Quantitative and Network Analysis of the Washington DC Metro System

Varun Jani
Virginia Tech
Alexandria, Virginia, United States
jvarun@vt.edu

Richard James Heiman
Virginia Tech
Alexandria, Virginia, United States
ricjames@vt.edu

Kevin Lin
Virginia Tech
Alexandria, Virginia, United States
kevinlin04@vt.edu

Benjamin Sullivan
Virginia Tech
Alexandria, Virginia, United States
bens21@vt.edu

Abstract

The DC WMATA Metro System, comprising both heavy-rail and bus, is an important part of the average DC metropolitan citizen's life. It serves as a key transportation method, providing a cheap and sustainable alternative to commuting without a personal vehicle. However, like any publicly-funded institution (and particularly public transportation), it faces a lean budget. Decision and policy makers in WMATA are under intense scrutiny to make efficient choices to better serve the metropolitan area. This analysis aims to assist in these choices by providing key insights into ridership data via using statistical analysis, machine learning and graph theory. Specific objectives include determining optimal bus routes, analyzing bus route efficiency, and building a network topology of the rail network. The analysis provides conclusions about the validity of uneven travel, effectiveness of adding bus routes, and provides insight into key stations of the rail network.

Keywords

Public Transport, Network Science, Analysis, Modeling, Prediction, Traffic, Passenger Volume

ACM Reference Format:

Varun Jani, Richard James Heiman, Kevin Lin, and Benjamin Sullivan. 2025. Quantitative and Network Analysis of the Washington DC Metro System. In *Proceedings of Make sure to enter the correct conference title from your rights confirmation email (Conference acronym 'XX)*. ACM, New York, NY, USA, 8 pages. https://doi.org/XXXXXXX.XXXXXXX

1 Introduction

The Washington Metropolitan Area Transit Authority (WMATA), or DC Metro, serves as a crucial system for the nation's capital, yet it suffers from significant inefficiencies that hinders its potential as the primary mode of transportation within the District of Columbia, Maryland and Virginia (DMV) area. Despite being a critical infrastructure for daily commutes and social connectivity, the network suffers from uneven station demand and suboptimal routing which often makes private automobiles a more attractive alternative. These operational deficits result in lower ridership and subsequent funding gaps, creating a sort of feedback loop that stagnates overall system improvement. This project aims to identify and address these challenges by analyzing the metro network's current state to find opportunities to enhance efficiency, accessibility and equity across the DC Metro system and the area in general.

Figure 1: Map of DC Metro Rail System [1]

To provide actionable solutions, we use an analytical approach that leverages data provided to us by WMATA, which includes

Conference acronym 'XX, Washington D.C.

ACM ISBN 978-1-4503-XXXX-X/2018/06
https://doi.org/XXXXXXX.XXXXXXX

ridership and geographic data. We focus on diagnosing bus route efficiency to visualize geospatial demand "hotzones", mapping temporal rail traffic patterns to pinpoint peak usage windows. Additionally, we look into developing a synthetic Origin-Destination (OD) matrix that models complex passenger flows across the network. Furthermore, by building a graph-based network topology of the physical rail system, we were able to isolate critical hubs and identify structural vulnerabilities. Through this combination of exploratory data analysis, network science, and predictive modeling, we have identified and determined recommendations for improvement. These recommendations range from optimizing bus-to-rail conversion across the network to relieve specific bottlenecks as well as enable data-driven decision making for improving urban mobility across the network.

The first approach was to optimize the bus routes using a regression approach. The cost function to be minimized is the sum of the squares of the travel times between any two stations in the network. This includes stations on different routes for which a bus transfer is required. The travel time is defined as the time between when a person leaves one station and arrives at the other. Therefore, the travel time accounts for time spent waiting at a bus stop when transferring from one bus to another. The optimal routes will be compared to the original six bus routes. Then, the entire process is repeated with additional buses, 7, 8, 9, and 10 buses in total and the cost function and average travel times are compared to determine the extent to which adding new buses increases overall efficiency. The results include the optimal routes with this approach, the cost function, and the average travel time between any two stops in the network.

