Mobility Vitality: Assessing Neighborhood Similarity Through Transportation Patterns In New York City

Dan Qiang
Platial Analysis Lab, Department of Geography, McGill University, Montréal, Canada

Grant McKenzie
Platial Analysis Lab, Department of Geography, McGill University, Montréal, Canada

Abstract

Though numerous studies have examined human mobility within an urban environment, few have explored the concept of urban vitality purely through the lens of urban transportation. Given the importance of different modes of transportation within a city, such analysis is necessary. In this short paper, we introduce the novel concept of mobility vitality by integrating human mobility and urban vitality, offering a multilayered framework to assess the degree of transportation and mobility within and between regions. The mobility patterns of three transportation modes, namely subway, taxicab, and bike-share, are first examined independently. These patterns are then aggregated to form the composite measure of static mobility vitality. Through this measure, we evaluate similarities between neighborhoods. Our results observed significant spatial differences in the travel patterns of three transportation modes on weekdays and weekends. Moreover, neighborhoods with high static mobility vitality have relatively similar mobility patterns. Ultimately, this approach aims to find neighborhoods with imbalanced transportation infrastructure or inadequate public.

2012 ACM Subject Classification Information systems → Geographic information systems; Applied computing → Transportation

Keywords and phrases mobility vitality, mobility similarity, transportation, bike-sharing, taxi, subway, New York City

Digital Object Identifier 10.4230/LIPIcs.GIScience.2023.61

Category Short Paper

1 Introduction

In 1961, the urban activist Jane Jacobs introduced the concept of *urban vitality* as a qualitative measure of a city's pulse [2]. The idea suggests that varying tempos of human activities and pedestrian flow can all be employed to differentiate regions. For decades, most of the research related to this concept was done using qualitative surveys, demographic studies, and narrative analysis. The difficulties with such approaches are costly and labor-intensive and are prone to subjective biases. The recent dramatic growth of publicly accessible activity and mobility data has set the stage for alternative approaches to assessing urban vitality.

Despite a large body of literature targetting the extraction of individual human mobility patterns and their accompanying impact variables [1, 7], little attention has been paid to urban dynamics characterized purely by individual movement. Recently, a growing number of research teams have focused on temporal characteristics of mobility to better understand urban vitality [3]. For instance, Sulis et al. [6] examined smart-card rail trips to assess spatiotemporal variation in urban vitality in London. They produced a set of three dynamic properties, namely the number of people, the continuity, and the fluctuations of this presence over particular intervals of time. Similarly, Zeng et al. [9] created a new index to measure urban vitality based on records from a bicycle-sharing system. Further work has demonstrated that *lively* regions of a city correlate with taxi drop-off locations [10]. A variety of research has shown that urban vitality/vibrancy can be measured through data ranging from social check-ins and points of interest to trajectories and mobile phone data [4, 8].

Though progress is being made, research focused exclusively on mobility as a measure of urban vitality is lacking [11]. In exploring the vitality of different parts of a city through a mobility lens, one is able to identify the impact that access to different mobility modes, has on city cohesion. Furthermore, a combination of mobility signatures can be used as a measure through which different regions of a city can be compared [5]. Urban and transportation planners can use such a measure to better understand the impacts of policy decisions on the vibrancy and vitality of the city as a whole. Through integrating human mobility with urban vitality, we proposed the novel concept of *mobility vitality*, serving as a multilayered framework to evaluate the degree of transportation and mobility within a space. In this preliminary work, we aim to address the following two research questions (`RQ`).

- `RQ1` Can a region, *e.g.,* neighborhood, be quantified by the mobility patterns of different modes of transportation that exist and traverse the region? Furthermore, do these patterns vary by mode and region?
- `RQ2` Can mobility vitality, as represented by a combination of mobility patterns, be used to compare and differentiate regions within the same city?

We address these questions through an analysis of three different modes of transportation within New York City (NYC). As the most densely populated city in the United States, NYC's transportation ecosystem is both complex and extensive. The scale of our analysis is neighborhoods within the five boroughs of NYC and the extent of analysis varies based on the service area of each transportation system.

