TRR
JOURNAL OF THE TRANSPORTATION RESEARCH BOARD

Research Article

Transportation Research Record
2020, Vol. 2674(1) 260–271

DOI: 10.1177/0361198119899389
journals.sagepub.com/home/trr
SAGE

Investigating Complementary and Competitive Relationships between Bikeshare Service and Public Transit: A Spatial-Temporal Framework

Ying Song[1] and Yuchuan Huang[1]

Abstract
Public transit offers many socioeconomic and environmental benefits but often suffers from the first/last-mile problem. The emergence of bike-sharing systems promises to provide first/last-mile access to transit stops/stations and increase the use of transit. However, it is simply unknown whether bikeshare can help transit, or whether it may instead compete with transit by offering people another travel mode to avoid the waiting time entirely. Recent studies have examined the spatial relationship between bikeshare and transit but have ignored important temporal aspects such as the timing of planned trips and transit schedules. This paper develops a framework with procedures and methods for investigating the potential competitive and complementary relationships between bikeshare and transit systems from a spatial-temporal perspective. The paper applies this framework to the Nice Ride bikeshare service and Metro Transit in the Twin Cities, Minnesota, as a case study. The results suggest both complementary and competitive relationships that are not exclusive from each other. The general patterns vary across different regions in the study area and are affected by the underlying neighborhood characteristics. The results provide novel insights to the complex interactions between bikeshare and transit systems and can support operation and planning practices.

Public transit offers many societal and environmental benefits: it promises to mitigate auto-dependency and promote sustainable transportation development while providing access to various resources and opportunities (*1–3*). A critical challenge in the effective provision of public transit is the first-mile and last-mile problem. In essence, people who want to choose transit often do not do so because the distance to the bus stop or train station is too great (*4, 5*). At the same time, it is too costly or impractical to invest in and build transit systems that are close to all homes, jobs, schools, and other destinations.

Bikeshare is an on-demand transportation service that allows users to access and use bicycles for a fee (*6*). Integrating bike-sharing services with existing transit systems can potentially solve the first/last-mile problem, expand service areas and times of current transit systems, and promote public health (*7–10*). Studies have indicated positive impacts of bikeshare on transit worldwide (*10–14*). Yet, some surveys and studies also illustrate an unfortunate irony—bikeshare may act as a competitor to transit instead of complementing it (*15, 16*). Thus, the relationship between bikeshare and transit is complex, and it is important to investigate this relationship and its effects on the current assessment and future planning of multimodal transportation systems.

Recently, studies have adopted various methods to advance our understanding on how bikeshare services interact with transit systems. When bikeshare and transit use the same smart card system, studies select transactions for bikeshare and transit trips and investigate users' travel behavior (*17*). However, not all regions have this type of integrated payment system and not all users use a smart card to pay for their trips. Therefore, a few studies have conducted surveys to understand users' preference on travel modes (*10, 14*, *17*), or focus on the spatial distributions of bikeshare and transit systems and use ridership data to examine modal shift patterns (*15, 18–22*). These studies provide useful insights into the relationship between transit and bikeshare. However, they often ignore the vitally important temporal aspect, in that

[1]Geography, Environment and Society, University of Minnesota – Twin Cities, Minneapolis, MN

Corresponding Author:
Ying Song, yingsong@umn.edu

most transit users have a personal schedule that must dovetail with transit schedules, and bikeshare users must face the additional challenge of building this into riding time.

This paper develops a framework from a spatial-temporal perspective to systematically investigate interactions between bikeshare and transit systems. The paper defines complementary and competitive relationships between bikeshare and transit systems considering both their spatial locations and temporal profiles, and applies different methods to examine their interactions. For the competitive relationship, the paper matches each bikeshare trip with transit trips that have similar spatial and temporal profiles. For the complementary relationship, the paper groups all bikeshare and transit stations/stops based on their spatial locations and investigates the patterns of bikeshare and transit usages within each subgroup. To illustrate the framework and methods, the paper uses Nice Ride bikeshare and Metro Transit in the Twin Cities, USA as a study case. The paper also discusses potential applications and future research directions.

Background

K-*Nearest Neighbor*

The *k*-nearest neighbors (k-NN) classification is a non-parametric method for classification that assigns the class to an unclassified observation based on its *k*-nearest neighbors in the training set (*23*). Compared with other classification methods, the *k*-NN classification does not require a prior defined/derived classifier to be used for all incoming unclassified observations (*24, 25*). It was first developed by Fix and Hodges to analyze a discriminant when no reliable parametric estimates of probability densities are available (*26*), and has been modified for various applications such as gene expression in biology (*27, 28*), image processing in computer science (*29, 30*), and spatial analysis in geographic information science (*31, 32*).

The *k*-NN classifier is based on the specified training samples and some similarity matrix. Let x_i be an input new sample with p features, $\mathrm{x}_i = (x_{i1}, x_{i2}, \ldots, x_{ip})_{i=0,1,\ldots,n}$ and x_j be a sample in the training set, $S = \{(x_{j1}, x_{j2}, \ldots, x_{jp})\}_{j=0,1,\ldots,N}$, the similarity of x_i and x_j is defined by a distance function $\mathrm{d}(\mathrm{x}_i, \mathrm{x}_j)$ such as Euclidian and Manhattan distances. The *k*-nearest neighbors of x_i are the k samples in the training set NN_i whose distances to x_i are smaller than any of the rest $(N-k)$ samples $(S-NN_i)$.

$$\max\{\mathrm{d}(\mathrm{x}_i, \mathrm{x}_j), \mathrm{x}_j \in NN_i\} \leqslant \min\{\mathrm{d}(\mathrm{x}_i, \mathrm{x}_j), \mathrm{x}_j \in (S-NN_i)\} \tag{1}$$

After finding the *k*-nearest neighbors of an input new sample, the *k*-NN classifier labels the input sample with its nearest neighbors or adopts some decision rule to assign a class to the input sample based on the classes of its nearest neighbors. An example is to find the five closest restaurants and show them on the map to the user, which is a common task in location-based services. This paper applies the *k*-NN classification algorithm to search for transit stops/stations near bikeshare stations and create subgraphs of stops/stations in two systems that are well connected (within the threshold distance).

