DEPARTMENT OF TRANSPORTATION

Mobile-Device Data, Non-Motorized Traffic Monitoring, and Estimation of Annual Average Daily Bicyclist and Pedestrian Flows

Raphael Stern, Principal Investigator
University of Minnesota

June 2024

Research Project
Final Report 2024-19

LRRB
LOCAL ROAD RESEARCH BOARD

To request this document in an alternative format, such as braille or large print, call 651-366-4718 or 1-800-657-3774 (Greater Minnesota) or email your request to ADArequest.dot@state.mn.us. Please request at least one week in advance.

Technical Report Documentation Page

1. Report No.	2.	3. Recipients Accession No.	
MN 2024-19			
4. Title and Subtitle Mobile-device data, non-motorized traffic monitoring, and estimation of annual average daily bicyclist and pedestrian flows		5. Report Date June 2024 6.	
7. Author(s) Simanta Barman, Michael W. Levin, Gregory Lindsey, Michael Petesch, Suzy Scotty, Raphael Stern		8. Performing Organization Report No.	
9. Performing Organization Name and Address University of Minnesota 500 Pillsbury Dr. SE Minneapolis, MN 55406		10. Project/Task/Work Unit No. #2022006 11. Contract (C) or Grant (G) No. (c) 1036342 (wo) 14	
12. Sponsoring Organization Name and Address Minnesota Department of Transportation Office of Research & Innovation 395 John Ireland Boulevard, MS 330 St. Paul, Minnesota 55155-1899		13. Type of Report and Period Covered Final Report 14. Sponsoring Agency Code	
15. Supplementary Notes http://mdl.mndot.gov/			
16. Abstract (Limit: 250 words) People who walk and bike are the most vulnerable road users. However, understanding where they walk and bike requires continual data monitoring. Traditional methods rely on physical sensors in the infrastructure to detect the presence of pedestrians and bicyclists. However, these are expensive to deploy and only detect road users at the specific locations they are deployed. Instead, this study develops methods to use mobile phone based GPS data to estimate the number of bicyclists and pedestrians, and applies this methodology to the Twin Cities Metro area in Minnesota. The developed methodology is able to estimate average pedestrian and bicyclist volumes with relatively high accuracy.			
17. Document Analysis/Descriptors Pedestrians, Cyclists, Geospatial data, Smartphones		18. Availability Statement No restrictions. Document available from: National Technical Information Services, Alexandria, Virginia 22312	
19. Security Class (this report) Unclassified	20. Security Class (this page) Unclassified	21. No. of Pages 155	22. Price

Mobile-device data, non-motorized traffic monitoring, and estimation of annual average daily bicyclist and pedestrian flows

Final Report

Prepared by:

Simanta Barman
Michael Levin
Greg Lindsey
Raphael Stern
University of Minnesota

Michael Petesch
Suzy Scotty
Minnesota Department of Transportation

June 2024

Published by:

Minnesota Department of Transportation
Office of Research & Innovation
395 John Ireland Boulevard, MS 330
St. Paul, Minnesota 55155-1899

Table of Contents

List of Figures

List of Tables

1 Introduction

1.1 Introduction

People who walk and bike are the most vulnerable users of a transportation system. Construction and maintenance of sufficient infrastructure for the safe and comfortable traveling experience for these users is very important. To effectively allocate the limited amount of available budget to create the necessary infrastructure, understanding pedestrian and bicyclist traffic flow patterns is vital.

Historically, MnDOT and Minnesota counties have lacked estimates of bicyclist and pedestrian traffic on trunk highways (TH) and country state aid highways (CSAH). MnDOT began monitoring bicyclist and pedestrian flows at more than 25 locations across the state in about 2016. However, monitoring of non-motorized traffic through these monitoring counters is challenging and time consuming. The accurate estimation of traffic flow across the entire transportation network requires the deployment of a high number of monitoring sites. Conducting traffic monitoring on that scale is cost prohibitive and likely not feasible.

One way to work around the challenge of limited pedestrian and bicycle data, and the high cost of data collection is to rely on mobile data sources. With the development of technologies like Bluetooth and GPS, new data sources are becoming available that can be leveraged to get a good estimation travel data like roadway (or link) flows. With more people using these technologies (e.g., on smart phones, smart watches, etc.), the data sources are also getting richer in terms of the amount of data that is available. Traffic data from these sources can be extracted inexpensively compared to manual data collection. However, these data sources are still very new, and the accuracy of the data is unknown. Moreover, the quality of such data may vary spatially, depending on the prevalence of the technology that collects mobile-source data. For example this project has shown that urban locations where more people use these technologies and network service is good provides better quality of data compared to rural locations where data is more sparse.

This main objective of this project was to estimate annual average daily bicyclists and pedestrians (AADB; AADP) using sparse pedestrian and bike data collected from mobile phones. This project used mobile data to inexpensively obtain estimates of these traffic-flow measures. Trip distribution models and network route choice models to predict road flows of pedestrians and bicyclists were constructed, and filtering techniques were used to improve the models to get estimates of AADB and AADP that matches with the available baseline traffic counts. As part of this project we have created an interactive map that shows the estimated AADB and AADP on each road segment in the Twin Cities network.

1.2 Economic Benefits

Effective development of infrastructure for pedestrians and bicyclists will result in substantial economic benefits. For the development and maintenance of this infrastructure, a good understanding of pedestrian and bicyclist traffic flows is vital. One major use of this active transportation volume data will be to help prioritize infrastructure on projects to expand on and build local networks. These infrastructure changes can create the following economic benefits.

1.2.1 Job creation

The American Association of State Highway and Transportation Officials (AASHTO) commissioned 2012 study found that transportation projects for greenways, sidewalks and bikeways created more jobs than any other type of projects at 17 jobs per $\$1$ million spent. The 2018 Benchmarking Report from the League of American Bicyclists show how economic development and job creation in communities are positively affected by bicycle tourism. Political Economy Research Institute reported in their 2012 study that pedestrian and bicyclist infrastructure projects created more than 11.41 jobs while road-only projects had the lowest level of job creation at 7.75 jobs per $\$1$ million invested.

1.2.2 Reduction in health-related costs

A 2019 report from Rails-To-Trails Conservancy reports a saving of $\$630$ to $\$1437$ on health costs due to walking and biking. Infrastructure development and providing facilities to make sure walking and biking is maximized would result in more than $\$2$ billion per year savings in health related costs in Minnesota even with very conservative estimates. American Public Health Association and Urban Design 4 Health, Inc.

estimated in a 2010 report that the national health costs for obesity/overweight to be $\$142$ billion. They also reported the annual cost of traffic crashes to be $\$180$ billion in 2008. PeopleForBikes and Alliance for Biking and Walking reported biking can improve the physical fitness of employees resulting up to 32% fewer sick days, 55% lower health costs, and 52% increased productivity.

1.2.3 Economic growth

According to Smart Growth America and National Complete Streets Coalition projects that support walking, biking and moving actively using assistive devices increased property and sales tax revenue by up to 10 times while requiring less than 75% of the budget compared to motorized traffic-focused transportation projects. Multiple Studies have found that non-motorists spend more per month than those who drive. Rails-to-Trails Conservancy in their 2019 reported found that modest public investment in completing trail and active transportation networks will result in safe and seamless walking and biking routes, improved health, economic growth, social connectivity, access to jobs, education and culture. They predicted an annual economic return of $\$73.8$ billion. In the best case scenario, they predicted these benefits to nearly double to more than $\$138.5$ billion annually. Several studies reported that investment in non-motorized infrastructure resulted in increased real estate values.

According to AAA, the average annual cost of owning and operating a new motorized vehicle in 2022 was $\$10,728$ and the cost of biking was only $\$308$, according to a 2020 figure from the League of American Bicyclists and the Sierra Club. Smart Growth America in their 2015 Safer Streets, Stronger Economies report presented that their Complete Street projects that were designed to provide safe access to all users including pedestrians and bicyclists averted $\$18.1$ million in collision and injury costs in one year. Cost of pedestrian crashes in the highest priority area Minnesota was $\$580.3$ millions in $2016-2018$. Researchers at UC Berkeley found that Nice Ride Minnesota bikesharing riders spent an average of $\$1.25$ per week on new economic activity. This bikesharing system also resulted in approximately $\$29,000$ of new economic activity per season in the Twin Cities. Numerous other reports and papers have presented the economic benefits that have resulted from an improved and efficient infrastructure for pedestrians and bicyclists. This project will result in better understanding of pedestrians and bicyclists traffic patterns. Insights from those traffic patterns are crucial for designing effective infrastructure that will result in such economic benefits.

1.2.4 Summary of the benefits

The completed work presented in this report will directly benefit MnDOT by providing data-driven tools to assess walking and biking demand in the Twin Cities area. This approach could also be extended to other areas in Minnesota to provide a better picture of where people are walking and biking. Equipped with this information, MnDOT can make data-informed decisions to make walking and biking safer and more enjoyable, which may have additional, broader, benefits to society. For example, consumer expenditures on automobile purchase, ride-sharing or taxi service costs, public transit fares, exercise equipment cost, and gym membership can be decreased dramatically if walking and cycling can be improved. Health-related costs, job creation, property value increase, automobile ownership cost reduction etc. are also some of the other benefits that can be observed by increasing pedestrian and bicycle movement. Some of the non-monetary benefits are presented below.

- High potential option value for non-motorized transportation facilities that provide more transportation mode choice.
- Social equity benefits for those with limited driving ability due to low income, disability etc.
- Increase in socialization, recreation, excercise etc.
- Non-motorized travel is less stressful than driving, which contributes to mental health improvement.
- Energy conservation and pollution reduction.
- Improvement in land-use efficiency.

Table 1.1 shows the monetary benefits that can be gained from converting motorized traffic into pedestrian and bicyclist traffic, according to 2021 report by the Victoria Transport Policy Institute.

According to MnDOT, the installation, operation and maintenance costs over a five-year period was $\$560,000$ with 32 active monitoring sites. With the use of mobile-sourced data, enough data needed for traffic flow estimation can be collected inexpensively. The total budget for this study is $\$165,829$ to develop a methodology and software tool to predict sufficiently accurate traffic flow using mobile data sources, which is considerably less than the budget required to collect all non-motorized traffic data manually. Thus, the additional cost of collecting mobile-source data to recompute AADB and AADP is relatively small, compared to the cost of installing additional sensors.

