International Journal of Engineering & Technology, 7 (2) (2018) 708-713

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET
doi: 10.14419/ijet.v7i2.9915
Research paper

Analysis and visualization of airport network to strengthen the economy

Saleena. P [1], P. K. Swetha [1], D. Radha [1*]

[1] Department of Computer Science & Engineering, Amrita School of Engineering, Bengaluru2, Amrita Vishwa Vidyapeetham, India
**Corresponding author E-mail: d_radha@blr.amrita.edu*

Abstract

The world's eminent airports are directly or indirectly connected to many other airports. Every airport is considered as a node and the route can be considered as edge connecting them. The work analyzes the USA airport network using different centrality measures of social network analysis. The centrality measures calculated on airport network help in identification of certain characteristics of the airports. Some of the characteristics are like the busiest airport and the airports which influence trade, alternate path, fastest route, nearest airports, etc. The characteristics helps to find the designated airports meant for improving the economy. The results of this paper say about the prominent communication and connections among the airports in the U.S.A. The tools used for the analysis are UCINET 6 and NetDraw.

Keywords: *Airport Network; Centrality Measures; Designated Node; Social Network Analysis; Visualization.*

1. Introduction

Airport network has an important role in improving the economy of the country. The analysis of such network results in designating an airport for better economy. There are various tools and ways to analyze the airport network like the other networks. The airport network can be analyzed in the way a social network is analyzed. This type of analysis also leads to better understanding of the characteristics of the airport and identification of designated airport for better economy for the country.

Social network has an important role in the communication in the present world. As a human, everybody is socially connected with each other in many aspects. So in day to day life scenario, social networks have more important roles. Social networks analyze the relationships between the organizations, groups or individuals by using a set of methods and the empirical structure expressed in the form of network. Some interesting social networking sites are Facebook, Pininterest, Twitter, Linkedin etc. So by taking these networks and analyzing its functionality with graph theory is termed as social network analysis [SNA]. The analysis of SNA includes both visual and mathematical views of relationships. Other than social networks, there are some networks, called Biological networks, Geo physical networks, Ecological networks and financial networks. Protein Interaction Networks, Metabolic Network, Gene Networks are types of Biological networks. The Internet and World Wide Web can be under Interaction networks. There are some social networks, which are widely used by common people. These networks connect the people in the world. Some of the modern and much used social networks that are meant for various purposes are Facebook which builds the connections with the people and organization and share the information like photos, videos, views, advertise etc… [1]. Quora is a question and answer network site which helps the user to create the profile and can post the questions for which they don't know the appropriate answers. Flickr is one of the photo sharing and photo management application. The users create their profile and can share their best pictures to the world which are captured by camera or can share privately with some known persons. Whatsapp is one of the popular instant messaging service applications for the smart phone where users can send messages, pictures and other types of media to others. Users can also make phone and video call to others. Classroom2.0 is one of the professional social networks designed for teachers to communicate with other professionals all over the world and can share their materials to others. The Math Forum is an educational social network designed for the students who are interested in mathematics. In this, students interact with others on academic projects and research via blogs and forums. My Place at Scrapbook.com is a hobby focused social network designed for users who have interest on scrap booking. In this site, users create their profile and share their craft works and can learn by watching the videos posted by other users. LinkedIn is a social network and is designed for business and employment services like posting jobs and job seekers posting their CVs. Users can create profiles and connections to each other's and can invite anyone to be connected.

The social networks are analyzed for various centrality measures to understand the importance or characteristics of people or users in the network. The various centrality measures are degree centrality, Betweenness centrality, closeness centrality Eigen vector centrality etc...