Cosine Similarity

The cosine similarity is a measure of similarity between two non-zero vectors based on the cosine of the angle between the two vectors. It has been applied to various fields such as image retrieval (*33, 34*), text mining (*35, 36*), and location analysis (*37, 38*) that need to deal with multiple attributes simultaneously. Let A_1 and A_2 be two vectors of attributes in the *n*-dimensional space, the cosine similarity, $\cos(\theta)$, can be represented using the Euclidean dot product and magnitude:

$$\begin{aligned}\cos(\theta) &= \cos(\mathrm{A}_1, \mathrm{A}_2) = \frac{\mathrm{A}_1 \cdot \mathrm{A}_2}{\|\mathrm{A}_1\|\|\mathrm{A}_2\|} \\ &= \frac{\sum_{i=1}^{n} \mathrm{A}_{1i}\mathrm{A}_{2i}}{\sqrt{\sum_{i=1}^{n} \mathrm{A}_{1i}^2}\sqrt{\sum_{i=1}^{n} \mathrm{A}_{2i}^2}}\end{aligned} \tag{2}$$

The cosine similarity is usually used for positive space where attributes A_{1i} and A_{2i} always take positive values. If A_1 and A_2 have the same orientation/direction, $\cos(\theta) = 1$ and the two vectors are most similar to each other. If A_1. and A_2 are orthogonal/perpendicular to each other, $\cos(\theta) = 0$ and the two vectors are viewed as maximally dissimilar. The cosine similarity is independent of the magnitude of the two vectors, and therefore provides great flexibility while measuring similarity between two vectors of attributes. This paper applies cosine similarity to compare bikeshare and transit usages across different periods of the day, and uses it to investigate the potential effects of bikeshare services on transit ridership.

Methodology

The paper considers both the complementary and competitive relationship between bikeshare and transit systems, and examines the relationship from a spatial-temporal perspective. Figure 1 shows the general framework with major steps and methods. Please note that the complementary and competitive relationships are not exclusive; that is, a bikeshare trip can support first/last-mile travel (supplementary) and replace transit trip

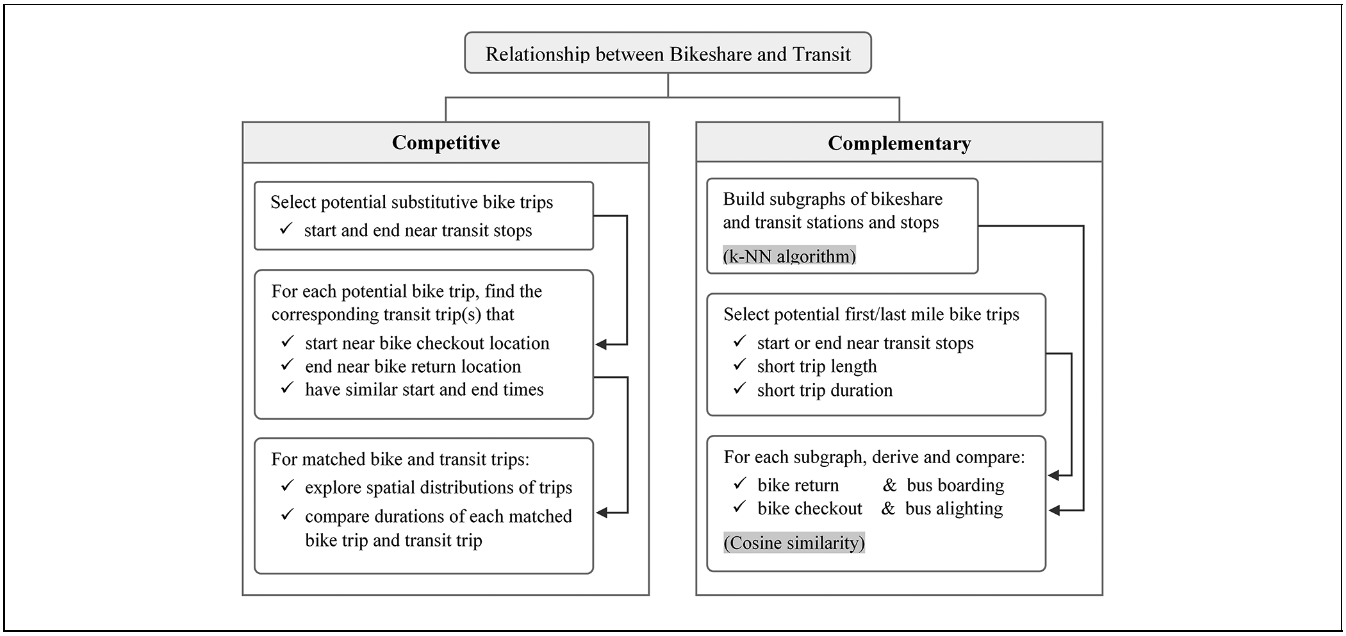


Figure 1. A framework to examine relationship of bikeshare and transit usages in space and time.

(competitive) at the same time. For instance, the user may use the bikeshare service to transfer between a transit station and a bus stop, whereas the user can also use the local bus to make the transfer. In this case, the bikeshare trip provides last-mile access for the first transit trip and first-mile access for the second bus trip (supplementary), and also replaces the transfers using buses because of its flexibility (competitive).

Competitive Relationship

For the competitive relationship, the analysis starts with spatial matching of stop and station locations and continues with temporal matching of the recorded bikeshare and transit trips. If a bikeshare trip starts and ends near a transit stop/station and has a similar temporal profile as a recorded transit trip, the bikeshare trip is considered to be a competitive trip. Figure 2*a* shows the criteria for spatial and temporal matching. Analytically, a bikeshare trip BT_i is considered as a potential substitution of a transit trip TT_k if:

1) bikeshare trip origin BS_{i1} is near transit trip origin TS_{k1}, $d(BS_{i1}, TS_{k1}) \leqslant \text{maxDist}$
2) bikeshare trip destination BS_{i2} is near transit trip destination TS_{k2}, $d(BS_{i1}, TS_{k1}) \leqslant \text{maxDist}$
3) bikeshare and transit trips start with similar time, $|BS_{i1}(t) - TS_{k1}(t)| \leqslant \text{maxDuration}$

AND bikeshare and transit trips end with similar time, $|BS_{i2}(t) - TS_{k2}(t)| \leqslant \text{maxDuration}$ After finding all pairs of bikeshare and transit trips, we first visualize the spatial distribution of potential competitive bike trips to identify popular origins and destinations for such trips. Then, trip durations are compared for each of the matching pairs, $\Delta t = (BS_{i2}(t) - BS_{i1}(t)) - (TS_{k2}(t) - TS_{k1}(t))$, to see whether users choose to use the bikeshare service because it takes less time than using public transit.

Complementary Relationship

For the complementary relationship, the analysis starts with classifying stops/stations into well-connected subgraphs, continues with selecting all potential first/last-mile trips, and ends with comparing bikeshare and transit usages within each subgraph. To define subgraphs, we first search for the k-nearest transit stops/stations $\{TS_k\}_i$ of each bikeshare station BS_i and create a virtual edge $e(BS_i, TS_k)$ between BS_i and one of its nearest neighbors TS_k. Then, each subgraph is defined as a subset of bikeshare stations and transit stops/stations that is linked by virtual edges and separated from all other stops/stations.