2 Related Research

As cities struggle with congestion, environmental degradation and spatial inequality, rail system are increasingly being viewed as the groundwork for sustainable urban development. This is especially true for the DMV area and its increasing population over the past 10 years [6], leading to more and more congestion on the road. Related literature and research indicates that simply laying more tracks and expanding in that manner is insufficient. It is important to identify where to build new rail lines. As found in the "Baltimore Study" [7], adding new rail lines along high-density corridors can definitely ensure maximum efficiency and utility as it results in frequency services, reduced wait times and make rail travel an attractive option for commuters. However, a critical step is the placement of new stations. The authors found that stations have to be strategically located at key points to make public transport convenient and a strong alternative to driving. These key points include locations where either a large number of people start or end their journeys, such as residential areas, business districts, shopping centers etc. Some other interesting ideas in this paper include the introduction of measures such as congestion pricing, restricted parking and pedestrian-only zones, to encourage more people to switch to public transport.

Looking at more specific research into WMATA, Troung et al. [9] introduces an interesting approach that combines Principal Component Analysis (PCA) and hierarchical clustering to model the passenger flow. These techniques help identify latent features within the DC metro network which allows them to show the time-series of passenger inflow and outflow. This classifies the stations based on temporal usage patterns rather than just as static locations. Their approach to predict passenger volume at stations was highly effective and outperforms similar efforts attempted in the past. One caveat to their approach is that every single data piece was completely anonymous, resulting in no proper way of identifying whether its the same person entering and exiting at two different stations. In summary, their approach allowed for high-accuracy prediction of passenger volumes which served as the foundation for community analysis and flow optimization within the DC Metro System.

A paper from Saadat et al. [10] in 2020 addresses the problem of network structure and bottleneck analysis by modeling the DC metro system as a graph and using that graph for network topology analysis. This graph, containing 91 stations and a 140 links, provides analytical insight into the system, as it shows signs of a scale-free phenomena, a topic we covered in class. Their evaluation focuses on two primary failure events, the failure of a metro station and the failure of a metro segment between stations. Specifically, their methodology utilizes topological metrics to pinpoint critical stations and segments that would be potential bottlenecks during said failure events, thus offering insight into these bottlenecks and allowing policymakers to best mitigate them before they become an even bigger problem.

Another avenue explored in literature to determine uneven demand and route inefficiencies has been through applied graph theory to model metro systems as a complex network of nodes and edges where nodes are denoted as stations and edges as tracks. Stoilova and Stoev (2015) [11] explore this idea by defining indicators such as degree of routing, connectivity of the route, average length per link (this takes into account the number of routes), intensity and density of the route. A total of 10 indicators are used for their analysis. They then employ the Statistical Package for Social Science (SPSS) software to carry out their study with a cluster analysis, forming dendograms and for overall classification of the metro network. This allows them to define the state and structure of the network, thus helping them group the metro network for further evaluation. Similarly, other papers have also used such graph based approaches for quantitative assessment of their metro systems. One example is from Barman and Mishra (2024) [3], who use this to study the structure of metro networks in India, specifically in Indian cities, by employing a graph-based theoretical approach and demonstrate that metro networks can be analyzed using vertex connectivity and edge centrality to help identify bottlenecks where any sort of failure would severely impact the flow of the metro.

Finally, the application of Machine Learning (ML) and Artificial Intelligence (AI), especially for urban mobility in smart cities, is rapidly evolving from retrospective analysis to real-time prediction. An example of this was explored in the paper [8] where collaboration between Virginia Tech and WMATA utilized RNNs to predict On-Time Performance (OTP) for current passengers in real-time. Their goal was to communicate delays to passengers before they enter the system in an effort to address unpredictability as that is a major cause of ridership decline. Chen and Zhang (2024) [4] sort of extended on this approach as they considered other factors such as interaction cost between transportation models, energy consumption and environmental impact. They built a modified Teaching-Learning Based Optimization (TLBO) algorithm for enhanced optimization and a hybrid Artificial Neural Network-Recurrent Neural Network (ANN-RNN) for improved system adaptability. Additionally, forecasting has always been an open-ended problem and can be applied to a lot of domains. Models such as the Long Short-Term Memory (LSTM) help develop more accurate forecasting models as these models can account for more variables compared to your traditional statistical or ML models. Zhang et al. (2019) [15] proposes a deep learning architecture that combines residual network (ResNet), graph convolutional network (GCN) and LSTM to forecast short-term passenger flow in urban rail transit. Their hybrid model is called ResLSTM. They include indicators such as inflow, outflow, graph-network topology as well as weather conditions and air quality as part of their model and use 10, 15 and 30 minute time intervals to conduct short-term passenger flow forecasting. Their study is specific to the Beijing subway system containing 17 lines and 276 subway stations. They find that including indicators like weather conditions and air quality have considerable influence on prediction precision, and as per their claim, they are probably the first paper to explore air quality as an indicator. However, they also highlight limitations such as the model being a "black box", which makes model interpretability poor and is a common problem when dealing with ML models like ResNet and LSTMs.

