STATISTICAL ANALYSIS OF THE INDIAN RAILWAY NETWORK: A COMPLEX NETWORK APPROACH*

Saptarshi Ghosh†, Avishek Banerjee, Naveen Sharma
Sanket Agarwal, Niloy Ganguly‡

Department of Computer Science and Engineering, IIT Kharagpur
Kharagpur 721302, India

Saurav Bhattacharya

Department of Computer Science and Engineering, NIT Durgapur
Durgapur 713209, India

Animesh Mukherjee

Complex Systems Lagrange Lab., ISI Foundation, Torino 10133, Italy

(Received March 28, 2011)

In this paper, we study the Indian Railway Network (IRN) as a weighted complex network in which railway stations are considered as the nodes, while the weights on edges represent the number of trains directly linking two stations. Apart from the small-world characteristics and exponential distributions of node-degrees and edge-weights in the IRN, we explore the correlations of the amount of traffic with the topology of the network. The traffic in the IRN is found to be accumulated on interconnected groups of stations and is concentrated in the links between large stations. We also identify the most important stations in the IRN with respect to connectivity and traffic-flow, which help to identify some of the probable points (and regions) of congestion in the network. The study indicates several guidelines to improve the performance of the IRN.

DOI:10.5506/APhysPolBSupp.4.123
PACS numbers: 89.75.–k, 89.75.Fb, 89.40.Bb

---

* Presented at the 2nd Summer Solstice International Conference on Discrete Models of Complex Systems, Nancy, France, June 16–18, 2010.

† `saptarshi@cse.iitkgp.ernet.in`

‡ `niloy@cse.iitkgp.ernet.in`

1. Introduction

Transportation networks are among the most important building blocks in the economic development of a country. The structure and performance of transportation networks reflects the ease of travelling and transferring goods among different parts of a country, thus affecting trade and other aspects of the economy. In the recent years, complex network analysis has been used to study several transportation networks. These include airport networks (for instance, the airport network of China [1], airport network of India [2] and the world-wide airport network [3, 4]), urban road networks [5, 6] and railway networks [7, 8, 9, 10, 11].

Some commonalities have been observed in the topological properties of almost all transportation networks, such as small-world properties. On the other hand, certain topological properties, such as the cumulative degree distribution, have been found to differ widely — power-laws for Indian airport network [2] and world-wide airport networks [4], two-regime power-laws for the China airport network [1] and US airport network [12] as opposed to the exponential degree distributions of the railway networks of India [10] and China [9].

It is to be noted that several different models have been used in literature to study transportation networks, and the observed topological properties often depend on the way the network is modelled. Most studies, including the ones referred to above, adopt a common network model where two nodes (airports or stations) are linked by an (undirected) edge if there exists a direct connection (flight or train) between the two nodes[1]. On the other hand, a *directed* network model was used in [11] to study the Chinese railway network, where the in-degree and out-degree of a node (station) were defined as the number of trains arriving at the station and the number of trains departing from the station respectively; the degree distribution of this network was observed to be a power-law. Transportation networks have also been modelled as bipartite networks (*e.g.* [8,13]) and weighted networks (*e.g.* [2, 4]).

Railways are one of the most prominent modes of transportation in many countries across the world and the complex topological properties of railway networks of different geographical regions have attracted the attention of the research community. The fractal structure of the railway network in Seoul was studied in [7] — the fractal dimension of the network was found to increase with time; also a comparison between the fractal dimension of the ensemble of stations and that of the railway lines was proposed as a measure of the quality of the transportation system. The underground (subway) railway networks of Boston and Vienna were studied as bipartite station-train networks in [8] — several topological metrics of the networks were

---

[1] The same model is used in this paper as well, as detailed in Section 2.

measured and compared with the corresponding theoretical predictions for random bipartite graphs using a generating function formalism. Various topological properties of the Chinese railway network have been studied in [9,11], whereas [13] used a weighted representation of the Chinese railway network to propose a metric to quantify the dependence of a station on another.

The Indian railway network (IRN) is one of the largest and busiest railway networks in the world, handling massive numbers of passengers and quantities of goods daily. Railways are the most popular means of long-distance transportation in India, hence the IRN is often described as the backbone of this nation's economy. However, analysis of the structure of the IRN has received considerably less attention, as compared to the railway networks of the European countries and China. To the best of our knowledge, the only study of the structure of the IRN from a networks perspective was in 2003 by Sen *et al.* [10], where the IRN was represented as a network of stations, two of which were linked by an edge if a train halted at both the stations. Hence the network considered in [10] was unweighted, and an edge simply indicated the presence of a train linking two stations.

