Course Project Report

COMP 6917 (Fall 2017)

Asma Javaid - 201790860
Mesam Timmar - 201692692

US Airports Centralities

ABSTRACT

This project analyzes the United States airport network. In this project, we determined the most important airports by considering domestic and international flights to and from United Stated airports. For this analysis we implemented the techniques proposed in [1] to the airports network dataset from year 2016 and first quarter of 2017. After obtaining the required information we identified most imporatant airports in this network. Finally we extented this work to study the variation in importance of airports in different months.

1 INTRODUCTION

In this report we will first look into some of the related work done on airports network. After this we will see the methodology we used for analysis of the network in our project. Then we describe the data set and source of this data. Then we will briefly compare the objectives proposed by us and the actual achieved goals. After this we will examine the results we obtained by executing our program on these datasets. For discussion of our results we will start with results for closeness centrality, then we will discuss the most imporant airports in terms of In Degree and Out Degree centralities. Finally we will see the variation in the importance of the aiports with changing months.

2 RELATED WORK

Lots of work has been done previously on the centrality measures of airport network of different countries including US airports in different context. Like in [2] using graph theoretical approach authors of this paper not only find critical airports which have high connections but also find the airports which have high betweenness and closeness centrality. Authors of [2] analysis the importance and significance of US airports for the first quarter of year 2010 by using the algorithms that are used in network science. Some work have been done on Indian Airports Network where they used the weighted network to examine the importance of the airports. Our work is different form these previous projects as we have used the technique proposed in [1] to study in importance of weighted airports network.

3 METHODOLOGY

Since node degree and strength are both indicators of the level of involvement of a node in the surrounding network, it is important to incorporate both these measures when studying the centrality of a node [1]. So, in an attempt to combine both degree and strength, they introduce a tuning parameter alpha, which determines the relative importance of the number of ties compared to tie weights. They proposed a new formula for degree centrality for the nodes associated with weights using the turning parameter α by

$$C_D^{w\alpha}(i) = k_i \times \left(\frac{s_i}{k_i}\right)^{\alpha} = k_i^{(1-\alpha)} \times s_i^{\alpha}$$

They also proposed a generalized formula for closeness and betweenness centrality using this turning parameter alpha as

$$C_C^{w\alpha}(i) = \left[\sum_j^N d^{w\alpha}(i,j)\right]^{-1}$$

and

$$C_B^{w\alpha}(i) = \frac{g_{jk}^{w\alpha}(i)}{g_{jk}^{w\alpha}}$$

4 DATA SET

Number of Nodes: 1484 **Number of Edges: 106857**
Nodes: Airports

Description of Dataset:
We obtained data set for this project from website of U.S. Bureau of Transportation Statistics. This is the data for all the flights to and from all airports of US for year 2017 (we obtained data for first 4 months of 2017). In this data we have 1484 nodes. Nodes are airports including all US airports and also other international aiports with which there is a flight to and from US. There are 106857 directed edges. There is an edge from node A to node B for a particular month if there is atleast one flight form airport A to aiport B in that month. These edges will have weights. These weights will depend on type of centrality measure we are interested in for example in case when we want to know the importnace

of an aiport with respect to number of passengers travelled then we will consider weight of edge from A to B as the number of passegers that took flight from airport A to aiport B. We also worked on the dataset of complete year 2016. We used this data for comparing the importance of airports in different months of an year.

5 GOALS ACHIEVED

We proposed that we will achieve following goals:

- Centrality of an aiport based on number of airports its connected to and the passengers travelling through that route.
- Closeness Centrality of all airports considering inverse of distance as weight.
- Betweeness Centrality of all airports aslo by considering inverse of distance as weight.
- We will see trends in the importnace of an airport in terms of number of passengers travelling through that airport and number of connections with other airports for different months of an year.

What we actually achieved

- Centrality of an aiport based on number of airports its connected to and the passengers travelling through that route.
- Closeness Centrality of all airports considering inverse of distance as weight.
- We saw trends in the reletive importnace of an airport in terms of number of passengers travelling to that airport for different months of an year [1].

We didn't analyse the betweeness centrality as we wanted to focus more on the variation in importance of airports. Geting information about the betweeness centralities of airports would not have lead us to any specific knowledge we were targeting in this project.