2. Literature review

Social network analysis is the process of analyzing the social structures with the help of graph theory and its network structure. So networks consist of nodes and edges which is the connection between the nodes. Understanding the relationship between the nodes and by analyzing the edges or ties, more information can be observed about the social network. There are so many advantages with respect to social analysis. If anybody wants to know more about any communities and their strategies, these types of analysis help to provide more information. The analysis results try to figure out the

Copyright © 2018 Saleena. P et. al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

weak and strong entities in the given network easily and this can be proceeded to other measurements flawlessly.
The analysis of social networks helps to determine and solve problems regards with it. Review of Social networks in mapping flows of a knowledge get analyzed by a case study of pharmaceutical company [2]. A recommendation system to social network by avoiding Sybil attack and ensured with privacy described by analyzing social network with Automated Dynamic Grouping system model [3]. An anonymization for a given social network can be achieved by the method of restricting access to sensitive data by analyzing its closeness centrality [4].
Analysis of social media network and its visualization is done by UCINET tool and NetDraw [5]. This paper discussed about various social networks in the real world and their applications. UCINET has facility to characterize social networks.
Researchers showed interest on analyzing all types of networks related to real life problems. An analysis of wheat trade [6] construes bargaining power of countries in export and import perspective. Here actors represent the countries and in-degree and out-degree represents the import and export respectively of a specific country. Bonacich measure expresses the power of a country in terms of wheat trade and flow centrality used to find the flow of wheat among countries. Similarly, the social network in online forum (BBS) of university analysis [7] contains the visualization of BBS network by matrix form and a graphical representation called sociogram. And also it showed the analysis of the network by centralities of a node and the whole network using data mining methods.
A large number of measures can be applied on any given social network. The main measurements done in this area are the Node centrality of a Road network by Degree, Closeness and Betweenness centrality [8]. Here centrality is used to rank the nodes of the Road network. Degree centrality refers to the ways in which roads are connected at a junction point, Closeness indicates the shortest path and global connectivity of the network and Betweenness indicates the control behavior of one node which has no direct connectivity to a particular node. Analysis of Urban Rail Network with Centrality Characteristics [9] is explained with degree based index that defines the travel activities of a traveler who can reach station, Betweenness index describes the ability of a station to control the travel activities and closeness index describes efficacy of a station. A world city network of cities [10] explains about the measurements on a Chinese air transportation which includes density analysis to predict the connections between major air hubs, Subgroup analysis by the CONCOR method with density matrix and cluster dendrogram.

2.1. Types of data sets

Datasets are available in plenty for analyzing social networks. The data sets are like yeast network, co-author ship networks, transport network, friendship communities etc. Since the focus of this paper is to understand about transportation, some of the transport network repositories are:

- Pajek [11] which consist of U.S airport network on the analysis is done.
- Tore Opsahl [12] also contains a dataset of 500 busiest airports in U.S with one mode and two mode networks.
- Free Geographic Information System (GIS) data in which transport and communication networks like open flights, world port index, Capitaine European train station etc. datasets are available.
- CTA Transportation Networks contains highway data, railroad network, inter mode network, waterway network etc.
- SNAP dataset consist of road networks.
- Advocacy Monitoring of Women and Children Health through Social Data [13] used Facebook and Twitter for the source of data.