Table 1.1: Monetary benefits for converting motorized traffic to pedestrian or bicycle traffic per day.

Benefits	Per Mile	Per Commuter
Congestion Reduction	$0.162	$0.81
Roadway Cost Savings	$0.041	$0.2025
Vehicle Cost Savings	$0.2025	$1.0125
Parking Costs Savings	–	$ 4.05
Air Pollution Reduction	$0.081	$0.41
Noise Pollution Reduction	$0.041	$0.2025
Energy Conservation	$0.041	$0.2025
Traffic Safety Benefits	$0.041	$0.2025
Total		$7.0925

Victoria Transport Policy Institute summarizes the benefits and costs from pedestrian and bicyclist traffic as follows in Table 1.2.

Table 1.2: Benefits and costs from pedestrian and bicyclist traffic as identified by the Victoria Transport Policy Institute.

User benefits	Increased user convenience, comfort, safety, accessibility and enjoyment
Option value	Benefits of having mobility options available in case they are ever needed
Equity objectives	Benefits to economically, socially or physically disadvantaged people
Fitness and health	Improved public fitness and health
Reduced Vehicle Travel	Benefits from reduced motor vehicle ownership and use
Vehicle cost savings	Consumer savings from reduced vehicle ownership and use
Avoided chauffeuring	Reduced chauffeuring responsibilities due to improved travel options
Congestion reduction	Reduced traffic congestion from automobile travel on congested roadways
Reduced barrier effect	Improved active travel conditions due to reduced traffic speeds and volumes
Roadway cost savings	Reduced roadway construction, maintenance and operating costs
Parking cost savings	Reduced parking problems and facility cost savings
Energy conservation	Economic and environmental benefits from reduced energy consumption

Pollution reductions	Economic and environmental benefits from reduced air, noise and water pollution
Land Use Impacts	Benefits from support for strategic land use objectives
Pavement area	Can reduce road and parking facility land requirements
Development patterns	Helps create more accessible, compact, mixed, infill development (smart growth)
Economic Development	Benefits from increased productivity and employment
Increased productivity	Increased economic productivity by improving accessibility and reducing costs
Labor productivity	Improved access to education and employment, particularly by disadvantaged workers
Shifts spending	Shifts spending from vehicles and fuel to goods with more regional economic value
Support specific industries	Support specific industries such as retail and tourism
Costs	Costs of improving non-motorized traffic conditions
Facilities and programs	Costs of building non-motorized facilities and operating special programs
Vehicle traffic impacts	Incremental delays to motor vehicle traffic or parking
Equipment	Incremental costs to users of shoes and bicycles
Travel time	Incremental increases in travel time costs due to slower modes
Accident risk	Incremental increases in accident risk

Obtaining an accurate estimate of traffic flow is crucial to developing infrastructure and facilities to gain the benefits described above.

1.3 Summary

The principal objective of this project was to develop a new method and procedures for modeling and visualizing AADB and AADP on individual roads or intersections. Because of the high expenses and privacy concerns of developing travel behavior models, which require manual traffic monitoring and conducting

household surveys, this project relies mostly on mobile-sourced data. However, the mobile sourced-data contains errors, the degree of which may differ among different locations and types of roads. Therefore, we have developed a methodology to combine mobile-sourced data with a trip behavioral model to correct erroneous data and estimate the true AADB and AADP.

2 Review of Relevant Literature

2.1 Introduction

Active travel modes, such as walking and biking, are important to address many issues in modern society, such as auto dependence, air pollution, climate change, obesity, and physical and mental health. However, the share of these active travel modes has been decreasing in recent decades (McGuckin and Fucci, 2018). Better understanding the spatial and temporal distribution of pedestrian and bicyclist traffic is necessary for scholars, planners, and engineers to promote active travel modes. In contrast to vehicular monitoring programs, the investment in pedestrian and bicyclist traffic monitoring programs has been very limited (Federal Highway Administration, 2001). Only three states (Minnesota, Colorado, and North Carolina) have established monitoring programs (Ohlms et al., 2019). Thirteen states have various efforts in progress (Ohlms et al., 2019), however, that do not include nonmotorized traffic monitoring.

To address the lack of comprehensive pedestrian and bicyclist traffic volume data worldwide, scholars historically have utilized the limited observed pedestrian and bicyclist count data with methods such as expansion factors and direct demand models to estimate the pedestrian and bicyclist traffic volumes for all segments in a network (Turner et al., 2017). Expansion factors show the ratio between short-time (e.g., hour and day) and long-time (e.g., year) pedestrian and bicyclist traffic and can be used to transfer short-time traffic into long-time one. Direct demand models establish the correlation between pedestrian and bicyclist traffic and multiple types of factors, such as built environment characteristics, traffic facility features, weather, etc., and use these correlations to predict traffic (see Section 2.4.2 for more discussion).

In recent years, researchers have begun to use crowdsourced mobile data, mainly from cellular signal or smartphone applications, when estimating pedestrian and bicyclist volumes. Crowdsourced mobile data are defined as mobile data collected from "a group of individuals of varying knowledge, heterogeneity, and number, via a flexible open call, the voluntary undertaking of a task (Estellés-Arolas and González-Ladrón-

de Guevara, 2012). They are collected through either cellular signal or applications of mobile devices. The mobile devices have a large number of users across both time and space. Compared with traditional count datasets, which are collected by a limited number of sensors or people, crowdsourced mobile data have a higher spatial and temporal coverage. The cost of data collection is also lower for crowdsourced mobile data as they are mostly by-products of the services provided by mobile devices. Although new approaches for using crowdsourced mobile data have been proposed, and vendors with proprietary models for estimating bicyclist and pedestrian traffic volumes exist, routine procedures for using crowdsourced data have yet to be developed and standardized. Given the rapid evolution in the field and the objectives of this project, it is useful to review related studies and provide a synthesis of crowdsourced mobile data and related methods for estimating bicyclist and pedestrian traffic.

In this report, through a systematic literature review approach, we aim to comprehensively describe how crowdsourced mobile data have been applied to estimate pedestrian and bicyclist traffic volumes. Specially, we address the following research questions: (1) what crowdsourced mobile data have been applied and (2) what methods have been used to carry out pedestrian and bicyclist traffic forecasting.

The rest of this chapter is organized as follows. We illustrate our method of systematic literature review in Section 2.2. We compare different types of crowdsourced mobile data in Section 2.3. We present our review results in Section 2.4. In Section 2.5, we conclude our findings for this report.

2.2 Method

We applied a systematic review approach in this study. A systematic literature review differs from a traditional literature review in that it attempts to identify all papers written on a topic using explicit inclusion and exclusion criteria and to characterize the state of knowledge based on comprehensive assessment of evidence. Traditional literature reviews are more likely to be purposeful and selective, with sources identified to make particular theoretical or practical points.

Systematic or evidence-based reviews have several key steps (Xiao and Watson, 2019). First, we defined our research topics or questions. In this study, our research questions were, “What crowdsourced mobile data and what methods have been used to carry out pedestrian and bicyclist traffic forecasting?” Second, we produced a protocol for conducting the review. The protocol specified the search criteria, selection criteria, and coding categories. Third, we searched several databases of literature with the keywords related

to our research topic. Fourth, we screened the search results and included or removed studies based on the selection criteria. Fifth, we extracted the information defined by the coding categories after thoroughly reviewing the selected studies. Finally, we analyzed the results and reported our findings.

We defined three sets of keywords related to our research question. The first set includes pedestrian, bicyclist, and bike. These keywords identify the travel modes we want to study. The second set includes demand, traffic, volume, AADB (Annual Average Daily Bicyclists), and AADP (Annual Average Daily Pedestrians). These keywords are different terms or measures used to describe pedestrian or bicyclist traffic. The last set of keywords include forecasting, estimation, and prediction. They are closely related to traffic forecasting. We did not define keywords related to crowdsourced mobile data as we wanted to include more studies in the search process. We selected those using crowdsourced mobile data in the step of screening. Some less-used keywords were not included, for example, cyclists or non-motorized travel. We did this for two reasons. First, these keywords are closely related to the keywords we have defined in the protocol. Therefore, studies using these keywords often showed up in the search results. Second, we also included additional studies when we found they are related to our study during the review process. We searched four databases, including Web of Science, Google Scholar, Academic Search Premier, and PubMed. We searched the keywords in the title, abstracts, and keywords of the studies. All the results were searched before September 30th, 2021.

Screening search results included two rounds (Figure 2.1). In the first round, we skimmed the titles, abstracts, and keywords of the studies and only included (1) studies published in English; (2) journal articles, conference proceedings, book sections, theses, and research reports; and (3) studies that propose a method or apply a method to predict pedestrian and/or bicyclist traffic with one or more datasets. After this round, we found 172 studies from the search results. In the second round of screening, we reviewed the title, abstract, and introduction of the searched items and removed replicate studies, giving priority to journal articles over reports when they included the same findings. Second, we removed articles that are not closely related to the topic, for example, studies on bike sharing system. Third, we removed literature review studies. Although we excluded literature reviews from the coding process we still included them in this report where relevant, including in the introduction and in our discussion of results. Finally, we removed studies that did not incorporate or use crowdsourced data. After this round, we had only 12 studies left for further review and analysis.

