**Analysis and Visualization of Delhi Metro Rail System**

**Abishek Sriram S**

*1VIT, Chennai, India*

*abisheksriram.senthil2019@vitstudent.ac.in*

**Abstract**

The national capital, Delhi has a high population. The population density of the national capital is also significantly high. As the population is high, many people own vehicles (two-wheelers and four-wheelers) leading to traffic congestion/traffic jams and inability to commute from one location to another. Through the analysis of the Delhi Metro Rail Network, we can predict the most congested places(nodes) and suggest some measures to reduce the congestion and make it easy to commute. Traffic reduction benefits all other aspects connected with it such as environmental pollution, time effectiveness, fuel wastage, etc. It also promotes sustainability for the future generations.

The analysis shows the centrality measures such as degree centrality, betweenness centrality of the whole network and centrality with respect to a single node. The computation is done by taking every station in the Delhi metro map as nodes and the connections as edges.

Using the above conventions, a network is formed and the centrality measures are computed for further analysis. Based on the results, the node with the highest centrality measure is the most congested network, meaning that many people commute in that network and that node has to be decongested. The traffic from these nodes can be diverted to those nodes which have low value of centrality measure. This process is repeatedly done in these nodes till the centrality value of the network is significantly lesser than the initially obtained centrality value.

Measures such as edge betweenness and betweenness of the nodes can be used to identify the key station in the network.Eigenvalue centrality can be used to support our conclusions.

Division of network into communities or modules in our scenario can be used to analyse subsets of the network.

All simulations, calculations and visualizations were done using the Gephi tool. The Delhi Metro map was downloaded from the metro website on the internet and the corresponding nodes and edges list was prepared for visualization.

**Literature Review**

Paper Citation: Yasir Tariq Mohmand, Aihu Wang, "Complex Network Analysis of Pakistan Railways", *Discrete Dynamics in Nature and Society*, vol. 2014, Article ID 126261, 5 pages, 2014. <https://doi.org/10.1155/2014/126261>

This paper analysed the Pakistan Railways and measured the betweenness and degree centrality of the network. The potential inferences made were the identification of congestion points in the network and the stations or group of stations driving the network by providing interconnectivity.

Paper Citation: Li Wang, Min An, Limin Jia, Yong Qin, "Application of Complex Network Principles to Key Station Identification in Railway Network Efficiency Analysis", *Journal of Advanced Transportation*, vol. 2019, Article ID 1574136, 13 pages, 2019. https://doi.org/10.1155/2019/1574136

This paper provides a structured and sequential application of complex network analysis to identify the stations with greater significance . It proposed a new method to identify the efficiency of a transport network by identifying the key station based on two network models of RPN and TNF.

Paper Citation: Ghosh, Saptarshi & Banerjee, Avishek & Sharma, Naveen & Agarwal, Sanket & Ganguly, Niloy & Bhattacharya, Saurav & Mukherjee, Animesh. (2011). Statistical analysis of the Indian Railway Network: A complex network approach. Acta Physica Polonica B, Proceedings Supplement. 4. 10.5506/APhysPolBSupp.4.123.

This paper analysed the Indian Railway System various network centrality measures. THe network is a weighted network, various measures such as eigenvalue centrality and betweenness centrality were used in identification of most important stations in the network.A correlation between the traffic and network topology using the degree distribution of the network were established.These results along with connectivity and traffic-flow were used to identify the congestion in the network.

Paper Citation:Kushal Kanwar, Harish Kumar, Sakshi Kaushal, Complex network based comparative analysis of Delhi Metro network and its extension,Physica A: Statistical Mechanics and its Applications,Volume 526,2019,120991,ISSN 0378-4371,

<https://doi.org/10.1016/j.physa.2019.04.227>.

This paper studied the network characteristics based Comparative analysis of Delhi Metro and its extension. It proposed two hybrid network attack protocols to analyse the vulnerabilities of the network by performing betweenness centrality measures of all edges in the network.

