THE HONG KONG POLYTECHNIC UNIVERSITY
香港理工大學

LSGI 545
Urban Informatics
January to April, 2019

Optimization of New York City Subway System

Group 4

ZHONG Lin	**(18077112g)**
WU Jiawei	**(18077211g)**
ZOU Chengchang	**(18075742g)**

By Christopher D. HIGGINS

15/04/2019

ABSTRACT

New York City, honored as an international metropolis, has the largest rapid transit system in the world. Subway system is the most significant component of public transit. However, under the trend of urbanization, New York City is experiencing continual and accelerating city expansion indicated as increasing population and residential dispersion, the existing transit system gradually does not meet the demands and needs to be extended and optimized. Our project is to dedicated to the improvement and optimization of city subway system.

INTRODUCTION

New York City Subway System has been constituted by 36 lines, 27 services and 472 subway stations so far, which run through the five boroughs of New York City, Bronx, Brooklyn, Manhattan, Queens and Staten Island.

As the increasing immigrants and residential dispersion, New York City is experiencing great extent of city expansion. As we check the subway map, there are some residential area and industrial area far away from subway's service area especially the area in the northwest side of Staten Island and east side of Queens.

Consequently, according to the map pattern, we proposed the hypothesis that new subway stations in East Queens to extend the previous line. In addition, a new subway line can be designed to connect The Bronx and East Queens running through the East River.

Based on the current condition of New York City subway network, we employed some techniques like Network Analysis, Accessibility Analysis to confirm that our hypothesis is right. We used the Gravity-type (GRAV) accessibility indices to do the Accessibility Analysis to calculate the accessibility of the locations which are suitable to build subway stations. Accessibility value is based on the factor of population and land use in our project. According to the original subway network and analysis results, the position of new subway stations is located to extend the previous line and the 37th subway line is delineated.

We sincerely hope that residents in New York City can access more improved infrastructure, enjoy the convenience of transit system and have high-quality life.

DATA AND METHODS

1. Data

The data which used in the present report include subway stations, subway lines, census information, land use situation, pedestrian network, community districts. They are downloaded in [NYC open data] at [https://opendata.cityofnewyork.us]. The data of subway information is provided by Metropolitan Transportation Authority (MTA) in 2013, and the data of census information is provided by Department of City Planning in 2010. In this project, we will mainly use the data about subway station, subway route, census and land use.

2. Method

Our analytical method can be outlined as follows (Figure 1): Firstly, we use network analysis to find out the areas which are not in the coverage of the service of subway system. Secondly, we calculate the cost value according to the weight of land use and population density, and using accessibility analysis to find suitable sites to build subway stations. Lastly, incorporating new sites to be the extension of previous subway lines or setting up new subway line.

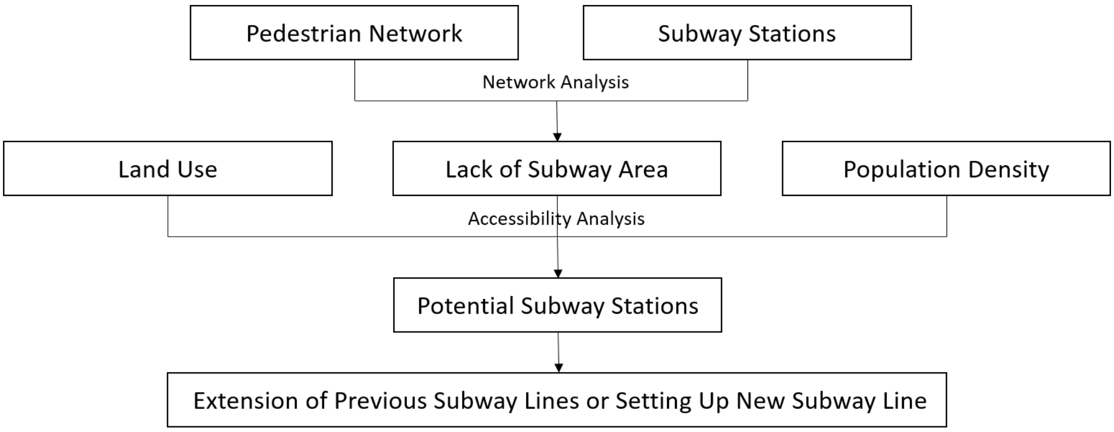

Figure 1. Analytical Structure

2.1. Network analysis

Network Analysis in GIS is based on the mathematical sub-disciplines of graph theory and topology. Any network consists of a set of connected vertices and edges. Graph theory describes, measures, and compares graphs or networks. Topological properties of networks are: connectivity, adjacency, and incidence. These properties serve as a basis for analysis.