We extracted several types of information from each of the 12 studies, including general information

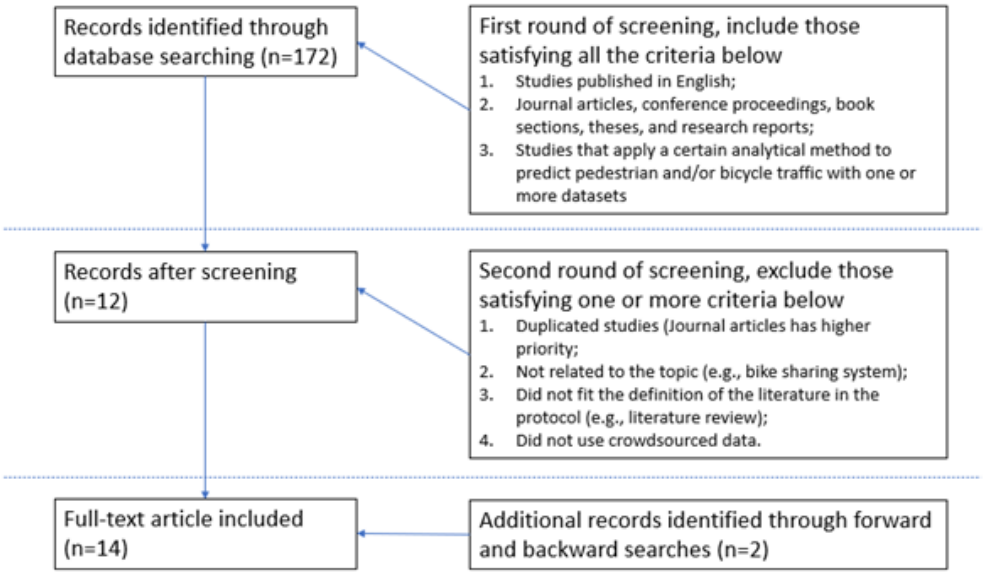


Figure 2.1: Searching and Screening results

about the studies (e.g., authors, year, and title), study locations, travel modes, application of their results, types of crowdsourced data, information about pedestrian or bicyclist count data, pros and cons of the crowdsourced data, and methods. During this review process, we found two additional related studies in the reference lists. Therefore, we include 14 studies in this review (Figure 2.1).

2.3 Crowdsourced mobile data sources and comparison

Traditional count data are either automatically collected by sensors or manually counted by people. Crowdsourced mobile data are defined as mobile data collected from "a group of individuals of varying knowledge, heterogeneity, and number, via a flexible open call, the voluntary undertaking of a task (Estellés-Arolas and González-Ladrón-de Guevara, 2012)." Three types of crowdsourced mobile data have been discussed in the literature: fitness application data, application location-based service data, and cellular signal data (Table 2.1).

Fitness application data are the data collected from smartphone fitness applications. One popular example of these smartphone fitness applications is Strava[1]. While the main function of these fitness applications is to help users record their GPS (Geographical Positioning System) trajectories during physical activities, the information of these trajectories can be utilized to support pedestrian and bicyclist traffic vol-

[1] https://www.strava.com/ (accessed 12.21.2021)

Table 2.1: Comparison among different types of crowdsourced mobile data and traditional data

	Advantages	Disadvantages
Traditional observed count data	• Representation of the whole population • Accurate count data	• Small coverage of geographical area and temporal range • High cost of data collection • Malfunction of facilities
Fitness application data	• Large coverage in geographical area and temporal range • Low cost of data collection	• Biased toward runners and cyclists
Application location -based service data	• Large coverage in geographical area and temporal range • Low cost of data collection	• Biased toward smartphone users
Cellular signal data	• Large coverage in geographical area and temporal range • Low cost of data collection	• Biased toward mobile phone users • Low spatial precision

ume estimation. For example, Strava provides three purchasable licenses based upon which the trajectory data will be aggregated to intersections, road segments, and Origin-Destination pairs (Lee and Sener, 2020). The aggregated data include trip numbers by different directions, age groups, and genders. Some fitness applications do not provide aggregated data but GPS trajectories directly, such as Mon ResoVelo (Strauss et al., 2015). In this case, data users need to aggregate the GPS trajectories by themselves which involves a heavy work of data processing. Many studies have compared the Strava data with observed count data directly (Lee and Sener, 2021). Lee and Sener summarized the correlations between Strava and observed bicyclist counts and found that the correlations range from 0.36 to 0.83 (Lee and Sener, 2021), most of which are larger than 0.6. Compared with traditional count data, fitness application data have a larger geographical and temporal coverage and lower cost of data collection (Lee and Sener, 2021). However, as their users are mostly runners and cyclists, fitness application data bias toward these populations. In addition, among these subpopulations, there is disproportionate representation of people wealthy enough to purchase the devices on which they can be used.

Application location-based service data are the geographical location data collected from smartphone applications. When people use these applications, their GPS locations, user ID, and the corresponding time will be uploaded to the cloud servers (Nishi et al., 2014). For example, when using Yelp to search a nearby restaurant, the location and time of the user will be recorded. The data aggregated from multiple smartphone applications can provide valuable support for pedestrian and bicyclist volume estimation. When dealing with these data, data users need to develop algorithms to identify walking or biking trips from all activities recorded by the applications, which is a challenging task. Compared with traditional observed count data, application location-based service data cover a larger geographical area and a longer temporal

range and have a lower cost when collecting data. However, these data are biased toward smartphone users.

Cellular signal data are geographical location data collected from mobile devices when they connect to cellular networks or move across the cell tower boundaries (Lee and Sener, 2020). Cellular service providers record and maintain these data for operation and billing purposes. Similar to application location-based services, analytical algorithms such as statistical regression or machine learning are necessary to detect walking or biking trips from all activities. Compared with traditional count data, cellular signal data have a more through coverage in terms of geographical area and temporal range. The cost of the corresponding data collection is lower. However, cellular signal data have a lower spatial precision, ranging from 200 to 1000 meters. This limitation makes it difficult to recognize walking or cycling trips from these data, as these trips are usually very short. In addition, cellular signal data are biased toward mobile phone users.

Some companies, such as StreetLight and Cuebiq, purchase data from multiple sources and generate pedestrian and bicyclist trip information for their customers (Lee and Sener, 2020). For example, StreetLight provides StreetLight pedestrian or bicyclist index of each OD pair between traffic zones or through traffic analysis zones by specific time unit (e.g., hour or day). The StreetLight index is a pedestrian or bicyclist trip "count" estimated by its algorithm. Additional trip information includes travel time, length, speed, and circuity. Simple demographics of travelers, such as education and ethnicity, also can be offered. The models used to generate these estimates tend to be proprietary, which means that users do not know exactly how the estimates were produced. This characteristic potentially can limit the ability of analysts to develop custom applications and can create dependencies on particular vendors. One study compared the StreetLight bicyclist index with observed bicyclist counts of 32 locations from six cities in Texas and found that the correlations between these two data are 0.62 and 0.69 on weekdays and weekends respectively (Turner et al., 2020). In the same study, the scholars also compared the StreetLight bicyclist index with bicyclist miles traveled calculated with Strava trip counts. The established a linear regression, in which the StreetLight bicyclist index is the dependent variable and the Strava based bicyclist miles traveled is the independent variable. They found that the correlation between these two variables is 0.94 (Turner et al., 2020).

Both StreetLight and Strava data are currently available to the Minnesota Department of Transportation (Table 2.2). Strava data are available from 2016 to 2021. Strava provides the number of pedestrian or bicyclist counts based on Strava users' activities for each road segment in the state of Minnesota. Road segments generally include all road classifications, such as local roads and bikeways. Minnesota contains

Table 2.2: Comparison between Strava and StreetLight

Data	Time period	Area coverage	Each observation contains
Strava	2017–2021	Minnesota (any sub-area is selectable)	Pedestrian or bicyclist count for each road segment during a specified time unit (year, month, day, and hour)
StreetLight	2016–2021	Minnesota (any sub-area is selectable)	StreetLight pedestrian or bicyclist index through traffic analysis zones or of each OD pair between traffic analysis zones during a specified time unit (all week, weekdays, weekends, different time periods of a day)

more than 1 million road segments (1,327,746). Strava aggregates these counts into different time periods including year, month, day, and hour.

StreetLight data are available for the period from 2016 to 2021. StreetLight provides StreetLight pedestrian and bicyclist index through traffic analysis zones or of each OD pair between traffic analysis zones in the state of Minnesota. These traffic analysis zones can be defined by the users and input into the system. For example, the Twin Cities metropolitan area has 3030 traffic analysis zones, which are defined by the Metropolitan Council. In addition, StreetLight aggregates these counts into several time units, including all days (Monday to Sunday), weekdays (Monday to Thursday), and weekends (Saturday to Sunday). StreetLight also provides counts from different time periods of each day, including all day (12am–12am), early AM (12am–6am), peak AM (6am–10am), mid-day (10am–3pm), and late PM (7pm–12pm).

2.4 Results

We present the results of our review in three parts. We first present descriptive information about each of the 14 selected studies. Then, we review in detail each of the works that incorporate crowdsourced data in direct demand models, the most common approach reported by researchers to date. We summarize specification of each of the direct demand models including, in addition to the crowdsourced measure, the results of other independent variables. Finally, we briefly discuss the other two methods published in the literature: aggregation analysis and Strava user rate expansion.

2.4.1 General review

Authors, study locations, travel mode, and application are summarized in Table 2.3. The earliest study was published in 2014. Eleven studies focus on the context of the North America. Thirteen studies focus on the travel mode of bicycling. As to the applications of the studies, 10 studies are for traffic volume estimation

Table 2.3: General information of the studies

Author	Study location	Travel mode	Application/Purpose
Nishi et al. (2014)	Japan	Pedestrian	Traffic volume estimation
Strauss et al. (2015)	Montreal, Canada	Bicyclist	Crash analysis
Jestico et al. (2016)	Victoria, BC, Canada	Bicyclist	Traffic volume estimation
Sanders et al. (2017)	Seattle, US	Bicyclist	Traffic volume estimation
Lißner et al. (2018)	Dresden, Germany	Bicyclist	Traffic volume estimation
Roll (2018)	Eugene, Springfield, Coburg, OR, US	Bicyclist	Crash analysis and health analysis
Roy et al. (2019)	Maricopa County, AZ, US	Bicyclist	Traffic volume estimation
Kwigizile et al. (2019)	Ann Arbor and Grand Rapids, MI, US	Bicyclist	Traffic volume estimation
Saad et al. (2019)	Orange County, US	Bicyclist	Crash analysis
Dadashova and Griffin (2020)	TX, US	Bicyclist	Traffic volume estimation
Dadashova et al. (2020)	TX, US	Bicyclist	Traffic volume estimation
Lin and Fan (2020)	Charlotte, US	Bicyclist	Traffic volume estimation
Camacho-Torregrosa et al. (2021)	Spain	Bicyclist	Crash analysis
Nelson et al. (2021)	Boulder, Ottawa, Phoenix, San Francisco, Greater Victoria, North America	Bicyclist	Traffic volume estimation

and four for crash analysis and/or health analysis.