3. Social network analysis

Social network analysis [SNA] is the mapping and measuring of relationships and flows between people, groups, organizations, computers or other information/knowledge processing entities. The nodes in the network are the people and groups while the links show relationships or flows between the nodes. SNA has its origins in both social science and in the wide fields of network analysis and graph theory. SNA is a unique perspective of how society works. Instead of focusing on individuals and their attributes it perturbs on relations between individuals, groups or social institutions. SNA has found enactment in many domains beyond social science like business domain, low enforcement agencies, civil society organizations, network operators etc... Key players of social network analysis are degree centrality, Betweenness centrality, Closeness centrality, Katz centrality, Eigenvector centrality, PageRank centrality, Percolation centrality, Reach centrality, Information centrality etc. Researchers engage social network analysis in the study of computer-supported collaborative learning because of its unique capabilities.
Degree centrality is the number of unswerving connections a node has in a given network. The degree can be interpreted in terms of the prompt risk of a node for catching whatever is gurgling through the network. In the case of a directed network (where ties have direction), usually elucidate two separate measures of degree centrality, namely in degree and out degree. In case of an airport network it is the total number of flights scheduled to and from a certain airport to all other airports in that transport network. Betweenness Centrality is a centrality measure of a vertex within a graph. Betweenness centrality quantifies the number of times a node acts as a bridge along the shortest path between two other nodes. A node with high Betweenness has great impact over what flows in the network stipulating important links and single point of failure. In case of an airport transport network it is the influence of a particular airport over all other enduring in a given network structure. Closeness Centrality is the degree of a discrete node near to all other solitary nodes in a network (directly or indirectly). The pattern of the direct and indirect ties entitles the nodes to any other node in the network more swiftly than other nodes and it has the shortest paths to all others. In case of an airport networks it means the nearest distance to other airports and which airport can outstretch to all other airport in a less distance or in less time. Eigenvector centrality is a measure of the impact of a node in a network. Katz centrality is an appendage to degree centrality and it appraises the path connection between all the nodes, while the handout of distant nodes is castigated. Page rank centrality is the number of outbound links in a directed graph. The Percolation Centrality is expounded for a given node, at a given time, as the proportion of a shortest path between a pair of nodes that go through that node. Reach centrality is the estimate that counts the number of nodes that each one can reach in some defined steps to all other nodes. Information centrality is the gauge that outline the information flow from one node to other nodes in a shortest path.
In order to analyze the social structure, an airport transportation network of United States of America (U.S.A) from Pajek dataset [11] is used. The network consists of 332 nodes and 4252 ties. This is a directed one-mode network where nodes constitute top most airports in U.S.A and ties entitle the connection between them by a flight. An edge exists if there is a direct flight scheduled between the two airports. In this dataset 332 topmost airports in U.S.A are considered. By using the social analysis tools like UCINET and NetDraw various measurements are applied on the U.S.A transport network. UCINET is a tool to analyze the social network with help of measurements like centrality measures, group behaviors etc... NetDraw is a visualization tool which comes along with UCINET. The measurements applied on this dataset are Degree centrality, Closeness centrality, Betweenness centrality, Information centrality, Eigen vector analysis and Reach centrality.

4. Results and analysis

The results of each analysis such as Degree centrality, Betweenness centrality, Closeness centrality, Eigen vector centrality and Reach centrality by using UCINET tool is described in detail in this section.

4.1. Degree centrality

Degree centrality of this dataset shows the number of direct flights boarded or arrived from or to a particular airport. Here Chicago O'hare Intl airport has the highest degree centrality among all other airports that is of 139.