3 Methodology

3.1 Data Collection and Enrichment

We received two sets of ridership data directly from WMATA for both rail and bus. The rail ridership data gives the average boardings and alightings per hour, date, and station. Rail data was provided for the years 2022-2025. The bus data provided is partitioned by stop, time of day, weekday or weekend, and the direction of the route; for each row ridership is provided as average load, max load, the sum of passengers on, the sum of passengers off and the sum load. Rail ridership data is collected by faregate sensors, which are set off when customers interact with them [13]. Bus ridership data is collected similarly, with sensors on each door [12].

Ridership data, for both rail and bus, was further enriched utilizing WMATA's free and publicly available developer API. For rail, using the station info endpoints provided, we were able to programatically gather and add information for station name and location. Similarly, for buses, we downloaded route data and joined with the ridership data to add route names and descriptions.

Utilizing the same API, we created an entirely new data-set of station-station distance info, providing both cardinal distance and railtime distance.

4 Experiments

The first experiment was using linear regression to determine the optimal bus routes, using the sum of the squares of the travel times between any two stations as an objective function. An undirected graph was constructed to represent the connectivity of the stations, with weighted edges that represent times between stations. The graph also incorporated weighted edges to represent time spent waiting at a station for a transfer. We first identified the 10 heaviest

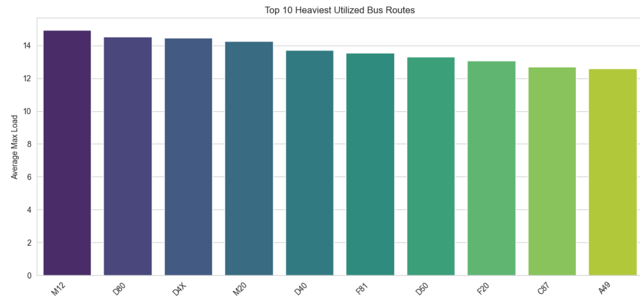

Figure 2: Heaviest Utilized Bus Routes

utilized bus routes, shown in Figure 2 which provides us the average maximum load for those stations.

Our next set of analysis was split into three distinct phases. First, we performed a geospatial analysis of the bus network. We filtered out invalid GPS coordinates, cleaning up the data, and then used Kernel Density Estimation (KDE) to create heatmaps. This allowed us to see where the demand actually clusters physically, which is more convenient to visualize and better than just having a list of route names.

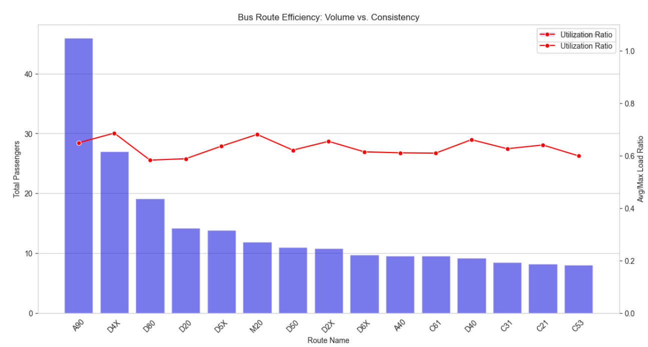

Figure 3: Bus Route Efficiency: Volume vs. Consistency

Second, for the bus efficiency analysis, we calculated a "utilization ratio" by dividing the average load by the maximum load, as shown in figure 3. This provides some insight into "how full is the bus on average compared to its peak" and doesn't just rely on the total number of riders. This is necessary because their totals might be the same, but this ratio allows us to determine how different a route that is always full is compared to a route that is empty half the time and overcrowded during the other half.

Finally, for the rail network, we built a topology graph using the NetworkX library. Due to lack of station data that defines the line sequence, we manually defined physical connections between the stations. The algorithm used was a Gravity Model algorithm. Specific data of passengers is required to estimate the proper flow. Since we didn't have this data, this model estimates the flow between two stations based on how many people enter from lets say station Ballston-MU and leave at station Rosslyn.

