A Multilayer Mobility Network Approach to Inferring Urban Structures Using Shared Mobility and Taxi Data

Abstract
Developing data-driven approaches to understanding urban structures is important for urban planning. However, it is still challenging to combine different transport datasets into a unified framework and reveal the dynamics of urban structures with the emergence of shared mobility. In this study, we propose two empirical multilayer networks to infer and profile urban structures. First, a temporal network is constructed using traditional taxi data over years to reveal the urban structures. Second, a multimodal network is constructed using shared mobility and traditional taxi data over a year to reveal the urban structures. The proposed networks are tested in New York City using a large volume of shared bike, shared vehicle, and traditional taxi data. The multilayer network centralities and community detection enable us to profile the characteristics of the urban flows and urban structure. The analytical results allow us to acquire a better understanding of urban structures from a multilayer perspective and also provide a geocomputation framework that is useful for urban and geographic researchers.

Keywords: Urban structure, shared mobility, multimodal transport, multilayer network.

1. Introduction
Urban structure refers to the spatial arrangement of land use in urban areas. It has been a subject of interest for geographers and urban planners to explain the urban structure based on social demographics or environmental settings (Rodrigue et al, 2009). In recent decades, the knowledge about urban structure has extended due to massive individual-level and high-frequency mobility data (Zhong et al., 2014; Sarkar et al., 2017; Zhang et al., 2018; Yildirimoglu & Kim, 2018). Based on the new datasets and analytical approaches, one can observe and infer urban structures that are formed by diverse types of travel behavior reflected by travel flows, namely orientation-destination (OD) data. Similar to spatial borders formed by physical configurations (e.g., rivers and mountains), the compound effect of facilities and travel purposes differentiates one zone from another, which works as an underlying structure (Jiang & Yao, 2010). Hence, mobility data are important to capture the travel behavior that emerged due to the underlying urban structure, and vice versa.

Relations and interactions between places play a critical role in the process of inferring the urban structure (Rodrigue et al, 2009). To capture such interactions, transport flows and mobility patterns have been mostly exploited from a network perspective (Barthélemy, 2011; Zhong et al., 2014; Zhang et al., 2018). In the context of network theory, the spatial structure of cities is inferred by modeling travel behavior using graphs. A graph that represents places as nodes and travel flows between nodes as edges can be partitioned into subgraphs, each of which is a collection of similar nodes (e.g., similar places in urban context). In practice, the projection from the network structure to the urban structure has been empirically tested in large-scale communication networks (Ratti et al., 2010) and mobility networks (Zhong et al., 2014; Zhang et al., 2018). However, it should be noted that there are two major challenges that remain in existing works. First, different mobility data may unveil the urban structure from different interaction perspectives, but single source data are insufficient overall for evaluating the urban structure. Second, those studies that compared travel patterns of multiple transport data treat and analyze each type of transport flow separately. Advancing the approach to better integrate and compare multiple data sources for urban structure inference is necessary to tackle the increasing complexity of travel behavior. The travel flows of different transport modes at different times are the multiple facets/layers existing in the same urban space, which will contribute to a more comprehensive understanding of urban structures.

In this study, we adopt the advanced definition and methods of network theory, i.e., multilayer network analysis, to represent multidimensional travel flows and to understand multilayer urban structures. A multimodal network model consisting of both shared mobility and traditional taxi data in a year and a temporal network model consisting of taxi data over six years are proposed and analyzed to explore the urban structures. Instead of layer-by-layer analysis, we integrate multiple datasets in a multilayer network, and the node centralities and community detection in this context make it possible to compare the feature differences among urban locations (i.e., nodes in network). The novelty of the approach is demonstrated in a case study in New York City, which aims to make new contributions to the following research questions: (1) what are the travel patterns in the multilayer mobility network consisting of shared mobility for traditional taxis, (2) what urban structures are inferred in terms of multilayer place importance, and (3) what urban structures are inferred and varying in the different layers (i.e., transport modes and years) of a multilayer network? The approach used in this study is a new adoption of network theory in the field of urban structure analysis considering shared mobility, which can be further used as a geocomputation method when studying other urban issues.

The remainder of this work is structured as follows. Section 2 presents the related work to justify the feasibility of analyzing urban structures based on mobility networks and summarizes existing knowledge on shared mobility and the advantages of multilayer network analysis. Section 3 briefly introduces the study area and multisource transport data. Section 4 presents the methodology of defining the empirical multilayer network models for the proposed questions and explains the techniques used for analysis. Section 5 presents the results, and we discuss and conclude the entire study in Section 6.

2. Related work

The inference of urban structures based on travel patterns has a long history in the geography research (Handy, 1996). However, the methods and insights are still limited due to the lack of large-scale location data. In recent decades, ubiquitous GPS-enabled sensing technologies have made positioning data increasingly available at the individual level. Such an empirical dataset makes it possible to observe and analyze human mobility and the underlying urban structure at a finer resolution. Abundant work has utilized positioning data to investigate travel behavior using spatial-temporal perspectives (Tao et al., 2014; Luo et al., 2017; Li et al., 2019) and transport mode choices (Paulssen et al., 2014; Klinger & Lanzendorf, 2016; Li et al., 2019). Several early studies adopted network analysis to understand the urban structure from the spatial interactions extracted from mobility data (Zhong et al., 2014; Sarkar et al., 2017; Zhang et al., 2018). Similarly, researchers in the time geography field have argued that mobility-related big data make it more feasible and effective to link travel patterns and urban structures (Chen et al., 2016). Although these studies reveal urban structure projects based on the properties or topology of mobility networks, the multidimensional complex interactions cannot be effectively characterized from the analysis of monoplex (i.e., single-layer) network. In addition, network communities detected in single-layer networks are not directly comparable, making it difficult to analyze the variation of interaction patterns and the associated urban structure.

Mobility-related big data of traditional transportation (e.g., travel records of buses and taxis) have been widely used to proxy urban flows and reveal urban structures (Zhong et al., 2014; Zhang et al., 2018). However, limited research has extracted urban flows and related urban structures using shared mobility data. In particular, the possible changes in the travel context due to the emergence of the new travel modes should be quantitatively modeled and analyzed. In recent years, the rise of shared mobility services (e.g., shared taxis and shared bikes) has occurred in many cities due to the wide use of smartphones and seamless information exchange platforms (Cannon & Summers, 2014). These services have been reported as one of the major travel options besides traditional transportation systems (Cannon & Summers, 2014; Wallsten, 2015). The automatic collection of shared mobility data has facilitated urban dynamics research in various topics. Most literatures are dedicated to evaluating the benefits of shared transportation for traffic conditions (Shmueli et al., 2015; Alexander & González, 2015; Li et al., 2016) and exploring travel patterns (Qian et al., 2015; Hochmair, 2016; Xu et al., 2019). The travel patterns of shared taxis significantly differ from the traditional patterns in supplementing distant commuting while the travel patterns of shared bikes are highly associated with public transport stations (Shen et al., 2017; Cao et al., 2019).

Follow-up studies consider and compare multiple urban flows of traditional and shared transportation to study travel behavior. The changes are significant in some ways. For example, shared bikes particularly reduce bus ridership and work more efficiently in dense areas than traditional taxis (Campbell & Brakewood, 2017; Faghih-Imani et al., 2017). There is very limited research on shared mobility flows using graph analysis. Yang et al. (2019) applied graph-based analysis to shared bike mobility data and quantified the temporal changes in travel structures, providing empirical evidence on urban metabolism theory. The study shows that network theory is quite important for understanding urban structures while the lack of considering multiple mobility flows in the same framework may not be as complex as the real situation. A more comprehensive framework to handle the complexity of multidimensional mobility data to understand travel behavior and urban structure is on the agenda.

Urban space bears various travel interactions at the same time, which means that multiple distinct networks exist and interact simultaneously. The fundamental advantage of network theory is that the network model and metrics effectively represent the processes and dynamics in real-world cases (e.g., Barabasi, 2005; Newman, 2006). In the context of cities, Batty (2013) proposed a new paradigm known as "*The new science of city*" that emphasizes the importance of considering flows among different entities in urban analytics. However, in nature, real-world systems are composed of multidimensional interactions (e.g., cooccurrence or interdependent interactions), and a single-layer network may provide biased conclusions for systems that consist of subsystems (Battiston et al., 2017). Although city and travel transit are similar to such complex systems, early studies focused more on single networks due to the limited development in network science (Ferber et al. 2005; Dimitrov & Ceder 2016). A multilayer network considers the co-occurrence of multiple relationships into the topology, which fits the real-world scenario and the topic of this paper better.

Recent developments in multilayer networks have made substantial progress. Besides extending the network definition by adding 'layer' and 'aspect', some network measures (e.g., centralities) are extended to analyze the structure of a multilayer network. With the new framework and techniques, Parshani et al. (2011) examined the robustness of the two-layer network structure of port and airport networks worldwide. Halu et al. (2014) model Indian air and train transportation as a multilayer network and find that the heterogeneity of the two networks enables good navigability performance of the global network. In recent years, a number of studies have applied multilayer work in various scenarios such as predicting epidemic transmission (Zhao et al., 2014), accessibility models (Strano et al., 2015; Aleta et al., 2017), social-physical travel behavior (Hristova et al., 2014 May; Gao & Tian, 2019), and urban structures (Yildirimoglu & Kim, 2018). According to the existing literatures, multilayer network analysis has great potential in modeling human mobility networks while empirical studies on urban structures are quite limited.

3. Study area and datasets

New York City (NYC), which has a population of approximately 8.4 million, is one of the largest megacities in the United States. A total of 264 taxi zones are the spatial unit in this study (Figure 1). Newark, a taxi zone outside the NYC, is not included in the following analysis, this paper focuses on the zones within five NYC boroughs. Taxi zones are the official boundaries for pricing and analyzing taxi trips and therefore are naturally suitable to analyze shared vehicles. Since the taxi zones are relatively dense in the Manhattan area, where shared bikes play the same important role as other transport modes in this region (Faghih-Imani et al., 2017), we argue that the selected spatial unit is suitable to capture human mobility using shared bikes. Using the same spatial unit, the taxi zones are used to extract the interzone urban flows of different layers (e.g., modes and years) and regarded as nodes in the multiplex network, which will be further discussed in Section 4.

