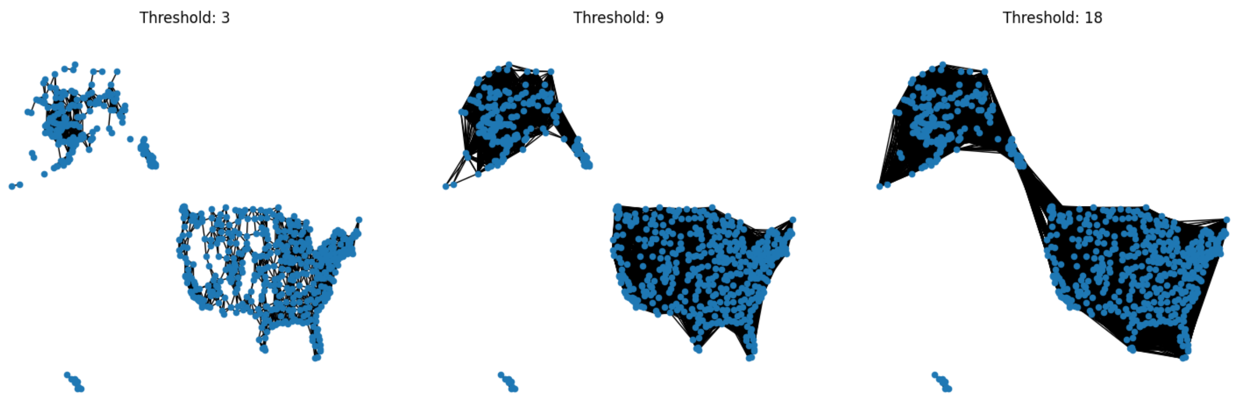


Figure 29: Random Spatial Graph Analysis using varying values for Distance Threshold to assess Network Connectivity for USA.

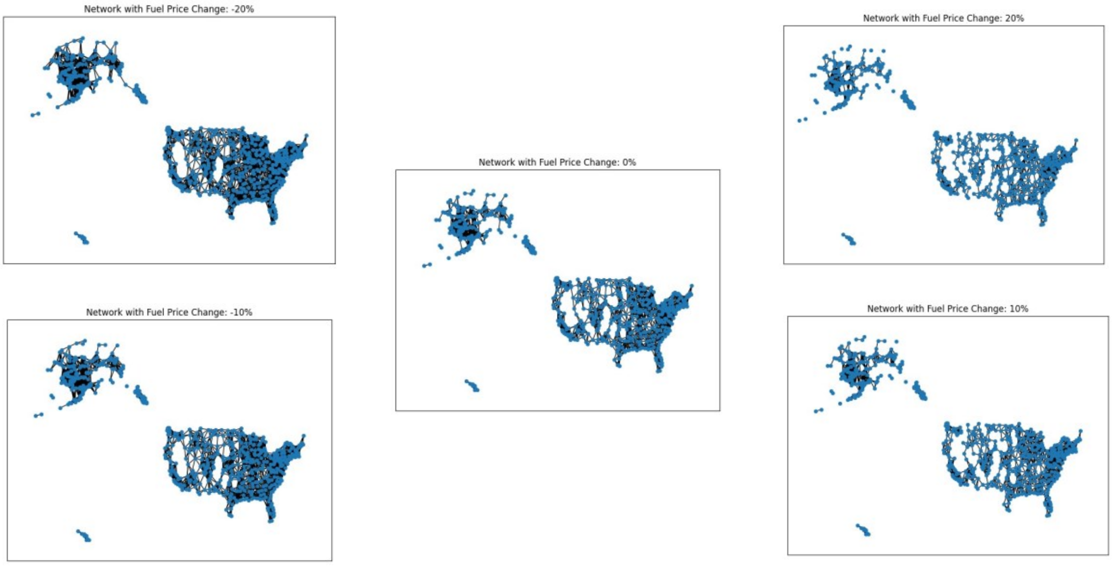


Figure 30: Using Random Spatial Graph analysis to visualize the impact of fuel prices on network connectivity using a base distance threshold of 3 for USA.