4.1 Exploratory Data Analysis

The results of the first experiment indicate that finding the optimal route is an intractable problem, so a heuristic is used. After being prompted multiple times to make stops that contain multiple routes so that people can transfer between buses to reach a station on a different route, the LLM failed to do this and instead provided six disjoint routes so that every station would only be assigned to one route. This also means that the cost function is underestimated, because the cost function only represents pairs of stations on the same route and not in the entire network. Additionally, the average time between stations calculation was also based on only stations in the same route.

As expected, the cost function and travel time decreased significantly with the addition of more buses. However, the travel time is suspiciously low for 9 and 10 buses, showing that the travel time only represents stops on the same route. Taking 3 minutes to travel to any two stations is not realistic, and that was the result for 10 stations. Another reason the numbers could be so low is that the LLM counted traveling from a station to itself, which is adding a zero to the dataset to drive down the average. Figure 4 represents

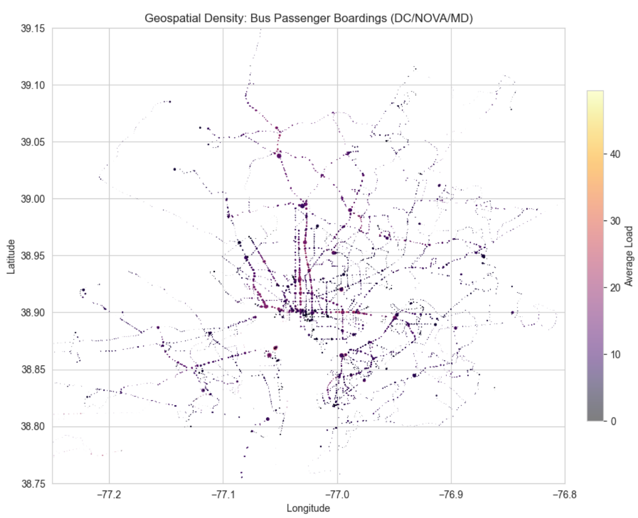

Figure 4: Bus Passenger Boardings Density Map

the geospatial distribution of bus demand across the DMV area metropolitan area. The visualization uses a weighted density approach, where the intensity of the color is directly proportional to the average load. Large, lighter-colored nodes suggest major transit hubs, as there seems to be frequent turnover. On the other hand, large, dark nodes suggests potential capacity bottlenecks where there is high demand but even higher passenger average load. Figure 5 aggregates the total boardings and alightings for every hour and shows us the same two spikes we see in figure 6 below. This verifies the data, since every passenger who enters must eventually exit and as the graph shows, we have an almost identical line for boardings and alightings. This is a great chart of capacity planning in order to improve the the overall DC metro system. Figure 6 visualizes the daily versus hourly pattern of the rail network, mapping average boardings to show distinct usage behaviors. One thing that

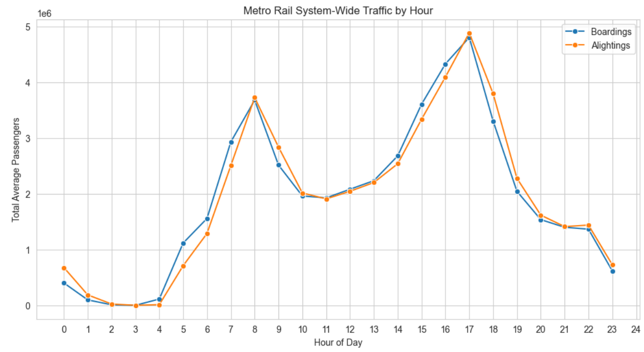

Figure 5: Metro Rail System-Wide Traffic by Hour

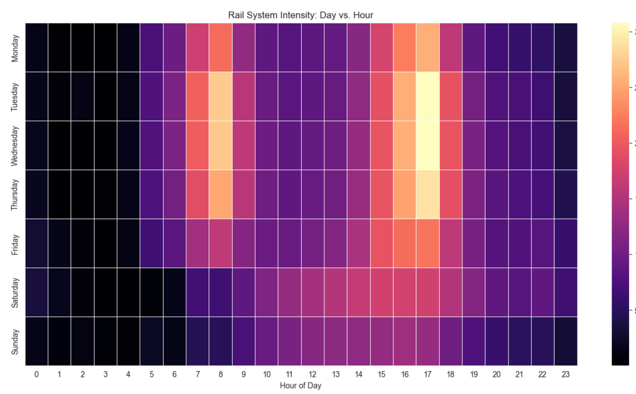

Figure 6: Rail System Intensity Day vs. Hour

immediately sticks out is the daily commuter pattern on weekdays, specifically during the morning (7 AM to 9 AM) and evening (4 PM to 6 PM). This heatmap and flow is vital for operational planning as it helps identify saturation levels and also helps analyze where adjustments are most needed to manage the massive inflow and outflow of passengers.

