Towards a Connective, Sustainable Transportation: A Study of Relationship between Bike Sharing Service and Public Transit in Minneapolis-St. Paul Area

A THESIS

SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL

OF THE UNIVERSITY OF MINNESOTA

BY

Yuchuan Huang

IN PARTIAL FULFILLMENT OF THE REQUIERMENTS

FOR THE DEGREE OF

MASTER OF ARTS

Advisor: Ying Song

August, 2020

©Yuchuan Huang 2020

Acknowledgements

I would like to extend my sincere thanks, gratitude and appreciation to all those who have guided and assisted me in the completion of this degree.

Firstly, I would like to thank my advisor, Ying Song, for all the guidance and support she has provided throughout the three years' studies. From her I learned to be a mature researcher, and I will keep working on that.

In addition, I really appreciate Eric Shook and Chen-Fu Liao for serving on my committee and providing me with their insightful and valuable feedback, especially during this unusual summer.

Finally, I would like to thank my parents for their long-time support throughout my education. I love you.

Abstract

Public transit offers many socioeconomic and environmental benefits but often suffers from the first/last-mile problem. Bike sharing service is designed and expected to provide first/last-mile access to transit, have a positive integration with public transit, and together contribute to a connective and sustainable transportation ecosystem. However, the relationship between existing bike sharing service and public transit is complex and ambiguous. This thesis proposes a data-driven framework with procedures and methods to investigate the competitive and complementary relationship between bike sharing and public transit systems. It defines relationships analytically and uses the criteria to detect pairs of bike sharing and transit trips correspondingly. Then, it examines the properties of paired trips and possible reasons. The thesis applies this framework to the Nice Ride bike sharing service and Metro Transit system in the Minneapolis-St. Paul Area, as a case study. The results suggest that competitive relationship exists, but only constitutes a small portion of all bike sharing trips when the spatio-temporal criteria are strict. The study of complementary relationship detects the potential first/last-mile bike trips and suggests that complementary relationship may exist and have unique spatio-temporal patterns. The correlation between bike sharing and transit ridership does not show significant competitive or complementary relationship in general, suggesting that these two systems tend to operate relatively independently from each other in the Twin Cities. However, evidence for competitive relationship can be found in several small regions. The results provide novel insights into the complex interactions between bike sharing and public transit systems and can support operation and planning practices. Since the relationships are purely defined using ridership data, we need to integrate more data to further validate our method in the future.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Background

Sustainable mobility has become a key theme in modern urban transportation planning, where moving faster and further is no longer the only goal and promoting green, healthy and equitable access to each and everyone becomes the core themes (Banister, 2008). Public transit, i.e., bus and subway/light-rail system, is proved to an effective approach to achieving this goal. Public transit offers many societal and environmental benefits: it promises to mitigate auto-dependency and provide access to various resources and opportunities (Anderson, 2014; Besser & Dannenberg, 2005; Litman, 2020). However, a critical challenge in improving public transit performance and increasing transit ridership is the first/last-mile problem (Boarnet et al., 2017; Lesh, 2013; Tilahun et al., 2016). In essence, people who want to choose transit sometimes do not use it because the distance to the bus stop or light-rail station is too great. At the same time, it is too costly or impractical to invest in and build transit systems close to all homes, jobs, schools, and other destinations.

Shared mobility, the shared use of the bicycle, scooter or other modes, is expected to be a possible solution to the first/last-mile problem (Chong et al., 2011; Moorthy et al., 2017; Scheltes & de Almeida Correia, 2017; S. Shaheen & Chan, 2016; Shen et al., 2018). Of all the shared modes, bike sharing service has become an indispensable component of shared mobility. In the year of 2018, 45.5 million trips were taken on shared bikes in the U.S.(NACTO, 2018). In addition to the promised benefits such as avoiding vehicular congestion, reducing carbon footprint and more opportunities for healthy physical activities, bike sharing service could potentially solve the first/last-mile problem by integrating with the existing transit system and expanding its service area and time (Adnan et al., 2019; Chen et al., 2020; Fan et al., 2019; Liu et al., 2012; S. A. Shaheen, 2012; S. A. Shaheen et al., 2010).

Bike sharing systems are usually planned to spatially connect to other public transit modes (Griffin & Sener, 2016; Krizek & Stonebraker, 2011). According to the Bureau of Transportation Statistics (Firestine, 2016), in the year of 2016, 77.0% of all 3378 bike sharing stations across 104 U.S. cities connect to another scheduled public transit mode within one block. Transit bus is the most typical connection, with 74.9% of bike sharing stations located a block or less from a bus stop. In 2019, the city of Minneapolis launched a mobility hub pilot program (Rasp et al., 2019). The mobility hub was designed to be a place where people can connect to multiple modes of transportation, including transit, shared scooters and Nice Ride bicycles, to make their trips as safe, convenient and reliable as possible. In general, bike sharing systems are born with the nature of having connections with public transit systems.

Studies have indicated positive impacts of bike sharing services on public transit systems (Fishman et al., 2013). T. Ma et al. (2015) found that public transit ridership was positively associated with bike sharing ridership at the station level. A 10% increase in annual bike sharing ridership contributed to a 2.8% increase in average daily Metrorail ridership in Washington D.C.. Jin et al. (2019) showed that the usage of bike sharing in Beijing, China did not result in a reduction in overall transit ridership, and transit transferring behaviors were highly correlated with the bike sharing usage, which suggests that bike sharing could interact with public transit in a cooperative manner. However, some surveys and studies also illustrate an unfortunate irony: bike sharing may act as a competitor to transit instead of complementing it. Campbell and Brakewood (2017) found a significant decrease in daily bus ridership along routes that are near bike sharing in New York City, in comparison to routes that are not. Graehler et al. (2019) conducted a longitudinal analysis across 22 metropolitan areas and found that the introduction of bike sharing in a city is associated with an increase in light and heavy rail ridership, but a 1.8% decrease in bus ridership in average. Therefore, the relationship between bike sharing and public transit can be complex.

1.2 Summary of Goals and Contributions

The above discussion illustrates the importance and challenges of determining the relationship between bike sharing service and public transit. That is, people design and expect that bike sharing service could have a positive integration with public transit and together contribute to a sustainable transportation system, while the relationship between the two systems could be ambiguous and complex. Thus, the overall goal of this thesis is to *develop a framework to first define and then determine or discuss the relationship between bike sharing service and public transit*. In essence, I hope that this work could provide a perspective for current deployment assessment and future multi-modal transportation planning.

The contributions of this thesis are as follows:

Explicit relationship definition I provide explicit definitions for two relationships of interest: competitive and complementary relationship. The relationships are both defined from a spatial-temporal perspective along with the analytical methods to detect and describe each relationship.

Data-driven framework I propose a data-driven framework to investigate the relationships. Compared with many works based on user surveys or simulations, my method uses historical ridership data so it provides an objective view on the actual interactions between bike sharing and transit systems.

Practical framework I apply the proposed methods on real data and show that my methods are theoretically and practically workable. I also integrate the result with the demographic context and demonstrate the interpretability of my work.

The following parts of this thesis are organized as follows. Chapter 2 reviews the related literature in the field. Chapter 3 defines the competitive and complementary relationship between bike sharing and public transit systems and proposes analytical methods to unveil the relationship from historical ridership. In Chapter 4 I apply the proposed methods to the case of Minneapolis-St. Paul metro area and discuss the results. Chapter 5 concludes the thesis and outlines some future research directions. I also give my suggestion on better integrating bike sharing service with public transit based on the study in the Twin Cities area.

Chapter 2

Literature Review

Studies have adopted various methods to advance our understanding on how bike sharing services interact with public transit systems. This chapter summarizes related works with a focus on their methods and discusses the results drawn from different works. It also reviews general spatial-temporal analysis methods and discusses why we do not apply those methods directly. Figure 2.1 shows the taxonomy of methods in related works.

The majority of related works are based on user survey data instead of driven by actual usage data (Adnan et al., 2019; Fan et al., 2019; Martin & Shaheen, 2014; S. A. Shaheen et al., 2011; S. Shaheen et al., 2013; Zhang & Zhang, 2018). Questions are designed and asked, either online or offline, to understand respondents' travel preference and modal shift in adopting bike sharing. S. A. Shaheen et al. (2011) took survey with about 1,000 respondents in Hangzhou, China and found that bike sharing was capturing modal share from bus transit, walking, autos, and taxis. Approximately 30% of respondents had incorporated bike sharing into their common commute. S.

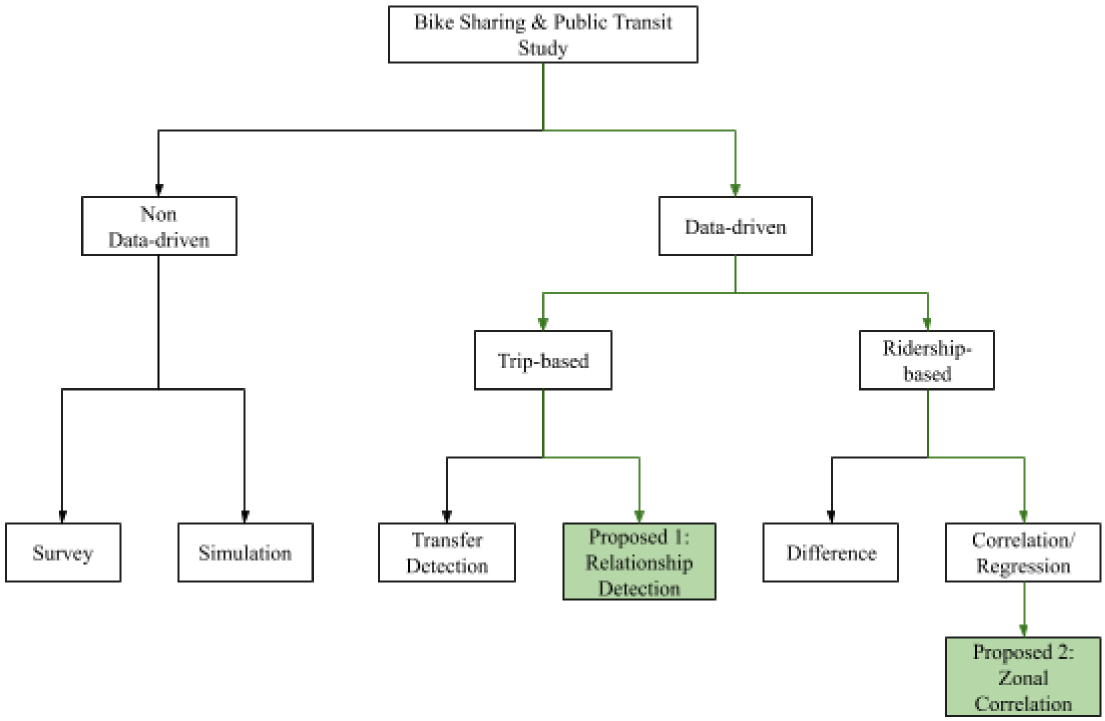


Figure 2.1: Related Works Taxonomy

Shaheen et al. (2013) conducted an online survey in Twin Cities and 15% of the respondents reported increasing rail usage in adopting bike sharing service. Martin and Shaheen (2014) evaluated another survey in Minneapolis and found that the modal shift toward rail extended to the urban core, while the shift for bus transit was more dispersed. Comparing with survey from Washington D.C., the authors suggested that bike sharing may be more complementary to public transit in small to mid-size cities and more substitutive of public transit in larger and denser cities. Zhang and Zhang (2018) used data from the 2017 National Household Travel Survey in U.S. where respondents were asked how many times they had used public transit and bike sharing in the last 30 days. The authors concluded that public transit usage is significantly positively associated with bike sharing usage in U.S..

Another common method is modeling/simulating(Hong et al., 2016; Jäppinen et al., 2013; Lu et al., 2018; Yang et al., 2018). In these studies, relationship between the two systems is usually unveiled by the increased spatial accessibility or the decreased travel time facilitated by integrating bike sharing system with public transit network. Hong et al. (2016) developed a utility-based model in conjunction with choice behavior strategy and found that bike sharing program successfully increased the accessibility of public transit. Jäppinen et al. (2013) modeled the travel times by public transit compared with public transit extended with shared bikes, in the Greater Helsinki, Finland. The trip simulation between popular Origin-Destinations showed that the launch of bike sharing system could reduce public transportation travel times on average by more than 10% in the study area, which is 6 minutes per individual trip. Similarly, Yang et al. (2018) developed a spatial network model of bus network and bike sharing system in Ningbo, China and the trip simulation showed that short distance biking can significantly reduce the average transfer times and average path length of passengers' trips.

With more bike sharing data being collected and becoming accessible, data-driven approaches emerge that utilize historical ridership and operational data. The most insightful data is collected when bike sharing and public transit share the same payment system and user database. X. Ma et al. (2018) detected bike-metro transfer trips from smart card transaction data. The detected transfer trips indicated that the integration of bike sharing and metro exists but the amount of transfer indicated that such integration was infrequently and impromptu.

However, not all cities have such integrated payment system and not all users use a smart card to pay for their trips. When integrated data is not available, most data-driven studies develop ridership-based methods, where scholars focused more on the number of passengers at each station/stop. A common manner is to investigate the difference of one ridership before and after another travel mode is available. Difference-in-Difference (DID) method is often applied in this kind of works. Campbell and Brakewood (2017) divided bus routes into control and treatment groups in New York City based on if they are located in areas that received bike sharing infrastructure. The results indicated that every thousand bike sharing docks along a bus route is associated with a 2.42% fall in daily bus trips on routes in Manhattan and Brooklyn. X. Ma et al. (2019) used a DID method to evaluate the impact of free-floating bike sharing service on bus ridership in Chengdu, China. The result showed that the emergence of bike sharing led an increase of bus ridership on route level, while the increase on bus stop level is insignificant. Gu et al. (2019) conducted similar "before-and-after" study on a new open metro line in Suzhou, China and found that most bike sharing ridership within the metro's catchment had largely increased since the introduction of the new line. Saberi et al. (2018) conducted a comparative analysis of bike sharing spatial mobility patterns before, during, and after a case of Tube strike in London. The analysis indicated that the disruption in public transit increased the total number of bike sharing trips by 85% and the duration of trips also was also increased by 88%.

In addition to the ridership differences before and after, another common focus is the correlation between bike sharing ridership and transit ridership. T. Ma et al. (2015) conducted an ordinary least squares regression analysis in Washington D.C., using transit ridership as dependent variable, and bike sharing ridership and other socio-demographic factors as independent variables. The analysis found that public transit ridership was positively associated with bike sharing ridership at the station level. A 10% increase in annual bike sharing ridership contributed to a 2.8% increase in average daily transit ridership. Ji et al. (2018) used smart card transaction to identify metro-bike transfer trips and then developed a Geographically Weighted Poisson Regression (GWPR) model to explore the relationship between transfer volume and several socio-demographic variables. The modeling and spatial visualization results showed that riding distance is negatively correlated with metro-bike transfer demand.

To sum up the above works, different studies of different cities get different results. The relationship was described by the words "competitive", "substituting" or "negative" for bike sharing decreasing the public transit usage; and "complementary", "supplement" or "positive" for bike sharing increasing public transit usage. In some studies the relationship is onefold while in other studies the relationship is multi-fold and the relationship can vary between center and outskirt of cities, between bus and metro systems and between route-level and stop-level.

Despite the efforts and contributions of previous works, the relationship between bike sharing service and public transit is still not well answered, especially when we are interested in how bike sharing solves first/last-mile problem of public transit. Non-data-driven methods are good for pre-deploy estimation but not always sufficient for post-deploy assessment. Smart card transaction data gives direct information about bike-transit transfer, but is only applicable to urban areas facilitated with it. And there is also a concern of privacy which makes such data hard to share among researchers and practitioners. Ridership-based studies help people understand how bike sharing ridership reacts to the change of public transit ridership, but that correlation is not strong enough to suggest any factual passenger transfer between the two systems. Therefore, my thesis attempts to develop a framework that can provide a comprehensive view on spatial-temporal relationship between transit and bike-sharing systems using ridership and operational data that do not include personal identities and are commonly available.

General spatial-temporal analysis methods have been widely applied to the studies of bike sharing. (1) Visual exploration is a common approach to get preliminary insights into the relationship and has been used in most studies. For instance, interactive visualization systems are developed to explore the dynamics of bike sharing ridership (Oliveira et al., 2016; Oppermann et al., 2018). In addition to the basic spatial-temporal information, some systems also include functions to extract and visualize mobility patterns (Moncayo-Martínez & Ramirez-Nafarrate, 2016; Yan et al., 2018). (2) Spatial statistics methods, such as spatial autocorrelation and regression, are also applied in the study of bike sharing (Christian et al., 2019; Feng & Wang, 2017; Ji et al., 2018; Wu & Chang, 2016). These works often focus on the correlation between bike sharing ridership with socio-demographic and build environment factors. (3) With bike sharing data becoming available, more and more works applied data mining or machine learning methods to bike sharing studies (Etienne & Latifa, 2014; O'Brien et al., 2014; Vogel et al., 2011; Zhang et al., 2018). In these works, data mining methods like classification are used to detect and describe mobility patterns. And machine learning methods are used to predict bike demands, primarily for rebalancing the bikes so as to provide reliable services (Ashqar et al., 2017; Lin et al., 2018; Wang & Kim, 2018; Xu et al., 2018).

These existing spatial-temporal analysis methods provide useful insights into the patterns of bike sharing usage and their relationship to geographies of the people and study areas. However, they often focus on one system at a time and cannot be directly applied to study the spatial-temporal relationships between two systems quantitatively. Therefore, in the following methodology chapter, we first propose a quantitative definition of relationship, and then we discuss the detected relationship from a spatio-temporal perspective.

Chapter 3

Method

This chapter proposes methods to study the relationship between bike sharing service and public transit, from both trip-based and ridership-based perspective. Competitive and complementary relationships between bike sharing trips and transit trips are defined first. The relationships are detected and examined from a spatial-temporal perspective. In addition to the study of individual trips, a study of the ridership at bike stations and transit stops is conducted next. Figure 3.1 shows the general method framework with major steps.

3.1 Data Schemas and Notations

The methods are developed based on the typical available data. The data of bike sharing service is organized as two main parts, bike trip and bike station. Bike trip data consists of when and where each trip starts and ends, and the type of user. Bike station data consists of the exact latitude and longitude of each station.

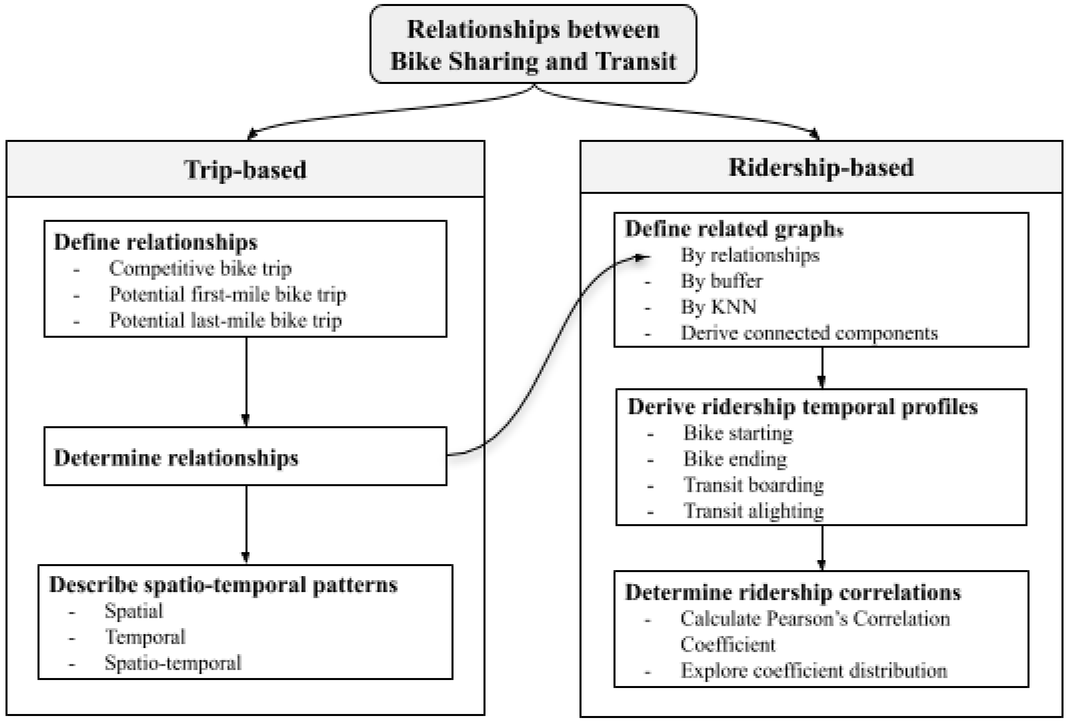


Figure 3.1: Method Framework

The data of public transit also contains two parts, Automatic Passenger Counting (APC) data and transit stop data. APC data is recorded by Automatic Passenger Counter, which is an electronic sensor mounted in bus doorways counting passengers boarding and alighting; the corresponding information on bus stop location and time is acquired from the Automatic Vehicle Location (AVL) system. APC data consists of the basic trip identification, when and where the transit vehicle stops by and how many passengers board and alight the transit at such stop. The transit stop data consists of the exact latitude and longitude of each stop. Figure 3.2 shows the explicit schema of each data source respectively.

For the convenience of description, a record of bike sharing trip bt_i will be denoted

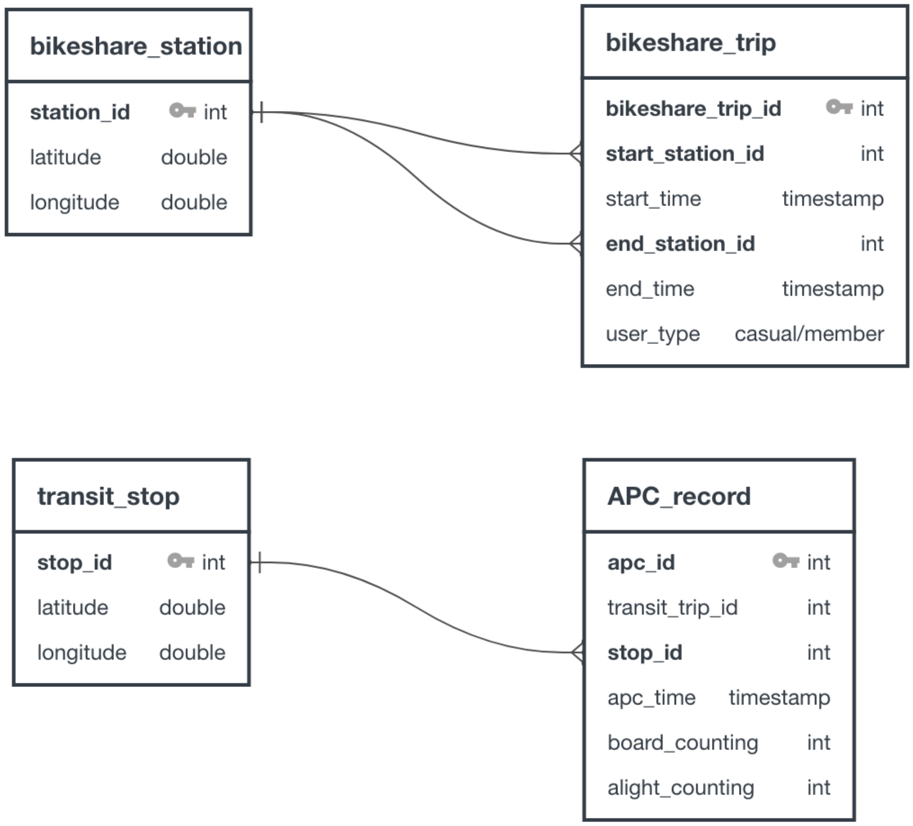


Figure 3.2: Source Data Schema

as

$$bt_i(start\_stn_i, start\_time_i, end\_stn_i, end\_time_i, user)$$

, where $start\_stn_i$ and end_stn_i are the start and end stations and $start\_time_i$ and end_time_i are the start and end time of the trip. A record of APC data apc_i will be denoted as

$$apc_i(trip_i, stop_i, apc\_time_i, board_i, alight_i)$$

, where $trip_i$ is the ID of the transit trip, $stop_i$ is the ID of the transit stop where this APC data is recorded and apc_time_i is the recording time, $board_i$ and $alight_i$ are the counting of boarding and alighting passengers of this transit trip at this transit stop.