Paper Citation: Yingying Xing , Jian Lu , and Shendi Chen

"Weighted Complex Network Analysis of Shanghai Rail Transit System"

Volume 2016 |Article ID 1290138 | https://doi.org/10.1155/2016/1290138

In this paper, the Shanghai RTS, as well as passenger flows, is investigated by using

complex network theory. Both the topological and dynamic properties of the RTS network

are analyzed and the largest connected cluster is introduced to assess the reliability and

robustness of the RTS network. Simulation results show that the distribution of nodes

strength exhibits a power-law behavior and Shanghai RTS network shows a strong weighted rich-club effect.

Paper Citation: E. Frutos Bernal , A. Martín del Rey and P. Galindo Villardón

"Analysis of Madrid Metro Network: From Structural to HJ-Biplot Perspective"

Appl. Sci. 2020, 10(16), 5689; https://doi.org/10.3390/app10165689

This paper analyzes Madrid metro network using the most important centrality measures,

some structural coefficients and robustness indicators. Once the results of the network analysis were obtained, they were inspected using multivariate analysis techniques. Specifically, the HJ-Biplot was employed to analyze correlation between centrality measures and also to classify stations according to their centrality.

Paper Citation: Wei Wang, Kaiquan Cai, Wenbo Du, Xin Wu, Lu (Carol) Tong, Xi Zhu, Xianbin Cao

"Analysis of the Chinese railway system as a complex network"

https://doi.org/10.1016/j.chaos.2019.109408

This paper conducts a series of complex network analyses on the Chinese railway network (CRN), which is a small-world network with two-regime power-law distributions. This paper also encapsulates CRN into a multi-layer infrastructure with a core-periphery structure via the ‘‘weighted k-core decomposition” method. The reasonability of cities contained in the core layer is generally consistent with the hub set identified by the gravity model.

Paper Citation: Mohieddin Jafari and Sayed Mohammad Fakhar

"Network Centrality Analysis of Tehran(Iran) Urban and Suburban Railway System"

https://arxiv.org/ftp/arxiv/papers/1802/1802.06219.pdf

In this paper, they have reconstructed the network model of TUSRS(Tehran Urban and Suburban Railway System) and tried to find central nodes (i.e. stations) within TUSRS network. The central stations are inferred based on several well-known measures.

Finally, we compared the potential of all lines and some major stations individually.This complex transportation system contains 168 links between each station pair and 20 cross-section and Y-branch stations among all eight lines.

Paper Citation: BRUNEL Julien, MARLOT Grégoire and PEREZ Maria

“Measuring Congestion in Rail Sector: The French Experience”

<http://www.wctrs-society.com/wp-content/uploads/abstracts/rio/selected/3315.pdf>

In this paper, they have visualised the dataset from an internal database of Réseau ferré de France. They calculate the congestion based on the delay of the trains. An economic analysis of congestion in rail transport is seen. They calculate the direct and indirect effect of congestion in the network.

Paper Citation: Shaopei Chen, Dachang Zhuang - “Evolution and Evaluation of the Guangzhou Metro Network Topology Based on an Integration of Complex Network Analysis and GIS”

In this paper they have analyzed the metro network of Guangzhou for providing the scientific basis and significant decision-making support to the planning and operation management of GMN. They have proposed various calculations based on the topological structures. The analysis is done based on the degree centrality and node betweenness centrality.

**Proposed Work**

**Data Extraction**

For this project,we compiled the metro network datasets of New Delhi City. Since we wanted network dataset to generate the network we had to manually transcribe the network then generate the graphs.Our source of information includes the network route map and list of stations provided by DMRC(Delhi Metro Rail Corporation).This involved identifying each station in the network and assigning a unique ID to each station. More information about the traffic of passengers using the metro network and geographical significance of each station might be provided for better estimation of the network.

In this project we considered the network of stations to be unweighted, undirected. N is a set of stations containing the unique ID mapping to the station label or the station name.Each station in the DMR(Delhi Metro Railway) is a node in our network . Based on the information source we complied and generated the edges. Edge simply represents a train or sequence of trains connecting each pair of stations.

**Analysis and Measures**

As of now, we have the Graphical representation of the Delhi Metro Railway.

We have used Gephi software for analysing and visualizing the network.

Through this, we have got some of the Network properties of our system like the Average Degree, the Diameter, Density, Modularity, the Average Path Length as well as the

Clustering Coefficient.

