

**PSG COLLEGE OF TECHNOLOGY, COIMBATORE 641004**

**Department of Computer Science & Engineering**



**19ZO02 – SOCIAL AND ECONOMIC  
NETWORK ANALYSIS**

**Airport Route Analysis**

**By**

**19Z316 – GOWTHAM S**

**19Z349 – SRIKANTH**

**20Z461 – CHANDRAPRAKASH J**

**20Z464 – RITHISH B**

**20Z465 – SRINIVASAN**

**BE CSE G2 (2019 – 2023)**

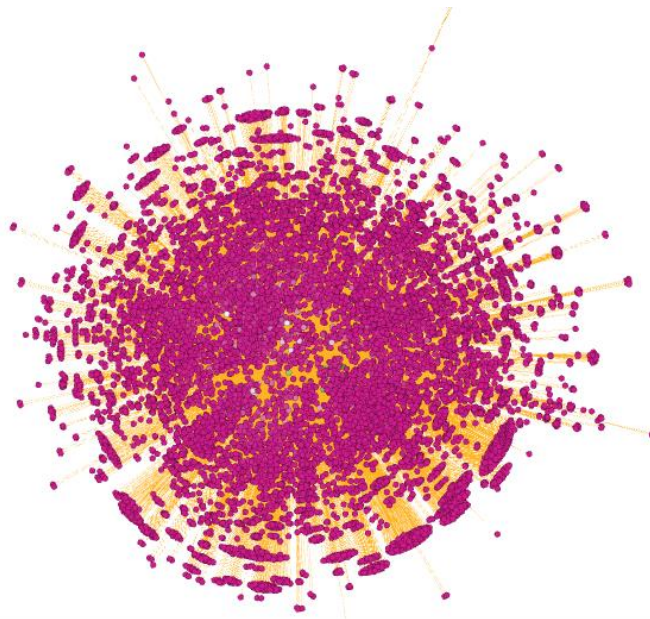
## **PROBLEM STATEMENT**

The analysis of the air-transport route dataset which focuses on

- Shortest path from source to destination
- Airport with the highest out-going flights
- Airport with the highest number of incoming flights
- Airport which occurs in between two airports to make possible connective flights
- The busiest route
- The possible shortest path from source to destination
- 

## **DATASET DESCRIPTION**

The dataset consist of two csv files. The graph is representation of routes of aeroplane in cities of United States and Canada. The first csv file consists 7698 entries which is the details about airport. The second CSV file consists of 67663 entries which are the details about the route between airport to airport. The node of the network is the airport Id which is involved in the analysis of the airway route. The edge between the nodes is formed using the route between the airport.



*Figure.1. Network of analysis of route.*

## **TOOLS USED**

### **Gephi**

Gephi is an open-source software for network visualization and analysis. It helps data analysts to intuitively reveal patterns and trends, highlight outliers and tells stories with their data. It uses a 3D render engine to display large graphs in real-time and to speed up the exploration. Gephi combines built-in functionalities and flexible architecture to:

- explore
- analyse
- spatialize
- filter
- cluster
- manipulate
- export

all types of networks.

Gephi is based on a visualize-and-manipulate paradigm which allow any user to discover networks and data properties. Moreover, it is designed to follow the chain of a case study, from data files to nice printable maps.

### **NetworkX**

NetworkX is a set of Python-based tools for building, modifying, and researching the composition, dynamics, and operation of complicated networks. Large complicated networks that are represented as graphs with nodes and edges are studied using this method. We can load and store complex networks using networkx. We are able to create a wide variety of random and conventional networks, analyze their structure, create network models, create new network methods, and even draw them. Classes for graphs that permit numerous edges between any two nodes are provided by networkx. You can add the same edge more than once using the Multigraph and MultiDiGraph classes, possibly with different edge data. For some applications, this can be quite useful, however, many algorithms do not work well on such networks.

## **CHALLENGES FACED**

- The dataset which is gathered is in gzip format and we face difficulty in the conversion of the gzip to CSV
- The dataset consists of different attributes like Airport ID, Name, City, Country, IATA, ICAO, Latitude, Longitude, Altitude, Timezone, DST, Database Timezone, Type, and Source. It can't be loaded into gephi directly for the visualization of the graph. So, we used Airport ID as the node and with the help route file, edges are created between the nodes.
- Visualization of the dataset remains a huge process because of the large number of connections between the nodes.

