Journal of Transport Geography 97 (2021) 103226

Contents lists available at ScienceDirect

Journal of Transport Geography

journal homepage: www.elsevier.com/locate/jtrangeo

Unravel the impact of COVID-19 on the spatio-temporal mobility patterns of microtransit

Yirong Zhou [a], Xiaoyue Cathy Liu [b,*], Tony Grubesic [c]

[a] Department of Civil & Environmental Engineering, University of Utah, 110 Central Campus Dr. RM 1650, Salt Lake City, UT 84112, United States of America
[b] Department of Civil & Environmental Engineering, University of Utah, 110 Central Campus Dr. RM 2137, Salt Lake City, UT 84112, United States of America
[c] Research School of Information, The University of Texas at Austin, 1616 Guadalupe St Suite #5.202, Austin, TX 78701, United States of America

ARTICLE INFO

Keywords:
Microtransit
COVID-19
Spatio-temporal analysis
Eigendecomposition
K-clique percolation

ABSTRACT

Shared mobility is an essential component of the larger sharing economy. Ride-hailing, bike-sharing, e-scooters, and other types of shared mobility continue to grow worldwide. Among these services is microtransit, a new transport mode that extends transit coverage within a region. Mobile devices enable microtransit services, aggregating riders and using real-time routing algorithms to group customers traveling in similar directions. Meanwhile, the newly emerged coronavirus, COVID-19, has radically reshaped the ridership behavior of all transit services, including microtransit. While existing research evaluates the performance of microtransit pilot programs before the pandemic, there is no information concerning the spatio-temporal pattern of microtransit activities under the impact of COVID-19. The purpose of this paper is to apply eigendecomposition and *k*-clique percolation methods to uncover the spatio-temporal patterns of microtransit trips. Further, we used these approaches to identify underlying communities using data from a pilot program in Salt Lake City, Utah. The resulting research offers insight into how COVID-19 altered travel behavior. Specifically, eigendecomposition delineated the homogeneity and heterogeneity of travel patterns across temporal dimensions. We identified first mile/last mile trips as a major source of variance in both pre- and post-COVID periods and that transit-dependent users prove to be inelastic despite the threat of COVID-19. The k-clique percolation method detected possible community formations and tracked how these communities evolved during the pandemic. In addition, we systematically analyzed overlapping communities and the network structure around shared nodes by using a clustering coefficient. The workflow developed in this research broadly is generalizable and valuable for understanding the unique spatio-temporal patterns of microtransit. The framework can also help transit agencies with performance evaluation, regional transport strategies, and optimal vehicle dispatching.

1. Introduction

Sustainable transport systems are crucial for interregional and intraregional mobility, commerce, and the socioeconomic stability of the communities they serve. Public transit is an essential component of modern multimodal transportation systems but suffers from various pressures (Wei et al., 2018; Zhou et al., 2020). For example, increasing operating costs and decreasing ridership continue to stress public transit systems throughout the United States (FTA, 2019). While many public systems continue to struggle, microtransit is emerging as an agile alternative for personal mobility. Microtransit is technology-enabled shared transportation that operates in-between fixed-route transit and ride-hailing. It leverages rider aggregation and routing algorithms to group customers traveling within the service zone in similar directions in real-time. Moreover, microtransit often expects customers to walk a short distance to common pick-up/drop-off locations. Thus, the service is transit-like but more nimble when compared to traditional public transit. As a technology-enabled on-demand service, microtransit shares many similarities with other services such as ride-hailing and paratransit. The platform collects requests from personal devices like smartphones then dynamically dispatches available vehicles to fulfill those requests. However, one substantial difference exists. The design of microtransit allows for integration into the current public transit system. Users often take advantage of microtransit to complete first mile/ last mile connections to the larger transit system.

For similar on-demand services, studies of dispatching algorithms

* Corresponding author.
E-mail addresses: Yirong.Zhou@utah.edu (Y. Zhou), cathy.liu@utah.edu (X.C. Liu), grubesic@austin.utexas.edu (T. Grubesic).

https://doi.org/10.1016/j.jtrangeo.2021.103226
Received 14 June 2021; Received in revised form 21 October 2021; Accepted 3 November 2021
Available online 9 November 2021

are numerous. For example, there have been many tests of different optimization algorithms and heuristics for ride-sharing (Agatz et al., 2011; Agatz et al., 2012; Aissat and Oulamara, 2014). Agatz et al. (2011) proposed an optimization framework based on a rolling horizon strategy to solve the dynamic ride-share problem. The authors tested this using a simulation environment based on the field data from the Atlanta Regional Commission (Agatz et al., 2011). Agatz et al. (2012) systematically reviewed the issues in ride-sharing and assessed relevant optimization models. Aissat and Oulamara (2014) modeled dynamic ride-sharing with intermediate locations and presented one enumerate algorithm and two heuristics to solve it. They later tested the approaches on real networks consisting of 3.5 million nodes and 8.7 million edges. Chen et al. (2019) designed an agent-based model to simulate dynamic ride-sharing in a multimodal network then tested it on the classic Sioux Falls network.

For microtransit, the global launch of pilot programs during the past five years (Haglund et al., 2019; Westervelt et al., 2018) seeks to provide (and improve) first/last mile connections to fixed transit stops and stations, replace underperforming bus routes, provide coverage in areas without fixed-route service, and extend the hours of operation for existing bus services. Many studies analyze the overall performance of microtransit pilot programs (Haglund et al., 2019; Volinski (2019) or microtransit vehicles (Ongel et al., 2019). For example, Haglund et al. (2019) systematically evaluated the performance of the *Kutsuplus* pilot in Helsinki, Finland. The evaluation framework used aggregated measures and spatio-temporal metrics, including the average annual number of passengers, annual price per journey, user age class, distribution of hourly departure/arrival trips for analysis. Volinski (2019) provided a case-based review and synthesis of more than 20 transit agencies that had implemented or intended to launch a microtransit service. This review included evaluations of underlying motivations, planning, design, marketing strategies, technology, and performance metrics. Ongel et al. (2019) evaluated the impacts of novel vehicle technologies on vehicle acquisition costs. This evaluation included lifecycle and end-of-life cost estimates for electric microtransit vehicles and conventional buses operating in Singapore.

Meanwhile, a new challenge has emerged for public transit. COVID-19, a novel coronavirus disease, became a global pandemic in, 2020. Nearly 90% of the American adults reported that COVID-19 impacted their personal lives, and 44% of them claimed their lives had changed dramatically (Pew Research Center, 2020). Due to its collective nature, public transit has been hit even harder (Liu et al., 2020; Wilbur et al., 2020; Yi et al., 2021). In New York City, the average subway and commuter rail ridership declined by 80%, and bus ridership dropped by 50% (Gao et al., 2020a, 2020b) in the first week of July 2020 compared to 2019. In Washington DC, subway and bus ridership declined by 90% and 75%, respectively, by the end of March, 2020 (WMATA, 2020) compared to their typical values. In Utah, three major public transit modes - bus, FrontRunner, and TRAX - have witnessed a massive decline in the total ridership upon the pandemic outbreak (Dillman and Posvistak, 2020). The week after the state of emergency was declared, average ridership has declined by 56% compared to the previous week (Dillman and Posvistak, 2020). Similarly, there was a substantial downturn in microtransit use throughout Utah after the COVID-19 outbreak.

Previous studies help deepen our understanding of microtransit and its dynamics. However, there is very little work concerning the spatio-temporal patterns of microtransit trips and the causal factors contributing to these patterns, especially under the impacts of COVID-19. The purpose of this research is to leverage trip data from a microtransit pilot in the State of Utah for developing a methodological framework that unravels the spatio-temporal patterns of microtransit activities in the region. The framework utilizes eigendecomposition to uncover the rhythms and structures of microtransit trips. Using 7-months of microtransit data, we constructed the spatiotemporal patterns of microtransit activities in pre- and post-COVID periods, respectively. Then, we systematically analyze how these patterns deviate from the average pattern in both periods and what possibly caused such variation. We use eigendecomposition to unravel the hidden temporal structures and k-clique percolation theory to explore the potential spatial communities formed in the service region. Also, for both periods, we intend to determine which locations are connected, how strong the connections are, what roles shared nodes (by different communities) play in different network structures, and how patterns evolve as COVID-19 progresses.

This study is important for three reasons. First, because microtransit helps fill gaps between fixed-route systems and ride-hailing, it is essential to understand its underlying spatio-temporal patterns for a community. In short, does the service provide connectivity to the places that people want to go and when they want to get there? Second, the costs of delivering microtransit services can be substantial, and there are no guarantees that the service will attract riders. For example, the now-bankrupt *Bridj* service in Kansas City served only 1480 riders during its year of operation, with the Kansas City Area Transportation Authority (KCATA) spending $1.5 million to subsidize the service. Considering that the first ten rides were free for users of Bridj, this translated into a subsidy of $1000 per ride (Schmitt, 2018). Thus, there are real financial implications for communities offering microtransit services. Developing a framework that can provide the geospatial intelligence required for improving system performance is crucial for service sustainability. Third, while the influence of COVID-19 on microtransit is easily observable in Utah, there is no analysis of the overall effects. This research focuses on the underlying travel patterns associated with microtransit and their changes during the pandemic. Our findings could help transit agencies understand the decline in microtransit ridership and the relationships between public health crisis and microtransit demand.

We organize the remainder of this paper as follows. Section 2 presents a literature review where we discuss elements of the impacts of COVID-19, spatio-temporal analysis, the application of eigendecomposition, and k-clique percolation methods. Section 3 describes the data used in this study and its pre-processing. Section 4 presents our methodological framework for uncovering the spatio-temporal patterns of microtransit activities. Finally, we offer the results in Section 5 and conclude with a summary of our findings and key contributions of the study.

2. Literature review

2.1. COVID-19 related analysis

COVID-19 has been transforming current society in various aspects. Recent studies have found that COVID-19 exerts a substantial impact on the global economy (McKibbin and Fernando, 2020), education (Pragholapati, 2020), mental health (Xiong et al., 2020), public transit (Liu et al., 2020; Wilbur et al., 2020), and many other domains. Researchers have also been studying the reasons why these impacts have taken place. Take public transit as an example. Before the pandemic, a significant motivation behind choosing public transit was self-interest (e.g., low cost, reliability, healthy lifestyle) and environmental consciousness (Hoang-Tung et al., 2015). However, after the pandemic, concerns for public health may dominate, resulting in usage reduction. Studies have also found that the declines in ridership are uneven across social groups (Liu et al., 2020; Wilbur et al., 2020). For example, Wilbur et al. (2020) found that in Nashville, Tennessee, public transit ridership declined up to 19% more in high-income neighborhoods than in low-income neighborhoods. Similarly, Tan et al. (2020) discovered that, under the influence of COVID-19, in cities with overall low public transit ridership, commuters consist of a higher percentage of the lower-income population due to the transit-dependent nature.

In sum, much of the most recent research focuses on COVID-19's influence on the public transit systems, in general. However, very few studies focus on emerging modes, such as microtransit. Furthermore,

according to Utah Transit Authority (UTA), in Utah, 95% of microtransit trips are paid using UTA tickets (UTA on Demand, 2021) which can be considered a strong indication of transit dependency among microtransit users.

2.2. Related spatio-temporal analysis

Microtransit offers demand-responsive service to customers to initiate any trip start and end within a designated service area (FTA, 2021). This structure is similar to many mobility-based services, including bike-sharing and car-sharing programs. Given the fact that mobile devices enable all these programs and one can retrieve a large amount of trip data from these services, there have been a myriad of studies to analyze spatio-temporal trip patterns (Alonso-González et al., 2018; Dong et al., 2018; Xu et al., 2019). For example, Xu et al. (2019) studied the spatio-temporal patterns of bike-sharing in Singapore. They associated the use patterns with built environment indicators such as floor area ratio (FAR) of residential buildings, FAR of commercial buildings, and land use mixture. In other work, Dong et al. (2018) studied both the service patterns and individual behavior patterns of internet-based ride-sharing services based on the record provided by DiDi, Inc. They applied a non-negative matrix factorization method and cluster analysis to study the spatial, temporal, and spatio-temporal elements of ride-sharing trips, as well as divisions of commuting styles and detour patterns. When it comes to microtransit analysis, Haglund et al. (2019) demonstrated the average journeys in specific periods during the day and the distribution of hourly departure/arrival trips based on the microtransit project *Kutsuplus*. While useful, this analysis was limited because it did not address the processes concerning underlying pattern formation. Similarly, Alonso-González et al. (2018) studied the distribution of generalized journey time across service areas in both aggregated and disaggregated levels using empirical data from a pilot program named "Breng flex" in the Netherlands. Again, while previous studies mainly focused on assessing how demand-responsive transport can benefit the current transportation network, we are not aware of any studies that use actual network data and analytics to examine the spatio-temporal patterns of microtransit usage. This research fills the gap by quantifying the evolving patterns of microtransit activities and elucidating the underlying reasons for such patterns.

2.3. Eigendecomposition

Principle Component Analysis (PCA) or Eigendecomposition is a descriptive tool that extracts the major sources of variance to identify directions of variation in the dataset (Abdi and Williams, 2010), reducing dimensionality and facilitating prediction. For example, Horner and Grubesic (2001) applied PCA to generate an index of derived demand for rail usage based on the local demographics in Columbus, Ohio. Nagendra and Khare (2003) applied PCA to analyze an extensive multivariate data set containing traffic, emission, and meteorological data collected from an intersection in Delhi. Specifically, they utilized the loading matrix to assess the significance and correlation of meteorological and traffic characteristics variables. Melman et al. (2021) applied PCA to determine the driving metrics most strongly associated with fuel use. This information helped predict fuel consumption based on driving behavior measurement. The authors used PCA on a matrix of 110 variables $\times$ 4617 laps to extract the significant sources of variance in scores and loadings. Wang et al. (2021) aimed to discover if personal habits or lack of accessibility dictated why some Salt Lake City residents rarely walk or cycle. During the process, they applied PCA to reduce the dimensionality of 40 attitudinal variables for better representation. Xu et al. (2019) incorporated eigendecomposition and seven built environment indicators to explore the spatio-temporal patterns of bike-sharing in Singapore.

In sum, these existing studies demonstrated the usability of PCA or eigendecomposition to extract sources of variance in different scenarios. This research will explore its applicability for analyzing the spatio-temporal patterns of microtransit activities.

2.4. K-clique percolation

Broadbent and Hammersley (1957) were the first to introduce percolation theory. It describes the behavior of a network when existing linkages get eliminated. One can use percolation theory as a framework for community detection and measuring broader percolation phenomena or transition (Grimmett, 1999; Jiang et al., 2018; Li et al., 2015a, 2015b) where the core structure of the original network is preserved. Various fields use percolation concepts and measures, including chemistry, physics, and material sciences (Achlioptas et al., 2009; Goffri et al., 2006; Majdandzic et al., 2014). That said, traditional percolation theory has its limitations. For example, where community detection methods are concerned, traditional percolation theory can only assign one node to a specific community. To explore overlapping communities on a large scale, Palla et al. (2005) proposed an efficient method to uncover the modular structure of complex networks (Palla et al., 2005), referred to as clique percolation or *k*-clique percolation theory.