Table 3.1 lists all the notations used in this chapter with each description respectively.

Notation	Description
$station_i(id_i, lat_i, lon_i)$	A bike station
id_i	Bike station ID
lat_i	Bike station latitude
lon_i	Bike station longitude
$bt_i(start\_stn_i, start\_time_i,$ $end\_stn_i, end\_time_i, user)$	A bike trip record
$start\_stn_i$	Start station ID
$start\_time_i$	Start time
end_stn_i	End station ID
end_time_i	End time
$user$	User type
$stop_i(id_i, lat_i, lon_i)$	A transit stop
id_i	Transit stop ID
lat_i	Transit stop latitude
lon_i	Transit stop longitude
$apc_i(trip_i, stop_i, apc\_time_i, board_i, alight_i)$	An APC record
$trip_i$	Transit trip ID
$stop_i$	Transit stop ID
apc_time_i	APC recording time
$board_i$	Boarding passenger counting
$alight_i$	Alighting passenger counting
$competitive(bt_i, apc_m, apc_n)$	A competitive relationship case
bt_i	Bike trip
apc_m, apc_n	Transit trip segment
$complementary_{first}(bt_i, apc_j)$	A complementary first-mile trip
$complementary_{last}(bt_i, apc_j)$	A complementary last-mile trip
bt_i	Bike trip
apc_j	APC record
$maxCloseDist$	Maximum distance for bike station being near transit stop
$maxCloseTime$	Maximum time difference for two events happening at similar time
$maxWalkDist$	Maximum distance people prefer to walk in one commuting trip
$maxWalkSpeed$	Maximum speed human prefer to walk

$maxBikeDist$	Maximum biking distance for first/last-mile trip
$maxBikeDuration$	Maximum biking duration for first/last-mile trip
$avgBikeSpeed$	Average biking speed
$eu\_dist(p_i, p_j)$	Euclidean distance between points
$walk\_dist(p_i, p_j)$	Walking distance between points
$bike\_dist(p_i, p_j)$	Biking distance between points
$subgraph_k(Stations_k, Stops_k, Edges_k)$	Subgraph of bike stations and transit stops
$Stations_k$	Set of bike stations
$Stops_k$	Set of transit stops
$Edge_k$	Set of edges
$e(station_i, stop_j)$	Edge between $station_i$ and $stop_j$
$start\_tp_i, end\_tp_i$	Starting/ending temporal profile for $station_i$
$board\_tp_j, alight\_tp_j$	Boarding/alighting temporal profile for $stop_j$
$start\_stp_k, end\_stp_k, board\_stp_k, alight\_stp_k$	Starting/ending/boarding/alighting temporal profile for $subgraph_k$

Table 3.1: Notations

3.2 Competitive and Complementary Relationships

Competitive relationships between bike sharing trips and public transit trips are determined by whether a bike sharing trip potentially replaces a transit trip; complementary relationships are determined by whether the bike sharing trip could provide first/last-mile transit access. Please note that the competitive and complementary relationships are not exclusive; that is, a bike sharing trip can be a first/last-mile trip (complementary) and substitute a transit trip (competitive) at the same time.

For instance, the user may use the shared bike to transfer between two transit stops, whereas the user can also use the local bus to make the transfer. In this case, the bike sharing trip provides last-mile access for the first transit trip and first-mile access for the second transit trip (complementary), and also replaces the transfers using buses (competitive).

3.2.1 Competitive Relationship

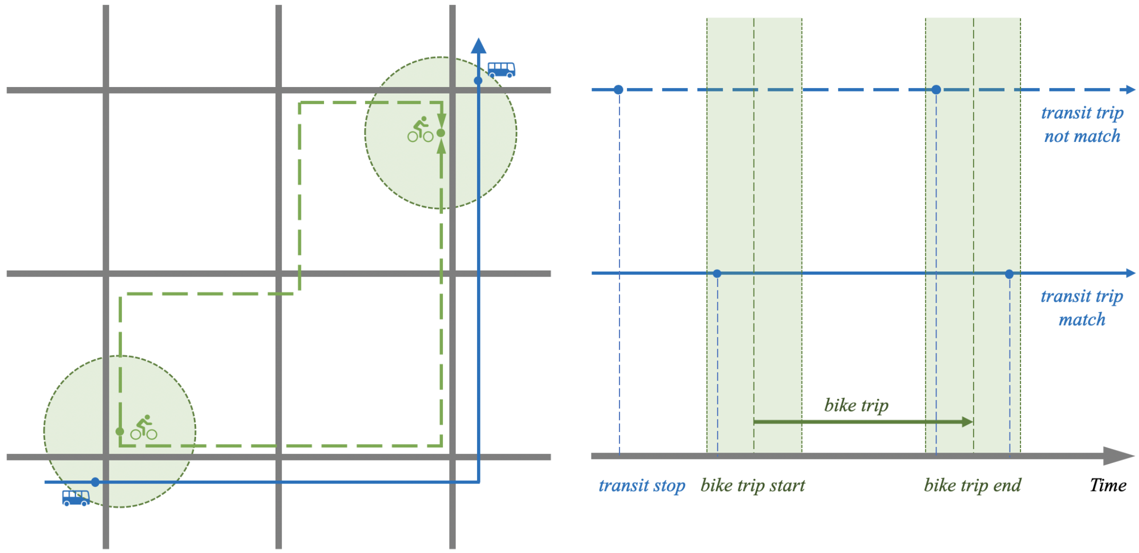
Conceptually, if a bike sharing trip starts and ends near a transit stop respectively and has a similar start and end time with a corresponded transit trip segment, the bike trip is considered to be a competitive trip of such transit trip segment. Unlike bike trips, being a transit trip segment does not mean that there was an actual transit passenger boarding at one transit stop and alighting at the other one. It only means that the transit also has the capability to carry passengers between the two areas at that time. Figure 3.3 shows the spatio-temporal relationship of an example competitive bike trip. Note that being a competitive trip does not mean the bike trip exactly follows the transit route.

Analytically, a case of competitive relationship is denoted as

$$competitive(bt_i, apc_m, apc_n)$$

, which means a bike sharing trip bt_i is a substitution of the transit trip segment from apc_m to apc_n. $competitive(bt_i, apc_m, apc_n)$ exists if:

1. bike trip's origin $start\_stn_i$ is near transit trip segment origin $stop_m$,



(a) Spatial (b) Temporal

Figure 3.3: Example of Competitive Bike Trip

$eu\_dist(start\_stn_i, stop_m) \leq maxCloseDist$

2. bike trip's destination end_stn_i is near transit trip segment destination $stop_n$,

   $eu\_dist(end\_stn_i, stop_n) \leq maxCloseDist$

3. bike trip and transit trip start at similar time,

   $|start\_time_i - apc\_time_m| \leq maxCloseTime$

4. bike trip and transit trip end at similar time,

   $|end\_time_i - apc\_time_n| \leq maxCloseTime$

5. apc_m and apc_n belong to a same transit trip,

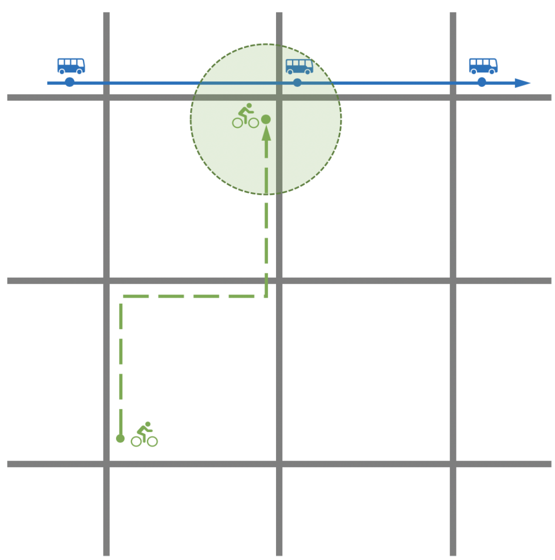
   $trip_m = trip_n$

6. bike trip and transit trip are of the same direction,

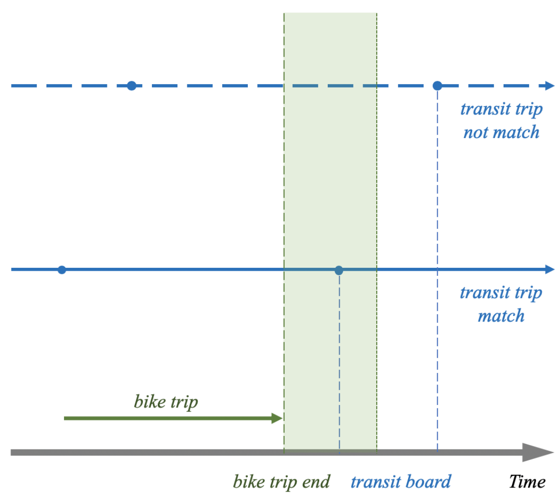
   $apc\_time_m < apc\_time_n$

3.2.2 Complementary Relationship

Conceptually, if a bike sharing trip meets the spatio-temporal criteria of first/last-mile trip, which means the trip is of relatively short distance and duration, and is approaching a transit stop prior to a passenger boarding at that stop (first-mile trip) or leaving a transit stop after a passenger alighting at that stop (last-mile trip), the bike trip is considered to be a complementary trip of such transit trip. Please note that the complementary relationship and first/last-mile trip defined here are only potential: due to the fact that there is no universal customer ID across the two systems, there is no way to confirm the real transferring behavior. Figure 3.4 is an example first-mile and Figure 3.5 is an example last-mile bike trip.



(a) Spatial

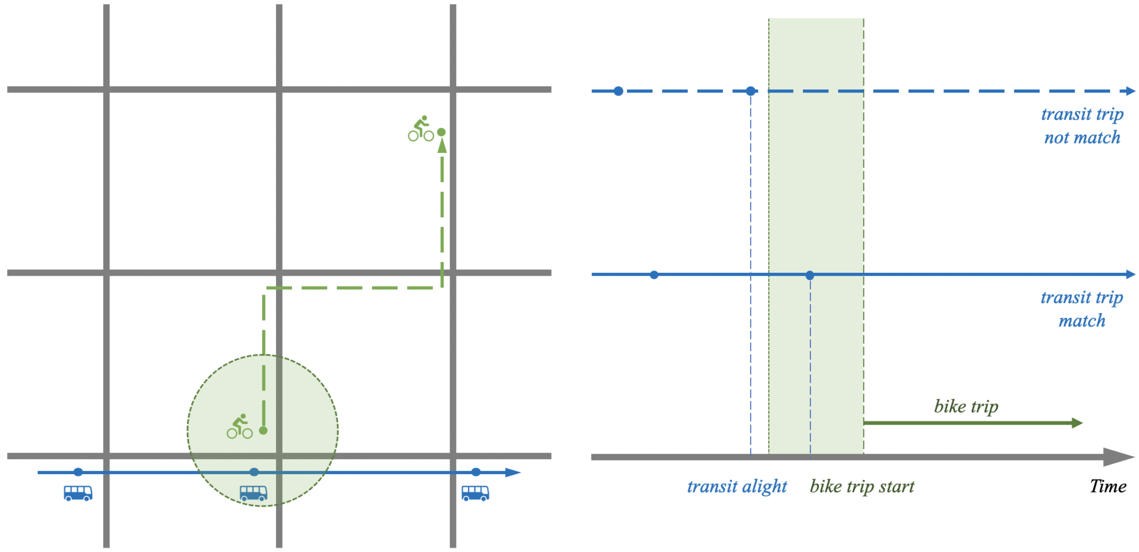


(b) Temporal

Figure 3.4: Example of First-mile Bike Trip

Analytically, a case of first-mile complementary relationship is denoted as

$$complementary_{first}(bt_i, apc_j)$$



(a) Spatial (b) Temporal

Figure 3.5: Example of Last-mile Bike Trip

, which means bike sharing trip bt_i meets the spatio-temporal criteria of first-mile trip of the boarding passengers of apc_j.

$complementary_{first}(bt_i, apc_j)$ exists if:

1. bike trip bt_i is of short distance,

   $bike\_dist(start\_stn_i, end\_stn_i) \leq maxBikeDist$

2. bike trip bt_i is of short duration or above average speed,

   $end\_time_i - start\_time_i \leq maxDuration$, OR

   $bike\_dist(start\_stn_i, end\_stn_i)/(end\_time_i - start\_time_i) \geq avgBikeSpeed$

3. bike trip's destination end_stn_i is near transit stop $stop_j$,

   $walk\_dist(end\_stn_i, stop_j) \leq maxWalkDist$

4. bike trip ends shortly before transit vehicle stops,

   $0 < apc\_time_j - end\_time_i \leq maxCloseTime$

5. there are passenger(s) boarding at transit stop,

   $board_j > 0$

6. the walking speed needed to get to transit stop from bike station is under human walking speed capability,

   $walk\_dist(end\_stn_i, stop_j)/(apc\_time_j - end\_time_i) \leq maxWalkSpeed$

Similarly, a case of last-mile complementary relationship is denoted as

$$complementary_{last}(bt_i, apc_j)$$

, which means bike sharing trip bt_i meets the spatio-temporal criteria of last-mile trip of the alighting passengers of apc_j. $complementary_{last}(bt_i, apc_j)$ exists if:

1. bike trip bt_i is of short distance,

   $bike\_dist(start\_stn_i, end\_stn_i) \leq maxBikeDist$

2. bike trip bt_i is of short duration or above average speed,

   $end\_time_i - start\_time_i \leq maxDuration$, OR

   $bike\_dist(start\_stn_i, end\_stn_i)/(end\_time_i - start\_time_i) \geq avgBikeSpeed$

3. bike sharing trip's origin $start\_stn_i$ is near transit stop $stop_j$,

   $walk\_dist(start\_stn_i, stop_j) \leq maxWalkDist$

4. bike sharing trip starts shortly before transit vehicle stops,

   $0 < start\_time_i - apc\_time_j \leq maxCloseTime$

5. there are passenger(s) alighting at transit stop,

   $alight_j > 0$

6. the walking speed needed to get to bike station from transit stop is under human walking speed capability,

   $walk\_dist(start\_stn_i, stop_j)/(start\_time_i - apc\_time_j) \leq maxWalkSpeed$

In the experiment, the data shows that the bike trip duration of a certain OD distance varies from several to tens of minutes. Therefore, in the definitions above, it is reasonable to limit the potential first/last-mile bike trips to the ones of both short distance and duration. The duration criterion is aimed to identify the trips that go relatively directly from the start station to the end station rather than including some intermediate stops for activities along its way, which is a crucial feature for first/last-mile trips. However, it is not reasonable to set an arbitrary duration threshold for all trips within the distance threshold. Hence, I use speed in addition to duration to detect the relatively direct bike trips. It is based on the observation that bike sharing customers usually ride at a similar speed, regardless of gender and age, and the biking route affects most on the duration. Therefore, a bike trip is more likely to be direct if it is not slower than the common, average speed, i.e., $avgBikeSpeed$.

3.2.3 Summary of Relationship Definitions

Competitive and complementary relationships are defined from both spatial and temporal perspectives. I use different distance metrics for the spatial criteria while examining the relationships. First, $maxCloseDist$ is a threshold for a bike station being close to a bus stop and vice versa. The threshold is used in the competitive relationship definition. Since no person or bike actually transfers between bike station and bus stop in competitive relationship, the distance is defined in Euclidean distance. Second, $maxWalkDist$ is a threshold for the distance people willing to walk in a bike-transit transfer. The threshold is used in the complementary relationship definition. Since the distance represents the walking transfer between bike station and transit stop in a first/last-mile trip, it is defined in walking network distance. Third, $maxBikeDist$ is a threshold for the first/last-mile bike trip distance. It is defined in biking network distance.

3.3 Ridership Relationship

In addition to defining relationships and detecting corresponding trips in 3.2, the ridership at bike stations and transit stops is investigated next. This step is designed to (1) validate the detected competitive and complementary relationships, especially to further discuss the detected potential first/last-mile trips; (2) detect the potential significant correlations between the two systems in terms of ridership ; (3) explore whether the individual competitive/complementary trips affect ridership relationships. Intuitively, if the first-mile relationship is significant among some bike stations and transit stops, there might be a positive ridership correlation between bike trip ending and transit boarding; similarly, if the last-mile relationship is significant, there might be a positive correlation between transit alighting and bike trip starting. If the ridership correlation is statistically significant, it is reasonable for us to make a stronger statement about the existence of such relationship. It also allows us to examine the two relationships at the same time and investigate the potential existence of duo-relationships.

3.3.1 Related Graph

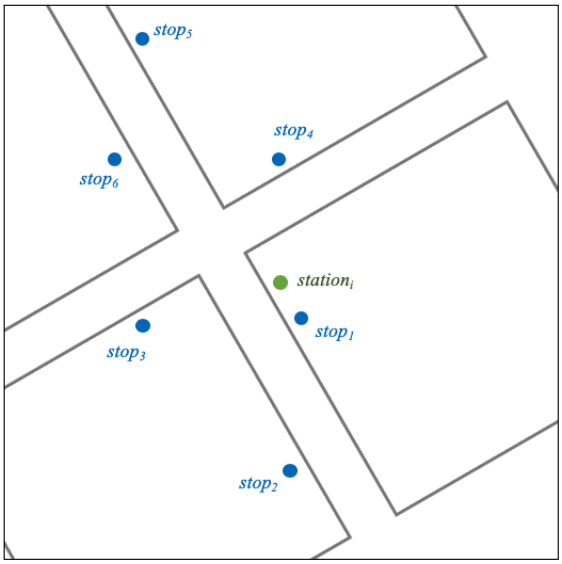
To investigate the ridership relationships between bike stations and transit stops, it is crucial to determine which station(s) and stop(s) are potentially related. Using the cases of competitive and potential first/last-mile relationship detected in 3.2, a relationship-based graph is built first. Then, two other graphs are built using buffer search and K-Nearest Neighbors (KNN) (Cover & Hart, 1967) search as comparisons. These three kinds of graphs are denoted as $G_{relation}$, G_{knn} and G_{buffer}, where

$$G_{relation} = (V, E_{relation})$$

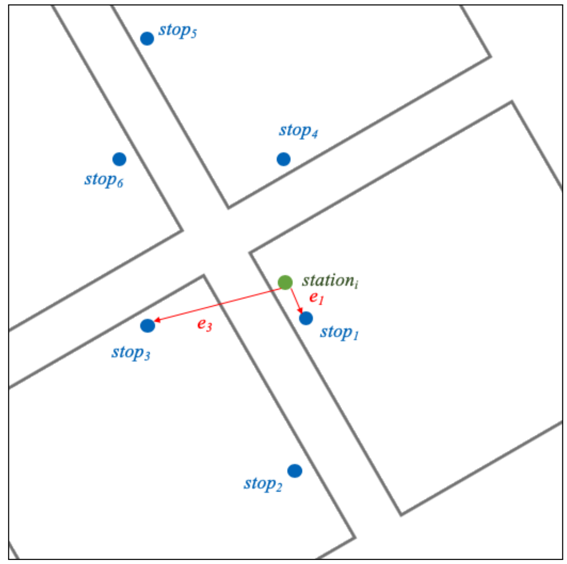
$$G_{buffer} = (V, E_{buffer})$$

$$G_{knn} = (V, E_{knn})$$

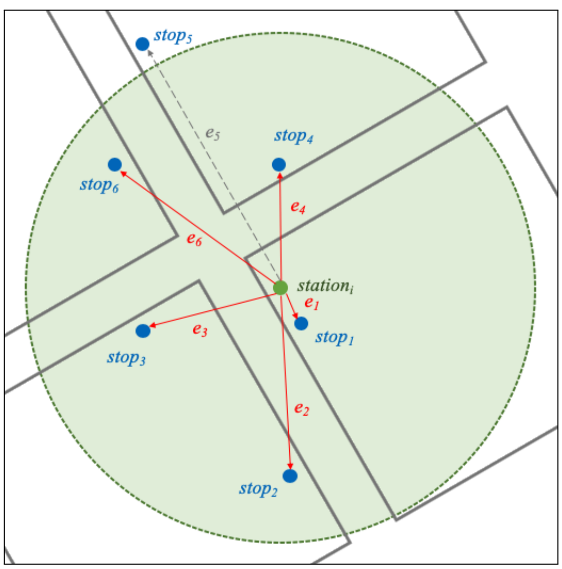
These three graphs have same set of vertices $V = Stations \cup Stops$, which is the union of bike station set and transit stop set. $E_{relation}$, E_{buffer} and E_{knn} are edges derived from relationships, buffer search and KNN search respectively. Connected components ("Component (graph theory)", n.d.; Hagberg et al., 2008) are derived for each graph. Each connected component is considered to be potentially inner-related and is studied as a whole. Figure 3.6 is an example for these three kinds of graphs under a synthetic spatial layout shown in Figure 3.6a.



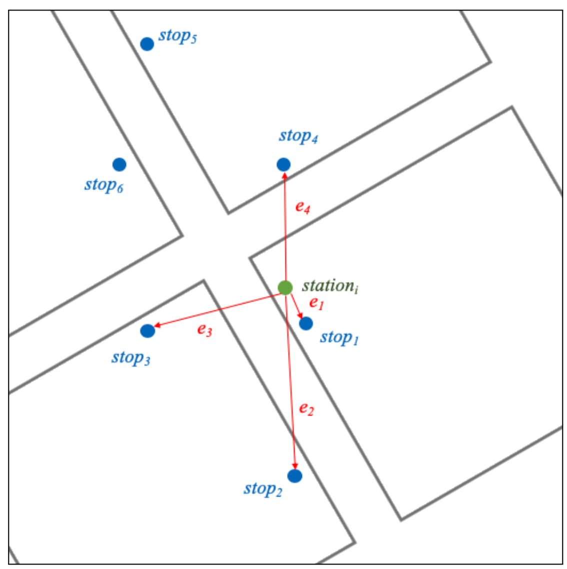
(a) Synthetic Layout



(b) Relationship-based



(c) Buffer-based



(d) KNN-based, when k=4

Figure 3.6: Graphs Generated by Different Methods

The synthetic layout in (a) is a road intersection where the grey blocks represent street blocks. Supposing relationship cases are between $(station_i, stop_1)$ and $(station_i, stop_3)$, so edges $e(station_i, stop_1)$ and $e(station_i, stop_3)$ are included in (b). All stops within given radius are included in (c). The k nearest stops of $station_i$ are included in (d), where $k = 4$.

Relationship-based Graph $E_{relation}$ is made up with edges derived from all the relationship cases detected in 3.2. From a competitive relationship case $cpt_i(bt_i, apc_m, apc_n)$, two edges are derived, $e(start\_stn_i, stop_m)$ and $e(end\_stn_i, stop_n)$. Similarly, a complementary first-mile relationship case $complementary_{first}(bt_i, apc_j)$ generates one edge $e(end\_stn_i, stop_j)$ and a last-mile relationship case $complementary_{last}(bt_i, apc_j)$ generates one edge $e(start\_stn_i, stop_j)$. As shown in Figure 3.6b, supposing the relationships are detected between $(station_i, stop_1)$ and $(station_i, stop_3)$, then edges $e(station_i, stop_1)$ and $e(station_i, stop_3)$ are included in $E_{relation}$.