However, a transportation network is specified not only by its topology of connections between stations, but also by the dynamics of the traffic-flow taking place in the network. Such networks display a large heterogeneity in the capacity of the connections; for instance, a significantly larger number of trains can be expected to link two major stations compared to that linking less important stations. Thus, in order to get a complete description of transportation networks, it is essential to take into account the amount of traffic-flow along the connections. Representing the amount of traffic on different links by edge-weights can yield observations that might be undetected by metrics based on topological information alone, as was demonstrated for the world-wide airport network in [4]. Hence, in this paper, we study the IRN as a *weighted* network of stations (nodes), where the weight of an edge indicates the number of trains linking two stations.

The present scenario in the transportation sector in India gives further motivation for a detailed analysis of the IRN — it is a commonly voiced opinion among economists that the current transportation network in India is too weak to meet the demands of the country's rapidly growing economy [14]. For instance, factors such as congestion, high traffic between major cities exceeding the planned capacity [15] and over-utilized railway tracks are resulting in trains having to travel at reduced speeds and carry lesser amounts of freight, thus increasing the cost and time of transportation. In this situation, a detailed understanding of the network-structure and traffic-flow is essential to identify the possible problems in the IRN; such a study can help in adopting effective extension policies in future and a better planning of the railway budget.

In this paper, along with measuring the basic topological properties of the present IRN, we also study the correlations of the amount of traffic with the network topology. In view of the recent criticism of the IRN (as discussed above), we also identify some of the potential points of congestion, which act as bottlenecks in transportation. We find that, along with the railway stations serving the major metropolitan cities, several relatively smaller stations (which have lesser resources such as railway tracks and platforms) are also potential points of congestion because of their geographical location, and hence more resources need to be allocated to these stations in order to efficiently handle the growing amounts of traffic. Thus our study suggests several guidelines for improving the performance of the IRN.

The rest of the paper is organized as follows. In Section 2, the representation of a railway network as a weighted network is described along with details on the method of collecting data on the IRN. The topological properties of the IRN are discussed in Section 3, while the analysis of the important nodes in the IRN is presented in Section 4. Section 5 concludes the paper.

2. Network construction

Two different, but related, approaches have commonly been adopted in the literature to represent a railway network as a complex network. In the context of a railway network, a train-route is a sequence of stations at which a train following that route is scheduled to halt. A railway network can be represented as a *bipartite train–station network* [8, 13] with one set S of nodes representing stations and the other set T of nodes representing the train-routes; there is an edge between $s \in S$ and $t \in T$ if and only if station s is a scheduled halt in the train-route t.

The more commonly used representation of a railway network is a network consisting of only station nodes, where two stations s_i and s_j are connected by an edge if there exists at least one train-route directly linking the two stations (in other words, if there exists at least one train-route such that both s_i and s_j are scheduled halts on that route). This representation is frequently used [9, 10, 2] to model different transportation networks since it directly captures some key facts on the connectivity of nodes (stations or airports) — for instance, the neighbours of a given station s_i are precisely those stations which can be reached from s_i by boarding a single train, while the shortest distance between an arbitrary pair of stations s_i and s_j is the minimum number of different trains that one needs to board to travel from s_i to s_j. In a weighted version of this *station–station network* representation, the weight of the edge between s_i and s_j is the number of train-routes on which both these stations are scheduled halts.

The station–station network representation can be derived from the bipartite train–station network by constructing a *one-mode projection* of the bipartite network over the station nodes, in which two stations $s_1, s_2 \in S$ are connected by an edge if they are linked to a common node $t \in T$ in the bipartite network (as depicted in Fig. 1). The weight of the edge linking s_1 and s_2 in the projection is thus the number of distinct nodes $t \in T$ to which both s_1 and s_2 are connected in the bipartite network (this is analogous to the number of train-routes on which both s_1 and s_2 are scheduled halts). This is the approach we use in this paper to construct the weighted station–station network representation of the IRN.