In Table 2.4, we report the modeling information of the selected studies. The studies varied in their use of crowdsourced data in modeling traffic in terms of analytic units, dependent variables, counts and sources of crowdsourced estimates, sample sizes, modeling approach or technique, and measures of evaluation. Specifically, 12 studies used road segment and four studies used analysis zone or intersection as the analytic unit. Eight studies estimated AADB. Six studies estimated daily, hourly, or short-time pedestrian or bicyclist volume. As to the main data sources, 13 studies used both observed count and data from smartphone fitness applications (Strava and Mon ResoVelo). One applied the location-based service data collected from Yahoo smartphone applications. The sample sizes, which indicate the number of locations or number of locations by time units, range from 14 to 8813. Generally, studies that estimated short-time volumes have larger sample sizes. Direct demand models (see Section 2.4.2 for more discussion) dominate the estimation methods: 12 studies used this approach. The R^2 of these models ranges from 0.48 to 0.8, meaning that the models can explain about half to more than three-quarters of the variation in actual, observed counts. The mean absolute percentage error (MAPE) in estimation ranges from 29% to 36%. One study applied aggregation analysis. One study used Strava user rate expansion. See Section 2.4.3 for more discussion about these two methods.

Table 2.4: Model information of the studies

Nishi et al. (2014)	Analysis zone	Hourly pedestrian volume	Smartphone users' location data accumulated on Yahoo App.		Aggregation analysis	
Strauss et al. (2015)[‡]	Intersection Road segment	AADB	Observed bicyclist count Mon ResoVelo GPS trip data	Signalized intersection model (638) Non-signalized intersection model (438) Cycle track model (70) Bicycle lane model (14) No facility model (36)	Direct demand model	R^2: 0.7 R^2: 0.58 R^2: 0.52 R^2: 0.76 R^2: 0.48
Jestico et al. (2016)	Road segment	Daily bicyclist volume (7–9 am and 3–6 pm combined)	Observed bicyclist count Strava bicyclist count	612	Direct demand model	
Sanders et al. (2017)	Road segment	AADB	Observed bicyclist count Strava bicyclist count	46	Direct demand model	R^2: 0.62

Table 2.4: Model information of the studies (continued)

Author (Year)	Analysis unit	Dependent Variable	Main data	Sample size	Estimation method	Evaluation[†]
Lißner et al. (2018)	Road segment	AADB	Observed bicyclist count	Strava bicyclist count	Direct demand model	MAPE: 36%; R^2:0.75
Roll (2018)	Road segment	AADB	Observed bicyclist count Strava bicyclist count	52	Direct demand model	R^2: 0.75
Roy et al. (2019)	Road segment	AADB	Observed bicyclist count Strava bicyclist count	44	Direct demand model	R^2: 0.64
Kwigizile et al. (2019)	Road segment	Hourly bicyclist volume	Observed bicyclist count Strava bicyclist count	1520	Direct demand model	
Saad et al. (2019)	Intersection	Daily bicyclist volume	Observed bicyclist count Strava bicyclist count	171	Direct demand model	R^2: 0.80
Dadashova and Griffin (2020)	Road segment	Daily bicyclist volume	Observed bicyclist count Strava bicyclist count	8813	Direct demand model	MAPE: 29%
Dadashova et al. (2020)	Road segment	AADB	Observed bicyclist count Strava bicyclist count	100	Direct demand model	R^2: 0.75; MAPE: 29%

Table 2.4: Model information of the studies (continued)

Author (Year)	Analysis unit	Dependent Variable	Main data	Sample size	Estimation method	Evaluation[†]
Lin and Fan (2020)	Road segment	Short-time period bicyclist volume	Observed bicyclist count Strava bicyclist count	Direct demand model	R^2: 0.61	
Camacho-Torregrosa et al. (2021)	Road segment	AADB	Observed bicyclist count Strava bicyclist count	Strava user rate expansion		
Nelson et al. (2021) [§]	Road segment Intersection Road segment Road segment Road segment	AADB Observed bicyclist count Strava bicyclist count	Boulder model (15) Ottawa model (1058) Phoenix model (35) San Francisco model (53) Greater Victoria model (54)	Direct demand model		

[†]Valuation indices only include R^2 and MAPE (Mean Absolute Percentage Error), which can be compared across models.

[‡]Strauss et al. (2015) estimated models for signalized intersections, non-signalized intersections, cycle track, bicycle lane, and road segments with no facility, respectively.

[§]Nelson et al. (2021) estimated models for Boulder, Ottawa, Phoenix, San Francisco, and Greater Victoria, respectively.

2.4.2 Direct demand model

Among the 14 studies, 12 applied direct demand models to estimate bicyclist traffic with crowdsourced data and other data sources. Direct demand models are based on the assumption that pedestrian or bicyclist traffic is correlated with several types of variables and that these correlational relationships can be estimated through statistical modeling or machine learning techniques. Analysts then can use the estimated models to predict the pedestrian or bicyclist traffic. The independent variables used in the 12 studies include crowdsourced mobile data (C), built environment (BE), socio-demographics (SD), traffic facility (TF), time (T), weather (W), and other variables (O). Equation (2.1) below is the conceptual direct demand model.

$$\text{Traffic} = f(C, BE, SD, TF, T, W, O) \tag{2.1}$$

where f specifies a modeling function, which could be a statistical model (e.g., ordinary least squared model or generalized linear model) or machine learning model (e.g., random forest). Specific measures for each variable are discussed in the following sections.

Determining what factors to be included is very important to establish a good direct demand model. After reviewing the 12 studies using direct demand models, we summarized their results in Tables 2.5 to 2.11. These tables summarize the directions of the linear relationships between traffic and significant variables with a plus sign (+) for a positive relationship and a minus sign (-) for a negative relationship. They also list the variables that were considered during the modelling process but were not significant or not included in the final modes. In this case, a character x is used after the variables. It is worth noting that all the 12 studies focus on bicyclist traffic volume estimation. In addition, because machine learning models did not provide significance levels for the independent variables and only one study applied machine learning models, this report did not include the result of machine learning approaches in the summary. In one study, scholars compared the performance of several models including both statistical models and machine learning approaches and found that the random forest machine learning approach had the best performance (Kwigizile et al., 2019).

The results related to crowdsourced data are summarized in Table 2.5. All 12 studies applied crowdsourced data as independent variables to help estimate or predict total bicyclist volumes. A commonly used form of crowdsourced mobile data is an aggregated bicyclist count on road segment. For intersections,

Strauss et al. (2015) used the bicyclist count from linked road segments with different bicycle facilities, and Saad et al. (2019) used total bicyclists entering the intersections from all linked road segments. All crowdsourced bicyclist count variables are significant and have positive relationships with bicyclist traffic except one study did not provide significance level (Table 2.4). In addition, Nelson et al. (2021) used percentage of crowdsourced trips that are commutes as one independent variable in their models. In their city models, three models showed that this variable is positively correlated with bicyclist traffic. In addition to the results summarized in Table 5, two studies also found that models with crowdsourced mobile data have higher prediction performance and fewer independent variables compared with those without crowdsourced mobile data (Sanders et al., 2017; Roll, 2018). These results confirm the substantial importance of crowdsourced mobile data in estimating bicyclist traffic despite the fact that they are biased and represent only a fraction of total bicyclist traffic.

Ten of the 12 papers included built environment variables in their final direct demand models (Table 2.5). Among these studies, we identified five categories of built environment variables: population measures, network design, land use, destination accessibility, and transit supply. Several observations can be drawn from these results. These built environment variables generally are measured for spatial units around the actual count locations, the assumption being that the magnitude of counts is associated with these variables. These spatial units could be census areas (e.g., census block group and census tract) or buffer areas (e.g., half-mile buffer within count location).

- Population measures include total population, population density, and population accessibility, which were included in seven studies. Among these seven studies, only one study (Nelson et al., 2021) showed that population density is positively correlated with bicyclist traffic.

- Land use variables are mostly related to commercial, institutional, residential, and recreational land use. These variables generally have two measurements. One is to indicate the general land use type of the spatial units around the count locations (in terms of dummy or categorical variables). The other is to measure the area of different land use types in spatial units (in terms of numerical variables). Different land use types have different capabilities to produce and attract pedestrians and bicyclists and, thus, are correlated differently with these traffics. They were considered in five studies. One study (Kwigizile et al., 2019) found that, relative to commercial land use, institutional land use and residential land use are positively correlated with bicyclist traffic.

Table 2.5: Direct demand models: Correlation of Crowdsourced mobile data with bicyclist traffic volume

Author (Year)	Crowdsourced bicyclist count on road segment*	Percentage of crowdsourced trips that are commutes
Strauss et al. (2015)		
• Signalized intersection model	Road segment with no cycling facilities (+) Road segment with bicycle path (+) Road segment with cycle track (+)	
• Non-signalized intersection model	Road segment with no cycling facilities (+) Road segment with bicycle path (+) Road segment with cycle track (+)	
• Cycle track model	+	
• Bicycle lane model	+	
• No bicycle facility model	+	
Jestico et al. (2016)	+	
Sanders et al. (2017)	+	
Lißner et al. (2018)	+	
Roll (2018)	NP†	
Roy et al. (2019)	+	
Kwigizile et al. (2019)	+	
Saad et al. (2019)	All road segments linked to the intersection (+)	
Dadashova and Griffin (2020)	+	
Dadashova et al. (2020)	+	
Lin and Fan (2020)	+	
Nelson et al. (2021)		
• Boulder model	+	-
• Ottawa model	+	+
• Phoenix model	+	+
• San Francisco model	+	-
• Greater Victoria model	+	+

*In most studies, the time period of the crowdsourced trip number on road segments were consistent with the dependent variables. However, in Sanders et al. (2017), the time periods were not matched due to the data availability.

†NP denotes no p-value, indicating that the variable was included in the final model, but the study did not provide the corresponding p-value.