Furthermore, we will be calculating some other network properties as well to go along

with the centrality measures. Then we will form some conclusions based on the analysis done so far and try to provide some specific details on the areas of improvement.

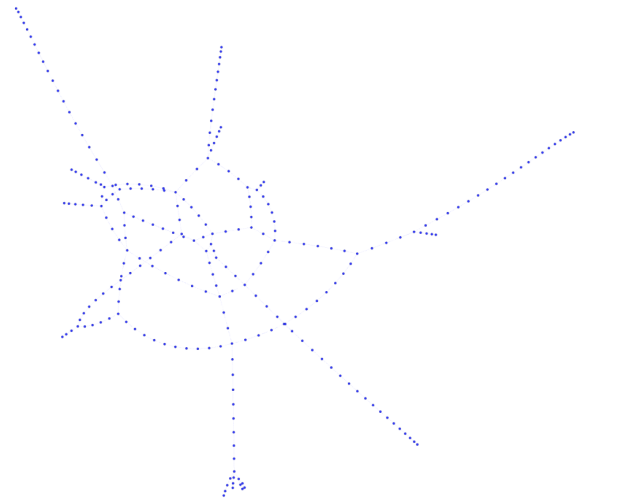
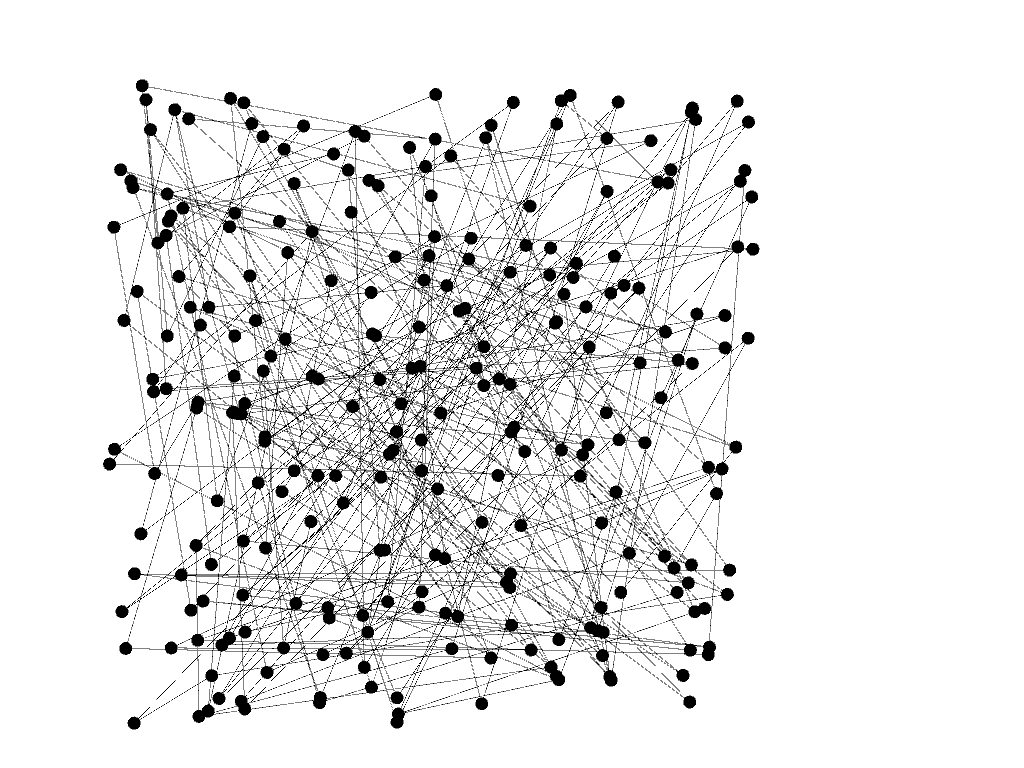
There are lots of conclusions we can draw from our visualization. For instance,based on the degree centrality of the network a comment on the structure of the network and the topology of the network can be made. Similarly, if we consider the Diameter and the Average Path length ,the relative comparison between these fields can help us draw conclusions on the spread of network.Also, if we look at the Clustering Coefficient, a lower clustering coefficient means a weaker linked network. Modularity divides the network into modules. As a result, the division which gets the maximum modularity gives us the more important community.

