COMPLEX NETWORKS OF THE AIR PASSENGER TRAFFIC IN MEXICAN AIRPORTS

Olivia Sashiko Shirai Reyna[(a)], Idalia Flores De La Mota[(b)]

[(a),(b)]Facultad de Ingeniería, UNAM
[(a)]sashikosr@outlook.com, [(b)] idalia@unam.mx

ABSTRACT
Nowadays, the air passenger traffic has been increasing, becoming an excellent, viable and reachable option for many people. This causes that airports may require an efficient organization to serve both, the companies that use the facilities and the passengers. In addition, it is important to consider that the amount of information that is generated may not be easy to analyze, sometimes because the managers don´t know all the information that they have or they don´t know how much this information can help the business.Therefore, in this work, we perform an analysis of the information obtained from some national and international airports of the Mexican Republic, using the methodology of Complex Networks. Also, with the results obtained, we will seek to put forward improvements in the service of this type of facilities, and the infrastructure.

Keywords: complex networks, visibility, time series, airpassengers.

1. INTRODUCTION
Since 2000, a record of all the passengers that arrive and depart at Mexican airports from domestic and international flights have been kept, and have not been analyzed because some of these data are not open data. Some of the data that may concern will be the number of airlines that operate in Mexican airports, the number of passengers, the number of routes, the number of available and taken seats, delays on the flights and more information all of these at domestic and international flights.

The main problem is that the strengths and weaknesses of Mexican airports as a whole systemare not known or identified, causing, among other things, losses in business opportunities and system saturation. For this reason, it is important to perform an initial diagnosis. Then according with the results of this analysis we model the system.Then we built mathematical models of the real data using complex networks, in order toput forward improvements for market strategies.

In order, to improve overall business strategies, it is necessary to apply statistical models and mathematical models to analyze, find patterns and predict data behavior.In this work, we will use different mathematical techniques such as network theory, complex networks, statistics, simulation and time series.

2. STATE OF THE ART
Some papers about complex networks and transportation are the next ones: “A dynamical model for air transportation network” written by Zanin M. et. al. (2009), this article talks about air transportation networks a how we can model this networks. Another article, from the same authors (written by Zanin M. and Lillo F.) is “Modelling the Air Transport with Complex Networks: a short review”, this article talks about how to model air transportation as complex networks. Finally, Boccaletti S. et.al. (2014)wrote a summary of many applications of complex networks, but the one´s that we were interested in was the applications of complex networks in transportation systems, they talk about transportation systems and specifically air transportation.These articles were the basement for our research since they give insights on the use of complex networks for air transport problems.

3. METHODOLOGY
For all this analysis, we used the R software, which is an open source programming language and software environment for statistical computing and graphics.We used specific packages for this work, such as,igraph, networks, tkrplot, sand, sna and others, this software allow us to generate graphs/networks, compute different network metrics like clustering or transitivity, different centrality metrics, plot networks and more functions.

We have the monthly information of the air passenger traffic in Mexican airports (domestic, international and total passengers) from January 2000 to March 2017.First, we analyzed one airport, just to prove that all our hypothesis are true. We used the information of Acapulco´s airport, so, with this information we will analyze the data with statistical techniques like time series analysis, then with the obtained results we will do some statistical models we will do some predictions of the air passenger traffic for the next years to analyses the strengths and the weaknesses of all the system in the future. Finally, simulation is used for the validation of the statistical results as well to analyze scenarios.

First, we use all the data that we have, so, we plot this information as a time series. The next figures show the different time series for domestic, international and total passengers of Acapulco´s Airport (respectively).

