Vienna Subway Network Analysis

Maddukuri Nivas[1],Mattapally Sai Nithin[2],Kovvuri Uday Surya Deveswar Reddy[3]

[1] VIT,Chennai,India

[2] VIT,Chennai,India

[3] VIT,Chennai,India

Email:

maddukuri.nivas2019@vitstudent.ac.in,

mattapally.sainithin2019@vitstudent.ac.in,

udaysurya.deveswar2019@vitstudent.ac.in

Abstract

Austria's city, Vienna, has a high population. The population thickness of Vienna is also essentially high. As the populace is high, numerous individuals possess vehicles driving leads to traffic jams and the failure to commute from one location to another. Through the examination of the Vienna Subway Network, we are able to foresee the foremost congested places(nodes) and propose a few measures to decrease the blockage and make it simple to commute. Traffic decrease benefits all other angles associated with it such as time viability, fuel wastage, etc. It moreover advances maintainability for future eras.The examination shows the centrality measures such as degree centrality, betweenness centrality of the full network, and centrality with regard to a single hub. The computation is done by taking each station within the Vienna subway outline as hubs and the associations as edges.Utilising the above traditions, an arrangement is formed and the centrality measures are computed for further analysis. Based on what comes about, the hub with the most elevated centrality degree is the foremost congested network, meaning that numerous individuals commute in that network which hub must be decongested. The traffic from these nodes can be redirected to those nodes which have a low esteem of centrality degree. This process is more than once done in these hubs till the centrality esteem of the network is altogether lesser than the initially gotten centrality value.Measures such as edge betweenness and betweenness of the nodes can be utilised to distinguish the key station within the network. Eigenvalue centrality can be utilised to back our conclusions. Division of networks into communities in our situation can be utilised to analyse subsets of the network. All simulations, calculations, and visualisations were about to be done by utilising the Gephi tool. The Vienna Subway data was taken from the Kaggle. The dataset consists of four attributes that provide info about the stations and their connections.

Introduction

A train operation plan is the one of the most important aspects of metro systems and directly affects the efficiency and service level of transportation organizations. In metro systems , passengers pay particular attention to the travel time reliability (TTR), which measures the reliability of rail services [1]. In order to achieve more viability and effectiveness in metro stations, the network from one place to other places needs to be analysed to identify the central nodes (or places) and the most congested places in the network. In a number of studies, complex network analysis has been used to measure the importance of stations in transit networks [2]. In particular, centrality measurements, network diameter and connectivity measures are used to analyze different transport networks such as buses, metros [3,4], rails [5,6] and airways. Transit-oriented development (TOD) is for the most part caught on as a successful urban plan for empowering the utilisation of public transportation [8].

Centrality measures are essential to understanding networks, or graphs, such as degrees, betweenness, closeness, page rank and eigenvalue centrality, which score the structural importance of stations in the network. Centrality measurements are characterised for the evaluation of the network of the organize. For the purpose of authenticity, limits for the most extreme reasonable distance to be travelled by train are presented within the definition of the centrality [7]. Network analysis can be utilised to analyse the structure of connections in social systems, the structure or process of altering common wonders, or indeed natural frameworks. To survey the network's connectivity, the network pointers more continually utilised are clustering coefficient and average path length.

Temporal network models in some cases make the solid presumption that interactions are immediate. For numerous sorts of spatially inserted framework, this presumption overlooks the impact that space has in obliging the structure of the network [9]. The network analysis may be valuable for redistributing techniques of the overcrowded stations and moving forward the organization of keeping up the system [10]. The distinguishing proof of centrality advantage stations and regions will offer assistance to create choices based on the TOD improvement points of interest, possibilities, and differences of distinctive locales, and give a logical premise for the formation of urban natural renewal strategies within the inventory optimization size [11].

This study will take the Vienna subway network as a case study. Based on the centrality and network examination of the complex network, it'll make a difference to recognize the centermost hub (or place), most congested places, and places that are not at all active within Vienna. The distinguishing proof of the centrality advantage of stations and locales will offer assistance to creating choices based on the travel-oriented development (TOD) improvement preferences, possibilities, and contrasts of diverse destinations, and provide a logical premise for the arrangement of urban natural re-establishment procedures within the stock optimization arrangement. This investigation makes a difference to contribute towards moving forward the security and capacity of the rail network, making rail travel more appealing for travellers, moving forward industry practice, and informing policy improvement. By distinguishing the foremost congested places, the examination makes a difference to see into arrangements to supply way better administrations for the individuals and it is utilised to redirect the trains from one place to another based on the results of centrality measures.

Literature Review

Ruihua Xu, Fangsheng Wang, Feng Zhou 2021, [1], This paper focuses on developing an efficient train operation plan that is based on journey time reliability. To do so, a data-driven automated fare collection (AFC) system is set up to determine station travel time reliability and analyse the train operation plan over time. They also used a clustering approach based on the SOM neural network. To demonstrate all of this, they took the Beijing metro network and optimised the existing train operation plan.

Bernal, Elisa & Rey, Ángel & GALINDO VILLARDÓN,2020,[2] The authors used the 243-station Madrid metro system. Raw data, HJ-Biplot, Cluster analysis, and Anova are the four steps in their process flow. They took a Madrid network with seven variables and utilised HJ-Biplot to perform multivariate analysis between the various centrality indices. Then, to classify metro stations into homogeneous groupings, a cluster analysis was used. Finally, the famous ANOVA approach was used to examine the relationship between the groups.

Kuhan Jeyapragasan, Gita Krishna & Yash Maniyar,2019,[3], They used cities such as Beijing, Vienna, London, and Delhi in this work. For each of these cities, they used GraphRNN to create a graph and then used Multivariate Bernoulli to construct the graph-level RNN as a Gated Recurrent Unit. After that, they used the Louvain algorithm for community detection, which takes advantage of O's complexity (nlogn). Finally, they calculated basic statistics such as the number of stations with betweenness centrality and so on.

