CS-GY 6313 / CUSP-GX 6006: Data Visualization: Final Project Proposal

Qingyuan Feng

qf2093

qf2093@nyu.edu

Abstract

This project aims to analyze and visualize the connectivity and accessibility of New York City's subway network through the lens of graph theory, treating subway line segments as vertices and connections between stations as edges. By incorporating weighting parameters such as connectivity degrees and centrality measures (utilizing eigenvalues or models like Google's PageRank), the study seeks to offer a detailed examination of the subway network's structure. The adoption of visual representation techniques, such as Force-Directed Layout or Circular Fixed Layout, aims to render the complex network both aesthetically pleasing and easy to interpret. Through this approach, the project endeavors to provide a comprehensive analysis of the network, highlighting crucial hubs and connections and offering insights into the subway system's operational and structural dynamics.

1. Introduction

New York City's subway system, an iconic symbol of urban mobility, is currently facing significant challenges, including overcrowding, aging infrastructure, and frequent service disruptions. Despite its expansive reach and the critical role it plays in the daily lives of millions, the subway system suffers from a foundational lack of planning that dates back to its inception. When the first lines were constructed in the early 20th century, the rapid expansion and evolving demands of the city were not fully anticipated, leading to a network that, while extensive, is now struggling to meet the needs of a modern, bustling metropolis. This historical oversight has led to a pressing need for innovative solutions to enhance the system's connectivity and accessibility, ensuring it can sustain the city's growth and dynamism into the future.

2. Related Work

The dataset: Subway Lines

The "Subway lines" dataset from NYC Open data. covers the line-string geometry information of NYC's subway network.

The dataset: MTA Subway Stations

Point geometry for all the subway station in NYC.

The sample works

A NYC subway connection graph using Delaunay triangulation

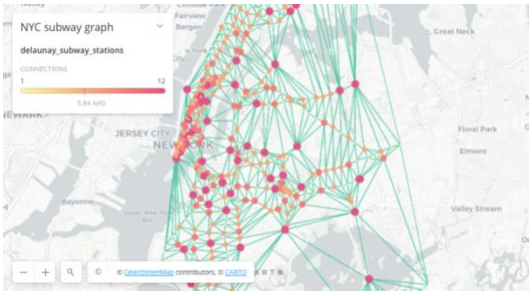

Figure 1. Sample Net work after reconstructing the subway grid by vertices and edges based on connectivity

3. Design and Implementation

Design

In designing my project, I focus on applying graph theory to model the subway system, considering each station as a vertex and each connection as an edge. I plan to assign weights to these edges based on various factors like distance, travel time, and passenger traffic, enabling a detailed analysis of the system's connectivity and efficiency. My visualization approach includes the Force-Directed Layout for an organic, structure-revealing view, using centrality measures and algorithms like eigenvalues or Google PageRank to highlight key stations and connections. In the situation that the Force-Directed Layout working well, I will try to fix the lay out to Fixed Layout for a more orderly, pattern-focused representation. After done the work on network analysis, I can do the heat map about accessibility of each zip code of NYC.

Implementation Plan

My implementation plan starts with collecting comprehensive data on the subway system, including station details and

connectivity (Subway lines). Next, I'll build the graph model, incorporating edge weights that reflect centricity and degrees of vertices (connectivity of stations). Based on the graph, I can make the fixed layout and heat map if needed.

4. Demonstration Plan

The demonstration plan for the project will showcase a series of static visual representations that elucidate the complex connectivity and operational characteristics of New York City's subway system. The primary feature will be a set of detailed network graphs, including a Force-Directed Layout and a Fixed Layout, which will illustrate the subway system's structure with nodes representing stations and weighted edges depicting connections. The inputs for these visualizations are the analyzed data sets: the "Subway Lines" providing line-string geometry and the "MTA Subway Stations" offering point geometry for stations. The outputs will be high-resolution images of the network graphs, with variations in node size and edge thickness corresponding to different centrality measures and connectivity degrees, effectively highlighting key hubs and routes within the system. Additionally, a heatmap will be presented, providing a visual summary of subway accessibility across NYC's zip codes. Each visualization will be accompanied by a narrative that interprets the data, emphasizing significant findings such as critical transit hubs, potential bottlenecks, and areas with varying levels of service accessibility. The demonstration will also include a comparison of the subway's current state with historical and potential future configurations to reflect on the system's evolution and to propose data-driven improvements. This cohesive and well-organized presentation aims to deliver a clear and aesthetically appealing narrative that is easily interpretable and offers valuable insights into the subway system's dynamics.

5. Timeline

• By April 4th: Data cleaning and filtering.

• By April 7th: data processing: doing the line and point graph and design the model for the weight for each station if have time.

• By April 14th: data processing: Done the weighting for each station, design the Fixed-Layout.

• By April 21st: data processing: Heat map and handout.

6. References

English, Jonathan. “Derailed: The Postwar End of New. York City Subway Expansion.” Journal of Urban History., vol. 47, no. 4, 2021.