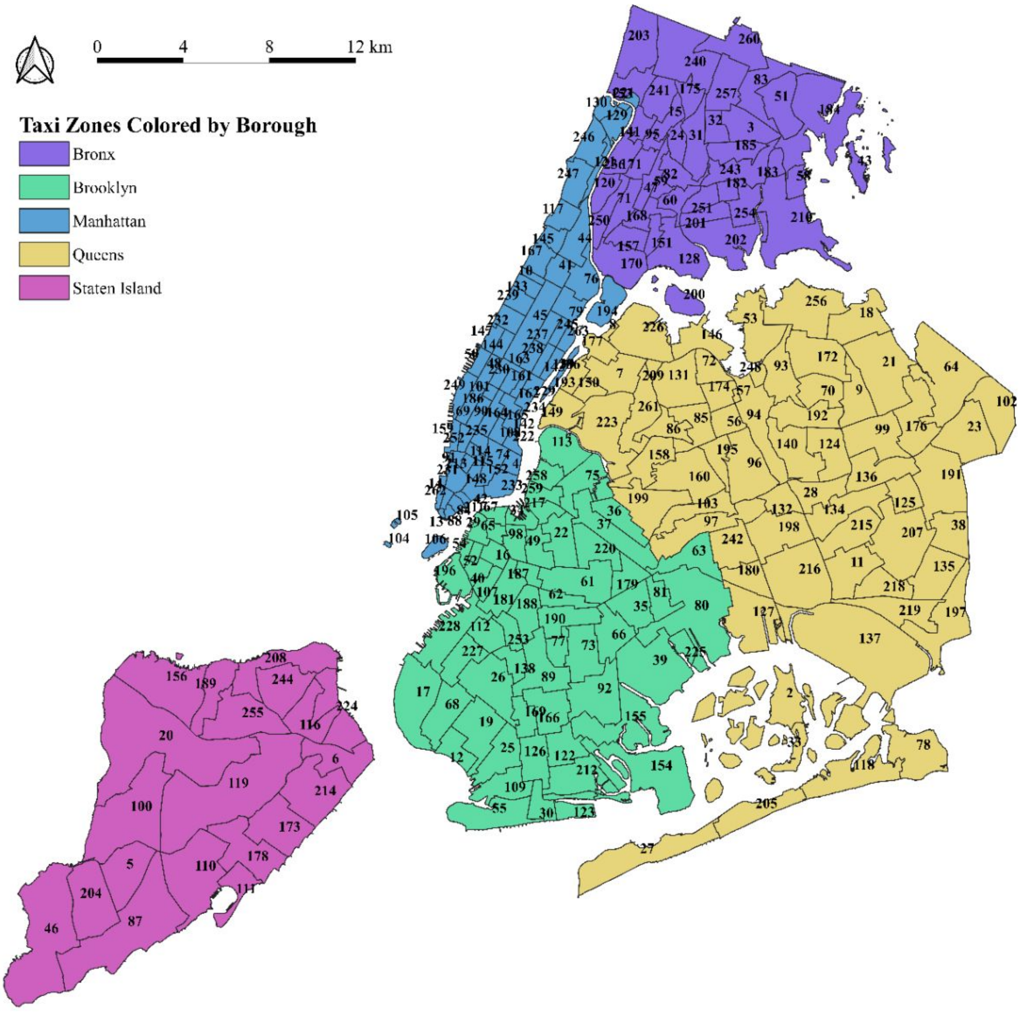


Figure 1. Study area: Taxi zones in the New York City Boroughs

Several open-data initiatives make it easier to access human mobility and transport datasets. Specifically, three types of trip data are collected from the NYC Taxi and Limousine Commission (TLC)[1], namely, yellow taxi, green taxi, and for-hired vehicle (FHV) data. As strong competitors to traditional public transport, shared mobility services have become important alternatives for daily travel in cities. Such datasets are able to capture human mobility on a large scale with a fine resolution and better spatial coverage. Yellow taxis are allowed to pick up passengers in any zone while green taxis are mainly allowed to serve outer boroughs (i.e., districts outside Manhattan). Yellow taxis and green taxis are integrated and treated together as traditional transport in this study. FHV data contain the trips made by taking a shared vehicle (e.g., Uber and Lyft). An FHV is distinguished from traditional taxis in that drivers and passengers are freer to choose each other and choose orientation-destination pairs via online platforms. In addition to shared vehicles, shared bikes are another popular mode for commuting in NYC. We collected bike data from the City Bike data portal[2], which is the largest bike sharing system in the city.

The temporal scheme for data collection and analysis is the monthly snapshot of 6 incremental years to decrease the computational complexity. Specifically, the August data of each year from 2013 to 2018 are collected. We choose these years because the

[1] https://www1.nyc.gov/site/tlc/about/tlc-trip-record-data.page
[2] https://www.citibikenyc.com/system-data

market share of shared mobility services started to increase during the period, which may bring changes to the context that was originally dominated by traditional taxis. The useful attributes of different types of data, including the pick-up and drop-off time and location of each trip, are similar (Tables 1-4).

Few data cleaning and preparation steps are conducted when generating mobility networks. First, data rows with missing values or values in inconsistent formats are deleted. Duplicated data rows are also deleted. Because the format of location information in yellow and green taxis is different from 2016 onwards, trip data with coordinates are required to be spatially joined to be projected to the taxi zone level. This step results in OD trips of different data from zone to zone, and such alignment of the spatial units is important because zones will serve as the same set of nodes in the multiplex transit network.

As this study aims to detect urban structures from the overall multilayer interaction patterns, the variation in hours or weeks is not included. Hence, the number of trips in each month of the year is aggregated at the individual level for any pair of zones, which will determine the interaction strength of the edges in each layer (i.e., intralayer edges). We believe that the multimodal data spanning from 2013 to 2018 are sufficient to cover different aspects of the interaction patterns for investigating urban structures.

Table 1. Sample of yellow taxi dataset

Before 2016					
pickup_datetime	**dropoff_datetime**	**pickup_longitude**	**pickup_latitude**	**dropoff_longitude**	**dropoff_latitude**
2013/8/28 18:03	2013/8/28 18:05	-74.007819	40.724951	-74.006129	40.735067
2013/8/31 0:26	2013/8/31 0:35	-74.00044	40.732387	-74.005396	40.711376
…	…	…	…	…	…
2013/8/29 9:26	2013/8/29 9:29	-73.931406	40.760204	-73.920704	40.756749
From 2016 onwards					
tpep_pickup_datetime	**tpep_dropoff_datetime**	**PULocationID**	**DOLocationID**		
2018/8/1 8:29	2018/8/1 8:35	100	90		
2018/8/1 10:07	2018/8/1 10:17	234	234		
…	…	…	…		
2018/8/1 1:21	2018/8/1 1:23	48	143		

Table 2. Sample of green taxi dataset

Before 2016					
lpep_pickup_datetime	**Lpep_dropoff_datetime**	**Pickup_longitude**	**Pickup_latitude**	**Dropoff_longitude**	**Dropoff_latitude**
2013/8/28 6:02	2013/8/28 6:13	-73.92948914	40.75686264	-73.92989349	40.75658035
2013/8/26 16:56	2013/8/26 17:05	-73.95521545	40.8044014	-73.97678375	40.7918396
…	…	…	…	…	…
2013/8/31 18:34	2013/8/31 18:39	-73.94673157	40.83132553	-73.94012451	40.84090805
From 2016 onwards					
lpep_pickup_datetime	**lpep_dropoff_datetime**	**PULocationID**	**DOLocationID**		
2018/8/3 7:34	2018/8/3 7:43	95	28		
2018/8/3 22:13	2018/8/3 22:17	255	255		
…	…	…	…		
2018/8/2 22:32	2018/8/2 22:39	65	106		

Table 3. Sample of shared bike dataset

starttime	stoptime	start station latitude	start station longitude	end station latitude	end station longitude
2018/8/2 12:52	2018/8/2 13:00	40.73705	-73.99009	40.73650	-73.97809
2018/8/1 18:54	2018/8/1 19:03	40.74916	-73.99160	40.76009	-73.99462
…	…	…	…	…	…
2018/8/2 13:08	2018/8/2 13:16	40.69128	-73.94524	40.70317	-73.94064

Table 4. Sample of FHV dataset

Pickup_DateTime	DropOff_datetime	PUlocationID	DOlocationID
2018/8/23 23:36	2018/8/24 0:01	255	249
2018/8/1 22:33	2018/8/1 22:42	230	90
…	…	…	…
2018/8/2 11:07	2018/8/2 11:31	50	90

4. Methodology

Considering the emergence of shared mobility services, this paper aims to infer the urban structure from a multilayer network perspective, namely, layers defined by temporal changes and multiple transport modes. The number of aspects of the two multilayer networks in this paper is equal to 1. The following sections explain the key definitions and methods for constructing and analyzing the multilayer mobility network for urban structure analysis.

4.1. Yearly change rate

Besides the network analysis, we first calculate a statistical metric, the yearly change rate, to depict the fundamental aspect of how the number of trips of different transport modes changes over time. This metric provides first-order characteristics in each zone, which is helpful in examining the long-term trend of behavioral changes that may be related to the urban structure. For each zone and each transport mode, the yearly change rate is calculated using Equation 1.

$$Yearly\ Change\ Rate\ (YCR) = \left(\left(\frac{E_{Trip}}{B_{Trip}}\right)^{1/Y} - 1 \right) * 100 \quad (1)$$

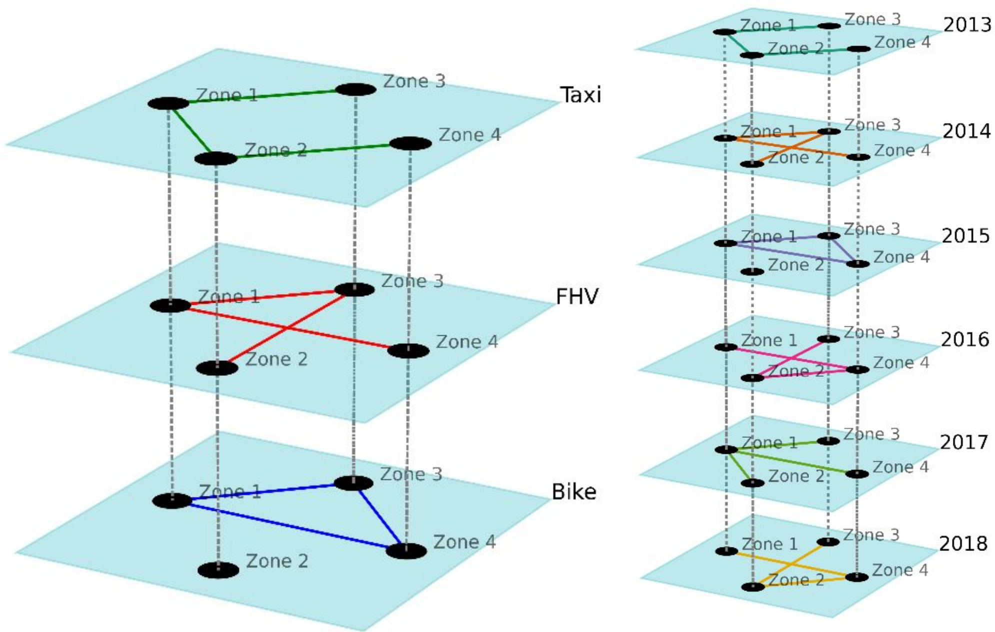
where E_{Trip} is the total number of trips in the last year of a given period, B_{Trip} is the number of trips in the first year of a given period, and Y is the total number of years. For example, given a five-year period of a transport mode in a zone, B_{Trip} is the trip volume of the first year (Year 1), and E_{Trip} is the trip volume of the fifth year (Year 5). It is worth noting that we include the part minus 1 to represent the decrease in trip volume more intuitively by producing a negative value through this equation.