Buffer-based Graph To build a buffer-based graph, for a bike station $station_i$, all transit stops within a radius of $maxWalkDist$ are considered to be potentially related stops. As shown in Figure 3.6c, for a given radius, $stop_1$, $stop_2$, $stop_3$, $stop_4$, $stop_6$ are within the buffer of $station_i$, then edges $e(station_i, stop_1)$, $e(station_i, stop_2)$, $e(station_i, stop_3)$, $e(station_i, stop_4)$ and $e(station_i, stop_6)$ are included in E_{buffer}.

KNN-based Graph To build a KNN-based graph, for a bike station $station_i$, its k nearest transit stops $(stop_1, stop_2, ..., stop_k)$ are queried first. For any $stop_j \in (stop_1, stop_2, ..., stop_k)$, edge $e(station_i, stop_j)$ will be pruned if $walk\_dist(station_i, stop_j) > maxWalkDist$. As shown in Figure 3.6d, when $k = 4$, the 4 nearest transit stops of $station_i$ are $stop_1, stop_2, stop_3, stop_4$, so edges $e(station_i, stop_1)$, $e(station_i, stop_2)$, $e(station_i, stop_3)$ and $e(station_i, stop_4)$ are included in E_{knn}.

3.3.2 Ridership Temporal Profile

The ridership of bike sharing has two topics: bike-start and bike-end; and the ridership of transit has two topics: transit-board and transit-alight. The ridership of each bike station or transit stop is managed in the form of temporal profile, which is a $|\mathbb{D}| \times |\mathbb{H}|$ matrix, where each element represents the ridership at that station/stop at a given temporal resolution. $\mathbb{D}$ is the set of days of study, and $\mathbb{H}$ is the set of equal length time periods of a day. $|\mathbb{D}|$ and $|\mathbb{H}|$ are the cardinalities of $\mathbb{D}$ and $\mathbb{H}$. $\mathbb{H}$ is determined based on the temporal resolution of study. For example, if the resolution is set to one hour, then $\mathbb{H} = (00:00 - 00:59, 01:00 - 01:59, ..., 23:00 - 23:59)$. $\mathbb{D} = (d_1, d_2, ...d_m)$ and $\mathbb{H} = (h_1, h_2, ...h_n)$ are used for the following description.

For a bike station $station_i$, it has two temporal profiles $start\_tp_i$ and end_tp_i for starting ridership and ending ridership respectively.

$$start\_tp_i = \begin{pmatrix} start_{(i,d_1,h_1)} & start_{(i,d_1,h_2)} & ... & start_{(i,d_1,h_n)} \\ start_{(i,d_2,h_1)} & start_{(i,d_2,h_2)} & ... & start_{(i,d_2,h_n)} \\ ... & ... & ... & ... \\ start_{(i,d_m,h_1)} & start_{(i,d_m,h_2)} & ... & start_{(i,d_m,h_n)} \end{pmatrix}$$

$$end\_tp_i = \begin{pmatrix} end_{(i,d_1,h_1)} & end_{(i,d_1,h_2)} & ... & end_{(i,d_1,h_n)} \\ end_{(i,d_2,h_1)} & end_{(i,d_2,h_2)} & ... & end_{(i,d_2,h_n)} \\ ... & ... & ... & ... \\ end_{(i,d_m,h_1)} & end_{(i,d_m,h_2)} & ... & end_{(i,d_m,h_n)} \end{pmatrix}$$

, where $start_{(i,d,h)}$ represents the total starting ridership at the station $station_i$, on the day d, in the time window h; and $end_{(i,d,h)}$ represents the total ending ridership at the station $station_i$, on the day d, in the time window h, where $d \in \mathbb{D}$ and $h \in \mathbb{H}$.

Similarly, for a transit stop $stop_i$, it has two temporal profiles as well, $board\_tp_i$ for boarding ridership and $alight\_tp_i$ for alighting ridership.

$$board\_tp_i = \begin{pmatrix} board_{(i,d_1,h_1)} & board_{(i,d_1,h_2)} & ... & board_{(i,d_1,h_n)} \\ board_{(i,d_2,h_1)} & board_{(i,d_2,h_2)} & ... & board_{(i,d_2,h_n)} \\ ... & ... & ... & ... \\ board_{(i,d_m,h_1)} & board_{(i,d_m,h_2)} & ... & board_{(i,d_m,h_n)} \end{pmatrix}$$

$$alight\_tp_i = \begin{pmatrix} alight_{(i,d_1,h_1)} & alight_{(i,d_1,h_2)} & ... & alight_{(i,d_1,h_n)} \\ alight_{(i,d_2,h_1)} & alight_{(i,d_2,h_2)} & ... & alight_{(i,d_2,h_n)} \\ ... & ... & ... & ... \\ alight_{(i,d_m,h_1)} & alight_{(i,d_m,h_2)} & ... & alight_{(i,d_m,h_n)} \end{pmatrix}$$

, where $board_{(i,d,h)}$ represents the total boarding ridership at the stop $stop_i$, on the day d, in the time window h; and $alight_{(i,d,h)}$ represents the total alighting ridership at the stop $stop_i$, on the day d, in the time window h, where $d \in \mathbb{D}$ and $h \in \mathbb{H}$.

3.3.3 Ridership Relationship and Correlation

Bike stations and transit stops are grouped into connected components in 3.3.1. For each subgraph of connected component, aggregated temporal profiles are generated for each of the topics: bike-start, bike-end, transit-board and transit-alight.

Analytically, a subgraph is denoted as $subgraph_i(Stations_i, Stops_i, Edges_i)$, where

$Stations_i$ is the set of bike stations, $Stops_i$ is the set of transit stops and $Edges_i$ is the set of edges between them. Then the temporal profiles for $subgraph_i$ are $start\_stp_i$, end_stp_i, $board\_stp_i$, $alight\_stp_i$, where

$$start\_stp_i = \sum_{m \in Stations_i} start\_tp_m$$

$$end\_stp_i = \sum_{m \in Stations_i} end\_tp_m$$

$$board\_stp_i = \sum_{n \in Stops_i} board\_tp_n$$

$$alight\_stp_i = \sum_{n \in Stops_i} alight\_tp_n$$

We then use the Pearson correlation coefficient ("Pearson correlation coefficient", n.d.) to investigate the ridership relationship between bike sharing and transit systems within each subgraph. Four pairs of topics are investigated: (bike-start, transit-board), (bike-start, transit-alight), (bike-end, transit-board), (bike-end, transit-alight).

For example, subgraph temporal profile $start\_stp_i$ and $board\_stp_i$ have correlation coefficients calculated for $\mathbb{D} = (d_1, d_2, ...d_m)$

$$P_i^{start-board} = (\rho_{(i,d_1)}^{start-board}, \rho_{(i,d_2)}^{start-board}, ..., \rho_{(i,d_m)}^{start-board})$$

. The Pearson correlation coefficient for $d \in \mathbb{D}$ is

$$\rho_{(i,d)}^{start-board} = \frac{\sum_{h \in \mathbb{H}} (start_{(i,d,h)} - sta\bar{r}t_{(i,d)})(board_{(i,d,h)} - boa\bar{r}d_{(i,d)})}{\sqrt{\sum_{h \in \mathbb{H}} (start_{(i,d,h)} - sta\bar{r}t_{(i,d)})^2 \sum_{h \in \mathbb{H}} (board_{(i,d,h)} - boa\bar{r}d_{(i,d)})^2}}$$

$$\bar{start}_{(i,d)} = \frac{1}{|\mathbb{H}|} \sum_{h \in \mathbb{H}} start_{(i,d,h)}$$

$$\bar{board}_{(i,d)} = \frac{1}{|\mathbb{H}|} \sum_{h \in \mathbb{H}} board_{(i,d,h)}$$

.

Chapter 4

Result and Discussion

The data of Minneapolis-St. Paul area in the year of 2017 is used as a case study. The bike sharing data is provided by Nice Ride[1], the bike sharing service provider in the Twin Cities area. The transit data is provided by Metro Transit[2], the primary public transportation system in the Twin Cities area.

4.1 System Overview

In the year of 2017, the system of Nice Ride bike sharing service was open from April 3 to November 5, with 460,718 trips recorded among 202 available stations. Nice Ride has users of two types: member, who has an account with Nice Ride with annual or monthly membership; and casual user, who purchased the pass at the station. Being a member of Nice Ride means that a rider does not need to pay for each particular bike trip. Economically, the more trips a member rides, the lower cost per trip is.

[1] www.niceridemn.com
[2] www.metrotransit.org

During the same time period, the Metro Transit bus system collected 73,477,261 APC data records among 13,782 transit stops, involving 33,245,899 person-time boardings and 33,171,510 person-time alightings. Table 4.1 lists some basic information of the two systems. Figure 4.1 shows the spatial distribution of bike sharing stations and transit stops.

Information	**Value**
# of bike sharing stations	202
# of bike sharing trips	460,718
# of casual trips	170,646(37.0%)
# of member trips	290,070(63.0%)
# of transit stops	13,782
# of APC records	73,477,261
# of boarding person-time	33,245,899
# of alighting person-time	33,171,510
study case dates	2017-04-03 ∼ 2017-11-05
# of days	217

Table 4.1: Basic Systems Information

Figure 4.2 and 4.3 show the counting of trip starting and ending at each bike station. Station-level speaking, bike stations in the University of Minnesota campus area and around Lake Bde Maka Ska have the highest (over 10,000) starting and ending counting, which indicates the two main purposes of bike sharing trips are potentially commuting to/from UMN campus and recreation around the lake. More specifically, the 20 (10.0%) bike stations in the University of Minnesota neighborhood are involved in 137,167 (29.8%) bike sharing trips. And the 20 (10.0%) bike stations near lakes, rivers and parks areas are involved in 94,538 (20.5%) trips. The bike stations in downtown Minneapolis do not have the highest starting/ending counting, partially because the ridership is dispersed by the dense distribution of stations. The ridership density at that area is still high and is gradually decreasing from the center

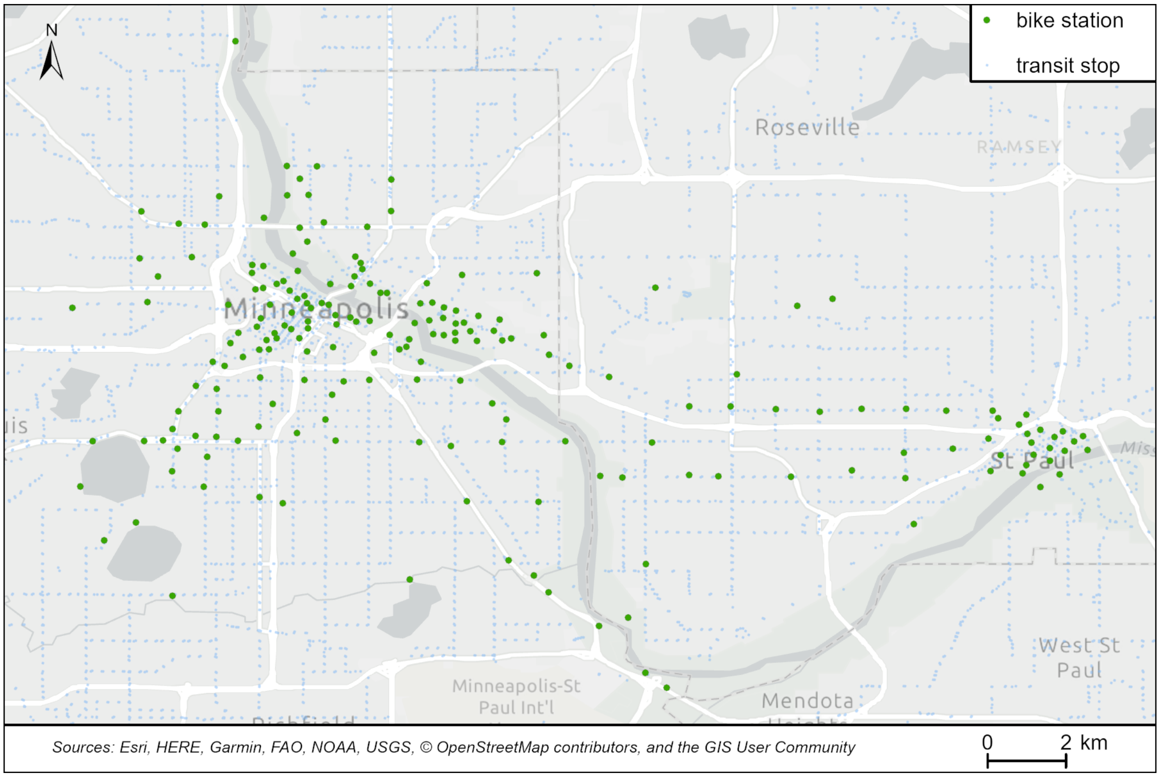


Figure 4.1: Bike Stations and Transit Stops

to the surrounding area.

Figure 4.4 shows the top 20 origin-destinations (OD) of all bike sharing trips. One thing noticeable is that more than half of the top 20 ODs are round trip, which means the origin and destination of the bike trip is the same. Of all bike trips, 55,306 (12.0%) of them are round trips. And the round trips are significantly concentrated at the bike stations near lakes (Bde Maka Ska and Harriet), parks (Minnehaha Falls, Minnehaha Creek and Como), and the Mississippi River. Again, it indicates that one of the main purposes of bike sharing trips in the Twin Cities area is recreation, which has little relationship with the transit system, and is not of the interest of this thesis.

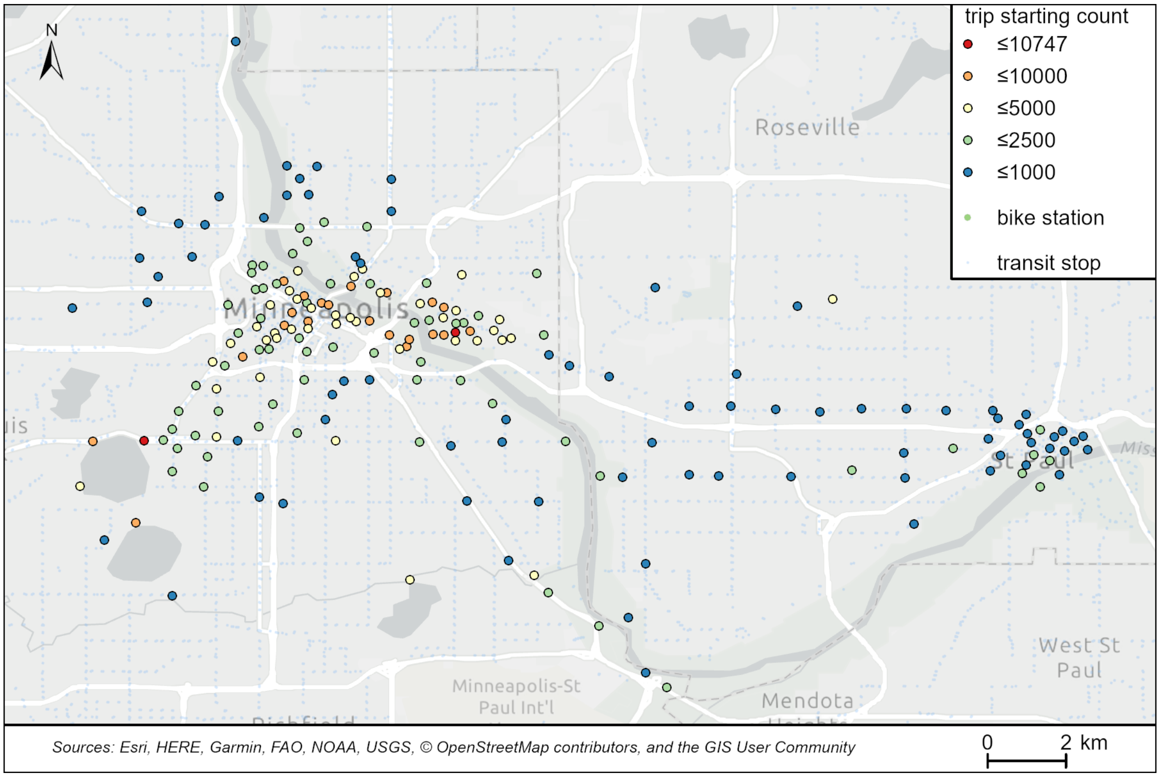


Figure 4.2: Bike Trip Starting Counting

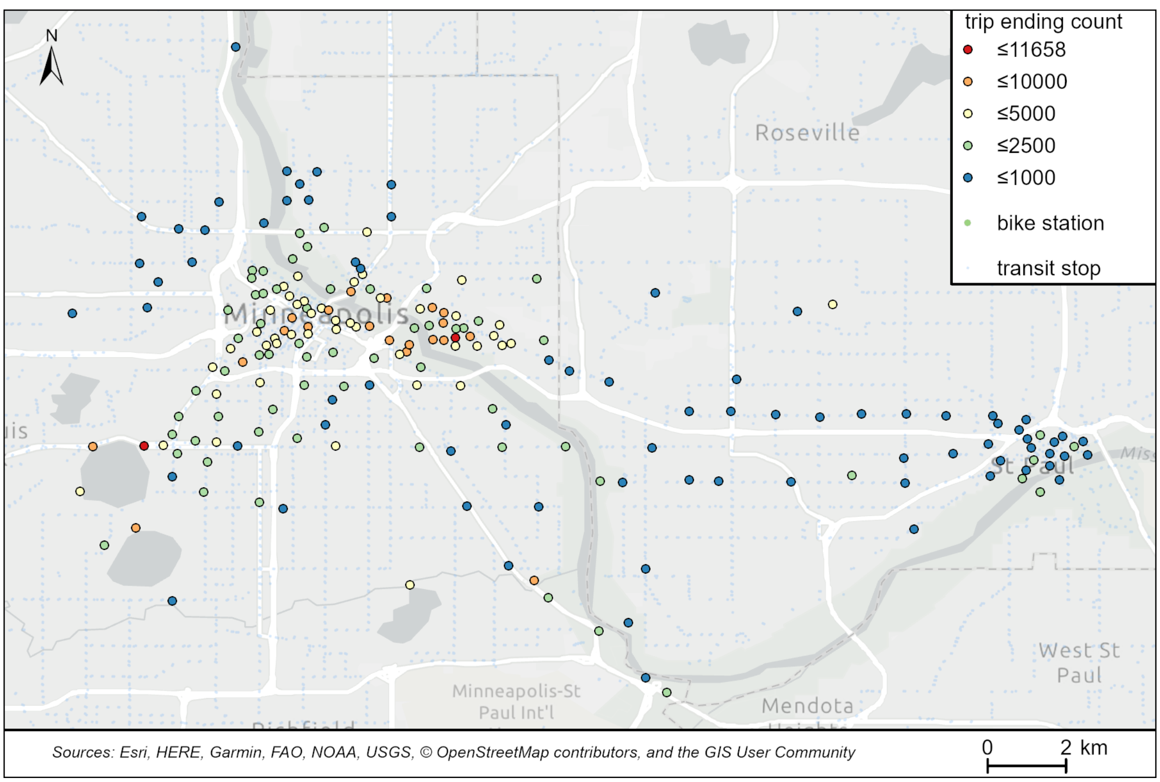


Figure 4.3: Bike Trip Ending Counting

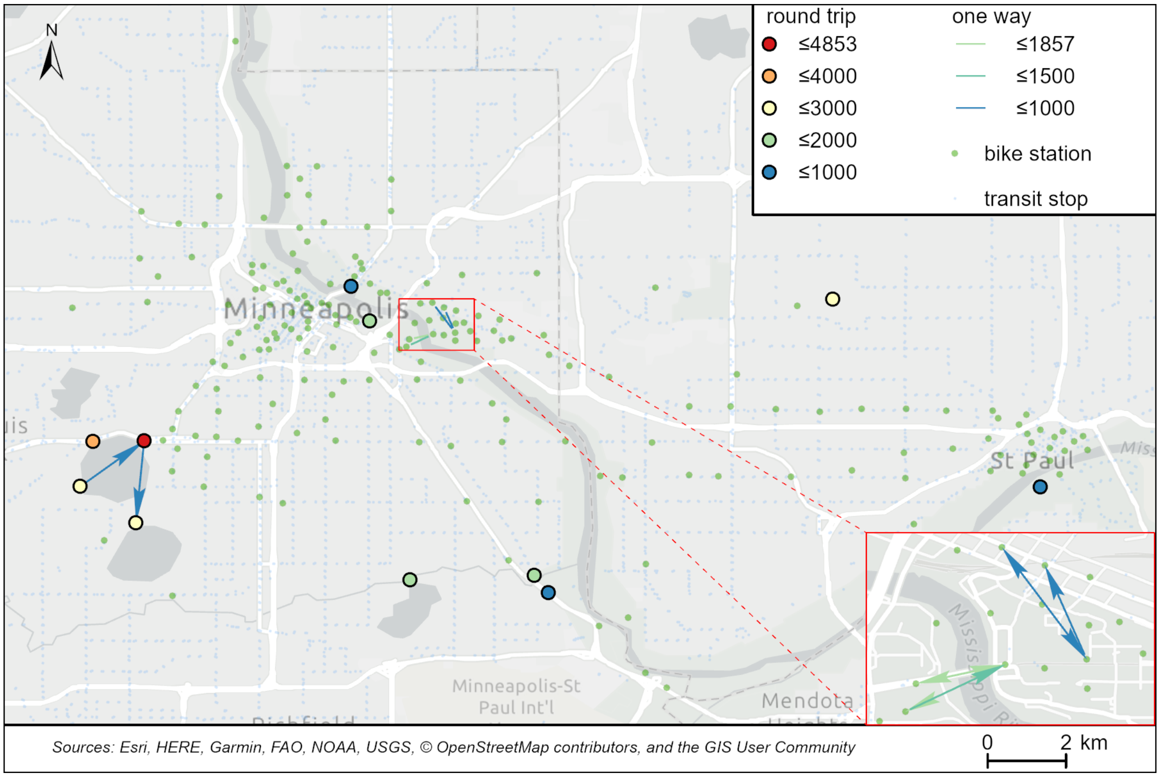


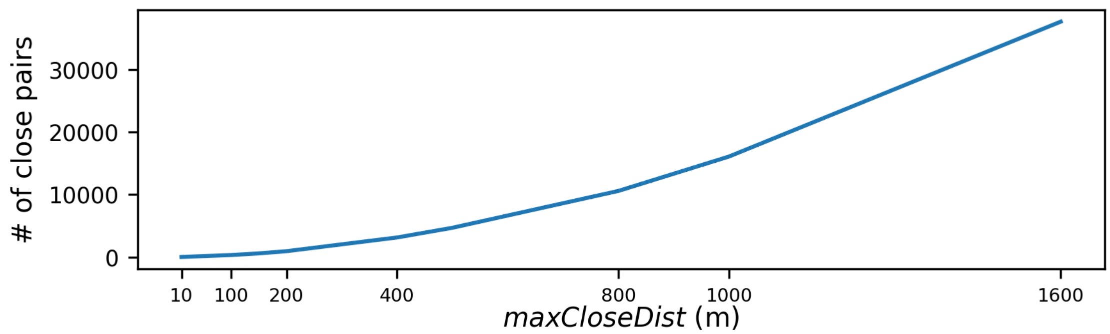
Figure 4.4: Bike Trip OD Counting

4.2 Study 1: Competitive Trips

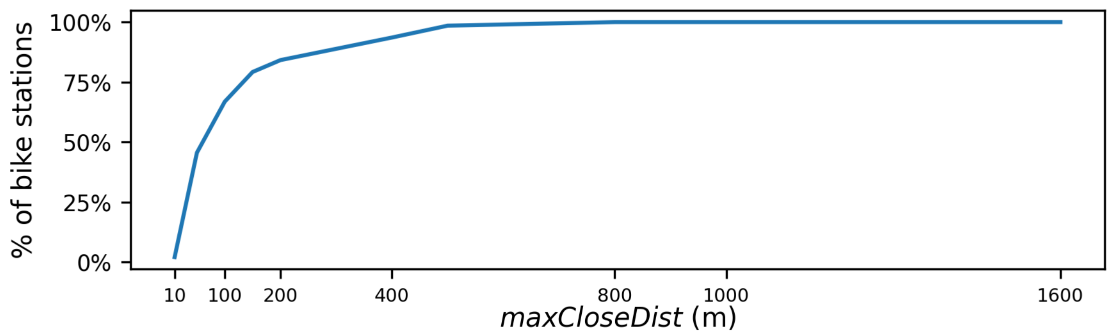
4.2.1 Parameters Determination

$maxCloseDist$ and $maxCloseTime$ are needed to detect the competitive bike sharing trips. $maxCloseDist$ represents the maximum distance of which bike stations and transit stops are considered to be spatially close to each other, which is in Euclidean distance. Figure 4.5a shows the number of close pairs of bike station and transit stop, and Figure 4.5b shows the percentage of bike stations involved in the close pairs. As the figures suggest, all of the 202 bike stations are within half mile (800 meters) of at least one transit stop, and the number of close pairs still grows after half mile.