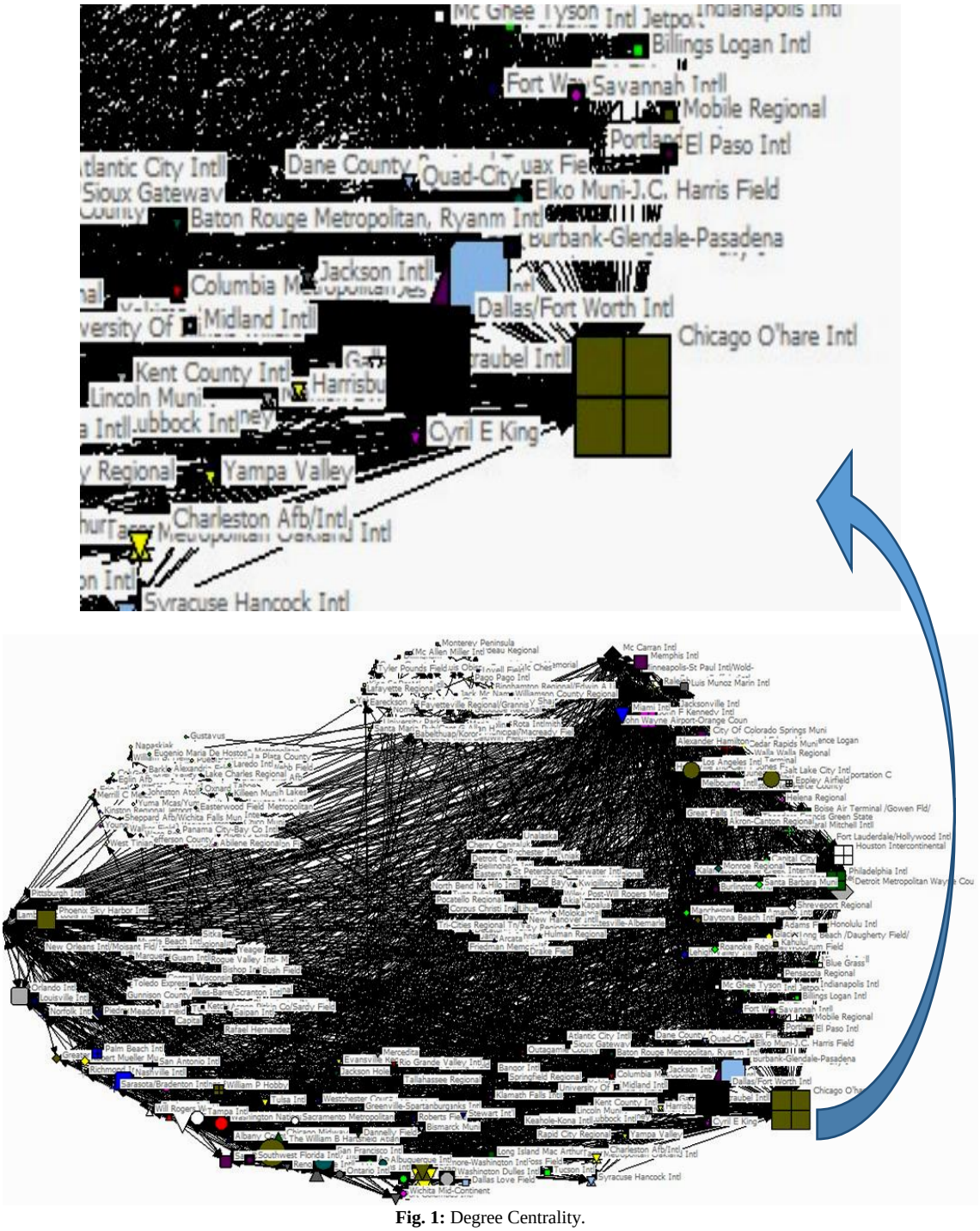


Fig. 1: Degree Centrality.

Which means total of 139 flights are boarded or departure from or to Chicago O'hare Intl. The degree centrality of whole dataset and the zoomed view of one portion are shown in the Fig.1. The node which is in larger size has the highest degree in the network i.e., Chicago O'hare Intl.

4.2. Betweenness centrality

In this dataset, Chicago O'hare Intl airport has the highest Betweenness centrality which shows that this airport has the high influence within that network. The change or disorder of this airport leads to disrupt communication between other airports, which are considered in this dataset. It acts as an intermediate airport for many airports in the U.S.A i.e..., without this airport the flights which have connection to it leads to take long path to reach the destination or it won't have any path to the destination. Therefore, this airport plays a vital role to reach the destination airport with shortest path. Fig.2 show values of the Betweenness centrality for the top 7 airports from the U.S.A transport dataset. The highest Betweenness value of the network is 11376.625, which is calculated using formula 1[14].

This formula indicates to find out how many pairs of individual nodes would have to go through in order to reach one another in the minimum number of nodes.

$$C_B(i) = \sum_{j<k}^{n} g_{jk}(i)/g_{jk} \qquad (1)$$

where $g_{jk}(i)$ represents the total number of shortest path between node j and node k that passes through node i. The term g_{jk} represents the total number of shortest paths between nodes j and k. The visualization of Betweenness centrality is given in Fig.3.