6 RESULTS

6.1 CLOSENESS

We used the methodology described above to compute the closeness centralities of all the airports. In this case we have considered the distances between the airports as the weights of the directed edges. We obtained results for different alpha values. Following table shows the results. For alpha = 0, it porduces the same outcome as binary distance measure. When

[1] for this case we considered the flights from Janaury 2016 to Decemberl 2016.

alpha = 1, the outcome is the same as the one obtained with Dijkstras algorithm. It will produce the same distance score for paths with different number of intermediary nodes if the wieghts on the edges adds up to same value. For alpha < 1, it assigns the path with the greatest number of intermediary nodes the longest distance. Hence, for alpha < 1, a path with lest intermediate nodes i.e direct flights would be favored over a longer path with more intermediate nodes. Conversely, when alpha > 1, the impact of additional intermediary nodes is relatively unimportant compared to the strength of the ties and paths with more intermediaries are favored.

$C_D^{W\alpha}$ when α =			
0	0.5	1	1.5
Atlanta, GA	Chicago, IL	Bethel, AK	Nunapitchuk, AK
Chicago, IL	Newark, NJ	Nunapitchuk, AK	Kasigluk, AK
Los Angeles, CA	Atlanta, GA	Napaskiak, AK	Bethel, AK
Miami, FL	New York, NY	Kasigluk, AK	Napaskiak, AK
Dallas/Fort Worth, TX	White Plains, NY	Napakiak, AK	Port Graham, AK
Newark, NJ	Detroit, MI	Atmautluak, AK	Nanwalek, AK
Houston, TX	Philadelphia, PA	Akiachak, AK	Atmautluak, AK
New York, NY	Teterboro, NJ	Kwethluk, AK	Napakiak, AK
Denver, CO	Indianapolis, IN	New York, NY	Akiachak, AK
Las Vegas, NV	New York, NY	Akiak, AK	Kwethluk, AK

6.2 IN DEGREE

We considered number of passengers travelling to a particular aiport from another airport as the weight of the directed edge between these two airports. The table below shows the importance of aiports in terms of number of passengers travelling towards them and also considered the number of imediate neighbors of the airport in measurement of the strength of that airport. If alpha parameter is between 0 and 1, then having a high degree is taken as favorable, whereas if it is set above 1, a low degree is favorable.

$C_D^{W\alpha}$ when α =			
0	0.5	1	1.5
Atlanta, GA	Atlanta, GA	Atlanta, GA	Atlanta, GA
Chicago, IL	Los Angeles, CA	Los Angeles, CA	Los Angeles, CA
Los Angeles, CA	Chicago, IL	Chicago, IL	Chicago, IL
Miami, FL	Dallas/Fort Worth, TX	Dallas/Fort Worth, TX	Dallas/Fort Worth, TX
Dallas/Fort Worth, TX	New York, NY	New York, NY	Denver, CO
Newark, NJ	Denver, CO	Denver, CO	New York, NY
New York, NY	Miami, FL	San Francisco, CA	San Francisco, CA
Houston, TX	Newark, NJ	Orlando, FL	Phoenix, AZ
Denver, CO	Las Vegas, NV	Las Vegas, NV	Orlando, FL
Washington, DC	Houston, TX	Phoenix, AZ	Las Vegas, NV

6.3 OUT DEGREE

We considered number of passengers travelling from a particular aiport to another airport as the weight of the directed edge between these two airports. The table below shows the importance of aiports in terms of number of passengers travelling from an airport and also considered the number of imediate neighbors of the airport in measurement of the strength of that airport. If alpha parameter is between 0 and 1, then having a high degree is taken as favorable, whereas if it is set above 1, a low degree is favorable.

$C_D^{W\alpha}$ when α =			
0	0.5	1	1.5
Atlanta, GA	Atlanta, GA	Atlanta, GA	Atlanta, GA
Chicago, IL	Los Angeles, CA	Los Angeles, CA	Los Angeles, CA
Miami, FL	Chicago, IL	Chicago, IL	Chicago, IL
Los Angeles, CA	Dallas/Fort Worth, TX	Dallas/Fort Worth, TX	Dallas/Fort Worth, TX
Dallas/Fort Worth, TX	New York, NY	Denver, CO	Denver, CO
New York, NY	Denver, CO	New York, NY	New York, NY
Newark, NJ	Miami, FL	San Francisco, CA	San Francisco, CA
Houston, TX	Newark, NJ	Las Vegas, NV	Phoenix, AZ
Denver, CO	Las Vegas, NV	Orlando, FL	Orlando, FL
Washington, DC	Houston, TX	Phoenix, AZ	Las Vegas, NV

Both in degree and out degree results show that Atlanta International Airport is the most important airport in all situations followed by Chicago OHare International and Los Angeles International Airport.