$maxCloseTime$ represents the maximum time difference between two events that



(a) Number of Close Pairs

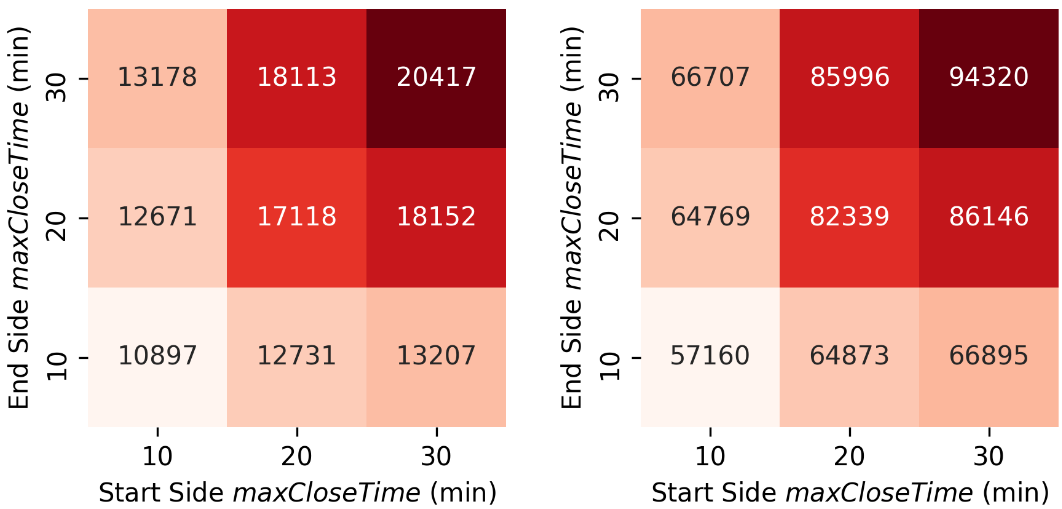


(b) Percentage of Bike Stations

Figure 4.5: $maxCloseDist$ Sensitivity

are considered to happen at similar time. In competitive relationship definition, the two events are bike trip starting and transit boarding on the start side, and bike trip ending and transit alighting on the end side. Figure 4.6 shows how many bike sharing trips are detected as competitive trips using different $maxCloseTime$. $maxCloseDist$ is set to 100 meters in Figure 4.6a and 200 meters in Figure 4.6b. When $maxCloseDist$ is set to 400 meters, the detection runs too long to get the result. As the figures suggest, the amount of competitive trips has a similar sensitivity to $maxCloseTime$ on both start and end side.

In this study, I detect the strict, deterministic competitive relationship. $maxCloseDist$ is set to 100 meters, which is also the length of the short edge of a common city block in the Twin Cities area. *maxCloseTime* is set to 10 minutes, which is the common time interval of a bus route.



(a) $maxCloseDist = 100m$ (b) $maxCloseDist = 200m$

Figure 4.6: $maxCloseTime$ Sensitivity

4.2.2 Competitive Trips Result and Discussion

When $maxCloseDist$ is set to 100 meters and $maxCloseTime$ is set to 10 minutes at both start and end side, 16,848 pairs of bike sharing trip and transit trip segment are detected as competitive relationship cases. 10,897 (2.4% of all) bike sharing trips are involved in the 16,848 cases.

According to the detected cases, the characteristics of competitive trips are significantly different from all bike trips in many ways. (1) The ratio of member trips in competitive trips (82.4%) is higher than the ratio in the entire bike trips (63.0%). (2) Spatially speaking, as starting/ending counting of competitive trips shown in Figure 4.7 and 4.8 and OD counting shown in Figure 4.9, the hot bike stations for compet-

Parameter	Value
maxCloseDist	100 m
maxCloseTime	10 min
Result	
# of competitive relationship cases	16,848
# of competitive bike trips	10,897 (2.4%)
# of casual bike trips	1,916 (17.6%)
# of member bike trips	8,981 (82.4%)

Table 4.2: Competitive Relationship Detection Parameters and Result

itive trips are stations in the University neighborhood and North Loop. Comparing with starting/ending counting of all bike trips shown in Figure 4.2 and 4.3, the bike stations in center Downtown Minneapolis are not top starting/ending stations anymore. (3) Temporally speaking, competitive trips are different from the entire trips in both day-of-week and duration distributions. As shown in Figure 4.10, for the entire bike trip, more than 30% of them happens on the weekends. However, for competitive trips, the proportion goes down to less than 15%. Also, trips within 5 minutes make up about 15% of the entire bike trips, while the proportion is 41% for competitive trips as shown in Figure 4.11. Figure 4.12 indicates that the starting time of competitive bike trips shares a similar time-of-day distribution with the entire trips, but still, its peaks in the morning, noon and evening are more significant. Because of the significant differences in user composition, spatial and temporal characteristics, it is fair to say that substituting transit trip with bike sharing trip does exist. Although it makes up only 2.4% of the entire trip, it has distinct spatial and temporal features.

A comparison between competitive bike trips with corresponding transit trips may explain why some passengers would ride a bike rather than take transit. An intuitive thought would be people prefer to ride a bike because bikes travel faster than buses,

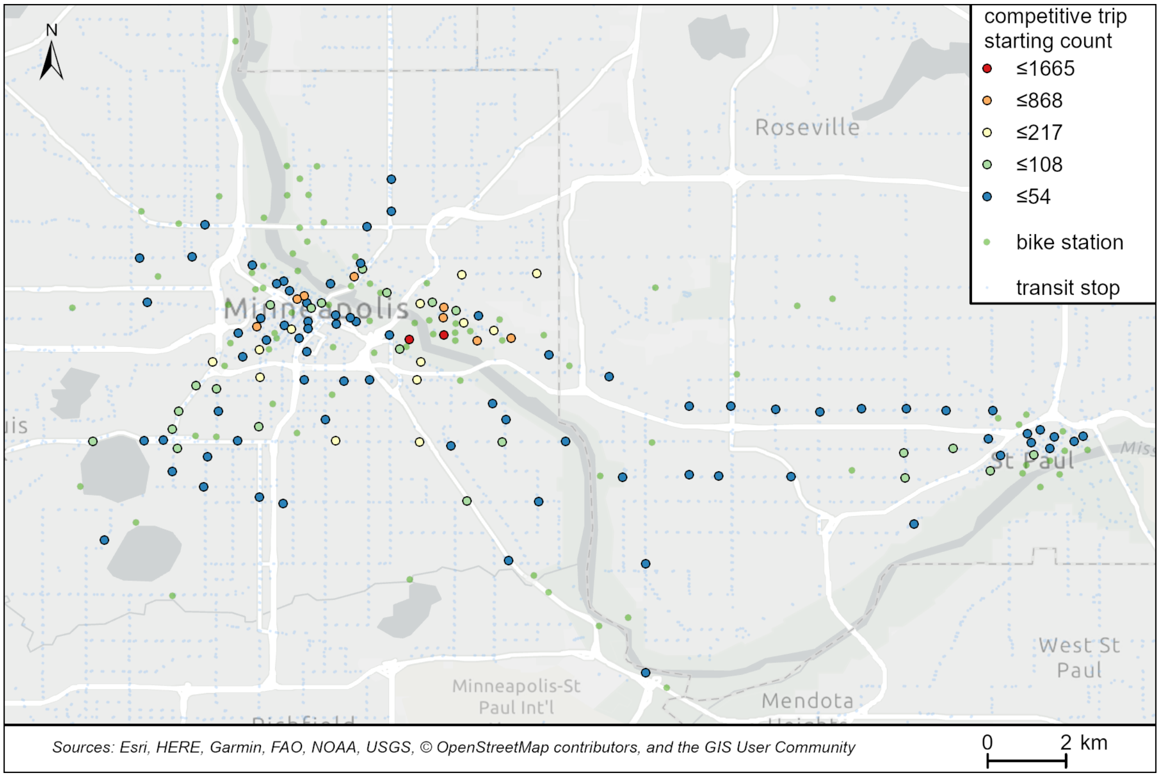


Figure 4.7: Competitive Trip Starting Counting

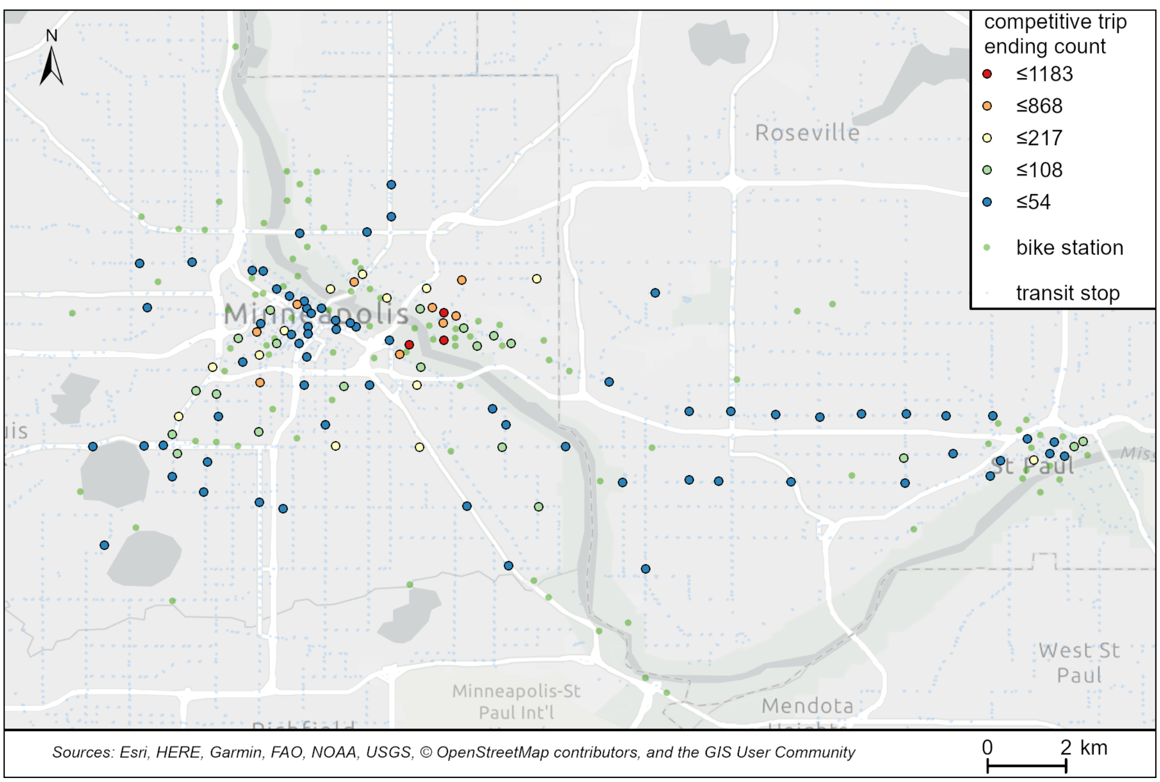


Figure 4.8: Competitive Trip Ending Counting

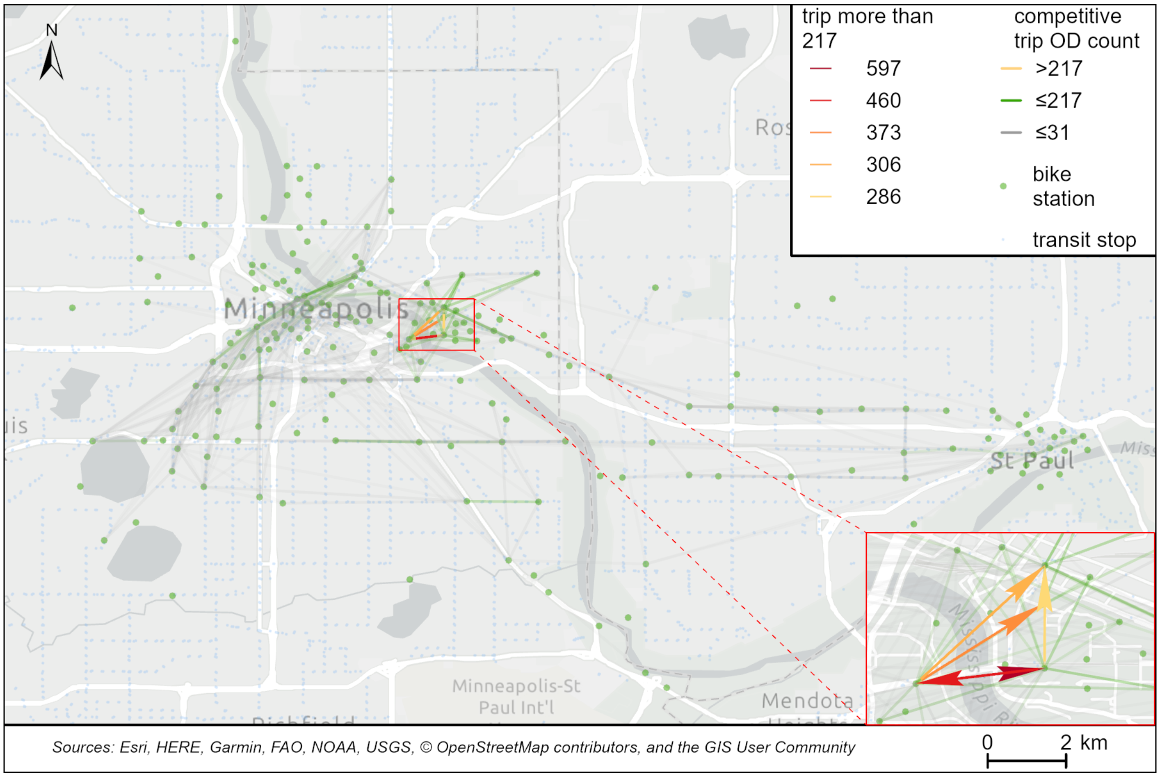
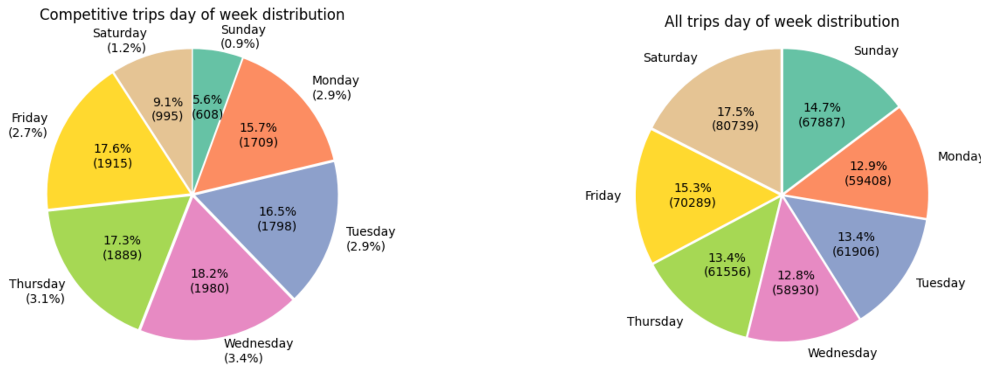


Figure 4.9: Competitive Trip OD Counting

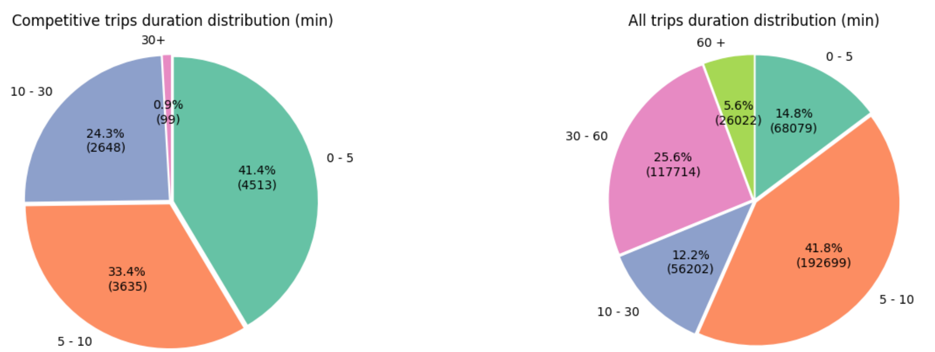
especially when there is traffic congestion. Figure 4.13a shows the duration differences between competitive bike trips with corresponding transit trips. Strictly speaking, only one-third of the competitive trips are faster than corresponding transit trips. Usually, people are not strictly sensitive to the time difference, so if a category of "similar" is added to the comparison, the proportion of bike trips faster than transit trips is even lower. Figure 4.13b shows the result of comparison under three kinds of criteria, 2 minutes, 5 minutes and 20% of corresponding transit duration. For example, under the criteria of 2 minutes, only the bike trips 2 minutes or more faster than corresponding transit trips will be regarded as faster trips. No matter under what kinds of criteria, bike trips faster than transit are fewer than the trips slower than transit. This indicates that traveling faster may be one of the reasons for replacement,



(a) Competitive Trips (b) All Trips

Figure 4.10: Competitive Trip Day of Week Distribution

In Figure 4.10a, the percentage within each pie means the percentage of daily competitive trips of all competitive trips; the percentage outside of each pie means the percentage of competitive trips of all trips on each day. For example, 15.7% of competitive trips are on Monday, which are 2.9% of all Monday trips.



(a) Competitive Trips (b) All Trips

Figure 4.11: Competitive Trip Duration Distribution

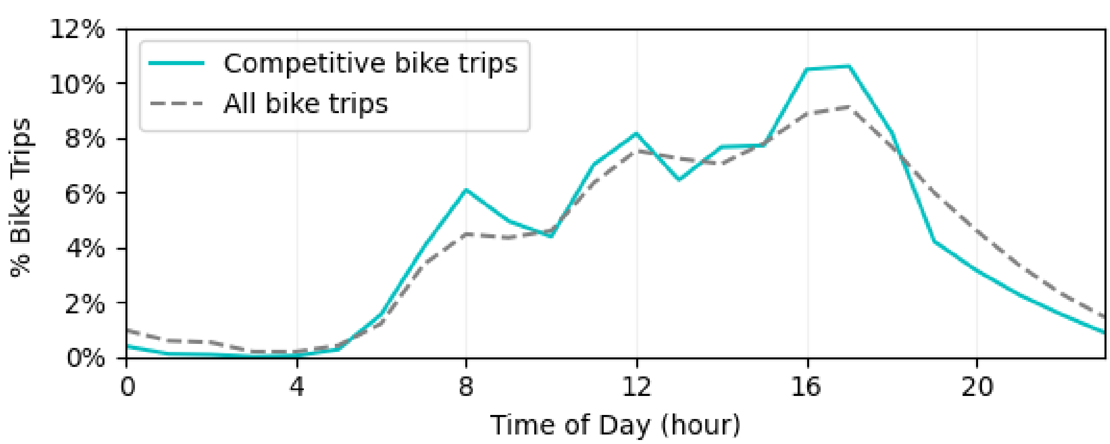
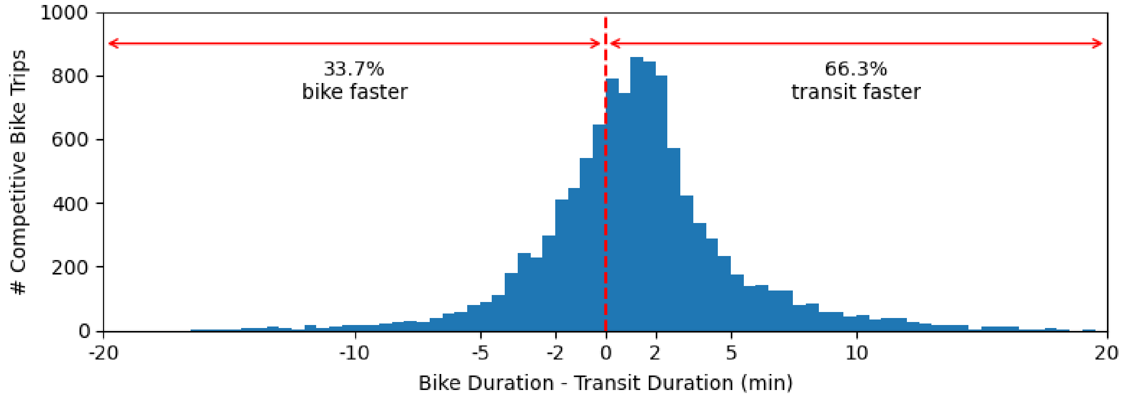


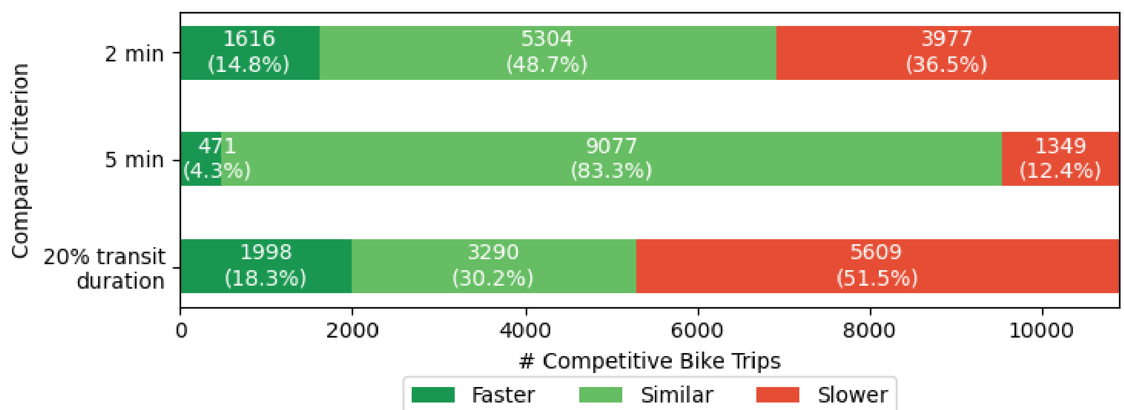
Figure 4.12: Competitive Trip Time of Day Distribution

but it is not always the fact. However, this only compares the duration of bike and transit trips. If we take the time people waiting for the transit into consideration, which is about 5 minutes on average, the result will be contrary. This indicates that it is not the bike trip itself that is faster than the transit trip, but the time people spend on a bike trip is shorter than the time people spend on a transit trip.

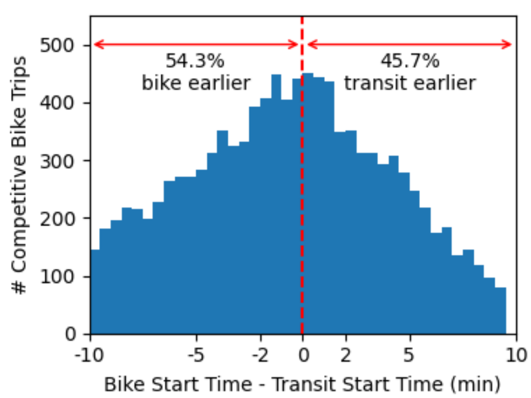
I then compare the starting/ending time between the competitive bike trip and its corresponding transit trip. As Figure 4.13c and 4.13d shown, more than half of the competitive bike trips start earlier than corresponding transit, but only fewer than half bike trips actually arrive earlier than transit. This indicates that the temporal high availability of bike sharing may be a reason why passengers take bike instead of transit. When there is a need to travel and both bike sharing and transit are options, usually bike sharing is immediately available, while people need to wait for a bus to come. Thus there is a possibility that, for short distance travel, people tend to start to move instead of staying and waiting for a bus, especially for the Nice Ride members who do not need to pay for the particular ride.