		1 Betweenness	2 nBetweenness
118	Chicago O'hare Intl	11376.625	20.831
8	Anchorage Intl	9288.358	17.007
261	Dallas/Fort Worth Intl	8366.774	15.320
201	San Francisco Intl	5146.297	9.423
47	Seattle-Tacoma Intl	5069.276	9.282
182	Lambert-St Louis Intl	4453.663	8.155
255	The William B Hartsfield Atlan	3904.955	7.150

Fig. 2: Betweenness Centrality Values.

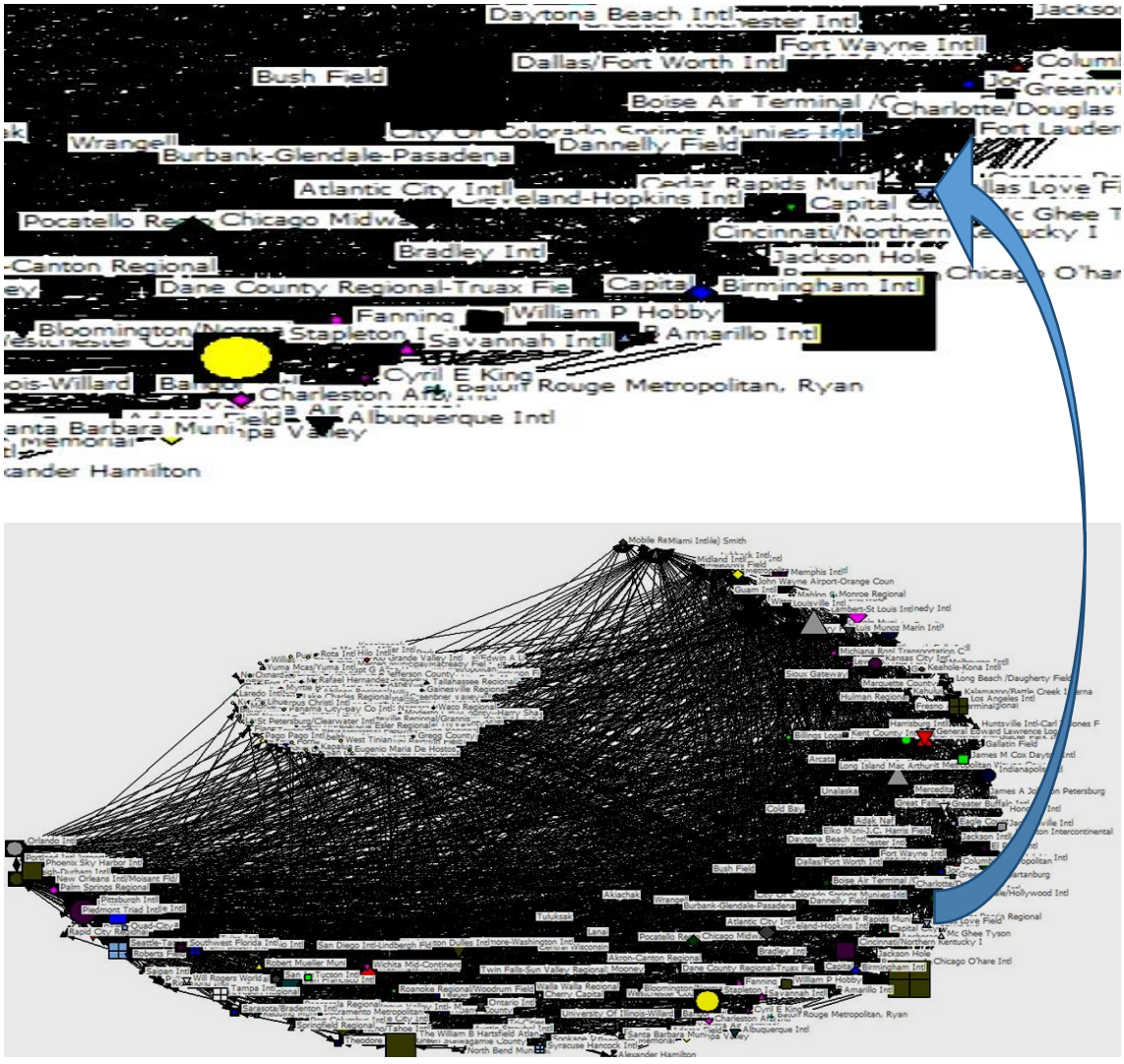


Fig. 3: Betweenness Centrality Visualization.