2 Data and Analysis

To start, three data sets representing three very different modes of transportation were collected. These include bike-share, subway (rail), and taxicab data. We restricted our analysis to May 2019, cleaned the data to remove errors, and aggregated the month of data to days in a typical week. We use this week as a representative sample of transportation usage in NYC. May was chosen due to the limited holidays, historically decent weather, and fewer people on summer vacation. For micro-mobility, we accessed data for the widely used bicycle-sharing system, *Citi Bike*[1]. Citi Bike is a privately operated docking station-based bike-sharing system. Citi Bike trip data include the start and end times of each trip as well as the origin and destination stations. For mid-sized transportation, we accessed trip data for *yellow taxis*[2]. The yellow taxi trip records include fields capturing pick-up and drop-off dates, times, and locations. For mass transit, we analyzed turnstile data of the *NYC subway system*[3]. These data report an accumulated number of entrances and exits, per station at a four-hour temporal resolution. All data were cleaned to remove erroneous trips (e.g., those that were one minute in length, 200 miles, etc).

Next, we intersected the trip data with the NYC neighborhood boundaries[4] to assign trip volume for each of the three services to each neighborhood in NYC. The assigned volume includes both origins (entries) and destinations (exits). More specifically, the numbers of

[1] `https://citibikenyc.com/system-data`
[2] `https://data.cityofnewyork.us/Transportation/2019-Yellow-Taxi-Trip-Data/2upf-qytp`
[3] `https://data.ny.gov/Transportation/Turnstile-Usage-Data-2019/xfn5-qji9`
[4] `https://data.cityofnewyork.us/City-Government/2020-Neighborhood-Tabulation-Areas-NTAs-Tabular/9nt8-h7nd`

origins and destinations were combined to determine the final trip volume. For the subway turnstile, the total number of entries and exits for every turnstile within a station was summed. For example, there are four control areas in the "Cortlandt St." station and each control area has 10 turnstiles. The trip volume for that station was calculated as the sum of all passengers through the 40 turnstiles. The trip data were then divided by the populations of their respective neighborhoods. This process was straightforward for the bike-share and subway turnstile data as they are represented as point geometries. The taxicab trip data, however, is reported by polygonal taxi zone[5] (TZ). A dasymetric mapping approach was used to allocate taxicab trip origins and destination TZ data to the NYC neighborhood boundaries.

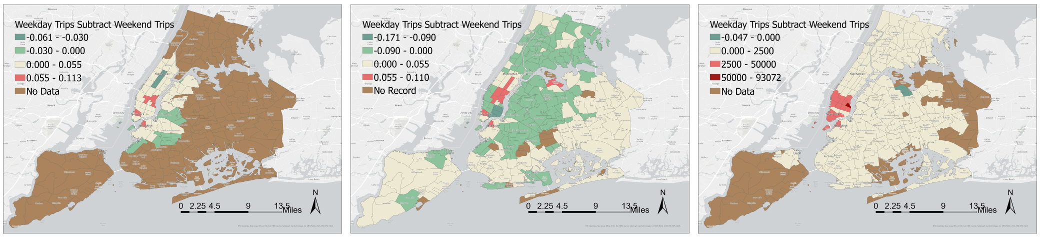
To address `RQ1`, our static[6] *mobility vitality* measure was generated by summing the individual transportation mobility patterns across each region, producing a single value for each neighborhood. While we took an "equal weights" approach here, the measure is designed to allow a user to adjust the importance (weights) of each individual transportation mode in the overall mobility vitality result, depending on their interests. Given this measure of mobility vitality, we then examined how such a measure could be used to better understand the vitality and variability of mobility services within a city such as NYC. To start, we averaged the mobility vitality measure for each neighborhood by weekday and weekend. This allowed us to subtract weekend mobility vitality from weekdays to better identify temporal variations in mobility and differentiate neighborhoods based on prototypical commuting behavior. Finally, we examined mobility vitality as a measure on which to identify similarities between neighborhoods based purely on how inhabitants and visitors use different transportation systems. To address this `RQ2`, we used Jensen-Shannon divergence (JSD), a method for assessing the (dis)similarity of two probability distributions. In our case, we took the trip volume for each day of the week of our three transportation modes as a *distribution*. Having one distribution for each neighborhood allowed us to assess the similarity between all neighborhood pairs. We then identified the neighborhoods that were most similar to all other neighborhoods and those that were most unique.