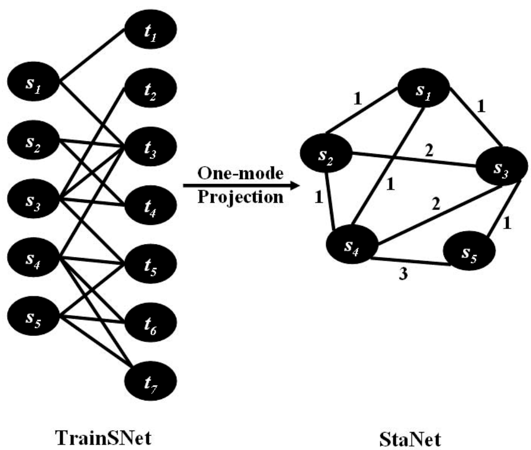


Fig. 1. Obtaining a weighted station–station network (StaNet) by one-mode projection of bipartite train–station network (TrainSNet).

The IRN is a dense network where the total number of stations and train-routes are of the order of tens of thousands. In this study, we consider only the 'express' train-routes and other long-distance train-routes (leaving out 'local' or suburban routes which traverse relatively short distances around major cities), and only those stations which are scheduled halts on at least one such train-route. We collected the data of the train-routes from the official website of Indian Railways (`www.indianrail.gov.in`) in May 2010. The website host information of 1072 express train-routes and 3041 stations which are scheduled stops on at least one such train-route. Since almost all train-routes in the IRN are bidirectional, the station–station network of the IRN is assumed to be undirected. We constructed the bipartite train–station network from the collected IRN data, and obtained the weighted station–station network as a projection of the train–station network. The next section presents the topological properties of the station–station network model of the IRN.

3. Topological analysis of the IRN

This section discusses the topological properties of the present-day IRN which is represented as a weighted station–station network. The network comprises of $N = 3041$ nodes (stations) and $E = 181,208$ edges representing presence of a direct link among stations. The average degree of the network is thus $2E/N = 119.177$ which indicates the average number of stations reachable from an arbitrary station via a single train. The network exhibits the small-world properties as already observed in [10]; the average shortest path length, measured as the average number of edges separating any two nodes in the network, is 2.53 which is very small compared to the network size N.

3.1. Degree and strength distributions

The degree distribution $p(k)$ of a network is defined to be the fraction of nodes having degree k in the network. Thus, if there are N nodes in a network and n_k of them have degree k, we have $p(k) = n_k/N$. The cumulative degree distribution $P(k)$, defined as the fraction of nodes having degree at least k, *i.e.*

$$P(k) = \sum_{i=k}^{\infty} p(i)$$

is preferred for analysis in practice, because the degree distribution is often noisy and there are rarely enough nodes having high degrees to get good statistics in the tail of the distribution, whereas the cumulative distribution effectively reduces the number of statistical errors due to the finite network size [16].

The degree of a node in the station–station network is the number of stations that can be reached from the given station via a single direct train, hence the node-degree is a measure of the connectivity of a station. The cumulative degree distribution $P(k)$ of the station–station network of the IRN (Fig. 2(a)) is observed to be an exponentially decaying distribution having the approximate fit $P(k) \sim \exp(-\alpha k)$ with $\alpha = 0.0082$; however, it deviates from the exponential nature for larger k. This exponentially decaying nature of the degree distribution for the IRN agrees with observations in [10]. The deviation for large degrees can be attributed to the high cost of adding links in the station–station network (in order to link a given station to a new neighbour, a new train-route needs to be introduced or a new station needs to be introduced in an existing train-route).

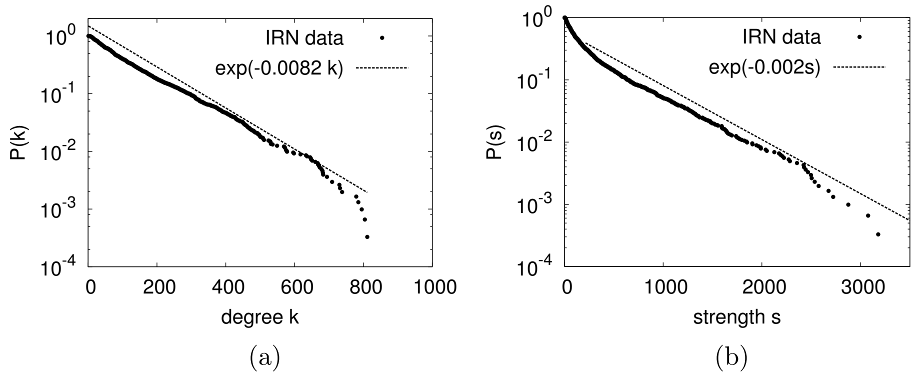


Fig. 2. (a) Cumulative degree distribution of the IRN. (b) Cumulative strength distribution of the IRN (both in semi-log scale, along with the exponential fits).