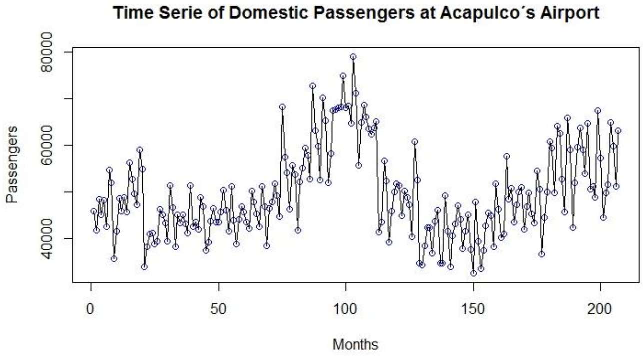

Figure 1: Time Serie Domestic Passengers Acapulco´s Airport

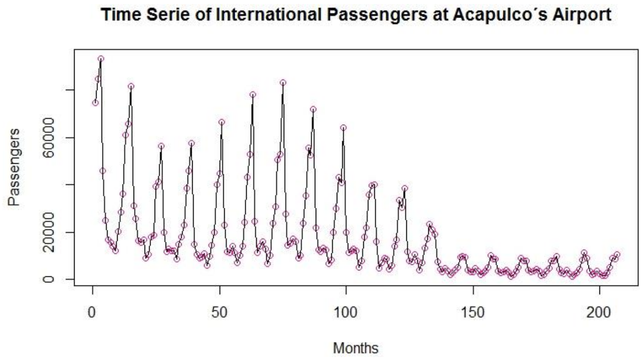

Figure 2: Time Serie International Passengers Acapulco´s Airport

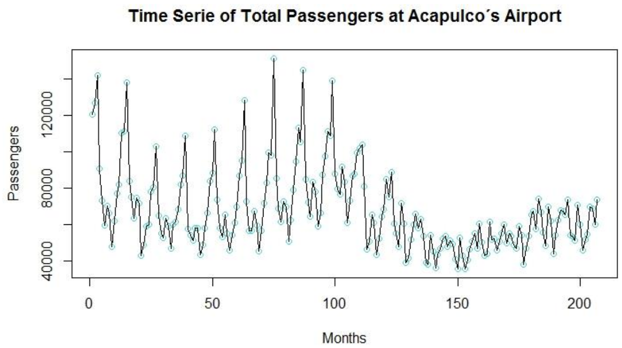

Figure 3: Time Serie Total Passengers Acapulco´s Airport

Now, once we have the plots of the time series we use a technique to transform this time series into complex networks.

3.1. Visibility Algorithm

As we described a little bit, we will transform the time series into complex networks. We use the visibility algorithm to do this transformation (Lacasa L. et al 2008).

The main goal of this algorithm is to map a time series into a network, so, we want to study to which extent the techniques and focus of network theory are useful as a way to characterize time series.

Another important thing is that this network inherits several properties of the time series.

The criterion of this algorithm is to establish two arbitrary data values $(t_a y_a)$ and $(t_b,\ y_b)$that will havevisibility, and consequently will become two connected nodes ofthe associated graph, if any other data $(t_c,\ y_c)$ placed between them fulfills:

$$y_c < y_b + (y_a - y_b)\frac{t_b - t_c}{t_b - t_a} \quad (1)$$

So, we apply this algorithm to our time series and we obtain the next graphs.