- Destination accessibility indicates the presence of or the distance to certain activity destinations. Six studies considered these variables. Distance to downtown is negatively correlated with bicyclist traffic (Strauss et al., 2015). Presence of university is positively correlated with bicyclist traffic (Sanders et al., 2017). This is consistent with the negative relationship between distance to educational institution and bicyclist traffic (Nelson et al., 2021). Distance to green space generally has a negative relationship with bicyclist traffic (Roll, 2018; Nelson et al., 2021). Distance to bike parking and distance to seashore have negative correlations with bicyclist traffic (Nelson et al., 2021). Distance to residential area, however, has mixed relationships with bicyclist traffic in the literature. Roll (2018) found that residential area is negatively correlated with bicyclist traffic. However, Nelson et al. (2021) showed that it is positively correlated with bicyclist traffic in their Phoenix model.

- Network design and transit supply variables were included in two studies. Network design variables indicate the connectivity of the transport network in the areas where the count locations locate. Similarly, transit supply variables measure the supply of transit stops or routes in the areas where the count locations locate. None of these variables was significantly correlated with bicyclist traffic.

Table 2.6: Direct demand models: Correlation of built environment characteristics with bicyclist traffic volume

Strauss et al. (2015)					
• Signalized intersection model			Distance to downtown (-)		
• Non-signalized intersection model			Distance to downtown (-)		
• Cycle track model					
• Bicycle lane model					
• No bicycle facility model					
Jestico et al. (2016)	Population density (x)				
Sanders et al. (2017)		Number of commercial properties (x)	Presence of university (+) Distance to university (x) Presence of school (x)	Ratio of intersections to dead-end streets (x)	Transit frequency (x)
Lißner et al. (2018)					
Roll (2018)	Population accessibility (NP)		Retail accessibility (NP)	Local street miles (NP) Number of intersections (NP) Network centrality (Bike path) (NP)	
Roll (2018)	Population density (x)		Distance to residential area (-) Distance to green space (-) Distance to commercial area (x)		

Table 2.6: Direct demand models: Correlation of built environment characteristics with bicyclist traffic volume (continued)

Author (Year)	Population measure	Land use	Destination accessibility	Network design	Transit supply
Kwigizile et al. (2019)		Land use (Reference: Commercial): Institutional (+) Residential (+) Recreational (x)			
Saad et al. (2019)					
Dadashova and Griffin (2020)	Population density (x) Total population (x)	Land area (x)	Presence of certain places of interest (x)		
Dadashova et al. (2020)	Population density (x) Total population (x)	Land area (x)			
Lin and Fan (2020)	Total population (x)	Residential land use (dummy) (x) Commercial land use (dummy) (x)			
Nelson et al. (2021)					
• Boulder model	Population density (x)		Distance to green space (x) Distance to residential area (x) Distance to commercial area (x) Distance to trailhead (x)		

Table 2.6: Direct demand models: Correlation of built environment characteristics with bicyclist traffic volume (continued)

Author (Year)	Population measure	Land use	Destination accessibility	Network design	Transit supply
• Ottawa model	Population density (+)		Distance to green space (x) Distance to residential area (x) Distance to commercial area (x) Distance to water body (x)		
• Phoenix model	Population density (x)		Distance to green space (-) Distance to residential area (+) Distance to green space (x) Distance to commercial area (x)		
• San Francisco model	Population density (x)		Distance to green space (-) Distance to bike parking (-) Distance to seashore (-) Distance to residential area (x) Distance to commercial area (x) Distance to educational institution (x)		
• Greater Victoria model	Population density (x)		Distance to green space (-) Distance to educational institution (-) Distance to residential area (x) Distance to commercial area (x)		

Traffic facility characteristics include five types of variables, including geometric and road classification, control and flow, sign and marking, visibility, and road condition (Table 2.7). Ten studies considered these variables in their direct demand models. We summarized several findings from Table 2.7.

- Geometric and road classification are the traffic facility variables modeled most often (9 studies). Road segment slope is negatively correlated with bicyclist traffic (Jestico et al., 2016; Nelson et al., 2021). Bicycle facility, such as cycleway and trail, is generally positively correlated with bicyclist traffic (Sanders et al., 2017; Nelson et al., 2021; Dadashova et al., 2020; Kwigizile et al., 2019). However, there are several exceptions. Lin and Fan (2020) found that the presence of a bike lane is negatively correlated with bicyclist traffic. Also, Nelson et al. (2021) found that presence of protected/off-street bike lane is negatively correlated with bicyclist traffic in the Greater Victoria model. Suggested bike route has mixed correlations with bicyclist traffic. Lin and Fan (2020) showed that this variable is negatively correlated with bicyclist traffic, while Nelson et al. (2021) found that this variable is positively correlated with bicyclist traffic in the Ottawa model.

- Control and flow variables only include speed limit related variables. These variables were considered in six studies. The speed limit is negatively correlated with bicyclist traffic (Jestico et al., 2016; Roll, 2018) [24].

- Sign and marking related variables were only considered in one study, which is the presence of bicycle markings on the road segment. However, the relationship is not significant.

- Visibility-related variables include presence of parking and street lighting. Three studied included these variables. Among them, Jestico et al. (2016) showed that presence of on-street parking is negatively correlated with bicyclist traffic.

- Road condition-related variables were considered in two studies. However, none of these studies found significant correlations.

Table 2.7: Direct demand models: Correlation of traffic facility characteristics with bicyclist traffic volume

Strauss et al. (2015)					
Jestico et al. (2016)	Road segment slope (-) Pavement width (x)	Posted speed limit (Reference: 20km/h): 50 km/h (-) 40 km/h (-) 30 km/h (+)		Presence of on-street parking (-)	
Sanders et al. (2017)	Number of bike lanes (+) Trail (dummy) (x) One-way street (dummy) (x) Presence of bicycle lane or cycle track (x) Maximum slope (x) Average slope (x)		Presence of bicycle markings (x)		
Lißner et al. (2018)					
Roll (2018)	Presence of bike lane (dummy) (NP) Presence of bike boulevard (dummy) (NP) Road classification (Reference: Path): Collector (NP) Minor Arterial Street (NP)				
Roll (2018)		Average speed limit (-)			

Table 2.7: Direct demand models: Correlation of traffic facility characteristics with bicyclist traffic volume (continued)

Author (Year)	Geometric and road classification	Control and flow	Sign and marking	Visibility	Road condition
Kwigizile et al. (2019)	Bike facility (Reference: sidewalk/Shoulder): Shared lane (+) Bike lane (+) Trail (+)				
Saad et al. (2019)	Intersection size (+)				
Dadashova and Griffin (2020)	Open street map road classification (Reference: Highway path): Cycleway (+) Highway primary (x) Highway secondary (x) Highway tertiary (x) Footway (x) Inside shoulder width (x) Medium width (x) Non-motorized facility buffer width (x) Non-motorized facility type (x) Non-motorized facility width (x) Number of lanes (x) Outside shoulder width (x) Surface width (x)	Posted speed limit (x)		Presence of parking (x) Street lighting (x)	Pavement condition (x) Pavement type (x) Shaded area of road segment (x)

Table 2.7: Direct demand models: Correlation of traffic facility characteristics with bicyclist traffic volume (continued)

Author (Year)	Geometric and road classification	Control and flow	Sign and marking	Visibility	Road condition
	RHiNO road classification (x)				
Dadashova et al. (2020)	Open street map road classification (15 = Primary; 21 = Secondary; 31 = Tertiary; 32 = Residential; 72 = Path; 81 = Cycleway; 91 = Footway): Primary (+) Secondary (+)	Posted speed limit (x)	Presence of parking (x) Street lighting (x)	Pavement condition (x) Pavement type (x) Shaded area of road segment (x)	
Lin and Fan (2020)	Bike lane (dummy) (-) Off-street path (dummy) (+) Suggested bike routes (dummy) (-) Road segment length (x) Number of through lanes (x) Road segment slope (x) Signed bike lane (dummy) (x) Suggested bike routed with low comfort (x) Greenway (dummy) (x) Tertiary (+) Residential (+) Path (+) Cycleway (+) Footway (+) Inside shoulder width (x)	Posted speed limit (x)			

Table 2.7: Direct demand models: Correlation of traffic facility characteristics with bicyclist traffic volume (continued)

Author (Year)	Geometric and road classification	Control and flow	Sign and marking	Visibility	Road condition
	Median width (x) Non-motorized facility buffer width (x) Non-motorized facility type (x) Non-motorized facility width (x) Number of lanes (x) Outside shoulder width (x) Surface width (x) RHiNO road classification (x)				
Nelson et al. (2021)					
• Boulder model	Road segment slope (-)	Average speed limit (x)			
• Ottawa model	Road segment slope (-) City suggested route with no bicycling infrastructure (+) Number of traffic lanes (x) Bicycling infrastructure type (x)	Average speed limit (x)			
• Phoenix model		Average speed limit (-)			

Table 2.7: Direct demand models: Correlation of traffic facility characteristics with bicyclist traffic volume (continued)

Author (Year)	Geometric and road classification	Control and flow	Sign and marking	Visibility	Road condition
• San Francisco model	Presence of protected/off-street bike lane (+) Presence of painted bike lane with no physical protection (+) Presence of signed bike route (+) Road segment slope (-) Other number of traffic lanes (x) Road classification (x)	Average speed limit (-)			
• Greater Victoria model Presence of protected/off-street bike lane (-) Presence of signed bike route (+) Road segment slope (+) Other number of traffic lanes (x)	Average speed limit (-)				

The results of socio-demographics were summarized in Table 2.8. Socio-demographics include income, race, gender, age, education, and other variables. Similar to built environment variables, sociodemographic characteristics are measured for the spatial units in which the count locations are located. They were examined in nine studies. We summarized several observations from these correlations.

- Income variables include medium household income and number of households in certain income groups for spatial units surrounding the count locations. Six studies considered income in their direct demand models. Generally, studies found that median household income is negatively correlated with bicyclist traffic (Roll, 2018) [24]. One exception is that Dadashova et al. (2020) found that number of households with income larger than 200K dollars is positively correlated with bicyclist traffic.