We then go on to visualize the spatial patterns in ridership, as seen in figure 7 where we generate a weighted density map using Kernel Density Estimation (KDE). The five highest volume stops are shown in the figure to help identify the specific nodes that are driving the most demand within these highly dense clusters. These resulting "hotspots" align with major transit hubs and confirm that population size and demand go hand in hand. Figure 8 visualizes the estimated volume of passenger trips between different station pairs, creating a Origin-Destination (OD) matrix where the axes represent entry and exit points. The station IDs are used for a cleaner visualization. This approach helps model complex network dynamics and shows key commuter flows, and is especially useful given we don't have individual ticket tracking data available for this study. A topological level analysis is done to identify critical structural bottlenecks. Given graph theory concepts, the network connectivity is determined using a graph layout with edges as track segments and nodes as stations. Figure 9 represents this graph, where the size of each node corresponds to its betweenness centrality score and

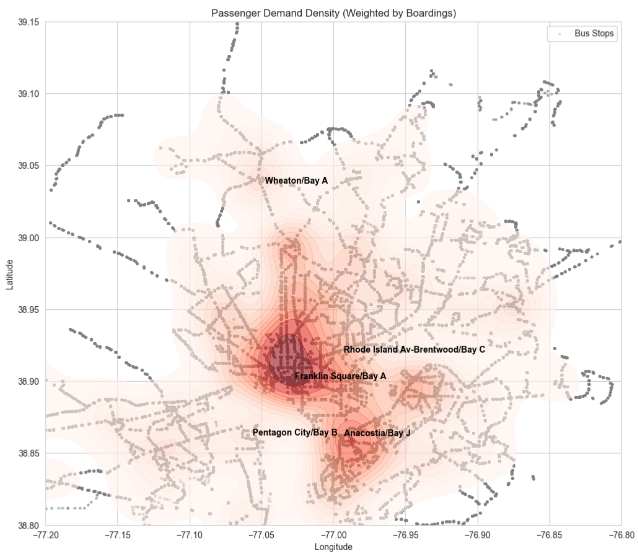

Figure 7: Bus Passenger Demand Density

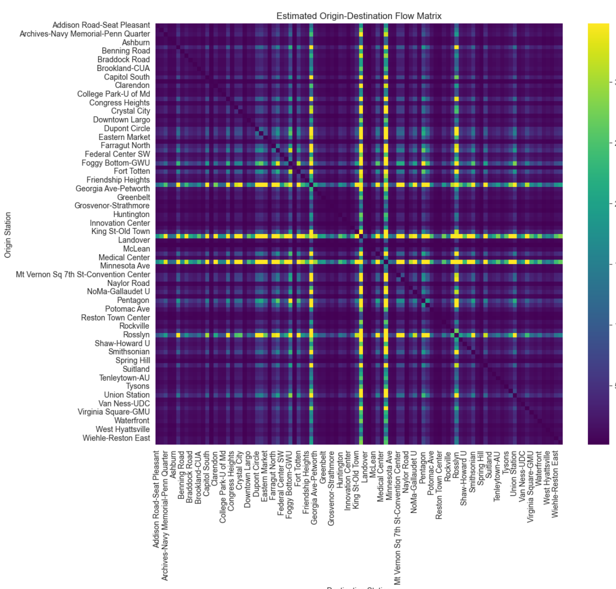

Figure 8: Bus Estimated Flow Matrix

larger nodes indicate stations that lie on a high number of shortest paths, thus making them critical points of failure or congestion.

4.2 Experimental Setup

In order to evaluate the efficiency and connectivity of the DC Metro system, we designed a multi-phase experimental framework which consisted of: data ingestion, bus route optimization, and rail network topology analysis. All experiments were implemented using

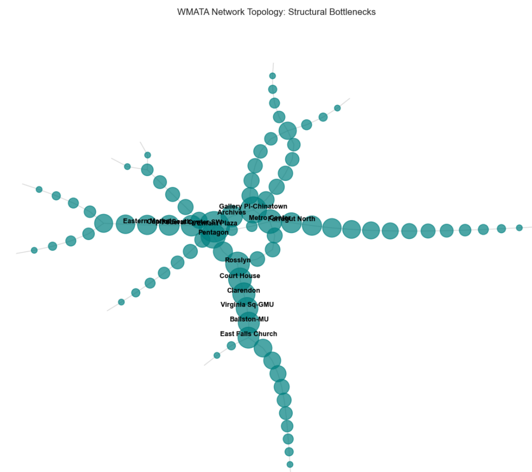

Figure 9: Rail Network Topology - WMATA

Python and we leveraged libraries such as *Pandas* for data manipulation, *NetworkX* for graph theory operations, and *Folium* to better understand and visualize the geospatial data.

