

## Research Article

# Weighted Complex Network Analysis of Pakistan Highways

**Yasir Tariq Mohmand and Aihu Wang**

*School of Business Administration, South China of University of Technology, Guangzhou 510640, China*

Correspondence should be addressed to Aihu Wang; [bmawang@scut.edu.cn](mailto:bmawang@scut.edu.cn)

Received 19 July 2013; Accepted 13 October 2013

Academic Editor: Zhen Jin

Copyright © 2013 Y. T. Mohmand and A. Wang. 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.

The structure and properties of public transportation networks have great implications in urban planning, public policies, and infectious disease control. This study contributes a weighted complex network analysis of travel routes on the national highway network of Pakistan. The network is responsible for handling 75 percent of the road traffic yet is largely inadequate, poor, and unreliable. The highway network displays small world properties and is assortative in nature. Based on the betweenness centrality of the nodes, the most important cities are identified as this could help in identifying the potential congestion points in the network. Keeping in view the strategic location of Pakistan, such a study is of practical importance and could provide opportunities for policy makers to improve the performance of the highway network.

## 1. Introduction

With the expansion of world economy and international trade, the role of transportation infrastructure is of crucial importance to the development of a country and is an important indicator of its economic growth. Land, water, and air transportation forms the backbone of an economy by supporting the movement of information, people, and goods. Previously, econometric models were applied to estimate the effect of infrastructure on the international trade; however, such a study did not provide sufficient insight into infrastructure itself. It is during the recent few years that different transportation systems are studied in much more detail using the complex network analysis. Key studies include but are not limited to the US, China, India, and worldwide air networks [1–4], Chinese and Indian railway networks [5, 6], public transportation in Singapore [7] and Poland [8], the underground transportation in Boston [9], and the Indian highway network [10]. The advances that were incorporated in these studies are enormous and revealed many new topological features of the networks. Small world property is one such feature displayed by most of the transportation networks. However, not all the networks are similar when it comes to degree distribution. The airport networks of US, China, and India show a power law degree distribution; exponential degree distribution is observed in Chinese and

Indian railway networks and the Polish public transportation whereas neither power law nor normal in the case of Indian highway network.

Pakistan lies at the crossroads between South Asia, Central Asia, and Western Asia. The location provides Pakistan with a valuable opportunity to enhance its economy by providing logistic routes and services to the landlocked central Asian countries and to act as a bridge. On the contrary, the logistic network of Pakistan is largely inadequate with total length of the national highways standing at roughly 8,780 kilometers, accounting for only 3 percent of the entire road network but handling 75 percent of the road traffic in the country [11]. Pakistan's logistics mostly rely on the road network where 96 percent of the national freight traffic is carried on road networks. This large percentage of road freight is a result of the failure of Pakistan Railways freight operations, which have been at halt since 2011 and briefly resumed in 2012. Poor and unreliable road infrastructure combined with traffic congestion in major cities result in 2–4 times longer freight journeys than they would in Europe. These factors hinder Pakistan's ability to integrate into the global supply chain which requires just-in-time delivery. According to the World Bank statistics, the poor performance of the logistic sector of Pakistan is estimated to cost the economy 4–6 percent (8–12 billion US \$) of the national GDP every year. In this scenario, it is of utmost importance to study

TABLE 1: Statistical properties of Pakistan national highway network.

Property	Value
Nodes, $N$	266
Edges, $M$	4802
Average path length, $(l)$	2.49
Average clustering coefficient, $C$	0.81
Average weighted clustering coefficient, $C^w$	0.82
Diameter, $d$	5
Average degree	36.1
Degree range	(1, 109)
Average weight, $(w)$	515.7
Weight range	(110, 1175)
Average strength, $(s)$	18,617.9
Strength range	(150, 69,996)
Assortativity, $r$	0.48

and analyze the topological features of the Pakistan National Highway Network (PNHN). Such a study would not only help in highlighting the existing bottlenecks of the PNHN but also provide opportunity for policy makers to design efficient modes of extension to the network and reducing the amounting congestions.