One can apply *k*-clique percolation theory to uncover the structure of communities based on a similar percolation transition when *k*-cliques organize into a gigantic community as the threshold exceeds a critical point (Derényi et al., 2005; Li et al., 2015a, 2015b; Palla et al., 2006). Thus, *K*-clique can explore large-scale and highly overlapped networks (such as social networks) without breaking the network into isolated clusters. It also naturally lends itself to the microtransit analysis in our study as popular origins or destinations (such as transit stations, supermarkets, residential buildings) are likely to be the "pivot node" shared by multiple communities.

3. Data

3.1. Data source

UTA partnered with Via transportation to launch a microtransit pilot program beginning November 2019 in South Salt Lake. Salt Lake County funded the project. This on-demand, shared-ride pilot is designed to expand access to the transit service throughout the service zone, improve mobility for all users, and provide a quality customer experience. UTA conducted this pilot to see whether microtransit provides a valuable and cost-effective service and whether a future deployment of microtransit service is possible. The project experienced early success, meeting most goals and objectives before the COVID-19 pandemic, including ridership, cost, customer rating, and vehicle hours traveled.

The microtransit program serves about 65 mile^2 in the cities of Bluffdale, Draper, Herriman, Riverton, and South Jordan in Utah. Fig. 1 shows the service area, including seven TRAX and FrontRunner Stations which are the main components of UTA's rail system. The program provides a corner-to-corner service in the region from 6:00 to 22:00 on weekdays only. It allows riders to be picked up and dropped off within a walkable distance (on average 0.1 miles) from their chosen origins and destinations.

The trip data used in this research contains features related to each trip request, including rider ID, pick-up/drop-off coordinates, pick-up/drop-off time, trip duration, trip distance, number of passengers, ride cost, payment type, request source, customer ratings, and wheelchair accessibility. We present the detailed description of the full feature set in Table 1. The study period spans January 1st, 2020 to July 31st, 2020, encompassing 31,199 trips from 1569 unique users. Among these 31,199 trips, there are 2472 unique pick-up points and 2317 unique drop-off points. Apart from the seven TRAX and FrontRunner Stations in Fig. 1, common pick-up/drop-off locations include apartment complexes, single-family houses, supermarkets, churches, educational institutions, private companies, and personal businesses (e.g., pet stores and liquor stores), among others.

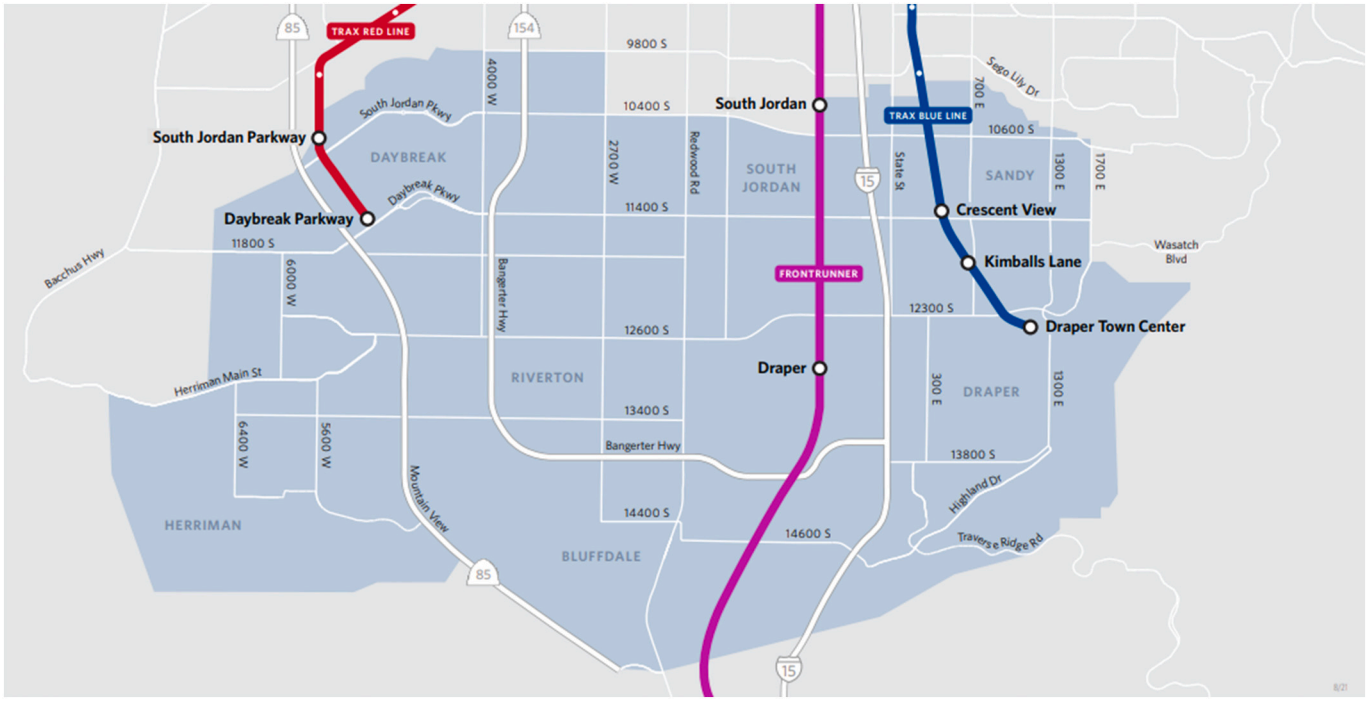

Fig. 1. Service area of UTA on Demand with Via (Source: UTA on Demand, 2021).

Table 1
Feature descriptions of raw data.

Features	Levels of Measurement	Units	Range
Rider ID	Nominal		
Pick-up lat	Interval	Decimal degree	[40.464, 40.568]
Pick-up long	Interval	Decimal degree	[−112.071, −111.83]
Drop-off lat	Interval	Decimal degree	[40.463,40.568]
Drop-off long	Interval	Decimal degree	[−112.071, −111.83]
Pick-up time	Interval		[01/01/2020 08:24:00, 07/31/2020 21:00:00]
Drop-off time	Interval		[01/01/2020 08:48:00, 07/31/2020 21:06:00]
Trip duration	Ratio	Minute	[0, 437.817]
Trip distance	Ratio	Mile	[0.077, 11.104]
Num. of passengers	Ratio		[1, 5]
Ride cost	Ratio	Cent	[0, 1250]
Payment type	Nominal		UTA ticket, Apple Pay, Credit card, Free, Google pay, Ride credit, Waived
Request source	Nominal		App, Call center
Customer rating	Ordinal		1, 2, 3, 4, 5
Wheelchair Accessibility	Nominal		0: wheelchair-accessible vehicle 1: non-wheelchair-accessible vehicle

3.2. COVID-19 in Utah

The first confirmed COVID-19 case in Utah was reported on March 6th, 2020. The patient was aboard the Grand Princess cruise ship and was later diagnosed (Utah Department of Health, 2020). This event, however, does not fully represent the community spread of COVID-19 in Utah. Fig. 2 presents the daily average confirmed COVID-19 cases in each months and daily average microtransit trips in different periods between January 1st and July 31st, 2020. Fig. 2a demonstrates how COVID-19 progressed within the service area and the entire State of Utah, respectively. The service area as mentioned above consists of Bluffdale city, Draper city, Herriman city, Riverton city, and South Jordan city. Fig. 2b shows the daily average microtransit trip counts within the study periods (January, February, March 1st – March 13th, March 14th – March 31st, April through July). We split March into two periods to highlight the comparison before and after the state of emergency declaration. (See Figs. 3 and 4.)

As seen in Fig. 2a, at the state level, the daily average confirmed COVID-19 cases have increased dramatically since March 2020, with the largest increase during June 2020. However, within the service area, during March through May, the confirmed cases remained at a comparatively low level. Nevertheless, the cases increased sharply in June and July, suggesting a wide community spread of COVID-19.

If we compared the trend of microtransit trips in Fig. 2b, there is some consistency. Before March 13th (state of emergency declaration date), daily average microtransit trips increased steadily as the user base of the pilot program was growing. However, between March 14th and March 31st, the daily average trip count dropped dramatically from 325.4 trips per day to 94.4 trips per day. The trip count reached its lowest level in April and May, at 73.4 trips per day and 80.3 trips per day, respectively. The trend in March through May suggests the significant impacts of COVID-19 on microtransit activities. Users were less inclined to use microtransit or stopped using it completely. Trip counts started to increase slightly in June and July, marking the recovery of microtransit activity or overall public transit usage. In the meantime, as we observed in Fig. 2a, a notable increase in COVID-19 cases happens during the same period in the service area.

3.3. Preliminary processing

Given the nature of corner-to-corner service, the actual origin of a trip is most likely within a walkable distance from the pick-up point (available in the trip data). The same applies to the drop-off locations. That said, the trip data shows a wide dispersion of pick-up and drop-off points, where multiple points might lead to identical origins and

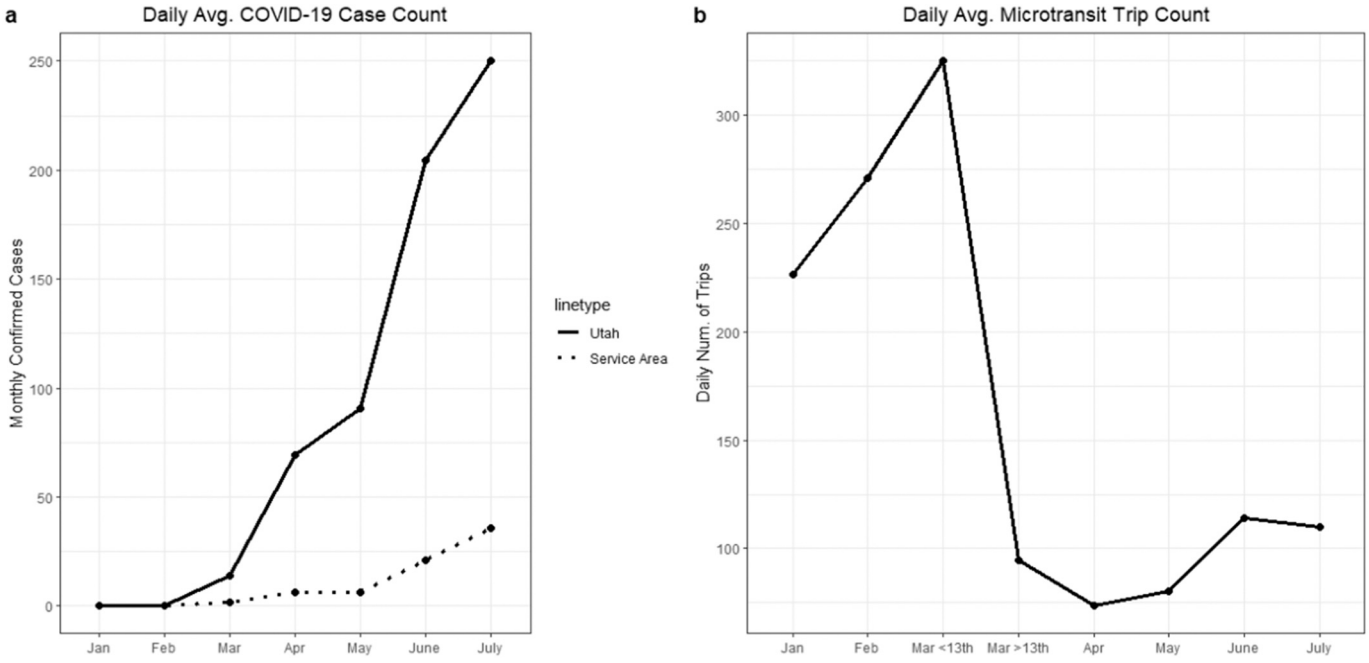

Fig. 2. Trends of COVID-19 Cases and Microtransit Trips.

destinations. Thus, this research applies hierarchical agglomerative clustering (HAC) to separately aggregate close-by pick-up and drop-off points to facilitate subsequent analysis (Murtagh and Contreras, 2012). Agglomerative hierarchical clustering is a standard clustering method based on the distance of clusters. It is a bottom-up approach that assigns each unit to its cluster at the beginning. Then clusters are iteratively merged with their closest neighbors.

When applying HAC, one must determine the pruning parameter – height - representing the number of iterations (merging clusters). Larger values for height (iterations) yield fewer clusters. For this analysis, we select a value of 300 so that the average distance within clusters is close to 0.1 miles. This value is the average walking distance assumed by Uber Pool service (Pachal, 2018). Figs. 3 and 4 highlight the distributions of pick-up and drop-off locations, respectively.

After merging pick-up and drop-off locations, we further divide the dataset into two periods, with a cut-off date of March 13th, 2020, corresponding to the declaration of the COVID-19 outbreak as a national emergency in the United States. This date and data partition helps distinguish the potentially heterogeneous spatio-temporal patterns as a result of the pandemic. Upon screening of missing values, the pre-COVID period consists of 17,980 trips, and the post-COVID period consists of 13,188 trips.

4. Methodology

4.1. Eigendecomposition

To uncover the spatiotemporal trip pattern for the microtransit program, we developed a methodological framework to delineate the variation and the homogeneity/heterogeneity in trips across spatial and temporal dimensions. Specifically, we employ eigendecomposition to achieve this. As detailed previously, eigendecomposition is well suited

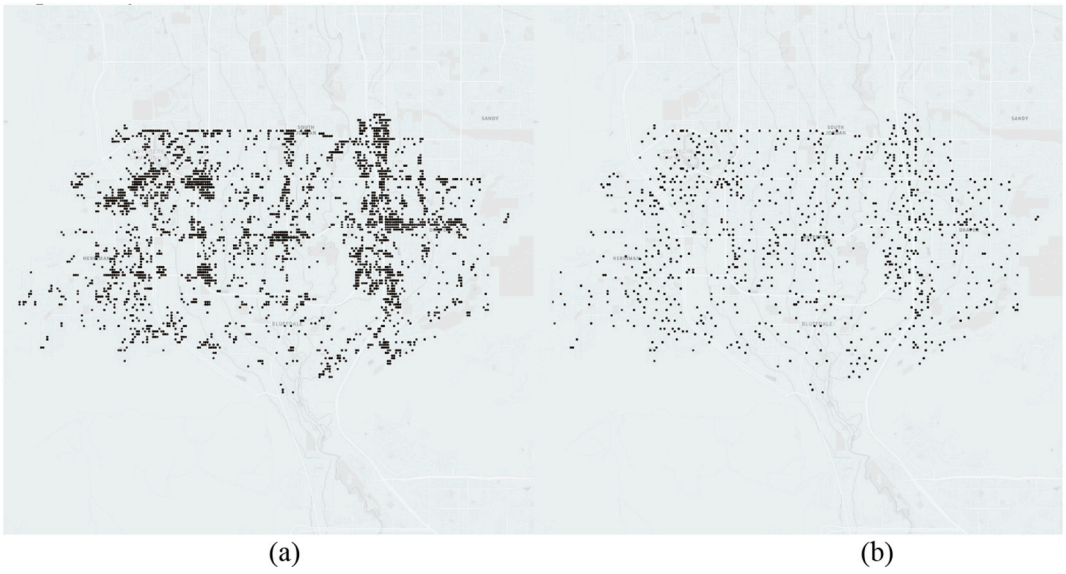

Fig. 3. Distribution of trip pick-up locations (a) before and (b) after merging.