According to the network analysis in ArcGIS, we can calculate the 20-minute service areas of existing subway location, and then we can find out which district are lack of subway stations.

2.2. Accessibility analysis

Accessibility measures in general use the impedance effect of distance, time, or generalized transport costs and the spatial distribution of urban opportunities to produce numerical indices of accessibility for each location in a study area

In this study, we choose the Gravity-type (GRAV) accessibility indices to calculate the accessibility of the locations which are suitable to build subway stations. Gravity-type (GRAV) accessibility indices, are derived by weighing the opportunities in an area by a measure of attraction and discounting each opportunity by a measure of impedance. Modified Gaussian function, a kind of impedance function will be used.

This is the function of modified Gaussian function:

$$\mathrm{A}_j = \sum_{i=1}^{n} W_j e^{-\frac{d_{ij}^2}{v}}$$

W_j is the weighted area of location j, d_{ij} is the travel time in minutes between location *i* and *j*, and the summations are for all *j* from a single-origin *i*, *v* is set to 20 in this study.

2.3. Weighted calculation

In order to calculate the accessibility of locations, the first step is to calculate the weighted area W_j. We calculate the weighted area is based on the values of population density and land use of the location.

$$W_j = w_1 V_p + w_2 V_l$$

W_j is the weighted area of location j, w_1 is the weight of population density, w_2 is the weight of land use area, and V_p is the value of population density, V_l is the value of land use area. w_1 and w_2 are showed in table 1.

Table 1. Indicators Weights and Description

Indicators	Rank order	Resulting weights	Description
Population Density	2	0.4	More efficient public transport is required for higher population density
Land Use Area	1	0.6	More efficient public transport is required for larger commercial and residential areas

The value of land use area is calculated by different categories area of land use. This is the calculation formula:

$$\mathrm{V}_l = \sum_{i=1}^{n} w_i A_i$$

V_l is the value of land use area, w_i is the weight of land use category *i* (Table 2), and A_i is the area of land use category *i*.

Table 2. Indicators Weights and Description

Indicators	Category	Rank order	Resulting weights
Land Use Area	Commercial & Office Buildings	2	0.16
	Industrial & Manufacturing	6	0.09
	Mixed Residential & Commercial Buildings	1	0.2
	Multi-Family Elevator Buildings	3	0.15
	Multi-Family Walk-up Buildings	4	0.12
	One & Two Family Buildings	5	0.1
	Open Space & Outdoor Recreation	7	0.05
	Parking Facilities	7	0.05
	Public Facilities & Institutions	7	0.05
	Transportation & Utility	10	0.03
	Vacant Land	11	0

This is the code we used in Python to calculate the weighted area:

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import geopandas as gpd
# Read the Data sheet from the shapefile
landuse = gpd.read_file("F:\\POLYU\\urban informatics\\Group\\data\\landuse_join.shp")
census = gpd.read_file("F:\\POLYU\\urban informatics\\Group\\data\\census_join_2.shp")

landuse_cat = {'Commercial & Office Buildings':0.16, 'Industrial & Manufacturing':0.09,
'Mixed Residential & Commercial Buildings':0.2, 'Multi-Family Elevator Buildings':0.15,
'Multi-Family Walk-up Buildings':0.12, 'One & Two Family Buildings':0.1, 'Open Space &
Outdoor Recreation':0.05, 'Parking Facilities':0.05, 'Public Facilities &
Institutions':0.05, 'Transportation & Utility':0.03, 'Vacant Land':0}
landuse_name = list(landuse_cat.keys())
Value = []
Landuse_value = []

for i in range(len(census.Geography)):
    landuse_all = landuse.loc[landuse.Geography == census.Geography[i], :]
    landuse_value = 0

    for j in landuse_name:
        landusetype = landuse_all.loc[landuse_all.LandUseCat == j, :]
        area = sum(landusetype.Shape_Area)
        areavalue = landuse_cat[j] * area
        landuse_value += areavalue
    Landuse_value.append(landuse_value)
```