The rest of the paper is as follows. Section 2 discusses the network construction methodology and Section 3 lists the topological properties of the network, whereas Section 4 concludes the paper.

## 2. Network Construction

Data for the movement of passengers were provided by the National Transport Research Center (NTRC) and the Provincial Public Transport Authorities of Pakistan. The data was transformed and travel for each day was represented as a weighted graph  $G$  with  $N$  nodes and  $M$  edges, an associated adjacency matrix  $A = [a_{ij}]$ , and a weight matrix  $W = [w_{ij}]$  representing the number of passengers travelling between locations  $i$  and  $j$  in a single day. It is important to note that the objective of the study is not studying the underlying physical structure of the networks but of the movement of people between the different nodes. As such, when we say two nodes  $i$  and  $j$  are connected,  $a_{ij} = 1$ , we mean that there is at least one passenger travelling between these two nodes during the day. Such a representation has already been used to represent public transportation in Singapore [7].

## 3. Topological Properties

Table 1 provides all computed network statistics, from basic network properties such as the number of nodes and edges to the more complex metrics such as weighted clustering and assortativity.

**3.1. Basic Properties.** The network has 266 nodes with 4802 edges. The average shortest path length (the minimum number of edges passed through to get from one node to another)

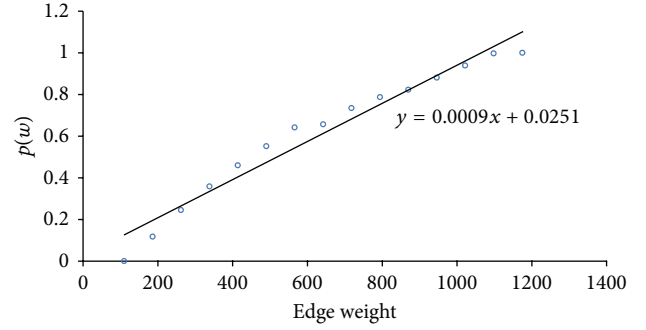


FIGURE 1: Weight distribution.

between one node and all other nodes of the network is calculated using the following equation:

$$D = \sum_{s,t \in V} \frac{d(a,b)}{N(N-1)}, \quad (1)$$

where  $V$  correspond to the set of nodes in the network,  $d(a,b)$  is the shortest path from  $a$  to  $b$ , and  $N$  is the total number of nodes in the network. A small average path length ( $l = 2.49$ ) means that there is travel between almost all cities in Pakistan, regardless of the geographical distance. The network also features small diameter (maximum path length of a network)  $d = 5$ . These observed path lengths are similar to those found for the transportation routes for cities in Poland [9] and Singapore [7]. The edge weight distribution is plotted in Figure 1. The range of travel between cities varied greatly,  $w \in (135, 1100)$ . From the data, it is evident that the flow of passengers is directed from nodes of low degree towards nodes of high degree. Karachi, Hyderabad, Lahore, Islamabad, Rawalpindi, Faisalabad, and Peshawar are a few node examples with high degree and also handle most of the flow.

**3.2. Degree and Strength Distributions.** The degree of a node, a measure of its connectivity, is defined as the fraction of nodes with degree  $k$  in the network. In our case, the degree is defined as the number of cities that can be reached from a given city via a single route. For a given node  $i$ , the degree can be represented using

$$k_i = \sum_j^N a_{ij}. \quad (2)$$

The average degree of the whole graph can be obtained using the following equation:

$$\langle k \rangle = \frac{1}{N} \sum_i^N k_i = \frac{1}{N} \sum_i^N \sum_j^N a_{ij}. \quad (3)$$

Subsequently, a node's strength is simply the sum of the weights on the edges incident upon it and is given by

$$s_i = \sum_j^N a_{ij} w_{ij}. \quad (4)$$

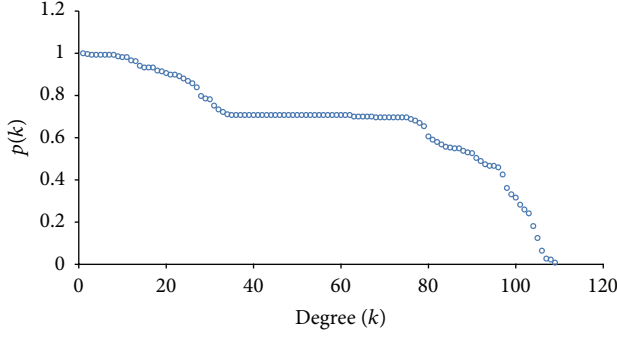


FIGURE 2: Degree distribution.