- Race variables were considered in three studies, in each case measured by the percentage of white population. These studies showed mixed correlations between percentage of white population and bicyclist traffic. Nelson et al. (2021) found that percentage of white population is negatively correlated with bicyclist traffic in the Boulder model. However, both Roll (2018) and Kwigizile et al. (2019) showed that percentage of white population is positively correlated with bicyclist traffic.

- Gender variables were included in four studies as the percentage of male or female population. These studies showed mixed results about the correlations between gender variables and bicyclist traffic. For example, Nelson et al. (2021) found that percentage of female population is negatively correlated with bicyclist traffic in the Boulder and Greater Victoria models. However, they also found that this relationship is positive in the San Francisco model. Kwigizile et al. (2019) also found a similar positive relationship in their study.

- Age variables were considered in five studies. These studies mainly applied median age in their models. Nelson et al. (2021) found that median age is negatively correlated with bicyclist traffic in the San Francisco model. However, they also found this relationship is positive in the Greater Victoria model.

- Education variables were examined in three studies. Only one study found significant correlations. Nelson et al. (2021) found a positive relationship between percentage of population with at least high school degree and bicyclist traffic in the San Francisco model. However, they also found that this relationship is negative in the Greater Victoria model.

- Other sociodemographic variables were considered in five studies. One most explored variable is percentage of bicyclists. Kwigizile et al. (2019) found that percentage of bicyclists/works is positively correlated with bicyclist traffic. Nelson et al. (2021) also found a similar relationship, in which percentage of male bicyclists is positively correlated with bicyclist traffic in the Ottawa and San Francisco models. However, Nelson et al. (2021) also found that this relationship is negative in the Greater Victoria model.

Table 2.8: Direct demand models: Correlation of socio-demographics with bicyclist traffic volume

Strauss et al. (2015)						
Jestico et al. (2016)						
Sanders et al. (2017)	Medium household income (x)					
Lißner et al. (2018)						
Roll (2018)					Student accessibility (NP)	
Roll (2018)	Median household income (-)	Percentage of white population (+)		Median age (x)	Percentage of high school educated population (x)	Percentage of veterans (x) Percentage of population who commute to work with bicyclists (x)
Kwigizile et al. (2019)		Percentage of white population (+)	Percentage of male population (-)			Percentage of bikers/workers (+)
Saad et al. (2019)						

Table 2.8: Direct demand models: Correlation of socio-demographics with bicyclist traffic volume (continued)

Author (Year)	Income	Race	Gender	Age	Education	Other
Dadashova and Griffin (2020)	Number of households in income groups (x)		Total female population (x) Total male population (x)	Female population in age groups (x) Male population in age groups (x)		Household density (x) Total number of household (x)
Dadashova et al. (2020)	Number of households with >200K income (+) Number of households in other income groups (x)		Total female population (x) Total male population (x)	Female population in age groups (x) Male population in age groups (x)		Household density (x) Total number of household (x)
Lin and Fan (2020)	Median household income (x)			Median age (x)		

Table 2.8: Direct demand models: Correlation of socio-demographics with bicyclist traffic volume (continued)

Author (Year)	Income	Race	Gender	Age	Education	Other
Nelson et al. (2021)						
• Boulder model	Median household income (x)	Percentage of white population (-)	Percentage of female population (-)	Median age (x)	Percentage of population with at least high school degree (x) Percentage of population with at least college degree (x)	Crime density near bike count location (+) Percentage of veterans (x)
• Ottawa model	Median household income (-)	Percentage of white population (x)	Percentage of female population (x)	Median age (x)	Percentage of population with at least high school degree (x) Percentage of population with at least college degree (x)	Percentage of male bikers (+) Unemployment rate (-) Number of pedestrians at intersection (+)

Table 2.8: Direct demand models: Correlation of socio-demographics with bicyclist traffic volume (continued)

Author (Year)	Income	Race	Gender	Age	Education	Other
• Phoenix model	Median household income (-)	Percentage of white population (x)	Percentage of female population (x)	Median age (x)	Percentage of population with at least high school degree (x) Percentage of population with at least college degree (x)	
• San Francisco model	Median household income (-)	Percentage of white population (x)	Percentage of female population (+)	Median age (-)	Percentage of population with at least high school degree (+) Percentage of population with at least college degree (x)	Percentage of veterans (+) Percentage of male bikers (+)

Table 2.8: Direct demand models: Correlation of socio-demographics with bicyclist traffic volume (continued)

Author (Year)	Income	Race	Gender	Age	Education	Other
• Greater Victoria model	Median household income (-)	Percentage of white population (+)	Percentage of female population (-)	Median age (+)	Percentage of population with at least high school degree (-) Percentage of population with at least college degree (x)	Percentage of male bikers (-)

Time variables were measured in hour, day, month, and year, which indicate when the corresponding count was collected. They were included in four studies that modeled bicyclist count during for periods less than one year (e.g., hour or day). Kwigizile et al. (2019) found that, compared to morning hours (12am–5:59am), other times in the day have higher bicyclist traffic. Lin and Fan (2020) showed a similar result. In addition to this, Lin and Fan (2020) found that weekdays have lower bicyclist traffic than weekends. Jestico et al. (2016) found that, relative to January, May, July, and October have higher bicyclist traffic in Victoria, Canada. One study considered year variable but no significant result was found (Dadashova and Griffin, 2020).

Weather variables, which indicate the weather condition when the count was collected, were considered in three studies. All these studies found significant relationships. Generally, better weather condition is correlated with higher bicyclist traffic. Kwigizile et al. (2019) found that relative humidity is negatively correlated with bicyclist traffic. Specifically, 1 percentage point increase in humidity is associated with 1.1% decrease of bicyclist traffic. Dadashova and Griffin (2020) tested three weather variables in their models. They found that temperature is positively correlated with bicyclist traffic while wind speed and precipitation are both negatively correlated with bicyclist traffic (related elasticities not provided). Nelson et al. (2021) found that counts conducted in the cold winter months are 52% smaller than those in other months.

Other variables are mainly related to traffic exposure, safety and geography indicator, which were considered in five studies. Nelson et al. (2021) found that AADT is negatively correlated with bicyclist traffic in the Ottawa model. Saad et al. (2019) found a similar result showing that ratio between vehicular traffic and Strava user count is negatively correlated with bicyclist traffic. However, Saad et al. (2019) also found that total vehicular traffic entering the intersection is positively correlated with bicyclist traffic. Nelson et al. (2021) found that bicyclist crash density is positively correlated with bicyclist traffic in the Boulder, Phoenix, and San Francisco models.

In summary, variables generated from crowdsourced data have significant and positive relationships with bicyclist traffic. Also, crowdsourced data can improve the prediction performance of the direct demand models. Besides crowdsourced data, some variable from the other types were found to have important relationships with bicyclist traffic as well. Those variables include destination accessibility of built environment characteristics, geometric and road classification and control and flow of traffic facility characteristics, income and race of socio-demographics, time, and weather. However, some variables that originally appeared to be significantly correlated with bicyclist traffic, such as total population and population

Table 2.9: Direct demand models: Correlation of time with bicyclist traffic volume

Author (Year)	Hour	Day	Month	Year
Strauss et al. (2015)				
Jestico et al. (2016)			Month (Reference: January): May (+) July (+) October (+) Other month indicators (x)	
Sanders et al. (2017)				
Lißner et al. (2018)				
Roll (2018)				
Roll (2018)				
Kwigizile et al. (2019)	Time of a day (Reference: 12am–5:59am): 6am–9:59am (+) 10am–2:59pm (+) 3pm–7:59pm (+) 8pm–11:59pm (+)			
Saad et al. (2019)				
Dadashova and Griffin (2020)		Day of week (Considered as random effect, NP)	Month of year (Considered as random effect, NP)	Year (x)
Dadashova et al. (2020)				
Lin and Fan (2020)	Time of day (Reference: 0:00–05:59) Hour (06:00-–08:59) (+) Hour (09:00-–14:59) (+) Hour (15:00-–17:59) (+) Hour (18:00-–19:59) (+) Hour (20:00-–23:59) (+)	Weekday (-)		
Nelson et al. (2021)				

Table 2.10: Direct demand models: Correlation of weather with bicyclist traffic volume

Author (Year)	Weather
Strauss et al. (2015)	
Jestico et al. (2016)	
Sanders et al. (2017)	
Lißner et al. (2018)	
Roll (2018)	
Roll (2018)	
Kwigizile et al. (2019)	Relative Humidity (%) (-)
Saad et al. (2019)	
Dadashova and Griffin (2020)	Temperature (F) (+) Wind speed (ml/hour) (-) Precipitation (inch) (-)
Dadashova et al. (2020)	
Lin and Fan (2020)	
Nelson et al. (2021)	
• Boulder model	
• Ottawa model	Counts conducted in winter months (dummy) (-)
• Phoenix model	
• San Francisco model	
• Greater Victoria model	

Table 2.11: Direct demand models: Correlation of other variables with bicyclist traffic volume

Author (Year)	Other variables
Strauss et al. (2015)	
Jestico et al. (2016)	
Sanders et al. (2017)	
Lißner et al. (2018)	
Roll (2018)	
Roll (2018)	AADT (x) Kwigizile et al. (2019)
Saad et al. (2019)	Total entering vehicular count (+) Ratio between vehicular volume and Strava user count (-)
Dadashova and Griffin (2020)	Average daily traffic (x) City indictor (x) Count type (Temporary or permanent) (x)
Dadashova et al. (2020)	City indicator (x)
Lin and Fan (2020)	
Nelson et al. (2021)	
• Boulder model	Bicyclist crash density (+) AADT (x)
• Ottawa model	AADT (-) Bicycling comfort and safety level (High or Medium) (+) Percentage of trucks crossing intersections (-)
• Phoenix model	Bicyclist crash density (+) AADT (x)
• San Francisco model	Bicyclist crash density (+)
• Greater Victoria model	

density, were mostly found to have no significant relationships after considering crowdsourced mobile data. This result indicates that crowdsourced mobile data may absorb the effect of some independent variables in the direct demand models as they are correlated with each other.

2.4.3 Other Methods

Besides the direct demand model, two methods were applied in the literature: aggregation analysis and Strava user rate expansion. Each method was applied in one study, respectively.