In the Delhi Metro Railway network, the Betweenness Centrality is the most important centrality measure. This is because, if the betweenness centrality is high, it means that that particular station is more important because more trains will be passing through that station.

It also means that there are more stations connected to that particular Station.

Thus, based on these deductions, we will be able to suggest measures to reduce the congestion, traffic, and pollution prevalent in the current Delhi Metro Railway System.

**Results and Discussion**

**Figure 1**: Graph Representation of Delhi Metro Railway **Figure 2:** Graph Representation of Delhi Metro Railway (Yifan hu Layout)

Yifan hu Layout algorithm is a combination of a multilevel force-directed algorithm and octree technique.

**Network Properties**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Average Degree | Diameter | Density | Modularity | Avg. Path Length | Clustering Coefficient |
| 2.123 | 66 | 0.008 | 0.836 | 19.665 | 0.010833 |

**Degree**

The Degree (k) of a node is the number of nodes directly connected to it. Our interpretation of degree would be the number of stations connected directly to a given station. The average degree is 2.123, this possibly represents the linear connectivity of the stations in the network. We can conclude that most of the stations in this network are connected in a linear fashion and the number of interconnections or junctions is less.

**Diameter**

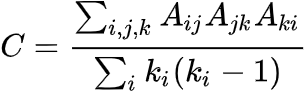
The diameter of a network is the largest shortest path between any pair of nodes in the network. It describes the maximum number of edges required to get from one point on the subway to another over all possible trips.

**Average Path Length**

The average path length along with diameter of a network can be used to estimate how high or how low the travel time is going to be.The average path length in our case is relatively smaller than diameter of the network.This indicates that the estimate travel time would considerably less.If the edges were weighted based on the distance of the edge much precise conclusions would be made.

**Clustering Coefficient**

The Global Clustering Coefficient of a network is based on the triplets of nodes.A low clustering coefficient indicates weak links in the network.This weak links can be helpful in division of networks.Clustering coefficient is calculated using the adjacency matrix.



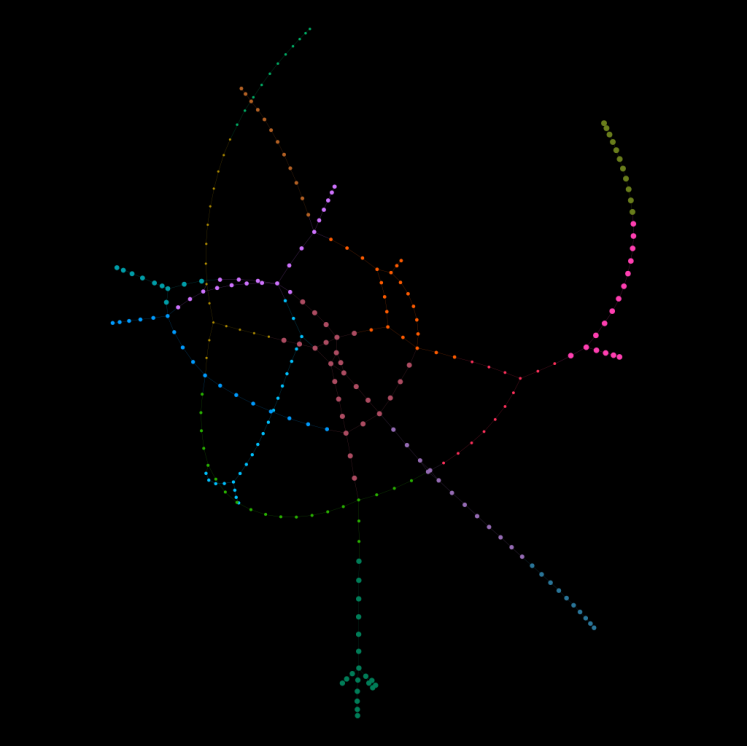
where



The global clustering coefficient for this network was calculated to be 0.010833.This low coefficient indicates the presence of weak links in the network.High clustering coefficient generally indicates the presence of “small world” phenomenon, but in our network the average path length and low clustering coefficient and the linear structure of the network creates a differentiation between social networks and transport network.