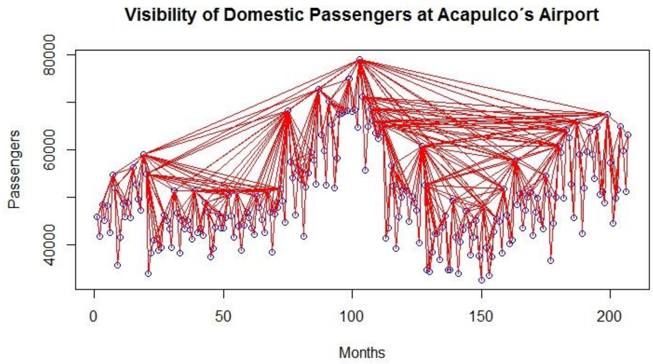

Figure 4:Visibility Domestic Passengers Acapulco´s Airport

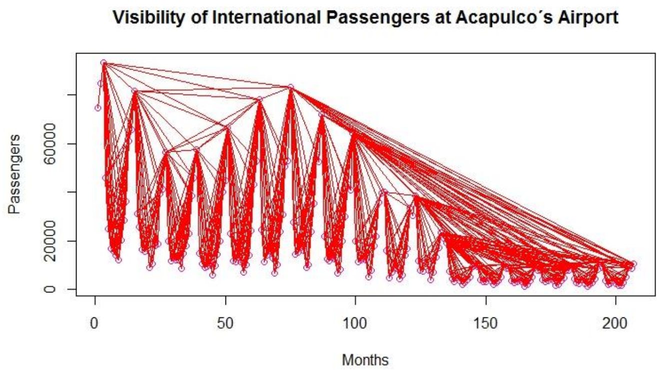

Figure 5:Visibility International Passengers Acapulco´s Airport

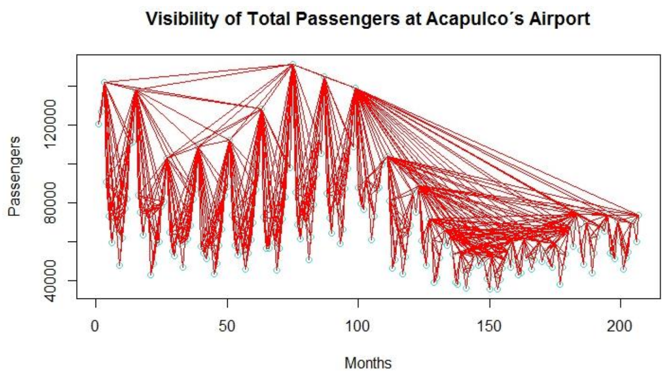

Figure 6:Visibility Total Passengers Acapulco´s Airport

We can easily check that by means of the present algorithm, theassociated graph extracted from a time series is always:

1. Connected: each node sees at least its nearest neighbors (left and right), the first and the last one at least see one.
2. Undirected: the way the algorithm is built up, there is nodirection defined in the links.
3. Invariant under affine transformations of the series data: the visibility criterion is invariant under rescaling of both horizontaland vertical axes, and under horizontal and verticaltranslations.

After, we applied the visibility algorithm, we can obtain the visibility graph, so now, we plot the time series as complex networks, that´s what we will see in the next figures.