For the bus route optimizations, we modeled the system as a graph where nodes represent stations and weighted edges represent travel times. To identify structural vulnerabilities in the current rail system, we modeled physical rail lines as an undirected graph using *NetworkX* module in Python.

To examine the efficiency of rail routes, we utilized station-to-station info from the WMATA API. This data broke down the distance in miles and in minutes of rail time from one station to another. From there, we calculated the speed of each station-to-station segment, in miles per hour, filtering down to the ten fastest and ten slowest. The top ten of each is filtered to remove duplicate reverse segments. To and from nodes for each top segment are larger in size, to highlight their location.

4.3 Results

Our analysis found that the most used Bus Routes were M12, D80, and D4x. These had average maximum loading times that were significantly higher than the average, suggesting that they serve high density corridors that are underserved by rail. We also found 5 structural bottlenecks, that being: L'Enfant Plaza, Gallery Pl-Chinatown, Pentagon, Rosslyn, and the Metro Center. Their betweenness and degree centrality are shown in Table 1.

Compiling the data from our route optimization, we see steady decreases in both objective values and average travel times per pair as the number of routes increases. When looking at Figure 12, we see that the objective values had the biggest decrease when going from 6 routes to 7 routes (59,976,040 to 30,056,888) and second biggest from 7 routes to 8 routes (30,056,888 to 5,492,654). Conversely, for the average times, we see from Figure 13 that the steepest decline actually occurred from 7 routes to 8 routes (25.63 minutes to 8.06 minutes) and the second biggest drop was 6 routes to 7 routes (34.97 minutes to 25.63 minutes). 11 routes does result in a large,

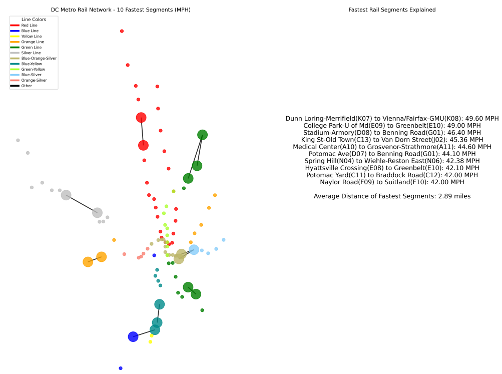

Figure 10: Ten Fastest Rail Segments

Figure 11: Ten Slowest Rail Segments

Table 1: Betweenness and Degree Centrality

Station	Betweenness Centrality	Degree Centrality
L'Enfant Plaza	0.548830	0.052083
Gallery Pl-Chinatown	0.402156	0.041667
Pentagon	0.353947	0.031250
Rosslyn	0.334942	0.031250
Metro Center	0.332822	0.041667

unexpected spike in both time and objective values, but it is more of the exception rather than the norm. The other spikes, such as ones at 14 routes and 17 routes are fare less noticeable. Beyond that, the decreases become more minor and are less noticeable, stabilizing after 12 routes to consistently stay below 2 minutes.

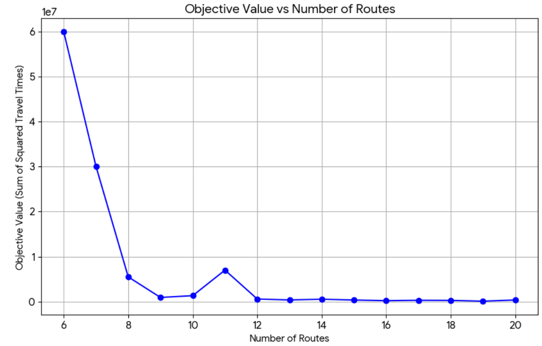

Figure 12: Objective Values (Global Sum of Squared Travel Times)