Fig. 4. Distribution of trip drop-off locations (a) before and (b) after merging.

for this as it is good at uncovering hidden structures of spatiotemporal patterns. Compared to other variation extraction methods, such as factor analysis and analysis of variance, eigendecomposition is advantageous because it does not assume any distribution, error term, or underlying statistical model. Eigendecomposition generates a series of directional vectors (i.e., principal components or PCs) based only on the data, each of which best explains variation while remaining orthogonal to other components. One orders the PCs by the amount of variance that they explain. For example, the first PC explains the most variance; the second PC explains the second most variance, and so on. Eigendecomposition is often used in statistical modeling to reduce dimensionality by projecting the data into a few PCs. One can also use eigendecomposition to explore the inherent variation within datasets. Our study uses eigendecomposition to unravel the deviation from (or resemblance of) the average pattern of microtransit usage within the service area in both pre- and post-COVID periods.

We assume that microtransit arrival and departure patterns differ from each other. Therefore, all trips are further categorized into four groups: (a) pre-COVID departure; (b) pre-COVID arrival; (c) post-COVID departure; and (d) post-COVID arrival. During the study period, over 800 pick-up and drop-off locations resulted from 31,168 unevenly distributed trips within the service area after merging close-by locations. For example, among the 892 trip pick-up locations, the most popular pick-up location generated 2633 trips, while over 50% of pick-up locations generated less than 12 trips each. To ensure unbiased analysis, we use traffic analysis zones (TAZ) to aggregate the pick-up and drop-off locations and uncover the trip patterns' geographic dimensions. Among the 163 TAZs within the service area, those that produced over 87 trips (median) and attracted over 85 trips (median) are labeled as active TAZs and used in this study. These 77 TAZs accounted for more than 91.7% of the total microtransit trips. 16,522 trips occurred pre-COVID, and 12,062 trips occurred post-COVID.

Using pre-COVID departure trips as an example, we illustrate the basic structure of our eigendecomposition. We formulate departure trips as follows:

$$\boldsymbol{N} = \begin{pmatrix} \boldsymbol{n}_{1,6} & \cdots & \boldsymbol{n}_{1,21} \\ \vdots & \ddots & \vdots \\ \boldsymbol{n}_{77,6} & \cdots & \boldsymbol{n}_{77,21} \end{pmatrix} \# \tag{1}$$

In matrix $\boldsymbol{N}$, $\boldsymbol{n_{i,j}}$ represents the total number of departure trips originated from TAZ $\boldsymbol{i}$ ($\mathbf{1} \le \boldsymbol{i} \le \mathbf{77}$) and in hour $\boldsymbol{j}$ ($\mathbf{6} \le \boldsymbol{j} \le \mathbf{21}$). We ensure that the 16 variables (6:00– 22:00) contribute equally to the variance maximizing exercise during eigenvalue decomposition by normalizing matrix $\boldsymbol{N}$ by dividing each row by its summation $\boldsymbol{N}'$. Correspondingly, $\boldsymbol{n_{i,j}}'$ represents the hourly percentage of departure trips during the pre-COVID period. Note that:

$$\sum_{j=6}^{j=21} \boldsymbol{n_{i,j}}' = \mathbf{1} \forall \boldsymbol{i} \in [\mathbf{1},\mathbf{77}] \# \tag{2}$$

We then average hourly departure trips across all TAZs to obtain the average temporal pattern of departure trips within the service area. Let $\boldsymbol{d_j}$ denote the average departure trip percentage in TAZ $\boldsymbol{j}$.

$$\boldsymbol{d}_j = \sum_{i=1}^{i=77} \boldsymbol{n_{i,j}}' \Big/ \mathbf{77}\, \boldsymbol{j} \in [\mathbf{6},\mathbf{21}] \tag{3}$$

D = $\{\boldsymbol{d_6}, \boldsymbol{d_7}, \boldsymbol{d_8}, \ldots, \boldsymbol{d_{21}}\}$ represents the average temporal pattern of microtransit activities across the 77 active TAZs. We measure the deviation of microtransit trips from this average pattern with matrix $\boldsymbol{M}$, which we construct subtracting $\boldsymbol{d_j}$ from column $\boldsymbol{j}$ of matrix $\boldsymbol{N}$:

$$\boldsymbol{M} = \begin{pmatrix} \boldsymbol{n}'_{1,6} - \boldsymbol{d_6} & \cdots & \boldsymbol{n}'_{1,21} - \boldsymbol{d_{21}} \\ \vdots & \ddots & \vdots \\ \boldsymbol{n}'_{77,6} - \boldsymbol{d_6} & \cdots & \boldsymbol{n}'_{77,21} - \boldsymbol{d_{21}} \end{pmatrix} \tag{4}$$

We perform eigendecomposition by first calculating the covariance matrix, $\boldsymbol{S}$, where:

$$\boldsymbol{S} = \frac{\mathbf{1}}{\boldsymbol{n}-\mathbf{1}} \boldsymbol{M}^{\boldsymbol{T}} \boldsymbol{M} \# \tag{5}$$

As a result, we can derive eigenvectors $\boldsymbol{v_6}, \boldsymbol{v_7}, \boldsymbol{v_8}, \ldots, \boldsymbol{v_{21}}$ with the corresponding eigenvalues $\boldsymbol{\lambda_6}, \boldsymbol{\lambda_7}, \boldsymbol{\lambda_8}, \ldots, \boldsymbol{\lambda_{21}}$where $\boldsymbol{v_i s}$ are all column vectors. Also, the principal components that explain a large portion of the total variance (summation of eigenvalues) can help interpret the overall patterns of microtransit activities. Thus, by applying eigendecomposition, we can uncover the spatio-temporal patterns of microtransit activities at different locations and examine how they deviate or resemble the average pattern in pre-COVID and post -COVID periods.

4.2. *k*-clique percolation

k-clique percolation is a variant of the traditional percolation theory. Percolation theory aims to discover how networks behave when removing nodes or links. It is a method to gradually break down large

networks into smaller connected clusters or sub-networks. This method works because the removal of key nodes or links in any network creates disconnections. Generally, percolation theory starts with a weighted or unweighted graph and ends at a graph containing only isolated nodes. The nature of percolation theory makes it suitable for identifying strongly connected subgraphs in the network, also known as community detection (Fortunato, 2010). Most community detection algorithms, including percolation, classify one node into one cluster or community *only*. However, *k*-clique percolation identifies overlapping communities by assigning specific nodes to *multiple* communities. This methodological attribute is essential for understanding microtransit because many first mile/last mile trips will make the transit station a pivot node that belongs to multiple communities. This local transport context is the primary reason why we use *k*-clique percolation in this research. Moreover, the concept of *k*-clique is a good analogy to individuals' microtransit travel patterns. People frequently travel between a limited number of locations within a service region.

Fig. 5 illustrates a 7-node unweighted and undirected network to highlight how k-clique percolation works. There are three key definitions:

Definition 1. *k*-clique is a fully connected or complete subgraph of *k* nodes.

Definition 2. Adjacent *k*-cliques are *k*-cliques that share exactly *k*-1 nodes.

Definition 3. A *k*-clique community is the union of all possible adjacent *k*-cliques.

The first step of *k*-clique percolation is to identify *k*-cliques. Here we choose $k = 3$ for illustration. There are three 3-clique in the example network: a – b – c, b – c – d, and c – f – g. The next step is to identify possible adjacent *k*-cliques which make up communities. Fig. 6 highlights the results with each node corresponding to its assigned community (by number). For example, Node c belongs to both community 1 and 2, while node e does not belong to any community. For the weighted graph, there exists one additional intermediate step. In traditional percolation theory, a probability threshold determines whether two nodes are connected or not. Similarly, for the *k*-clique percolation method, we define an *intensity* threshold to determine whether a k-clique can be part of adjacent *k*-cliques. Specifically, intensity is a measure of the strength of connectivity of a clique. We define intensity as the geometric mean of the link weights of a certain *k*-clique:

$$\boldsymbol{Intensity} = \left(\sum_{1 \le m \le n \le k} \boldsymbol{W}_{m,n} \right)^{\frac{2}{k(k-1)}} \tag{6}$$

We denote the intensity threshold by $\boldsymbol{I}$, and we do not consider cliques whose intensity is below the threshold value for adjacent cliques. In Fig. 7, we detail a weighted and undirected graph. The number on the link represents the link weight. Among the three 3-cliques identified in Fig. 5, two of them have an intensity value of 0.187 (a-b-c and c-g-f), and one of them has an intensity of 0.1 (c-b-d) according to Eq. (6). When one sets $\boldsymbol{I}$ as 0.01, all three cliques survive the threshold. According to Definition 2, clique a-b-c and c-b-d are considered adjacent. Thus, there are two communities detected in the network. One contains node a, b, c, and d, and the other contains nodes c, f, and g, where node c belongs to both communities and node e is an isolated node (Definition 3). When $\boldsymbol{I} = \mathbf{0.18}$, cliques a – b – c and c –g – f survive, but we eliminate clique c-b-d. Likewise, cliques a-b-c and c-g-f now make two separate communities while nodes e and d are isolated nodes. When $\boldsymbol{I} = \mathbf{0.2}$, no clique survives. This type of network structure measurement is how k-clique percolation works for given values of $\boldsymbol{I}$ and $\boldsymbol{k}$. However, an outstanding challenge remains – namely, determining optimal values for $\boldsymbol{k}$ and $\boldsymbol{I}$.

The criterion of choosing $\boldsymbol{k}$ and $\boldsymbol{I}$ is to find the most highly structured communities possible. Although different $\boldsymbol{k}$ and $\boldsymbol{I}$ values can generate various optimal subgraphs, there is still a need to set a global rule such that the overall structure and pattern can be analyzed. Derényi et al. (2005) highlight one way to do this. Specifically, as one continues to remove links in a graph, a gigantic component containing a large portion of the existing nodes will eventually emerge. Thus, a rule of thumb for determining $\boldsymbol{I}$ for each $\boldsymbol{k}$ (typically 3 to 6) is to choose an $\boldsymbol{I}$ that is smaller than the value that allows the emergence of the gigantic component (Derényi et al., 2005; Pallas et al., 2005). This value of $\boldsymbol{I}$ is considered the critical point for this network at a specific $\boldsymbol{k}$. One can argue that the size of communities at the critical point follows a power-law. One can estimate this critical point of $\boldsymbol{I}$ by calculating the ratio between the size (number of nodes) of the largest community and the second-largest community. In Palla et al. (2005), the ratio is set to 2 so that there is neither a gigantic community that slanders the details of the network nor are there too many small communities making the network poorly structured.

Moreover, to further study the connectivity of local cohesiveness, we apply the concept of a clustering coefficient (Barrat et al., 2004). As a measurement of inner-connectivity in the network, one can apply the clustering coefficient in two ways: 1) the clustering coefficient of a single node and 2) the average clustering coefficient of all nodes in the

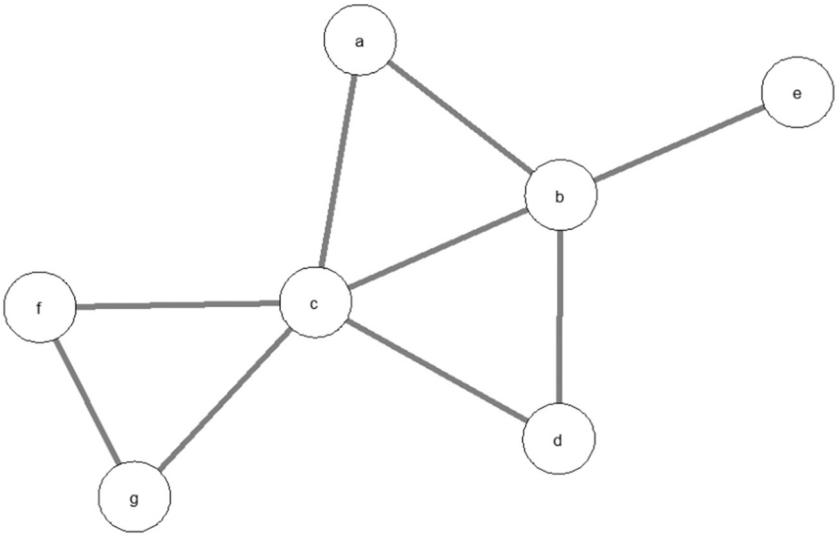

Fig. 5. An unweighted and undirected graph with 7 nodes.

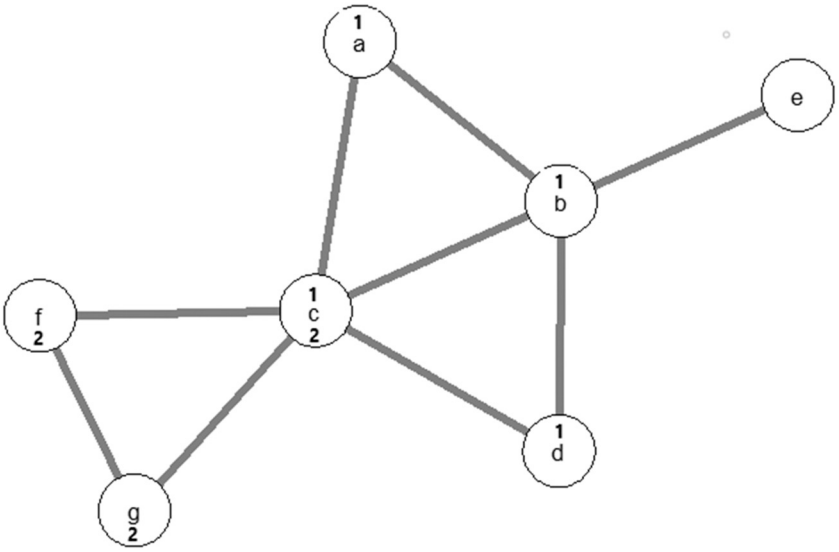

Fig. 6. Result of 3-clique percolation.