## CONTRIBUTION OF TEAM MEMBERS

Roll No.	Name	Contribution
19Z316	GOWTHAM S	The busiest route , the possible shortest path from source to destination
19Z349	SRIKANTH	Airport which occurs in between two airports to make possible connective flights
20Z461	CHANDRAPRAKSH J	Airport with the highest number of incoming flights
20Z464	RITHISH B	Airport with the highest out-going flights
20Z465	SRINIVASAN	Shortest path from source to destination

## ANNEXURE I : CODE

Code link : [https://github.com/johnDC14/SENA/blob/main/airway\\_route\\_analysis.ipynb](https://github.com/johnDC14/SENA/blob/main/airway_route_analysis.ipynb)

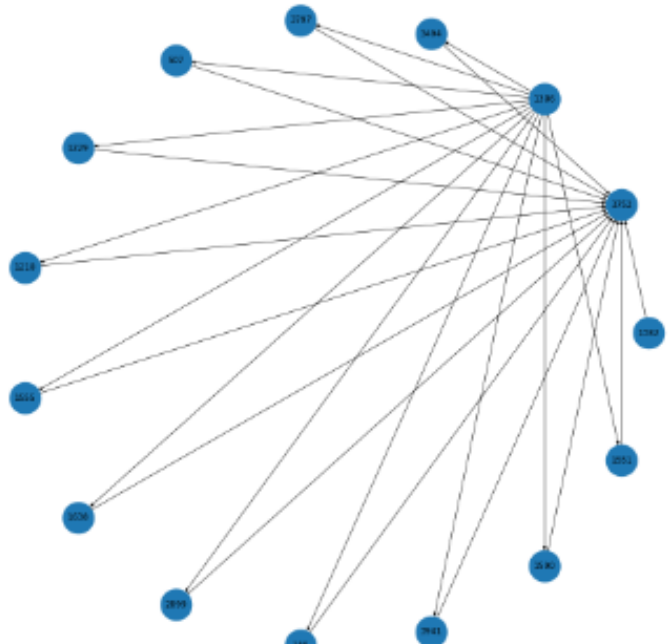
## ANNEXURE II : SNAPSHOTS OF OUTPUT

```
Enter the Source City : paris
Enter the destination City :philadelphia
path - 1
Paris Airport - Charles de Gaulle International Airport to
Philadelphia Airport - Philadelphia International Airport

path - 2
Paris Airport - Paris-Orly Airport to
Athens Airport - Eleftherios Venizelos International Airport to
Philadelphia Airport - Philadelphia International Airport
```



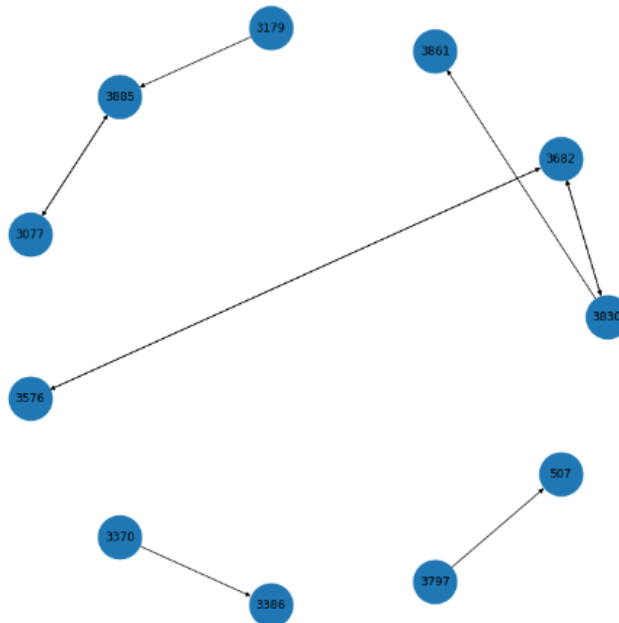
a) Shortest path



b) possible shortest path

`['Frankfurt', 'Paris', 'Amsterdam', 'Istanbul', 'Atlanta']`

c) Airport with the highest in-coming and out-going flights



d) *Top five busiest route*

"Ted Stevens Anchorage International Airport"  
"Los Angeles International Airport"  
"Charles de Gaulle International Airport"  
"Dubai International Airport"  
"Frankfurt am Main Airport"

---

e) *Top five airport acts as a bridge for connective flights*

## **REFERENCES**

- Dataset: <https://www.kaggle.com/datasets/divyanshrai/openflights-airports-database-2017>
- Dataset: <https://www.kaggle.com/datasets/divyanshrai/openflights-route-database-2014>
- <https://gephi.org/users/>
- <https://networkx.org/>
- <https://openflights.org/data.html>
- <https://networkx.org/documentation/stable/tutorial.html>
- <https://stackoverflow.com/questions/31028815/how-to-unzip-gz-file-using-python>