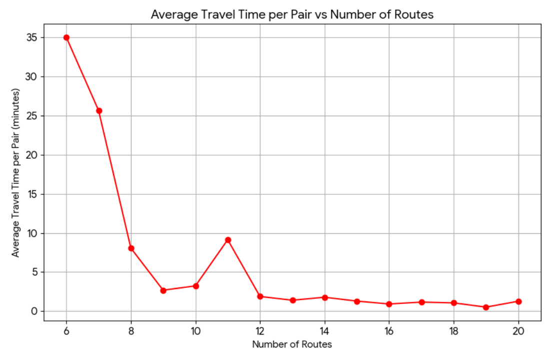

Figure 13: Average Travel Times per Pair

Table 2: Objective Values and Average Travel Times by Number of Routes

Number of Routes	Objective Value	Average Travel Time (min)
6	59,976,040.00	34.97
7	30,056,888.00	25.63
8	5,492,654.00	8.06
9	926,696.00	2.69
10	1,321,036.00	3.24
11	6,997,936.00	9.13
12	580,742.00	1.89
13	373,604.00	1.41
14	539,082.00	1.79
15	352,934.00	1.30
16	208,946.00	0.93
17	293,412.00	1.17
18	268,468.00	1.07
19	97,714.00	0.53
20	357,336.00	1.28

5 Discussions

The most concerning finding is that the Betweenness Centrality Score of 0.54, L'Enfant Plaza is in 54% of all normalized shortest paths between stations. This makes it a key conduit in the network and can be catastrophic if it were to ever shut down. It can be said that the impacts can be mitigated by routing to adjacent stations, but being that it also has the highest Degree Centrality, it is also connected to many stations. This means that there is a high likely hood that it serves as a major transfer point. When combining these two factors together, it becomes evident that a failure at this station would drastically impact service and the ability for people to get to their destination on time. Mitigating factors to these central stations can be found in quick avenues of what is known as "micromobility", the infrastructure created from commuting options such as bike or e-scooter rentals [5].

Analyzing the rail speed, it is both unexpectedly slow. In January of 2024, WMATA updated its policy to remove the limit of 59 mph on rail lines which would both improve transit time and reduce cost [2]. However, this seems to have little to no meaning when considering the top average speeds reached in these segments do not approach this former limit, with the fastest being 49 MPH. Furthermore, rail speeds for central areas drop a magnitude of 10 times slower than the fastest. According to a 2019 study by researchers at the School of Traffic and Transportation at Beijing Jiatong University, "when the interval distances are 900 m and 1700 m, the recommended design speeds are 80km/h and 100km/h, respectively", which translates to 49 MPH for 0.55 miles and 62 MPH for 1 mile [14]. The average speeds of segments are nowhere near that recommended amount. Increasing the rail speeds would lead to a cascade of benefits for the system at large, not least of which being reduced transit time. Looking at the density of slow vs fast segments, there is a clear distinction that the slowest segments are in the center and the fastest segments are near the edges.

We also found that increasing the routes had a clear and consistent decrease in both the Global Sum of Squares of Travel Times and the Average Travel Time per pair. However, it should be noted that there were diminishing returns going from 9 to 10 routes. It is abundantly clear that even going from 8 routes into 9 routes has big impacts (cutting average travel times by more than half) so the operational cost of additional busses, drivers, and maintenance is very likely worth it. But going beyond 9 results in less than a minute average travel time per pair saved, so the additional cost of a 10th route outweighs the marginal gain in efficiency. That being said, there is a world where the 10th route is worth it, or the 9th route isn't worth it. The key takeaway from this point is that there is a diminishing return, and adding more routes will not always be worth it.

6 Conclusion

This study evaluated the structural integrity and operational efficiency of the DC Metro system. From our analysis, we found that travel demand is rather uneven and the current network doesn't fully fit the patterns. Our bus route optimization experiments showed that adding more routes can substantially reduce travel time, although only to a certain point before diminishing returns. Our rail network topology also showed a small set of transfer stations to form the backbone of the entire system. A disruption at any of these stations could cause a ripple effect through the entire system. The rail speed analysis highlights a core issue that the rail itself is slow, and is nowhere near the recommended speed guidelines. Increasing speed, especially for short distances, should be a priority moving forward.

Further work could include modeling either the rail or bus system with implemented fixes, to highlight the efficacy of certain improvements. For example, rail speed could be modeled at the "recommended guideline"; from there, predict metrics on transit time, potential operating costs, and any risks. Utilizing a forecasting such as this would assist decision makers with prioritizing fixes.

We hope that the results we have generated can give WMATA insight into where weaknesses lie in the system and enable improvement across the DC Metro system.