After defining subgraphs, the potential first/last-mile bike trips are selected. Figure 2*b* and *c* show the criteria conceptually. A bikeshare trip is consider a potential complementary if it starts or ends near a transit stop/station and has relatively short distance and short duration. Analytically, a bikeshare trip BT_i is potentially a complementary trip if:

1) bikeshare trip origin BS_{i1} is near a bus stop TS_{k1}, $d(BS_{i1}, TS_{k1}) \leqslant \text{maxDist}$ OR bikeshare trip

destination BS_{i2} is near a bus stop TS_{k2}, $d(BS_{i2}, TS_{k2}) \leqslant \text{maxDist}$

2) bikeshare trip origin and destination are within 1 mi, $d(BS_{i1}, BS_{i2}) \leqslant 1\text{mile}$
3) bikeshare trip has a short duration, $|BS_{i2}(t) - BS_{i1}(t)| \leqslant \text{maxDuration}$

After finding potential supplementary bike trips, bikeshare and transit usages are compared within each subgraph. Four types of passenger flows are considered within each subgraph: bike checkout (unlock), bike return, transit boarding, transit alighting. We obtain the total number of each type within a certain time interval τ (e.g., an hour) and create four $DN \times TN$ matrices, BC^g, BR^g, TB^g, TA^g such that:

1) DN is the number of calendar days
2) TN is the number of temporal units within a calendar day (e.g., $TN = 24$ for $\tau = 1hr$)
3) bc^g_{dt} is the number of bike checked out during the temporal unit t at the day d within graph g
4) br^g_{dt} is the number of bike returned during the temporal unit t at the day d within graph g
5) tb^g_{dt} is the number of transit boarding during the temporal unit t at the day d within graph g
6) ta^g_{dt} is the number of transit alighting during the temporal unit t at the day d within graph g

These four $DN \times TN$ matrices are visually and quantitatively compared for each subgraph. First, each matrix is visualized by assigning different colors to each cell based on its value, temporal patterns examined in one day and across all days, and such patterns compared across four matrices. For instance, if bike returns and transit boarding have similar temporal patterns across days, it is likely that the bike trips provide first-mile supports to the transit trips during those days. In contrast, if there are no obvious similar patterns, it is less likely that the bikeshare service provides strong supports to transit.

Then, the cosine similarity is applied to quantify the similarities between two matrices. To do this, we represent one row of a matrix as a TN -dimension vector, calculate the cosine similarity of two vectors in two matrices for the same day, and derive summary statistics of the cosine similarities for all days. For instance, if we would like to compare the bike returns BR^g and transit boarding TB^g in the subgraph g, we define two vectors $BR^g_d = \left(br^g_{d1}, br^g_{d1}, \ldots, br^g_{dTN}\right)$ and $TB^g_d = \left(tb^g_{d1}, tb^g_{d2}, \ldots, tb^g_{dTN}\right)$ that represent the temporal patterns of bike returns and transit boarding within a day d, respectively. Then, we calculate the cosine similarity $\cos\left(BR^g_d, TB^g_d\right)$ for $d \in [1, DN]$ and obtain summary statistics of all calculated similarities (e.g., mean, variation and outliers). If the cosine similarities tend to take large values (between 0 and 1), it indicates that the bike returns and transit boarding have similar temporal patterns across most days. If the cosine similarities have large variations, it indicates that the bike returns and transit boarding do not have similar temporal patterns. And if the cosine similarities tend to take small values, it indicates that the bike returns and transit boarding have quite distinct temporal patterns and are less likely to be complementary.

Results

Implementation and Study Design

The methodology was implemented using the programming language Python and its four supported modules: (i) pyproj for coordinate transformations, (ii) scipy for data queries, (iii) pandas for data manipulations and visualizations, and (iv) sciki-learn for machine learning algorithms (k-NN classifier and cosine similarity). Functions and modules were developed and managed using Jupyter Notebook, an open-source web application that allows users to create and share scripts and results. Maps were created using QGIS, an open-source GIS application that supports viewing, editing, and analysis of geospatial data. A GitHub repository was created to share the scripts at https://github.umn.edu/huan1531/bikeshare-with-transit/tree/master/result.

The method is demonstrated using Twin Cities, Minnesota, USA as the study area. The bikeshare data are provided by Nice Ride, which was operated by the nonprofit organization Nice Ride Minnesota and is currently in partnership with Motivate to provide full services. The Nice Ride data are downloadable at https://s3.amazonaws.com/niceride-data/index.html (accessed in July 2019) in General Bikeshare Feed Specification (GBFS) format. Data for the 2017 season from April to November, which records over 460,718 bikeshare trips among 202 bike stations, were cleaned and selected. The transaction information of each trip was mainly used: (i) the checkout time and station, (ii) the return time and station, and (iii) the longitudes and the latitudes of bikeshare stations. The transit data are provided by Metro Transit, which is the primary public transportation system in Twin Cities. The Automatic Passenger Counting data in 2017 were primarily used, during the same period as the bikeshare data. The data contain 73,477,261 records for 33,245,899 person-time at 5,002 transit stops/stations. Each record describes the number of boarding and alighting at a bus stop or light rail station, the corresponding arrival and departure time, and the location of the stop or station. The base map provided by Google Map is also used to illustrate the results.

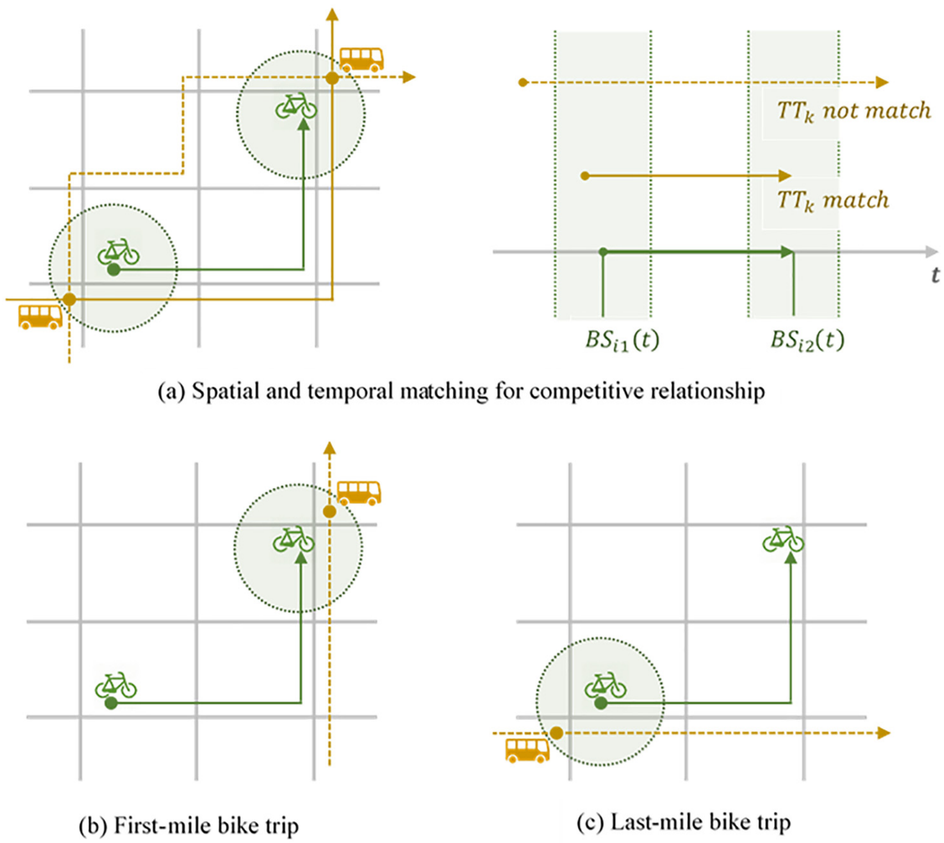


Figure 2. Conceptual definitions of competitive and complementary bike trips.