4.2. Generation of transport multiplex network

Network analysis has been widely used in analyzing movement data and understanding zone-to-zone behavior (Zhong et al., 2014; Sarkar et al., 2017; Zhang et al., 2018). The recent advancement in network representation and analysis, the multilayer network, has also proven to be effective at including multiple transport entities or relationships in the same framework (Ding et al., 2018; Yildirimoglu & Kim, 2018). Inspired by these works, this study aims to extend the knowledge on multidimensional travel behavior using a multilayer network framework, which is more accurate for describing complex transport systems. The general form of a monoplex (single-layer) network is $G = (V, E)$, where V is the set of nodes and $E \subseteq V \times V$ is the set of edges connecting each pair of nodes. For the multilayer network, we follow the definition made by Kivelä et al. (2014), i.e., a multilayer network $M = (V_M, E_M, V, L)$. Edges are allowed to exist between any possible pairs of nodes, i.e., $E_M \subseteq V_M \times V_M$. The links between nodes within a layer are called intralayer edges, and the links across layers are called interlayer edges. Layers of d aspects are represented by L, where $L = \{L_1, L_2, \ldots L_d\}$. That is, for each aspect, there can be multiple layers. For each layer, there are nodes belonging to the same set V. For example, an aspect of layers can represent multiple shared mobility modes while at the same time another aspect of layers represents multiple public transport modes.

In this paper, the multilayer network inferred from empirical transport data is known as a multiplex network (Nicosia et al., 2013). The aspect equals 1 in this work, which means that networks integrated in the same framework belong to the same type (i.e., same year or same mode). Both the orientation-destination direction and trip volume are considered in the multiplex network model and in the following analysis. In short, direction is considered in multiplex PageRank and modularity, and the number of trips among locations determines the intralayer link weights. For the interlayer links, we use categorical coupling, which represents the adjacency of a node to itself in other layers, to connect layers (Mucha & Porter, 2010; Kivelä et al., 2014). The most common weight of the interlayer for the multiplex network is used, and the weight is set to 1 (Kivelä et al., 2014).

A multiplex network is a subset of the general multilayer network, and it has been used to study multisource transport data in several literatures (Cardillo et al., 2013; Yildirimoglu & Kim, 2018). Compared to the generalized definition of a multilayer network, a multiplex network reduces to one aspect, i.e., $L = \{L_1\}$. In the network sequence $\{(V_\alpha, E_\alpha)\}_{\alpha=1}^{\beta}$, nodes set in different layers are usually the same, i.e., $V_\alpha = V_\beta$ for all α and β. It is known as an edge-colored graph because it contains the same set of nodes but different sets of intralayer edges in each layer (Kivelä et al., 2014). Multimodal transport modes and temporal slices can be effectively represented by multiplex networks. In general, the following generated networks are directed and weighted by the travel flows between zones. Different layers share the same set of nodes (i.e., NYC taxi zones) while intralayer links represent different types of human mobility interactions between locations.



(a) (b)

Figure 2. Thematic representation of the multiplex networks in this study: (a) Multimodal; (b) Temporal. Note that each layer is a directed graph determined by the OD direction, and intralayer links are weighted by the OD volume. The common interlayer link weights equal 1 in this study. The multiplex network can effectively integrate the occurrence of multi-flows among the same set of locations for analysis.

To investigate the impact of on-demand shared transportation, the first multiplex network is constructed based on multimodal transit: $M_{Mode} = (V_m, E_m, L_m)$, where $L_m = \{Taxi, FHV, Bike\}$ and $E_m \subseteq V_m \times V_m$. Time is fixed in 2018, and different modes (i.e., taxis, bikes, and FHVs) are regarded as layers. The transit interactions among the same set of zones are therefore considered together in the same framework (Figure 2a). Based on M_{Mode}, the experiment focuses on revealing different travel behaviors related to both traditional transport (i.e., taxis) and newly emerged modes (i.e., FHVs and shared bikes). In contrast, the second generated multiplex network only involves traditional transport and taxis but focuses on the possible change in travel behavior over time: $M_{Time} = (V_t, E_t, L_t)$, where $L_t = \{2013, 2014, \ldots 2018\}$, and $E_t \subseteq V_t \times V_t$. From 2013 to 2018, the taxi transit network in each year serves as a layer (Figure 2b). In this period, the context of NYC transport gradually changed due to the emergence of shared bikes and shared vehicles. As the context changes, the temporal multiplex network of taxis is useful to observe whether there is a longitudinal change in travel behavior.

4.3. Centrality metrics

Centrality metrics have been widely used to evaluate the importance of nodes. Centrality can be regarded as a basic characteristic of a network structure as it indicates the levels of heterogeneity of node properties. In a human mobility network, nodes with higher centrality may indicate a transportation hub that bears more daily transit and activities, which also help to describe the urban structure. In this paper, the degree centrality and PageRank centrality are calculated. The degree centrality is the basic metric, and PageRank has been used in transport and urban networks to measure the attractiveness of a location (Ding et al., 2009; Agryzkov et al., 2016; Jia et al., 2019). Although PageRank centrality can be calculated in various domains such as informatics (Page et al., 1999), biology (Yu et al., 2017), and human mobility and transport studies (Wen, 2015; Xu et al., 2017; Zhou & Qiu, 2018), the network context is still a single layer. The extension of centrality measures from monoplex networks to multiplex networks is still in its infancy (Battiston et al., 2014; Halu et al., 2013). This study depends on the implementation of multiplex centrality measures of MuxViz (De Domenico et al., 2015a), specifically multilayer degree centrality (De Domenico et al., 2013) and multiplex PageRank centrality (De Domenico et al., 2015b).

PageRank is one of most popular algorithms to rank node importance in graphs and was proposed by one of the cofounders of Google (Page et al., 1999). PageRank measures a node's (e.g., website's) importance based on its outbound links. In the urban transport context, this metric reflects how a location attracts outbound interactions from other locations, which can be extracted from massive transit data. Therefore, the multiplex PageRank extends the capability of this metric in a multilayer context. Specifically, in this study, this metric indicates the attractiveness of a location considering multimodal flows.

For a node i, it can be calculated as the summation of degree k_i^{α} of each layer, which is only suitable for a multilayer network without interlayer links. In this paper, interlayer links are assumed to exist, that is, the zones on each layer are connected to their counterparts on other layers. Therefore, another improved definition of degree centrality considering the presence of interlayer links is used in our work:

$$k_i = M^{i\alpha}_{j\beta} U^{\beta}_{\alpha} u^{j} \quad (2)$$

where k_i is the aggregated multilayer degree centrality of node i, $M^{i\alpha}_{j\beta}$ is the adjacency matrix containing the relationship between node i on layer α and node j on layer β, L is the total number of layers, N is the total number of unique nodes, u^j is a first-order tensor in which all elements equal 1, and $U^{\beta}_{\alpha} = u_{\alpha} u^{\beta}$ is a second-order tensor in which all elements equal 1.

Early work on generalizing the PageRank centrality to a multilayer network was performed by Halu et al. (2013). However, their metric is mainly feasible for a two-layer empirical network due to the complex layer dependence. Here, we rely on the multiplex PageRank proposed by (De Domenico et al., 2015b). The key idea of PageRank is to explore the network using the random walk equation, which produces a transition matrix that defines 'walk behavior'. In a multiplex network, the PageRank centrality of node *i* is defined as:

$$\omega_i = \Omega_{i\alpha} u^{\alpha} = \sum_{\alpha=1}^{L} \Omega_{i\alpha} \quad (3)$$

where ω_i is the aggregated PageRank centrality of node *i*, and $\Omega_{i\alpha}$ is the eigenvector of tensor $R^{i\alpha}_{j\beta}$.

$$R^{i\alpha}_{j\beta} = \tau T^{i\alpha}_{j\beta} + \frac{(1-\tau)}{NL} u^{i\alpha}_{j\beta} \quad (4)$$

where τ is the walking rate that is normally set to a constant value (e.g., 0.85), $T^{i\alpha}_{j\beta}$ is the transition tensor containing the jumping probabilities between pairs of nodes in any layer, N is the total number of unique nodes, L is the total number of layers, and $u^{i\alpha}_{j\beta}$ is a 4^{th}-order tensor in which all elements equal to 1.

By adopting these two centrality measures on a multimodal transit network and temporal taxi transit network, the importance of taxi zones will be evaluated in both layer-by-layer and aggregated manners.

4.4. Community detection in multilayer network

In single-layer network analysis, community detection has been proven to be efficient in characterizing travel preferences from mobility networks (Zhong et al., 2014; Liu et al., 2015). Given a network, the community detection process partitions and groups nodes in a manner that maximizes the intergroup distance and minimizes the intragroup entity distance. In a mobility network generated by urban flows, a community is a cluster of locations with similar interaction (e.g., in/out flows) patterns. Given a multimodal network, communities across nodes and layers can be understood as the variation of interaction patterns across places and different transport modes. The interpretation of the community in the temporal multiplex network is similar, indicating variations in interaction patterns across locations and different times. After projecting the network community to geographical space, the interaction-associated urban structure and its dynamics can be examined.

