Project Report : ELL 880
Social Network Analysis of Aviation Networks(India)

Chandragupta 2019EET2341
Anand 2019EET2340

January 18, 2021

Departement of Electrical Engineering
IIT Delhi, India

Contents

List of Figures

1 Abstract

We approached this topic in our efforts towards completion of our Course in Social Network Analysis as part of our Curriculum.Our problem statement included analysing entire Indian and a few International Airports and study their complete ecosystem.Based upon the data-set that we extracted using web scraping we also tried to study the various aspects of the network including Costly paths,Factors affecting new Airports conception, factors including their relationship with economics of sustenance,ecological impact and new idea conception. We found that almost 97 percent routes in India have less than 15 flights and only 3 percent have more than 15 flights. We found the centrality measures of various airports,their footfalls and found a big relationship with economies of the affiliated cities.It was found that bird ecosystem changed quite a lot due to coming up these huge airports on Grasslands and provided and alternate ecosystem for them.

2 Introduction

The main problem statement :

- We plan to prepare a network and carry out analysis of Air Traffic across the whole India and also of a select few cities of the world
- On this basis we will figure out which of the Airports are the most important ones based upon analysis of their centrality measure like degree and betweenness.
- We can also try to find out bridges in Air Traffic which might need to be consolidated with more flights between those routes and also determine if extra bridges need to be created to control impact of route inaccessibility.
- We also propose to find relation between the degree centrality of an airport and economic and political prosperity of the affiliated city.
- We also try and predict as to whether inter country relations can also be predicted based upon flight data availability

The approach : Data-set for Indian flights was nowhere to be found along with their co ordinates.We had initially claimed to scrap data over the internet and we wrote scraping code to get the data from easemytrip. We used the same technique to get data for airport coordinates for better visualization on gephi and used them to better show using the “map” and “geolayout” plugins. We also approached AAI for data on airport sustenance and flight statistics. We were provided pointers to some of the data already in the public domain. Based on this Airport data scraped we could figure out various centrality measures like betweenness centrality and closeness centrality of various airports and used data like footfalls we could co relate with the economic prosperity of the reason.

3 Data Collection

Since no open source data was available on current flight network, we decided to collect our own data. to collect data from Internet we use web-scrapping methods. we used selenium as a tool and www.easemytrip.com as the website to scrap. Selenium is open source library and it is close to what a user experience in any website, so working with it was quite easy. EaseMyTrip website is a

flight booking website and the main two reason behind scrapping this website was that it doesn't give any CAPCHE and all flight details between two airports is visible in single page contrary to other website where we have to click "load more" button