```
for k in range(len(census.Geography)):

    max_landuse_value = max(Landuse_value)
    min_landuse_value = min(Landuse_value)
    max_population = max(census.TtlPop2010)
    min_population = min(census.TtlPop2010)
    pop = (census.TtlPop2010[k] - min_population) / (max_population - min_population)
    land = (Landuse_value[k] - min_landuse_value) / (max_landuse_value - 
min_landuse_value)
    value = 0.4 * round(float(pop) , 4) + 0.6 * round(float(land) , 4)
    Value.append(value)

census['Value'] = Value
census['Landuse_value'] = Landuse_value
census.to_file("F:\\POLYU\\urban informatics\\Group\\data\\census_value.shp")
```

RESULTS

According to the network analysis, the lacking subway districts are found (yellow areas in Figure 2). They include Bayside-Little Neck, Canarsie–Flatlands, Flushing–Clearview, Willowbrook, Southeast Queens, Fresh Meadows, Port Richmond, Pelham-Throgs Neck and Jamaica neighborhoods.

MAP OF SELECTIVE COMMUNITY AREA

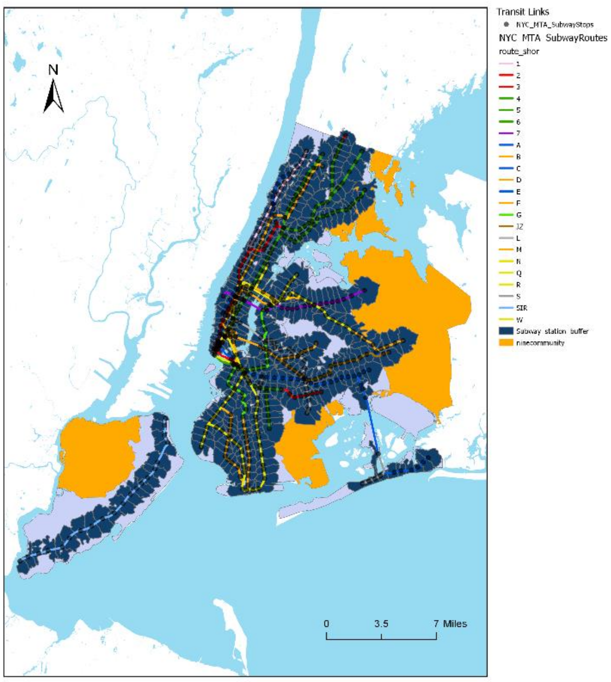

Figure 2. Map of Lacking Subway Neighborhoods