In practice, the most commonly used metric to be maximized is the modularity (Newman & Girvan, 2004). Despite abundant detection algorithms for monoplex networks, very few algorithms have been developed in multilayer network frameworks. Instead of extracting a community layer-by-layer, the multilayer network community detection algorithm detects the community simultaneously across layers. In this paper, we utilize the most used multiplex-Infomap algorithm (De Domenico et al., 2015a), which relies on the refined modularity proposed by Mucha et al (2010).

$$Q_{multilayer} = \frac{1}{2\mu}\sum_{ijsr}\left[\left(A_{ijs} - \gamma_s \frac{k_{is}k_{js}}{2m_s}\right)\delta_{sr} + \delta_{ij}\omega\right]\delta(g_{is},g_{js}) \tag{5}$$

where A_{ijs} is the intralayer edge weight between node i and node j at layer s; the k of a node is calculated by the sum of the weights of edges attached to this node; k_{is} represents the total strength (i.e., weighted by trips) for node i at layer s; k_{js} represents the total strength of node j at layer s; $k_{is} = \sum_j A_{ijs}$; $\mu = \frac{1}{2}\sum_{jr} k_{jr}$; $m_s = \frac{1}{2}\sum_{ij} A_{ijs}$; δ is the Kronecker delta function, which equals 1 if two variables are the same and 0 otherwise; g_{is} is the community label assigned to node i in layer s; γ_s is a resolution parameter set to 1 by default; and ω is the interlayer coupling weight from 0 to 1, which equals 1 in this study.

Most community detection algorithms in networks rely on the concept of modularity to compress data and find regularities (Grünwald & Grunwald, 2007). The optimization target is to find the partition structure that minimizes the communication length. Combining the refined modularity with the classic Infomap algorithm, multiplex-Infomap can search the community in a multiplex network. In contrast to single-layer community detection, multiplex-Infomap results in community labels that are comparable across different layers. For example, zones with the same community label in different layers (i.e., year or transport mode) indicate that their interaction travel behaviors are similar, indicating similar urban structures. Therefore, this technique is efficient in revealing the structure for the first experiment, which investigates the possible variation between traditional transport (i.e., taxis) and shared transportation (i.e., bikes and FHVs); and for the second experiment, which explores the variation in traditional taxis from 2013 to 2018.

5. Results

5.1. Overall trend of multimodal transit in NYC

In this section, we examine the basic characteristics of trips by taxis, bikes, and FHVs. First, the total number of trips in each year of different transport modes is plotted in Figure 3. A clear trend can be observed in terms of the variation over time. Particularly, significant variations are found in yellow taxis and FHVs, showing that the number of trips by yellow taxis has been decreasing since 2013 while that of FHVs is dramatically increasing. The trip volumes of green taxis and shared bikes are relatively low, and their variations are weak. Both green taxis and shared bikes are included due to their unique role in serving specific travel purposes. Green taxis were launched to supplement taxi services in the outer boroughs, and shared bikes are especially popular in Manhattan and its surrounding areas. Based on the yearly numbers of trips, a preference shift from traditional taxis to shared transportation is clearly witnessed.

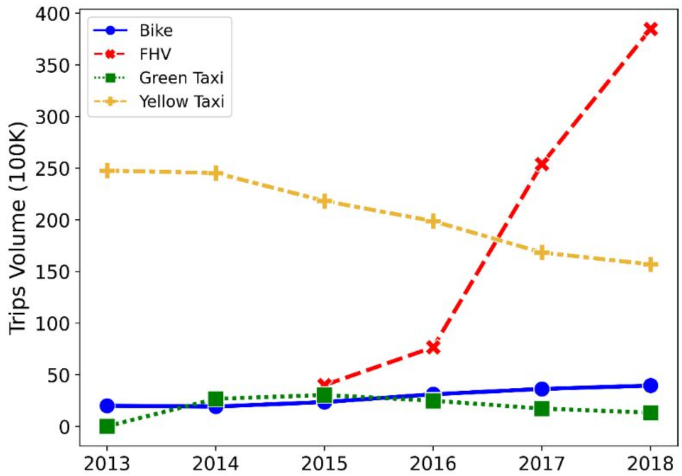


Figure 3. Trip variation of three transport modes over the years

We spatially join the OD trips of all transport modes to the zone level and visualize them on a map (Figure 4), which illustrates how taxis, bikes, and FHVs are differently used in space. Taxi and FHV services cover almost all NYC zones while shared bikes are mainly available in Manhattan and near-Manhattan zones in Brooklyn, Queens, and Bronx. The profound difference between taxis and FHVs is in the distribution of medium (i.e., line in blue) and high volume (i.e., line in red) trips. High trip volumes for taxis are mainly constrained within Manhattan (Figure 4a); in contrast, high trip volumes for FHVs are spread more widely in both Manhattan and the outer zones (Figure 4c). It seems that FHVs not only strengthen the commutes from outer zones to Manhattan but also makes the connections among outer zones stronger.

Taxi drivers may spend less time searching for customers in a more populated dense area (i.e., Manhattan) while the demand-match mechanism in FHVs makes FHVs more flexible to serve more areas. Based on Figure 4, it is obvious that FHVs play a more important role in supplementing the unbalanced supply in distant zones. Shared bikes presents another different spatial pattern, showing that heavy use is mostly in downtown Manhattan and the Brooklyn zones across the river. Additionally, a decent number of bike trips with long travel distances are observed, indicating that shared bikes might be a popular choice for commuting in these areas.

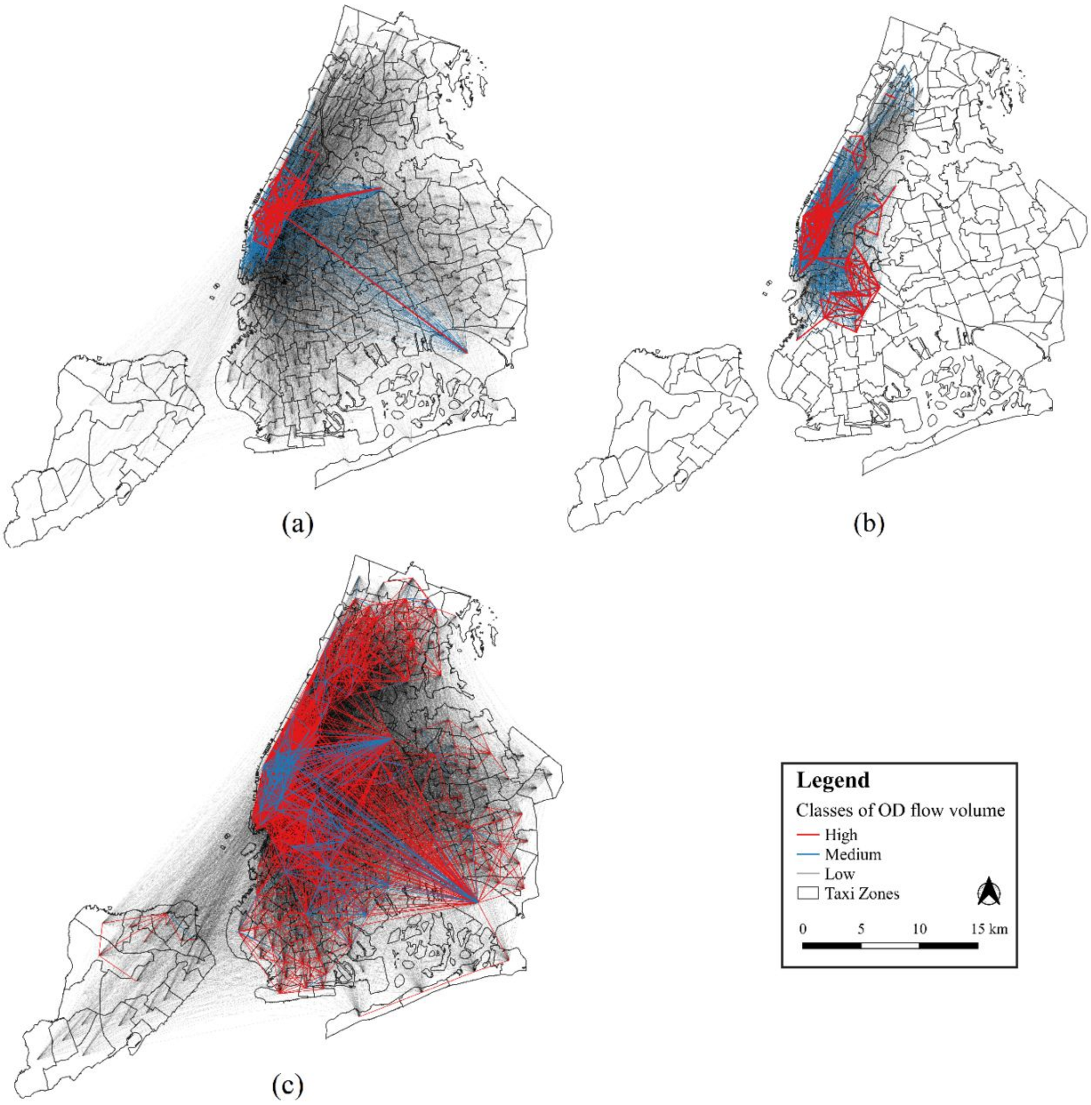
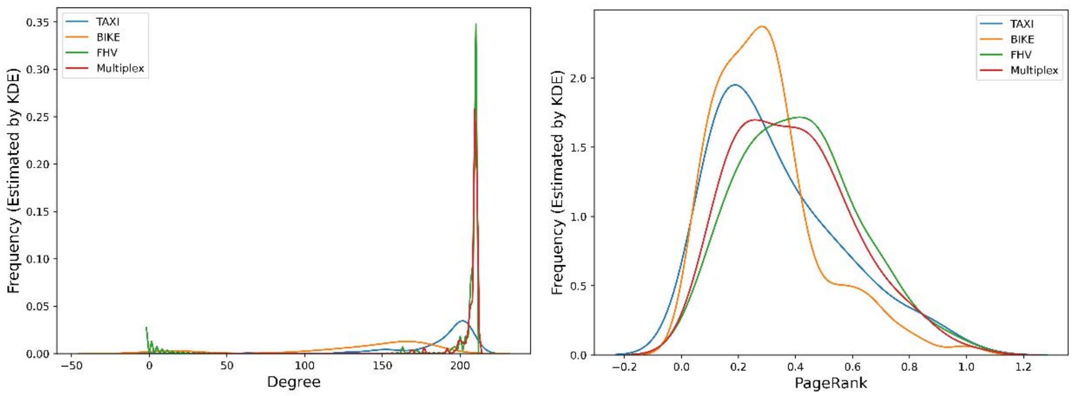


Figure 4. Distribution and classification of OD trips: (a) Taxi; (b) Bike; (c) FHVs