The average strength of the whole graph can then be obtained using the following equation:

$$\langle s \rangle = \frac{1}{N} \sum_i s_i = \frac{1}{N} \sum_i \sum_j a_{ij} w_{ij}. \quad (5)$$

The network possesses a high average degree of 36.1, indicating high connectivity among the nodes. The degree distribution of the network is presented in Figure 2. The distribution of node connectivity is neither normal nor power law. It is interesting to note that even though many other transportation networks reported different distribution fits, the highway network of Pakistan's close neighbor India was also neither normal nor power law [10]. The strengths reveal that although nodes with similar degree exist in the network, the traffic handled by each node is significantly different. We can average the strengths over all nodes with a given degree to get the strength spectrum (Figure 3). The spectrum illustrates a positive relationship between the degree and strength of nodes which means that the more connected a node is, the more traffic it handles.

**3.3. Topological Clustering.** The clustering coefficient of a node  $i$  is defined as the ratio of the number of links shared by its neighboring nodes to the maximum number of possible links among them. Simply put, the clustering coefficient is a measure of cohesiveness around a given node  $i$  and it is defined by the equation

$$C_i = \frac{2E_i}{k_i(k_i - 1)} = \frac{2}{k_i(k_i - 1)} \sum_{j,h} a_{i,j} a_{i,h} a_{j,h}, \quad (6)$$

where  $E_i$  is the number of edges between node  $i$ 's neighbors and  $2/k_i(k_i - 1)$  is a normalization factor equal to the maximum number of possible edges among the neighbors. Because of this normalization,  $C_i$  is in the interval  $[0, 1]$  where 0 and 1 indicate that none or all of node  $i$ 's neighbors are linked, respectively. The average clustering coefficient can thus be represented by the following mathematical expression:

$$\langle C \rangle = \frac{1}{N} \sum_{i=0}^N C_i. \quad (7)$$

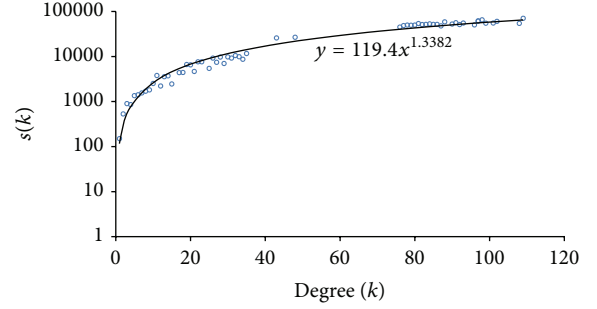


FIGURE 3: Strength as a function of degree.

Using the above equation, the average clustering coefficient ( $C$ ) of the network is calculated to be 0.81, indicating that the NHNP is a highly clustered network. This result is substantially higher than the value of an equivalent Erdős-Rényi random graph, ( $C_{ER}$ ) = 0.1. The clustering coefficient together with the small average path length (Section 3.1) indicates that the PNHN is a small world. The computed clustering coefficient of the network is greater than that of the Indian highway ( $C = 0.78$ ) and Railway ( $C = 0.69$ ) [6, 10], and that of location network of Portland ( $C = 0.06$ ) [12] whereas almost the same as that of the public transportation in Poland ( $C = 0.85$ ) and Chinese Railways ( $C = 0.83$ ) [5, 8].