Competitive Relationship

The analysis starts with the investigation of competitive relationship. As described in Figure 2*a*, bike trips were selected that start and end within a certain distance from transit stops/stations. Considering the common width of a city block, 100 m was used as the threshold distance, and 64,501 bikeshare trips (about 2.5% of all trips) were obtained that spatially match with transit stops/stations. Figure 3*a* and *b* show the spatial distributions of the origin and destinations of substitutive trips, respectively. Each dot is a bikeshare station, and the color indicates the average number of trips per week starting and ending at that bikeshare station. This reveals that the University of Minnesota (UMN) campus in Minneapolis is the most frequent origin and destination, in particular, from Dinkytown to Prospect Park neighborhoods where many students are living; and the two stations at the ends of the Washington Ave Bridge that connects the east and west campus of UMN. The Hennepin Avenue connecting downtown and uptown is another frequent origin and destination that may result from daily commute to and from downtown central business district areas. The two stations by Bde Maka Ska (the lake in uptown west) are also two frequent origins and destinations, likely for recreation purposes.

However, not all these bikeshare trips can temporally match to one (or several) transit trips. Ten minutes was used as the threshold time for temporal matching (as shown in Figure 2*a*) so that bikeshare trip and transit trip have taken place during similar time periods and had similar trip durations. In total 2,092 pairs of bikeshare and transit trips that are within 10 min from each other were obtained, and Figure 3*c* presents the difference between their trip durations. Some 82.22% of the matching pairs were within 5 min and are almost symmetrically distributed. This indicates that users may choose the bikeshare service because of its flexibility for short-distance trips. For matching pairs with quite different trip durations (over 10 min), the bikes trips often take more than 20 min, whereas the bus trips take less than 10 min (short trip distance). In this case, users might not use the bikeshare service merely for commuting purpose (such as recreation), or stop for some intermediate activities (such as picking up a package at a postal office).

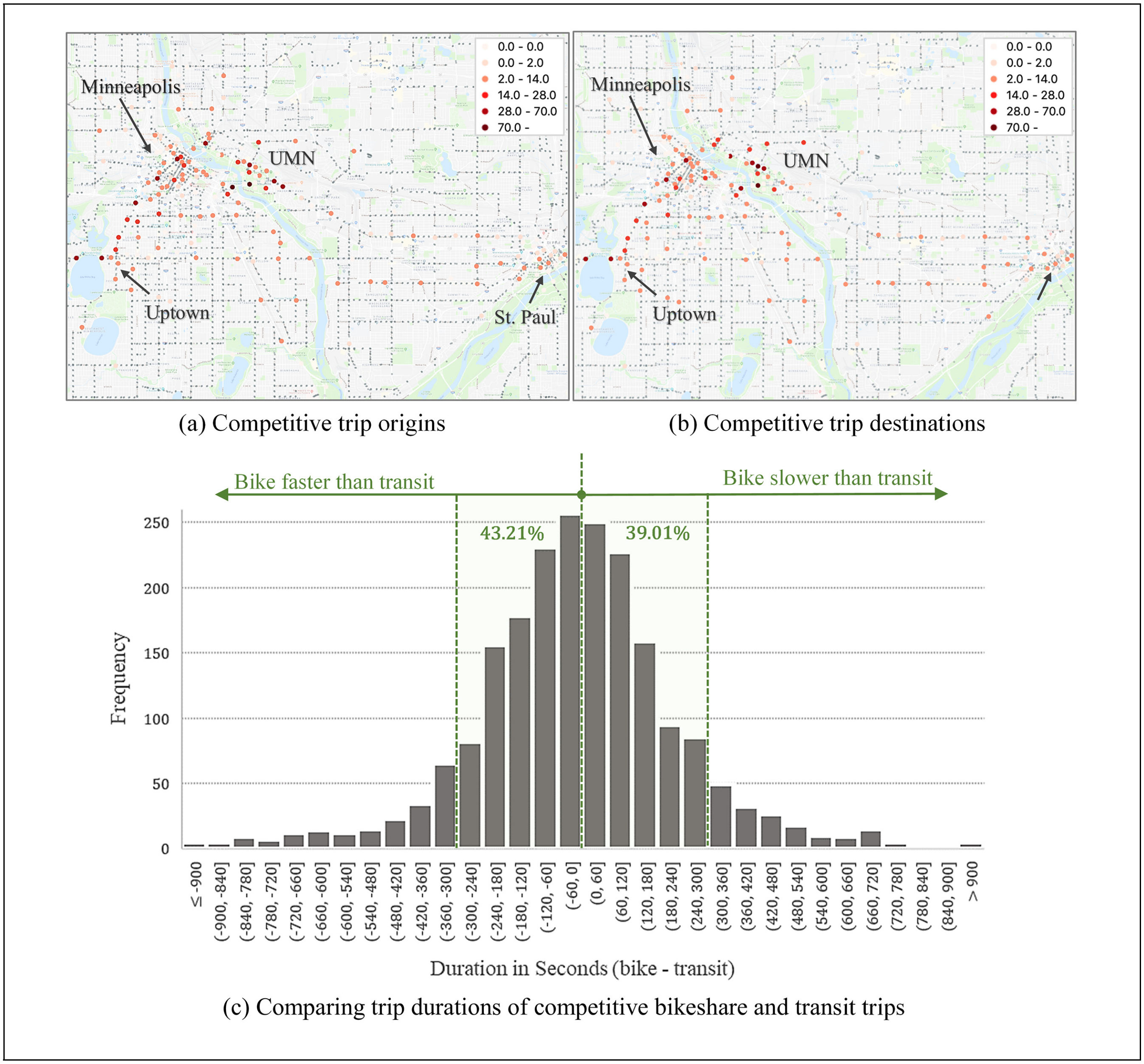
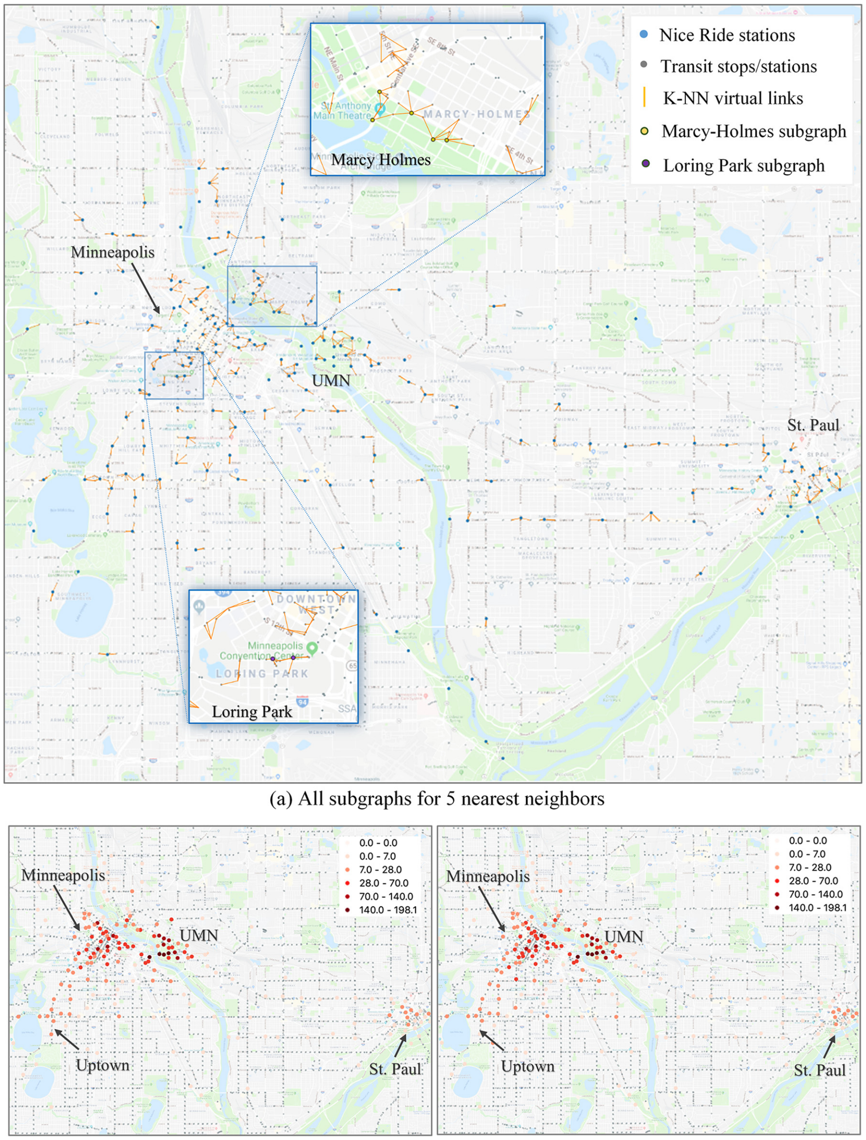