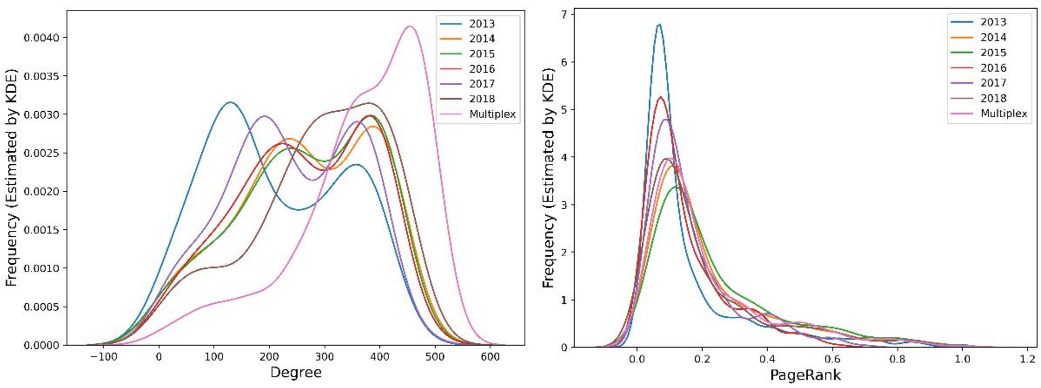
5.2. Node centrality in multimodal and temporal mobility networks

We first explored the basic law of travel behavior in the multilayer mobility network. Previous researchers observed a heavy-tailed distribution (e.g., power-law or exponential distribution) of displacement and elapsed time (Gonzalez et al., 2008; Liang et al., 2012; Zhao et al., 2015). Using network analysis, Zhong et al. (2014) also reported the heavy-tailed distribution of nodes (i.e., zones) centrality in different modes of the mobility network. However, the node centrality distribution in a multilayer mobility network is limited.

We calculate two centrality metrics, the multiplex degree and the multiplex PageRank, for both the multimodal network and temporal network. Then, we infer the empirical univariate distribution by using the Python package *distfit*, which fits 89 models and ranks the models based on the residual sum of squares (RSS). Overall, node centrality in these two multiplex networks does not follow a power-law or exponential distribution. Instead, the *beta* distribution is identified as the best model for the two centralities in a multimodal network based on the smallest RSS. In multimodal network, most zones (i.e., nodes) have a strong degree centrality and medium PageRank centrality (Figure 5a). The intralayer's centrality (i.e., taxis, bikes, and FHVs) and aggregated centrality (i.e., multiplex) have relatively similar distributions. A heavy left tail is found in the degree distribution while a slight right tail is found in the PageRank distribution. In temporal network, most zones (i.e., nodes) have a strong degree centrality and low PageRank centrality (Figures 5c & 5d). There is significant variance among the layers (i.e., 2013 to 2018) in terms of the degree centrality. More zones with higher degrees appear chronologically. The PageRank distributions, which all have slightly long right tails, are similar among layers.



(a) (b)



(c) (d)

Figure 5. KDE estimated distribution of node centralities: (a) degree in the multimodal network; (b) PageRank in the multimodal network; (c) degree in the temporal network; and (d) PageRank in the temporal network

Centrality metrics indicate the importance and vibrancy of a location, which may reflect the underlying urban structure that breeds such activities (Jia et al., 2019). We project the centralities of zones into geographical space (Figure 6), which helps obtain a better understanding of the polycentric structure of NYC (Zhong et al., 2014). The spatial distribution of multimodal network centralities is different when using the degree and PageRank (Figures 6a & 6b). Extreme high degree centralities are mainly found in the downtown area of Manhattan while almost all of Manhattan is identified as having high PageRank centralities. The possible reason for the heavily left skew of the degree is that the Manhattan zones are quite connected when considering multiple types of transport modes together (i.e., taxis, bikes, and FHVs). The degree might not be the best choice to describe the node importance at such a scale while PageRank is able to capture more variance in this well-connected network. Comparatively, the spatial distribution of the degree and PageRank centralities of the temporal network show similar patterns (Figures 6c & 6d). Manhattan is the most 'important' district in terms of the node centralities; moreover, some zones in Brooklyn and Queens show high values. The result of the multiplex centralities in New York City provides evidence on the polycentric urban structure reflecting the travel demand.

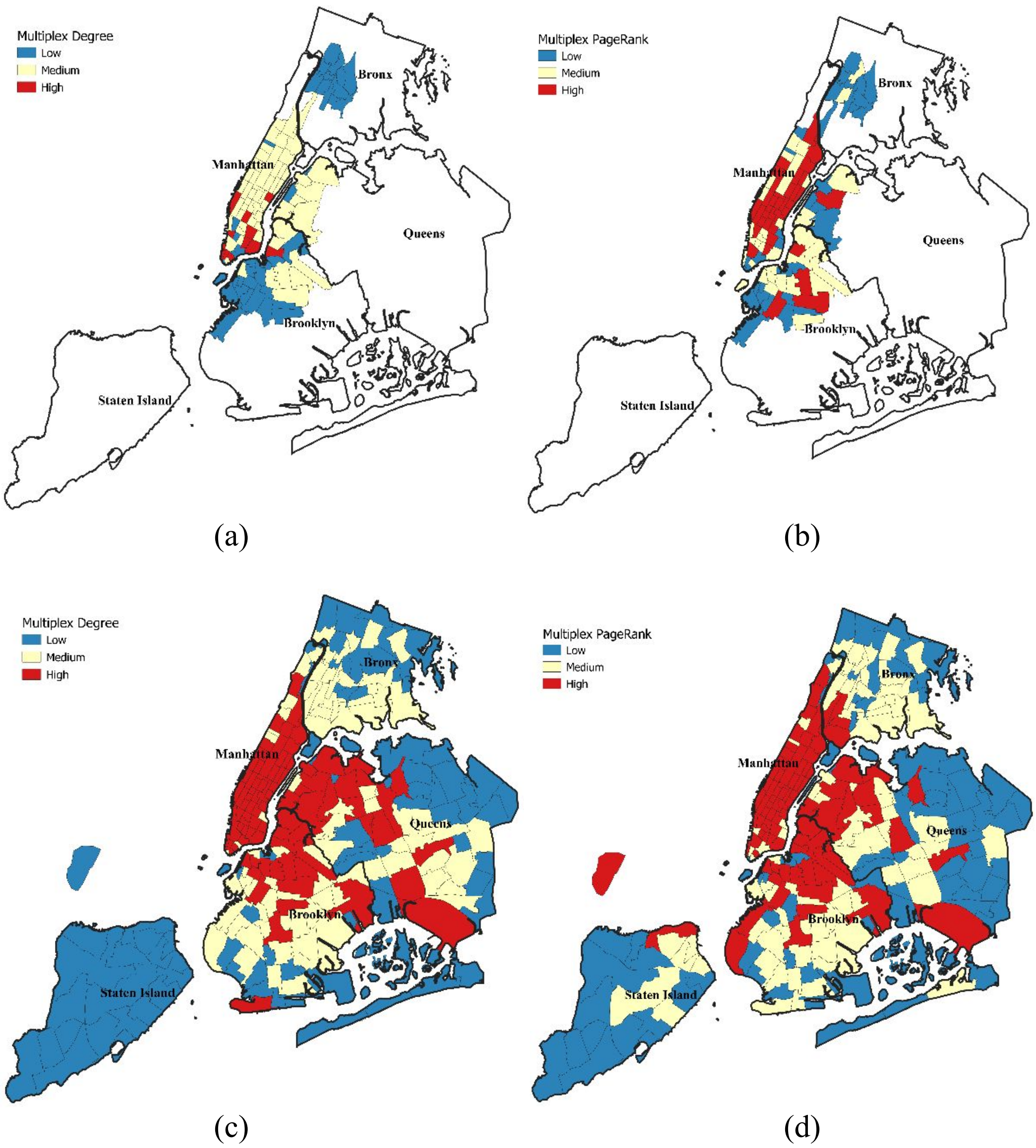
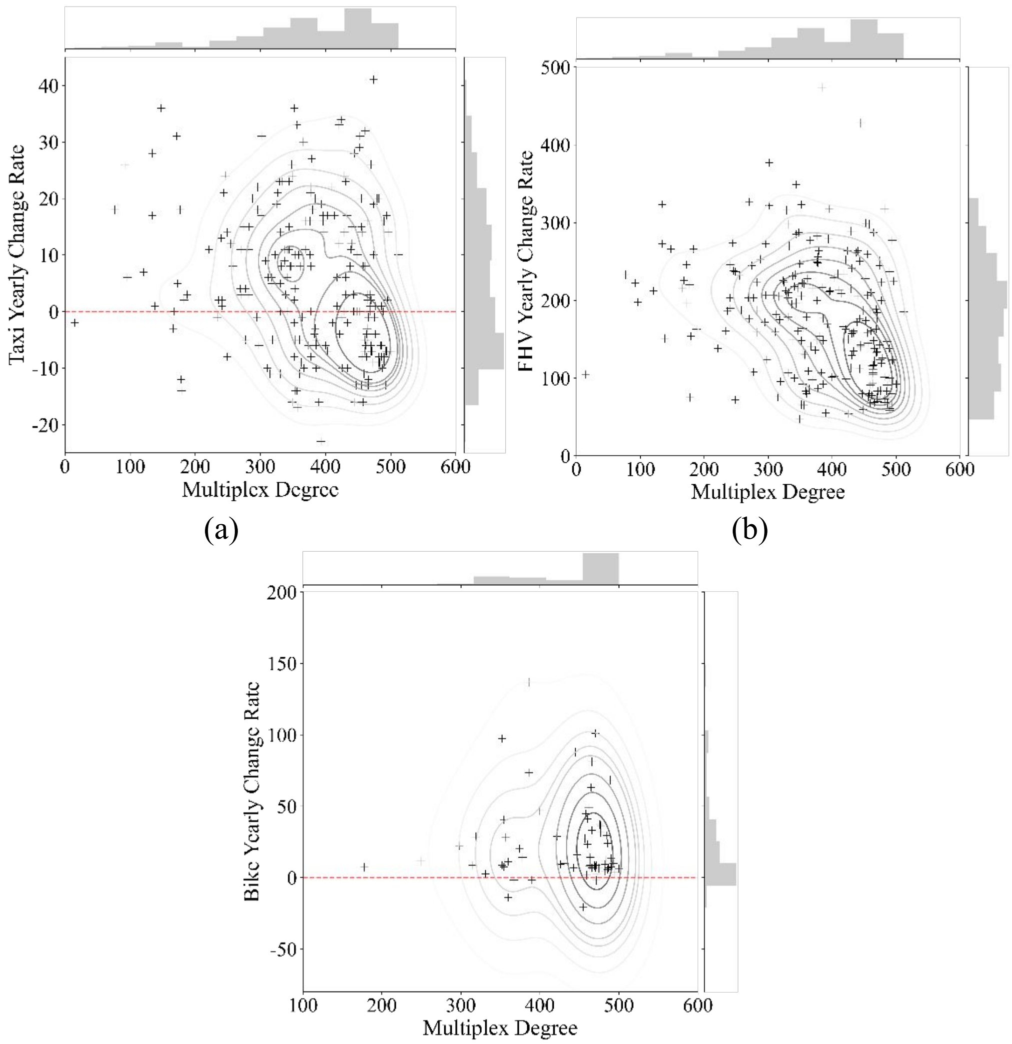


Figure 6. Visualization of the node centralities in geographical space: (a) Degree in the multimodal network; (b) PageRank in the multimodal network; (c) Degree in the temporal network; and (d) PageRank in the temporal network.