4.3. Closeness centrality

This measure perceives the closeness between the nodes or how long it takes to send/receive the information between the nodes. From the perspective of the U.S.A transport dataset, closeness centrality measure states that which airport is nearer to all other airports so that how fast the flight can reach the destination. Fig.4 depicts the closeness centrality values for some of the airports taken from the dataset. Chicago O'hare Intl airport has the highest closeness centrality in the given dataset which states that this airport is closer to many airports and the flights which departures from this airport takes shorter time to any other airports. The closeness centrality value for the airports is calculated using the following formula (2) [14]. Which is based on the length of the average shortest path between a node and all nodes in the given network.

$$C_c(i) = \left[\sum_{j=1}^{N} d(i,j)\right]^{-1} \qquad (2)$$

where d (i, j) indicates the distance between i and j nodes.

116	Klamath Falls Intll	0.327	0.657	0.347
117	Erie Intl	0.340	0.676	0.365
118	Chicago O'hare Intl	0.607	0.892	0.699
119	Bradley Intl	0.440	0.788	0.489
120	Cedar Rapids Muni	0.397	0.747	0.427

Fig. 4: Closeness Centrality Values.

4.4. Eigen vector centrality

This measure describes the behavior of a node in the network by allocating relative scores to all the nodes in the network based on connections with the high scoring nodes. Eigen vector justifies that

the node's importance is based on the number of connections to significant nodes. With respect to U.S.A transport dataset, Chicago O'hare Intl has higher Eigen vector value which means it is connected to most influenced airports in USA. Fig.5 depicts some of the Eigenvector values of the dataset. The Eigen vector measure of whole network is shown in the Fig.6.

116	Klamath Falls Intll	0.001
117	Erie Intl	0.001
118	Chicago O'hare Intl	0.207
119	Bradley Intl	0.070
120	Cedar Rapids Muni	0.012

Fig. 5: Eigen Vector Centrality Values.