**Modularity**

Modularity of a network is a measure of its structure.It measures the strength of division of the networks into communities or modules.Community detection in a network involves maximizing the modularity of the network.Using the algorithm proposed in [1] we were able to detect 16 communities in the network.



**Figure 3**: 16 communities obtained by maximizing modularity

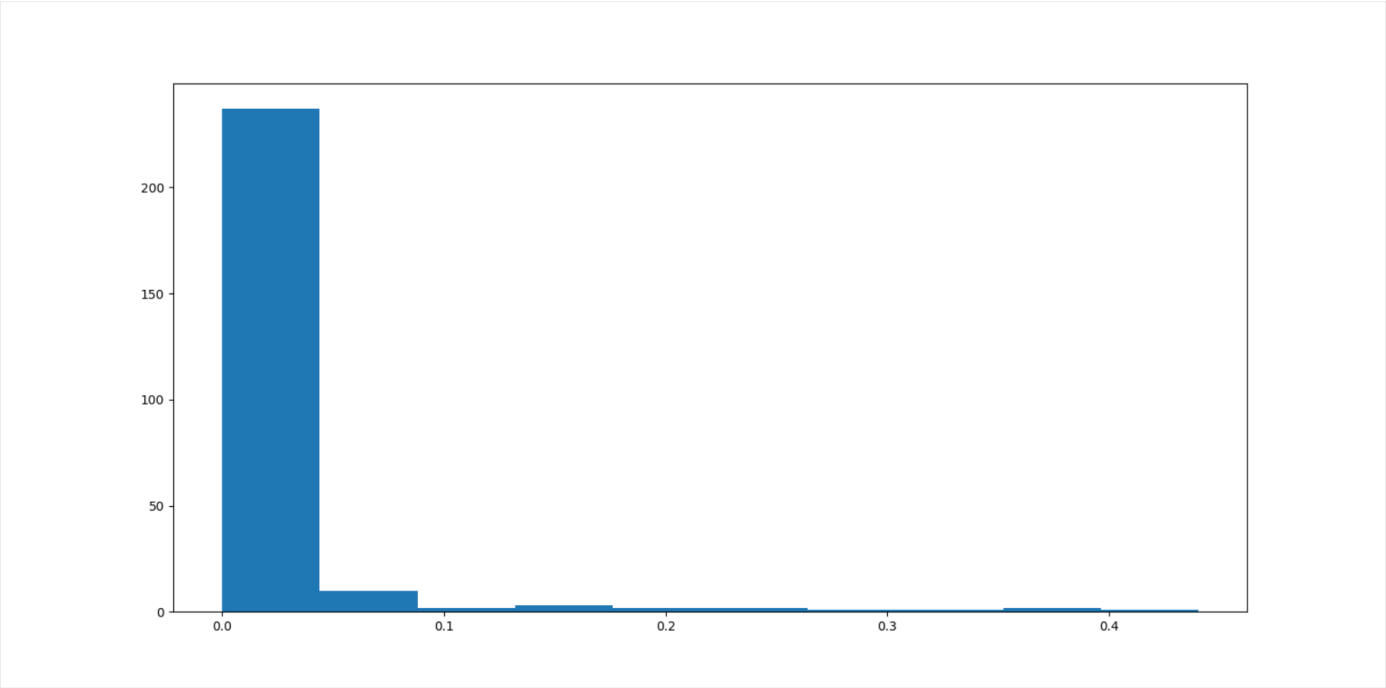
**Eigenvector Centrality**

Eigenvector Centrality is an important measure of any social network. It provides insights into which node is the most influential node in the network. In terms of information spread the most important nodes play a vital role. The importance of a node is decided by how important are its immediate neighbors.So we assume that the importance of each node is 1 at the start and iterate to obtain the eigenvalue Centrality. The real question is how important is eigenvector centrality to understand a transport metro.