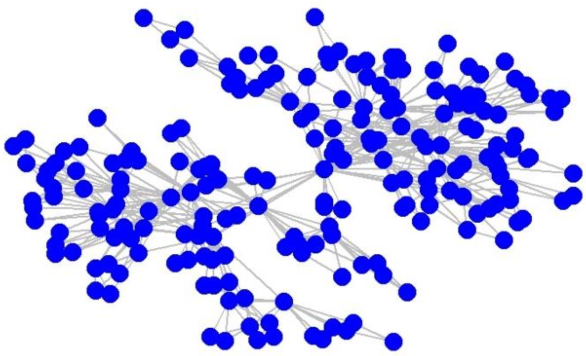

Figure 7:Network Domestic Passengers Acapulco´s Airport

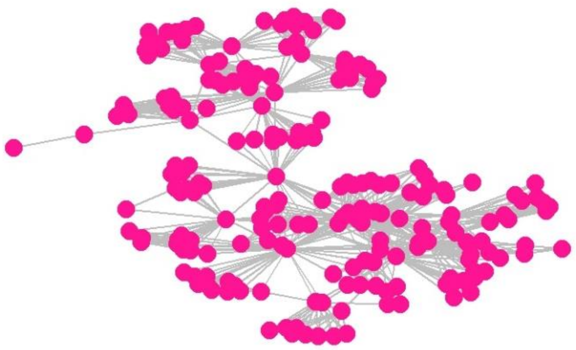

Figure 8:Network International Passengers Acapulco´s Airport

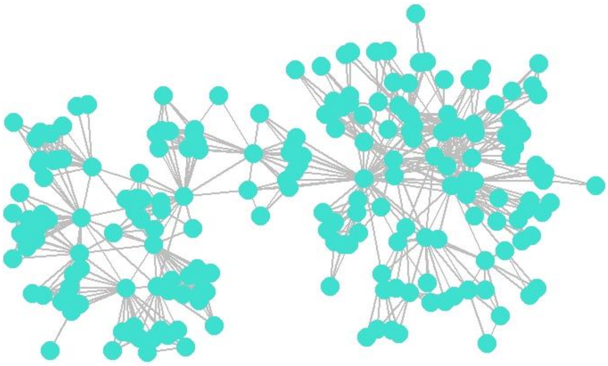

Figure 9:Network Total Passengers Acapulco´s Airport

3.2. Complex Networks

Acomplex networkis graph ornetwork with non-trivial topological features, with patterns of connection between their elements that are neither purely regular nor purely random. Such features include a heavy tail in the degree distribution, a high clustering coefficient, assortativityor disassortativity among vertices, community structure, and hierarchical structure. that do not occur in simple networks such as lattices or random graphsbut often occur in graphs modelling of real systems (Newman, 2010).

We will focusthis paper to study the different network metrics, such as the network connectivity. Network connectivity is also a kind of metric to discuss how well parts of the network connect to another. Related terms include network topology, which refers to the structure and makeup of the network as a whole structure.

4. ANALYSIS AND RESULTS

After we have our data as networks we can compute the different network´s metrics.

We compute and compare metrics of our three networks, such as the number of edges, minimum, maximum and mean degree, diameter, mean distance, the number of cliques, the density of the network, if its assortativity or disassortativity among nodes, the global clustering of the network, the mean local clustering, closeness centrality, degree centrality and betweenness centrality. We compute this metrics to study the topology of the networks and to understand and identify different patterns and to classify the networks in the different complex networks models that we have (Random networks, Scale-free networks and Small World networks).

Table 1: Metric´s Results

Results	Domestic	International	Total
Nodes	207	207	207
Edges	638	1013	801
Max. Degree	32	43	44
Min. Degree	2	1	2
Mean Degree	6.164251	9.78744	7.73913
Diameter	8	7	7
Mean Distance	3.663337	3.415224	3.597908
Cliques	8	9	9
Density	0.0299236	0.04751184	0.03756859
Assortativity	0.1387072	0.01292899	0.03675553
Global Clustering	0.3931777	0.504932	0.4548834
Mean Local Clustering	0.7447704	0.7810761	0.7675662
Closeness Centrality	0.3800197	0.3303979	0.3635828
Degree Centrality	0.1254163	0.161226	0.1760236
Betweenness Centrality	0.5545298	0.503284	0.5915958

First, we need to understand the numbers that we have in the Table 1. We notice that there´s a lot of difference between the minimum, maximum and the mean degree, which is because there are just few nodes with a lot of links and there are many nodes with just a few links. The diameter is similar in the three networks. The mean distance is small that means that we can go from any node from the network to another node in a few steps, in this case approximately 3 steps. The number of cliques in the network tells us the number of complete subgraphs in our network. The network density describes the portion of potential connections in a network that are actual connected, so that means that our networks don´t have a lot of actual connections.

To continue with our analysis, our three networks are assortative so that means that there´s a preference for a network´s nodes to attach to others that are similar in some way, soassortativity is often operationalized as a correlation between nodes. The global clustering coefficientis based on triplets of nodes and the local clustering coefficient of a node in a networkquantifies how close itsneighbors are to being a clique (Bollobás, 1998). In our networks the global clustering and the mean local clustering are high.