Aggregation analysis was used by Nishi et al. (2014) to compute the hourly pedestrian volume with smartphone application location-based service data. These data, which are generated by the Yahoo smartphone applications in Japan, include the coordinates, time, user ID, and accuracy of a specific location (Figure 2.2). Nishi et al. (2014) aggregated these locations into hourly pedestrian volume based on several steps.

1. Define the traffic analysis zones which are used to aggregate location data

2. Aggregate the location data into each traffic analysis zone and hour

Table 1: Examples of the location data

Latitude	Longitude	Accuracy	Time stamp	ID
34.8716480098989	135.661309193946	5.00	20130611000000	UID1
38.7213293603050	139.849629867952	30.00	20130611000000	UID2
43.0928880757489	141.371409950081	112.00	20130611000001	UID3
35.5574600559872	139.445981733805	14.88	20130611000000	UID4
35.7329128494678	139.670771645082	65.00	20130611000012	UID5
35.7846891648607	139.899813949837	10.53	20130611000001	UID6
35.6521350751540	140.026954640171	1414.00	20130611000002	UID7
35.6998792435734	139.841407947290	165.00	20130611000002	UID8

Figure 2.2: Examples of the smartphone application location-based service data (Nishi et al., 2014)

3. Remove static population based on the speed calculated from location and time

4. Use temporal and spatial distribution to impute missing data

Strava user rate expansion was proposed by Camacho-Torregrosa et al. (2021) to compute bicyclist traffic. Strava user rate indicates the ratio between Strava bicyclist count and observed bicyclist count in equation (2.2). The authors assumed that Strava user rate is consistent across time for each location. Accordingly, they applied Strava user rate to compute the bicyclist traffic based on several steps.

1.

2. Select several locations to collect actual bicyclist count

3. Calculate Strava user rate with Strava bicyclist count and observed bicyclist count with equation (2.2)

4. Apply Strava user rate to other locations

5. Use Strava user rate and Strava count to compute actual count for other locations with equation (2.3)

$$Stravauserrate = V_{Strava}/V_{Observed} \tag{2.2}$$

$$Bicyclecount = V_{Strava}/(Stravauserrate) \tag{2.3}$$

2.5 Conclusions

In this study, we applied a systematic literature review approach to explore the types of crowdsourced mobile data and methods that have been used to estimate the pedestrian and bicyclist traffic volume. After

searching the scholarly databases with the keywords related to our topics and screening the search results with several criteria in two rounds, we included 14 studies in our final review. We then extracted key information related to our topics from the selected studies.

Three distinct types of crowdsourced mobile data, including fitness application data, application location-based service data, and cellular signal data have been used in traffic volume estimation. Compared with traditional observed count data, crowdsourced mobile data have larger coverage of geospatial and temporal range and lower cost data collection. However, they are biased toward a certain population based on their mobile data sources. For example, fitness application data are biased toward runners and cyclists as they are collected from smartphone fitness applications such as Strava and Mon ResoVelo. In addition, cellular signal data have relatively lower spatial precision.

Researchers have applied three different methods to use crowdsourced mobile data to estimate pedestrian and bicyclist volumes. These methods include direct demand modeling, aggregation analysis, and Strava user rate expansion. The direct demand modeling technique is the frequently used approach in the literature. Our review results showed that crowdsourced mobile data not only have significant correlations with bicyclist count but also improve the model prediction performance. One study found that machine learning models have better performance than traditional statistical models. In addition to crowdsourced data, we found that destination accessibility of the built environment characteristics, geometric and road classification and control and flow of the traffic facility characteristics, income and race of the socio-demographics, time, and weather have significant correlations with bicyclist traffic.

Strava data have been used more than other data sources in the studies analyzed in this review. Strava, an activity tracking application for fitness-related activities, collects GPS trajectories of its users during their active travel activities and provides aggregated pedestrian and bicyclist traffic for each road segment, intersection, or OD pair. Pedestrian or bicyclist user counts provided by Strava have been shown to be highly correlated with observed counts (Lee and Sener, 2021). However, they are biased toward the users of fitness applications, such as runners and cyclists, who are not representative of the population as a whole. Many studies have applied Strava data in their direct demand models. As direct demand models usually use road segment or intersection as analysis unit, Strava data provide convenience as they aggregate data to these locations. In addition, as Strava data are biased toward certain populations, direct demand models can include other types of variables, such as built environment attributes, traffic facility characteristics, socio-demographics, time, and weather, to correct this bias.

StreetLight is another pedestrian and bicyclist traffic data provider. Although researchers have not published the models incorporated within StreetLight in the peer-reviewed literature, StreetLight estimates of pedestrian and bicyclist traffic are being used in planning and engineering applications by public agencies. StreetLight data mainly use two types of crowdsourced data, application location-based service data and cellular signal data, to generate pedestrian and bicyclist counts. One study indicates that StreetLight bicyclist count is highly correlated with observed count (Turner et al., 2020). Similar to Strava data, StreetLight input data are biased toward certain populations, such as cell phone users. StreetLight data provide StreetLight pedestrian and bicyclist index through traffic analysis zones or of each OD pair between different traffic analysis zones. Therefore, they are potentially useful when being applied in four-step models and methods that need traffic OD matrix as input.

3 Methodology and Implementation

3.1 Introduction

This chapter describes the methodology and implementation details of the trip prediction model and filtering algorithm. The trip prediction model consists of two main components: trip distribution model predicting non-motorized trips (both pedestrian and bicyclists), and a network route choice model that converts the origin-destination trips into link traffic flows. To generate the initial trip table or the origin-destination (OD) matrix we will use the gravity model with matrix rebalancing. Then using that OD matrix as the initial solution based on the partial link flow data, we will keep improving the OD matrix using an algorithm, details of which will be discussed later. We will also make use of the logit route choice model in the algorithm after finishing improving the OD matrix to generate an estimate of the link flows throughout the network. The trip distribution, route choice model and the algorithm to improve traffic flow estimation will be discussed in detail in the Methodology section. Implementation details will be provided in the Implementation section.

3.2 Methodology

3.2.1 Network Terminology

Any point where travellers may want to start or end their trips is called a node. A connection between two nodes is called an edge or a link. A directed graph or network is just a collection of nodes and edges with directions. We are going to model both the bicyclist and pedestrian networks as directed graphs. We also have access to the distance information from one node to another. In other words, we have access to the length information of each edge in the network in the unit of distance from the OpenStreetMaps dataset. This directed graph representation of networks will allow us to perform important operations and

algorithms to obtain the link flow esimates. These terminologies will be used throughout the report.

3.2.2 Notations and Definitions

First, we introduce the relevant notation and definitions required to present the methodology. Note that a brief summary of each methodological component is presented at the start of each of the following subsections, and a more detailed summary is presented next. Thus, the interested reader can understand the full details required to understand the methodology. However, the full methodological details are only provided for completeness and are not needed to use the presented results.

Define a directed graph $\mathcal{G} = (\mathcal{N}, \mathcal{E})$ where the sets of nodes and edges/links are represented by $\mathcal{N}$ and $\mathcal{E}$ respectively. Let $\mathcal{Z} \subseteq \mathcal{N}$ be the set of special zone nodes through which demands enter and exit the network. From the mobile sources dataset, we have access to the number of trips produced from and attracted to different zones. Let the number of trips produced from zone $r \in \mathcal{Z}$ be P_r and trips attracted to zone $s \in \mathcal{Z}$ be A_s. In this paper, we use the StreetLight data to get the number of trips produced from and attracted to different zones for both pedestrians and bicyclists. After we have the number of trips produced from and attracted to different zones, we can use Fratar (1954)'s method to find an initial OD matrix.

Trip Distribution (Fratar, 1954) using $\mathcal{G}$, P $= \{P_r \, \forall r \in \mathcal{Z}\}$, A $= \{A_s \, \forall s \in \mathcal{Z}\}$ and T $= \left\{T_{rs} \, \forall (r,s) \in \mathcal{Z}^2\right.$ would give the initial OD matrix $\tilde{\text{d}}$. Here, T_{rs} is the distance of the shortest path from node r to node s. Denote the initial OD matrix by $\tilde{d} = \left\{\left\{\tilde{d}_{rs} \, \forall s \in \mathcal{Z}\right\} \, \forall r \in \mathcal{Z}\right\}$.

As part of the process, we tried to estiamte the initial OD matrix so that it was consistent with the partially observable observed link flows. We assumed that the traffic count was available for only links $\bar{\mathcal{E}} \subseteq \mathcal{E}$. In other words, link flows were only available for some of the links and not all links of the network. We added this assumption to our model because none of the mobile sources datasets we found looked at provide link counts in all links. For some of the links, count data were missing. This could be due to many reasons including travelers losing internet connection, turning off GPS connection, the data was not stored for those links etc. We represented the partially observable link flows by $\tilde{\text{x}} = \left\{\tilde{x}_{ij} \, \forall (i,j) \in \tilde{\mathcal{E}}\right\}$ where $\tilde{\mathcal{E}} \subseteq \mathcal{E}$ is again the set of links where we have link flow data. For this work, we used the StravaMetro dataset to get the partial link flow data for both pedestrians and bicyclists.

To make any improvement to the estimated OD matrix, we first had to determine how consistent it is with the partially observable link flows. We needed an efficient route choice model to obtain link flows from a given OD matrix. Once we had a method to do that we were able to measure an OD matrix's con-

sistency with the partially observable link flows by looking at the differences between the link flows. We then minimized these differences between the link flows obtained from the OD matrix and the partially observable link flows. We also made sure that the OD matrix was still consistent with the production and attraction numbers.

During the process, we continually updated the OD matrix until the link flows obtained from the updated from the updated OD matrix matched the partially observed link flows or it could not be improved any further. After we had the final OD matrix used it to obtain the final link flows that are consistent with the partially observed link flows for all links and for which flow conservation holds. More details are discussed about each of these steps in the next sections.