3 Results and Discussion

For all three modes of transportation, there is greater mobility activity on weekdays than on weekends. For bike-share origins and destination points, the population-normalized mean values are 0.028 and 0.021, for weekdays and weekends, respectively. Similarly, the mean population-normalized taxi pick-up density on weekdays is 0.0143, while on weekends it is 0.0137. The subway turnstile data was much more pronounced with a population-normalized weekdays value of 1,407.86 and a weekend value of 839.91. These large values indicate that, for many of the neighborhoods within NYC, the number of subway passengers is several orders of magnitude higher than the residential population.

The weekday/weekend variation in normalized transportation trips is shown in Figure 1. In order to compare weekday trips and weekend trips, we delineated the legend on the maps by setting 0 as the dividing line in the class intervals. In both bike and taxi categories, those values greater than 0 and those less than 0 were separately averaged into two intervals. For the metro map, given the significantly higher number of weekday trips compared to weekend ones, only one level was established for values less than 0, while those greater

[5] `https://data.cityofnewyork.us/Transportation/NYC-Taxi-Zones/d3c5-ddgc`
[6] Static here refers to the fact that temporal variability was not included in this approach.



(a) Bike share. **(b)** Taxicab. **(c)** Subway turnstile.

Figure 1 Population-normalized weekend trip counts subtract from weekday trip counts, for three modes of transportation.

than 0 were evenly divided into three levels. In general, the Manhattan business district witnesses a predominance of weekday trips over weekend ones, with the intensity varying across transportation modes. Bike sharing predominantly favors weekdays, with Central Park being the exception. Conversely, neighborhoods encompassing recreational areas report higher weekend bike trip volumes. Taxi trips exhibit a starkly distinct pattern, with higher weekend volumes in both northern and southern Manhattan, notably in downtown neighborhoods near Queens. Subway data, however, shows a universal weekday preference, except in East Elmhurst and North Corona. A clear spatial clustering of neighborhoods with the greatest discrepancy between weekday and weekend trips is evident in downtown Manhattan.

The results of the equally-weighted static mobility vitality measure are shown in Figure 2. The operating region for the bike share service is the most spatially restrictive of our data and so all data sets were restricted to this analysis area. As one can see, the greatest degree of mobility vitality is in Central Park and the southeast corner of Manhattan. As one moves towards the east side of Brooklyn and north Harlem, the vitality gradually decreases.