The relationship between node (i.e., zone) centrality and trip variation is explored, indicating that the zone importance in the multilayer mobility network is a key indicator correlated to the yearly change rate. Specifically, the node centrality in the temporal network of traditional taxis is selected for comparison for two reasons. First, the multiplex network constructed from taxi data covers the entire city, which provides a more heterogeneous urban context for analyzing the correlation. Second, nodes with high multiplex centrality represent truly important zones for vehicle-based transit because the temporal network takes 6 years of transit patterns in the same framework for evaluation.

In each zone, the yearly change rate is calculated for taxis, FHVs, and bikes, respectively. In Figure 7, each point represents a zone, the X axis represents the multiplex degree, and the Y axis represents the yearly change rate of the specific transport mode. Contour lines are added by the KDE function to indicate the concentration patterns, and the dotted horizontal line divides the Y axis into positive or negative yearly change rates. Overall, both negative and positive change rates of traditional taxis are observed while the change rates of FHVs and shared bikes are almost all positive. Similar distribution patterns are observed for taxis and FHVs, which indicates that the trip variations of these two vehicle-based transportation systems have similar relations to zone importance. Zones with multiplex degrees of approximately 350 and 450 are dense in Figure 7a and Figure 7b. Yearly change rate of taxi and FHV are often lower in higher degree zones (i.e., 450). For traditional taxis, the yearly change rate is even negative in such high degree zones, which means that the number of trips is decreasing. The number of valid points for shared bikes is less than those for the other two modes because the bike data are only available in Manhattan and its near zones (Figure 7c). The 450-degree zones play an important role in shared bikes. The node centrality of the multiplex mobility network provides a useful indication of location importance, which is highly correlated to the variation of different transport modes.



(a) (b)

(c)

Figure 7. Relations between yearly change rate and multiplex node degree: (a) Taxi; (b) FHV; (C) Bikes

5.3. Revealing the variation in the urban structure based on multiplex community detection

We adopt the multiplex-Infomap algorithm to conduct community detection for multimodal temporal and multiplex networks. Zones with similar interaction patterns are grouped into a community and assigned a unique label, which is used to evaluate the travel behavior of nodes/zones of multimodal transport modes over time or modes. The variation or consistency of community labels directly indicates the variation of the overall urban structure.

The first experiment is conducted on the multimodal network, a directed and weighted zone-based interaction network of different modes in 2018 on three layers: {Taxi2018, FHV2018, Bike2018}. This experiment examines whether newly emerged shared transportation has zone-to-zone interactions similar to those of traditional taxis. In total, 5 communities were identified in the multimodal network, and labels (0-4) were given to the nodes on all layers.

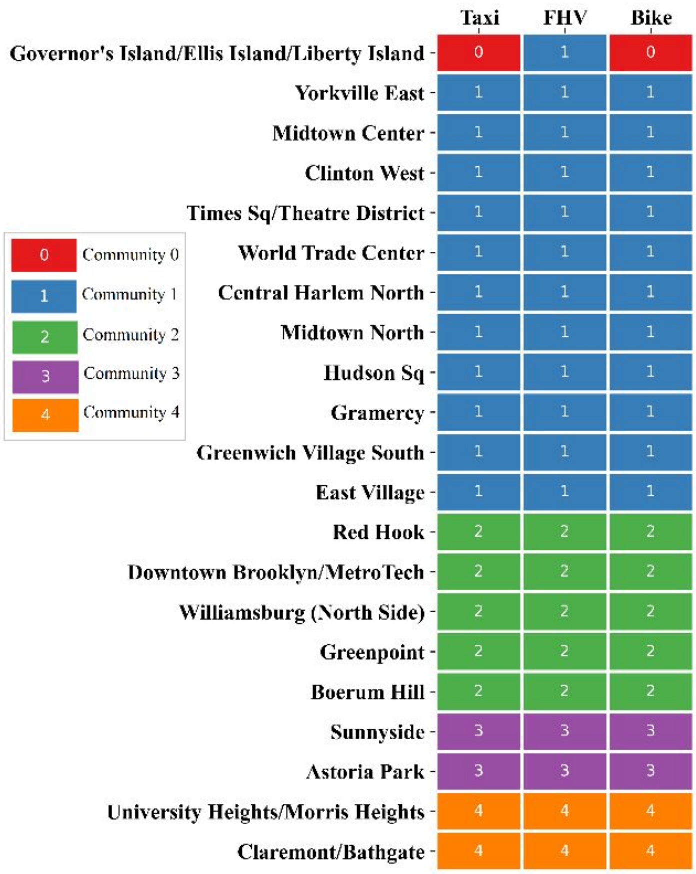


Figure 8. Community labels of zone across transport modes. Note that sample zones are displayed in this figure due to space limitations. There is actually 1 zone with varied community labels, 57 zones with all-1 community labels across modes, 27 zones with all-2 community labels, 8 zones with all-3 community labels, and 13 zones with all-4 community labels.

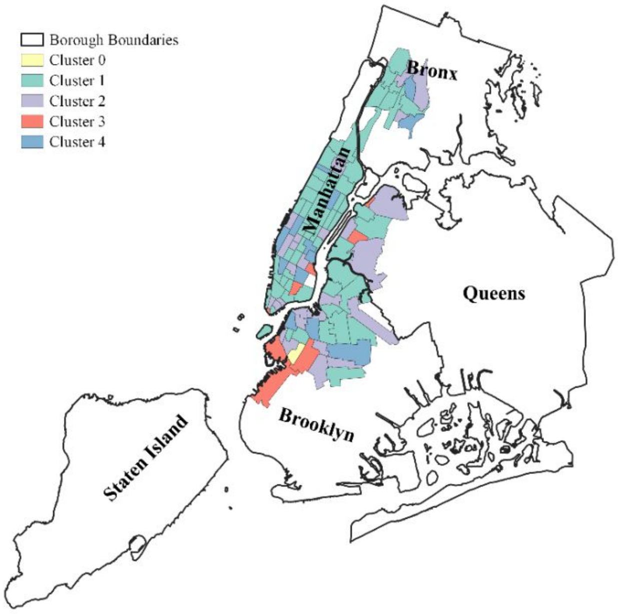


Figure 9. Spatial distribution of the communities in the multimodal network

A community detected in a mobility network indicates a set of nodes (i.e., zones) with similar interaction patterns. However, different from single-layer community detection, the obtained community labels in the multiplex network are comparable. For example, label 0 in the taxi layer and label 0 in the FHV layer indicate the same community role. Using this method, we are able to evaluate the similarity of the interaction behavior across transportation modes (i.e., layers). The community distribution of the multimodal network is shown in a matrix plot (Figure 8), where the Y axis represents zones (i.e., nodes) and X axis represent transport modes (i.e., taxi, FHV and bike layers). The spatial structure of the modal-network community is shown in Figure 9. The spatial clustering of cluster 1 is clearly observed in the central zones while clusters 2 and 3 are mostly located in distant areas.

Sampled zones are visualized in Figure 8 due to limited space, and the actual number of different types of zones is illustrated in the title note. The zone type is identified according to the combinations of community labels across layers. In Figure 8, the distribution of interaction behavior in different modes can be examined. Specifically, the three numbers in each horizontal line represent the community labels of a zone in each of the three layers (taxis, FHVs and bikes). If the three numbers are the same, it means the interaction patterns of the three transport modes are similar in this node (i.e., zone). Taking the second zone as an example, the community labels of the second zone (Yorkville East) are the same across different modes, which means that the interaction patterns among zones are similar regardless of whether taxis, FHVs, or bikes are chosen. In the first zone, *Governor's Island*, the community labels are {0, 1, 0}. This means that the interaction behaviors are similar in terms of taking taxis and bikes while different interaction patterns are shown for FHVs. Interestingly, the community labels of most zones are consistent among taxis, FHVs, and bikes, excluding Gowanus (Figure 9). This indicates that the new transportation modes retain the same travel patterns as traditional taxis; specifically, the interaction patterns in such a multimodal network are the same.

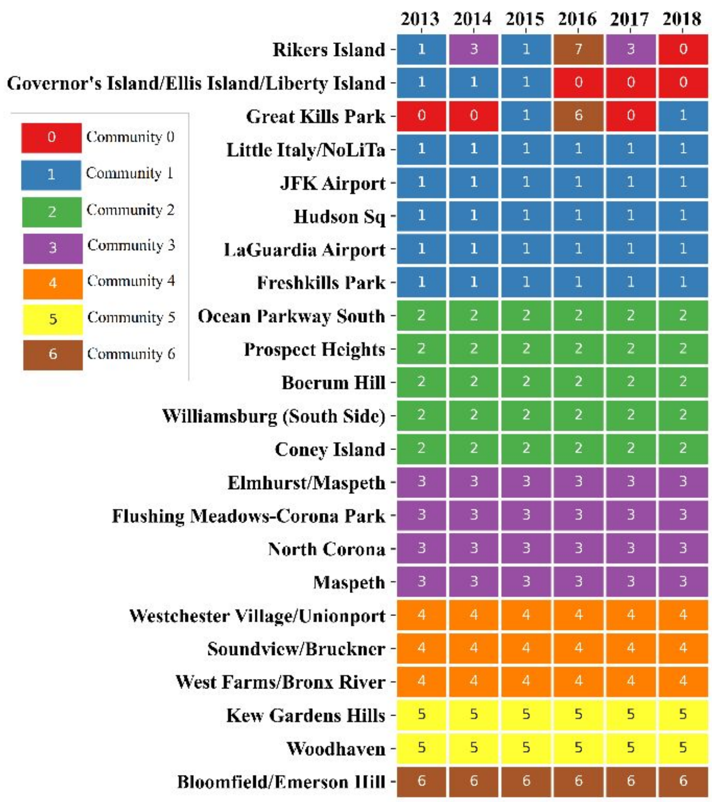


Figure 10. Community labels of zones across years. Note that sample zones are displayed in this figure due to space limitations. There are actually 3 zones with varied community labels, 65 zones with all-1 community labels, 62 zones with all-2 community labels, 29 zones with all-3 community labels, 46 zones with all-4 community labels, 37 zones with all-5 community labels, and 18 zones with all-6 community labels.