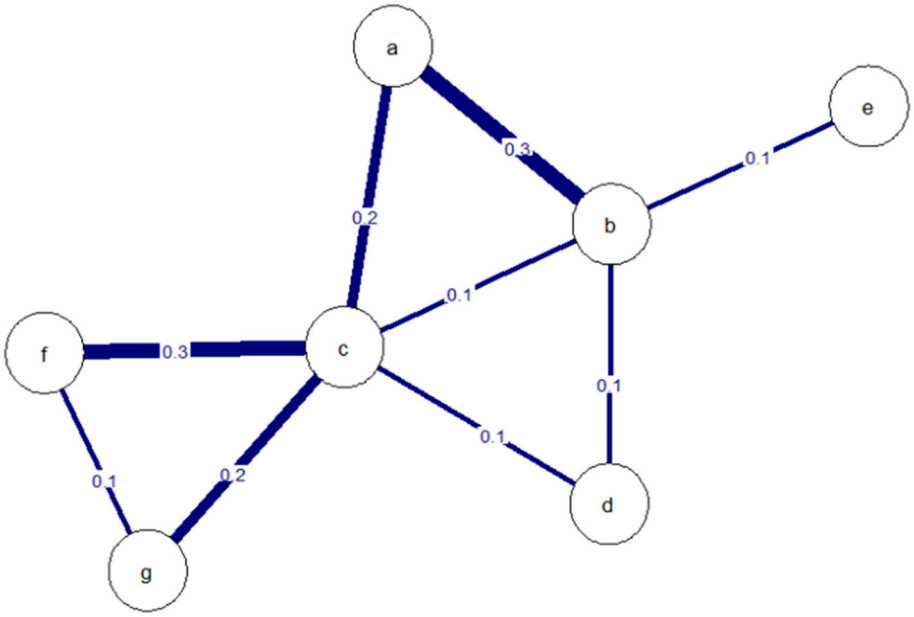

Fig. 7. Weighted graph with 7 nodes.

network. For weighted networks, one can use the clustering coefficient to evaluate the importance of clustered structures based on the aggregated link weights of the local triplets. In this instance, a triplet refers to three nodes connected by either two (e.g., open triplet, examples: a-b and b-c in Fig. 7) or three (e.g., closed triplet, example: a-b, b-c, and a-c in Fig. 6) undirected links. Closed triplets are known as a 3-clique. In a network where one denotes a complete node set as $\boldsymbol{N}$, the formula for calculating clustering coefficient for node $\boldsymbol{i}$ is:

$$\boldsymbol{C_i} = \frac{\mathbf{1}}{\boldsymbol{s_i}(\boldsymbol{n_i}-\mathbf{1})}\sum_{\boldsymbol{j,k}}\frac{(\boldsymbol{w_{ij}}+\boldsymbol{w_{ik}})}{\mathbf{2}}\boldsymbol{a_{ij}a_{jk}a_{ik}}\,\boldsymbol{j,k}\in\boldsymbol{N} \tag{7}$$

Where, $\boldsymbol{a_{ij}}$,$\boldsymbol{a_{jk}}$ and $\boldsymbol{a_{ik}}$ are binary indicators implying connection ($\boldsymbol{a_{ij}}$, $\boldsymbol{a_{jk}}$,$\boldsymbol{a_{ik}}=\mathbf{1}$) or disconnection ($\boldsymbol{a_{ij}}$,$\boldsymbol{a_{jk}}$,$\boldsymbol{a_{ik}}=\mathbf{0}$). $\boldsymbol{w_{ij}}$ is the weight between node $\boldsymbol{i}$ and $\boldsymbol{j}$. $\boldsymbol{n_i}$ is the degree of node $\boldsymbol{i}$ defined as $\boldsymbol{n_i}=\sum_j \boldsymbol{a_{ij}}$. $\boldsymbol{s_i}$ corresponds to node strength, which measures the network property of node $\boldsymbol{i}$ in terms of the weights obtained by extending the definition of node degree $\boldsymbol{n_i}$. We define this as:

$$\boldsymbol{s_i} = \sum_{\boldsymbol{j}}\boldsymbol{a_{ij}w_{ij}} \tag{8}$$

$\boldsymbol{C_i}$ accounts for all triplets formed around $\boldsymbol{i}$ in terms of average link weight. We normalize this by $\frac{1}{s_i(n_i-1)}$ so that $\mathbf{0}\le\boldsymbol{C_i}\le\mathbf{1}$. Larger values for $\boldsymbol{C_i}$, indicate more cohesiveness for nodes around $\boldsymbol{i}$. This structure means that the nodes around $\boldsymbol{i}$ are more tightly connected. We define the average clustering coefficient $\boldsymbol{C_{avg}}$ is defined as:

$$\boldsymbol{C_{avg}} = \frac{\sum_{\boldsymbol{i}}\boldsymbol{C_i}}{|\boldsymbol{N}|}\;\boldsymbol{i}\in\boldsymbol{N} \tag{9}$$

One can use $\boldsymbol{C_{avg}}$ to estimate the level of network overlap. Higher values of $\boldsymbol{C_{avg}}$ suggest more overlap.

In the context of microtransit activities, our goal is to identify communities in pre-COVID and post-COVID periods to examine the pattern

within each period and how the pattern evolves. Again, we treat each pick-up or drop-off location as nodes in the network. Since the customers using microtransit programs tend to travel between a limited number of locales multiple times, the outflow and inflow of most nodes are almost equal to each other. Hence, we construct the network as an undirected and weighted graph. We measure link weights as the ratio between the trip count on each link and the maximal trip count across all links. We apply k-clique percolation to the built network to uncover the underlying communities and how communities evolve. The preliminary analysis for both pre-COVID and post-COVID periods suggests a value of $\boldsymbol{k} = 3$ will provide a good baseline for measurement. When $\boldsymbol{k} \geq \mathbf{4}$, the number of k-cliques is very limited, and more than 75% of the nodes became isolated even when $\boldsymbol{I} = 0$. Also, less than 4% of the nodes will survive when reaching the critical point. In other words, $\boldsymbol{k} \geq \mathbf{4}$ eliminates a vast portion of the network before structure detection. After we fix $\boldsymbol{k}$ to 3, we explore a range of values for $\boldsymbol{I}$ to reveal different levels of information. When $\boldsymbol{I} = 0$, one can treat the system as an unweighted network since any 3-cliques can survive the threshold. In this case, we fully preserve the topological structure. However, as we increase the values of $\boldsymbol{I}$, cliques with less weight will continue to dissipate. Only the cliques with larger weights (more trips) will remain in the network.

5. Results and analysis

5.1. Eigendecomposition

There are 16,522 pre-COVID trips and 12,062 post-COVID trips in the 77 active TAZs used for analysis. These counts translate to an average of 229.5 and 86.2 trips per day, respectively. Such daily usage of microtransit demonstrates a sharp decrease in activities since the outbreak of COVID-19. Fig. 8 presents the overall temporal patterns (sum of each hour) for departure and arrival trips in pre- and post-COVID regimes, respectively. Both regimes show a two-peak distribution, one in the morning (7:00– 9:00) and one in the evening (16:00–18:00). Moreover, the departure vs. arrival trip patterns are very similar (with a small time lag) because the service region is limited, and the average trip duration is around 10 min. Apart from the overall decline in microtransit activities, it is interesting to note the pronounced drop in the morning peak (7:00–8:00) after the COVID-19 outbreak.

Table 2 shows the total variance and the respective portions of variance explained by the first four PCs. Here, we apply three empirical rules to decide upon a suitable number of PCs (Jolliffe, 2002; Xu et al., 2019):

1. The chosen PCs cumulatively explain more than 70% of the total variation.
2. The eigenvalues of the chosen PCs are larger than 0.7 times the average eigenvalue.
3. The chosen PCs are to the left of the elbow points of scree plot, which shows the eigenvalues for each PC.

Given these guidelines, we use the first four PCs for further analysis. Table 2 shows that the first four PCs explain more than 60% of the total variation for all four trip categories. While the results are largely homogeneous, each of the four categories displays its own temporal structure. For example, the variance of pre-COVID trips is smaller than post-COVID trips, generally speaking. This result suggests that microtransit usage is more consistent in the pre-COVID period, but it is more diverse (from place to place) in the post-COVID period.

Also, in the pre-COVID period, the variance of departure trips is significantly larger than that of arrival trips (by 21%). However, in the post-COVID period, their variance is quite similar. This result suggests a higher diversity in temporal patterns for departure trips when compared to arrival trips, pre-COVID. Nevertheless, this temporal signature disappeared in the post-COVID period. One possible explanation for this is that drop-off locations are often more connected (for trip purposes) than pick-up locations. This finding suggests that TAZs with similar social

Table 2
Summary of eigendecomposition.

	Total Variance	PC1	PC2	PC3	PC4
Pre / Departure	0.088	37.4%	17.7%	11.8%	10.1%
Pre / Arrival	0.073	36.4%	12.9%	11.1%	9.5%
Post / Departure	0.098	23.9%	16.1%	15.5%	10.4%
Post / Arrival	0.096	27.5%	15.8%	12.5%	9.1%

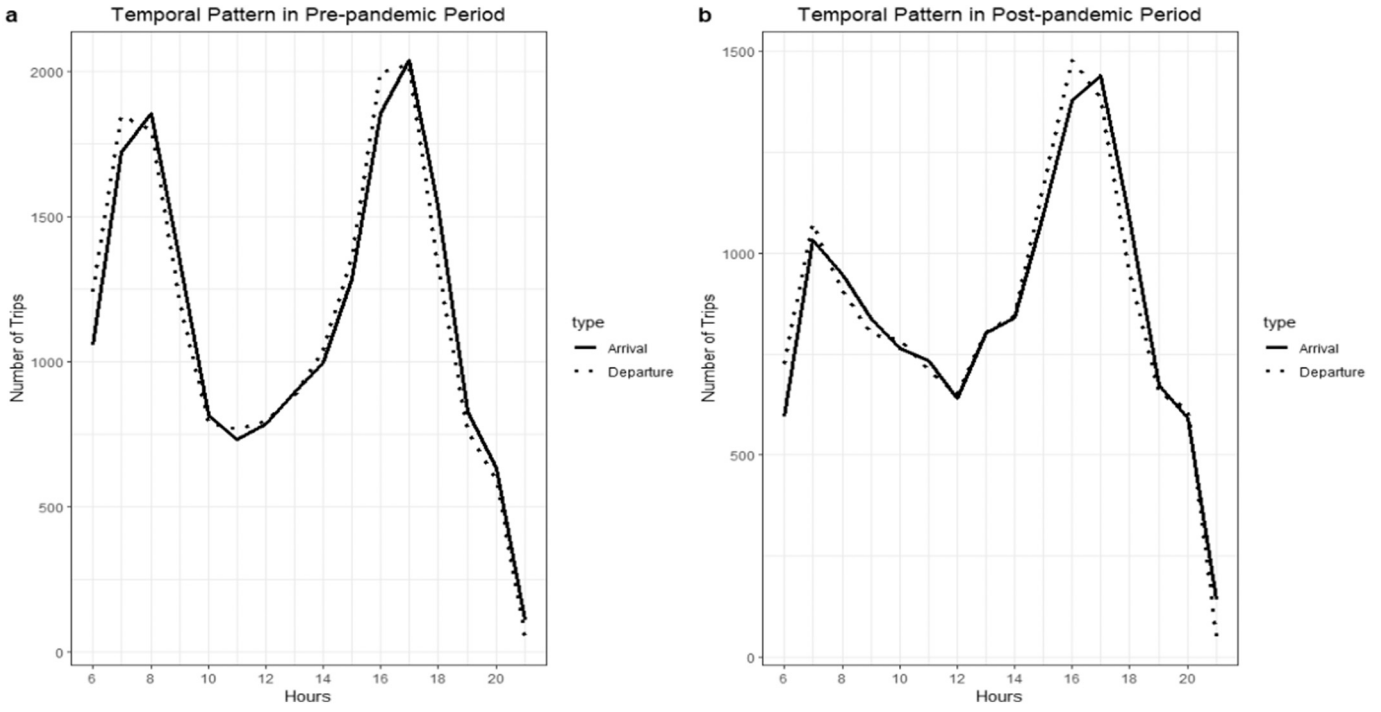

Fig. 8. Distribution of edges by the number of trajectories.

functions tend to generate similar microtransit usage patterns. As a result, the total variance is lower.

In this research, first mile/last mile trips are important. This type of trip accounts for more than 58% of total trips in the pre-COVID period, resulting in a more uniform arrival pattern. However, after the outbreak of COVID-19, the percentage of first mile/last mile trips continues to decrease. For example, in July, the relative number of trips dropped to 29%. Meanwhile, trip purposes have become much more diverse, making the variance of arrival trips close to departure trips.

The first several PCs can help summarize the spatio-temporal patterns of microtransit activities. Fig. 9 displays the first 4 PCs for pre-COVID trips. In Fig. 9, the x-axis represents the 16 input variables that correspond to each hour when the service is in operation, and the y-axis demonstrates the loadings of the first 4 PCs associated with all the input variables. In other words, the y-axis is the coefficient of the linear combination between PCs and input variables. For example, PC1 of pre-COVID arrival trips explains 37.4% of the total variation and shows a clear two-peak pattern at 7:00–10:00 and 17:00–19:00 (Fig. 9a). This result suggests that microtransit activity during these two time periods varies considerably, especially when compared to the uniformity found in other periods. Fig. 9e, also suggests that PC1 for departure trips explains 36.4% of the total variation and has a similar two-peak pattern, yet the graph shifts slightly to the left. The two-peak periods are 6:00–9:00 and 16:00–18:00. Such a pattern is consistent with the morning and evening peaks of transit usage, indicating that both arrival and departure trips have the most diversified spatio-temporal patterns during these two peak periods.

Of note are the results from Fig. 9a. In this instance, the coefficients of PC1 have opposite signs in the morning and evening peaks. This result suggests that if a TAZ attracts more microtransit trips during 7:00– 9:00, then arrivals trips are likely lower in that TAZ during 17:00– 19:00. In fact, of the top five TAZs that attract most trips in the morning peak, four of them attract a limited number of trips in the evening peak. The pattern holds true for departure trips as well. If more people departed certain TAZs during 16:00– 18:00, fewer customers tend to depart from those TAZs during 6:00– 9:00.

While the coefficients of PC1 support these results, PC2, PC3, and PC4 combine to explain 39.6% of the variation for arrival trips and 33.5% variation for departure trips. This explanatory power means that there are other significant temporal patterns that PC1 cannot explain (by itself) and that many TAZs do not follow the patterns shown by PC1. For example, PC2 explains 17.7% of the total variation for arrival trips, suggesting a notable shift in the morning and evening peaks (Fig. 9b). Further, Fig. 9b suggests that in some TAZs if more patrons arrive during 17:00– 18:00, fewer customers will arrive at those TAZs during 18:00–19:00. Fig. 9f highlights a similar pattern for departure trips. The variation of temporal patterns peaks during 7:00– 8:00 and 8:00– 9:00, yet in opposite directions. Moreover, PC3 and PC4 each explain around 10% of the total variation for both arrival and departure trips. The proximity of variance explained by PC1, PC2, PC3, and PC4 suggests a variety of underlying temporal patterns in pre-COVID microtransit activities, which, to some extent, uncovers the diversity of trip purposes.