Unlike the clustering coefficient, the weighted clustering coefficient  $C_i^w$  also takes into account the weights of edges; that is,

$$C_i^w = \frac{1}{s_i(k_i - 1)} \sum_{j,h} \frac{w_{i,j} + w_{i,h}}{2} a_{i,j} a_{i,h} a_{j,h}. \quad (8)$$

The average weighted clustering coefficient can thus be represented by the following mathematical expression:

$$\langle C^w \rangle = \frac{1}{N} \sum_i C_i^w. \quad (9)$$

If the weighted clustering coefficient is equal to the clustering coefficient of the network, ( $C^w = C$ ), it means that the weights are completely uncorrelated. However, in the case when the weighted clustering is greater or smaller than clustering coefficient, the result implies that clustering is formed by edges with larger weight or smaller weights, respectively. In our case,  $C^w > C$  implies that the clustering is formed by edges with larger weights (Figure 4). In this figure, the clustering and weighted clustering coefficients are averaged over all nodes with a certain degree to get the clustering spectrum.

Another important topological characteristic of a network that is examined is the degree-degree correlation between connected nodes. A given network is said to be assortative if the high-degree nodes have a tendency to connect to other high-degree nodes. Similarly, disassortative networks are where low-degree nodes tend to connect to high-degree nodes. Newman introduced a summary statistic for assortativity ( $r$ ) in 2002 [13], defined as the Pearson

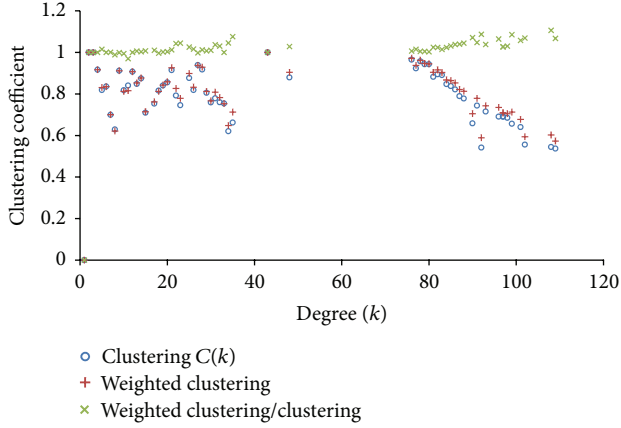


FIGURE 4: Clustering spectrum.

correlation coefficient of the degrees at either ends of an edge. Mathematically, this expression can be represented by the following equation:

$$r = \frac{1}{\sigma_q^2} \sum_{jk} jk (e_{jk} - q_j q_k), \quad (10)$$

where

$$q_k = \sum_j e_{jk}, \quad \sigma_q^2 = \sum_k k^2 q_k - \sum_k k q_k^2. \quad (11)$$

This statistic lies in between the range of  $[-1, 1]$ , where  $-1$  indicates a completely disassortative network and  $1$  indicates a completely assortative network. Examples of  $r$  greater than  $0$  include the Indian highway network [10] and ship transport network of China [14] whereas examples of  $r$  less than  $0$  include the Indian railway network [6], public transportation of Singapore [7] and Poland [8], and airport network of China [2] and India [3]. For PNHN, the observed topological assortativity is  $0.48$  (similar to the Indian highway network). A closer inspection of the degree correlations can be done using another measure, the average degree of the nearest neighbor,  $K_{nn}(k)$ , for nodes of degree  $k$ . Consider

$$k_{nn,i} = \frac{1}{k_i} \sum_{j=1}^N a_{ij} k_j. \quad (12)$$

If  $K_{nn}(k)$  increases with  $k$ , the network is assortative. If  $K_{nn}(k)$  decreases with  $k$ , the network is disassortative. The weighted version is given by

$$k_{nn,i}^w = \frac{1}{k_i} \sum_{j=1}^N a_{ij} w_{ij} k_j. \quad (13)$$

$k_{nn,i}^w \approx k_{nn,i}$  implies that the edge weights are uncorrelated with the degree of  $i$ 's neighbors. If the resultant weighted neighbor degree is greater than simple neighbor degree, ( $k_{nn,i}^w > k_{nn,i}$ ), then heavily weighted edges connect to neighbors with larger degree, while the opposite occurring