Chen, S.; Zhuang, D,2020,[4], This study uses Guangzhou's historical data, or the metro network's year-by-year network. They did this by using multiple Centrality measurements for each year and determining which areas' centrality was increasing. They discovered that from Kecun station to Jiahe wanggang station, both betweenness and degree centrality were moving northward. And jiahe wanggang station was the network's most popular node. From 1999 to 2018, a large number of network topology statistics were used for each year.

L. Calzada-Infante, B. Adenso-Díaz, S. García Carbajal,2020,[5], They took the European railway transport network and used CNA to construct two topological characterisations of the European international railway network. The first examines direct connections between cities using commonly used measures, and the results are then compared to the Chinese railway network. The second analysis discusses how synchronised timetables promote connectivity by allowing passengers to transfer between routes. For assessing the network's connectedness, centrality measures are defined.

Singh, H.P., Arora, S., Jain, A., Arora, N., Singh, A. and Pal, R.,2022,[6], This study looks at India's railway network, which has 8000 stations and is the fourth largest railway network in the world. As of 2014, the authors split the 926 significant stations into 16 zones to analyse this largest network. They generated centrality measures for each zone, then grouped the 16 zones into four categories (Central railway, South central railway, Northern railway, and Western railway) based on the greatest average degree and estimated P(k) for each of them. They also created the necessary hypotheses for each of the centrality measurements.

In Harold Soh, Sonja Lim, Tianyou Zhang 2010, [7], their study covers analysing the travel routes of the rail (RTS) and other public transportation systems in Singapore from a complex weighted systems viewpoint. They induced that the dynamical properties of a network may contrast essentially from its topological properties. In specific, the RTS network is topologically uninteresting; it is nearly completely associated, and hence shows tall clustering, with a little range and nearly unbiased assortativity. They watched that the activity streams on the RTS network take after power-law dissemination, i.eThe network has centre hubs that are exceptionally high.

Dou, M.; Wang, Y.; Dong, S 2021 [8] their study covers the examination of the transportation framework in Shanghai. They utilised the node to place the show and amplified it into the node place network to demonstrate which can be utilised to survey the degree to which the participation between the arrival and open transport as well as travel request. They gotten three key discoveries those are: TOD rules have as of now been executed in Shanghai, in spite of the fact that there's too an expansive jumble between the travel requests, and moment, a comparative investigation of the disparities between two classifications, the stations that require conceivable changes are distinguished. the encompassing environment of stations. Third, through the list of the carrying weight, the coordinating degree of travel characteristics and natural improvement can be quantitatively assessed, which gives a rule for the improvement of multi-center cities.

Williams Matthew J. and Musolesi Mirco 2016 [9], they proposed a model of Spatio-temporal paths in time-varying spatially embedded networks which captures the property that, as in many real-world systems, the interaction between nodes is non-instantaneous and governed by the space in which they are embedded. The authors presented centrality measures that assess the significance of a node as a basic bridge and its part in supporting spatiotemporally effective streams through the arrangement.

Jafari, M., & Fakhar, S.M. (2018) [10], authors analysed the Tehran Urban and Suburban Railway system by using centrality measures and they discovered the central stations in the network.

Shen, P. and Wang, X., 2021 [11], their study chooses all stations of the Hangzhou Metro to inquire about cases. Based on the centrality of the complex network, the study found that metro stations at each arrangement can be classified into four sorts of centrality focal points, and there are critical contrasts between these sorts. They inferred that a few stations play imperative network roles in totally different arranging stages.

Proposed Methodology

Data Extraction

For this project, our source of information includes the network route map and list of junctions provided by vienna public transport. This involves identifying each junction or station in a network. More information about the traffic of passengers using the subway 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. Each station in Vienna subway transport is a node in our network. And the edge simply represents a train travelling from the source to the destination of the stations.

Analysis and Measures

Since we have the data with us, Through this, we can get 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.

Moreover, we are going to be calculating a few other arranged properties as well to go along with the centrality measures. At that point we'll form some conclusions based on the investigation done so distant and attempt to supply a few particular subtle elements on the zones of advancement. There are parts of conclusions able to draw from our representation. For instance,based on the degree centrality of the arrangement a comment on the structure of the organise and the topology of the organise can be made. Essentially, on the off chance that we consider the Distance across and the Average path length ,the relative comparison between these areas can offer assistance to draw conclusions on the spread of organisation. Moreover, in the case we see at the Clustering Coefficient, a lower clustering coefficient implies a weaker connected organize. Measured quality partitions are arranged into modules. As a result, the division which gets the greatest seclusion gives us the more imperative community.

Results and Discussion

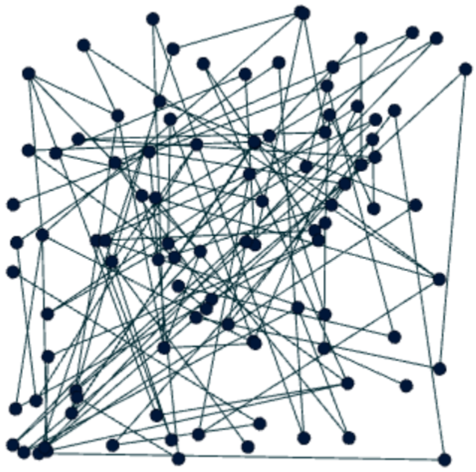

Figure 1: Graph Representation of Vienna Subway Network **Figure 2:** Graph Representation of Vienna Subway Network (Yifan hu Layout)

The Yifan Hu layout algorithm belongs to the category of force-directed algorithms, which includes the Force Atlas and Fruchterman Reingold algorithms.