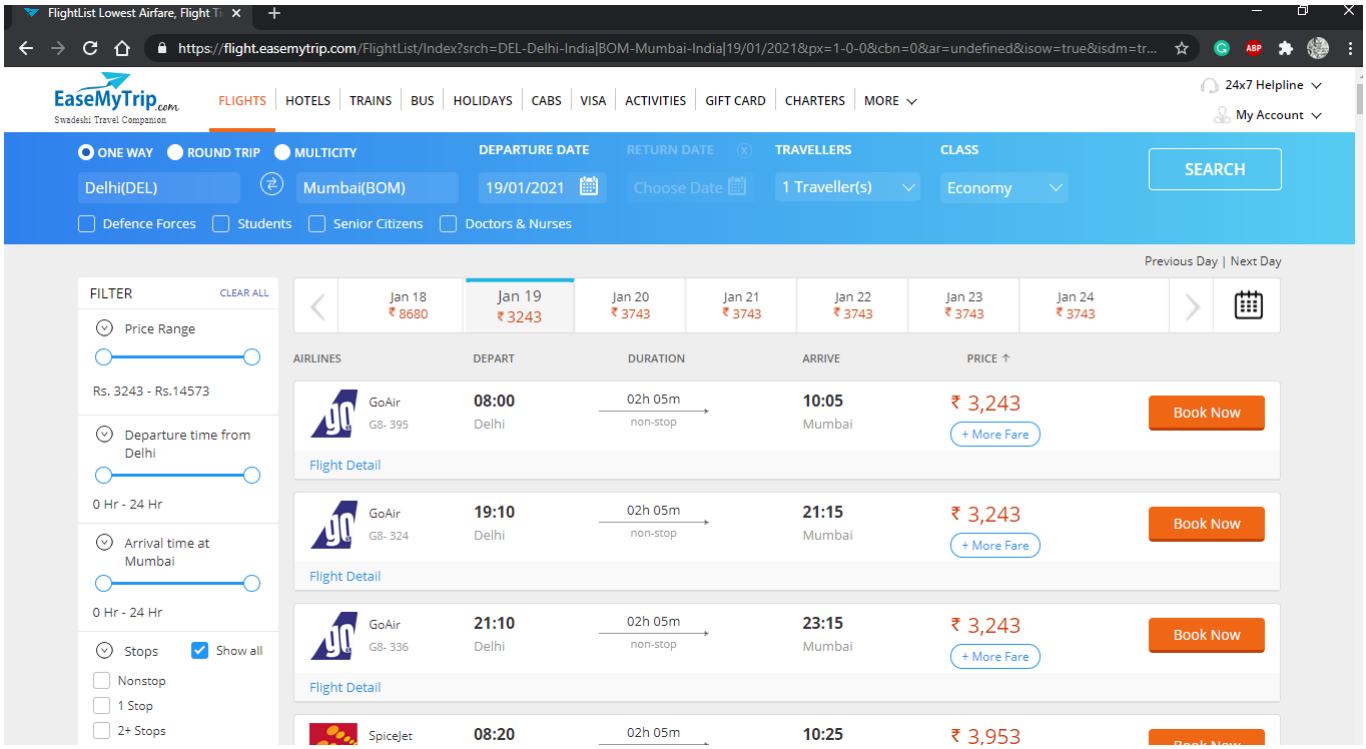


Figure 1: EaseMyTrip Website

3.1 Code and implementation detail

first we needed all list of airport with their code in India. we found that there are 137 airports in India. then we need to find all paths from each airports to another airports, that means we had to scrap 136 x 136 routes between airports. EaseMyTrip website use URL which consist of source , destination and date in specific format shown above in fig 1.

Listing 1: Python example

```
from time import sleep, strftime
from random import randint
from selenium import webdriver
from selenium.webdriver.common.keys import Keys
import smtplib

import csv

chromedriver_path = 'C:\\Users\\Chandragupta\\Desktop\\sna\\scraping2\

driver = webdriver.Chrome(executable_path=chromedriver_path)
sleep(2)
```

```
ease_my_trip="https://www.easemytrip.com/"
driver.get(ease_my_trip)
sleep(3)

Airports = ["IXA-Agartala", "AGX-Agatti Island", "AGR-Agra", "AMD-Ahme
    "IXD-Allahabad", "IXV-Along", "ATQ-Amritsar", "IXU-Aurangabad", "IX
    "BLR-Bangalore", "BEK-Bareli", "IXG-Belgaum", "BEP-Bellary", "BUP-
    "BHO-Bhopal", "BBI-Bhubaneswar", "BHJ-Bhuj", "BKB-Bikaner", "PAB-B
    "CBD-Car Nicobar", "IXC-Chandigarh", "CJB-Coimbatore", "COH-Cooch
    "DAE-Daparizo", "DAI-Darjeeling", "DED-Dehra Dun", "DEL-Delhi", "D
    "DHM-Dharamsala", "DIB-Dibrugarh", "DMU-Dimapur", "DIU-Diu", "GAY-
    "GUX-Guna", "GAU-Guwahati", "GWL-Gwalior", "HSS-Hissar", "HBX-Hubl
    "IDR-Indore", "JLR-Jabalpur", "JGB-Jagdalpur", "JAI-Jaipur", "JSA-
    "JGA-Jamnagar", "IXW-Jamshedpur", "PYB-Jeypore", "JDH-Jodhpur", "JH
    "IXQ-Kamalpur", "IXY-Kandla", "KNU-Kanpur", "IXK-Keshod", "HJR-Kha
    "KLH-Kolhapur", "KTU-Kota", "CCJ-Kozhikode", "KUU-Kulu", "IXL-Leh"
    "LUH-Ludhiana", "MAA-Madras (Chennai)", "IXM-Madurai", "LDA-Malda"
    "MZA-Muzaffarnagar", "MZU-Muzaffarpur", "MYQ-Mysore", "NAG-Nagpur"
    "NVY-Neyveli", "OMN-Osmanabad", "PGH-Pantnagar", "IXT-Pasighat", "
    "PNY-Pondicherry", "PBD-Porbandar", "IXZ-Port Blair", "PNQ-Pune", "
    "RJA-Rajahmundry", "RAJ-Rajkot", "RJI-Rajouri", "RMD-Ramagundam", "
    "REW-Rewa", "RRK-Rourkela", "RUP-Rupsi", "SXV-Salem", "TNI-Satna",
    "IXS-Silchar", "SLV-Simla", "SXR-Srinagar", "STV-Surat", "TEZ-Tezp
    "TRV-Thiruvananthapuram", "TRZ-Tiruchirapally", "TIR-Tirupati", "TC
    "BDQ-Vadodara", "VNS-Varanasi", "VGA-Vijayawada", "VTZ-Visakhapatna
```

The above Python code , we have imported selenium and use chrome web-browser. we also have make a list of airports exactly in the form use by URL. after that we have a function start_ease(), which take three argument city_from, city_to and date. the other part of code is given below.