9.2 Random Spatial Graph Analysis Simulating Impact of Fuel Prices on Connectivity for UK

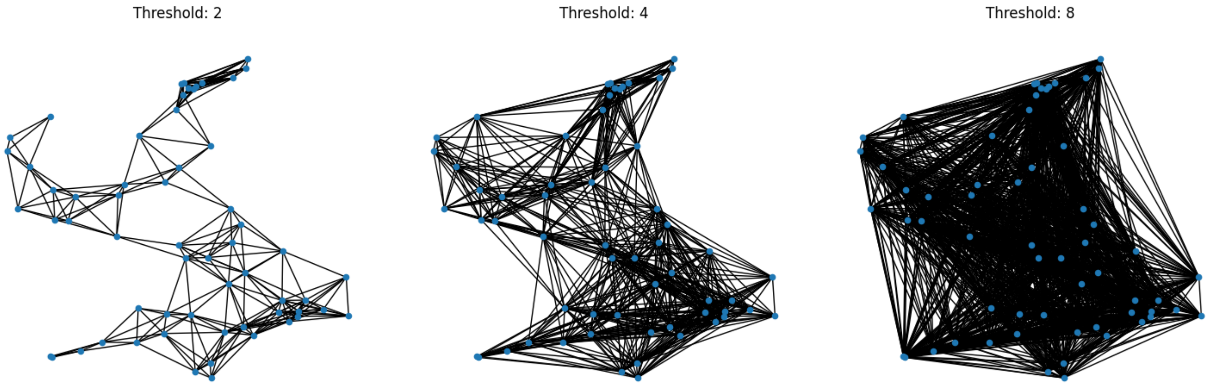


Figure 31: Random Spatial Graph Analysis using varying values for Distance Threshold to assess Network Connectivity for UK.

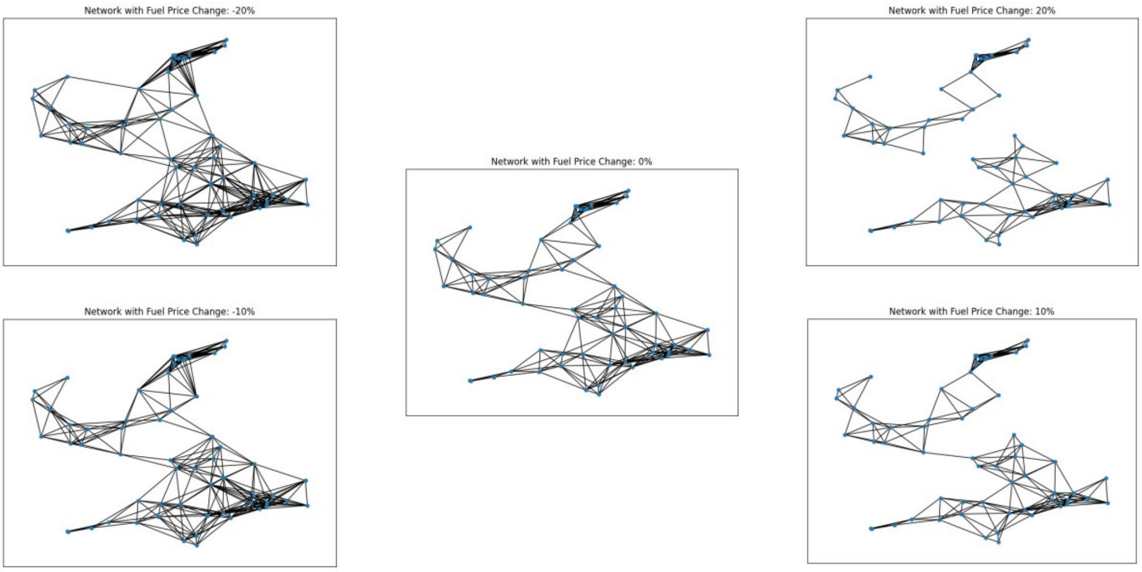


Figure 32: Using Random Spatial Graph analysis to visualize the impact of fuel prices on network connectivity using a base distance threshold of 2 for UK.