It may be noted that in contrast to the exponential degree distributions of most railway networks, the degree distributions of most airport networks [1, 2, 4] have been observed to be power-laws which can be explained by the preferential attachment growth model [17]. There can be several explanations for this variation, some of which are as follows.

First, there exist significant differences between the architecture of railway networks and that of airport networks. In an airport network, if two airports are connected by an air-route, it is rare for there to be an intermediate airport in the route. However, in a railway network, even if most train-routes are plausibly introduced between major end-stations *i.e.* high-degree nodes (in agreement to the preferential attachment model), several smaller stations are present between the terminal ones along the train-route, thus raising the degrees of the smaller stations as well. This may result in exponential degree distributions which are known to be more homogeneous compared to scale-free distributions [18]. Second, the networks having power-law degree distributions are characterized by the presence of a few hubs which are very high-degree nodes. A railway station can handle only a limited number of railway-tracks and trains (which limits the degree of the corresponding node in the network), while it is relatively easier for an airport to have direct connections with a large number of others; thus hubs are more likely to be present in airport networks than in railway networks[2].

The *strength*, or weighted degree, of a node in a weighted network is defined as the total weight of the edges adjacent to the node [4]. In the station–station network representation, the strength of a node (station) rep-

[2] For instance, each of the metropolitan cities in India, which need to have high connectivity with all parts of the country, are served by *multiple* railway stations in order to share the high amounts of traffic; this limits the degree of the individual nodes (stations) in the network.

resents the total number of different journeys that can be undertaken from that station (*i.e.* journeys to a different station or journey by a different train-route); hence, it is a measure of the available transportation from a station, which combines both the notions of connectivity and amount of traffic-flow (number of train-routes) through the station. For cities having large population and industrial production, the availability of transportation should match the high demands, hence the strength of such nodes should be high (along with high degree or connectivity). The distribution of node-strengths in the IRN (Fig. 2(b)) also exhibits an exponential nature similar to the degree distribution of the network.

3.2. Distribution of edge-weights

The edge-weights of the station–station network model the flow of traffic in the railway network — the weight w_{ij} of the edge between two station nodes i and j represents the number of train-routes which directly link both these stations; hence passengers (and freight) move more frequently along edges of higher weights. The analysis of edge-weights indicate a high level of heterogeneity in the traffic-flow in the IRN. The cumulative distribution of the edge-weights in the IRN (Fig. 3(a)) has an exponential fit $P(w) \sim \exp(-\alpha w)$ with $\alpha = 0.12$.

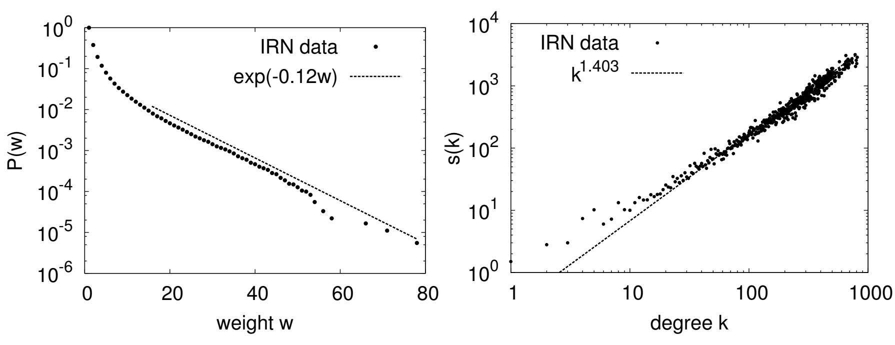


Fig. 3. (a) Cumulative distribution of edge-weights in the IRN, along with exponential fit (semi-log scale). (b) Average strength of nodes having degree k, as a function of k, along with the power-law fit (log–log scale).