Listing 2: Python example

```
def start_ease(city_from, city_to, date):
ease_my_trip = "https://flight.easemytrip.com/FlightList/Index?srch=" \
               + city_from + "-India|" \
               + city_to + "-India|" \
               + date \
               + "&px=1-0-0&cbn=0&ar=undefined&isow=true&isdm=true&lng=&ut
driver.get(ease_my_trip)
sleep(randint(5, 8))

flag = True
i = 1
while flag == True:
    xp_sections = '//*[@id="ResultDiv"]/div/div/div[3]/div[' + str(i) + ']
    i = i + 1
    sections = driver.find_elements_by_xpath(xp_sections)
    sections_list = [value.text for value in sections]
    if sections_list[0] == '':
```

```
            flag = False
            continue
        details = sections_list[0].split("\n")
        fill_pd(details)
        print(sections_list[0].split("\n"))


def start_scrap():
    for f in Airports:
        for t in Airports:
            if f == t:
                continue
            start_ease(f, t, "24/1/2021")


start_scrap()
```

the scrap data is stored in a CSV format. We found that only 83 airports running. Apart from data collection of flights, we also needed the longtitude and latitude of each airports for visualization purpose. we run web scrapping on gps-coordinate.net for this purpose. Data is collected for 24 Dec 2020.

4 Analysis and Visualization

For analysis and visualization purpose we use Gephi software. We study various measure of networks like degree distribution , centrality ,modularity classes, cost analysis etc.

4.1 Non-stop flights

Indian air network have currently 83 active airports which account for 2084 edges in between them. Air traffic of India is shown in fig 2. Size of the node text is base on the total degree of node. we found that the flights that have 1-stop or more is nothing more but the combination of these non-stop flights. Hence we have excluded those data from our further analysis.

4.2 Degree Distribution

In the study of graphs and networks, the degree of a node in a network is the number of connections it has to other nodes and the degree distribution is the probability distribution of these degrees over the whole network.The flight network is a directed graph which have In-degree i.e arrivals, Out-degree i.e departure and total degree. The distribution plots and data is shown in figure 3. From the degree distribution and data we can see that Delhi is the most busiest airport in India 56% more than second busiest airport(Bangalore)

4.3 Betweeness Centrality

Betweenness centrality measures the extent to which a vertex lies on paths between other vertices. Vertices with high betweenness may have considerable influence within a network by virtue of their