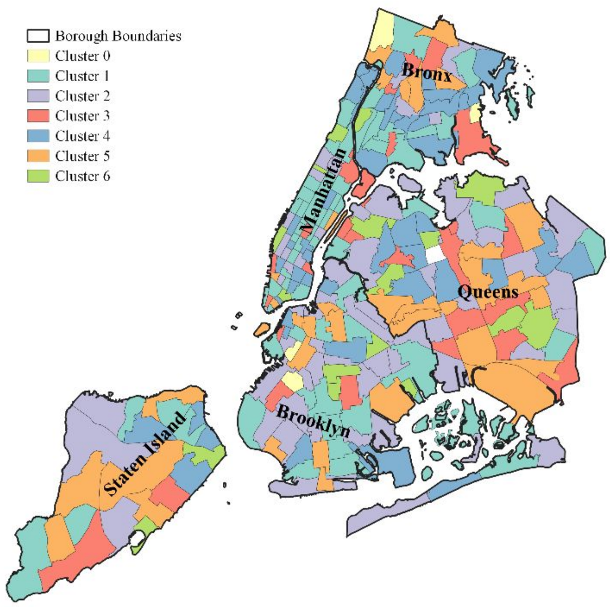


Figure 11. Spatial distribution of the communities in the taxi temporal network

Multiplex community detection is also conducted in the temporal multiplex network of traditional taxis. This experiment investigates whether the interaction patterns of traditional taxis vary from 2013 to 2018, the period when shared transportation was greatly expanding in the market. In Figure 10, we demonstrate the six communities identified across six years/layers, and the associated spatial structure is shown in Figure 11.

The community labels of traditional taxis are consistent across years in all zones except Country Club, Riverdale, Green-Wood Cemetery, and Gowanus. This result means that the interaction patterns of most zones remain stable in terms of traveling using traditional taxis. The result is similar to the findings in the first experiment on the multimodal network. That is, although there was dramatic variation in the market share among traditional taxis and shared transportation during these years, the interaction behavior at the zone level did not significantly change.

6. Conclusion

On-demand shared transportation (i.e., shared vehicles and shared bikes) has dramatically increased in recent years. In view of the advances in multilayer network analysis, this paper constructs empirical multiplex network models to explore how city zones are different from each other due to multidimensional urban flows. Specifically, this paper investigates travel patterns and associated urban structure in several ways.

Two centrality metrics, i.e., the multiplex degree and multiplex PageRank, are calculated in the multimodal network and temporal network. The centrality values attached to the spatial units reveal the hierarchical structure of location attractiveness. In a multimodal network, we found distinct differences between the uptown and downtown of Manhattan, which is also reported in the literature (Zhou et al., 2019). What is more interesting is the centrality spatial distribution in the temporal network across years, which shows that nearly the entirety of Manhattan and near-Manhattan zones all have the same high interaction flows. The possible reason is the differences in network layers defined in the multimodal network and temporal network. The results suggest that the selected layer significantly describes the dimension of variation, that is, the relatively large variation of the magnitude of the flow across transport modes while the relatively small variation of the flow across years. Both are valuable as the lens of multiflow urban structure while providing two different perspectives. The statistical distribution of network centrality is contrary previous studies that reported heavy-tailed human mobility patterns (Gonzalez et al., 2008; Liang et al., 2012; Zhao et al., 2015). In contrast, we observed a left skew of the degree and a slight right skew of PageRank. This pattern is particularly strong in the multimodal network (Figures 5a & 5b), which indicates that multimodal transit options make zones more connected to each other, resulting in a more balanced distribution of transit. Shared mobility may complement traditional taxis in distant areas but substitute in central areas (Kong et al., 2020).

Compared to the layer-by-layer analysis, the findings in community detection in multilayer analysis enable direct comparison in this study. Although the market share of traditional taxis has been greatly taken over by FHVs and shared bikes, the identified interaction behavior shows that most zones are consistent across years and across modes. This suggests that shared transportation, a strong competitor as a travel mode, does not change collective travel behavior from zone to zone and may instead be affected by socioeconomic factors. For example, the interaction between a residential zone and a working zone remains the same regardless of the transport mode the traveler takes. The consistency of community patterns across transport modes in NYC agrees with another study using agent-based simulation (Lokhandwala & Cai, 2018). In this literature, they quantify the traffic conditions considering both traditional and shared mobility services, suggesting that although shared mobility reduces traditional mode ridership, the overall service level remains the same. In this study, the consistency of a multimodal community can be understood as the stability of the overall interaction among places, although the ridership of each change.

Against initial expectation interaction variations of traditional taxis, community detection in taxi temporal networks shows consistent community labels across years. Although we select a period when shared mobility dramatically increases its market share, the stability of long-term interaction among places by taking taxis is observed in this study. A similar result is also reported in Riascos & Mateos (2020) using other network metrics. Our results may suggest that the long-term human mobility of taxis is generalizable in other cities, which may support the view that long-term taxi data are suitable for measuring the nature of human mobility. Another implication of a consistent community in a temporal network is that environmental factors may play a more influential role in changing taxi travel behavior rather than the emergence of shared mobility. As reported in Zhang et al. (2020), residential and commercial land uses have a significant impact on taxi ridership across many locations.

In this study, the relationship between the interaction pattern and land use was not systematically explored. However, we found some similar indications using network metrics. The results in Figure 7 provide similar indications that the factors of the urban context may determine the change in preferences for choosing modes instead of the emergence of shared transportation. In zones with 350 degrees, both taxis and FHVs grow in trip volume, although FHVs increase more rapidly. However, in 450-degree zones, the use of taxis decreases over time. The high degree zones indicate busy places such as Manhattan, and the drop in traditional taxis may be due to the feasibility and convenience of taking bikes or FHVs. These zones have a high network degree, which suggests that traffic jams may occur due to their high importance in a mobility network. In this context, shared bikes may even have greater roles in commuting. The results suggest rising travel demand and indicate that traditional taxis and shared vehicles do not have to be ‘competitors’ but can serve together to make distant transit more sufficient and diverse. We conclude that shared transportation influences travel choices in terms of ridership numbers due to its convenience in some areas, but it does not change the collective interaction patterns among zones compared to other modes.

There are some limitations of this work. Public transport flows are not considered when exploring the travel behavior and urban structure in this study. There are two major reasons. First, from our point of view, public transport is a fixed-route system that uses buses, metros, light rails, and other vehicles, which cannot reflect the on-demand mobility that uses mixed operating systems (e.g., offline and online). The on-demand mobility patterns would provide a special perspective to investigate the urban structure. Second, public transport data in the study area are currently unavailable. It would be great to use mixed datasets to explore the urban structure in future studies. However, our work will provide another empirical angle to understand urban dynamics. In addition, how these multiplex communities (e.g., urban structure) are associated with socioeconomic factors is also interesting but is out of the scope of this paper. The quantitative relation between these two is worthy of investigation in future studies.

References

Agryzkov, T., Tortosa, L., & Vicent, J. F. (2016). New highlights and a new centrality measure based on the Adapted PageRank Algorithm for urban networks. *Applied Mathematics and Computation, 291*, 14-29.

Aleta, A., Meloni, S., & Moreno, Y. (2017). A multilayer perspective for the analysis of urban transportation systems. *Scientific reports, 7*, 44359.
Alexander, L. P., & González, M. C. (2015). Assessing the impact of real-time ridesharing on urban traffic using mobile phone data. *Proc. UrbComp*, 1-9.
Barabási, A. L. (2005). Taming complexity. *Nature physics*, *1*(2), 68-70.
Barthélemy, M. (2011). Spatial networks. *Physics Reports*, *499*(1-3), 1-101.
Battiston, F., Nicosia, V., & Latora, V. (2017). The new challenges of multiplex networks: Measures and models. *The European Physical Journal Special Topics*, *226*(3), 401-416.
Batty, M. (2013). *The new science of cities*. MIT press.
Campbell, K. B., & Brakewood, C. (2017). Sharing riders: How bikesharing impacts bus ridership in New York City. *Transportation Research Part A: Policy and Practice*, *100*, 264-282.
Cannon, S., & Summers, L. H. (2014). How Uber and the sharing economy can win over regulators. *Harvard business review, 13*(10), 24-28.
Cao, M., Ma, S., Huang, M., Lü, G., & Chen, M. (2019). Effects of free-floating shared bicycles on urban public transportation. *ISPRS International Journal of Geo-Information*, *8*(8), 323.
Cardillo, A., Zanin, M., Gómez-Gardenes, J., Romance, M., del Amo, A. J. G., & Boccaletti, S. (2013). Modeling the multi-layer nature of the European Air Transport Network: Resilience and passengers re-scheduling under random failures. *The European Physical Journal Special Topics*, *215*(1), 23-33.
Chen, B. Y., Yuan, H., Li, Q., Shaw, S. L., Lam, W. H., & Chen, X. (2016). Spatiotemporal data model for network time geographic analysis in the era of big data. *International Journal of Geographical Information Science, 30*(6), 1041-1071.
De Domenico, M., Porter, M. A., & Arenas, A. (2015a). MuxViz: a tool for multilayer analysis and visualization of networks. *Journal of Complex Networks, 3*(2), 159-176.
De Domenico, M., Solé-Ribalta, A., Cozzo, E., Kivelä, M., Moreno, Y., Porter, M. A., Gómez, S. & Arenas, A. (2013). Mathematical formulation of multilayer networks. *Physical Review X, 3*(4), 041022.
De Domenico, M., Solé-Ribalta, A., Omodei, E., Gómez, S., & Arenas, A. (2015b). Ranking in interconnected multilayer networks reveals versatile nodes. *Nature communications, 6*(1), 1-6.
Dimitrov, S. D., & Ceder, A. A. (2016). A method of examining the structure and topological properties of public-transport networks. *Physica A: Statistical Mechanics and its Applications*, 451, 373-387.
Ding, R., Ujang, N., bin Hamid, H., Manan, M. S. A., He, Y., Li, R., & Wu, J. (2018). Detecting the urban traffic network structure dynamics through the growth and analysis of multi-layer networks. *Physica A: Statistical Mechanics and its Applications*, *503*, 800-817.
Ding, Y., Yan, E., Frazho, A., & Caverlee, J. (2009). PageRank for ranking authors in co-citation networks. *Journal of the American Society for Information Science and Technology*, *60*(11), 2229-2243.
Faghih-Imani, A., Anowar, S., Miller, E. J., & Eluru, N. (2017). Hail a cab or ride a bike? A travel time comparison of taxi and bicycle-sharing systems in New York City. *Transportation Research Part A: Policy and Practice, 101*, 11-21.
Ferber, C., Holovatch, Y., & Palchykov, V. (2005). Scaling in public transport networks. *Condens Matter Phys.* 8(141):225–234.