3.3. Strength-degree correlation

To investigate the relationship between the degree and strength (weighted degree) of nodes, we plot the correlation between degree k and the average strength $s(k)$ of nodes having degree k in Fig. 3(b). $s(k)$ increases rapidly with k, following a power-law behaviour $s(k) \sim k^{\beta}$, with $\beta = 1.403$. In the absence of correlations between the edge-weights and the degree of adjacent vertices, the strength of a vertex would be simply proportional to its degree, yielding $\beta = 1$ [4]. The higher value of β for the IRN implies that node-strengths are strongly correlated with node-degree in the IRN and the strength of nodes grow faster than their degrees. This indicates that introduction of new trains on existing routes (*i.e.* increasing the weights of existing edges, thus increasing the strength of nodes) is more common in the IRN compared to construction of new train-routes that link a station with new neighbours (*i.e.* increasing the degree of nodes). Similar trends have also been observed for the Chinese railway network [11].

3.4. Weight-degree correlation

The strength-degree relationship can also be characterized by the correlation of weight w_{ij} of the edge between nodes i and j, with the degrees k_i and k_j of the end-points i and j, as studied in Fig. 4(a). It is evident that the links between high-degree nodes (important stations having high connectivity) tend to have high values of traffic in the IRN. Such high-traffic links between the major cities are generally referred to as *trunk routes*.

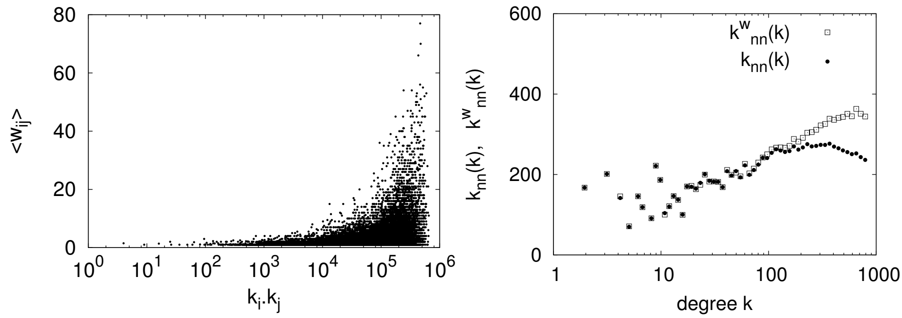
3.5. Degree-degree correlations

Another parameter used to investigate the network architecture is the correlation among degrees of neighbouring nodes, which can be observed from the average nearest-neighbour degree $k_{nn}(k)$ of nodes having degree k (Fig. 4(b)). It is observed that $k_{nn}(k)$ remains the same on the average over a significant range of degrees, implying the absence of major correlations among the nodes of different degrees. This behaviour of $k_{nn}(k)$ agrees with the results for the IRN in [10].

However, a completely different perspective is gained regarding the assortativity of the IRN by taking edge-weights into consideration. We use a weighted variant of the average nearest-neighbours degree, k_{nn}^{w}, as defined by Barrat *et al.* in [4]. For a given node i, $k_{nn,i}^{\mathrm{w}} > k_{nn,i}$ if the edges adjacent to i having the larger weights are connected to the neighbours (of i) having larger degree, and $k_{nn,i}^{\mathrm{w}} < k_{nn,i}$ in the opposite case. Analogously, the behaviour of $k_{nn}^{\mathrm{w}}(k)$ (the average weighted nearest-neighbour degree of nodes having degree k) indicates the weighted assortative or disassortative properties, taking into account the flow of traffic among the stations of the network.

Fig. 4(b) compares the variations of $k_{nn}(k)$ and $k_{nn}^{\mathrm{w}}(k)$ with degree k (using logarithmic binning of k-values for better visibility); $k_{nn}^{\mathrm{w}}(k)$ shows a pronounced assortative behaviour, implying that high-degree stations tend to connect with other high-degree stations, and the amount of traffic (weight) along such links between high-degree nodes tend to be high as well. Similar trends have also been observed for the world-wide airport network [4].

The topological assortativity coefficient, as defined by Newman [19], comes out to be 0.0813 for the IRN, indicating that the topology of the IRN is weakly assortative in nature. The definition by Newman was extended for weighted networks by Leung *et al.* [20]; the weighted assortativity coefficient for the IRN is observed to be 0.2378, indicating that a pronounced assortative behaviour when the traffic-flow is taken into consideration.



(a) (b)

Fig. 4. (a) Correlation of edge-weights and product of end-point degrees in the IRN (semi-log scale). (b) Average degree of nearest neighbours $k_{nn}(k)$ and average *weighted* degree of nearest neighbours $k^{\mathrm{w}}_{nn}(k)$ of nodes having degree k, using logarithmic binning of degrees (semi-log scale).