After getting the lacking subway neighborhoods, we calculate the value of weighted area of each location, and use the Accessibility Calculator tool in ArcGIS to calculate the accessibility of each location. Figure 3 is the result of accessibility analysis of each neighborhood.

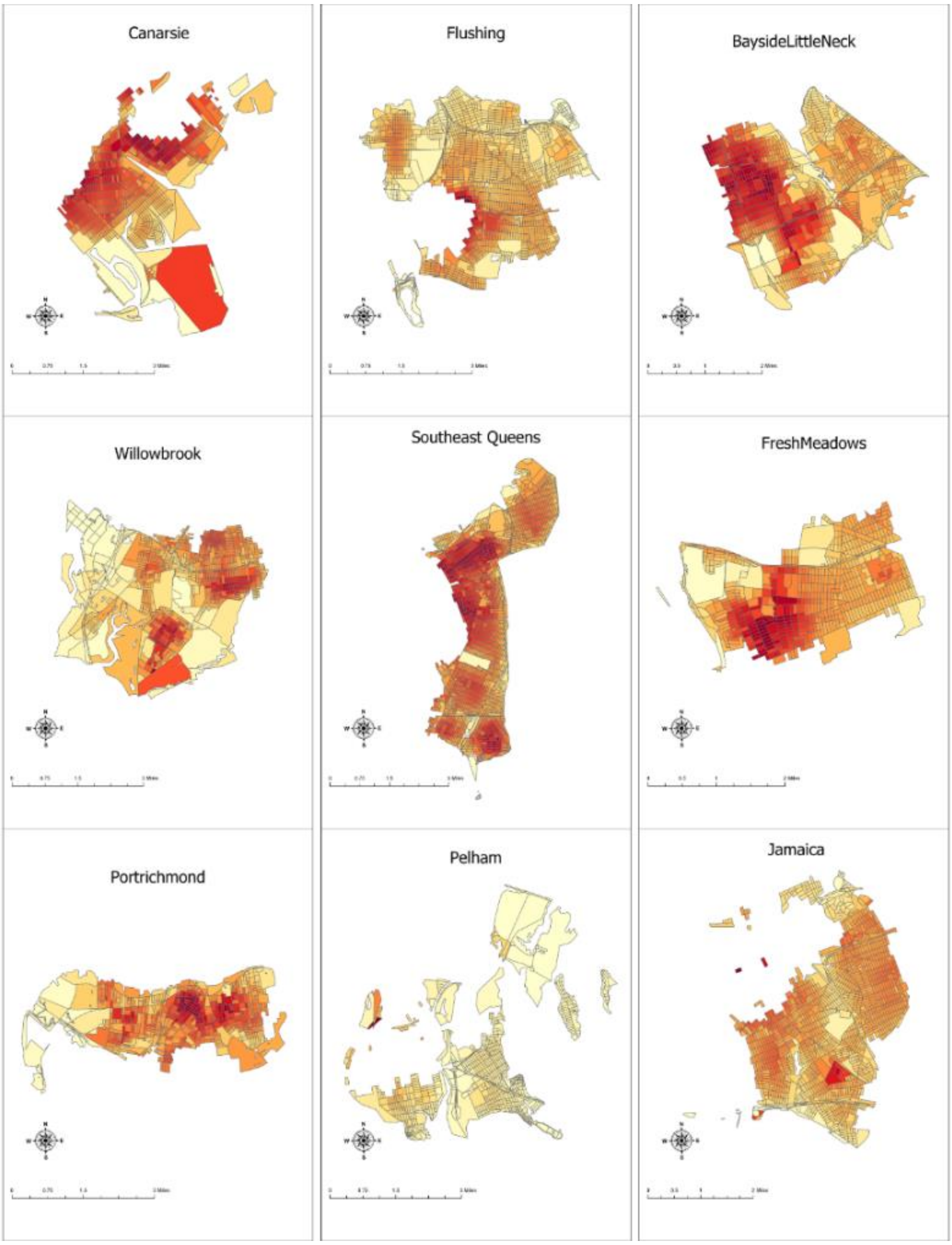

Figure 3. Accessibility Value of Lacking Subway Neighborhoods

According to the accessibility value of each location, we choose the largest value of accessibility of locations as the suitable new subway stations. The result is in the Figure 4. Figure 5 is the detail of the new subway line and new subway location in each neighborhood.

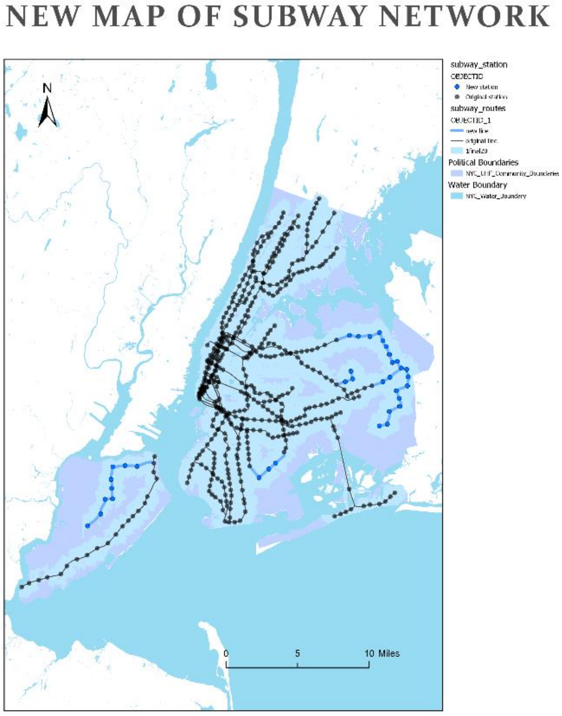

Figure 4. New Map of Subway Network

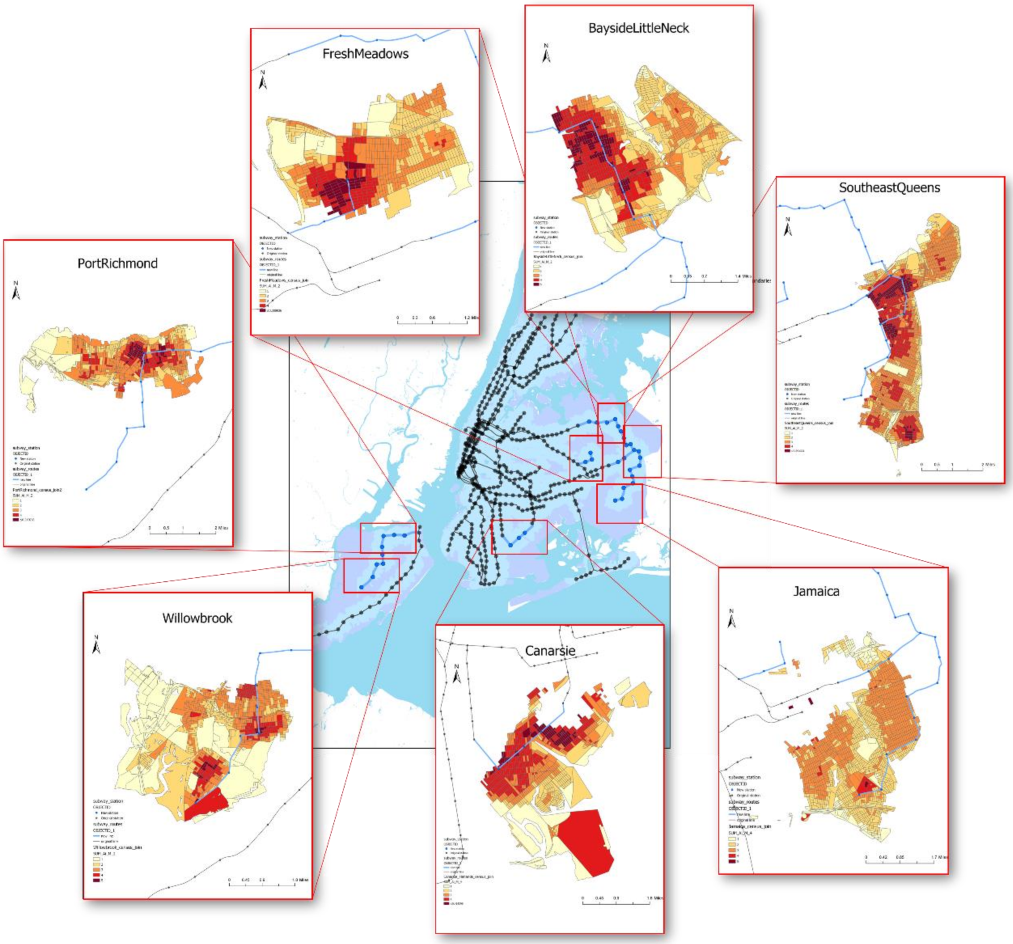

Figure 5. Optimization of New York City Subway System

CONCLUSION

After the data processing and data analysis for a period of time, our hypothesis is confirmed right and reasonable. The service of New York City Subway System does not meet the current travel demands of residents. The system needs to be improved and optimized.

The extension and optimization of New York Subway System can be concluded as these lists below.

(1) There are 36 subway stations newly designed in the New York city planning subway map, 24 stations in Queens, 9 stations in Staten Island, 3 stations in Brooklyn, totally 529 subway stations.

(2) As for subway line, in Staten Island, an extension line is delineated cross through 9 stations to connect the island from the southwest side to northeast side. In Queens, three subway lines are elongated eastward and even a new subway line is delineated from north to south in the east side of Queens. The subway line is also lengthened in the southeast side of Brooklyn.

REFERENCES

[1]. Mei, Z. X., Ou, Y. J., & Chen, S. M. (2011). Accessibility Evaluation of Guangzhou Subway Network Based on Space Syntax Model and GIS. In Applied Mechanics and Materials (Vol. 90, pp. 3292-3297). Trans Tech Publications.

[2]. Zhu, W. J., & Kan, X. D. (2012). Coordination analysis for urban underground space and rail transit based on space syntax. In Applied Mechanics and Materials (Vol. 170, pp. 1388-1391). Trans Tech Publications.

[3]. Alam, B., Thompson, G., & Brown, J. (2010). Estimating transit accessibility with an alternative method: evidence from Broward County, Florida. Transportation Research Record: Journal of the Transportation Research Board, (2144), 62-71.