Degree:

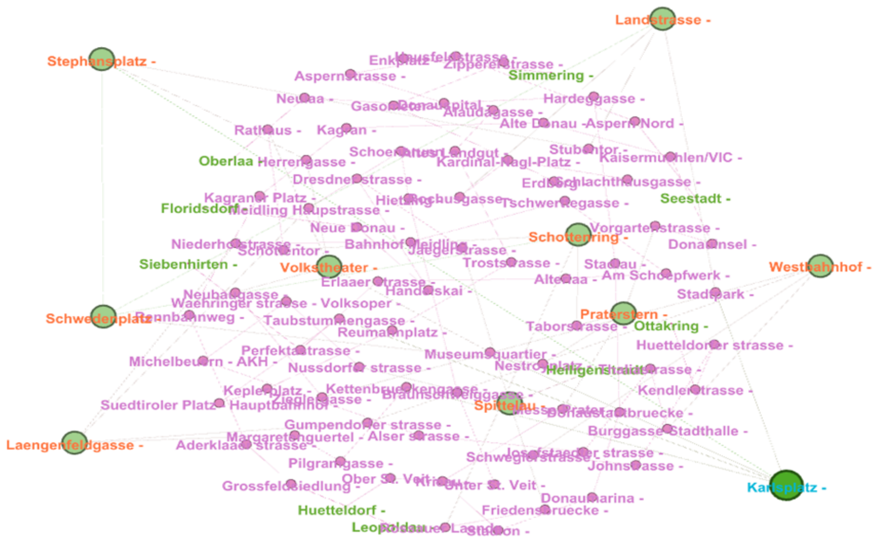

Figure 3: Vienna subway network based on degree of the station

Degree centrality appears in how numerous associations an individual has. They may be associated with lots of individuals at the heart of the network, but they might be distant off on the edge of the organisation. In our work, we can say that the nodes which are having the highest degree can be considered as important stations of Vienna. The average degree is 2.122. It infers that most of the stations have been connected to two other stations and by the average degree esteem we can say that there are very few junctions in the Vienna subway network.

Station	Degree
Karlsplatz	5
Schwedenplatz	4
Praterstern	4
Stephansplatz	4
Laengenfeldgasse	4
Schottenring	4
Westbahnhof	4
Spittelau	4
Landstrasse	4
Volkstheater	4

Table - 1: Degree of the top - 10 stations

From the table - 1 and figure - 3 we can say that Karlsplatz station has the highest degree among all the stations and it was considered as the junction of the vienna subway network. Because more stations are connected to Karlsplatz. So, the committee or the organisation can decide on whether they can improve the station or construct a small substation to partition the traffic.

Betweenness Centrality:

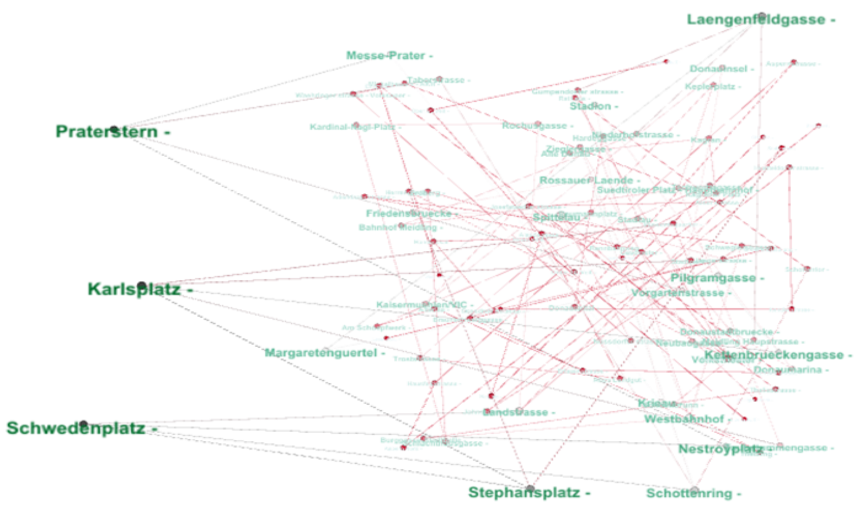

Figure 4: Vienna subway network based on Betweenness centrality of the station

Betweenness centrality could be a way of recognizing the sum of the impact a hub has over the flow of data in a graph. It is frequently used to discover hubs that serve as a bridge from one portion of a graph to another. It measures the number of times a hub acts as a bridge along the most limited way between two other hubs. A hub with higher betweenness centrality would have more control over the network since more data will pass through that hub.

Station	Betweenness Centrality
Karlsplatz	0.388406
Schwedenplatz	0.380799
Praterstern	0.380369
Stephansplatz	0.307775
Laengenfeldgasse	0.291273
Nestroyplatz	0.247816
Schottenring	0.24089
Kettenbrueckengasse	0.226518
Pilgramgrasse	0.218571
Margaretenguertel	0.211054

Table - 2: Betweeness centrality of the top - 10 stations

Above are the top 10 places according to the Vienna subway network. And we can see that Karlsplatz has more Betweenness centrality. And we can say that more shortest paths move from the Karlsplatz station. From table-1 and table-2 we can see that Stephansplatz and Praterstern have the same degree but different Betweenness centrality.

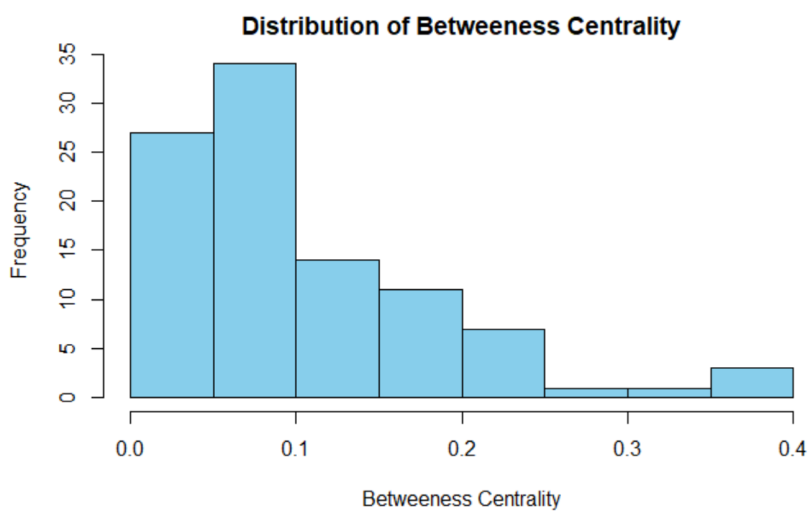

Figure 5: Distribution of Betweeness centrality

Figure - 5 shows the distribution of between centrality of every station. Where, we can see that betweeness centrality was more dense between 0 to 0.1 and was decreasing at a higher rate as the interval increases. One conclusion we may draw is that if the rate of change of distribution is linear, the relevance of each node will be nearly the same, reducing the network's overall susceptibility.