3.6. Clustering coefficient

Fig. 5(a) plots the average clustering coefficient $cc(k)$ of nodes having degree k as a function of k; $cc(k)$ remains at a constant value close to unity for small k and then shows an almost power-law decay at larger values of k. This observation, which agrees with results in [10], can be explained as follows. All stations on the same train-route are linked to form a clique in the station–station network. The smaller stations (having low degrees) in the IRN are served by very few train-routes, hence they are linked only to other stations on these train-routes (other nodes in the clique), thus resulting in a clustering coefficient near to unity for the nodes with low degrees. On the other hand, major stations (having high degrees) are served by a large number of train-routes, hence these stations are linked with other geographically distant stations in diverse parts of the country, which themselves do not tend to be connected, thus lowering the average clustering coefficient for nodes with higher degree.

It has been shown [21] that a power-law decay of $cc(k)$ with degree k is an evidence of hierarchical organization in a network, which implies that low-degree nodes belong to interconnected communities. Thus an inherent hierarchy is evident from the structure of the IRN.

For a weighted network, the clustering coefficient has been re-defined [4] to incorporate edge-weights, in order to take into account the importance of the clustered structure based on the amount of traffic actually found in the cluster. Analogous to $cc(k)$, $cc^{\mathrm{w}}(k)$ is defined as the *weighted* clustering coefficient averaged over all nodes of degree k. Fig. 5(b) compares the variations of $cc(k)$ and $cc^{\mathrm{w}}(k)$ with degree k; both versions have similar values for low degrees, however $cc^{\mathrm{w}}(k)$ lies consistently above the unweighted $cc(k)$ for intermediate and higher degrees, indicating that most of the traffic (*i.e.* edge-weights) in the IRN is accumulated on interconnected groups of high-degree nodes. Further, the variation of $cc^{\mathrm{w}}(k)$ is much more limited in the whole spectrum of k compared to that of $cc(k)$, implying that high-degree stations have a tendency to form interconnected groups with high-traffic links (trunk routes), thus balancing the reduced topological clustering.

The clustering coefficient C of the network, which is the average of the clustering coefficients for all nodes, is 0.733, while the corresponding weighted clustering coefficient C^{w} comes out to be 0.789. $C^{\mathrm{w}} > C$ again indicates that the major stations (high-degree nodes) form high traffic corridors among themselves.

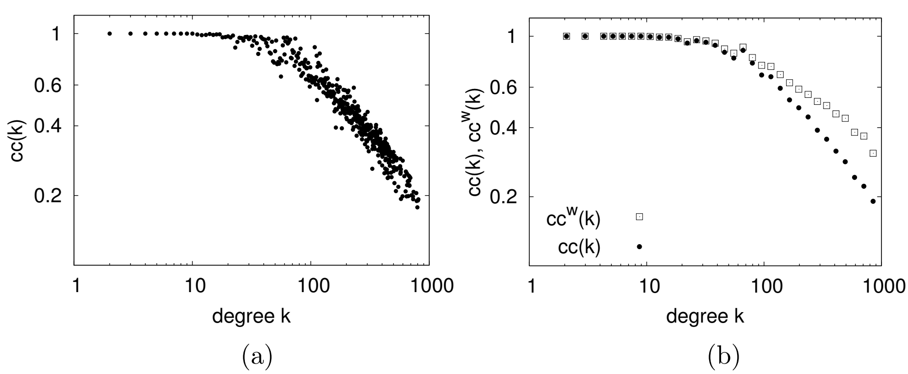


Fig. 5. (a) Average (unweighted) clustering coefficient $cc(k)$ of nodes having degree k, as a function of k. (b) Average unweighted and weighted clustering coefficients as function of degree, using logarithmic binning of degrees (log–log plots).

From the above discussions, it is evident that considering the edge-weights in the station–station network of the IRN has led to a more complete reflection of the properties of the network, compared to what can be obtained from the network topology alone. This justifies our motivation of studying the IRN as a weighted network. The practical implications of the results obtained in this section in context of the IRN are discussed later in Section 5.