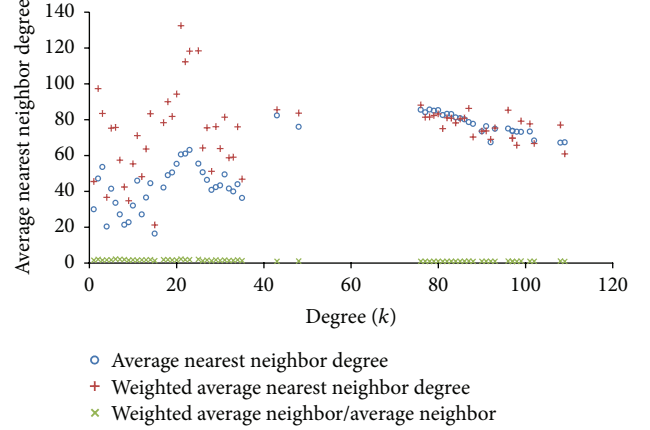
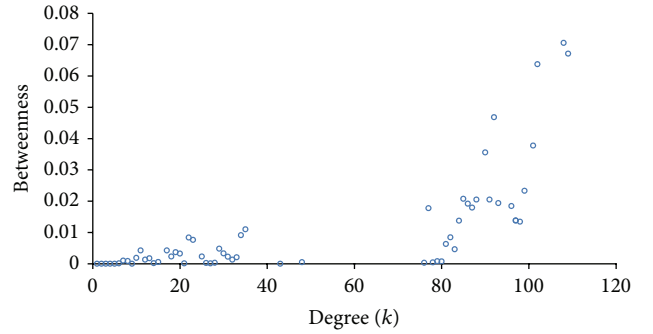
FIGURE 5: Average degree of nearest neighbors of nodes with degree  $k$ .

FIGURE 6: Betweenness centrality.

when  $k_{nn,i}^w < k_{nn,i}$ . The average degree of the nearest neighbor is represented in Figure 5 where  $K_{nn}(k)$  increases with  $k$  and  $k_{nn,i}^w > k_{nn,i}$  meaning that the heavily weighted edges connect to neighbors with larger degree.

**3.4. Betweenness Centrality.** The betweenness centrality measure is used to identify the nodes with high congestion [15, 16]. Betweenness centrality of a node  $i$  can be defined as sum of the fraction of all-pairs shortest paths that passes through  $i$ . Mathematically

$$c_B(i) = \sum_{s,t \in V} \frac{\sigma(s, t | i)}{\sigma(s, t)}, \quad (14)$$

where  $V$  is the set of nodes,  $\sigma(s, t)$  is the total number of shortest paths, and  $\sigma(s, t | i)$  is the number of shortest paths passing through  $i$  [17]. Betweenness centrality is presented in Figure 6. The top ten cities based on the betweenness centrality are listed in Table 2. Quetta and Peshawar, the trade corridor border cities, lead the list and handle maximum traffic followed by the port city of Karachi.

## 4. Conclusion and Future Work

Transportation networks, whether being land, air, or sea, communicate the development level of a country and can

TABLE 2: Top ten cities identified based on betweenness centrality.

Betweenness centrality	City
0.089434	Quetta
0.070581	Peshawar
0.067116	Karachi
0.052508	Naushahro Feroze
0.049318	Hoshab
0.046812	Mastung
0.044426	Hyderabad
0.038023	Loralai
0.037758	Sibi
0.034866	Jamshoro

rightly be described as forming the backbone of economic development. Along with other tools, complex network methodologies have been extensively used to analyze transportation networks. As an addition to the theory and application of complex networks, the weighted national highway network of Pakistan is analyzed using complex network theory. The PNHN is a highly clustered network where the degree distribution is neither normal nor power law. The small world properties and assortative mixing of the highway network are evident from the calculated properties. It is interesting to note that the topological properties of the PNHN are largely similar to those of the Indian highway network. Furthermore, using betweenness centrality, the cities with potential traffic congestion are also identified.