Closeness Centrality:

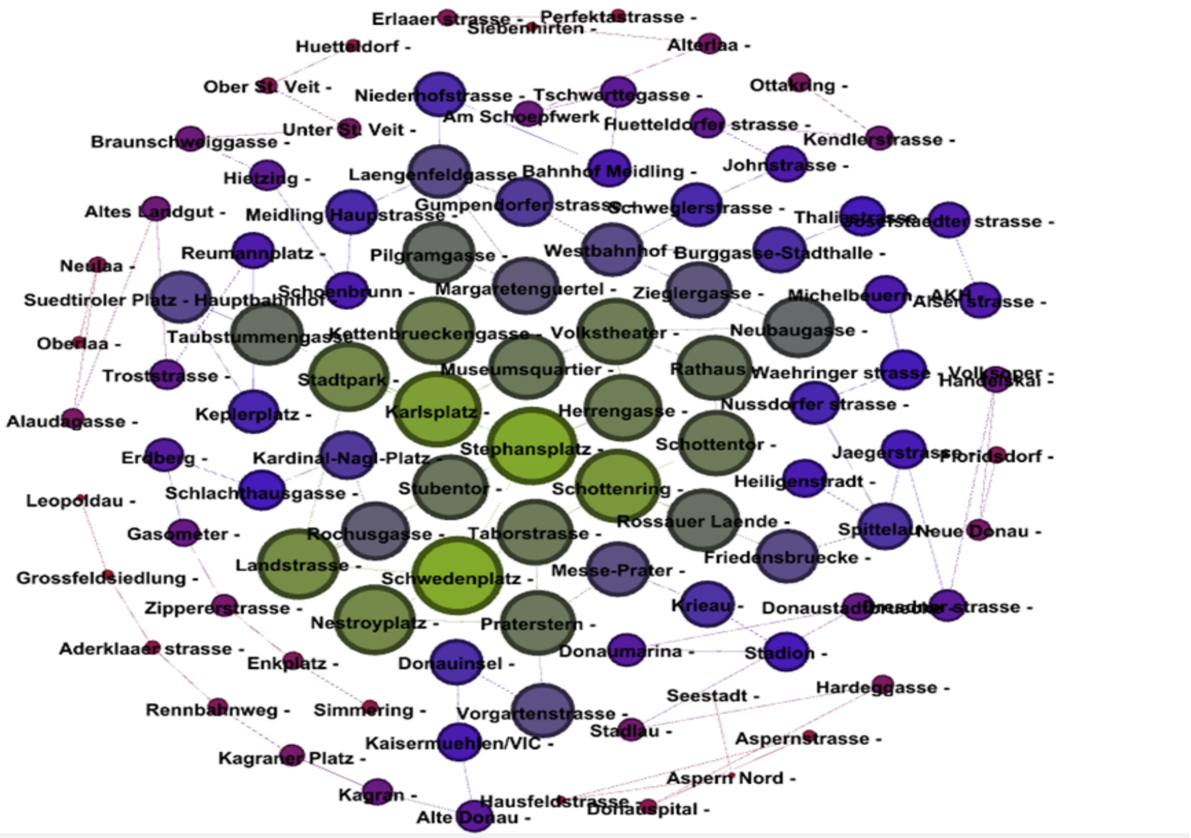

Figure 6: Vienna subway network based on Closeness centrality of the station

Closeness centrality could be a way of identifying nodes that are able to spread data exceptionally effectively through a network. The closeness centrality of a node measures its normal farness to all other hubs. Hubs with a high closeness score have the most limited separations from all other hubs. A node with a high closeness centrality would mean it has close bonding with other hubs.

Station	Closeness Centrality
Schwedenplatz	0.149461
Stephansplatz	0.148545
Karlsplatz	0.146084

Schottenring	0.143491
Landstrasse	0.139769
Nestroyplatz	0.138968
Stadtpark	0.138374
Kettenbrueckengasse	0.136236
Herrengasse	0.135286
Volkstheater	0.134722

Table - 3: Closeness centrality of the top - 10 stations

From table-3 and figure -6 we can see that Schwedenplatz has the highest closeness centrality. This means we can say that it has the closest bond with the other stations. And can improve the ways from there instead of Karlsplatz which has the highest degree. So in this way they can reduce the congestion.

Eigenvector Centrality:

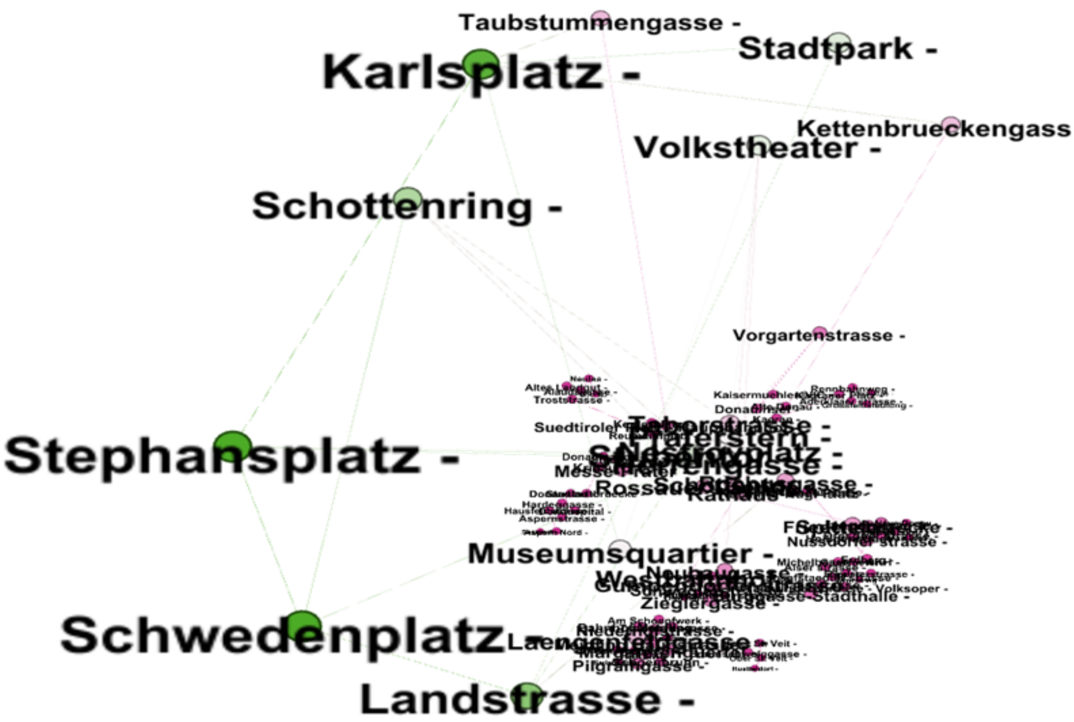

Figure 7: Vienna subway network based on Eigenvector centrality of the station

As each node is given a score or value, eigenvector centrality can be used to determine the level of influence that node has over the network. The greater the score, the more influential that node is. An eigenvector of the network adjacency matrix is used to assess how important nodes are in a network based on their connections. The score is based on the number of connections that a node will have to other nodes.

Station	Eigenvector Centrality
Schwedenplatz	1
Stephansplatz	0.998774
Karlsplatz	0.955384
Landstrasse	0.820444
Schottenring	0.71673
Stubentor	0.58848
Stadtpark	0.581754
Volkstheater	0.56719
Herrengasse	0.518479
Museumsquartier	0.511753

Table - 4: Eigenvector centrality of the top - 10 stations

From the table - 4 and Figure - 7 we can say that Schwedenplatz has the highest eigenvector centrality. Hence, Schwedenplatz has the most influence on the network. So, the committee or the organisation should take care of the crowd of that station because many people will come to that station because of its influence.

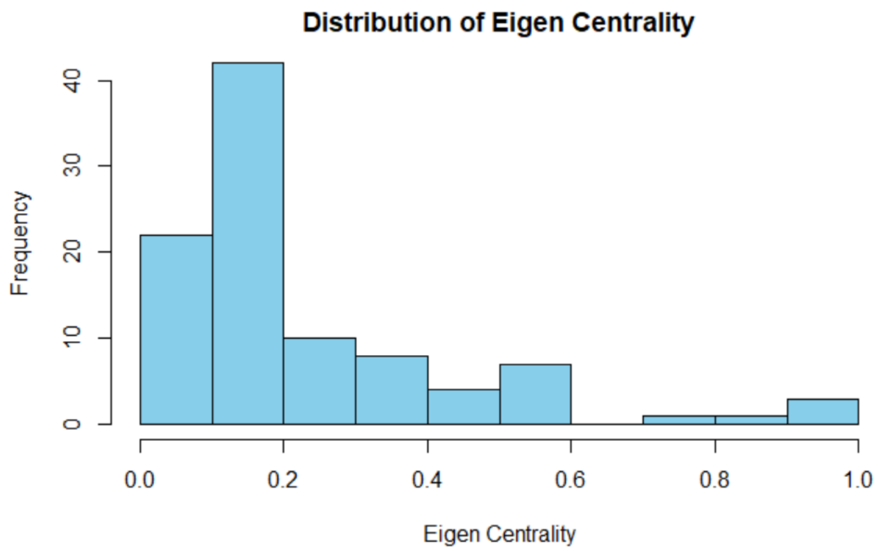

Figure 8: Distribution of Eigen Centrality

Figure - 8 shows the distribution of eigen centrality of every station in vienna subway network. We can observe that there was no much difference between Figure - 5 and Figure - 8 which shows the distribution of Betweeness and Eigen centrality respectively. Both the centrality measure are dense between 0 to 0.1. Therefore, we can conclude that there are only few influential nodes among the network. So, the best node will be more congested compared to other nodes.

Eccentricity:

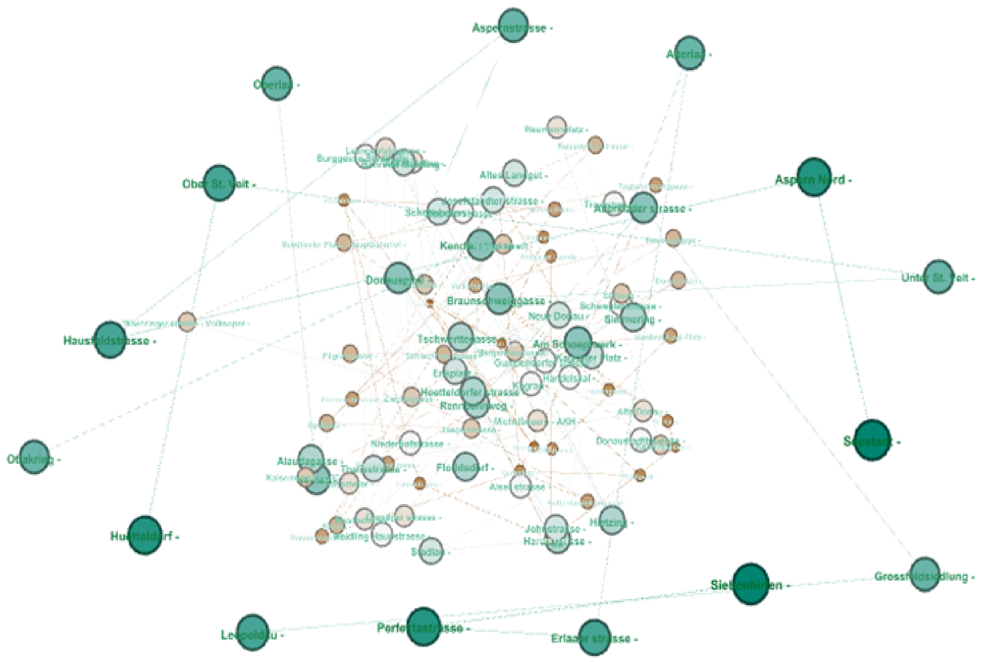

Figure 9: Vienna subway network based on Eccentricity of the station

A graph's eccentricity is defined as the maximum distance between the vertex and its neighbours. The maximum distance between a vertex and its neighbours is considered the eccentricity of a vertex. The reciprocal of eccentricity is used to measure a node's importance in social networks.

Station	Eccentricity
Siebenhirten	28
Seestadt	28
Perfektastrasse	27
Aspern Nord	27
Huetteldorf	27

Erlaaer strasse	26
Hausfeldstrasse	26
Ober St. Veit	26
Leopoldau	26
Alterlaa	25

Table - 5: Eccentricity of the top - 10 stations

From table - 5 and figure - 7 we can say that Siebenhirten has the highest Eccentricity. Which means it is the farthest stop or the end point of any line. This station was considered as the less popular station and there is a scope to increase in the future. Based on the revenue from this stop the organisation can decide whether to keep this station or remove it. Because, so many disadvantages like current cost, maintenance etc. For low revenue stations. Instead they can improve the station for the high centrality areas.

Page Rank:

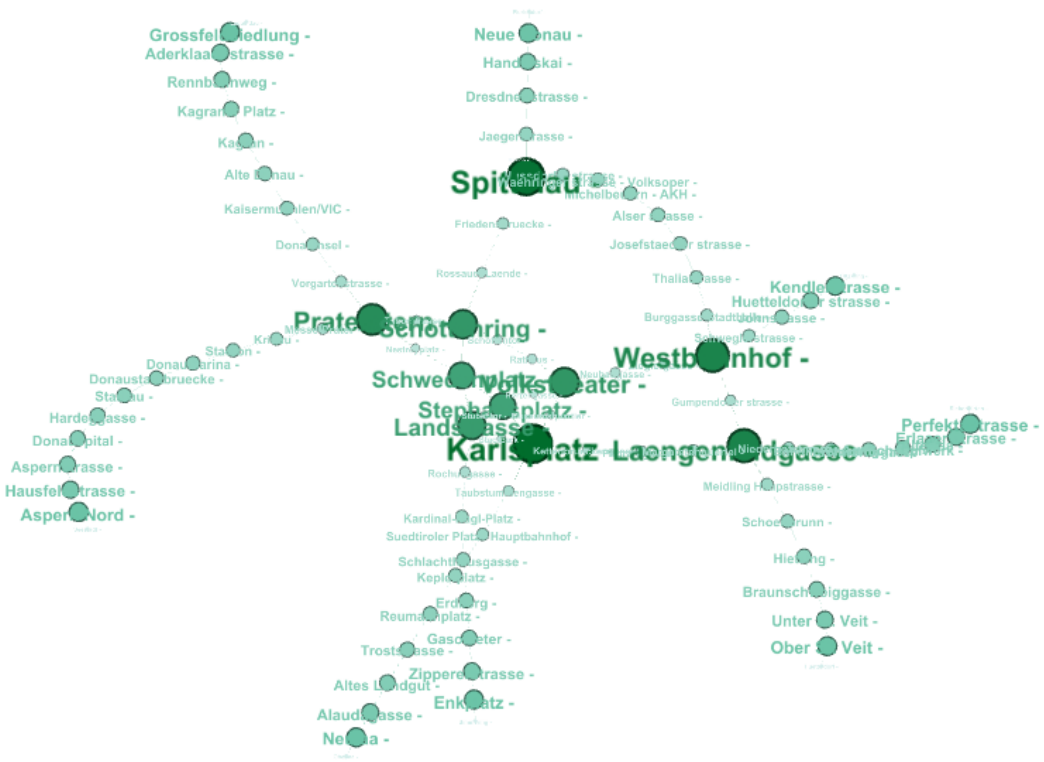

Figure 10: Vienna subway network based on Page Rank of the station

The PageRank calculation measures the significance of each node inside the graph, based on the number of approaching connections and the significance of the comparing source nodes. It is accepted in a few research papers that the distribution is equally partitioned among all reports within the collection at the start of the computational process. The PageRank computations require a few passes, called "iterations", through the collection to alter inexact PageRank values to more closely reflect the hypothetical genuine esteem.

Station	Page Rank
Karlsplatz	0.018633
Spittelau	0.018183
Laengenfeldgasse	0.016898
Westbahnhof	0.016839
Praterstern	0.01607
Volkstheater	0.015439
Schottenring	0.015286
Landstrasse	0.014949
Schwedenplatz	0.014421
Stephansplatz	0.014412

Table - 6: Page Rank of the top - 10 stations

Modularity:

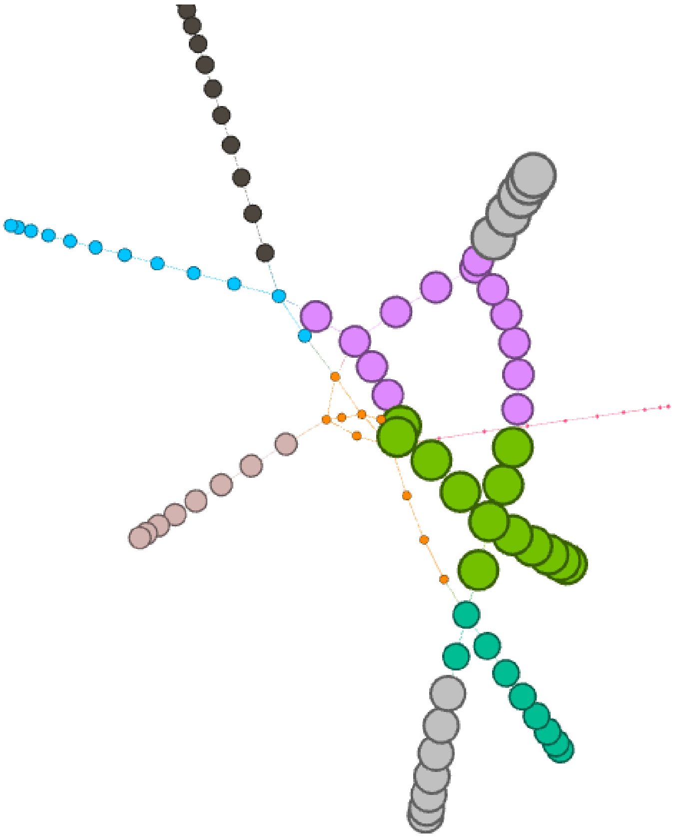

Figure 11: Vienna subway network based on the Communities

Modularity is a measure of a network's structure. It was used to detect the communities in the network. The concept of community detection has risen in network science as a strategy for finding bunches inside complex frameworks represented by a graph. It was outlined to measure the quality of the division of a network into modules. Networks with high measured quality have thick associations between the hubs inside modules but inadequate associations between hubs with totally different modules. We are able to identify 10 communities in the Vienna subway network.

Stations	Community
Oberlaa, Neulaa, Alaudagasse, Altes Landgut, Troststrasse, Reumannplatz, Keplerplatz, Suedtiroler Platz - Hauptbahnhof, Taubstummengasse	0
Karlsplatz, Stephansplatz, Schwedenplatz, Herrengasse, Stubentor, Landstrasse, Margarentenguertel, Piligramgasse, Kettenbrueckengasse, Stadpark	1
Nestroyplatz, Praterstern, Vorgartenstrasse, Donauinsel, kaisermuehlen/VIC, Alte Donau, Kagran, Kagrsner Platz, Rennbahnweg, Aderklaaer strasse, Grossfeldsiedlung, Leopoldau	2
Seestadt, Aspern Nord, Hausfeldstrasse, Aspernstrasse, Donauspita;, Hardeggasse, Stadlau, Donaustadtbruecke, Donaumarina, stadion, Krieau, Messe-Prater	3
Rochusgasse, Kardinal-Nagl-Platz, Schlachthausgasse, Erdberg, Gasometer, Zipperersstrasse, Enkplatz, Simmering	4
Huetteldorf, Ober St. Veit, Unter St. Veit, Braunschweiggasse, Hietzing, Schoenbrunn, Meidling Haupstrasse, Laengenfeldgrasse, Niederhofstrasse	5
Taborstrasse, Schottenring, Schottentor, Rathaus, Rossauer Laende, Friedensbruecke, Spittelau, Heiligenstradt, Josefstaedter strasse, Alser strasse, Michelbeuern - AKH, Waehringer strasse - Volksoper, Nussdorfer strasse	6
Siebenhirten, Perfektastrasse, Erlaaer strasse, Alterlaa, Am schoepfwerk,	7

Tschwerttegasse, Bahnhof Meidling	
Volkstheater, Museumsquarter, Ottakring, Kendlerstrasse, Huetteldorfer strasse, Johnstrasse, Schweglerstrasse, Westbahnhof, zieglergasse, Neubaugasse, Gumpendorfer strasse, Burggasse-Stadthalle, Thaliastrasse	8
Jaegerstrasse, Dresdner strasse, Handelskai, Neue Donau, Floridsdorf	9