6.4 IMPORTANCE ON MONTHLY BASIS

By computing In degree centralities with alpha value 1 for each month's data seperately we obtained the trends in the reletive importance of an airport in differnet months of the year. Following tables show the results we obtained.

January	Febuary	March	April
Atlanta, GA	Atlanta, GA	Atlanta, GA	Atlanta, GA
Los Angeles, CA	Los Angeles, CA	Los Angeles, CA	Los Angeles, CA
Chicago, IL	Chicago, IL	Chicago, IL	Chicago, IL
Dallas/Fort Worth, TX	Dallas/Fort Worth, TX	Dallas/Fort Worth, TX	Dallas/Fort Worth, TX
New York, NY	Denver, CO	New York, NY	New York, NY
Denver, CO	New York, NY	Denver, CO	Denver, CO
San Francisco, CA	San Francisco, CA	Phoenix, AZ	San Francisco, CA
Miami, FL	Las Vegas, NV	San Francisco, CA	Charlotte, NC
Phoenix, AZ	Phoenix, AZ	Orlando, FL	Las Vegas, NV
Las Vegas, NV	Miami, FL	Las Vegas, NV	Orlando, FL

May	June	July	August
Atlanta, GA	Atlanta, GA	Atlanta, GA	Atlanta, GA
Chicago, IL	Los Angeles, CA	Los Angeles, CA	Los Angeles, CA
Los Angeles, CA	Chicago, IL	Chicago, IL	Chicago, IL
Dallas/Fort Worth, TX	Dallas/Fort Worth, TX	Dallas/Fort Worth, TX	New York, NY
New York, NY	New York, NY	New York, NY	Dallas/Fort Worth, TX
Denver, CO	Denver, CO	Denver, CO	Denver, CO
San Francisco, CA	San Francisco, CA	San Francisco, CA	San Francisco, CA
Charlotte, NC	Seattle, WA	Seattle, WA	Seattle, WA
Las Vegas, NV	Las Vegas, NV	Las Vegas, NV	Las Vegas, NV
Seattle, WA	Charlotte, NC	Charlotte, NC	Charlotte, NC

September	October	November	December
Atlanta, GA	Atlanta, GA	Atlanta, GA	Atlanta, GA
Chicago, IL	Chicago, IL	Los Angeles, CA	Los Angeles, CA
Los Angeles, CA	Los Angeles, CA	Chicago, IL	Chicago, IL
Dallas/Fort Worth, TX	Dallas/Fort Worth, TX	Dallas/Fort Worth, TX	Dallas/Fort Worth, TX
New York, NY	Denver, CO	Denver, CO	Denver, CO
Denver, CO	New York, NY	New York, NY	New York, NY
San Francisco, CA	San Francisco, CA	San Francisco, CA	San Francisco, CA
Las Vegas, NV	Las Vegas, NV	Las Vegas, NV	Miami, FL
Seattle, WA	Charlotte, NC	Phoenix, AZ	Orlando, FL
Charlotte, NC	Seattle, WA	Charlotte, NC	Phoenix, AZ

7 CONCLUSION

Results we obtained shows that there are few airports whose reletive importance changes significantly over the year. For example importnace of Miami International Airport is increased during winter months December, January and Feburary. Atlanta International Airport stays the busiest/most important airport throughout the year followed by Chicago OHare International and Los Angeles International Airport. Generally there is no significance change in reletive importance of an airport with change in months except few cases.

References

[1] Opsahl, T., Agneessens, F. and Skvoretz, J., 2010. Node centrality in weighted networks: Generalizing degree and shortest paths. Social networks, **32**(**3**), pp.245-251.

[2] Sapre, M. and Parekh, N., 2011, June. Analysis of Centrality Measures of Airport Network of India. In PReMI (pp. 376-381)

[3] Conti, E., Cao, S. and Thomas, A.J., 2013. Disruptions in the US airport network. arXiv preprint arXiv:1301.2223.