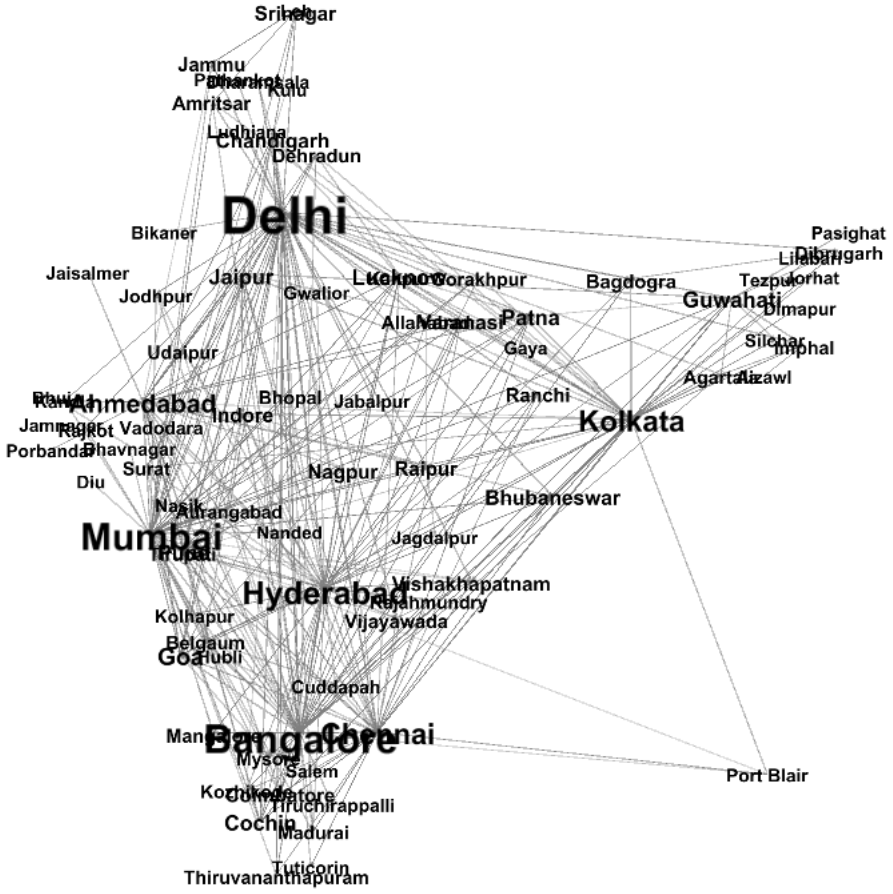


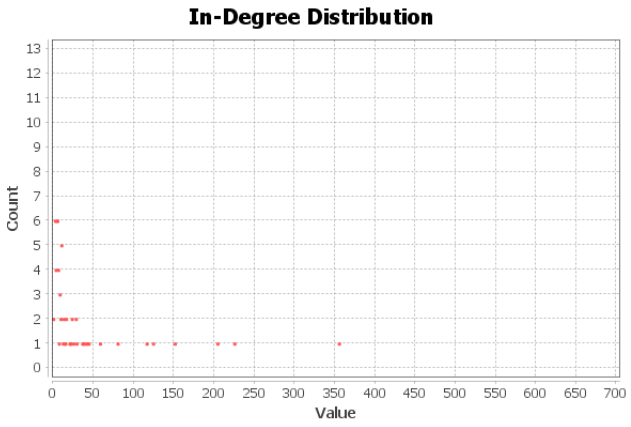
Figure 2: Non-stop flights

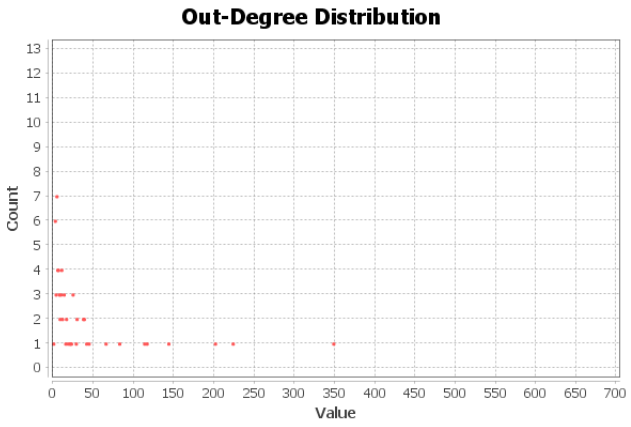
control over the connection between others. They are also the ones whose removal from the network will most disrupt communications between other vertices because they lie on the largest number of paths taken by flight.

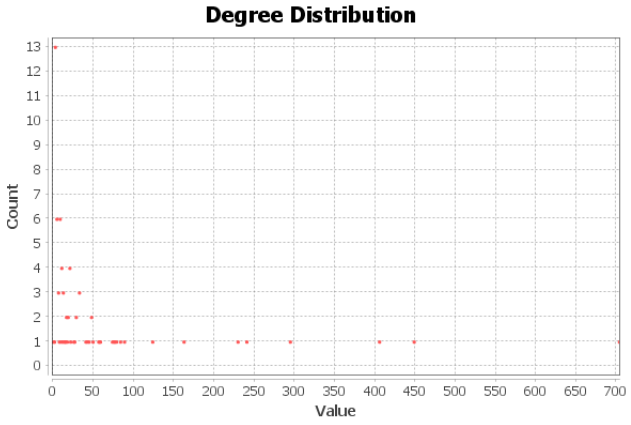
Our analysis show that Delhi have highest betweenness centrality which is 2.5 times more than second highest Mumbai. There are 7 to 8 airports with high betweenness centrality as shown in fig 4. These data shows how much Indian air network depend on these airport, any disturbance in there airport will have adverse effect on the network.

4.4 Modularity Class

Modularity measure the strength of division of a network into modules (also called groups, clusters or communities). Networks with high modularity have dense connections between the nodes within modules but sparse connections between nodes in different modules. Modularity is often used in optimization methods for detecting community structure in networks. We found main 3 modularity class in air network. These 3 modular class have high dense network within and sparse connection between other class. These three sub networks are shown in figure5. The node size represent betweenness of node in network. This show that Delhi and Mumbai has play key role in the its







Id	Degree	In-Degree	Out-Degree
Delhi	703	355	348
Bangalore	448	225	223
Mumbai	405	204	201
Hyderabad	294	151	143
Kolkata	240	124	116
Chennai	229	116	113
Ahmedabad	162	80	82
Goa	123	58	65
Guwahati	88	44	44
Lucknow	83	42	41
Patna	78	41	37
Jaipur	76	38	38
Pune	75	37	38

Figure 3: degree distribution

sub network. if these two airport get disturbed due to some reason the than it will adversely affect airports within sub-network compare to other airports in other sub-network. Similarly Bangalore and Kolkata play key role in sub-network and Hyderabad and Chennai in sub-network 3

4.5 Cost Analysis

we analysis the network on the basis of cost of flight. we found the routes which have high cost i.e greater than ₹ 10,000 and cheap i.e less than ₹ 3,000 as shown in figure 6. It can be seen that costly flights are the long routes and cheap flights are short routes.