4. Identifying major stations in the IRN

In this section, we identify the major stations in the IRN from the station–station representation of the network. Since the node-degree is a measure of the connectivity of a station, the nodes with high degrees are evidently important in the network (this measure of importance of nodes is known as degree centrality). The top 10 stations in the IRN based on node-degree are listed in Table I. These stations can be classified into two groups based on their geographical locations, as shown in Fig. 6(a):

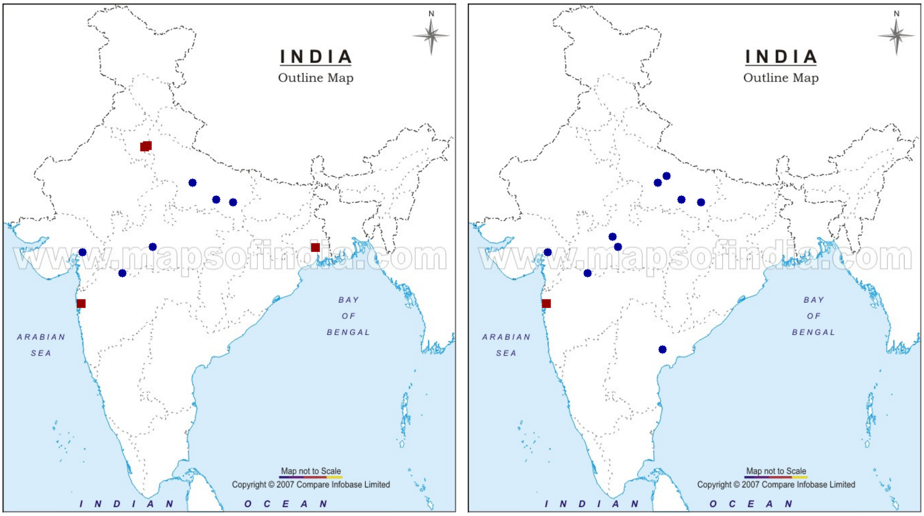
- stations that are located in close vicinity to the metropolitan cities in India (*e.g.* Howrah near Calcutta, Kalyan near Mumbai),
- stations that are located in the central parts of the country or at the meeting points of railway lines connecting different zones (for instance, the left-most circle in Fig. 6(a) is at Vadodara junction that is used by most train-routes linking the western zone of India with the southern, central and eastern zones).

Analogously, the nodes having high values of strengths (weighted degrees) are the ones which handle a high amount of traffic. Table I lists the top 10 stations in the IRN based on node-strength, and Fig. 6(b) shows their geographical locations.

TABLE I

Top 10 stations in the IRN on the basis of node-degree and node-strength. The stations located in vicinity of metropolitan cities are marked by (*).

Top stations w.r.t. degree	Top stations w.r.t. weights
Kanpur Central	Itarsi
Howrah (*)	Vijayawada
Kalyan (*)	Kanpur Central
Ghaziabad (*)	Vadodara
Itarsi	Mughal Sarai
Varanasi	Kalyan (*)
Vadodara	Bhusawal
Allahabad	Lucknow
Bhuwsawal	Bhopal
Hazrat Nizamuddin (*)	Allahabad



(a) (b)

Fig. 6. The top 10 stations in IRN based on (a) *degree*, (b) weighted degree or *strength*; Stations in the vicinity of the metropolitan cities in India marked with (red) squares, other stations marked with (blue) circles.

Interestingly, all except one of these stations are located in the central regions of the country or at the junction of railway lines connecting different zones. Though these stations handle large amounts of traffic, they often do not have as much resources (*e.g.* platforms, railway tracks) as the stations located in close vicinity of the metropolitan cities. For instance, Howrah, located near metropolis Calcutta and having the highest node-degree among metropolitan stations, has 23 platforms and 25 tracks, while the two stations with the highest node-strengths, Itarsi (located at the centre of the country) and Vijayawada (located on the lines linking south zone with east and north zones), have only 7 and 10 platforms respectively (as given in the Wikipedia articles on these stations)[3]. Hence these stations are potential points of congestion in the network.

Further, Fig. 6(b) shows that a majority of the stations with high strengths are limited to two specific regions — in the states of Uttar Pradesh and western parts of Madhya Pradesh. Comparing the locations of the metropolitan cities shown in Fig. 6(a) with Fig. 6(b), it is seen that these regions lie in between the metropolitan cities of India (between Calcutta and Delhi, and between Mumbai and Delhi respectively), and hence these

[3] It is to be noted that the stations in close vicinity to metropolitan cities, such as Howrah, handle a significant amount of suburban railway traffic as well, but our analysis does not take into account suburban train-routes (as stated in Section 2).

regions contain several trunk-routes linking the metropolitan cities. Large amounts of resource and manpower are required in these regions for efficient management of the excessive traffic.