Figure 3. Spatial distributions and durations of potential substitutive bikeshare trips.

Complementary Relationship

Next the complementary relationship is investigated (see Figure 2*b, c*). First, we search for the neighboring transit stops/stations of each bikeshare station using the *k*-NN classification algorithm, and build subgraphs of neighboring stops/stations. We test different values for the number of neighbors in the k-NN classifier, $k \in \mathbb{Z}^+$, and compare the total number of the subgraphs. $k = 5$ was chosen because the number of subgraphs does not change much for $k \in [5, 7]$ and we start to have many neighboring transit stops/stations outside a 1 mi radius for $k \geqslant 8$. Figure 4*a* shows all 133 subgraphs for the entire study area. Each subgraph corresponds to a spatial region that contains one or more bike stations connecting to their neighboring transit stops/stations within 1 mi. For Minneapolis downtown, UMN campus, and St. Paul downtown areas, one transit stop/station usually has more than one bikeshare station that could provide first/last-mile support. The bikeshare stations inside parks, along recreation trails, or both often do not have any transit stops/stations nearby. Therefore, these subgraphs reflect spatial proximities between bikeshare stations and transit stops/stations and the spatial distributions of these stops/stations. Investigating the



(a) All subgraphs for 5 nearest neighbors

(b) Complementary trip origins

(c) Complementary trip destinations

Figure 4. Spatial distribution of *k*-NN subgraphs and potential first/last-mile trips.

passenger flows from and to a subgraph can provide a more comprehensive view at system level than examining ridership at each bikeshare station individually.

After creating the subgraphs, potential supplementary bikeshare trips that satisfy all the spatial and thematic criteria were selected. The threshold distance of 100 m was used to define spatial proximity and select bikeshare trips starting or ending near transit stops/stations, which is the same threshold value used to define the substitutive trips. Values of 1 mi and 10 min were used as the maximum trip distance and duration, respectively, to exclude bikeshare trips that are not for commuting to/from transit stops/stations. There are 91,255 potential complementary trips given these settings. Figure 4*b* and *c* show the spatial distributions of trip origins and destinations, which are heavily concentrated around Minneapolis downtown and UMN campus areas. Many of the stations in these areas have more than 140 bike checkouts or returns per week on average, which may be stand-alone commuting trips or first/last-mile trips to transit stops/stations.

For each subgraph, bikeshare and transit usages were further examined and compared. For potential first-mile bikeshare trips, bike return and transit boarding patterns were compared. For the last-mile case, transit alighting and bike checkout patterns were compared. This paper