To deepen our understanding of the causal mechanisms for the patterns displayed in Fig. 9, we focus on the temporal pattern of first mile/last mile trips. In this study, the first mile/last mile trips either start from or end at any transit station identified in Fig. 1. Fig. 10 shows the percentage of first mile/last mile trips over total trips that start in each hour for the pre-COVID and post-COVID periods, respectively. The average value in Fig. 10 is the average percentage of total first mile/last mile trips over total trips for the pre-COVID and post-COVID periods, respectively. The percentage of first mile/last mile trips of both arrival and departure trips peaks between 6:00– 9:00 and 16:00– 19:00. This result means that many patrons travel to or from transit stations during those time periods. It is important to note that the first mile/last mile trip patterns are quite similar pre- and post-COVID. In other words, although there is an overall reduction (from 58.07% to 47.88%) in percentage, first mile/last mile trips still account for a large portion of the total trips. One can observe the same peaks in Fig. 9(a) and Fig. 9(e). This pattern implies that when the percentage of first mile/last mile trips is high, the variation in travel patterns is more extensive and vice versa. As we further examine PC3 and PC4 of arrival and departure trips, please note that the variation peaks between 6:00– 9:00 and 16:00– 19:00. These windows are when the percentage of first mile/last mile trips is above average (Fig. 10). It indicates that PC1, PC3, and PC4 all capture part of the variation caused by first mile/last mile trips.

Fig. 11 presents the eigendecomposition of post-COVID trips. One focus of the eigendecomposition is to study how COVID-19 has altered the temporal patterns of microtransit activities. For example, Fig. 11(a) and Fig. 11(e) illustrate a similar two-peak pattern for arrival and departure trips. For arrival trips, the largest variation occurs during 8:00–9:00 and 18:00– 19:00. For departure trips, the peak periods are 6:00–7:00 and 16:00– 17:00. Thus, one might suggest that specific patterns emerged after the outbreak of COVID-19. However, the variance explained by PC1 for arrival and departure trips decreased from 37.4% and 36.4% to 27.5% and 23.9%, respectively. This decrease in variance explained by PC1 suggests that the temporal patterns of post-COVID trips are more diverse. Also, different patterns can be found based on the PCs. For instance, PC2 for post-COVID arrival trips demonstrates a different one-peak pattern. Specifically, for some TAZs, trip arrival is most active and diverse during 13:00– 14:00 (Fig. 11(b)). That said, PC2, PC3, and PC4 do not exhibit such a pattern for pre-COVID arrival trips. The difference between the two eigendecomposition results suggests a substantial pattern transition after the outbreak of COVID-19.

Based on the results of eigendecomposition, the first mile/last mile trips are likely to be the primary source of variance, which PC1 explains in both pre- and post-COVID periods. Moreover, transit-dependent users remain inelastic despite the threats brought by COVID-19 (Fig. 10). However, the patterns are mutable. For example, there are reductions in microtransit activity along with first mile/last mile trips. In addition, there is a dispersive trend for pick-up and drop-off locations and the emergence of new travel patterns. In the next section, we analyze the coordinates of pick-up and drop-off locations to detect underlying community structures and compare the temporal patterns of these locations to the service area averages.

5.2. K-clique percolation

5.2.1. Pre-COVID period

There are 17,980 trips in the pre-COVID period from 1235 different users in 1163 pick-up and drop-off locations. The maximum number of trips is 348, while the minimum number is 1. We first vary $\boldsymbol{I}$ (Intensity thresholds) to see how communities progress. We present the evolving pattern of communities with increasing values of $\boldsymbol{I}$ in Fig. 12. Specifically, Fig. 12(a) displays the total number of communities in response to changes in $\boldsymbol{I}$. Fig. 12(b) demonstrates the size ratio of the largest community to the second-largest community. For higher values of $\boldsymbol{I}$, community formation decreases. This type of pattern is not always monotonic. For example, when $\boldsymbol{I}$ increases, cliques with lower levels of intensity will disappear. Simultaneously, these eliminations will dissolve large communities into smaller ones. For instance, in Fig. 12(b), the ratio is well above 2 when $\boldsymbol{I}$ is small, indicating a gigantic community, with a few small communities containing a minimal number of nodes. The formation of a gigantic community suggests a wide connection across the service region, apart from a few pick-up or drop-off locations. The vertical dashed line in Fig. 12 highlights the "critical" point, detailed in Section 4. Specifically, when $\boldsymbol{I}$ increases from 0, fewer 3-cliques remain in the network. Until $\boldsymbol{I}$ reaches 0.031, the gigantic community that contains most of the nodes breaks down into two medium-sized communities with a ratio of 4.05. Thus, we identify $\boldsymbol{I} = 0.031$ as the critical point for the pre-COVID microtransit network.

We further explore the characteristics of communities using ten unique values for $\boldsymbol{I}$: 0, 0.031, 0.047, 0.063, 0.079, 0.095, 0.111, 0.127,

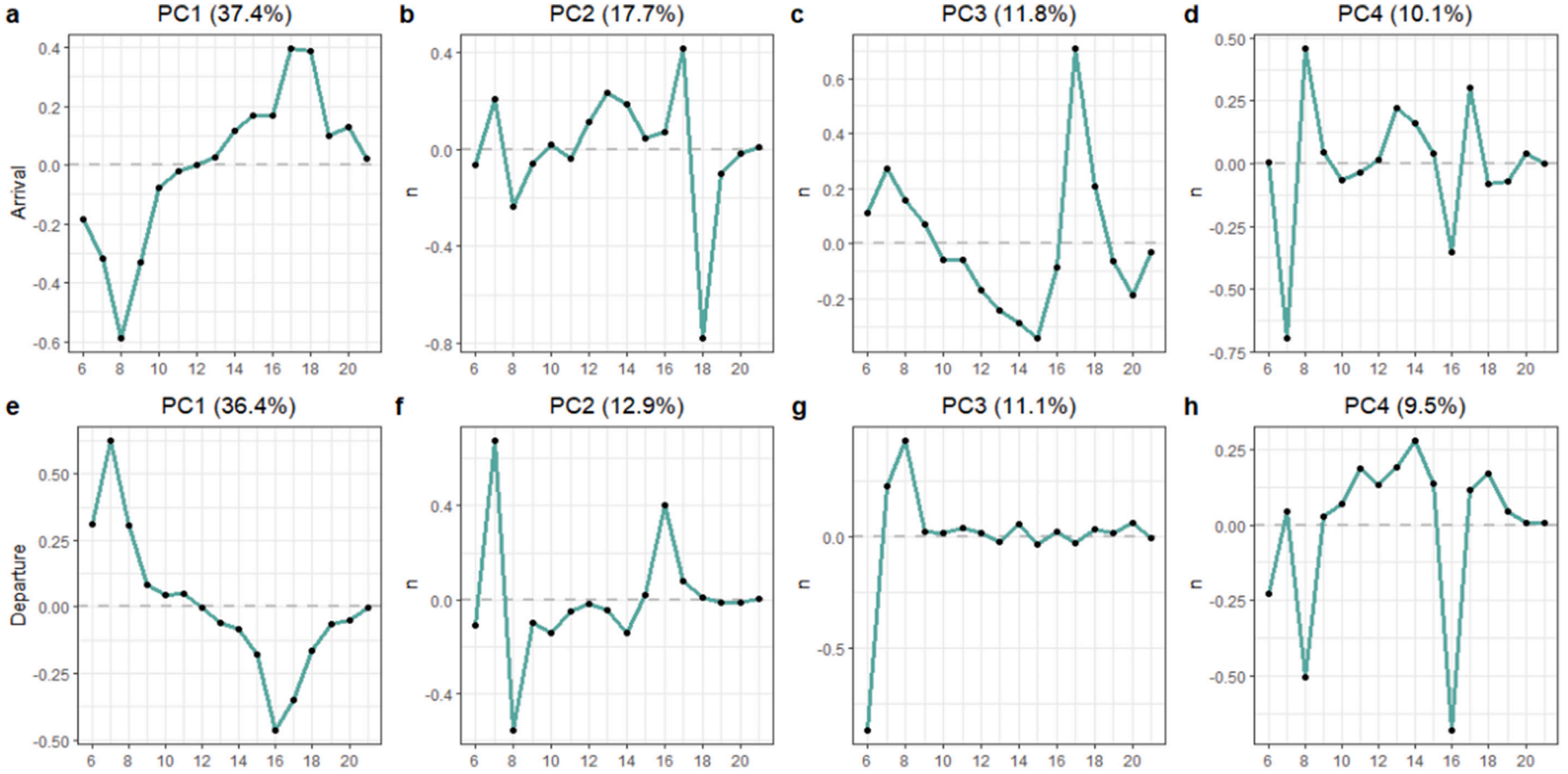

Fig. 9. Eigendecomposition of pre-COVID arrival and departure trips.

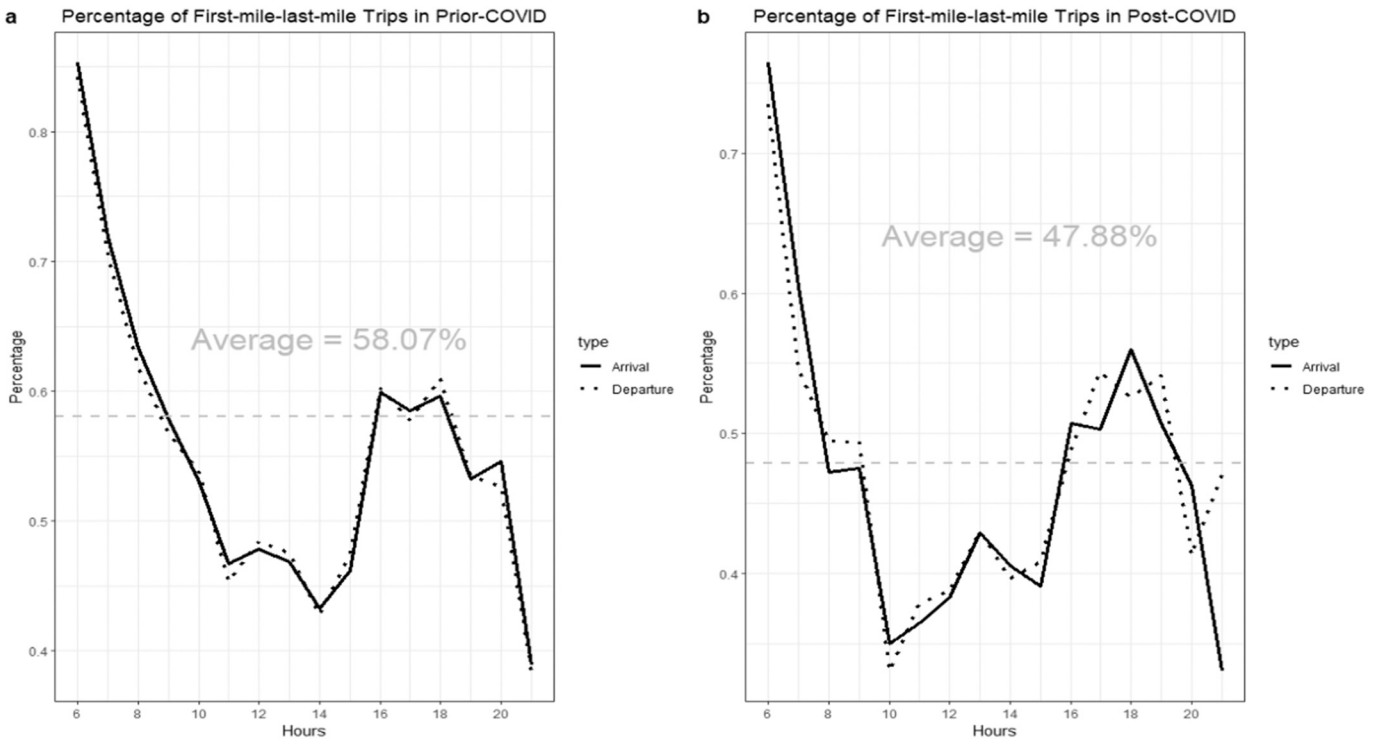

Fig. 10. Percentage of first mile/last mile trips.

0.143, 0.159. $I = 0$ represents the point when all 3-cliques remain in the network without any isolated node. $I = 0.031$ is the critical point. When $I = 0.159$, only one community remains in the network. The remaining seven values are interpolated at an increment of 0.016 from $I = 0.031$ to $I = 0.159$. We present these results in Table 3, along with the corresponding community attributes. When multiple communities have the same number of nodes, we break the tie by selecting the community with the highest trip rate. We select the two largest communities for illustrative purposes because their size ratio is an important indicator of community structure. Also, the two largest communities account for a majority ($\geq$ 85%) of the total trips. In Table 3, *Size* represents the number of nodes contained in the community. *Trips* are the indicator of the total amount of trips captured by the community. *FMLM* trips stand for first mile/last mile trips with a calculated percentage of total trips. *Users* are the unique number of customers who appeared across the trips. *Avg. Trips* equal to Trips divided by Users. *Duration* is the average trip length of all trips in the unit of minutes in that community.

In general, the evolution of communities follows a hierarchical structure. As values of I increase, 3-cliques with lower values of Intensity begin to disappear. However, When $I = 0$, all 3-cliques survive. Again, this turns the network into an unweighted and undirected graph. This scenario means that one can consider any link as only connected or disconnected. The communities detected when $I = 0$ are a good reflection of the connectivity across the region. There are 22 communities in the network, and the largest one is a gigantic community that contains 601 nodes. The remaining communities contain only 3 or 4 nodes (Table 3). It is worth mentioning that the remaining 22 communities include 626 nodes in total, leaving 537 nodes forming no 3-cliques. In other words, 537 nodes do not belong to any clique. These isolated nodes demonstrate a commuting pattern for *specific individuals*. In other words, trips concerning the 537 pick-up or drop-off locations are user-centric. The locations are not likely to be shared across different users.

When $I = 0.031$ (critical point), the gigantic component breaks down into two medium-sized communities (81 and 20). In this scenario, the community structure is strong, maximum information is present, yet it displays a distinguishable pattern. For example, Fig. 13 and Fig. 14 demonstrate the spatial distribution of the two largest communities and the shared nodes when $I = 0.031$. Fig. 13 suggests that the spatial pattern of the network includes a distribution of nodes from the entire region, accounting for 6 of 85.7% (6 of 7) of the major transit stations. However, a new spatial pattern emerges in Fig. 14. The second-largest community contains a congregation of 20 nodes from the west side of the service area. The corresponding trips gravitate to the Daybreak Parkway station, with 81% of trips being first mile/last mile trips. This result suggests that the pick-up and drop-off locations surrounding Daybreak Parkway station maintain a connection, which provides valuable information for possible vehicle dispatching and routing optimization.

Section 5.1 explored how first mile/last mile trips contribute to the spatio-temporal patterns of microtransit activities. Similarly, we observe a dominant percentage of first mile/last mile trips for different Is in Table 3. When $I = 0$, the largest community has 15,735 trips, with 62% being first mile/last mile trips, slightly higher than the average (58.07%). However, only 35% of trips occur outside of network communities, with the majority connecting isolated nodes. This difference suggests that first mile/last mile trips enhance the connectivity between popular pick-up and drop-off locations. Trips connecting isolated nodes are more user-specific and less dependent on transit. Further, at the critical point, first mile/last mile trips account for a large portion of the total trips in both the largest and second-largest communities with a minimum of 77% and a maximum of 100% (Table 3). Clearly, first mile/last mile trips play a crucial role in connecting the entire region.