|  |  |
| --- | --- |
| **Station name** | **Eigenvector Centrality** |
| Patel Chowk | 0.4403026468863414 |
| Rajiv Chowk | 0.38242542009362246 |
| Janpath | 0.3590355616989419 |
| Mandi House | 0.32084073064819996 |
| Khan Market | 0.28941444110324294 |
| Barakhambha Road | 0.253739565005332 |
| Central Secretariat | 0.23396181014140172 |
| New Delhi | 0.20282113955822056 |
| Udyog Bhawan | 0.18947012168131228 |
| Ramakrishna Ashram Marg | 0.16306931269499442 |

**Table 1:** Top 10 Stations with the most eigenvalue centrality

From Table1 we can identify that the most influential node in the network is Patel Chowk, but we can’t make conclusions based on only the eigenvalue centrality because we first need to understand what importance of a node in a transport network really means, we propose that the most important node in a transport network would the node that connects the network.So the betweenness of a node is to be considered while identifying the most important node in a transport network because in a transport network the subject of importance would be how well connected the network is rather than how important a node is.Figure 4 depicts the eigenvalue centrality of Delhi Metro Rail System. We can see that a distribution is concentrated in the 0 to 0.1 range and very dilute elsewhere.From this distribution we can conclude that the nodes with

****

**Figure 4:** Distribution of Eigenvalue Centrality among the Delhi Metro Rail System

**Betweenness Centrality**

Betweenness Centrality is a measure of centrality in a graph based on the shortest path or geodesic distances.In a network it represents a degree to which nodes stand between each other.The Betweenness Centrality of a node

v can be calculated as the summation of the ratio of the number of shortest paths between two nodes v to the number of shortest paths between two nodes including v in the path.



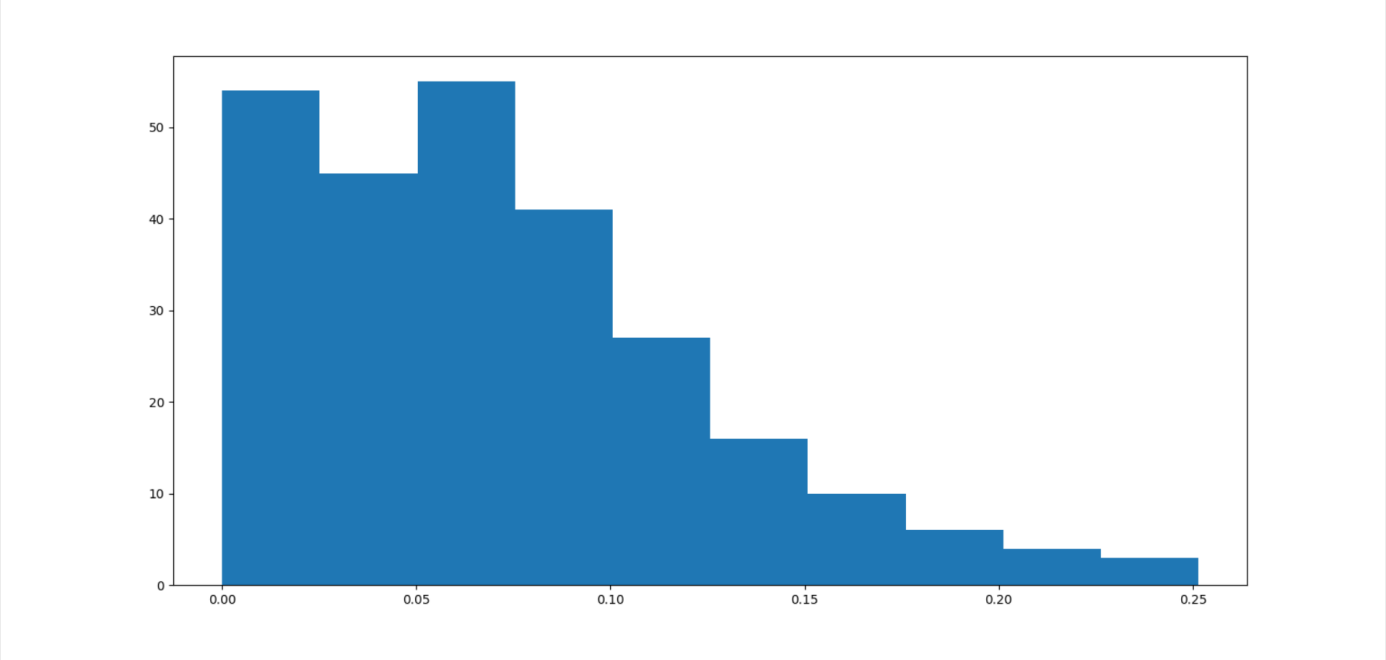
Betweenness Centrality plays an important role and is more important than eigenvalue centrality in the identification of the most important node in terms of network security in a transport network. The most important node in a transport network can be identified as the node which maintains the connectivity of the network and if this node and its corresponding edges are removed from the network various effects like increase in average shortest path, which inturn increases the cost of transportation and fuel consumption, and the breaking of network to self containing components which reduces the coverage of network takes place. But betweenness alone is not enough for such a conclusion because in calculation of betweenness centrality only the shortest path is considered but in real world scenarios to make the shortest path journey we might need to change trains as well which also make the junction point where this exchange point happens more vulnerable or important.