The beauty of complex network theory is that it is a powerful tool with limitless application possibilities. Although the analysis was performed taking daily average number of passengers as edge weights, it would also be interesting to conduct a much larger study using data for weeks or months. Similarly, subject to availability of data, weighted network analysis using movement of traffic or better yet the flow of goods in terms of TEU (twenty foot equivalent unit) from one city/district to other could provide useful insight into the logistics aspect of the network. Such a study would clearly highlight the network and its topological features in much more detail and help the policy makers to further enhance the infrastructure to achieve efficient flow.

## Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

## Acknowledgments

This research is supported by the 2011 Founded Project of the National Natural Science Foundation of China (71171084), the 2011 Research Fund for the Doctoral Program of Higher Education of China (20110172110010), and the Fundamental Research Funds for the Central Universities (2012, x2gsD2117850).

## References

- [1] A.-L. Barabási and R. Albert, "Emergence of scaling in random networks," *Science*, vol. 286, no. 5439, pp. 509–512, 1999.
- [2] W. Li and X. Cai, "Statistical analysis of airport network of China," *Physical Review E*, vol. 69, no. 4, Article ID 046106, 2004.
- [3] G. Bagler, "Analysis of the airport network of India as a complex weighted network," *Physica A*, vol. 387, no. 12, pp. 2972–2980, 2008.
- [4] R. Guimerà, S. Mossa, A. Turttschi, and L. A. N. Amaral, "The worldwide air transportation network: anomalous centrality, community structure, and cities' global roles," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, no. 22, pp. 7794–7799, 2005.
- [5] W. Li and X. Cai, "Empirical analysis of a scale-free railway network in China," *Physica A*, vol. 382, no. 2, pp. 693–703, 2007.
- [6] P. Sen, S. Dasgupta, A. Chatterjee, P. A. Sreeram, G. Mukherjee, and S. S. Manna, "Small-world properties of the Indian railway network," *Physical Review E*, vol. 67, no. 3, Article ID 036106, 2003.
- [7] H. Soh, S. Lim, T. Zhang et al., "Weighted complex network analysis of travel routes on the Singapore public transportation system," *Physica A*, vol. 389, no. 24, pp. 5852–5863, 2010.
- [8] J. Sienkiewicz and J. A. Holyst, "Statistical analysis of 22 public transport networks in Poland," *Physical Review E*, vol. 72, no. 4, Article ID 046127, 2005.
- [9] V. Latoraa and M. Marchiorib, "Statistical mechanics and its applications," *Physica A*, vol. 314, no. 1–4, pp. 109–113, 2002.
- [10] S. Mukherjee, "Statistical analysis of the road network of India," *Pramana Journal of Physics*, vol. 79, no. 3, pp. 483–491, 2012.
- [11] National Highway Authority of Pakistan, <http://nha.gov.pk/about-nha/background/>.
- [12] G. Chowell, J. M. Hyman, S. Eubank, and C. Castillo-Chavez, "Scaling laws for the movement of people between locations in a large city," *Physical Review E*, vol. 68, no. 6, Article ID 066102, 2003.
- [13] M. E. J. Newman, "Assortative mixing in networks," *Physical Review Letters*, vol. 89, no. 20, Article ID 208701, 2002.
- [14] X. Xu, J. Hu, F. Liu et al., "Empirical analysis of the ship-transport network of China," *Chaos*, vol. 17, Article ID 023129, 2007.
- [15] B. K. Singh and N. Gupte, "Congestion and decongestion in a communication network," *Physical Review E*, vol. 71, no. 5, Article ID 055103, 2005.
- [16] S. Mukherjee and N. Gupte, "Gradient mechanism in a communication network," *Physical Review E*, vol. 77, no. 3, article 036121, 12 pages, 2008.
- [17] U. Brandes, "On variants of shortest-path betweenness centrality and their generic computation," *Social Networks*, vol. 30, no. 2, pp. 136–145, 2008.