9.3 Random Spatial Graph Analysis Simulating Impact of Fuel Prices on Connectivity for China

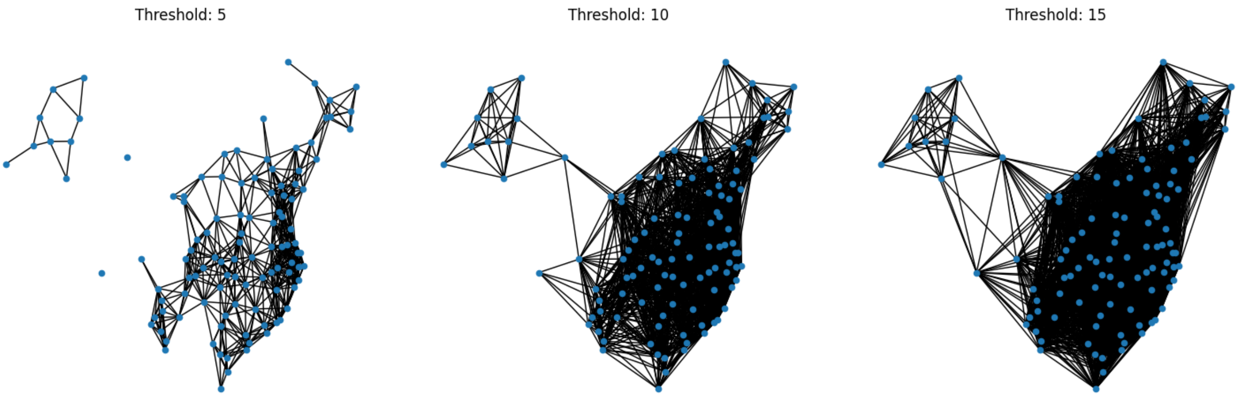


Figure 33: Random Spatial Graph Analysis using varying values for Distance Threshold to assess Network Connectivity for China.

Figure 34: Using Random Spatial Graph analysis to visualize the impact of fuel prices on network connectivity using a base distance threshold of 5 for China.

9.4 Random Spatial Graph Analysis Simulating Impact of Fuel Prices on Connectivity for Australia

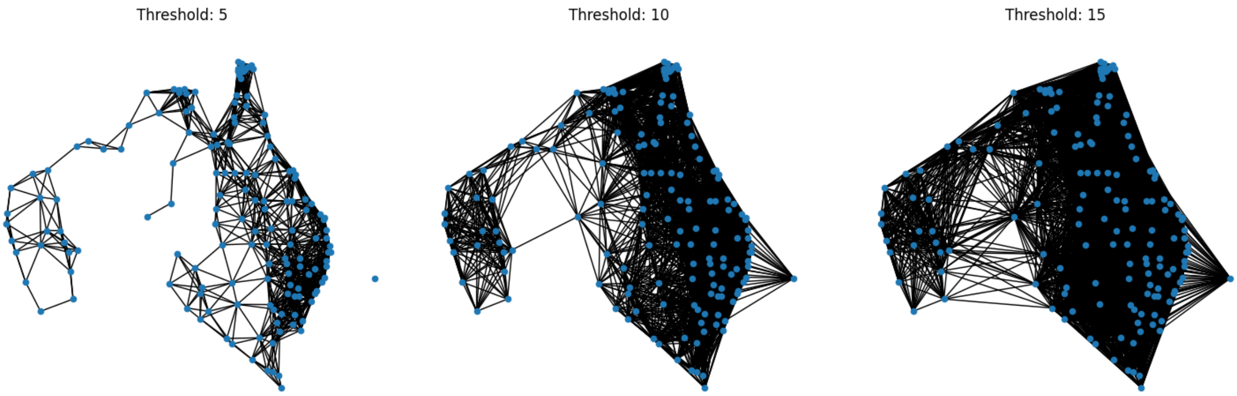


Figure 35: Random Spatial Graph Analysis using varying values for Distance Threshold to assess Network Connectivity for Australia.