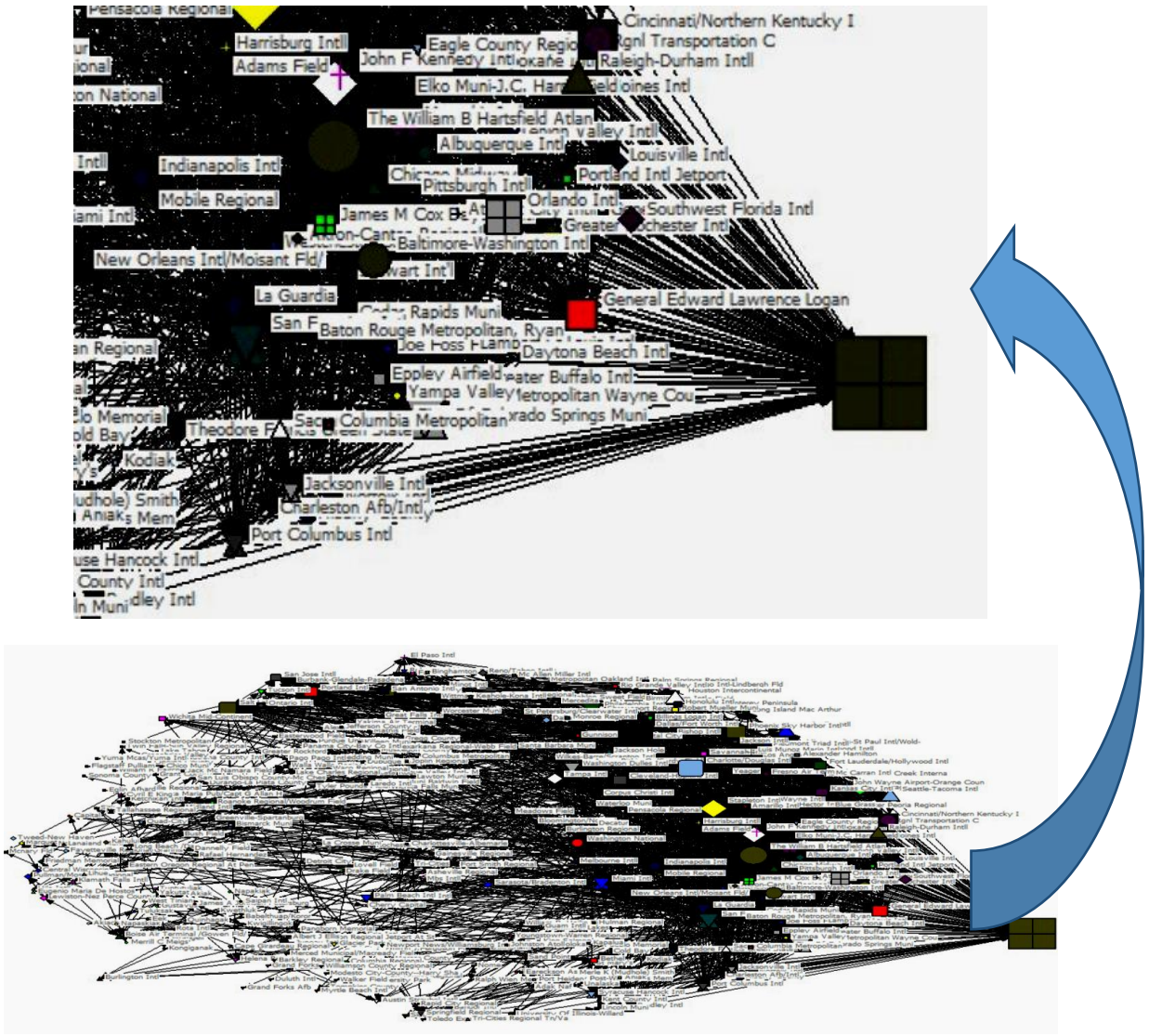


Fig. 6: Eigen Vector Centrality.

4.5. Reach centrality

It is a measure that counts the number of nodes that each one can reach in k or less steps to all other nodes. If the value of k=1; then this one will get equivalent to degree centrality. In U.S.A airport dataset's reach centrality is shown in the Fig.7.

		1 dwReach	2 NormdwRea
118	Chicago O'hare Intl	232.417	0.700
261	Dallas/Fort Worth Intl	217.667	0.656
255	The William B Hartsfield Atlan	208.500	0.628
182	Lambert-St Louis Intl	204.833	0.617
152	Pittsburgh Intll	202.200	0.609

Fig. 7: Reach Centrality Values.

4.6. Information centrality

This measure depicts the information flow from one node to other nodes by taking paths other than the shortest one. Since Betweenness centrality describes the mediator behavior of a node with shortest path; this might not get the influence of one node which uses paths other than the shortest path. Here in this transport network dataset, the information points to the flight and flow relates to its journey from one airport to other one. So information centrality in U.S.A airport network deals with the information of flights path other than shortest one. Fig.8 shows the information centrality measurement result of U.S.A airport dataset. Most of the nodes got higher information centrality value. Some of those airports are Portland Intl, Minneapolis-St Paul Intl/Wold, General Mitchell Intl, Greater Buffalo Intl, General Edward Lawrence Logan, Detroit Metropolitan Wayne Cou, Chicago O'hare Intl, Bradley Intl, Chicago Midway, Cleveland-Hopkins Intl, Elko Muni-J.C, Harris Field got the same Centrality value as 0.031 which shows the information flow of nodes with paths other than shortest one.