4.6 Air Routes

For route analysis we create new data-set using the data-set we collected. We calculate weight of each route as the number of flights in that route. We use weight to represent the busyness of route. We found that 96.78% of route have less than 15 flights and 3.22% of routes have more than 15 flights. The maximum flight in route was 31. These route are shown in figure7 .

4.7 Connectivity

Connectivity shows number of airports one airport is connected to through direct flight. This is given by out-degree of weighted network as shown in figure 8. Delhi is connected to most of the airports i.e 53 followed by Bangalore, Mumbai,Hyderabad with airports 47,45 and 44 respectively.

Id	Betweenness Centrality	Closeness Centrality	Clustering Coefficient
Delhi	2596.717851	0.733945	0.279715
Mumbai	1090.402924	0.683761	0.355967
Bangalore	933.390414	0.689655	0.346907
Kolkata	747.600544	0.615385	0.378976
Hyderabad	734.045035	0.672269	0.429783
Chennai	443.047981	0.615385	0.403771
Ahmedabad	399.54562	0.559441	0.470315
Guwahati	355.613255	0.536913	0.546769
Imphal	159.0	0.451977	0.6
Bhavnagar	159.0	0.412371	0.0
Chandigarh	84.554457	0.522876	0.458824
Pasighat	80.0	0.262295	0.5
Tezpur	79.0	0.352423	0.5
Raipur	19.898439	0.509554	0.616923
Goa	13.57048	0.526316	0.590258
Pune	7.141887	0.503145	0.646667
Cochin	4.963848	0.512821	0.601504
Vijayawada	3.37946	0.473373	0.529412
Belgaum	3.208379	0.45977	0.236364
Jaipur	2.181903	0.529801	0.617073
Indore	1.766197	0.519481	0.623016
Lucknow	1.493805	0.522876	0.648485
Bagdogra	0.603418	0.479042	0.629167

Figure 4: Betweenness Centrality

5 Various Aspects of civil aviation and analysis of the data

5.1 UDAN (Ude Desh ka Aam Naagrik)

UDAN (Ude Desh ka Aam Naagrik) is a brainchild of GOI wherein it promotes regional connectivity of Indian cities at very low prices.

- Total target current airport: 486

Statistics of UDAN last year			
UDAN flights Operated	Total number of Pax carried in UDAN flight	No. of UDAN Pax seats sold	No. of Non UDAN Pax seats sold
86,425	47,00,698	27,36,843	19,61,018

Criticisms of UDAN : It is very difficult to low cost Air Carriers to sustain flights to these regions as per the data released from AAI. As per one of the publicly available data we could see lot of sectors have not even started operating because subsidies provided to low cost airlines is still not helping them gain operable Profits.
Thus even if Bigger Aiports like Chennai are used to subsidise the aiports under regional connectivity they still need to be self reliant in short term to be able to continue to exist

5.2 Relationship between GDP and footfalls in airports of a region

For long airports are considered a middle-rich income group mode of transportation although in near future it might not be true.
We found that footfalls in airport has good relationship with its GDP and vice versa.

GDP and footfalls of a region		
City	Per Capita(In $)	Footfalls (m)
Mumbai	4700	30
New Delhi	4623	42
Kolkata	3718	15
Bengaluru	3375	23
Hyderabad	2353	14.8
Chennai	2514	14.21
Pune	2456	6.98
Ahmedabad	2771	7.6
Surat	2610	1.21
Visakhapatnam	2006	1.92

We plotted these data in matplotlib in a scatter plot and there was slight co relation between them just like a linear regression.In cant be the only factor as we know that there are other factors like cultural and historical preference of a place, demographics etc. Thus we can safely assume that these two factors are co-related.

5.3 International Relation Predictions based on Airport Flight Data

- Country relations can also be obtained from the data that we have scraped
- Poor relationships reflects less flights between countries.More than 130 fights operate between Indian cities and Tel Aviv wheras there is no flights between India and Pakistan. These are examples of strong liking and strong hatred towards each other

We all know by the principle of strongly connected graph that an unstable order will create problems like War. For example if in a triad we have "++-" relation then it will unstable as node having good relation with 2 nodes who have poor relationship amongst them will have a hard time.
Stable Triad :
There is no flight operations between India and Israel to Pakistan.This is a stable order and it never changes.
Unstable Triad:
Consider the case of UAE, Israel and India .India has better relations with both and needs to appease both countries to maintain relations from time to time. Ditto with India, US and USSR axis. Anytime India tries to do a deal with US or USSR the other one tries to sanction it.
Recent Example : India was dared with Sanctions when it dealt with Russia over S-400 'Triumf' Air Defence Systems.
India US and Russia Flight patterns show similar results. Although US and Russia flights are not no existent but they do not reflect true potential between two behemoth economies and seems more like a token gesture.