Gao, X., & Tian, L. (2019). Effects of awareness and policy on green behavior spreading in multiplex networks. *Physica A: Statistical Mechanics and its Applications*, *514*, 226-234.

Gonzalez, M. C., Hidalgo, C. A., & Barabasi, A. L. (2008). Understanding individual human mobility patterns. *nature*, *453*(7196), 779-782.

Grünwald, P. D., & Grunwald, A. (2007). *The minimum description length principle.* MIT press.

Halu, A., Mondragón, R. J., Panzarasa, P., & Bianconi, G. (2013). Multiplex pagerank. *PloS one*, *8*(10).

Halu, A., Mukherjee, S. and Bianconi, G., 2014. Emergence of overlap in ensembles of spatial multiplexes and statistical mechanics of spatial interacting network ensembles. *Physical Review E*, 89(1), p.012806.

Handy, S. (1996). Methodologies for exploring the link between urban form and travel behavior. *Transportation Research Part D: Transport and Environment*, *1*(2), 151-165.

Hochmair, H. H. (2016). Spatiotemporal pattern analysis of taxi trips in New York City. *Transportation research record, 2542*(1), 45-56.

Hristova, D., Musolesi, M., & Mascolo, C. (2014, May). Keep your friends close and your facebook friends closer: A multiplex network approach to the analysis of offline and online social ties. In *Eighth International AAAI Conference on Weblogs and Social Media*.

Jia, C., Du, Y., Wang, S., Bai, T., & Fei, T. (2019). Measuring the vibrancy of urban neighborhoods using mobile phone data with an improved PageRank algorithm. *Transactions in GIS, 23*(2), 241-258.

Jiang, B., & Yao, X. (Eds.). (2010). Geospatial analysis and modelling of urban structure and dynamics (Vol. 99). *Springer Science & Business Media.*

Kivelä, M., Arenas, A., Barthelemy, M., Gleeson, J. P., Moreno, Y., & Porter, M. A. (2014). Multilayer networks. *Journal of complex networks*, *2*(3), 203-271.

Klinger, T., & Lanzendorf, M. (2016). Moving between mobility cultures: what affects the travel behavior of new residents?. *Transportation, 43*(2), 243-271.

Kong, H., Zhang, X., & Zhao, J. (2020). How does ridesourcing substitute for public transit? A geospatial perspective in Chengdu, China. *Journal of Transport Geography, 86*, 102769.

Li, Y., Yang, L., Shen, H., & Wu, Z. (2019). Modeling intra-destination travel behavior of tourists through spatio-temporal analysis. *Journal of destination marketing & management*, *11*, 260-269.

Li, Z., Hong, Y., & Zhang, Z. (2016, September). An empirical analysis of on-demand ride sharing and traffic congestion. In *Proc. International Conference on Information Systems*.

Liang, X., Zheng, X., Lv, W., Zhu, T., & Xu, K. (2012). The scaling of human mobility by taxis is exponential. *Physica A: Statistical Mechanics and its Applications*, *391*(5), 2135-2144.

Liu, X., Gong, L., Gong, Y., & Liu, Y. (2015). Revealing travel patterns and city structure with taxi trip data. *Journal of Transport Geography, 43*, 78-90.

Lokhandwala, M., & Cai, H. (2018). Dynamic ride sharing using traditional taxis and shared autonomous taxis: A case study of NYC. *Transportation Research Part C: Emerging Technologies, 97*, 45-60.

Luo, X., Dong, L., Dou, Y., Zhang, N., Ren, J., Li, Y., Sun, L. & Yao, S. (2017). Analysis on spatial-temporal features of taxis' emissions from big data informed travel patterns: a case of Shanghai, China. *Journal of cleaner production*, *142*, 926-935.

Mucha, P. J., & Porter, M. A. (2010). Communities in multislice voting networks. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, *20*(4), 041108

Newman, M. E. (2006). Modularity and community structure in networks. *Proceedings of the national academy of sciences*, *103*(23), 8577-8582.

Newman, M. E., & Girvan, M. (2004). Finding and evaluating community structure in networks. *Physical review E*, *69*(2), 026113.

Nicosia, V., Bianconi, G., Latora, V., & Barthelemy, M. (2013). Growing multiplex networks. *Physical review letters*, *111*(5), 058701.

Page, L., Brin, S., Motwani, R., & Winograd, T. (1999). *The PageRank citation ranking: Bringing order to the web*. Stanford InfoLab.

Parshani, R., Rozenblat, C., Ietri, D., Ducruet, C., & Havlin, S. (2011). Inter-similarity between coupled networks. *EPL (Europhysics Letters)*, *92*(6), 68002.

Paulssen, M., Temme, D., Vij, A., & Walker, J. L. (2014). Values, attitudes and travel behavior: a hierarchical latent variable mixed logit model of travel mode choice. *Transportation*, *41*(4), 873-888.

Qian, X., Zhan, X., & Ukkusuri, S. V. (2015). Characterizing urban dynamics using large scale taxicab data. *In Engineering and Applied Sciences Optimization* (pp. 17-32). Springer, Cham.

Ratti, C., Sobolevsky, S., Calabrese, F., Andris, C., Reades, J., Martino, M., Claxton, R. & Strogatz, S. H. (2010). Redrawing the map of Great Britain from a network of human interactions. *PloS one, 5*(12), e14248.

Riascos, A. P., & Mateos, J. L. (2020). Networks and long-range mobility in cities: A study of more than one billion taxi trips in New York City. *Scientific reports, 10*(1), 1-14.

Rodrigue, J. P. (2020). The geography of transport systems. *Routledge*.

Sarkar, S., Chawla, S., Ahmad, S., Srivastava, J., Hammady, H., Filali, F., Znaidi, W. & Borge-Holthoefer, J. (2017). Effective urban structure inference from traffic flow dynamics. *IEEE Transactions on Big Data, 3*(2), 181-193.

Shen, J., Liu, X., & Chen, M. (2017). Discovering spatial and temporal patterns from taxi-based Floating Car Data: A case study from Nanjing. *GIScience & Remote Sensing*, *54*(5), 617-638.

Shmueli, E., Mazeh, I., Radaelli, L., Pentland, A. S., & Altshuler, Y. (2015, March). Ride sharing: a network perspective. In *International Conference on Social Computing, Behavioral-Cultural Modeling, and Prediction* (pp. 434-439). Springer, Cham.

Strano, E., Shai, S., Dobson, S., & Barthelemy, M. (2015). Multiplex networks in metropolitan areas: generic features and local effects. *Journal of The Royal Society Interface, 12*(111), 20150651.

Tao, S., Rohde, D., & Corcoran, J. (2014). Examining the spatial–temporal dynamics of bus passenger travel behaviour using smart card data and the flow-comap. *Journal of Transport Geography*, *41*, 21-36.

Wallsten, S. (2015). The competitive effects of the sharing economy: how is Uber changing taxis. *Technology Policy Institute*, *22*, 1-21.

Wen, T. H. (2015). Geographically modified PageRank algorithms: Identifying the spatial concentration of human movement in a geospatial network. *PloS one*, *10*(10), e0139509.

Xu, J., Li, A., Li, D., Liu, Y., Du, Y., Pei, T., Ma, T. & Zhou, C. (2017). Difference of urban development in China from the perspective of passenger transport around Spring Festival. *Applied Geography*, *87*, 85-96.

Xu, Y., Chen, D., Zhang, X., Tu, W., Chen, Y., Shen, Y., & Ratti, C. (2019). Unravel the landscape and pulses of cycling activities from a dockless bike-sharing system. *Computers, Environment and Urban Systems*, *75*, 184-203.

Yang, Y., Heppenstall, A., Turner, A., & Comber, A. (2019). A spatiotemporal and graph-based analysis of dockless bike sharing patterns to understand urban flows over the last mile. *Computers, Environment and Urban Systems, 77*, 101361.

Yildirimoglu, M., & Kim, J. (2018). Identification of communities in urban mobility networks using multi-layer graphs of network traffic. *Transportation Research Part C: Emerging Technologies*, *89*, 254-267.

Yu, X., Lilburn, T. G., Cai, H., Gu, J., Korkmaz, T., & Wang, Y. (2017). PageRank influence analysis of protein-protein association networks in the malaria parasite Plasmodium falciparum. *International Journal of Computational Biology and Drug Design*, *10*(2), 137-156.

Zhang, X., Huang, B., & Zhu, S. (2020). Spatiotemporal Varying Effects of Built Environment on Taxi and Ride-Hailing Ridership in New York City. *ISPRS International Journal of Geo-Information, 9*(8), 475.

Zhang, X., Xu, Y., Tu, W., & Ratti, C. (2018). Do different datasets tell the same story about urban mobility—A comparative study of public transit and taxi usage. *Journal of Transport Geography, 70*, 78-90.

Zhao, D., Li, L., Peng, H., Luo, Q., & Yang, Y. (2014). Multiple routes transmitted epidemics on multiplex networks. *Physics Letters A*, *378*(10), 770-776.

Zhao, K., Musolesi, M., Hui, P., Rao, W., & Tarkoma, S. (2015). Explaining the power-law distribution of human mobility through transportationmodality decomposition. *Scientific reports*, *5*(1), 1-7.

Zhong, C., Arisona, S. M., Huang, X., Batty, M., & Schmitt, G. (2014). Detecting the dynamics of urban structure through spatial network analysis. *International Journal of Geographical Information Science*, *28*(11), 2178-2199.

Zhou, J., & Qiu, G. (2018). China's high-speed rail network construction and planning over time: a network analysis. *Journal of Transport Geography*, *70*, 40-54.

Zhou, Y., Chen, H., Wu, Y., Wu, J., Li, J., & Chen, L. (2019, February). A Method for Analyzing Pick-Up/Drop-Off Distribution of Taxi Passengers' in Urban Areas Based on Dynamical Network View. In *2019 IEEE International Conference on Big Data and Smart Computing (BigComp)* (pp. 1-8). IEEE.