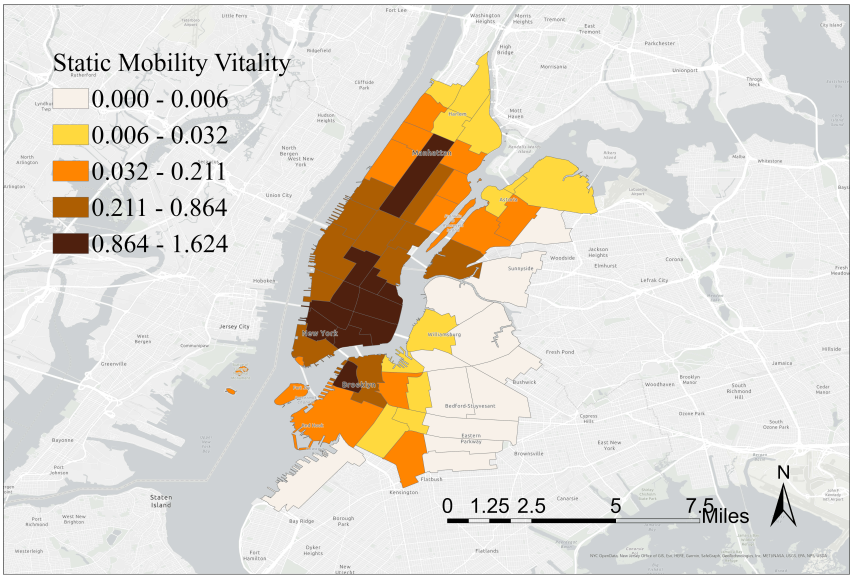


Figure 2 Static mobility vitality as calculated by summing the population-normalized trip volume from three different modes of transportation.

The results of our Jensen-Shannon divergence approach are shown in Figure 3. In this Figure, pink neighborhoods are the most unique neighborhoods in terms of mobility vitality, reporting the highest average JSD values. These include the Upper West Side (Central), Upper East Side-Yorkville, East Midtown-Turtle Bay, Williamsburg, Harlem (North), and Astoria (North)-Ditmars-Steinway. Among these, four are located in Manhattan, while Queens and Brooklyn each contain one. Comparing these results to the static mobility vitality map shown in Figure 2, it can be observed that the six most dissimilar neighborhoods are not the ones with the highest static mobility vitality. They belong to the lower-scoring group in terms of the three individual mobility patterns as well as static mobility vitality. This speaks to the influence of the temporal dimension on assessing the vitality of a city with respect to mobility. In our data, most neighborhoods with high static mobility vitality have relatively low divergence values, indicating that they tend to be similar to one another. Neighborhoods with low JSD values are in regions with high volumes of everyday traffic for each mode of transportation and their individual mobility patterns show little variation. The six most unique neighborhoods, as measured by our mobility patterns, are scattered throughout the city but share a common characteristic, they are all waterfront neighborhoods.

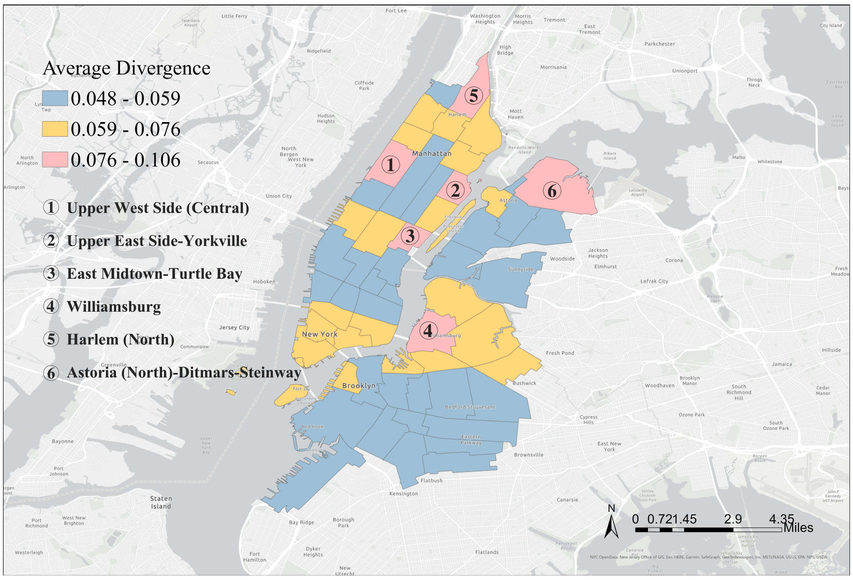


Figure 3 Unique and similar neighborhoods as measured through three modes of transportation using Jensen-Shannon divergence.

4 Conclusions and Next Steps

In this preliminary work, the concept of mobility vitality is proposed to measure the degree of transportation and mobility within a region. This work investigates mobility vitality patterns when different transportation starts or ends in the neighborhood and uses these patterns to identify the divergence between different neighborhoods within NYC. Not surprisingly, we found that mobility patterns are different on weekdays than on weekends. In most cases, the volume of trips in downtown neighborhoods is greater during weekdays than on weekends; however, taxicabs in some central business districts are the exception. Additionally, neighborhoods with excessive divergence are dispersed and more dissimilar neighborhoods often exhibit a high degree of clustering and high mobility vitality.