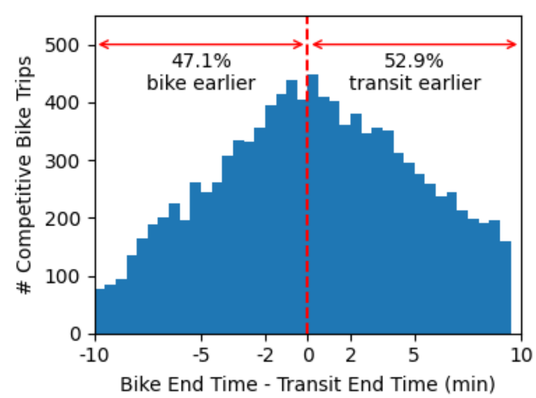
(a) Duration Difference



(b) Competitive Bike Trip Durations Compare with Transit Trip Durations



(c) Starting Time Difference



(d) Ending Time Difference

Figure 4.13: Compare Competitive Bike Trips with Corresponding Transit Trips

4.3 Study 2: Complementary Trips

4.3.1 Parameters Determination

$maxBikeDist$, $maxBikeDuration$, $avgBikeSpeed$, $maxCloseTime$, $maxWalkDist$ and $maxWalkSpeed$ are needed to detect the potential first/last-mile bike trips. $maxBikeDist$ represents the maximum distance of a bike sharing trip to be a candidate for a first/last-mile trip. In this study, I use the simple definition of first/last-mile trip and set the $maxBikeDist$ to one mile (1,600 meters). $maxBikeDuration$ and $avgBikeSpeed$ identify the first/last-mile trip candidates from temporal perspective. Figure 4.14 shows the Duration-Distance scatter of all bike trips within 20 minutes' duration. As Figure 4.14 shows, the trip duration of a certain distance varies from several minutes to tens of minutes. Therefore, in addition to the $maxBikeDist$ (orange solid line), I set $maxBikeDuration$ (red dash line) to 5 minutes to identify the short duration bike trips. However, a 5-minute duration is quite a hurry for relatively long distance (i.e., distance about 1,500 meters), so $avgBikeSpeed$ (red dot line) is used to include the trips longer than 5 minutes but still relatively fast. These trips are counted as still following the feature of first/last-mile trips, i.e., going directly from origin to destination. The $avgBikeSpeed$ is set to 3.33 m/s, which is commonly used for bike sharing trip mileage estimation ("Citi Bike System Data", n.d.), and also conforms to the density distribution of the trips' speed (see Figure 4.15).

In order to keep the consistency of the study, $maxCloseDist$ is still set to 100 meters based on the discussion in 4.2.1. $maxWalkDist$ is set to 400 meters, which is commonly used by transit planners as the distance people will walk to reach a bus

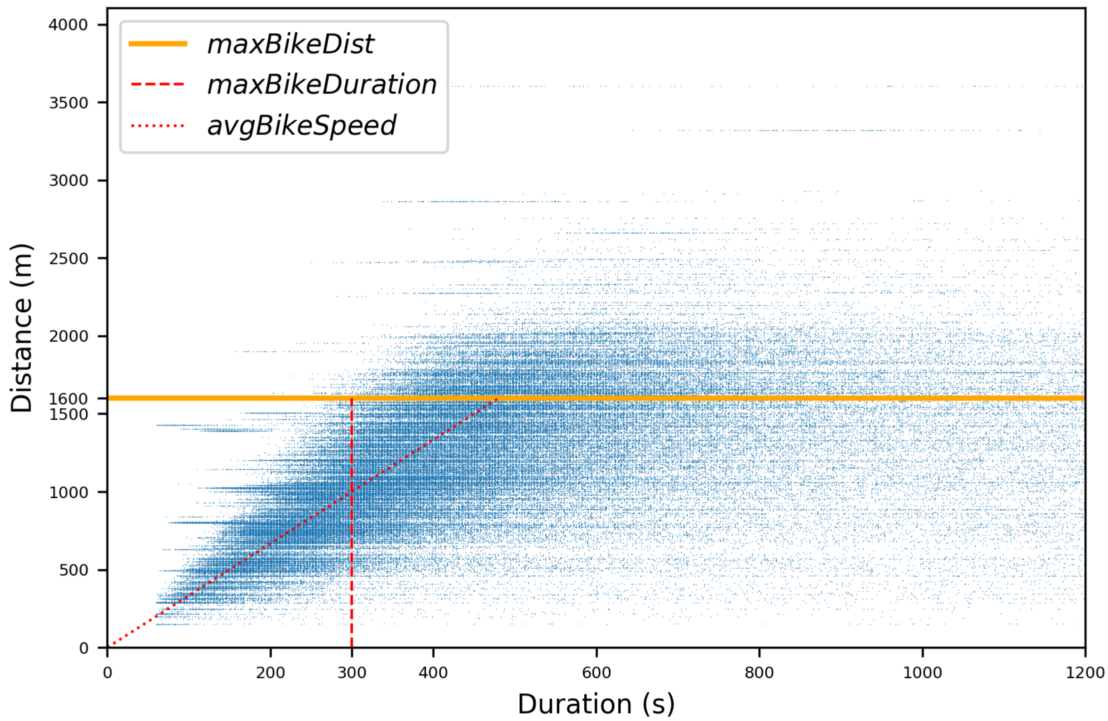
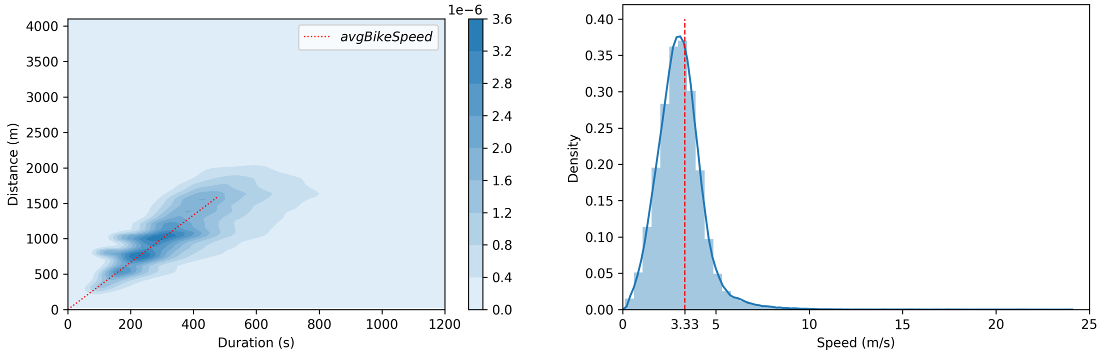


Figure 4.14: Duration-Distance Scatter of Bike Sharing Trips

stop (Walker, 2012), as well as 5 minutes' walk under the speed of 1.4 m/s, i.e., the human preferred walking speed (Browning et al., 2006). $maxWalkSpeed$ is set to 2.5 m/s, which is the common walking speed capacity (Minetti, 2000).



(a) 2D KDE of Duration-Distance Scatter (b) KDE of Speed

Figure 4.15: Bike Trip Speed Density Distribution

4.3.2 Complementary Trips Result and Discussion

When $maxBikeDist$ is set to 1,600 meters, $maxWalkDist$ is set to 400 meters and other parameters set as shown in Table 4.3, 253,768 pairs of bike sharing trips and APC records are detected as potential first-mile complementary cases, where 58692 (12.5% of all) bike trips are involved; 264,741 pairs are detected as potential last-mile complementary cases, where 59125 (12.8% of all) bike trips are involved.

Parameter	**Value**
$maxBikeDist$	1,600 m
$maxBikeDuration$	5 min
$avgBikeSpeed$	3.33 m/s
$maxCloseTime$	10 min
$maxWalkDist$	400 m
$maxWalkSpeed$	2.5 m/s
Result	
# of First-mile relationship cases	253,768
# of First-mile bike trips	58,692 (12.5%)
# of casual bike trips	3,718 (6.3%)
# of member bike trips	54,974 (93.7%)
# of Last-mile relationship cases	264,741
# of Last-mile bike trips	59,125 (12.8%)
# of casual bike trips	3,716 (6.3%)
# of member bike trips	55,409 (93.7%)

Table 4.3: Complementary Relationship Detection Parameters and Result

The detection rules for complementary relationship illustrated in 3.2.1 are the essential conditions of a first/last-mile trip, which means the potential first/last-mile trips are a superset of actual first/last-mile trips. According to the detected trips, it is fair to make the following deductions about first/last-mile bike sharing trips:

1. Bike sharing members are more likely to make first/last-mile bike trips. The ratio of member trips in potential first/last-mile are both 93.7%, higher than the ratio in competitive trips (82.4%) and in all bike trips (63.0%).

2. Spatially speaking, first/last-mile bike trips are more likely to happen in the University campus. Figure 4.16 to 4.19 show the starting and ending counting for potential first and last-mile bike trips at each station. The bike stations in the University are the top starting and ending stations for potential first/last-mile trips. Downtown Minneapolis is the second possible place for first/last-mile bike trips.

3. Temporally speaking, first/last-mile bike trips are more likely to happen on weekdays, at one of the peak hours (morning, noon, afternoon). Figure 4.20 shows the day-of-week distribution of potential first/last-mile bike trips. Nearly 90% of them happen on weekdays. Figure 4.21 shows the time-of-day distribution of potential first/last-mile bike trips. There are three significant peaks in the hour 8:00-9:00, 12:00-13:00 and 16:00-17:00, comparing with the all trips.

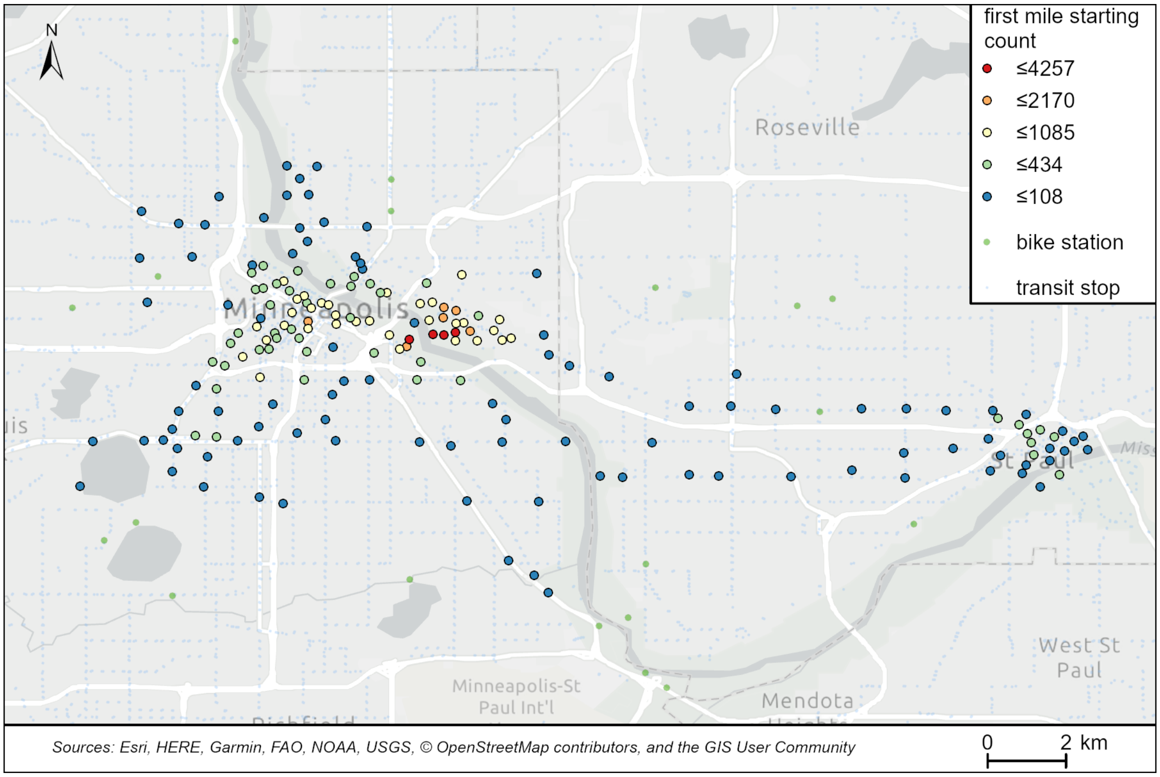


Figure 4.16: First-mile Trip Starting Counting

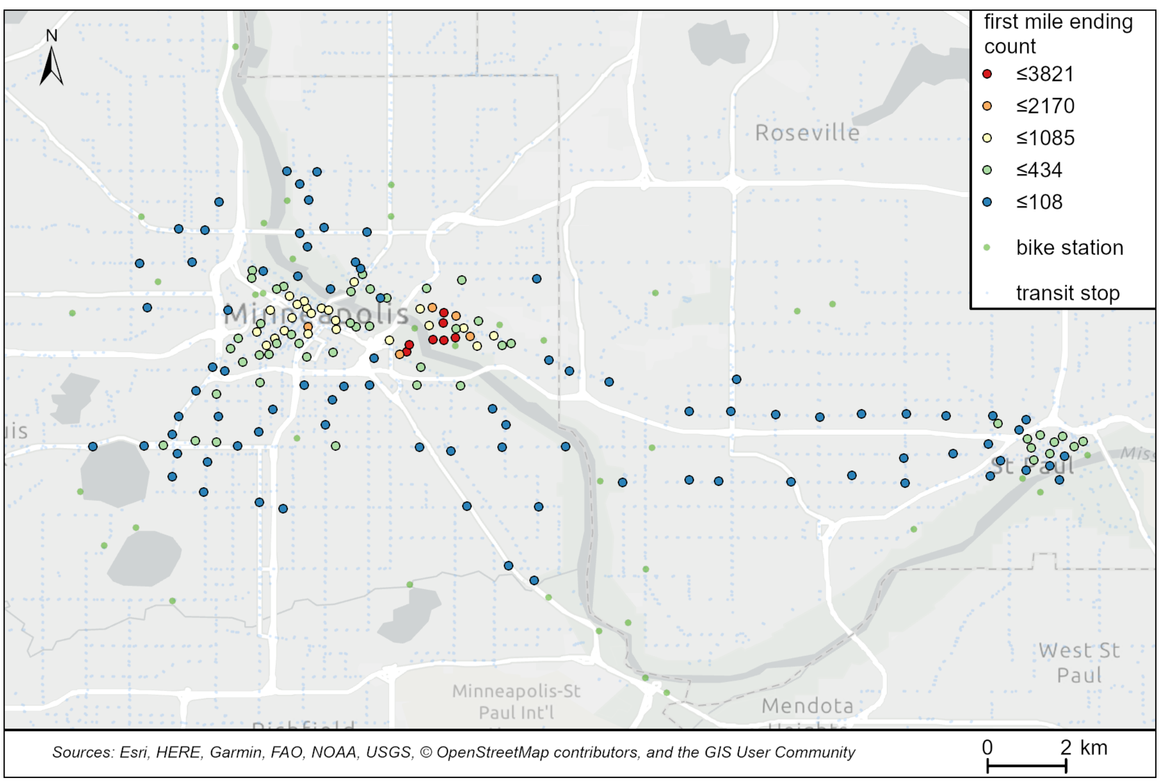


Figure 4.17: First-mile Trip Ending Counting

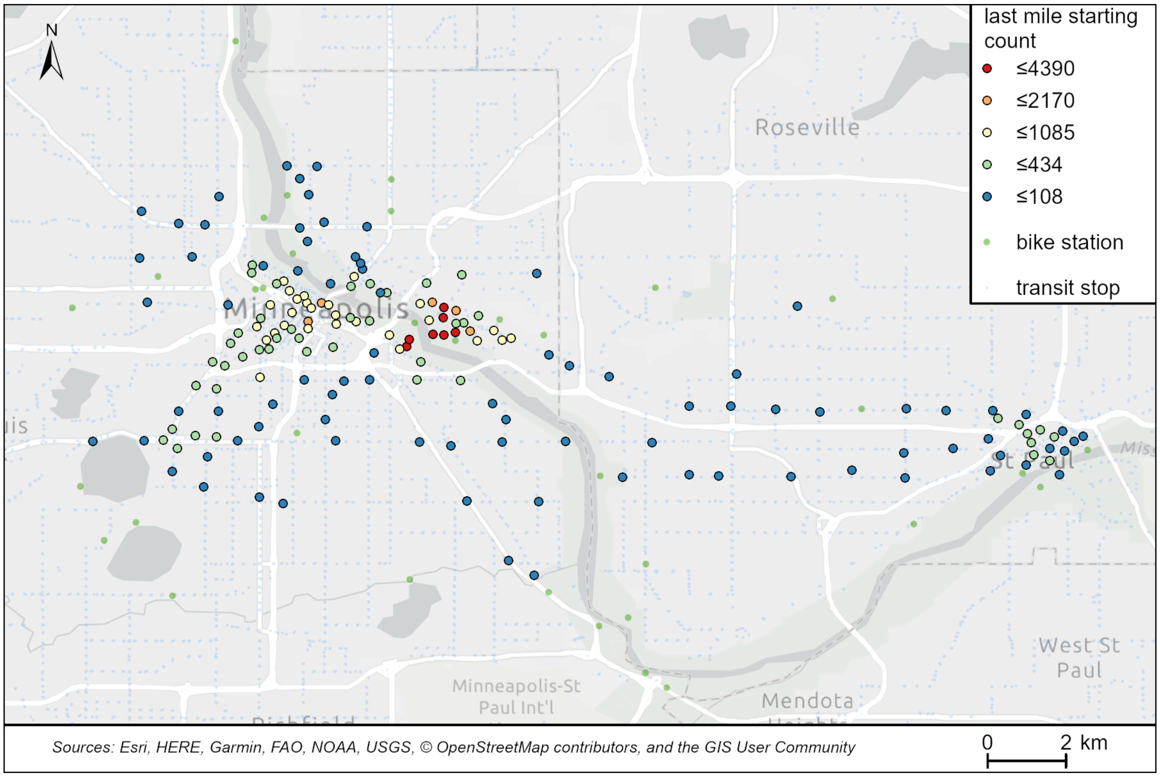


Figure 4.18: Last-mile Trip Starting Counting

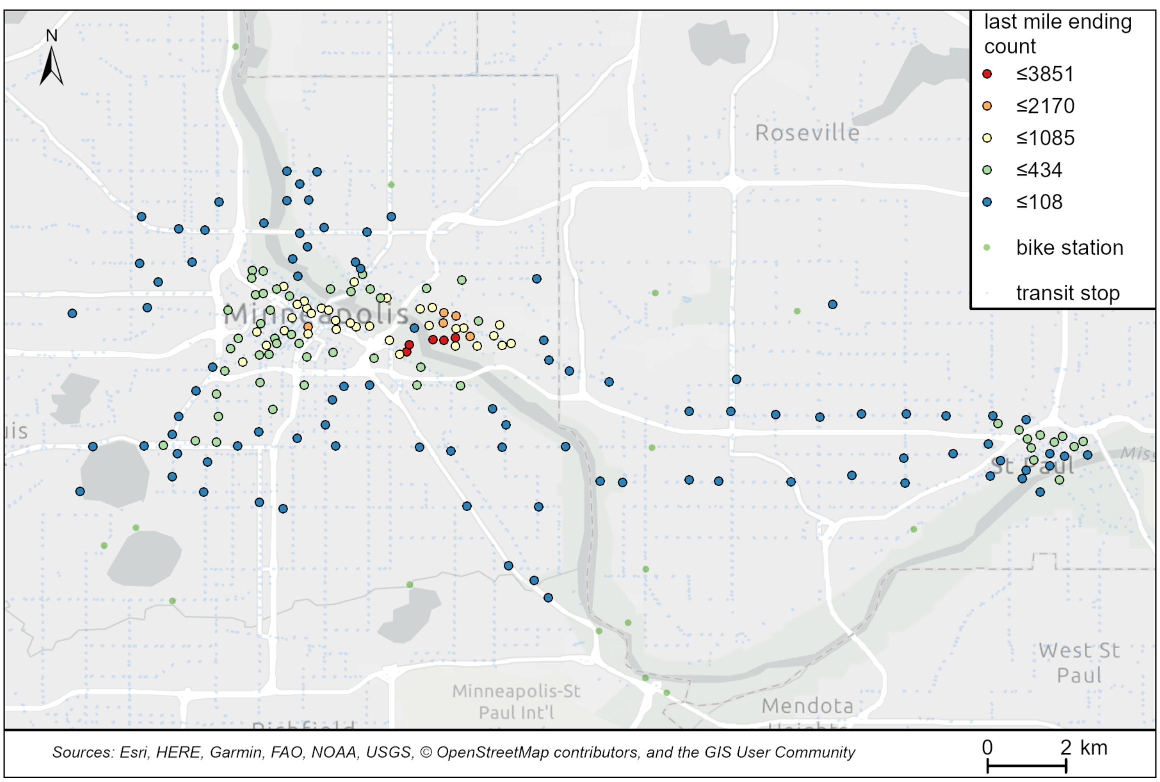
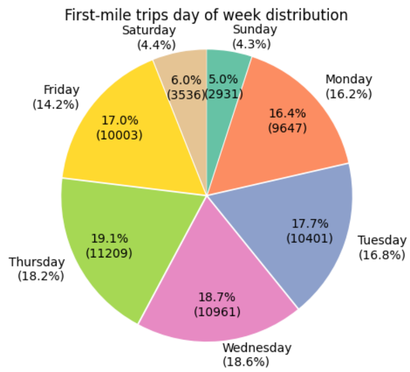
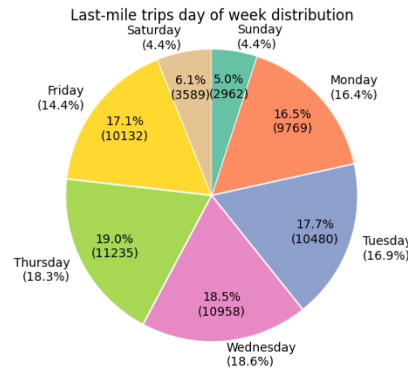


Figure 4.19: Last-mile Trip Ending Counting



(a) First-mile trips



(b) Last-mile trips

Figure 4.20: First/Last-mile Trip Day of Week Distribution

In Figures 4.20a and 4.20b, the percentage within each pie means the percentage of daily complementary trips of all complementary trips; the percentage outside of each pie means the percentage of complementary trips of all trips on each day. For example, 16.4% of potential first-mile bike trips are on Monday, which are 16.2% of all Monday trips.