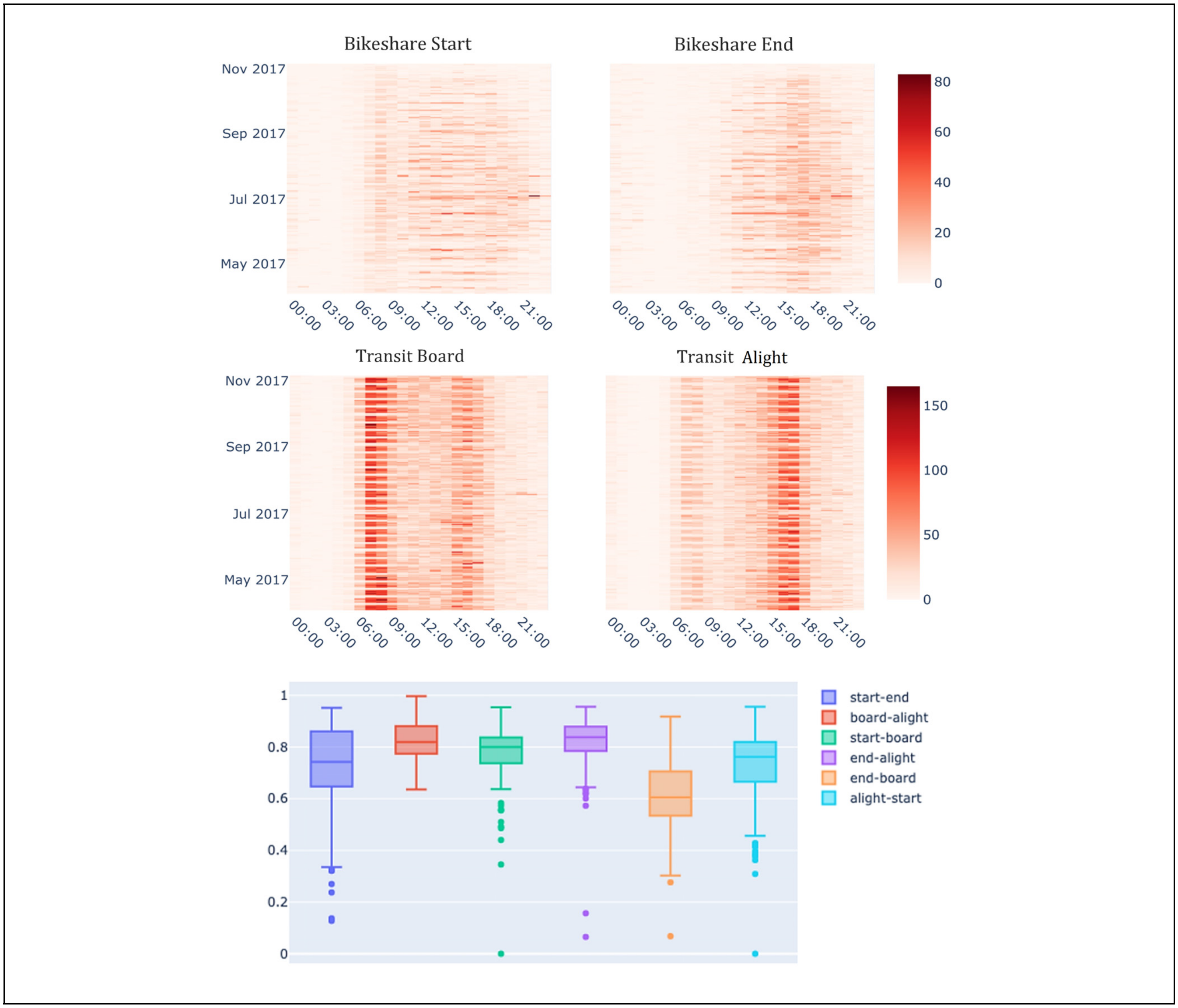


Figure 5. Bikeshare and transit usage comparisons (Marcy-Holmes).

uses 1 h as the time interval, τ = 1h and derives the four usage matrices. Each cell in the matrix represents the total number of bike or transit trips starting/ending in each subgraph. We visualize the four usage matrices and calculate the pair-wise cosine similarity for each of the 133 subgraphs. Figure 5 and Figure 6 show the comparison results for two example subgraphs for two different neighborhoods (see Figure 4*a*).

Figure 5 shows results for Marcy-Holmes, which is the first neighborhood of Minneapolis next to the UMN campus. The neighborhood is mainly residential, and about 88% of households are rental (2010 US Census). The bike checkout has a slightly high volume from 7 to 9 a.m. and has lower volumes than bike return during the rest of the day. The transit boarding has a morning peak from 7 to 9 a.m., and the transit alighting has an afternoon peak from 4 to 6 p.m. The pair-wise cosine similarities indicate that the bike return and the transit alighting have the most similar patterns with the least variations, whereas the bike return and the transit boarding have the least similar patterns with large variations. This suggests that the bikeshare is likely to serve as an alternative travel mode instead of first/last-mile support.

Figure 6 shows results for Loring Park, which is mainly residential and has large green space and extensive walking and cycling paths. The Minneapolis Convention Center is also in this subgraph region. The visualizations of the four matrices do not indicate any obvious patterns. The cosine similarities show that the transit boarding and alighting are quite similar to each other, both with slightly higher afternoon volumes than morning. All other pairs have similar but not quite the

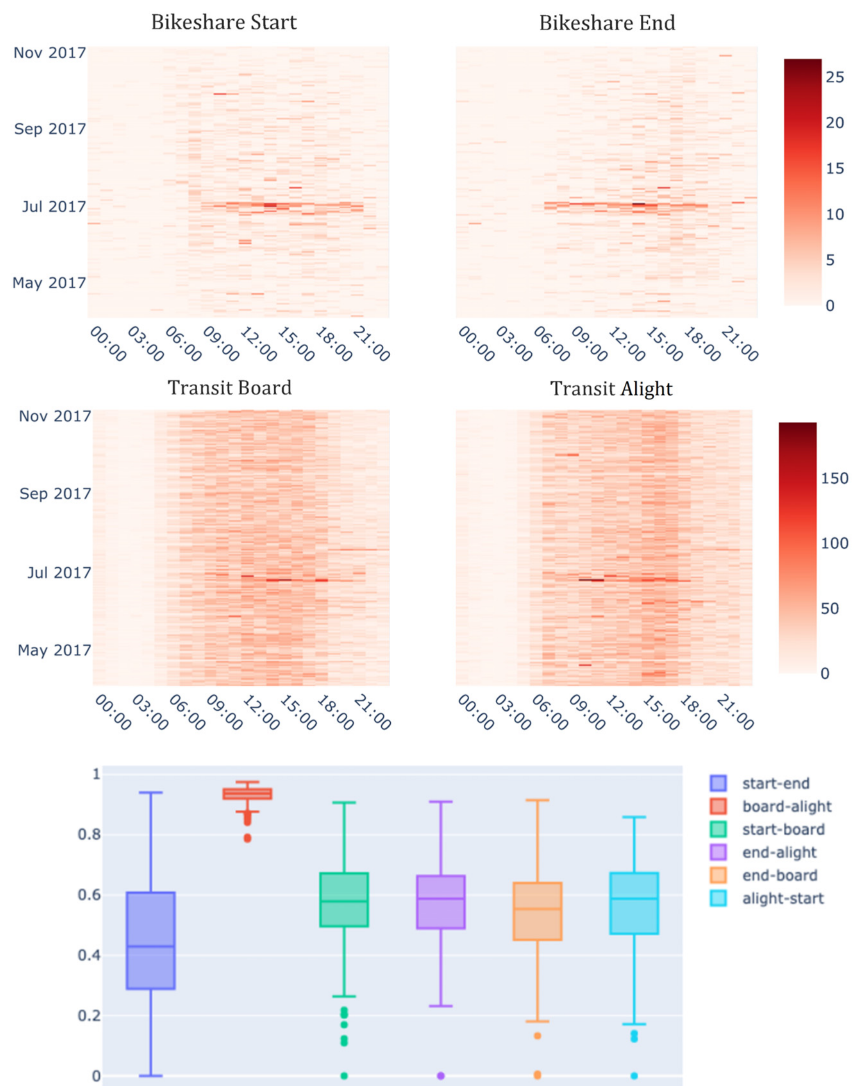


Figure 6. Bikeshare and transit usage comparisons (Loring Park).

same patterns. This is likely for two reasons: (i) users do not use bikeshare and transit for daily commute to/from work given the mixed land-use types, and (ii) the extreme high volume during July 4th holiday may affect the overall similarities. In sum, there is no strong evidence of a complementary relationship in the Loring Park region too.