We also computed the centrality metrics, the closeness centrality of a node is a measure ofcentralityin anetwork, calculated as the sum of the length of theshortest pathsbetween the node and allother nodes in the network. Thus, the morecentral a node is, the closer it is to all other nodes, so, in our networks the closeness centrality is not so high (Caldarelli et. al. 2012).The degree centrality is defined as the number of links incident upon a node (i.e., the number of ties that a node has), in our case, the networks have a low degree centrality (Newman, 2010). The betweenness centrality is a measure ofcentralityin a networkbased onshortest paths. For every pair of vertices ina connected network, there exists at least one shortest path between the vertices such that either the number of edges that the path passes through so, the betweenness centrality for each node is the number of these shortest paths that pass through the node; in our networks, we have a high betweenness centrality(Caldarelli et. al. 2012).

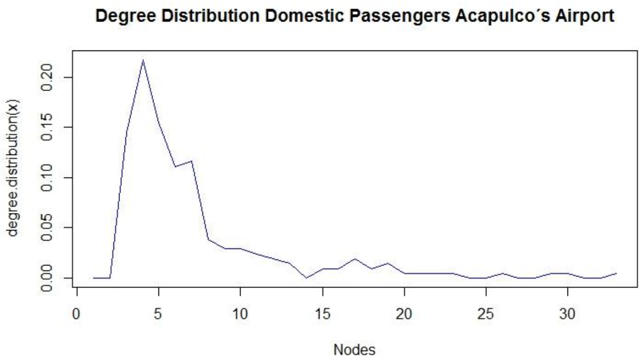

Figure 10:Degree Distribution Domestic Passengers Acapulco´s Airport

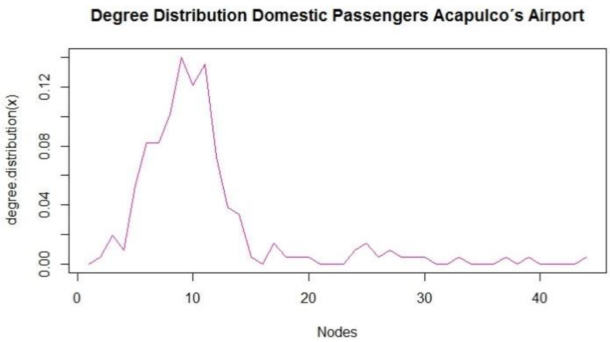

Figure 11:Degree Distribution International Passengers Acapulco´s Airport

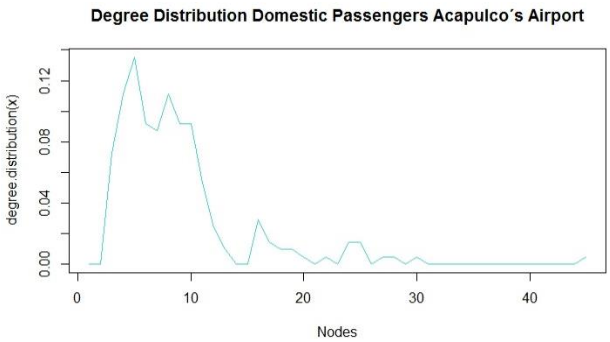

Figure 12:Degree Distribution International Passengers Acapulco´s Airport

With the degree distribution graphs above we notice that it seems that our networks have a power law distribution. Maybe it´s not so clear but there are a few nodes with a high degree and it decays fast, so, there are a lot of nodes with a low degree.

With all these results, we can conclude that our networks follow a scale-free model. The most notable characteristic in a scale-free network is the relative commonness of nodes with a degree that greatly exceeds the average. The scale-free property strongly correlates with the network's robustness to failure. It turns out that the major hubs are closely followed by smaller ones.This implies that the low-degree nodes belong to very dense sub-networks and those sub-networks are connected to each other through hubs (Barabási et. al. 2003).