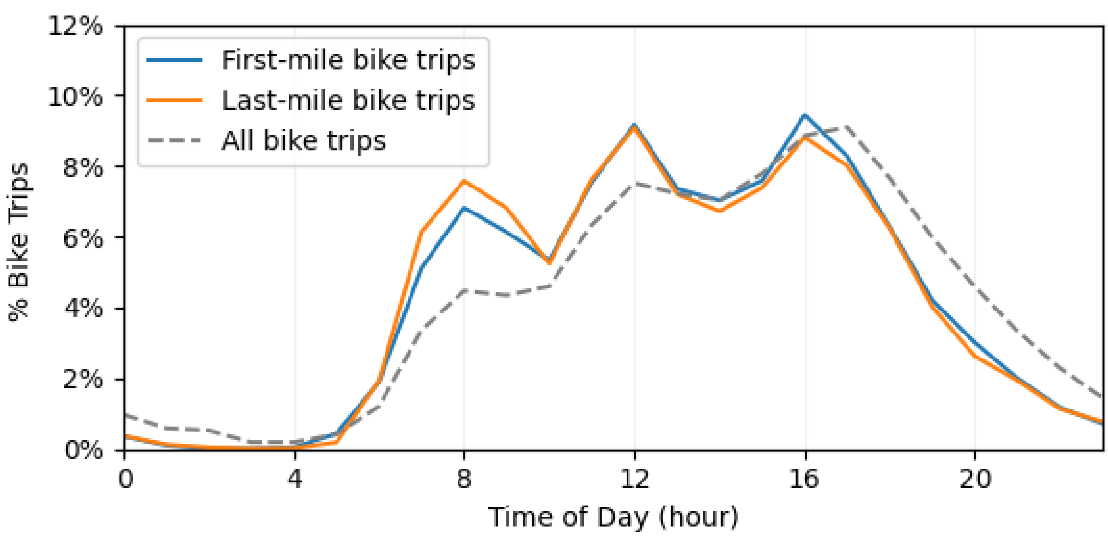


Figure 4.21: First/Last-mile Trip Time of Day Distribution

4.4 Study 3: Ridership Relationship

4.4.1 Related Graph

Graph-building methods proposed in 3.3.1 are applied to find potentially related bike sharing stations and transit stops. Figures 4.22 to 4.25 show the graphs built by the three kinds of methods and their variations. In these figures, lines represent edges between bike stations and transit stops, and polygons represent the convex hull of each connected component of the stations/stops and edges.

	Relationship-based	Buffer-based	Buffer-based*	KNN-based
# of subgraphs	75	75	131	139
# of bike stations	176	179	170	179
# of transit stops	1,090	1,289	1,050	614
# of edges	1,885	2,160	1,148	692

Table 4.4: Graph-building Result

According to the results, both the relationship-based graph and buffer-based graph have a large subgraph that covers the entire Downtown Minneapolis area. This is mainly due to the high density of both bike stations and transit stops in that area. Such large subgraphs may depress the effectiveness and interpretability in the following analysis of ridership correlations between subgraphs. The spirit of ridership relationship study is to reveal potential transferring behaviors from the correlation between bike ending and transit boarding, and transit alighting and bike starting. Supposing the majority of biking ridership in Downtown Minneapolis is in the North Loop area (northwest Downtown) while the majority of transit ridership is in the Loring Park area (southeast Downtown). The ridership in these two areas will dominate the cross-topic correlations of this subgraph. However, the bike endings in North Loop and transit boardings in Loring Park cannot indicate any transferring even if they are significantly correlated. Therefore, methods creating extremely large subgraphs should be refined or abandoned.

An alternative variation to the buffer-based method is to break down the large subgraphs using a smaller $maxWalkDist$, namely buffer-based* method. Figure 4.24 is an example where subgraphs with more than three bike stations are broken down by re-build the subgraphs with smaller $maxWalkDist = 200m$. As the figure shows, large subgraphs in Minneapolis and St.Paul downtown areas, the University neighborhood and Uptown area are divided into smaller subgraphs while the other subgraphs remain the same as the original division.

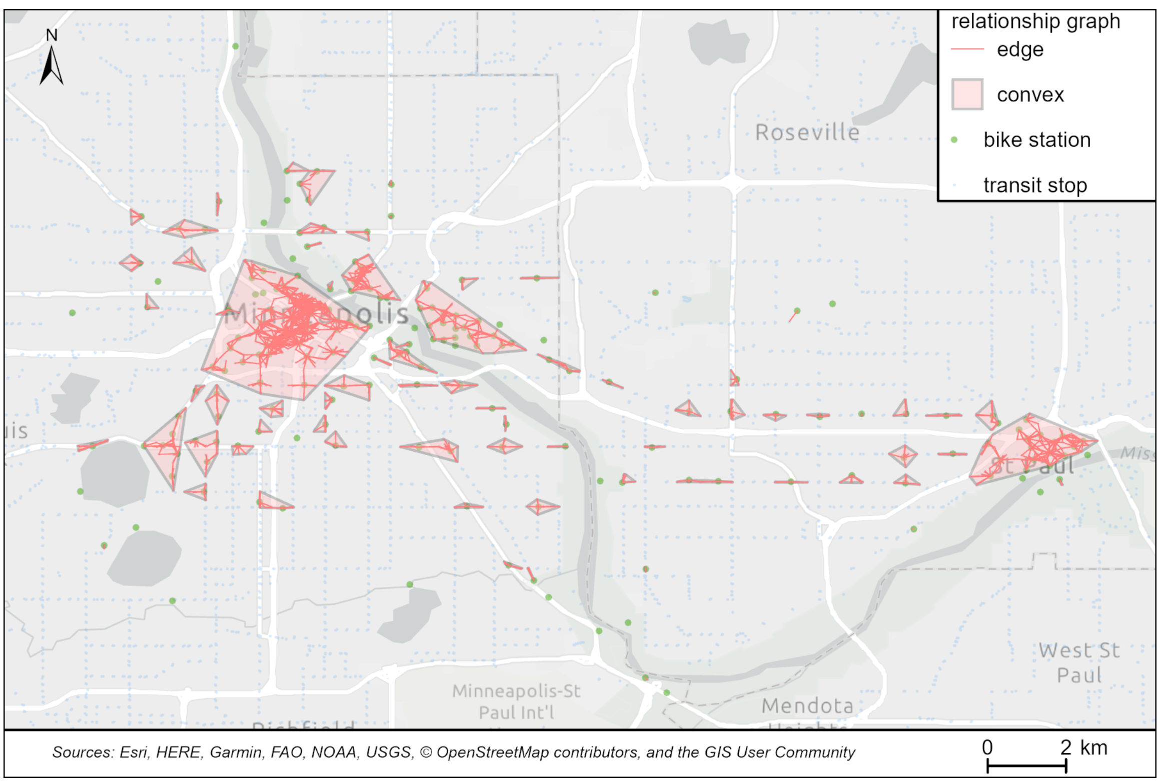


Figure 4.22: Relationship-based Graph

Several factors are considered and compared to evaluate the graphs built by dif-

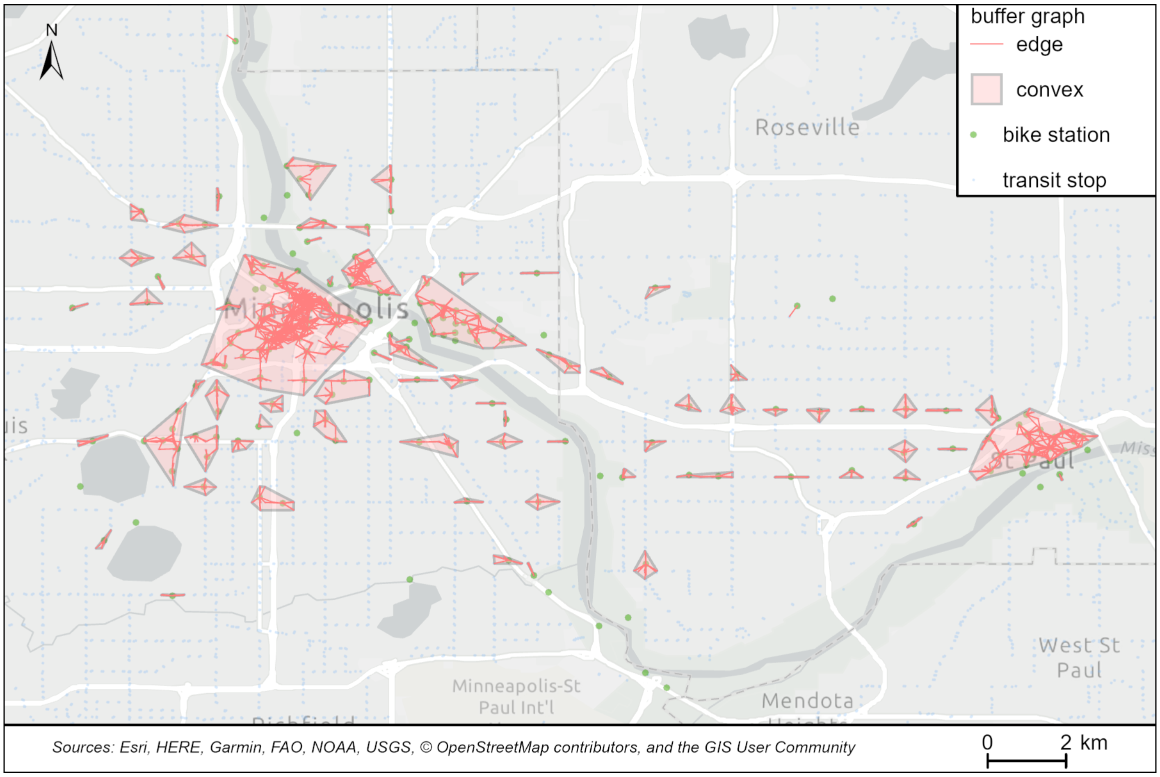


Figure 4.23: Buffer-based Graph, $maxWalkDist = 400m$

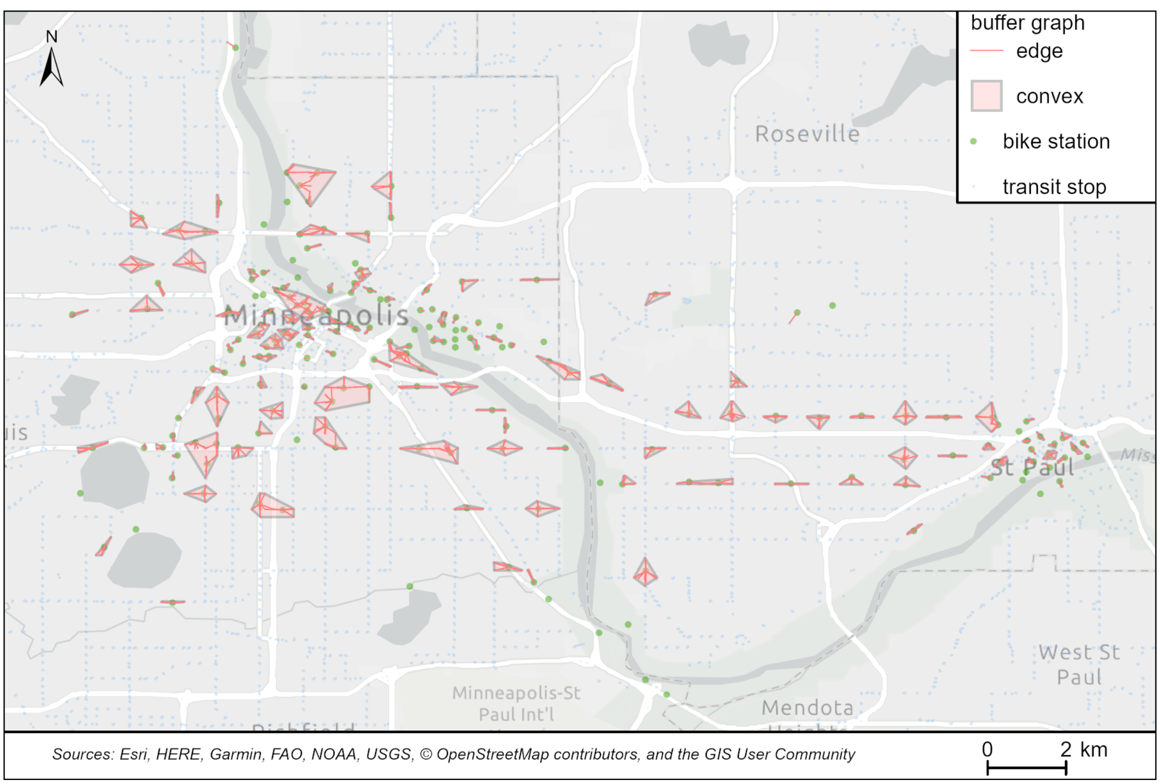


Figure 4.24: Buffer-based Graph with Decomposition
Connected components with more than 3 bike stations are re-built with $maxWalkDist = 200m$

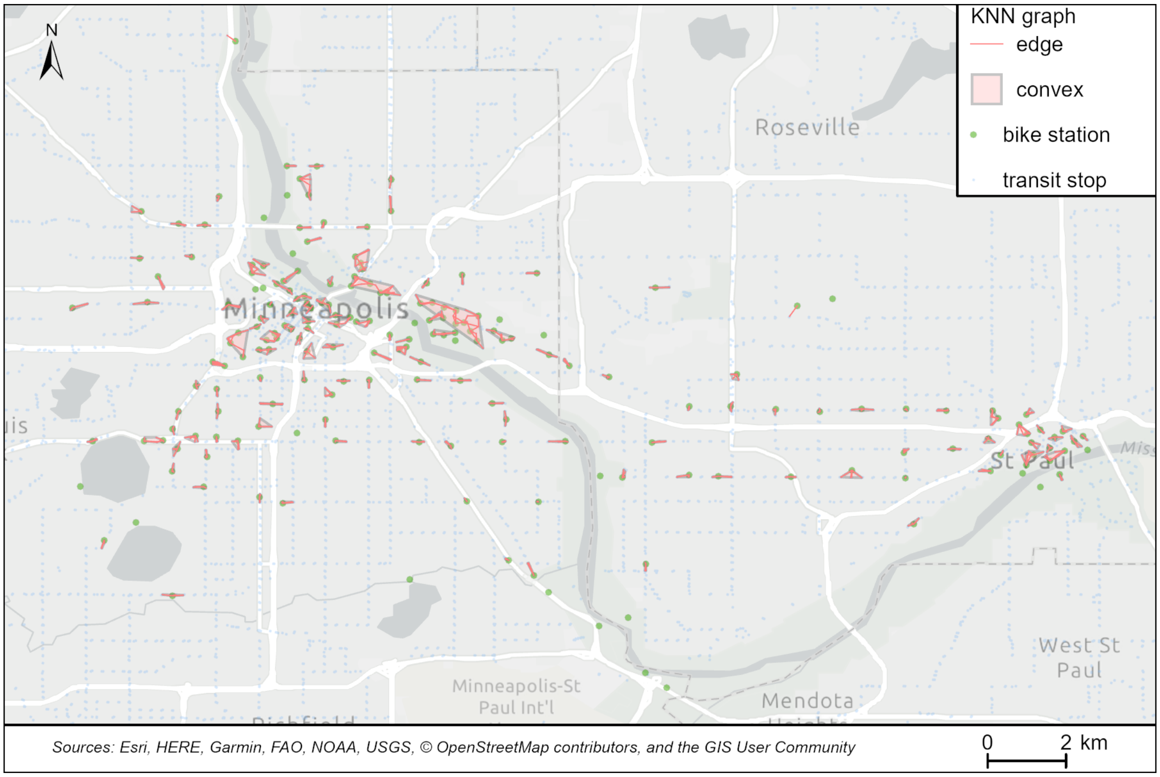


Figure 4.25: KNN-based Graph, K=4

ferent methods. Specifically, two metrics are defined to evaluate the correctness and compactness of the graphs. The correctness metrics take the detected competitive and complementary relationships in Chapter 4.2 and 4.3 as ground truth. The generated graph is more precise if it includes more edges of detected relationships and includes fewer edges with no corresponding relationship. Analytically, for edge set E_i of a generated graph G_i and edge set $E_{relation}$ derived from all relationships generated based on all detected complementary or competitive trips, the two correctness coefficients are:

$$CorrectCoef_1 = \frac{|E_{relation} - E_i|}{|E_{relation}|}$$

$$CorrectCoef_2 = \frac{|E_i - E_{relation}|}{|E_i|}$$

. $CorrectCoef_1$ measures how many detected relationships are missed in the graph, and $CorrectCoef_2$ measures how many edges have no corresponding relationship. The smaller the two coefficients are, the more precise the graph is.

The compactness metrics are based on the principle that integrating too many riderships of different patterns within each subgraph will compromise their original patterns. The compactness coefficient of bike-start ridership of a subgraph is defined as the averaged standard deviation of starting ridership across all bike stations in the subgraph. Therefore, the smaller the coefficient is, the more compact the ridership is within the subgraph. For a subgraph $subgraph_i(Stations_i, Stops_i, Edges_i)$, where $Stations_i$ is the set of bike stations and its cardinality is denoted as $|Stations_i|$,

$$start\_CompactCoef_i = \frac{\sum_{d\in\mathbb{D}}\sum_{h\in\mathbb{H}} start\_std_{(i,d,h)}}{|\mathbb{D}| * |\mathbb{H}|}$$

, where

$$start\_std_{(i,d,h)} = \sqrt{\frac{1}{|Stations_i| - 1}\sum_{s\in Stations_i}(start_{(s,d,h)} - sta\bar{r}t_{(i,d,h)})^2}$$

and

$$sta\bar{r}t_{(i,d,h)} = \frac{\sum_{s\in Stations_i} start_{(s,d,h)}}{|Stations_i|}$$

.

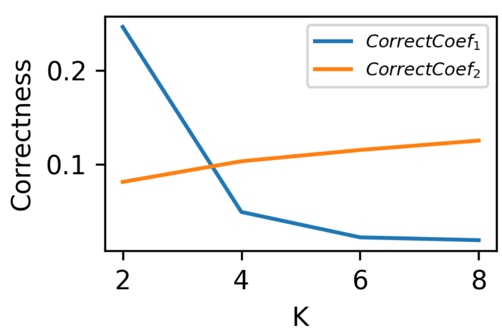
Table 4.5 compares the building process and evaluation of each method. (1) Relationship-based method needs no parameter; buffer-based method needs $maxWalkDist$ to do the range query; buffer-based* method inherits $maxWalkDist$ from buffer-based method, and it also needs a maximum station number to determine which subgraphs to be broken down and a smaller $maxWalkDist$ to do the re-build; KNN-based method needs K to run KNN query and $maxWalkDist$ to prune the long edges. (2) Among the methods and variations, relationship-based and buffer-based methods create super large subgraphs. (3) Transit stops of the same route's different directions are better to be included within the same subgraph, so the morning and evening peaks of the same group of passengers can be investigated together. Only KNN-based method can guarantee to include both directions of transit routes to the maximum extent. (4)Relationship-based method builds the most precise graph as it is purely based on relationships. (5) Buffer-based* method builds the most ridership-compact graph as it breaks down all the large subgraphs and KNN-based method builds the second compact graph as K limits the number of transit stops a bike station could link with.

	Relationship-based	Buffer-based	Buffer-based*	KNN-based
Parameters needed	0	1	3	2
Create large subgraphs	Yes	Yes	No	No
Guarantee both directions	No	No	No	Almost
$CorrectCoef_1$/ $CorrectCoef_2$	0.000/0.000	0.007/0.128	0.007/0.173	0.049/0.103
$start\_CompactCoef$/ $end\_CompactCoef$	0.107/0.108	0.113/0.114	0.099/0.099	0.103/0.101

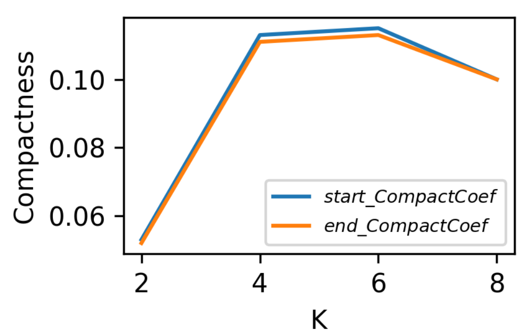
Table 4.5: Graph-building Methods Comparison

KNN-based graph is used for the following correlation study for the reasons that: (1) it requires only two parameters, which means less domain knowledge is required; (2) it does not create super large subgraphs; (3) it guarantees to include transit routes of both directions to the most extend; (4) it is relatively precise as it does not include many edges with no corresponding relationship and (5) it has relatively compact biking ridership which means the bike sharing customers of each subgraph are of similar characteristics.

K needs to be even to guarantee to include transit routes of both directions. The correctness and compactness coefficients are used to choose the appropriate K. Figure 4.26 show how the coefficients change as K is set to 2, 4, 6, and 8. More edges, with or without corresponding relationship, are included as K growing. As shown in Figure 4.26a, $CorrectCoef_1$ is going lower while $CorrectCoef_2$ is going higher. As Figure 4.26b shows, the subgraphs are compactest when K is small. The compactness coefficients are also relatively small when K is big because the subgraphs become large now and the effect of outliers is mitigated. Based on the discussion above, K is set to 4, so in most cases, one bike sharing station is associated with four nearest transit stops, which means two sets of transit routes of both directions.



(a) Correctness



(b) Compactness

Figure 4.26: Correctness and Compactness Sensitivity of K

4.4.2 Significant Correlations

For each subgraph derived from KNN-based method, Pearson Correlation Coefficients are calculated across bike sharing and transit ridership topics.

The range of Pearson Correlation Coefficient, i.e., $[-1, 1]$, is divided into 5 equal-length intervals $[-1, -0.6], (-0.6, -0.2], (-0.2, 0.2), [0.2, 0.6), [0.6, 1]$, representing "significant negative correlation", "moderate negative correlation", "weak negative/positive correlation", "moderate positive correlation", "significant positive correlation", respectively. Then the entire correlation between two ridership topics is determined by the interval with the highest amount. For example, for the Pearson Correlation Coefficients distribution shown in Figure 4.28b, coefficient intervals with highest amount are $[0.6, 1]$, $(-0.2, 0.2)$, $[0.2, 0.6)$, $[0.2, 0.6)$ for the four categories respectively, so the correlation levels are "significant positive correlation", "weak negative/positive correlation", "moderate positive correlation" and "moderate positive correlation". I use the quantile method here to convert continuous values to interval-based measures. It is also possible to use other categorization method as long as it is independent from input data and the empirical values.

Table 4.6 shows the number of subgraphs with non-significant and significant correlations between bike sharing ridership and transit ridership. Since no correlation shows as significantly negative, so in the following, significant correlation means a significant positive correlation. Most of the subgraphs show non-significant correlation between bike and transit ridership, which means that in most areas, competition and complementary relationships are not common enough to form any correlation pattern.

The most common significant correlation is the ones between bike-start and transit-alight, and the second common is the ones between bike-end and transit- alight. This indicates that in most cases of significant correlation, the bike sharing and transit trips are of the same direction, either both leaving the area (start-board) or arriving at the area (end-alight). This suggests that people in these areas usually use these two kinds of transport for the same purpose. However, this kind of concurrent in-flow and out-flow is not enough to suggest any competitive relationship, which is defined with both trip origin and destination.

Non-significant	start-board end-alight	start-board	start-board end-board	start-alight end-alight
124	8	5	1	1

Table 4.6: Significant Correlated Subgraph Counting

Figure 4.27 shows the non-significant (grey) and significant (color) correlated subgraphs on the map. According to the figures, the significant correlations located in Downtown Minneapolis, the University, Uptown lakefront and Marcy-Holmes community, which corroborates with the hotspots of detected competitive/complementary relationships. In the following part, each type of significant correlation will be explored and discussed.