7 Author Contributions

- Varun Jani: Developed exploratory data analysis (EDA) notebook with a focus on plotting density heatmaps, network topology, flow simulation, rail traffic intensity etc.
- Richard James Heiman: Sourced and enriched data from WMATA for analytic use. Ran a network analysis on rail speeds. Gathered supporting WMATA documentation.
- Kevin Lin: Developed a geographic gap analysis tool to suggest new rail stations and visualizes the results on a folium map.
- Benjamin Sullivan: Developed a route optimization program.

8 Data and Code Availability

Our code and data should all be available on this GitLab link. Please see README.md for more details.
GitLab Link: https://code.vt.edu/jvarun/dc-metro-analysis

Acknowledgments

We want to thank WMATA for their collection and sharing of data for public consumption, with specific kudos to the Department of Ridership and Mobility Analysis for providing un-aggregated data to us upon request in a short period of time. Their contribution enabled the success of this project.

References

[1] [n. d.]. System Map. https://www.wmata.com/schedules/maps/wmata-system-map.cfm.
[2] Washington Metro Transit Authority. 2024. Rail Program Update. In *Board Action/Information Summary*.
[3] Kamal Jyoti Barman. 2024. Analysis of metro network by applying graph theoretical notions. *J. Mech. Contin. Math. Sci.* 19, 4 (April 2024).
[4] Gen Chen and Jia Wan Zhang. 2024. Intelligent transportation systems: Machine learning approaches for urban mobility in smart cities. *Sustain. Cities Soc.* 107, 105369 (July 2024), 105369.
[5] Northern Virginia Transportation Commmision. 2025. Integrating Transit, Bicycling and Micromobility in Northern Virginia. *Northern Virginia Transportation Commission* (2025). https://novatransit.org/uploads/studiesarchive/2025NVTCTransitMicromobilityReport.pdf
[6] MACROTRENDS. 2025. Washington DC metro area population | historical data | chart | 1950-2025.
[7] Jose Antonio Panero, Travis Johnson, Sean Williams, Nii Okai Tetteh, and Ardeshir Faghri. 2024. Metropolitan rail transit improvement: The Baltimore study. *Current Urban Studies* 12, 04 (Dec. 2024), 790–814.
[8] Robert Randy, Soper, Bryse Flowers, Sandeep Kumar Bijinemula, Brian Mayer, Farnaz Khaghani, Jordan Holt, and Naren Ramakrishnan. [n. d.]. Real-time

forecasting of on-Time Performance in Metrorail systems. https://people.cs.vt.edu/naren/papers/wmata-OTP-18.pdf.
[9] Troung Robert, Gkountouna Olga, Pfoser Dieter, and Zufle Andreas. 2018. Towards a Better Understanding of Public Transportation Traffic: A Case Study of the Washington, DC Metro. *MDPI, Urban Science* (2018). https://www.mdpi.com/2413-8851/2/3/65
[10] Yalda Saadat, Bilal M. Ayyub, Yanjie Zhang, Dongming Zhang, and Hongwei Huang. [n. d.]. Resilience of Metrorail Networks: Quantification With Washington, DC as a Case Study.
[11] Svetla Stoilova and Veselin Stoev. 2015. An application of the graph theory which examines the metro networks. *Transport Problems* 10, 2 (2015). https://EconPapers.repec.org/RePEc:exl:1trans:v:10:y:2015:i:2:p:35-48
[12] WMATA. 2025. Interpreting Metrobus Tap and No-Tap Data. *WMATA* (07 2025). https://www.wmata.com/initiatives/ridership-portal/upload/FE-Bus-Metadata.pdf
[13] WMATA. 2025. Interpreting Metrorail Total Ridership Data. *WMATA* (07 2025). https://www.wmata.com/initiatives/ridership-portal/upload/FE-Rail-Metadata.pdf
[14] Yang Xu, Yun Bai, Jiajie Li, and Yuhe Zhou. 2019. Study on Design Speeds of Metro Inter-stations for Traction Energy Saving and Travel Time Reduction. *IOP Conference Series: Materials Science and Engineering* 688, 2 (nov 2019), 022055. doi:10.1088/1757-899X/688/2/022055
[15] Jinlei Zhang, Feng Chen, Zhiyong Cui, Yinan Guo, and Yadi Zhu. 2019. Deep-learning architecture for short-term passenger flow forecasting in urban rail transit. (Dec. 2019). arXiv:1912.12563 [eess.SP]

Received 16 December 2025