ACKNOWLEDGMENTS

To CONACYT scholarship for the first author, and DGAPA-PAPIIT Project ITl 02117 Accesibilidad y movilidad del transporte público urbano en la Ciudad de México, el caso de la delegaci6n Tlalpan.

REFERENCES

[1] Barabási A.-L., Albert R., 1999. Emergence of scaling in random networks.Science 286 (5439): 509-512.

[2] Barabási A.-L., Bonabeau E., 2003. Scale-Free Networks. Scientific American. 50-59.

[3] Boccaletti S., Latora V., Moreno Y., Chavez., Hwang D.-U., 2006. Complex networks: Structure and dynamics. Physics Reports. 424. 175-308.

[4] Bollobás B., 1998. Modern Graph Theory.Springer.

[5] Buchanan M., 2002. Nexus: Small Worlds and the Groundbreaking Science of Networks. New York, N.Y.: W.W. Norton.

[6] Caldarelli G., Catanzaro M., 2012. Networks: A Very Short Introduction. Oxford University Press, Oxford.

[7] Caldarelli G., Vespignani A.,2007. Large Scale Structure and Dynamics of Complex Networks: From Information Technology to Finance and Natural Science. World Scientific Publishing Co., Inc., USA.

[8] Dalgaard P., 2008. Introductory Statistics with R. Springer.

[9] Dorogovtsev S.N., 2010. Lectures on Complex Networks. Oxford University Press, Inc. New York, NY, USA.

[10] Kolaczyk E.D., Csárdi G., 2014. Statistical Analysis of Network Data with R. Springer, New York, USA.

[11] Lacasa L., Luque B., Ballesteros F., Luque J., Nuño J.C., 2008. From time series to complex networks: The visibility graph. PNAS. Vol. 5. No, 13. 4972-4975.

[12] Lacasa, L. et al., 2015. Network structure of multivariate time series. Sci. Rep. 5,15508; doi: 10.1038/srep15508.

[13] Lütkepohl H., 2005. New Introduction to Multiple Time Series Analysis. Springer.

[14] Newman M.E.J., 2010. Networks: An Introduction. Oxford University Press, Oxford.

[15] Newman M.E.J., 2003. The Structure and Function of Complex Networks. SIAM Review 45,167-256.

[16] Newman M.E.J., Barabási A.-L., Watts D.J., 2006. The Structure and Dynamics of Networks. Princeton University Press, Princeton, NJ, USA.

[17] Newman M.E.J., Peixoto T.P., 2015. Generalized Communities in Networks. Physical Review Letters 115, 08870. American Physical Society.

[18] Wei W. W. S., 2006. Time Series Analysis Univariate and Multivariate Analysis. Pearson.

[19] Zanin M., Lacasa L., Cea M., Crístobal S., 2009. A dynamical model for the air transportation network. Proceedings 23rd European Conference on Modelling and Simulation.

[20] Zanin M., Lillo F., 2013. Eur. Phys. J. Spec. Top. 215:5. doi:10.1140/epist/e2013-01711-9

AUTHORS BIOGRAPHY

Sashiko Shirai obtained her bachelor in Actuarial Science in the Faculty of Science of the UNAM. She is studying her Master in Operations Research in the Faculty of Engineering of the UNAM. She is an assistant professor at the Faculty of Engineering and the Faculty of Science. Her research interests lie in complex networks, statistics and operational research.

Idalia Flores received a Master with honors, being awarded the Gabino Barreda Medal for the best average of her generation, in the Faculty of Engineering of the UNAM, where she also obtained her Ph.D. in Operations Research. Dr. Flores is a referee and a member of various Academic Committees at CONACYT as well as being a referee for journals such as Journal of Applied Research and Technology, the Center of Applied Sciences and Technological Development, UNAM and the Transactions of the Society for Modeling and Simulation International. She is a full time professor at the Postgraduate Program at UNAM and her research interests lie in simulation and optimization of production and service systems, based on this research she has been editor of two books edited by Springer.