Relationship between Competitive and Complementary Relationships

The competitive and complementary relationships are parallel processes in our framework; therefore, it is worthy to check whether these two relationships are exclusive. To do this, labels are generated for each bikeshare trip, containing three digits. The first digit indicates whether the trip is competitive; the second digit indicates whether the bike trip provides first-mile transit access; and the third digit indicates whether the bike trip provides last-mile transit access. The digit "1" indicates an affirmative answer "Yes" and the digit "0" indicates a negative answer. For instance, the label "100" means the bike trip is competitive and provides neither first nor last-mile transit access, and the label "010" means that the bike trip only provides first-mile access and does not compete with transit. Table 1 provides the percentage of bike trips for each case. The table shows that out of all the potential competitive bikeshare trips (6.12%), over half of them (3.55%) are also providing both first and last-mile access to the transit, and the other 2.56% bike

Table 1. Different Types of Bikeshare Trips Considering their Relationships to the Transit

	Neither first/last	First mile only	Last mile only	Both first and last	Total
Is competitive	**2.56%**	0.00%	0.00%	**3.55%**	6.12%
Not competitive	64.22%	**9.60%**	**9.96%**	**10.10%**	93.88%
Total	66.78%	9.60%	9.96%	13.65%	100.00%

share trips only compete with the transit, but do not provide first or last-mile access to the transit. There are 29.66% of bikeshare trips that provide either first or last-mile access to the transit.

Conclusion

This paper develops a framework for investigating the interactions between bikeshare and transit systems from a spatial-temporal perspective. The framework considers both the competitive and the complementary relationships, and adopts different procedures and methods to investigate each relationship with respect to both spatial locations and temporal profiles of the two systems. Twin Cities in MN is used as an example study area, and the interactions between Metro Transit and Nice Ride bikeshare services are examined. Of the 460,718 bikeshare trips and 33,245,899 transit boarding and alighting records, 3.55% of trips are competitive and 29.66% of trips are complementary. More interesting, the competitive bikeshare trips are also potentially supporting both first and last access to transit. This suggests that the bikeshare service is more likely to support transit services than to compete with them. As for the spatial distributions of bikeshare trips, both competitive and complementary trips are concentrated in the Twin Cities downtown and UMN campus areas. This finding is intuitive because bike stations along or near biking trails tend to be used for recreation purposes as well as for commuting.

The current work can be expanded in three major directions: theoretical, empirical, and practical. First, the methods and parameter settings can be refined by: (i) considering both spatial proximity and other domain-driven knowledge such as the predefined community neighborhoods when we derive subgraphs, (ii) collecting additional data such as user surveys to determine the distance threshold and time window, (iii) considering the dockless bikeshare, scooter, and other services and using methods other than k-NN to define spatial proximity and derive subgraphs, and (iv) applying data mining methods such as hierarchical clustering algorithms (*39*) to classify and describe subgraphs based on the pair-wise cosine similarities of bikeshare and transit usage within each subgraph.

Second, the framework and findings can be further validated by applying them to other study areas with both transit and bikeshare services. The authors have created a GitHub repository to share the scripts and results with researchers, local governments, transit agencies, and private companies so that they can apply the framework directly to new study areas and datasets to investigate the relationships between share mobility services and public transit. Learning from various case studies can further refine the general methods and parameter settings in the current framework.

Third, the framework and findings can potentially be used in operation and planning practices. The bikeshare companies and the transit authority can use the spatial distribution of all subgraphs and the usage patterns within these subgraphs to identify possible needs to add/remove bike stations, transit stops/stations or both. The local government can use the results to identify areas with high volumes of transit and bike usage, and possibly allocate more resources if needed (e.g., bike and bus priority lane; bike rack inside the subgraph with many first/last-mile bikeshare trips).

Authors' Note

The supplementary data is provided by Metro Transit, the primary public transit operator in the Minneapolis–Saint Paul area (through the Traffic Observatory at University of Minnesota) and Nice Ride, the nonprofit bicycle sharing service, at Twin Cities, MN, U.S.

Author Contributions

The authors confirm the following contributions to the paper. Y. Song takes the lead on study design, data collection, and manuscript preparation. Y. Huang takes the lead on method development, implementation, and result generations. Y. Song and Y. Huang work together on analysis and interpretation of results and revision of the manuscript.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This paper is based upon research supported by Grant-in-aid of Research, Artistry, and Scholarship at University of Minnesota (UMN), a nonprofit organization.