Significant Start-Board Correlation Figure 4.28 is an example of significant start-board correlation with its ridership temporal profiles and all cross-topic Pearson correlation coefficients. According to the temporal profiles and correlation coefficients, the bike-start and transit-board riderships both have a single peak in the afternoon and significantly correlated. The bike-end and transit-alight riderships also both have

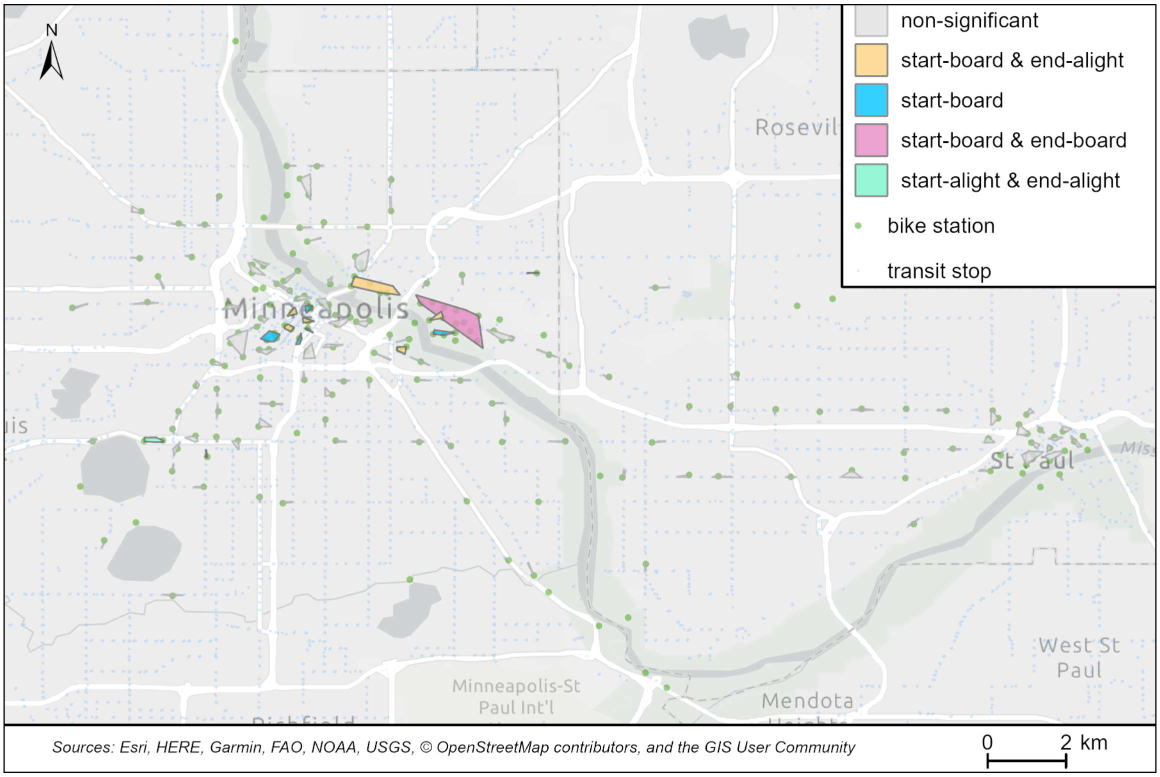


Figure 4.27: Significant Correlated Subgraphs

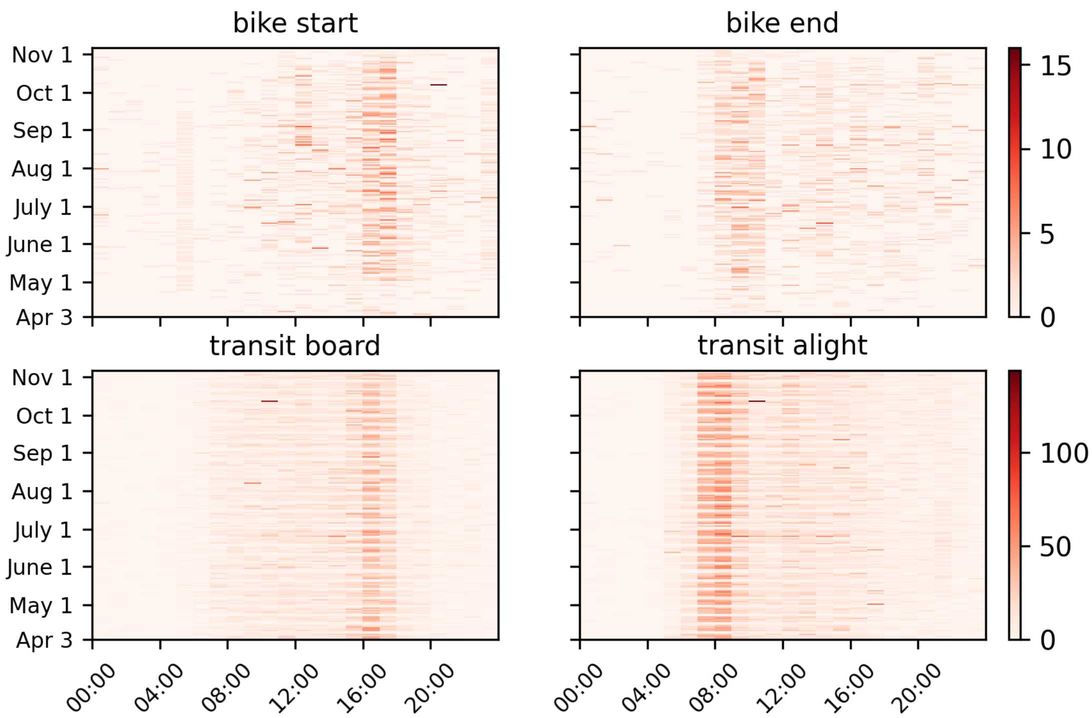
a single peak in the morning. However, their correlation is not significant. The reason is that the bike-end peak is about two hours later than the transit-alight peak and this lowers the correlation. This example is typical of other significantly start-board correlated subgraphs at the outer Downtown Minneapolis.

Based on the "morning-arriving-afternoon-leaving" pattern, there is a reasonable interpretation that the ridership may be mainly conducted by people who work in the Downtown area. The downtown-working people usually get off work and leave the area at about the same period of time, so there is a significant correlation in the afternoon peak. However, the outer Downtown includes not only business area but also universities and colleges (e.g., University of St. Thomas, Minneapolis College). Students usually have more flexible arriving time in the morning than office workers, depending on the time of their first classes. That is probably why the morning peak of bike-end shifts from the peak of transit-alight.

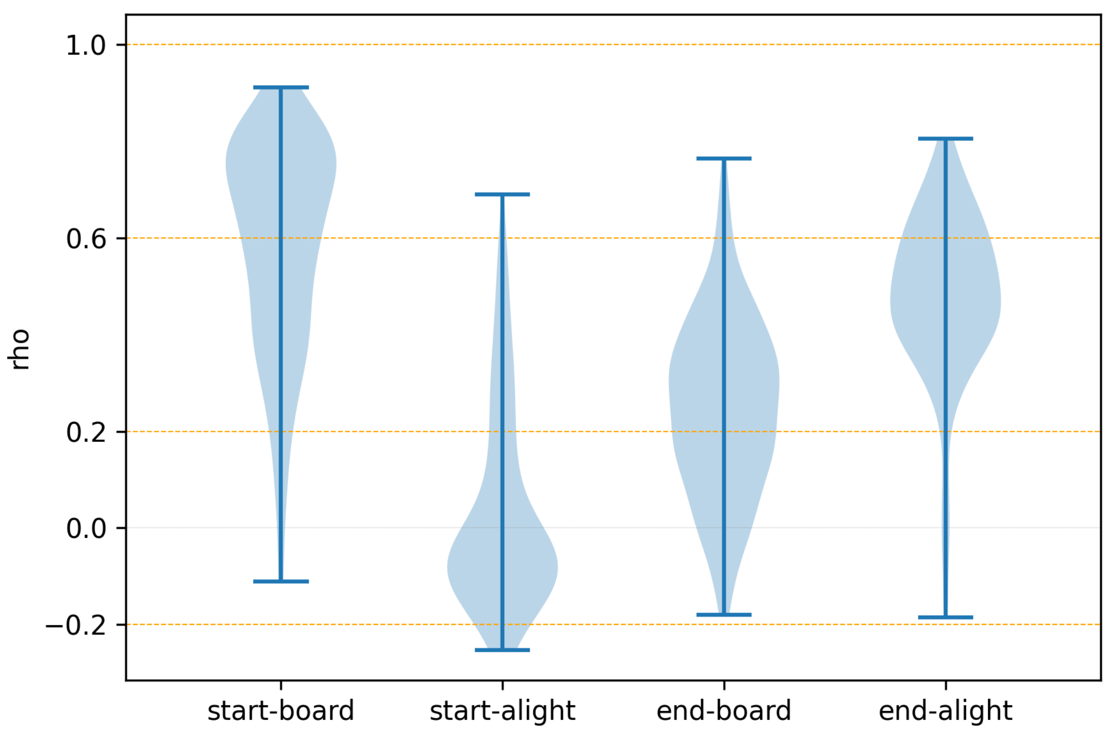
Significant Start-Board & End-Alight Correlation Figure 4.29 is an example of significant start-board and end-alight correlation. According to the temporal profiles and correlation coefficients, the bike-start and transit-board riderships both have afternoon peak; bike-end and transit-alight riderships both have morning peak, and they are nearly simultaneously. This kind of significant correlation mainly locates at central Downtown, where people are of the same group and they usually arrive and leave the area at the nearly same period of time.

Significant Start-Board & End-Board Correlation Figure 4.30 is an example of significant start-board and end-board correlation.This subgraph locates along the University Avenue and overlaps with both the University and Dinkytown Community. As the temporal profiles and correlation coefficients, the bike-start and bike-end ridership correlated with the transit-board transit, especially in the semester days. One of the reasons that the bike-start and bike-end ridership share a similar pattern is that this is the largest subgraph in the graph and more than 25% involved bike trips are actually within the subgraph. Therefore, the "bike-end-transit-board" correlation is not strong enough to suggest any clue for first-mile relationship.

Significant Start-Alight & End-Alight Correlation Figure 4.31 is an example of significant start-alight and end-alight correlation. According to the temporal profiles and correlation coefficients, both bike-start and bike-end riderships concentrate

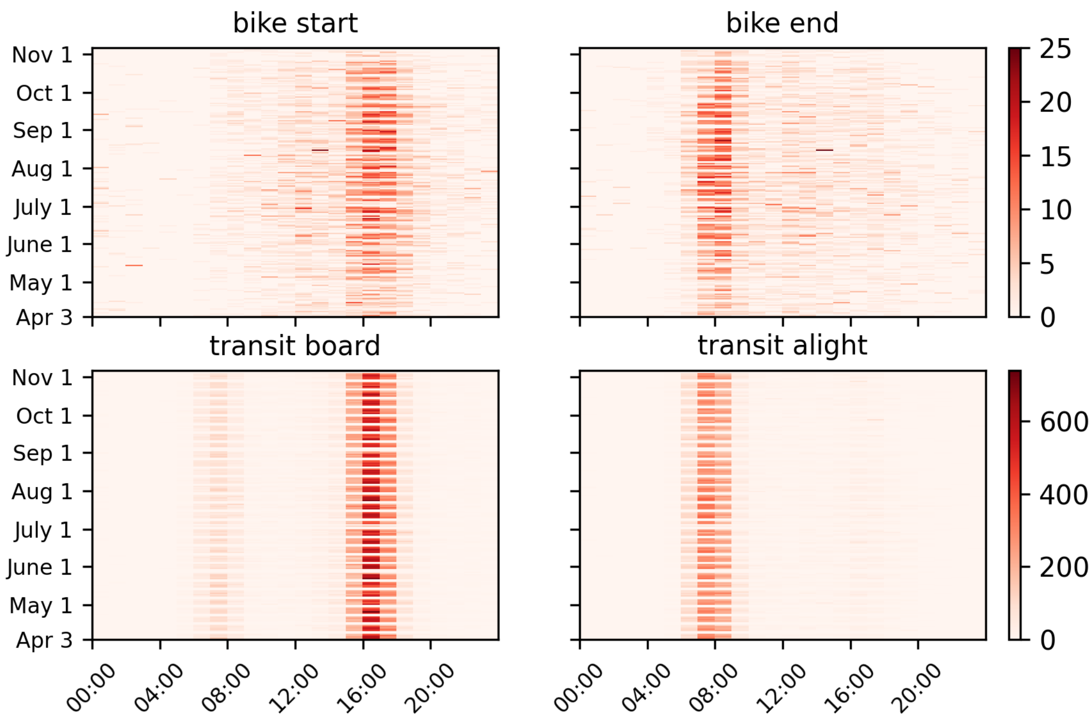


(a) Ridership Temporal Profiles

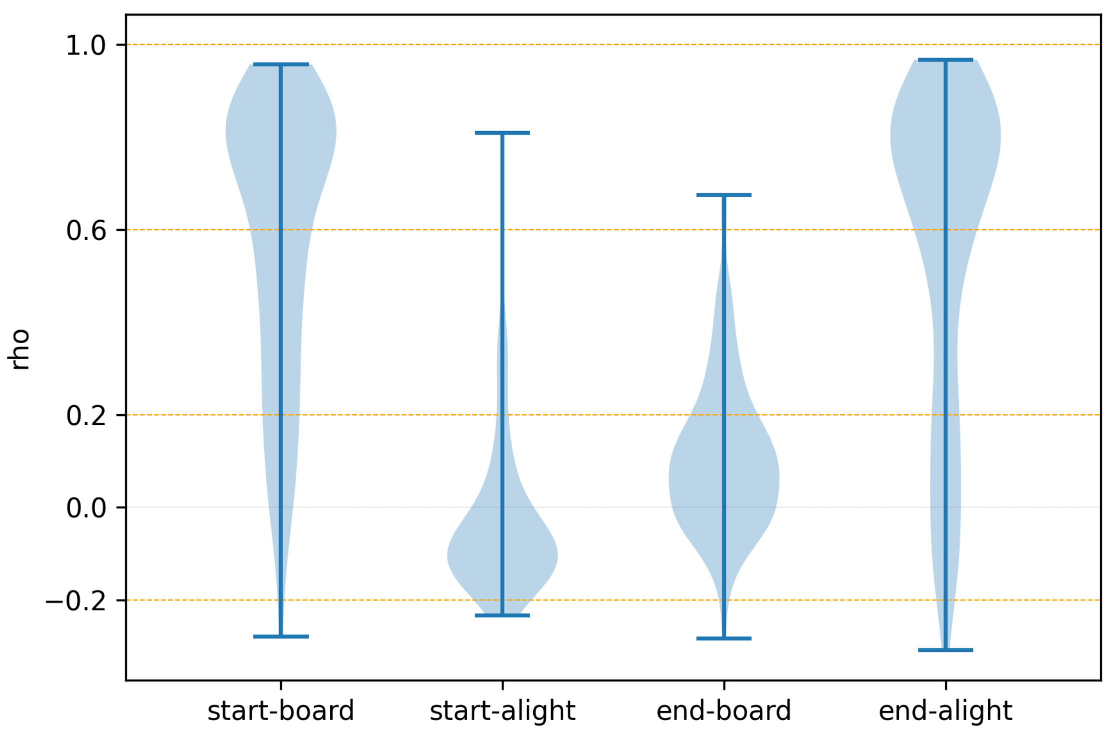


(b) Cross-topic Pearson Correlation Coefficient

Figure 4.28: Example of Significant Start-Board Correlation

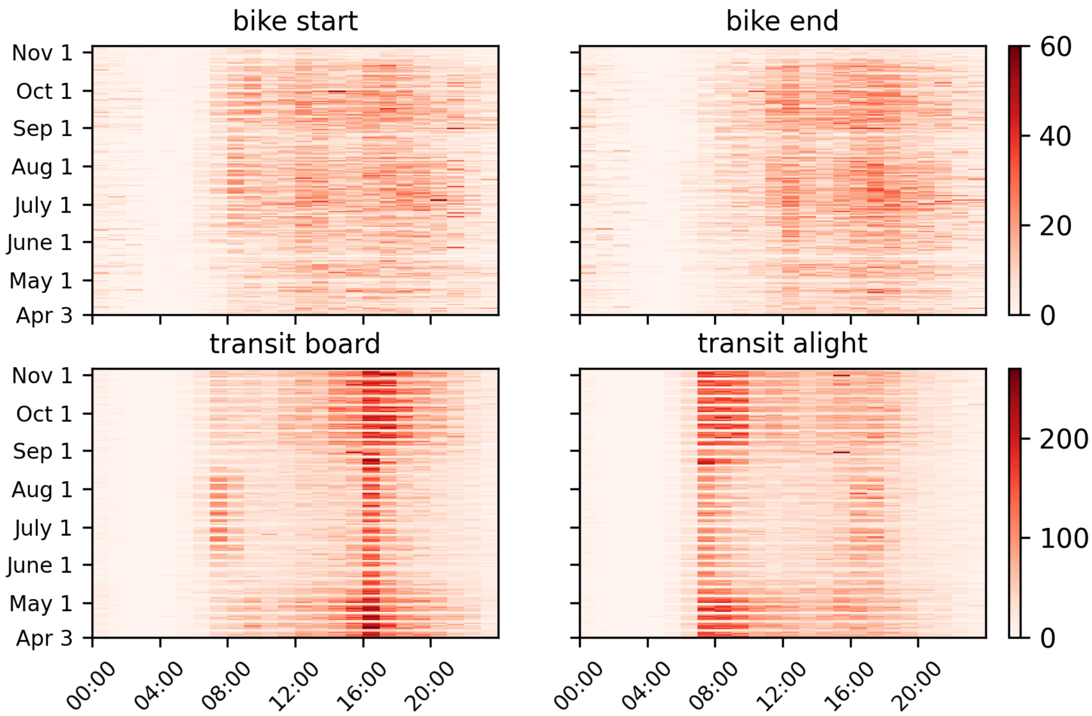


(a) Ridership Temporal Profiles

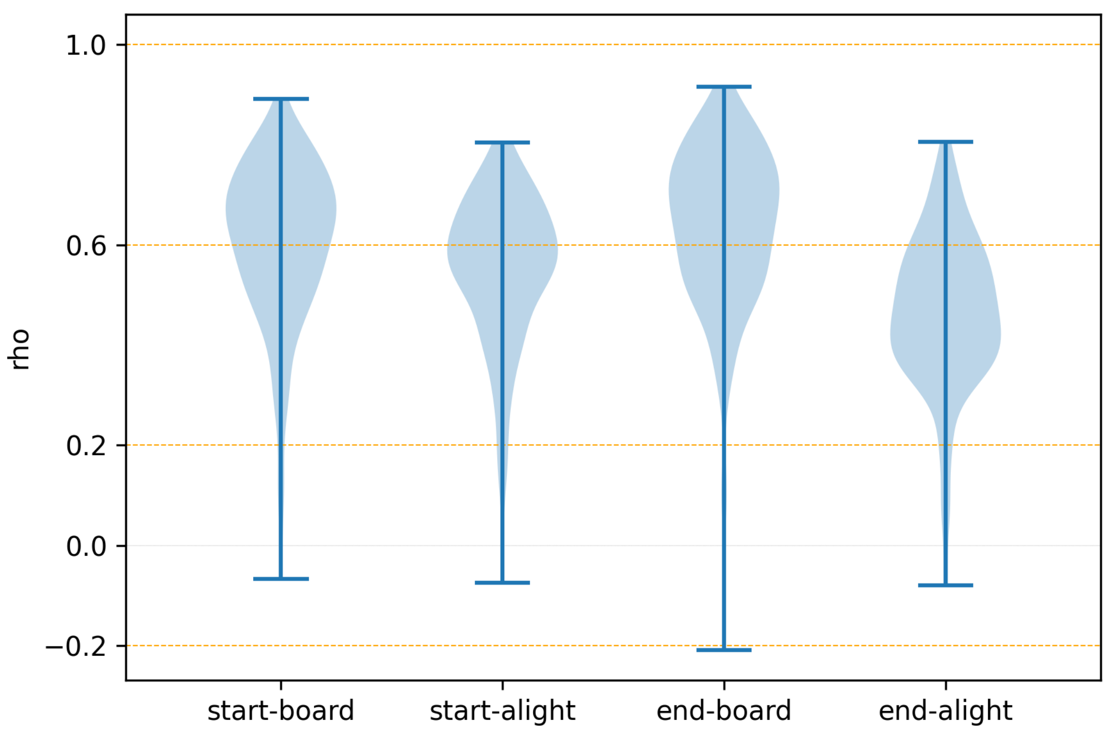


(b) Cross-topic Pearson Correlation Coefficient

Figure 4.29: Example of Significant Start-Board & End-Alight Correlation



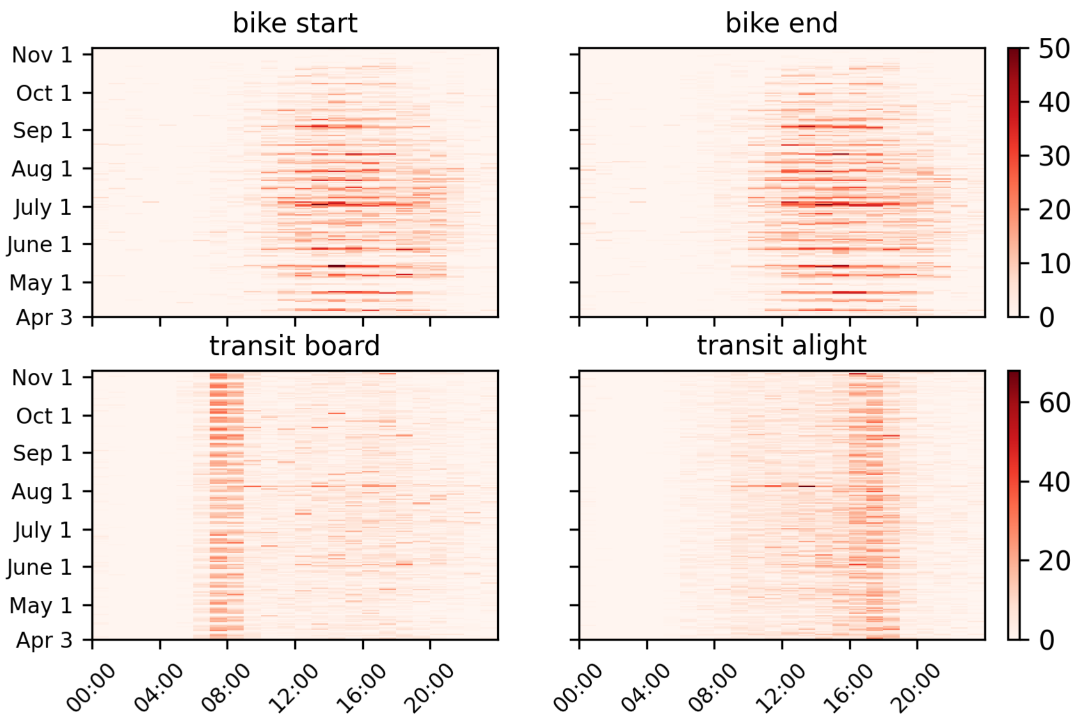
(a) Ridership Temporal Profiles



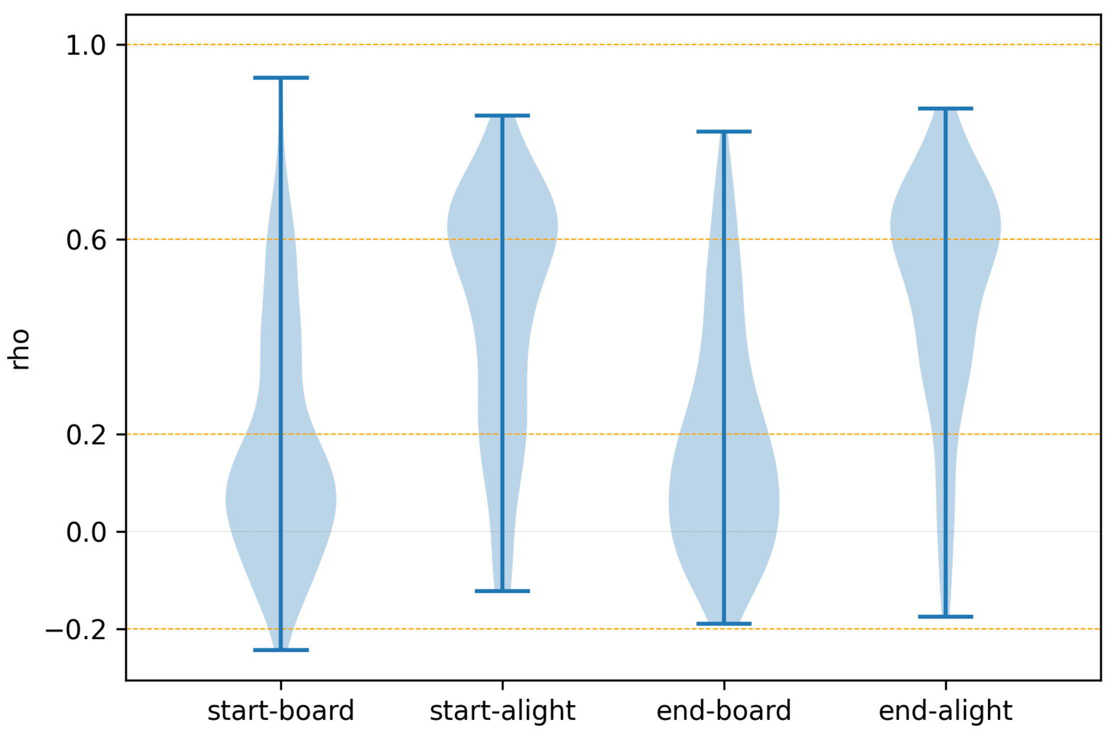
(b) Cross-topic Pearson Correlation Coefficient

Figure 4.30: Example of Significant Start-Board & End-Board Correlation

in the afternoon, which correlates with the afternoon peak of transit alight. This subgraph is located by the Lake Bde Maka Ska, where also several apartments and living houses lie. Based on the "morning-leaving-afternoon-arriving" pattern, the transit riderships are most likely conducted by people who live here. However, the bike trips mostly happen in the afternoon, which are most likely taken by people for recreation. Therefore, generally speaking, the bike riderships and transit riderships in this example are mainly conducted by two different groups of people, which suggests less relationship between bike sharing and public transit. However, the "transit-alight-bike-start" correlation and especially the slight increment in summer in both of them may indicate connections to some extent. It is possible that in the summer time, people take the transit to the lakefront and then rent a bike for recreation.