The next steps for this work will involve including additional modes of transit and assessing the robustness of our approach through varying types of transportation. Our current mobility vitality approach is meant as a “proof-of-concept” and further iterations will allow users to vary the weights depending on the question they are investigating. Last, the current analysis was conducted using data collected at a daily temporal resolution. We aim to examine the spatiotemporal characteristics of mobility vitality with a finer time granularity in the future.

The results of the analysis presented in this short paper are meant to offer a glimpse at the objective of generating a mobility vitality measure that represents the spatial and temporal dynamics of mobility within a city. Through developing such a measure, our aim is to empower urban and transportation planners with measures by which similarities and differences within a city can be identified. Planners and government agencies will be able to monitor how transportation policies can change vitality within a city and use such a measure to improve equitable access to transportation systems within the urban environment.

References

1 Ezgi Eren and Volkan Emre Uz. A review on bike-sharing: The factors affecting bike-sharing demand. *Sustainable Cities and Society*, 54:101882, 2020. doi:10.1016/j.scs.2019.101882.

2 Jane Jacobs. *The death and life of great American cities*. Random House, New York, 1961.

3 Chaogui Kang, Dongwan Fan, and Hongzan Jiao. Validating activity, time, and space diversity as essential components of urban vitality. *Environment and Planning B: Urban Analytics and City Science*, 48:1180–1197, 2021. doi:10.1177/2399808320919771.

4 Qian Li, Caihui Cui, Feng Liu, Qirui Wu, Yadi Run, and Zhigang Han. Multidimensional Urban Vitality on Streets: Spatial Patterns and Influence Factor Identification Using Multisource Urban Data. *ISPRS International Journal of Geo-Information*, 11(1):2, 2022. doi:10.3390/ijgi11010002.

5 Grant McKenzie and Daniel Romm. Measuring urban regional similarity through mobility signatures. *Computers, Environment and Urban Systems*, 89:101684, 2021. doi:10.1016/j.compenvurbsys.2021.101684.

6 Patrizia Sulis, Ed Manley, Chen Zhong, and Michael Batty. Using mobility data as proxy for measuring urban vitality. *Journal of Spatial Information Science*, 16:137–162, 2018. doi:10.5311/JOSIS.2018.16.384.

7 Yang Xu, Dachi Chen, Xiaohu Zhang, Wei Tu, Yuanyang Chen, Yu Shen, and Carlo Ratti. Unravel the landscape and pulses of cycling activities from a dockless bike-sharing system. *Computers, Environment and Urban Systems*, 75:184–203, 2019. doi:10.1016/j.compenvurbsys.2019.02.002.

8 Yihong Yuan and Martin Raubal. Analyzing the distribution of human activity space from mobile phone usage: an individual and urban-oriented study. *International Journal of Geographical Information Science*, 30(8):1594–1621, 2016. doi:10.1080/13658816.2016.1143555.

9 Peng Zeng, Ming Wei, and Xiaoyang Liu. Investigating the Spatiotemporal Dynamics of Urban Vitality Using Bicycle-Sharing Data. *Sustainability*, 106(1):1714, 2020. doi:10.3390/su12051714.

10 Bin Zhang, Shuyan Chen, Yongfeng Ma, Tiezhu Li, and Kun Tang. Analysis on spatiotemporal urban mobility based on online car-hailing data. *Journal of Transport Geography*, 82:102568, 2020. doi:10.1016/j.jtrangeo.2019.102568.

11 Zhonghao Zhang, Yusi Zhang, Tian He, and Rui Xiao. Urban Vitality and its Influencing Factors: Comparative Analysis Based on Taxi Trajectory Data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 15:5102–5114, 2022. doi:10.1109/JSTARS.2022.3183176.