5.4 Effects of airports on Bird Ecosystem

We had assumed at the beginning of this project that flights might have effect on bird migration but during our research we found this to not to be completely correct.

We went to websites ebird.org to find if there is any disturbance reported in places where airport frequencies have changed but found no such support.More over birds generally don't fly at airport cruising height so that's why they have most of their effect limited in vicinity of airports. There is no website from any Wildlife conservation or any government in official capacity which quotes these data.
The only effects are changes in energy conversion, time of food intake and increase in heart rate.

Disturbed reactions to flight reported for other waterfowl, great bustards, predatory birds and crows.

It has been found that bird hits cause damage to the tunes of billions of $ to aviation industry on upkeep and maintenance. Wetland reclamation efforts have also led to aggravated levels of dangers.Airports use techniques like hazing,bird scattering using loud noise and high frequency sounds and many other techniques

Ironically BIAL has advertised for an aviary for its new terminal so that bird population which has bewilderingly increased in its airports can be kept in check.
Case Study of Sparrows in Bengaluru as per a local daily :
It has been found that sparrows of Bengaluru have found a heavenly abode in BIAL(Bengaluru International Airport Limited) It was concluded that the grasslands on the opposite side of the airport takes care of the insect food that the sparrows may need to feed its young ones. Constant supply of food from the eateries offer an important source of food.

6 Conclusion

- Regional Airports in India are unsustainable because of extremely low footfalls and almost being outliers in civil aviation ecosystems.
- Most airports are financially unstable currently
- The breakeven point changes 0.8 million passengers in 2014-15 to 0.6 million passengers in 2016-17. Most regional airports in the sample have more than 0.5 million passengers per annum and this paves way for smaller and upcoming airports looking for incentive schemes to attract airlines.
- We feel that schemes like UDAN will surely bring about economic success to areas which were not connected well by air traffic before.
- Footfalls and GDP of a region follow a vicious circle.
- The GDP of a region permits region affords it a regional airport which in turn increases the GDP of a region through the fastest mode of accessibility which in turn helps push the GDP further.
- It helps take a region take the next leap in development and the faster the breakeven point in an airport development reaches better it is.

References

- https://www.sciencedirect.com/science/article/pii/S209252121730041X[[Analysis of the Air Transport Network Characteristics of Major Airports]
- A Set of Measures of Centrality based on Betweenness by L.C. Freeman
- https://www.aai.aero/sites/default/files/traffic-news/Aug2k19Annex3.pdf
- Breakeven Passenger Traffic for Regional Indian Airports K Chandrashekhar Iyera , Soumya Jain
- https://www.aai.aero/en/corporate/resources
- https://www.natureforever.org/blog/2014/01/21/vip-sparrows-of-bangalore-airport/

- https://en.wikipedia.org/wiki/List$_o f_I ndian_c ities_b y_G DP_p er_c apitattps : //www.fai.org/sites/default/files/1_a ircraft_e ffects_o n_b irds.pdf$