Table - 7: Stations in each community

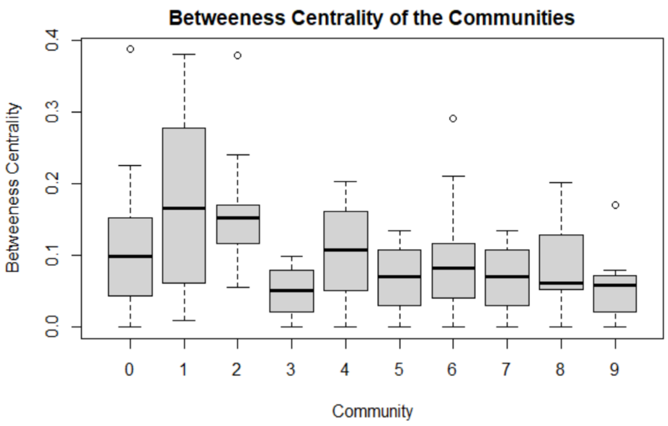

Figure 12: Boxplot of Betweeness centrality of every community

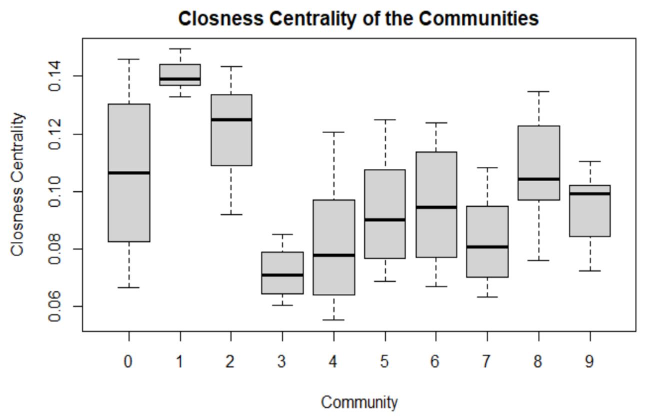

Figure 13: Boxplot of Closness centrality of every community

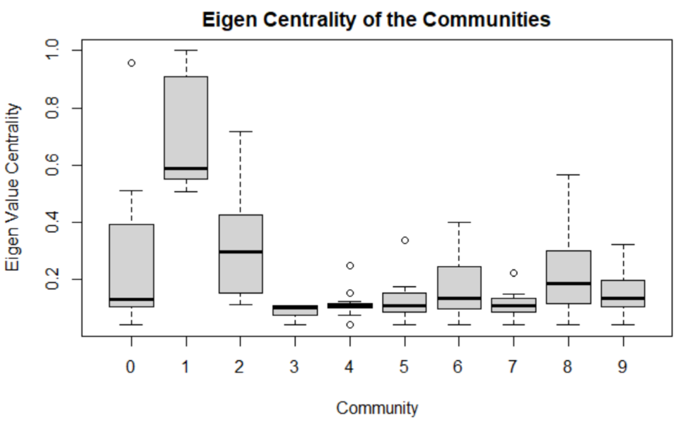

Figure 14: Boxplot of Eigen centrality of every community

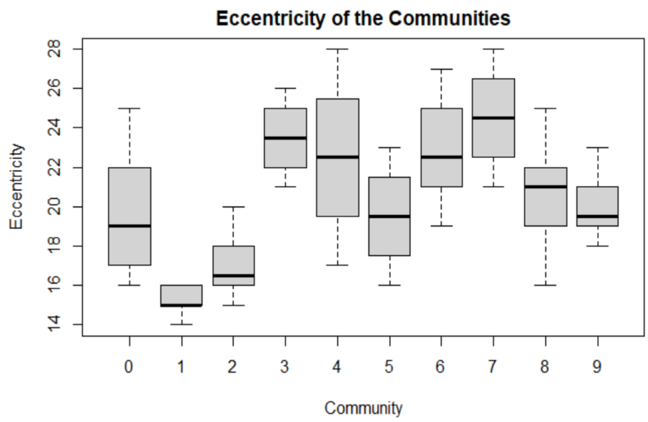

Figure 15: Boxplot of Eccentricity of every community

Figure - 12,13,14,15 shows the boxplots of the betweenness centrality, closeness centrality, Eigen centrality and Eccentricity of every community respectively. We can observe that Community - 1 has the highest Centrality measures. Whereas, community - 4 has the highest Eccentricity. Therefore, Community - 1 is the precious community in the identified communities.

Conclusion:

Network analysis gives the capacity to gauge complex designs of connections and the network structure can be dissected to uncover the centre highlights of the network. Our work consists of analysing the Vienna subway network and to interpret useful references. The network consists of 98 nodes (stations) and 104 edges (links). We identified the most congested places and the central station of Vienna. Among all stations there are a few stations like karlsplatz, Schwedenplatz, Stephansplatz. These stations can be considered as most stations. Karlsplatz, Splittelau, Laengenfeldgasse, Westbahnhof, Praterstern are considered as big central rail zones in the Vienna subway network.

References

[1]Ruihua Xu, Fangsheng Wang, Feng Zhou, "Metro Train Operation Plan Analysis Based on Station Travel Time Reliability", Journal of Advanced Transportation, vol. 2021, Article ID 8813461, 13 pages, 2021. https://doi.org/10.1155/2021/8813461

[2] Bernal, Elisa & Rey, Ángel & GALINDO VILLARDÓN, Mª Purificación. (2020). Analysis of Madrid Metro Network: From Structural to HJ-Biplot Perspective. Applied Sciences. 10. 10.3390/app10165689

[3]Kuhan Jeyapragasan, Gita Krishna & Yash Maniyar, (2019). An Analysis of Subway Networks using Graph Theory and Graph Generation with GraphRNN. Stanford publications 12 (2019).

[4]Chen, S.; Zhuang, D. Evolution and Evaluation of the Guangzhou Metro Network Topology Based on an Integration of Complex Network Analysis and GIS. Sustainability 2020, 12, 538. https://doi.org/10.3390/su12020538

[5]L. Calzada-Infante, B. Adenso-Díaz, S. García Carbajal, Analysis of the European international railway network and passenger transfers, Chaos, Solitons & Fractals, Volume 141,2020,110357,ISSN 0960-0779,https://doi.org/10.1016/j.chaos.2020.110357

[6]Singh, H.P., Arora, S., Jain, A., Arora, N., Singh, A. and Pal, R., SYSTEMATIC STUDY OF INDIAN RAILWAYS SUBNETWORK: ZONE SPECIFIC ANALYSIS

[7]Harold Soh, Sonja Lim, Tianyou Zhang, Xiuju Fu, Gary Kee Khoon Lee, Terence Gih Guang Hung, Pan Di, Silvester Prakasam, Limsoon Wong, Weighted complex network analysis of travel routes on the Singapore public transportation system,Physica A: Statistical

Mechanics and its Applications, Volume 389, Issue 24,2010,Pages 5852-5863,ISSN 0378-4371,https://doi.org/10.1016/j.physa.2010.08.015

[8]Dou, M.; Wang, Y.; Dong, S. Integrating Network Centrality and Node-Place Model to Evaluate and Classify Station Areas in Shanghai. ISPRS Int. J. Geo-Inf. 2021, 10, 414. https://doi.org/10.3390/ijgi10060414

[9]Williams Matthew J. and Musolesi Mirco 2016Spatio-temporal networks: reachability, centrality and robustnessR. Soc. open sci.3160196160196 http://doi.org/10.1098/rsos.160196

[10]Jafari, M., & Fakhar, S.M. (2018). Network Centrality Analysis of Tehran Urban and Suburban Railway System. ArXiv, abs/1802.06219

[11]Wu, F., Zhu, Y., Zhu, W., Wang, K., Shen, P. and Wang, X., 2021, February. The spatial distribution characteristics of the centrality advantage stations in Hangzhou metro planning (2005-2022). In IOP Conference Series: Earth and Environmental Science (Vol. 634, No. 1, p. 012110)