Table2 shows the list of top ten stations with the highest betweenness centrality among the entire network.Since Kashmere Gate has the highest Betweenness Centrality and it's also a junction. From the network map and station information provided by the Delhi Metro System, Kashmere is a transfer station between Red,Yellow and Violet lines of the Metro.This line each constitute a full trip i.e.with out changing trains we can travel through the entire line.

|  |  |  |
| --- | --- | --- |
| Station Name | Betweenness Centrality | Junction |
| Kashmere Gate | 0.25145282645282646 | 1 |
| Rajiv Chowk | 0.24163647163647195 | 1 |
| Hauz Khas | 0.22832194832194846 | 1 |
| Mayur Vihar Phase-1 | 0.225990495990496 | 1 |
| Dilli Haat-INA | 0.21516285516285516 | 1 |
| Botanical Garden | 0.21192951192951218 | 1 |
| Lajpat Nagar | 0.2037224037224038 | 1 |
| Golf Course | 0.19293139293139294 | 0 |
| New Delhi | 0.1916607266607269 | 1 |
| Noida City Center | 0.18684288684288686 | 0 |

**Table 2:** Top 10 Stations with the highest Betweenness Centrality among the Delhi Metro Rail System

Figure 5 shows the distribution of Betweenness Centrality of Delhi Metro Rail System.In contrast to the eigenvalue centrality distribution which is highly dense at lower values near 0.1, we can see that the distribution of Betweenness Centrality is distributed. But we can still see similarities among them. As the range increases, in both the distributions the number of stations falling in the range is decreasing and at a much higher rate. One conclusion we can make is that if the rate of change of distribution follows a linear fashion the importance of each node would be roughly the same hence the overall vulnerability of the network decreases.



**Figure 5:** Distribution of Betweenness Centrality of Delhi Metro Rail System

**References**

1. Blondel, Vincent & Guillaume, Jean-Loup & Lambiotte, Renaud & Lefebvre, Etienne. (2008). Fast Unfolding of Communities in Large Networks. Journal of Statistical Mechanics Theory and Experiment. 2008. 10.1088/1742-5468/2008/10/P10008.
2. Kushal Kanwar, Harish Kumar, Sakshi Kaushal, Complex network based comparative analysis of Delhi Metro network and its extension,Physica A: Statistical Mechanics and its Applications,Volume 526,2019,120991,ISSN 0378-4371.
3. Wasserman, S., & Faust, K. (1994). *Social Network Analysis: Methods and Applications* (Structural Analysis in the Social Sciences). Cambridge: Cambridge University Press. doi:10.1017/CBO9780511815478
4. Joseph Bryan, Glen Weisbrod - ASSESSING RAIL FREIGHT SOLUTIONS TO ROADWAY CONGESTION: FINAL REPORT (2006) .
5. Kumari, Menka. (2020). Evaluation of Mass Rapid Transit System (MRTS): A Case Study of Delhi. 10.1007/978-981-15-1205-6\_19.
6. Wei Wang, Kaiquan Cai, Wenbo Du, Xin Wu, Lu (Carol) Tong, Xi Zhu, Xianbin Cao,"Analysis of the Chinese railway system as a complex network"

https://doi.org/10.1016/j.chaos.2019.109408