We also explore 3-clique communities, their overlap, and levels of connectivity. We use the average clustering coefficient, C_{avg}, to measure local community cohesiveness. Without any processing, the C_{avg} of the original network is 0.169. This high value suggests significant overlap between communities in the pre-COVID period. Also, it means that when two communities overlap, they are likely to overlap with each other. Fig. 15 presents this trend graphically. As values of I increase, there is a removal of isolated nodes and 3-cliques with lower I values. This trend results in increasing values of C_{avg}, although there is some fluctuation. The results suggest that popular (i.e., large trip counts) pick-up and

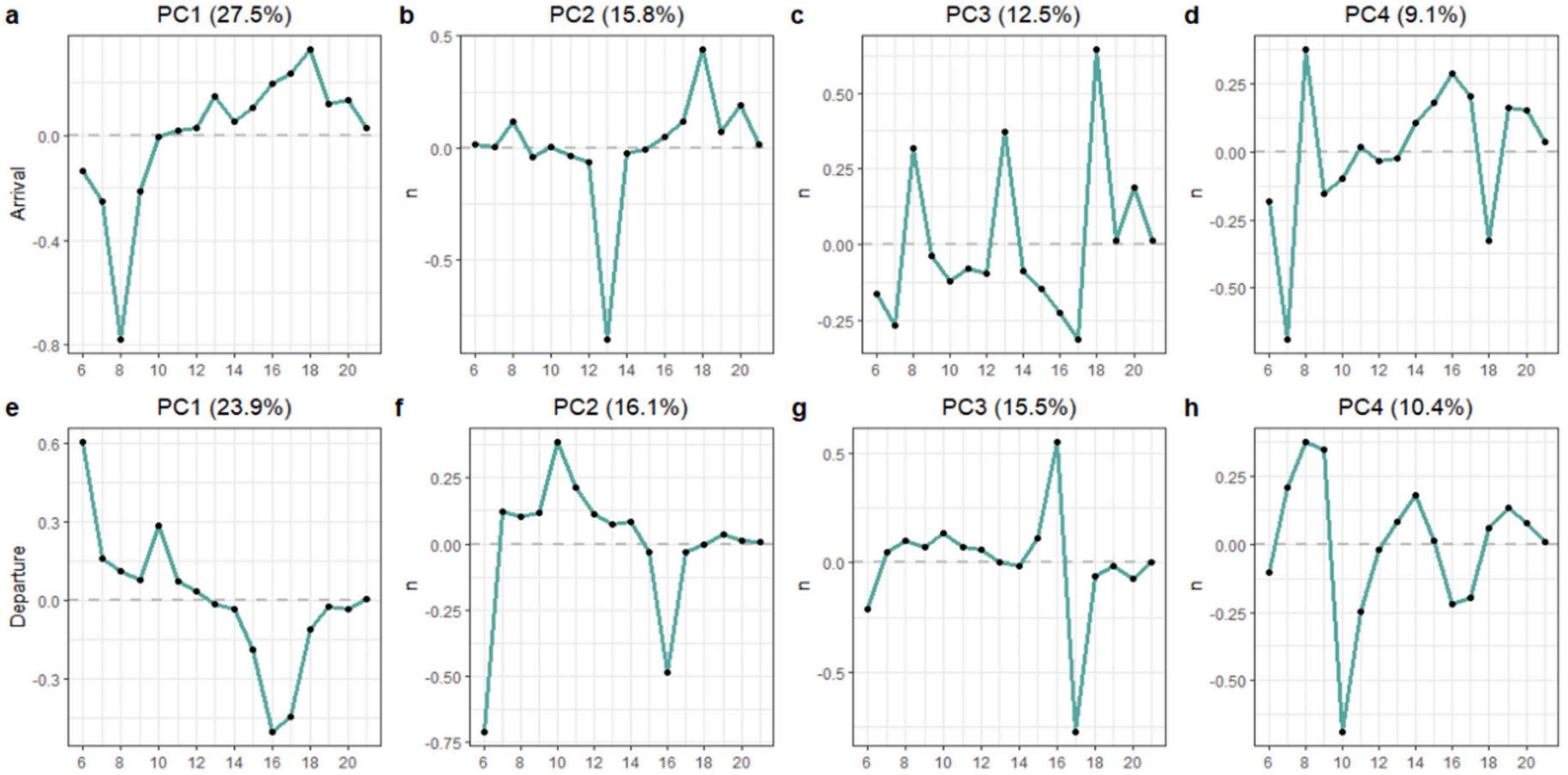

Fig. 11. Eigendecomposition of post-COVID arrival & departure trips.

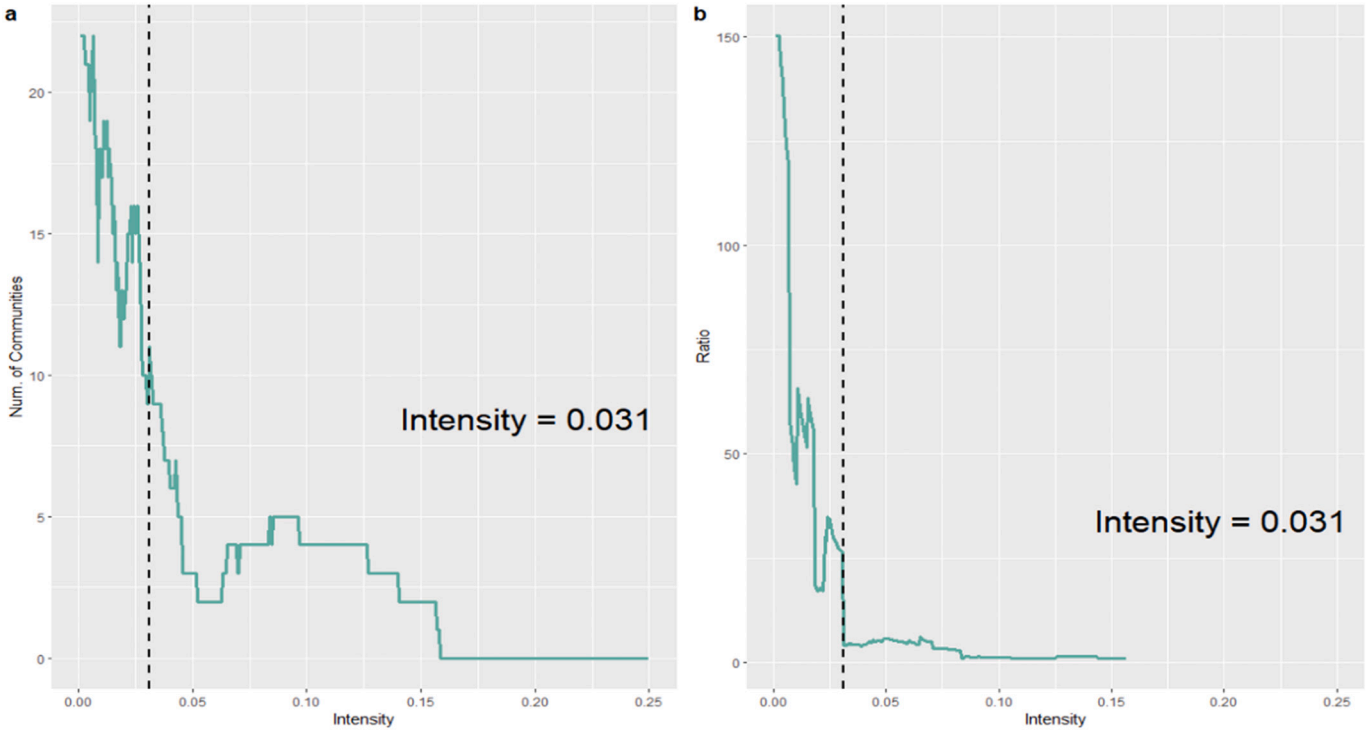

Fig. 12. Community summary under different Intensity thresholds: (a) the number of communities; and (b) the size ratio of the largest community to the second-largest community.

Table 3
The result of 3-clique percolation for the pre-COVID period.

I	Community	Size	Trips	FMLM Trips (% of total trips)	Users	Avg. Trips (trips/person)	Duration (minutes)
0	1	601	15,735	9818 (62%)	1039	15.1	10.3
	2	4	13	11 (85%)	5	2.6	12.3
0.031	1	81	4930	4330 (88%)	361	13.7	9.6
	2	20	699	564 (81%)	65	10.8	8.2
0.047	1	42	3128	2835 (91%)	237	13.2	8.8
	2	8	441	400 (91%)	38	11.6	8.6
0.063	1	26	2240	1992 (89%)	156	14.4	9.0
	2	6	336	319 (95%)	35	9.6	7.7
0.079	1	15	1410	1338 (95%)	117	12.4	9.4
	2	5	291	231 (79%)	16	18.2	6.0
0.095	1	5	471	466 (99%)	51	9.2	12.2
	2	4	321	306 (95%)	28	11.5	7.3
0.111	1	4	321	306 (95%)	28	11.5	7.3
	2	4	244	189 (77%)	11	22.18	6.2
0.127	1	4	321	306 (95%)	28	11.5	7.3
	2	3	409	409 (100%)	49	8.3	13.0
0.143	1	4	321	306 (95%)	28	11.5	7.3
	2	3	409	409 (100%)	49	8.3	13.0
0.159	1	3	239	224 (94%)	14	17.1	6.72
	2						

drop-off locations are tightly connected, resulting in high C_{avg}.

Community structure is also important to consider, especially at the critical point. When $I = \mathbf{0.031}$, there are nine shared nodes with clustering coefficients ranging from 0.017 to 0.32. Note that six out of these nine nodes with the lowest clustering coefficient correspond to major transit stations, while the top three correspond to an apartment, a private company, and a supermarket (Fig. 13 & Fig. 14). The average clustering coefficient of the six transit stations is 0.0275 - well below C_{avg} (0.169) of the original network. Moreover, among the six transit stations, Draper Station has the largest nodal degree, yet the smallest clustering coefficient, 0.017. In contrast, Kimballs Lane Station has the lowest nodal degree yet the largest clustering coefficient, 0.049. A similar pattern emerges for the top four nodes that are not transit stations. All have smaller nodal degrees and larger clustering coefficients, with a maximal of 0.32. This result suggests a different distribution pattern between transit stations vs. non-transit stations. For transit stations, a hub-and-spoke pattern best describes the system surrounding them. In short, traffic travels along the links connecting different nodes to a central distribution center (transit stations). However, those nodes themselves are not well connected. For apartments, private companies, and supermarkets, the underlying pattern better corresponds to a point-to-point structure. In these instances, riders can use microtransit to travel directly between those locations without going through a central hub.

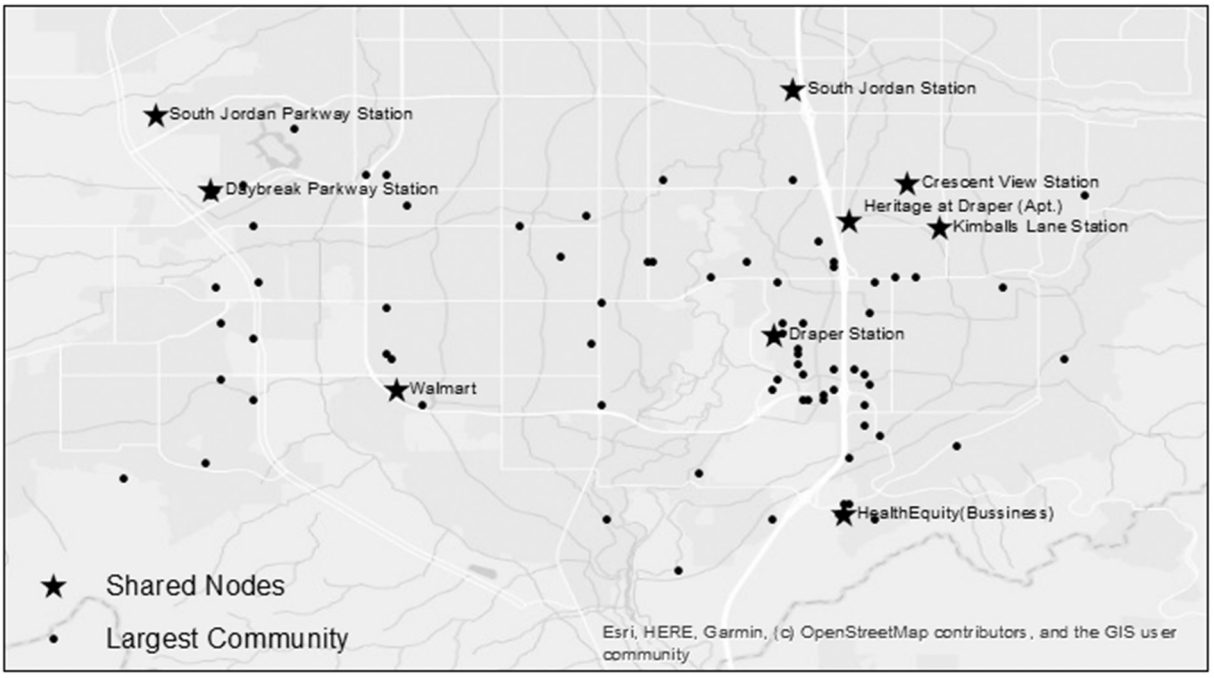

Fig. 13. The largest community and shared nodes when $I = 0.031$ in the pre-COVID period.

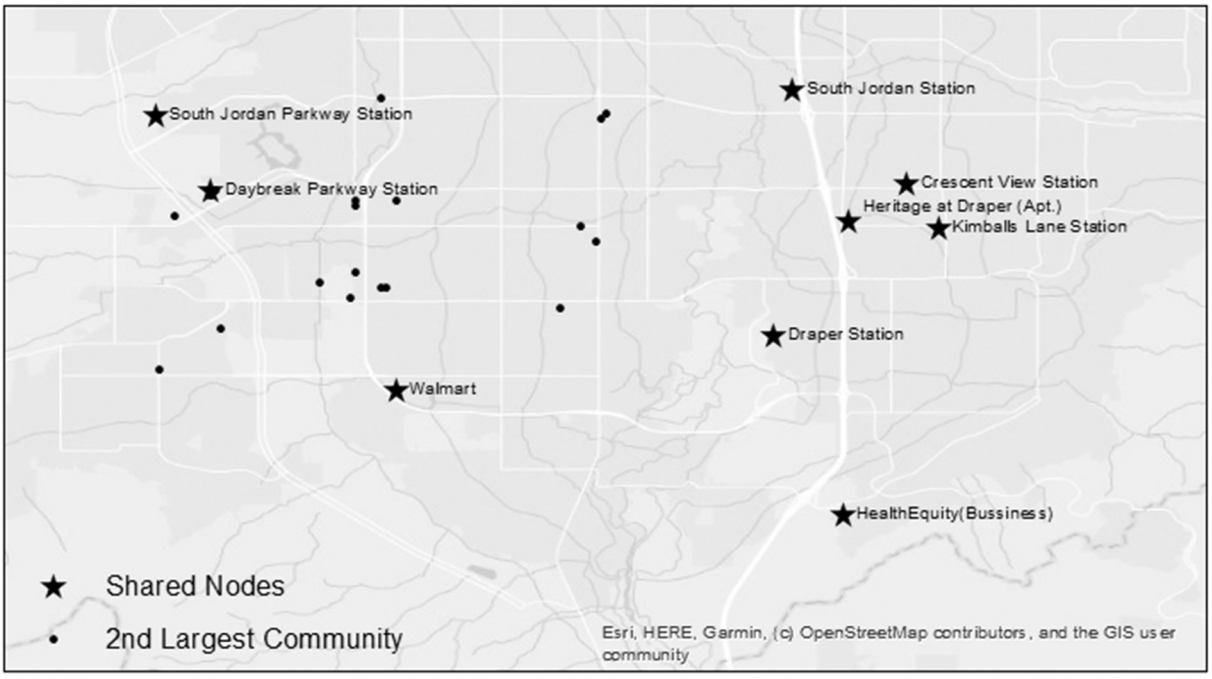

Fig. 14. The second-largest community and shared nodes when $I = 0.031$ in the pre-COVID period.