Figure 36: Using Random Spatial Graph analysis to visualize the impact of fuel prices on network connectivity using a base distance threshold of 5 for Australia.

9.5 Discussions

A significant relationship is identified between the distance thresholds in spatial graphs and the degree of connectivity in airline networks. Moreover, the study highlights the impact of geographical location on network changes. Airports exhibit varying responses in connectivity based on their geographical positions. This variation is attributed to the unique spatial distribution of airports and their accessibility.

The impact of fuel price changes on network connectivity can be clearly visualized in the spatial graph plots for each country. The general trend seen across all the countries is as follows:

- Network connectivity becomes more intense and connected when there is a 20% and 10% reduction in fuel price, suggesting cheaper fuel prices enables more network connectivity across short and long nodes (airports).
- Connections between nodes (airports) reduces as fuel prices increase by 10% and 20%, suggesting that higher fuel prices impact network connectivity, and direct connection across long nodes (airports) do not seem feasible anymore.

This trend suggests a network's adaptation to more favourable economic conditions, where longer, more costly routes become financially viable.

The study provides crucial insights into the potential impacts of fuel price fluctuations on the airline industry, particularly in terms of shaping future aircraft network designs. The observed sensitivity of airline networks to fuel price variations highlights the need for flexible and adaptive strategies in airline operations and network planning.

10. Conclusion

The comprehensive multi-scale network analysis of domestic flight data across the United States, the United Kingdom, China, and Australia has yielded insightful revelations into the dynamics of aviation networks. This study, underpinned by meticulously selected analytical metrics, has illuminated the intricate interplay of nodes within these networks, offering actionable insights for airline operators and policy makers. Crucially, the degree distribution, betweenness centrality, and assortativity metrics have emerged as powerful tools in unravelling the complex network structures, highlighting key nodes and connectivity patterns. These insights are particularly valuable in the context of airline operations and strategic planning, where understanding the network's topology can significantly impact decision-making processes. The findings have underscored the role of major hubs in maintaining network efficiency and the strategic importance of certain airports in facilitating effective traffic flow across the network.

Furthermore, the analysis has adeptly demonstrated the sensitivity of these networks to fuel price variations. By simulating different fuel price scenarios, the study has provided a clear visualization of how changes in fuel economics can influence network connectivity. The observed trends across various countries indicate that lower fuel prices enhance network connectivity, enabling economically viable longer routes, whereas higher fuel prices lead to reduced connectivity, especially impacting longer-haul routes. These results highlight the critical need for adaptive strategies in airline network planning and operations, underscoring the importance of economic considerations in the design and optimization of future aircraft networks. The study's findings offer a roadmap for airlines and policymakers to navigate the complexities of the aviation industry, ensuring more robust, efficient, and economically viable networks in the face of fluctuating fuel prices and other economic variables.

In conclusion, this multi-scale network analysis has not only provided a comprehensive overview of the current state of domestic airline networks in several key regions but has also laid the foundation for future explorations into this dynamic field. As the aviation industry continues to evolve in response to economic, technological, and environmental factors, such analyses will be invaluable in shaping the strategies and designs that will define the future of air travel.

The code used for the purpose of this study is maintained in my GitHub repo: https://github.com/juliangdz/Multi-Scale-Network-Analysis-of-Flight-Data.

11. References

1. ip2location. IATA/ICAO List (https://github.com/ip2location/ip2location-iata-icao).
2. Guo W, Toader B, Feier R, Mosquera G, Ying F, Oh SW, et al. Global air transport complex network: multi-scale analysis. SN Appl Sci. 2019 Jul 8;1(7):680.
3. Hagberg hagberg AA, -Los lanlgov, Schult DA, Swart swart PJ. Exploring Network Structure, Dynamics, and Function using NetworkX [Internet]. 2008. Available from: http://conference.scipy.org/proceedings/SciPy2008/paper_2