5. Discussion and conclusion

In this paper, we studied the Indian Railway Network as a weighted complex network of stations, where the edge-weights represent the amount of traffic between two stations. We observed that the IRN has exponential distributions of node-connectivity and traffic-flows. Also the major stations (high-degree nodes) tend to be linked among themselves and most of the traffic in the IRN flows among these high-degree nodes.

Our findings from the topological analysis of the IRN (considering network structure only) agree qualitatively with the findings in the only prior study of the IRN in 2003 [10]. Hence it is evident that the basic topological characteristics of the IRN, such as the degree distribution, degree correlations and clustering coefficient, have remained almost unchanged over the last decade. This corroborates the criticism that the IRN has *not* expanded structurally as much as was required to handle the increasing demands of traffic in the recent years.

The reasons for the recent concern over the IRN become even more apparent by our analyses combining the topology with the traffic-flow in the IRN. The node-strengths (weighted degree) grow faster compared to node-degrees in the IRN (Fig. 3(a)) implying that the construction of new links between stations has been significantly less than the introduction of new trains along existing links. Considering the limited capacity of links to handle trains, this shows the need for construction of new links among stations[4]. The correlation of edge-weights with the degrees of the adjacent nodes (Fig. 4(a)) corroborates another reported cause for concern in the present-day IRN — traffic on the trunk-routes between the large cities is far exceeding the planned capacity, which means that trains have to travel more slowly and the railway tracks wear out faster than intended [14, 15]. Hence new train-routes can be introduced to connect the larger cities; also, the tracks in the existing trunk-routes should be replicated to handle the large amounts of traffic. We also identify some of the smaller stations that handle large amounts of traffic (Fig. 6(b)). The infrastructure at these stations should be improved to ease the congestion in the network. Thus our study provides several guidelines for improving the performance in the IRN.

[4] This has been recognized by the Indian Railways authority as well, and it has been announced [15] that 25,000 kilometres of new railway-tracks would be constructed by 2020, which is far greater than the average rate of construction of tracks till now.

REFERENCES

[1] W. Li, X. Cai, *Phys. Rev. E* **69**, 046106 (2004).

[2] G. Bagler, *Physica A* **387**, 2972 (2008).

[3] R. Guimera, S. Mossa, A. Turtschi, L.A.N. Amaral, *Proc. Natl. Acad. Sci. USA* **102**, 7794 (2005).

[4] A. Barrat, M. Barthélemy, R. Pastor-Satorras, A. Vespignani, *Proc. Natl. Acad. Sci. USA* **101**, 3747 (2004).

[5] H. Lu, Y. Shi, *Tsinghua Sci. Technol.* **12**, 204 (2007).

[6] J. Sienkiewicz, J.A. Holyst, *Phys. Rev. E* **72**, 046127 (2005).

[7] K.S. Kim, L. Benguigui, M. Marinov, *Cities* **20**, 31 (2003).

[8] K. Seaton, L. Hacket, *Physica A* **339**, 635 (2004).

[9] C.-M. Liu, J.-W. Li, *Front. Phys. China* **2**, 364 (2007).

[10] P. Sen *et al.*, *Phys. Rev. E* **67**, 036106 (2003).

[11] W. Li, X. Cai, *Physica A* **382**, 693 (2007).

[12] C. Li-Ping *et al.*, *Chinese Phys. Lett.* **20**, 1393 (2003).

[13] Y. Wang *et al.*, *Physica A* **388**, 2949 (2009).

[14] *Have India's Clogged Rail Lines Stalled Progress?*, 2010, `http://www.ndtv.com/news/india/indias-clogged-rail-lines-stall-economic-progress-31977.php`

[15] *Indian Railways Vision 2020*, 2009, published by the Ministry of Railways, Government of India.

[16] M.E.J. Newman, *SIAM Rev.* **45**, 167 (2003).

[17] A.-L. Barabási, R. Albert, *Science* **286**, 509 (1999).

[18] H. Hu, X.-F. Wang, *Physica A* **387**, 3769 (2008).

[19] M.E.J. Newman, *Phys. Rev. Lett.* **89**, 208701 (2002).

[20] C.C. Leung, H.F. Chau, *Physica A* **378**, 591 (2007).

[21] E. Ravasz, A.L. Barabasi, *Phys. Rev. E* **67**, 026112 (2003).