5.2.2. Post-COVID period

There are 13,188 trips by 794 unique users (334 new) between 1040 pick-up and drop-off locations in the post-COVID period. The maximum number of trips is 221, while the minimum number is 1. We present the results of the 3-clique percolation in Fig. 16. In Fig. 16 (a), the same decreasing pattern with fluctuations is apparent. In Fig. 16 (b), the ratio is well above 2 when I is small, but the value has dropped slightly because of less microtransit activity in general. The vertical dashed line stands for the critical point in the post-COVID period.

The characteristics of communities is further explored at ten values of I in the same manner: 0, 0.066, 0.087, 0.108, 0.129, 0.150, 0.171, 0.192, 0.213, 0.234. Table 4 displays the results.

When $I = 0$, the network yields a gigantic component that contains 492 nodes while the remaining communities contain 6 or fewer nodes and 530 nodes in isolation. Compared to the pre-COVID period, there are more isolated nodes (51% vs. 46%). This result indicates that the user-centric pattern has become more distinct after the outbreak of COVID-19.

When $I = 0.066$ (critical point), two medium-sized communities (25 and 18) emerge from the gigantic network. Fig. 17 and Fig. 18 present

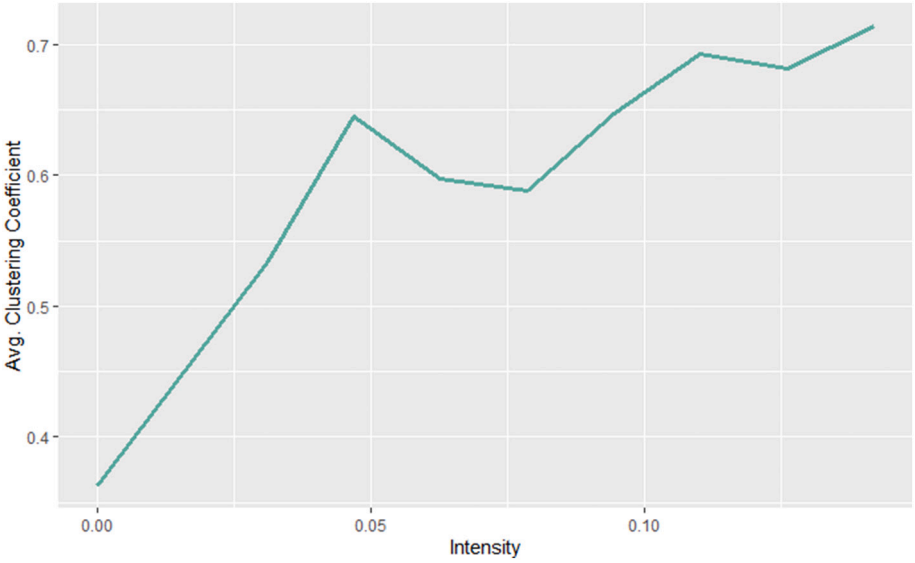

Fig. 15. The average clustering coefficient at different ***I***.

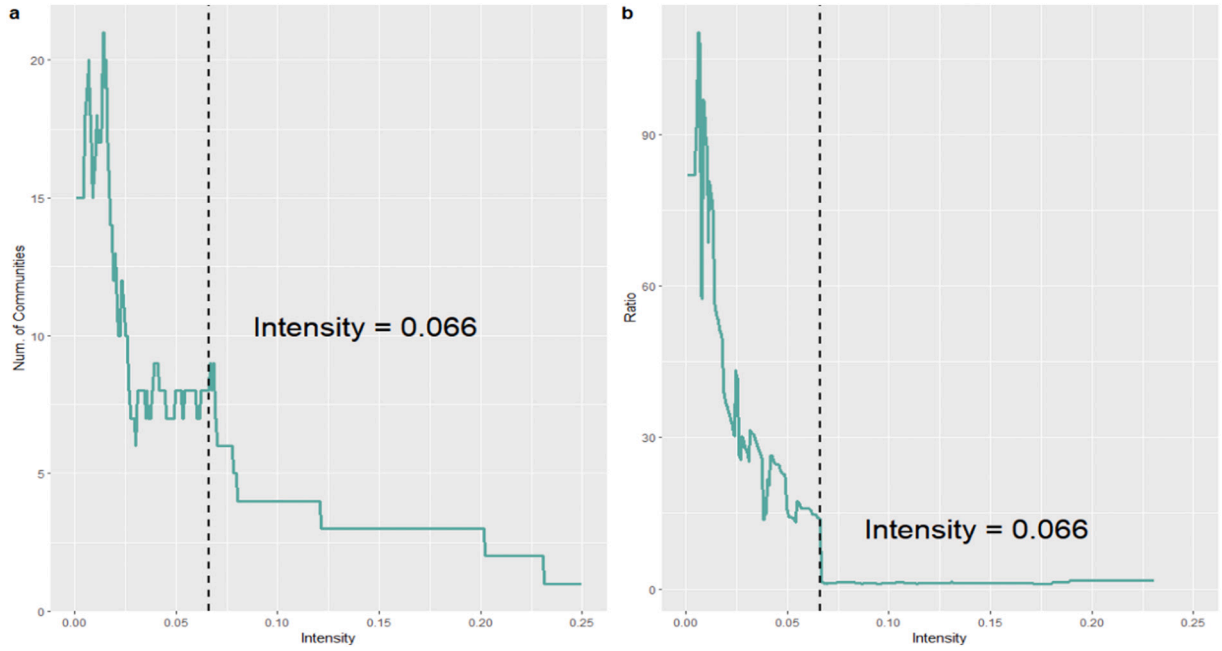

Fig. 16. Community summary under different *Intensity* thresholds in the post-COVID period: (a) the number of communities; and (b) the size ratio of the largest community to the second-largest community.

the spatial distribution of both and the shared nodes. There are some significant spatial changes worth noting. First, the spatial distributions of the two communities became less distinct. The largest community contains Draper Station, Draper Town Center Station, and Crescent View Station. The second community contains Draper Station, Crescent View Station, South Jordan Station, and Daybreak Parkway Station. The major takeaway here is that both communities span around major transit stations across the region. Also, their *Sizes* and *Trips* (in Tables 2 & 3) are similar, suggesting that these two communities are equally important in the post-COVID period.

In addition, for all values of ***I***, the number of users has decreased. This result is likely due to the overall decline in microtransit activity. However, the average number of trips per user increased compared to the pre-COVID period (Tables 2 and 3). There are two possible explanations. One is that less frequent users might have stopped using microtransit due to pandemic concerns. The other is that those who depend on microtransit remain riders (e.g., transit dependents and essential workers). Aside from these differences, similarities also emerge. For example, many trips remain in the network for different values of ***I,*** and the first mile/last mile trips account for a large percentage of total trips for both periods.

We are also interested in how the overlapping pattern changes. Fig. 19 presents the average clustering coefficients for different values of ***I***. The C_{avg} of the original network decreased from 0.169 (pre-COVID) to 0.129 (post-COVID). This result indicates that the level of network overlap decreased. At the critical point, we identify eight shared nodes. The three with the lowest clustering coefficient are transit stations with a minimal clustering coefficient of 0.029 (Draper Station) and a maximal clustering coefficient of 0.041 (Crescent View Station). The top five nodes contain four supermarkets and one residential node (Fig. 18) with

Table 4
The 3-clique percolation for the post-COVID period.

I	Community	Sizes	Trips	FMLM Trips (% of total trips)	Users	Avg. Trips (trips/person)	Duration (minutes)
0	1	492	10,948	5737 (52%)	610	17.9	9.5
	2	6	9	0 (0%)	2	4.5	8.4
0.066	1	25	1549	1520 (98%)	82	18.9	6.3
	2	18	1042	919 (88%)	81	12.9	24.3
0.087	1	11	820	809 (99%)	38	21.6	6.3
	2	11	817	712 (87%)	50	16.3	5.3
0.108	1	9	750	748 (99%)	30	25.0	6.3
	2	7	630	621 (99%)	35	18	5.3
0.129	1	7	630	621(99%)	35	18.0	5.3
	2	6	517	516 (99%)	24	21.5	6.3
0.150	1	6	601	600(99%)	34	17.7	5.3
	2	5	468	468 (100%)	18	26.0	6.3
0.171	1	6	601	600 (99%)	34	17.7	5.3
	2	5	468	468 (100%)	18	26.0	6.3
0.192	1	5	551	550 (99%)	30	18.4	5.3
	2	3	162	162 (100%)	13	12.5	5.2
0.213	1	5	551	550 (99%)	30	18.4	5.3
	2	3	157	157 (100%)	11	14.3	7.2
0.234	1	5	551	550 (99%)	30	18.4	5.3
	2						

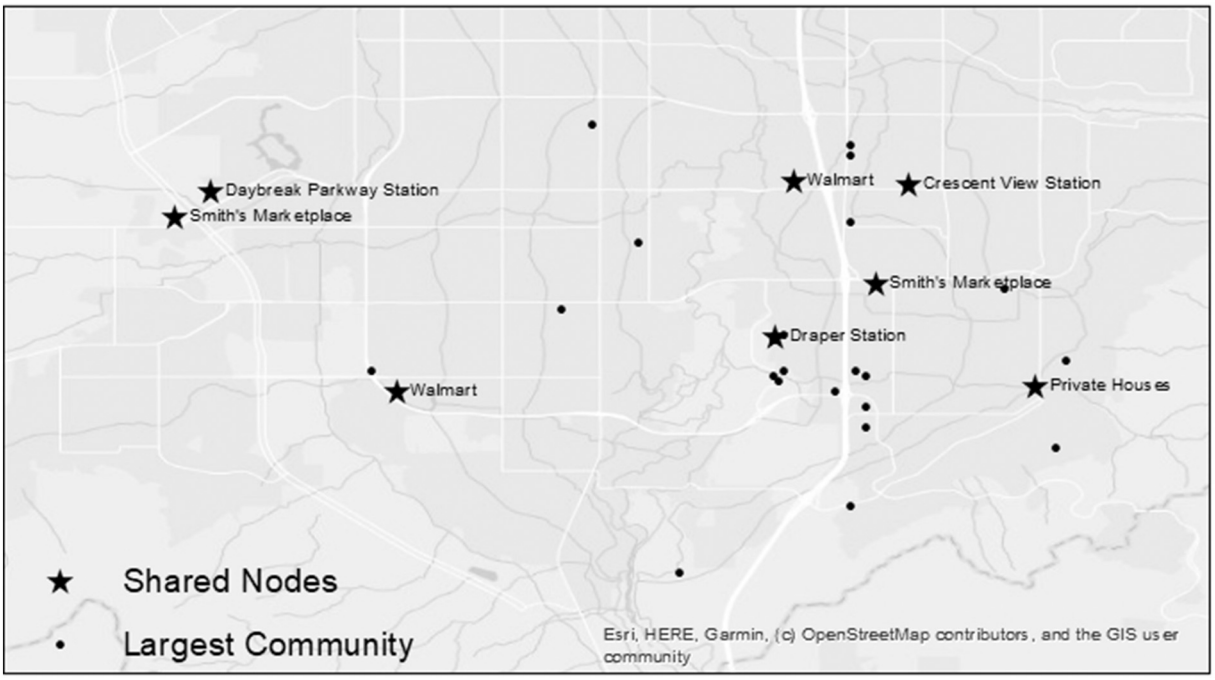

Fig. 17. The largest community and shared nodes when $I = 0.066$ in the post-COVID period.

a minimum coefficient of 0.073 and a maximal coefficient of 0.124. The hub-and-spoke distribution for transit stations and point-to-point distribution for supermarkets and residential locations were also present in the post-COVID period.

6. Conclusions

The worldwide expansion of microtransit continues, bringing many opportunities and challenges for service providers. As a new demand-responsive transport mode, it has the advantage of increasing service coverage, flexible routing, and enhancing transit accessibility. With proper design and execution, microtransit can supplement traditional transport methods to help reduce road pressure and create first mile/last mile connections. However, microtransit projects often suffer from poor marketing and operational efforts, making them less cost-effective than they should be. Further, there is limited information concerning the spatio-temporal patterns of microtransit activities. This research fills the gap by unraveling the spatio-temporal structures of microtransit activities utilizing empirical data from a pilot program in the Salt Lake City, Utah area, UTA Via. The results compare microtransit activity during pre- and post-COVID periods, highlighting how some use patterns persist and change after the outbreak.