References

1. Handy, S. L. Accessibility-vs. Mobility-Enhancing Strategies for Addressing Automobile Dependence in the US. Presented at the ECMT Round Table on Transport and Spatial Policies: The Role of Regulatory and Fiscal Incentives, RT124, Paris, 2002.
2. Litman, T. *Evaluating Public Transit Benefits and Costs*. Victoria Transport Policy Institute, 2002. www.vtpi.org. Accessed December 11, 2019.
3. Shatu, F., and M. Kamruzzaman. Investigating the Link Between Transit Oriented Development and Sustainable Travel Behaviour in Brisbane: A Case Control Study. *Journal of Sustainable Development*, Vol. 7, No. 4, 2014, pp. 61–70.
4. Boarnet, M. G., G. Giuliano, Y. Hou, and E. J. Shin. First/Last Mile Transit Access as an Equity Planning Issue. *Transportation Research Part A: Policy and Practice*, Vol. 103, 2017, pp. 296–310.
5. Chong, Z. J., B. Qin, T. Bandyopadhyay, T. Wongpiromsarn, E. S. Rankin, M. H. Ang, E. Frazzoli, D. Rus, D. Hsu, and K. H. Low. Autonomous Personal Vehicle for the First-and Last-Mile Transportation Services. *Proc., 2011 IEEE 5th International Conference on Cybernetics and Intelligent Systems (CIS)*, Qingdao, China, IEEE, New York, 2011, pp. 253–260.
6. National Academies of Sciences, Engineering, and Medicine. *Public Transit and Bikesharing*. The National Academies Press, Washington, D.C., 2018.
7. Pendall, R., E. Blumenberg, and C. Dawkins. *What if Cities Combined Car-Based Solutions with Transit to Improve Access to Opportunity?* Metropolitan Housing and Communities Policy Center. The Urban Institute, Washington, D.C., 2016. https://www.urban.org/research/publication/what-if-cities-combined-car-based-sol. Accessed December 11, 2019.
8. Sallis, J. F., L. D. Frank, B. E. Saelens, and M. K. Kraft. Active Transportation and Physical Activity: Opportunities for Collaboration on Transportation and Public Health Research. *Transportation Research Part A: Policy and Practice*, Vol. 38, No. 4, 2004, pp. 249–268.
9. Shaheen, S., and N. Chan. Mobility and the Sharing Economy: Potential to Facilitate the First-and Last-Mile Public Transit Connections. *Built Environment*, Vol. 42, No. 4, 2016, pp. 573–588.
10. Brons, M., G. Moshe, and R. Piet. Access to Railway Stations and its Potential in Increasing Rail Use. *Transportation Research Part A: Policy and Practice*, Vol. 43, No. 2, 2009, pp. 36–49.
11. Fishman, E., S. Washington, and N. Haworth. Bike Share's Impact on Car Use: Evidence from the United States, Great Britain, and Australia. *Transportation Research Part D: Transport and Environment*, Vol. 31, 2014, pp. 13–20.
12. Lansell, K. *Melbourne Bike Share and Public Transport Integration*. Master of Urban Planning Minor thesis. University of Melbourne, 2011.
13. LDA Consulting. *Capital Bikeshare 2011 Member Survey Report*. 2012. Washington, D.C.,: LDA Consulting.
14. Yang, T., H. Pan, and Q. Shen. *Bike-Sharing Systems in Beijing, Shanghai and Hangzhou and their Impact on Travel Behavior*. No. 11-3862. Department of Urban Planning, Tongji University, 2010.
15. Campbell, K. B., and C. Brakewood. Sharing Riders: How Bikesharing Impacts Bus Ridership in New York City. *Transportation Research Part A: Policy and Practice*, Vol. 100, 2017, pp. 264–282.
16. Shaheen, S., H. Zhang, E. Martin, and S. Guzman. China's Hangzhou Public Bicycle. *Transportation Research Record: Journal of the Transportation Research Board*, 2011. 2247: 33–41.
17. Ma, X., Y. Ji, M. Yang, Y. Jin, and X. Tan. Understanding Bikeshare Mode as a Feeder to Metro by Isolating Metro-Bikeshare Transfers from Smart Card Data. *Transport Policy*, Vol. 71, 2018, pp. 57–69.
18. Martin, E. W., and S. A. Shaheen. Evaluating Public Transit Modal Shift Dynamics in Response to Bikesharing: A Tale of Two US Cities. *Journal of Transport Geography*, Vol. 41, 2014, pp. 315–324.
19. Krizek, K., and E. Stonebraker. Assessing Options to Enhance Bicycle and Transit Integration. *Transportation Research Record: Journal of the Transportation Research Board*, 2011. 2217: 162–167.
20. Wang, X., G. Lindsey, J. E. Schoner, and A. Harrison. Modeling Bike Share Station Activity: Effects of Nearby Businesses and Jobs on Trips to and from Stations. *Journal of Urban Planning and Development*, Vol. 142, No. 1, 2015, p. 04015001.
21. Yang, X. H., Z. Cheng, G. Chen, L. Wang, Z. Y. Ruan, and Y. J. Zheng. The Impact of a Public Bicycle-Sharing System on Urban Public Transport Networks. *Transportation Research Part A: Policy and Practice*, Vol. 107, 2018, pp. 246–256.
22. Jäppinen, S., T. Toivonen, and M. Salonen. Modelling the Potential Effect of Shared Bicycles on Public Transport Travel Times in Greater Helsinki: An Open Data Approach. *Applied Geography*, Vol. 43, 2013, pp. 13–24.
23. Cover, T. M., and P. Hart. Nearest Neighbor Pattern Classification. *IEEE Transactions on Information Theory*, Vol. 13, No. 1, 1967, pp. 21–27.
24. Peterson, L. E. K-Nearest Neighbor. *Scholarpedia*, Vol. 4, No. 2, 2009, p. 1883.
25. Roussopoulos, N., S. Kelley, and F. Vincent. Nearest Neighbor Queries. *ACM Sigmod Record*, Vol. 24, No. 2, 1995, pp. 71–79.
26. Fix, E., and J. L. Hodges. *Discriminatory Analysis, Nonparametric Discrimination: Consistency Properties*. Technical Report 4. USAF School of Aviation Medicine, Randolph Field, Tex., 1951.
27. Pan, F., B. Wang, X. Hu, and W. Perrizo. Comprehensive Vertical Sample-Based KNN/LSVM Classification for Gene Expression Analysis. *Journal of Biomedical Informatics*, Vol. 37, No. 4, 2004, pp. 240–248.
28. Li, L., C. R. Weinberg, T. A. Darden, and L. G. Pedersen. Gene Selection for Sample Classification Based on Gene Expression Data: Study of Sensitivity to Choice of

Parameters of the GA/KNN Method. *Bioinformatics*, Vol. 17, No. 12, 2001, pp. 1131–1142.
29. Chapelle, O., P. Haffner, and V. N. Vapnik. Support Vector Machines for Histogram-Based Image Classification. *IEEE Transactions on Neural Networks*, Vol. 10, No. 5, 1999, pp. 1055–1064.
30. Zhang, H., A. C. Berg, M. Maire, and J. Malik. SVM-KNN: Discriminative Nearest Neighbor Classification for Visual Category Recognition. *Proc., 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06)*, New York, IEEE, New York, 2006, 2126–2136.
31. Pan, J., and D. Manocha. Fast GPU-Based Locality Sensitive Hashing for K-Nearest Neighbor Computation. *Proc., 19th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, Chicago, Ill., 2011, pp. 211–220.
32. Bui, D. T., Q. P. Nguyen, N. D. Hoang, and H. Klempe. A Novel Fuzzy K-Nearest Neighbor Inference Model with Differential Evolution for Spatial Prediction of Rainfall-Induced Shallow Landslides in a Tropical Hilly Area Using GIS. *Landslides*, Vol. 4, No. 1, 2017, pp. 11–17.
33. Perronnin, F., Y. Liu, J. Sánchez, and H. Poirier. Large-Scale Image Retrieval with Compressed Fisher Vectors. *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition,* San Francisco, Calif., IEEE, New York, 2010, pp. 3384–3391.
34. Babenko, A., and V. Lempitsky. Aggregating Local Deep Features for Image Retrieval. *Proc., IEEE International Conference on Computer Vision*, Santiago, Chile, IEEE, New York, 2015, pp. 1269–1277.
35. Mihalcea, R., C. Corley, and C. Strapparava. Corpus-Based and Knowledge-Based Measures of Text Semantic Similarity. *Proc., 2006 Association for the Advancement of Artificial Intelligence (AAAI)*, Boston, Mass., 2006, pp. 775–780.
36. Huang, A. Similarity Measures for Text Document Clustering. *Proc., 6th New Zealand Computer Science Research Student Conference*, Christchurch, New Zealand, 2008, pp. 9–56.
37. Liu, H., and M. Schneider. Similarity Measurement of Moving Object Trajectories. *Proc., 3rd ACM SIGSPATIAL International Workshop on GeoStreaming*, Redondo Beach, Calif., 2012, pp. 19–22.
38. Li, Q., Y. Zheng, X. Xie, Y. Chen, W. Liu, and W. Y. Ma. Mining User Similarity Based on Location History. *Proc., 16th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, Irvine, Calif., 2008.
39. Jain, A. K., and R. C. Dubes. *Algorithms for Clustering Data*. Englewood Cliffs, N.J.,: Prentice Hall, Vol. 6, 1988.