(a) Ridership Temporal Profiles



(b) Cross-topic Pearson Correlation Coefficient

Figure 4.31: Example of Significant Start-Alight & End-Alight Correlation

Chapter 5

Conclusion

My thesis proposes a framework with procedures and methods for investigating the complex relationships between bike sharing and public transit systems. This framework first defines and detects competitive and complementary relationships from spatio-temporal perspectives. After getting the potentially related bike stations and transit stops, the framework examines the correlations of bike sharing and transit ridership to verify and corroborate the detected relationship. I use the Minneapolis-St. Paul Area as a case study and apply the framework to the Nice Ride bike sharing and Metro Transit systems.

5.1 Case Study Summary

5.1.1 Competitive Relationship

In the most strict parameter setting, competitive trips make up only a small proportion of all bike sharing trips. The amount of competitive trips is more sensitive to the spatial criterion as more pairs of bike station and transit stop are considered to be close to each other when the distance threshold is larger. The spatial distribution of competitive trips is different from the distribution of all bike trips: competitive trips are only concentrated in the University of Minnesota, Minneapolis campus and are less concentrated in the center Downtown Minneapolis area. Temporally speaking, competitive trips mainly happen on weekdays and concentrate in morning, noon and evening peak hours, while there are a good amount of bike trips that are not competitive happening during weekends. The proportion of short duration trips is higher in competitive trips than in all bike trips when the temporal criterion is tight.

5.1.2 Complementary Relationship

The study of complementary relationship detects potential first/last-mile bike sharing trips. From the potential ones, it is reasonable to infer that the first/last-mile trips are more likely to happen in the UMN Minneapolis campus and the center Downtown Minneapolis, during the morning, noon and evening peak hours on weekdays.

5.1.3 Ridership Correlation

The correlation between bike sharing and transit ridership does not show a significant competitive or complementary relationship in general, suggesting that these two systems tend to operate relatively independently from each other in the Twin Cities.

The significant ridership correlation mainly exists between bike starting ridership and transit boarding ridership as well as between bike ending and transit alighting.

Both pairs of ridership are referring to competitive relationship, which further corroborates that competitive relationship does exist. However, the correlations between both pairs of ridership are positive instead of negative, which indicates that although the competitive relationship may exist, it is not significant.

The significant correlation referring to complementary relationship is rare in this study, which means that more factual evidence is still needed to assert that bike sharing is playing a complementary role to the public transit system in this case. Speaking from a ridership perspective, the complementary relationship may exist, but it is not general.

5.2 Operational Insights and Strategies

These insights on how Nice Ride working with Metro Transit in Minneapolis-St. Paul area can also provide some strategies to better facilitate first/last-mile by bike sharing.

1. More memberships. Both potential first/last-mile trips have 90+% of membership trips. To encourage more people to take shared bikes for first/last-mile, the city and Nice Ride could use strategies such as a reduced joint-membership price, integrated payment card, and more service coverage to draw more members.

2. The factor underneath membership is the riding cost. For a casual Nice Ride user, a single 30-minute ride takes the same as a 2.5-hour bus or light-rail trip. If the city wants to encourage people to take first/last-mile trips by shared bikes, potential methods include reducing the bike sharing fee for transit passengers

and waiving the fee for low-income and other under-representative groups who rely more on the transit services.

5.3 Future Work

For future work, if Metro Transit and Nice Ride could try to use an integrated payment system, it will be valuable to validate the potential first/last-mile trips detected using methods in this thesis with the smart card transaction data. The validation could further advance, even revise our understanding of the first/last-mile problem. It will also be valuable to apply this framework to other cities of different geographic and service features, which has both transit and bike sharing services, such as Chicago or New York City. Intuitively, some of the patterns may be similar, while others may be different across the cities. We can compare the different patterns with service characteristics such as system coverage and transit service frequencies and also the underlying geographic contexts such as neighborhood and socioeconomic features.

Bibliography

Adnan, M., Altaf, S., Bellemans, T., Yasar, A. u. H., & Shakshuki, E. M. (2019). Last-mile travel and bicycle sharing system in small/medium sized cities: user's preferences investigation using hybrid choice model. *Journal of Ambient Intelligence and Humanized Computing*, *10*(12), 4721–4731. https://doi.org/10.1007/s12652-018-0849-5 (cit. on pp. 2, 6)

Anderson, M. L. (2014). Subways, strikes, and slowdowns: The impacts of public transit on traffic congestion. *American Economic Review*, *104*(9), 2763–2796. https://doi.org/10.1257/aer.104.9.2763 (cit. on p. 1)

Ashqar, H. I., Elhenawy, M., Almannaa, M. H., Ghanem, A., Rakha, H. A., & House, L. (2017). Modeling bike availability in a bike-sharing system using machine learning, In *5th ieee international conference on models and technologies for intelligent transportation systems, mt-its 2017 - proceedings*, Institute of Electrical; Electronics Engineers Inc. https://doi.org/10.1109/MTITS.2017.8005700. (Cit. on p. 12)

Banister, D. (2008). The sustainable mobility paradigm. *Transport Policy*, *15*(2), 73–80. https://doi.org/10.1016/j.tranpol.2007.10.005 (cit. on p. 1)

Besser, L. M., & Dannenberg, A. L. (2005). Walking to public transit: Steps to help meet physical activity recommendations. *American Journal of Preventive Medicine*, *29*(4), 273–280. https://doi.org/10.1016/j.amepre.2005.06.010 (cit. on p. 1)

Boarnet, M. G., Giuliano, G., Hou, Y., & Shin, E. J. (2017). First/last mile transit access as an equity planning issue. *Transportation Research Part A: Policy and Practice*, *103*, 296–310. https://doi.org/10.1016/j.tra.2017.06.011 (cit. on p. 1)

Browning, R. C., Baker, E. A., Herron, J. A., & Kram, R. (2006). Effects of obesity and sex on the energetic cost and preferred speed of walking. *Journal of Applied Physiology*, *100*(2), 390–398. https://doi.org/10.1152/japplphysiol.00767.2005 (cit. on p. 46)

Campbell, K. B., & Brakewood, C. (2017). Sharing riders: How bikesharing impacts bus ridership in New York City. *Transportation Research Part A: Policy and Practice*, *100*, 264–282. https://doi.org/10.1016/j.tra.2017.04.017 (cit. on pp. 3, 9)

Chen, Z., van Lierop, D., & Ettema, D. (2020). Dockless bike-sharing systems: what are the implications? *Transport Reviews*, *40*(3), 333–353. https://doi.org/10.1080/01441647.2019.1710306 (cit. on p. 2)

Chong, Z. J., Qin, B., Bandyopadhyay, T., Wongpiromsarn, T., Rankin, E. S., Ang, M. H., Frazzoli, E., Rus, D., Hsu, D., & Low, K. H. (2011). Autonomous personal vehicle for the first- and last-mile transportation services, In *Proceedings* *of the 2011 ieee 5th international conference on cybernetics and intelligent systems, cis 2011*. https://doi.org/10.1109/ICCIS.2011.6070337. (Cit. on p. 2)

Christian, K., Cho, S. H., Kho, S. Y., & Kim, D. K. (2019). Bayesian models with spatial autocorrelation for bike sharing ridership variability based on revealed preference GPS trajectory data, In *Iet intelligent transport systems*, Institution of Engineering; Technology. https://doi.org/10.1049/iet-its.2019.0159. (Cit. on p. 11)

Citi Bike System Data. (n.d.). https://www.citibikenyc.com/system-data. (Cit. on p. 45)

Component (graph theory). (n.d.). https://en.wikipedia.org/wiki/Component%7B%5C_%7D(graph%7B%5C_%7Dtheory). (Cit. on p. 25)

Cover, T. M., & Hart, P. E. (1967). Nearest Neighbor Pattern Classification. *IEEE Transactions on Information Theory*, *13*(1), 21–27. https://doi.org/10.1109/TIT.1967.1053964 (cit. on p. 25)

Etienne, C., & Latifa, O. (2014). Model-based count series clustering for bike sharing system usage mining: A case study with the vélib' system of Paris. *ACM Transactions on Intelligent Systems and Technology*, *5*(3), 1–21. https://doi.org/10.1145/2560188 (cit. on p. 11)

Fan, A., Chen, X., & Wan, T. (2019). How Have Travelers Changed Mode Choices for First/Last Mile Trips after the Introduction of Bicycle-Sharing Systems: An Empirical Study in Beijing, China. *Journal of Advanced Transportation*, *2019*. https://doi.org/10.1155/2019/5426080 (cit. on pp. 2, 6)

Feng, Y., & Wang, S. (2017). A forecast for bicycle rental demand based on random forests and multiple linear regression, In *Proceedings - 16th ieee/acis international conference on computer and information science, icis 2017*, Institute of Electrical; Electronics Engineers Inc. https://doi.org/10.1109/ICIS.2017.7959977. (Cit. on p. 11)

Firestine, T. (2016). *BTS Technical Report: Bike-Share Stations in the United States* (tech. rep.). Bureau of Transportation Statistics. https://www.bts.gov/archive/publications/bts%7B%5C_%7Dtechnical%7B%5C_%7Dreport/april%7B%5C_%7D2016. (Cit. on p. 2)

Fishman, E., Washington, S., & Haworth, N. (2013). Bike Share: A Synthesis of the Literature. *Transport Reviews*, *33*(2), 148–165. https://doi.org/10.1080/01441647.2013.775612 (cit. on p. 3)

Graehler, M., Mucci, R. A., & Erhardt, G. D. (2019). Understanding the Recent Transit Ridership Decline in Major US Cities: Service Cuts or Emerging Modes?, In *98th annual meeting of the transportation research board*. (Cit. on p. 3).

Griffin, G. P., & Sener, I. N. (2016). Planning for bike share connectivity to rail transit. *Journal of Public Transportation*, *19*(2), 1–22. https://doi.org/10.5038/2375-0901.19.2.1 (cit. on p. 2)

Gu, T., Kim, I., & Currie, G. (2019). Measuring immediate impacts of a new mass transit system on an existing bike-share system in China. *Transportation Research Part A: Policy and Practice*, *124*, 20–39. https://doi.org/10.1016/j.tra.2019.03.003 (cit. on p. 9)

Hagberg, A. A., Schult, D. A., & Swart, P. J. (2008). Exploring Network Structure, Dynamics, and Function using NetworkX, In *Proceedings of the 7th python in science conference*, Pasadena, CA USA. (Cit. on p. 25).

Hong, Z., Mittal, A., & Mahmassani, H. S. (2016). Effect of Bicycle-sharing on Public Transport Accessibility: Application to Chicago Divvy Bicycle-sharing System, In *Transportation research board 95th annual meeting*, Washington DC, United States. (Cit. on p. 8).

Jäppinen, S., Toivonen, T., & Salonen, M. (2013). Modelling the potential effect of shared bicycles on public transport travel times in Greater Helsinki: An open data approach. *Applied Geography*, *43*arXiv z0037, 13–24. https://doi.org/10.1016/j.apgeog.2013.05.010 (cit. on p. 8)

Ji, Y., Ma, X., Yang, M., Jin, Y., & Gao, L. (2018). Exploring spatially varying influences on metro-bikeshare transfer: A geographically weighted poisson regression approach. *Sustainability (Switzerland)*, *10*(5), 1526. https://doi.org/10.3390/su10051526 (cit. on pp. 10, 11)

Jin, H., Jin, F., Wang, J., Sun, W., & Dong, L. (2019). Competition and Cooperation between Shared Bicycles and Public Transit: A Case Study of Beijing. *Sustainability*, *11*(5), 1323. https://doi.org/10.3390/su11051323 (cit. on p. 3)

Krizek, K. J., & Stonebraker, E. W. (2011). Assessing Options to Enhance Bicycle and Transit Integration. *Transportation Research Record: Journal of the Transportation Research Board*, *2217*(1), 162–167. https://doi.org/10.3141/2217-20 (cit. on p. 2)

Lesh, M. C. (2013). Innovative Concepts in First-Last Mile Connections to Public Transportation, In *Urban public transportation systems 2013*, Reston, VA, American Society of Civil Engineers. https://doi.org/10.1061/9780784413210.007. (Cit. on p. 1)

Lin, L., He, Z., & Peeta, S. (2018). Predicting station-level hourly demand in a large-scale bike-sharing network: A graph convolutional neural network approach. *Transportation Research Part C: Emerging Technologies*, *97*, 258–276. https://doi.org/10.1016/j.trc.2018.10.011 (cit. on p. 12)

Litman, T. (2020). *Evaluating Public Transit Benefits and Costs* (tech. rep.). Victoria Transport Policy Institute. (Cit. on p. 1).

Liu, Z., Jia, X., & Cheng, W. (2012). Solving the Last Mile Problem: Ensure the Success of Public Bicycle System in Beijing. *Procedia - Social and Behavioral Sciences*, *43*, 73–78. https://doi.org/10.1016/j.sbspro.2012.04.079 (cit. on p. 2)

Lu, M., Hsu, S. C., Chen, P. C., & Lee, W. Y. (2018). Improving the sustainability of integrated transportation system with bike-sharing: A spatial agent-based approach. *Sustainable Cities and Society*, *41*, 44–51. https://doi.org/10.1016/j.scs.2018.05.023 (cit. on p. 8)

Ma, T., Liu, C., & Erdoğan, S. (2015). Bicycle sharing and public transit: Does capital bikeshare affect metrorail ridership in Washington, D.C.? *Transportation Research Record*, *2534*, 1–9. https://doi.org/10.3141/2534-01 (cit. on pp. 3, 9)

Ma, X., Zhang, X., Li, X., Wang, X., & Zhao, X. (2019). Impacts of free-floating bikesharing system on public transit ridership. *Transportation Research Part*

D: Transport and Environment, *76*, 100–110. https://doi.org/10.1016/j.trd.2019.09.014 (cit. on p. 9)

Ma, X., Ji, Y., Yang, M., Jin, Y., & Tan, X. (2018). Understanding bikeshare mode as a feeder to metro by isolating metro-bikeshare transfers from smart card data. *Transport Policy*, *71*(August 2017), 57–69. https://doi.org/10.1016/j.tranpol.2018.07.008 (cit. on p. 8)

Martin, E. W., & Shaheen, S. A. (2014). Evaluating public transit modal shift dynamics in response to bikesharing: A tale of two U.S. cities. *Journal of Transport Geography*, *41*(9), 315–324. https://doi.org/10.1016/j.jtrangeo.2014.06.026 (cit. on pp. 6, 7)

Minetti, A. (2000). The three modes of terrestrial locomotion. *Biomechanics and biology of movement*, 67–78 (cit. on p. 46).

Moncayo-Martínez, L. A., & Ramirez-Nafarrate, A. (2016). Visualization of the mobility patterns in the bike-sharing transport systems in Mexico City, In *Ieee international conference on industrial engineering and engineering management*, IEEE Computer Society. https://doi.org/10.1109/IEEM.2016.7798198. (Cit. on p. 11)

Moorthy, A., De Kleine, R., Keoleian, G., Good, J., & Lewis, G. (2017). Shared Autonomous Vehicles as a Sustainable Solution to the Last Mile Problem: A Case Study of Ann Arbor-Detroit Area. *SAE International Journal of Passenger Cars - Electronic and Electrical Systems*, *10*(2). https://doi.org/10.4271/2017-01-1276 (cit. on p. 2)

NACTO. (2018). *Shared Micromobility in the U.S.: 2018* (tech. rep.). National Association of City Transportation Officials. https://nacto.org/shared-micromobility-2018/. (Cit. on p. 2)

O'Brien, O., Cheshire, J., & Batty, M. (2014). Mining bicycle sharing data for generating insights into sustainable transport systems. *Journal of Transport Geography*, *34*, 262–273. https://doi.org/10.1016/j.jtrangeo.2013.06.007 (cit. on p. 11)

Oliveira, G. N., Sotomayor, J. L., Torchelsen, R. P., Silva, C. T., & Comba, J. L. (2016). Visual analysis of bike-sharing systems. *Computers and Graphics (Pergamon)*, *60*, 119–129. https://doi.org/10.1016/j.cag.2016.08.005 (cit. on p. 11)

Oppermann, M., Möller, T., & Sedlmair, M. (2018). BikeSharingAtlas: Visual Analysis of Bike-Sharing Networks. *International Journal of Transportation*. http://eprints.cs.univie.ac.at/5855/ (cit. on p. 11)

Pearson correlation coefficient. (n.d.). https://en.wikipedia.org/wiki/Pearson%7B%5C_%7Dcorrelation%7B%5C_%7Dcoefficient%7B%5C#%7Dcite%7B%5C_%7Dnote-2. (Cit. on p. 30)

Rasp, E., DesRoches, C., & Lee, C. (2019). *Mobility Hubs* (tech. rep.). City of Minneapolis. http://minneapolismn.gov/publicworks/trans/WCMSP-220794. (Cit. on p. 2)

Saberi, M., Ghamami, M., Gu, Y., Shojaei, M. H. (, & Fishman, E. (2018). Understanding the impacts of a public transit disruption on bicycle sharing mobility patterns: A case of Tube strike in London. *Journal of Transport Geography*, *66*(November 2017), 154–166. https://doi.org/10.1016/j.jtrangeo.2017.11.018 (cit. on p. 9)

Scheltes, A., & de Almeida Correia, G. H. (2017). Exploring the use of automated vehicles as last mile connection of train trips through an agent-based simulation model: An application to Delft, Netherlands. *International Journal of Transportation Science and Technology*, *6*(1), 28–41. https://doi.org/10.1016/j.ijtst.2017.05.004 (cit. on p. 2)

Shaheen, S. A. (2012). Introduction: Shared-Use Vehicle Services for Sustainable Transportation: Carsharing, Bikesharing, and Personal Vehicle Sharing across the Globe. *International Journal of Sustainable Transportation*, *7*(1), 1–4. https://doi.org/10.1080/15568318.2012.660095 (cit. on p. 2)

Shaheen, S. A., Guzman, S., & Zhang, H. (2010). Bikesharing in Europe, the Americas, and Asia: Past, Present, and Future. *Transportation Research Record: Journal of the Transportation Research Board*, *2143*(1), 159–167. https://doi.org/10.3141/2143-20 (cit. on p. 2)

Shaheen, S. A., Zhang, H., Martin, E., & Guzman, S. (2011). China's Hangzhou Public Bicycle: Understanding Early Adoption and Behavioral Response to Bikesharing. *Transportation Research Record: Journal of the Transportation Research Board*, *2247*(1), 33–41. https://doi.org/10.3141/2247-05 (cit. on p. 6)

Shaheen, S., & Chan, N. (2016). Mobility and the sharing economy: Potential to facilitate the first-and last-mile public transit connections. *Built Environment*, *42*(4), 573–588. https://doi.org/10.2148/benv.42.4.573 (cit. on p. 2)

Shaheen, S., Martin, E., & Cohen, A. (2013). Public Bikesharing and Modal Shift Behavior: A Comparative Study of Early Bikesharing Systems in North America. *International Journal of Transportation*, *1*(1), 35–54. https://doi.org/10.14257/ijt.2013.1.1.03 (cit. on p. 6)

Shen, Y., Zhang, H., & Zhao, J. (2018). Integrating shared autonomous vehicle in public transportation system: A supply-side simulation of the first-mile service in Singapore. *Transportation Research Part A: Policy and Practice*, *113*, 125–136. https://doi.org/10.1016/j.tra.2018.04.004 (cit. on p. 2)

Tilahun, N., Thakuriah, P. V., Li, M., & Keita, Y. (2016). Transit use and the work commute: Analyzing the role of last mile issues. *Journal of Transport Geography*, *54*, 359–368. https://doi.org/10.1016/j.jtrangeo.2016.06.021 (cit. on p. 1)

Vogel, P., Greiser, T., & Mattfeld, D. C. (2011). Understanding bike-sharing systems using Data Mining: Exploring activity patterns, In *Procedia - social and behavioral sciences*, Elsevier. https://doi.org/10.1016/j.sbspro.2011.08.058. (Cit. on p. 11)

Walker, J. (2012). *Human transit: How clearer thinking about public transit can enrich our communities and our lives*. Island Press-Center for Resource Economics. https://doi.org/10.5822/978-1-61091-174-0. (Cit. on p. 46)

Wang, B., & Kim, I. (2018). Short-term prediction for bike-sharing service using machine learning, In *Transportation research procedia*, Elsevier B.V. https://doi.org/10.1016/j.trpro.2018.11.029. (Cit. on p. 12)

Wu, J. L., & Chang, P. C. (2016). A Prediction System for Bike Sharing Using Artificial Immune System with Regression Trees, In *Proceedings - 2015 iiai 4th international congress on advanced applied informatics, iiai-aai 2015*, Institute of Electrical; Electronics Engineers Inc. https://doi.org/10.1109/IIAI-AAI.2015.159. (Cit. on p. 11)

Xu, C., Ji, J., & Liu, P. (2018). The station-free sharing bike demand forecasting with a deep learning approach and large-scale datasets. *Transportation Research Part C: Emerging Technologies*, *95*, 47–60. https://doi.org/10.1016/j.trc.2018.07.013 (cit. on p. 12)

Yan, Y., Tao, Y., Xu, J., Ren, S., & Lin, H. (2018). Visual analytics of bike-sharing data based on tensor factorization. *Journal of Visualization*, *21*(3), 495–509. https://doi.org/10.1007/s12650-017-0463-1 (cit. on p. 11)

Yang, X. H., Cheng, Z., Chen, G., Wang, L., Ruan, Z. Y., & Zheng, Y. J. (2018). The impact of a public bicycle-sharing system on urban public transport networks. *Transportation Research Part A: Policy and Practice*, *107*(September 2016), 246–256. https://doi.org/10.1016/j.tra.2017.10.017 (cit. on p. 8)

Zhang, Y., Brussel, M. J., Thomas, T., & van Maarseveen, M. F. (2018). Mining bike-sharing travel behavior data: An investigation into trip chains and transition activities. *Computers, Environment and Urban Systems*, *69*, 39–50. https://doi.org/10.1016/j.compenvurbsys.2017.12.004 (cit. on p. 11)

Zhang, Y., & Zhang, Y. (2018). Associations between public transit usage and bike-sharing behaviors in the United States. *Sustainability (Switzerland)*, *10*(6), 1868. https://doi.org/10.3390/su10061868 (cit. on pp. 6, 7)