The results of the eigendecomposition suggest that first mile/last mile trips declined, but the hourly distribution remained nearly identical. This finding suggests transit dependency for many riders. However, several new patterns emerged (e.g., Fig. 9 and Fig. 11). Specifically, the explanatory power of PC1 decreased. This finding suggests a temporal dispersion trend for microtransit activity after the pandemic. Furthermore, by focusing on first mile/last mile trips, our results suggest that a higher percentage of first mile/last mile trips leads

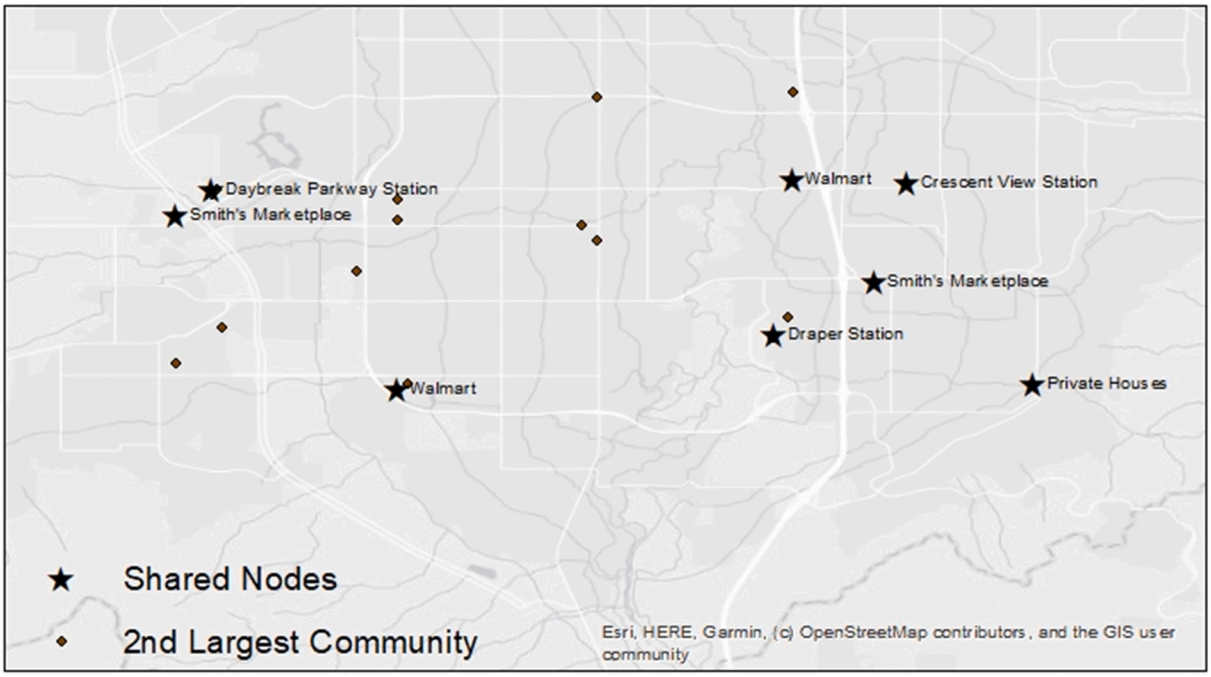

Fig. 18. The second-largest community and shared nodes when $I = \mathbf{0.066}$ in the post-COVID period.

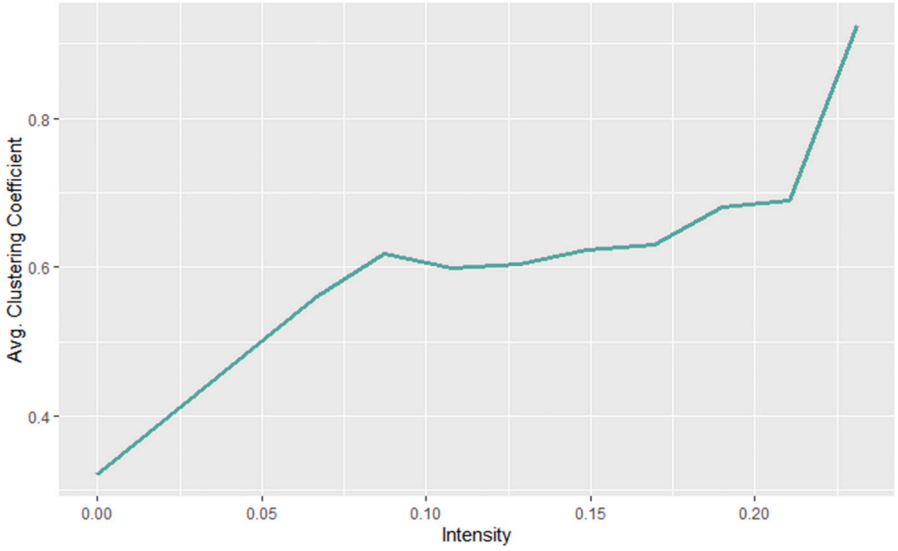

Fig. 19. The average clustering coefficient at different I.

to greater variation for PC1. This result means that first mile/last mile trips can be the major source of variation in both periods.

We further apply k-clique percolation to identify the communities in two networks (pre-COVID and post-COVID). We vary threshold values of I for each network. As I increases, the gigantic component begins to break down into medium-sized communities. During this process, the results suggest that first mile/last mile trips account for most of the total trips captured by the network. However, Tables 2 and 3 suggest a decrease in unique users and a heavy dependency on microtransit for specific users.

The application of a clustering coefficient revealed that network overlap decreased after the COVID-19 outbreak. Interestingly there was an emergence of two distinct communities, pre- and post-COVID. The communities surrounding transit stations take the form of hub-and-spoke systems, with transit stations serving as traffic distribution centers while the surrounding nodes are disconnected. For other popular locales, connections are more dispersed, demonstrating a point-to-point distribution style.

This work aimed to reconstruct and understand the spatio-temporal patterns of microtransit activity in portions of Salt Lake City, Utah. The framework is generalizable and can provide additional insights for UTA Via as it grows or inspires applications to pilot programs in other cities. By understanding the patterns and possible causal factors for microtransit network development, use, and underlying spatio-temporal patterns, one can enhance the transferability of microtransit programs without additional cost. Furthermore, spatio-temporal structures of microtransit usage reveal that the usage is uneven. For example, the connection between certain regions is significantly stronger than that of other places. Also, the two-peak temporal pattern demonstrates many variations resulting from first mile/last mile trips. Understanding these

structures of microtransit activity can help with possible customer segmentation and vehicle dispatching for all microtransit programs. Lastly, by comparing results between pre- and post-COVID periods, it is possible to inform transit agencies on people's behavioral changes and the evolution of their travel patterns to guide operational strategy adjustments further.

Acknowledgments

This article is based upon work partially supported by the National Science Foundation under Grant No. **2051226, and partially supported by** the Mountain-Plains Consortium (MPC) of the U.S. Department of Transportation University Transportation Center (MPC-608). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s).

References

Abdi, H., Williams, L.J., 2010. Principal component analysis. WIREs Comp Stat. 2, 433–459.
Achlioptas, D., D'Souza, R.M., Spencer, J., 2009. Explosive percolation in random networks. Science 323, 1453–1455.
Agatz, N., Erera, A.L., Savelsbergh, M.W., Wang, X., 2011. Dynamic ride-sharing: a simulation study in metro Atlanta. Procedia Soc. Behav. Sci. 17, 532–550.
Agatz, N., Erera, A., Savelsbergh, M., Wang, X., 2012. Optimization for dynamic ride-sharing: a review. Oper. Res. 223, 295–303.
Aissat, C., Oulamara, A., 2014. A priori approach of real-time ride-sharing problem with intermediate meeting locations. J. Artif. Intellig. Soft Comp. Res. 4.
Alonso-González, M.J., Liu, T., Cats, O., Van, N., Hoogendoorn, S., 2018. The potential of Demand-responsive transport as a complement to public transport: an assessment framework and an empirical evaluation. Transp. Res. Rec. 2672 (8), 879–889.
Barrat, A., Barthélemy, M., Pastor-Satorras, R., Vespignani, A., 2004. The architecture of complex weighted networks. Proc. Natl. Acad. Sci. 101 (11), 3747–3752.
Broadbent, S.R., Hammersley, J.M., 1957. Percolation processes: I. Crystals and mazes. In: Mathematical proceedings of the Cambridge philosophical society (Vol. 53, No. 3, pp. 629-641. Cambridge University Press.
Chen, Z., Liu, X.C., Wei, R., 2019. Agent-based approach to analyzing the effects of dynamic ride-sharing in a multimodal network. Comput. Environ. Urban. Syst. 74, 126–135.
Derényi, I., Palla, G., Vicsek, T., 2005. Clique percolation in random networks. Phys. Rev. Lett. 94 (16), 160202.
Dillman, M., Posvistak, C., 2020. COVID-19 and Public Transportation in Utah. Inquiry of the Public Sort.
Dong, Y., Wang, S., Li, L., Zhang, Z., 2018. An empirical study on travel patterns of internet based ride-sharing. Transp. Res. Part C: Emerg. Technol. 86, 1–22.
Federal Transit Administration (FTA), 2019. National Transit Summaries and Trends. https://www.transit.dot.gov/sites/fta.dot.gov/files/2020-12/2019-NTST-1-1_0.pdf.
Federal Transit Administration (FTA), 2021. https://www.transit.dot.gov/regulations-and-guidance/shared-mobility-definitions.
Fortunato, S., 2010. Community detection in graphs. Phys. Rep. 486 (3–5), 75–174.
Gao, J., Wang, J., Bian, Z., Ban, X.J., 2020a. The effects of the COVID-19 pandemic on transportation systems in new york city and seattle, USA. In: arXiv preprint. arXiv:2010.01170.
Gao, J., Bernardes, S.D., Bian, Z., Ozbay, K., Iyer, S., 2020b. Initial impacts of COVID-19 on transportation systems: a case study of the US epicenter, the New York metropolitan area. In: arXiv preprint. arXiv:2010.01168.
Goffri, S., Muller, C., Stingelin, N., Breib, D.W., Radano, C.P., Andreasen, J.W., Thompson, R., Janssen, R.A.J., Nielsen, M.M., Smith, P., Sirringhaus, H., 2006. Multicomponent semiconducting polymer systems with low crystallization-induced percolation threshold. Nat. Mater. 5, 950–956.
Grimmett, G., 1999. In: What is percolation?. In Percolatio. Springer, Berlin, Heidelberg, pp. 1–31.
Haglund, N., Mladenović, M.N., Kujala, R., Weckström, C., Saramäki, J., 2019. Where did Kutsuplus drive us? Ex post evaluation of on-demand micro-transit pilot in the Helsinki capital region. Res. Transp. Bus. Manag. 32, 100390.
Hoang-Tung, N., Kojima, A., Kubota, H., 2015. Travelers' motivations to use public transportation: a changing strength of pro-environmental motivation in deciding bus use intention. J. East. Asia Soc. Transp. Stud. 11, 1522–1534.
Horner, M.W., Grubesic, T.H., 2001. A GIS-based planning approach to locating urban rail terminals. Transportation 28 (1), 55–77.
Jiang, Y., Timmermans, H.J.P., Yu, B., 2018. Relocation of manufacturing industry from the perspective of transport accessibility – an application of percolation theory. Transp. Policy 63, 10–29.
Jolliffe, I.T., 2002. Choosing a subset of principal components or variables. In: Principal Component Analysis. Springer, pp. 111–149.
Li, M., Deng, Y., Wang, B.H., 2015a. Clique percolation in random graphs. Phys. Rev. E 92 (4), 042116.
Li, D., Fu, B., Wang, Y., Lu, G., Berezin, Y., Stanley, H.E., Havlin, S., 2015b. Percolation transition in dynamical traffic network with evolving critical bottlenecks. Proc. Natl. Acad. Sci. 112, 669–672.
Liu, L., Miller, H.J., Scheff, J., 2020. The impacts of COVID-19 pandemic on public transit demand in the United States. PLoS One 15 (11), e0242476.
Majdandzic, A., Podobnik, B., Buldyrev, S.V., Kenett, D.Y., Havlin, S., Stanley, H.E., 2014. Spontaneous recovery in dynamical networks. Nat. Phys. 34-38.
McKibbin, W., Fernando, R., 2020. The economic impact of COVID-19. In: Economics in the time of COVID-19, 45, 10.1162.
Melman, T., Abbink, D., Mouton, X., Tapus, A., Joost De, W., 2021. Multivariate and location-specific correlates of fuel consumption: a test track study. Transp. Res. Part D: Transp. Environ. 92, 102627.
Murtagh, F., Contreras, P., 2012. Algorithms for hierarchical clustering: an overview. WIREs Data Min. Knowl. Discov. 2, 86–97.
Nagendra, S.M., Khare, M., 2003. Principal component analysis of urban traffic characteristics and meteorological data. Transp. Res. Part D: Transp. Environ. 8 (4), 285–297.
Ongel, A., Loewer, E., Roemer, F., Sethuraman, G., Chang, F., Lienkamp, M., 2019. Economic assessment of autonomous electric microtransit vehicles. Sustainability 11, 648.
Pachal, P., 2018. Express Pool: cheaper rides, as long as you don't mind walking. In: Mashable. Accessed on: march 20th, 2021. https://mashable.com/2018/02/21/uber-express-pool/.
Palla, G., Derényi, I., Farkas, I., Vicsek, T., 2005. Uncovering the overlapping community structure of complex networks in nature and society. Nature 435, 814–818.
Palla, G., Derényi, I., Vicsek, T., 2006. The critical point of k-clique percolation in the Erdős–Rényi graph. J. Stat. Phys. 128 (1–2), 219–227.
Pew Research Center, 2020. Most Americans Say Coronavirus Outbreak Has Impacted Their Lives. Accessed on: August 15th, 2021. https://www.pewsocialtrends.org/2020/03/30/most-americans-say-coronavirus-outbreak-has-impacted-their-lives/.
Pragholapati, A., 2020. COVID-19 Impact on Students.
Schmitt, A., 2018. The story of "micro transit" is a consistent, dismal failure. In: Streetblog. Accessed on: April 4th , 2021. https://usa.streetsblog.org/2018/06/26/the-story-of-micro-transit-is-consistent-dismal-failure/.
Tan, S., Fowers, A., Tierney, D.K.L., 2020. Amid the pandemic, public transit is highlighting inequalities in cities. In: Washington Post. Available: https://www.washingtonpost.com/nation/2020/05/15/amid-pandemic-public-transit-is-highlighting-inequalities-cities/?arc404=true.
UTA On Demand, UTA, 2021. Accessed on: May 8th, 2021. https://www.rideuta.com/Services/UTA-On-Demand.
Volinski, J., 2019. Microtransit or general public Demand–response transit services: state of the practice. Transp. Res. Board 2019.
Wang, C.H., Chen, N., Tian, G., 2021. Do accessibility and clustering affect active travel behavior in Salt Lake City? Transp. Res. Part D: Transp. Environ. 90, 102655.
Wei, R., Liu, X., Ou, Y., Fayyaz, S.K., 2018. Optimizing the spatio-temporal deployment of battery electric bus system. J. Transp. Geogr. 68, 160–168.
Westervelt, M., Huang, E., Schank, J., Borgman, N., Fuhrer, T., Peppard, C., Narula-Woods, R., 2018. UpRouted: Exploring Microtransit in the United States. Eno Center for Transportation.
Wilbur, M., Ayman, A., Ouyang, A., Poon, V., Kabir, R., Vadali, A., Dubey, A., 2020. Impact of COVID-19 on public transit accessibility and ridership. In: arXiv preprint. arXiv:2008.02413.
WMATA, 2020. Metro and Covid-19: Steps we've taken. Accessed on: August 18th, 2021. https://www.wmata.com/service/status/details/COVID-19.cfm.
Xiong, J., Lipsitz, O., Nasri, F., Lui, L.M., Gill, H., Phan, L., McIntyre, R.S., 2020. Impact of COVID-19 pandemic on mental health in the general population: a systematic review. J. Affect. Disord. 277, 55–64.
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. Comput. Environ. Urban. Syst. 75, 184–203.
Yi, Z., Liu, X.C., Wei, R., Chen, X., Dai, J., 2021. Electric vehicle charging demand forecasting using deep learning model. J. Intell. Transp. Syst. 1–14.
Zhou, Y., Liu, X.C., Wei, R., Golub, A., 2020. Bi-objective optimization for battery electric bus deployment considering cost and environmental equity. IEEE Trans. Intell. Transp. Syst. 22 (4), 2487–2497.