117	Erie Intl	0.010
118	Chicago O'hare Intl	0.031
119	Bradley Intl	0.031
120	Cedar Rapids Muni	0.029
121	William B. Heilig Field	0.021

Fig. 8: Information Centrality Values.

5. Conclusion

The analysis of USA airport network using centrality measures like degree, Betweenness, closeness, reach, Eigen vector and information identifies the important airports and its characteristics that help in revamping the economy of the country. The analysis of the U.S.A transport dataset identifies Chicago O'hare Intl airport plays a vital role in the network. It has the highest degree, Betweenness, closeness, Eigenvector, reach and information centrality which shows that Chicago is one of the designated city in U.S.A and the airport Chicago O'hare Intl is a highest influenced one. It is the one which acts as a connection airport for most of the airports within USA and by accessing this helps in boosting the trade within and outside of the country. So Chicago is one of the Porsche cities that promotes the economic growth to the country. Analysis of such datasets on different centrality measures helps in identifying the importance of the airports in various aspects.

References

[1] A.Malathi, D. Radha,"Analysis and Visualization of Social Media Networks", IEEE International Conference on "Computational Systems and Information Systems for Sustainable Solutions" CSITSS 2016, 6-8 October 2016, RVCE, Bengaluru, India. https://doi.org/10.1109/CSITSS.2016.7779440.

[2] V Murale and G. Raju, P, Analyzing the Role of Social Networks in Mapping Knowledge Flows: A case of a Pharmaceutical Company in India, Knowledge Management E-Learning: An International Journal, vol. 6, no. 1, pp. 49-65, 2014.

[3] Shyamala CK, Hemaashri S, Swetha R. An improved recommendation system for social network,In: International Journal of Control Theory and Applications. (ICSCS). IJCT A International Science Press. 2015; 8(5):1903-1910.

[4] Shyamala C.K., Hemaashri, S, An enhanced design for anonymization in social networks, Indian Journal of Science and Technology, Volume 9, Issue 30, 2016, Article number 99069.

[5] Borgatti, S.P., Everett, M.G. and Freeman, L.C. 2002. Ucinet for Windows: Software for Social Network Analysis. Harvard, MA: Analytic Technologies

[6] CongcongWang1, "A Social Network Analysis of International Wheat Trade", 3rd International Conference on Information Management, Innovation Management and Industrial Engineering, 26-28 Nov 2016

[7] HuiqingNiu, "Social Network Analysis of University Online Forum", International Conference on Computational Aspects of Social Networks, Computational Aspects of Social Networks (CASoN), 2010 International Conference on Computational Aspects of Social Networks,26-28 Sept. 2010, Taiyuan, China

[8] Kyoungjin Park, AlperYilmaz, "A social Network Analysis Approach To Analyze Road Networks", ASPRS 2010 Annual Conference, San Diego, California, April 2010, pp.26-30

[9] TU Yingfei, "Centrality Characteristics Analysis of Urban Rail Network", IEEE International Conference on "Intelligent Rail Transportation proceedings", 30 Aug.-1 Sept 2013, pp.285 – 290

[10] Zeyun Li, SharifahRohayah Sheikh Dawood. (2016, Jul.). "World City Network in China: A Network Analysis of Air Transportation Network". Canadian Center of Science and Education, Modern Applied Science. [10] (10), pp. 213-223. "Available online: https://doi.org/10.5539/mas.v10n10p213.

[11] Vladimir Batagelj and Andrej Mrvar (2006): Pajek datasets. http://vlado.fmf.unilj.si/pub/networks/data/.

[12] "available online: https://toreopsahl.com/datasets/"

[13] G.R.Ramya and Sivakumar, P.B, Advocacy monitoring of women and children health through social data, Indian Journal of Science and Technology, vol, 9, 2016

[14] Newman, M. "Networks: An Introduction", Oxford University Press, Inc., New York, NY, USA newman 2010. https://doi.org/10.1093/acprof:oso/9780199206650.001.0001.