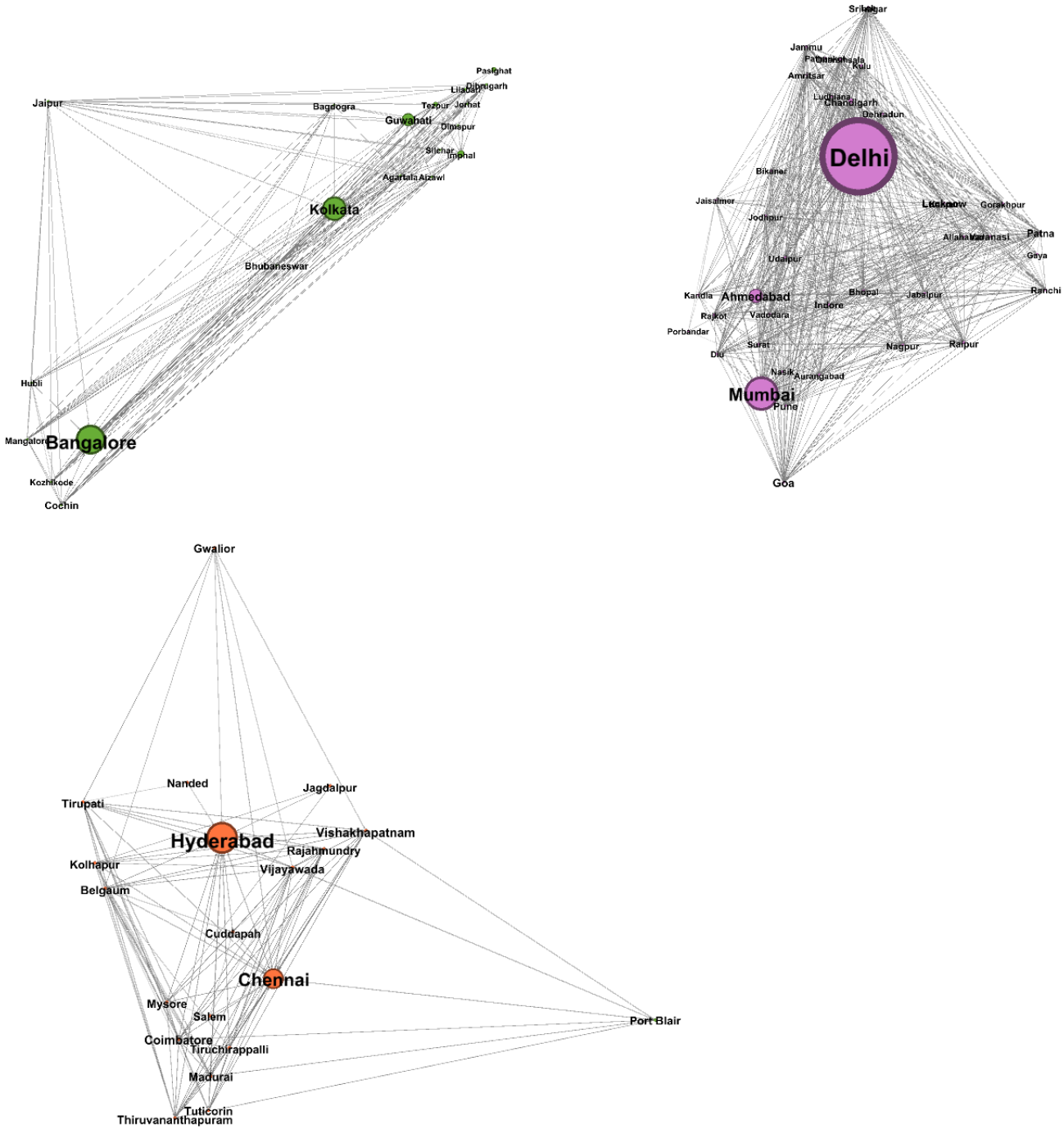


Figure 5: Modularity Class

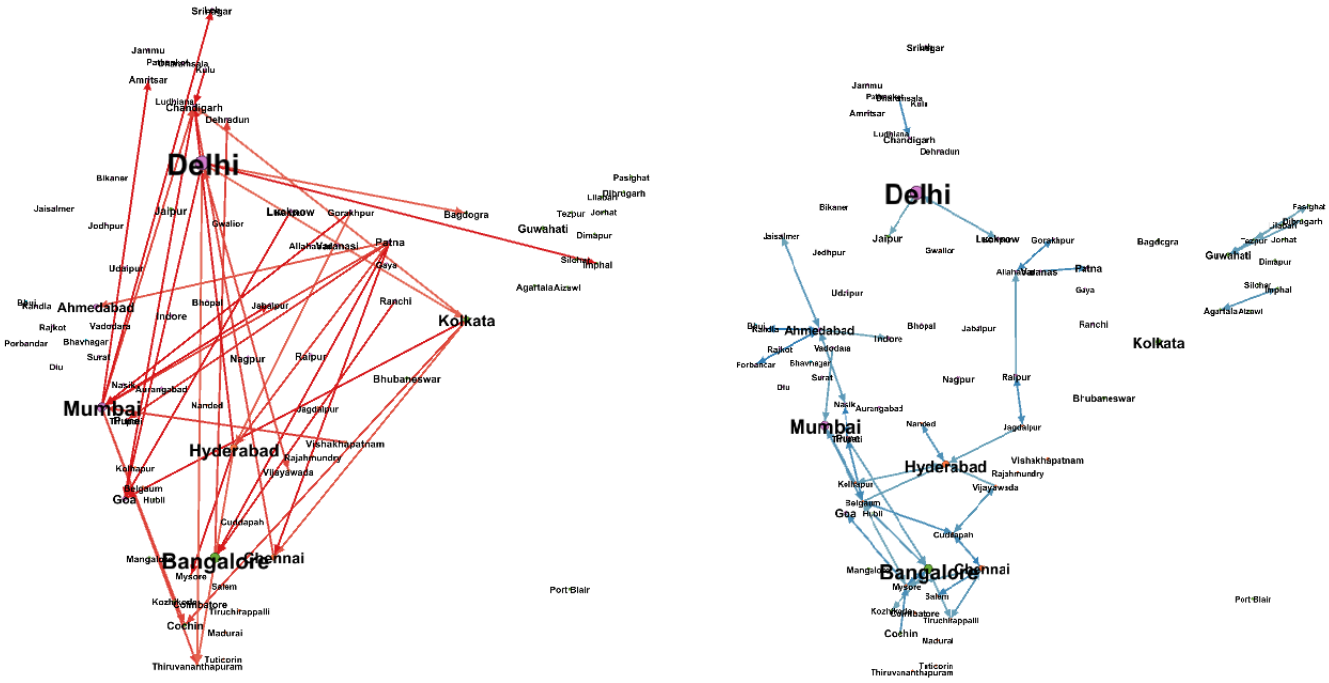


Figure 6: Flight cost: costly(left) & cheap(right)

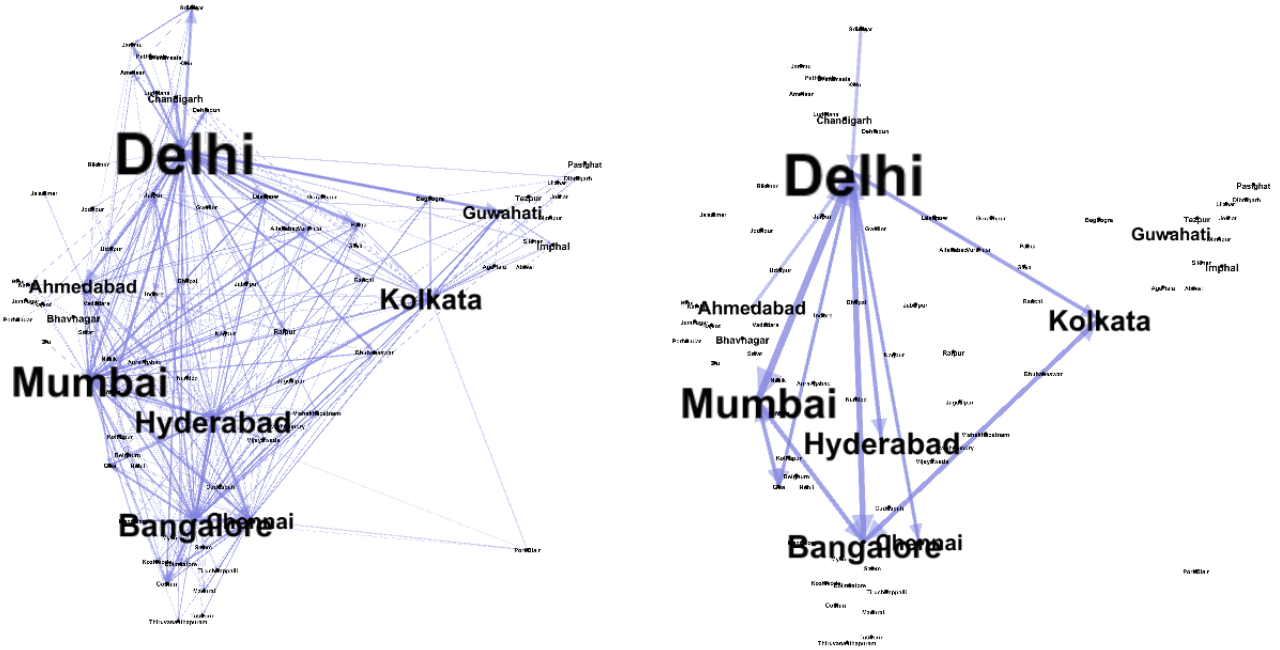


Figure 7: Busyness of Route: at-most 15 flight(left) , at-least 15 flights(right)

Id	Degree	In-Degree	Out-Degree
Delhi	107	54	53
Bangalore	95	48	47
Mumbai	90	45	45
Hyderabad	88	44	44
Chennai	68	35	33
Kolkata	63	31	32
Ahmedabad	52	26	26
Pune	29	15	14
Guwahati	26	12	14
Chandigarh	26	14	12
Goa	26	13	13
Jaipur	24	12	12
Lucknow	24	12	12
Indore	23	12	11
Patna	21	11	10
Cochin	18	9	9
Raipur	18	9	9
Varanasi	18	9	9
Nagpur	17	8	9
Bhubaneswar	14	7	7
Allahabad	14	7	7
Belgaum	14	7	7
Vishakhapatnam	14	7	7
Vijayawada	13	7	6
Coimbatore	12	6	6
Mysore	12	6	6
Surat	12	6	6
Dehradun	11	5	6
Bagdogra	11	6	5

Figure 8: Connectivity