3.2.3 Available Data

The data sources used for this project were selected based on a review of data sources commonly used in the literature (See Chapter 2 and Table 2.2 for a comprehensive comparison). Additionally, emphasis was placed on using data that MnDOT had readily available. For example, StreetLight and StravaMetro were two data sources that MnDOT had access to at the time that this work was completed, and are commonly used data sources in the literature. Data was used for 2021 and 2022 when conducting the presented analysis, though more recent data is now available and the presented method can be reapplied to new data as it becomes available.

The OpenStreetMaps dataset contained detailed information including geometry information about different paths both pedestrian and bicyclist. While bicyclists often need to use the same roads as motorized vehicles, pedestrians usually have more freedom to move because of the presence of sidewalks, crosswalks, and other unmarked paths. To capture this difference, OpenStreetMaps provided different networks for pedestrians and bicyclists. From StreetLight data, we had access to the number of trips produced and attracted to these different regions. Another type of data that we had access to is the partial link flow data. After looking at different mobile sources datasets, we found that none of them provide link counts in all the links of the network. The StravaMetro dataset provided link counts for some of the links in the network while the other links counts are missing. We used this partially observable link flow data along with the production and attraction numbers and the network structure to estimate the link flows in all the links of the network. The next subsections will present detailed methodology that uses all these mobile data sources to estimate link flows while maintaining consistency between aspects of traffic flow.

3.2.4 Route Choice Model

Once we had an OD matrix, the simple approach was to assign OD trips to the shortest paths in terms of travel time or some other cost. However, that approach had the following issues.

1. Multiple paths with equal or very similar travel times or costs.
2. Relying solely on travel time or some arbitrary cost to decide a path is not realistic because a traveller's decision to choose a path can be influenced by many different factors including familiarity, personal preference, convenience or infrastructure etc.
3. Perfect information about the travel times or costs at different links are not available to the travellers consistently.
4. Since its not guaranteed that people will always take the shortest path, all paths that can be used need to be identified.

Therefore, we used the most likely path flow formulation instead where the probability of a path to have some flow is maximized. One complexity that arose from this approach is that requiring totally acyclic paths is a strong assumption and in reality pedestrians' and bicyclists' probability of using any path including the cyclic paths are non-zero. If we had excluded all cyclic paths from the set of all possible paths then that results in excluding all such cyclic trips. This could have resulted in undercounts in the link flow estimates. Alternately, considering all cyclic paths in the set of all possible paths people could take included these low probability paths as well. Another reason for considering all possible paths including the cyclic paths was that routing of travelers can be done without explicit path enumeration details of which will be discussed later. This allowed us to solve the trip assignment problem more efficiently.

matrix.

Most Likely Path Flows to Logit Assignment

? showed that when all possible paths including the cyclic paths are considered and included in the reasonable paths set Π, then the segment substitution property holds. Using that property they also showed that the link flow x_{ij} for all link (i, j) in $\mathcal{E}$ could be calculated as follows:

$$x_{ij} = \sum_{(r,s)\in\mathcal{Z}^2} d_{rs} \frac{V_{ri}\left[\exp(-\theta t_{ij})\right] V_{js}}{V_{rs}} = \sum_{(r,s)\in\mathcal{Z}^2} d_{rs} \frac{V_{ri} L_{ij} V_{js}}{V_{rs}} \tag{3.1}$$

Let L be the adjacency matrix for the weighted graph $\mathcal{G}$ where for link (i,j) the weight or cost is $L_{ij} = \exp(-\theta t_{ij})$. Replacing t_{ij} with $\psi_i^r + t_{ij} - \psi_j^r$ to avoid taking the exponential of large number and to avoid numerical issues (?),

$$L_{ij}^r = \exp(-\theta(\psi_i^r + t_{ij} - \psi_j^r)) = \exp(\theta(\psi_j^r - \psi_i^r - t_{ij})) \tag{3.2}$$

where ψ_i^r was the cost-to-go label for node i set using Dijkstra's shortest path algorithm for one to all label setting. Note that a superscript r was added to L_{ij} to denote the source node for Dijkstra's algorithm. Since we replaced L_{ij} with L_{ij}^r to capture the effect of different sources in the weighted adjacency matrix, the V_{ij} values were different for different sources. This is because the the V_{ij} values depend on the L_{ij} values. To capture the effects of node labels using Dijkstra's algorithm from different sources we used superscripts to denote the source for both L_{ij} and V_{ij} values and rewrite eq. (3.1) as the following.

$$x_{ij} = \sum_{r\in\mathcal{Z}} x_{ij}^r(d) = \sum_{r\in\mathcal{Z}} \sum_{s\in\mathcal{Z}} d_{rs} \frac{V_{ri}^r L_{ij}^r V_{js}^r}{V_{rs}^r} \tag{3.3}$$

By the Chapman-Kolmogorov equations the n^{th} power of an weighted adjacency matrix gives the weight of walks or segments of length n. So, for all $r \in \mathcal{Z}$ origins the V^r's can be written as,

$$V^r = \sum_{n=1}^{\infty} (L^r)^n = L^r + L^r \sum_{n=1}^{\infty} (L^r)^n = L^r + L^r V^r \tag{3.4}$$

$$\implies V^r = L^r (I - L^r)^{-1} = [[V_{ab}^r \; \forall b \in \mathcal{N}] \; \forall a \in \mathcal{N}] \tag{3.5}$$

Here, V_{ab}^r gave us the total cost of using all paths from node a to b and the superscript r means that the value of the cost was calculated using the Dijkstra labels when the source node was r. Since we knew V^r's

and L^r's, x_{ij} could be written as a function of d,

$$x_{ij}(\mathsf{d}) = \sum_{(r,s)\in\mathcal{Z}^2} d_{rs} \frac{V^r_{ri} L^r_{ij} V^r_{js}}{V^r_{rs}} \tag{3.6}$$

Equation (3.6) allowed us to calculate the link flows for a given OD matrix efficiently. On the numerator, V^r_{ri} represents the total cost of taking any path from node r to node i and V^r_{js} represents the total cost of taking any path from node j to node s. L^r_{ij} represents the cost of taking link (i, j) when the source node was r. On the denominator, V^r_{rs} represented the total cost of taking any path from node r to node s. Therefore, the ratio $\frac{V^r_{ri} L^r_{ij} V^r_{js}}{V^r_{rs}}$ represented the probability of taking link (i, j) when the journey begins at node r. This equation also showed how to efficiently perform logit loading.

As we have a method to calculate the link flows from an OD matrix, we can now formulate an optimization problem to find an OD matrix that is consistent with production and attraction numbers and produces link flows that differs minimally from all the observed link flows. The next section will discuss how to do that.

3.2.5 Formulation

The data described in this section was available for both bicyclists and pedestrians. We had the network topology $\mathcal{G}$, the productions P and attractions A numbers for different zones from mobile sources data. We also had partial link counts $\tilde{x}_{ij}$ for some links $\tilde{\mathcal{E}} \subset \mathcal{E}$ from mobile sources data. However, we did not have link flows for all of the links. We also did not know how consistent the available link flows from the mobile data sources were with the available productions and attractions numbers. Therefore, we formulated an optimization problem to first estimate an OD matrix that satisfied the production and attraction consistencies. To ensure that the estimated OD matrix did not result in link flows that are too different from the partially observable link flows, we minimized the sum of squared errors between the estimated and partially observable link flows in the objective function. We estimated the link flows given a feasible OD matrix using the logit loading equation eq. (3.6) during the solution process of this optimization problem.

We formulated the estimation problem as the following non-linear program.

$$\min \quad f(\mathsf{d}) = \frac{1}{2} \sum_{(i,j)\in\tilde{\mathcal{E}}} \left(\sum_{(r,s)\in\mathcal{Z}^2} d_{rs} \frac{V^r_{ri} L^r_{ij} V^r_{js}}{V^r_{rs}} - \tilde{x}_{ij} \right)^2 \tag{3.7a}$$

$$P_r = \sum_{s\in\mathcal{Z}} d_{rs} \qquad \forall r \in \mathcal{Z} \tag{3.7b}$$

$$A_s = \sum_{r\in\mathcal{Z}} d_{rs} \qquad \forall s \in \mathcal{Z} \tag{3.7c}$$

$$d_{rs} \geq 0 \qquad \forall (r,s) \in \mathcal{Z}^2 \tag{3.7d}$$

Equation (3.7a) minimized the sum of squared errors between the partially observable link flows and the link flows from the OD matrix. Equations (3.7b) and (3.7c) ensured production and attraction consistency and eq. (3.7d) ensured that the OD matrix entries were non-negative. The decision variable of this optimization problem was the OD matrix.

Here, the constraints given in eqs. (3.7b) to (3.7d) are linear. Therefore, the feasible set $\mathcal{D}$ of the optimization program was convex where $\mathcal{D}$ is defined as

$$\mathcal{D} = \left\{ \mathsf{d} : \sum_{s\in\mathcal{Z}} d_{rs} = P_r \; \forall r \in \mathcal{Z}, \sum_{r\in\mathcal{Z}} d_{rs} = A_s \; \forall s \in \mathcal{Z}, d_{rs} \geq 0 \right\} \tag{3.8}$$

To solve this optimization problem using a solver, we first needed to calculate and store all the $\frac{V^r_{ri} L^r_{ij} V^r_{js}}{V^r_{rs}}$ values for all $(r,s,i,j) \in \mathcal{Z}^2 \times \mathcal{N}^2$. Therefore, the space complexity of storing these values would have been $\mathcal{O}(\mathcal{Z}^2\mathcal{N}^2)$. However, if we calculated these values as needed then we would only need to store L^r and V^r matrices and the space complexity would then be $\mathcal{O}(\mathcal{Z}\mathcal{N}^2)$. These values serve as coefficients associated with the different elements of the OD matrix. Therefore, calculating all these values individually could not have been avoided if we wanted to use a solver to solve this optimization problem.

However, to evaluate the objective function with a solution, calculations of these individual coefficients was not required. First, we represented the link flow estimates as a weighted adjacency matrix $x(\mathsf{d})$ where the elements represent the link flow values for different links. Using the eq. (3.6) and any solution OD matrix d, we could calculate these values for all the sources $r \in \mathcal{Z}$ in parallel and then take the matrix sum to get the final link flow estimates $x(\mathsf{d})$. Then we could create a similar matrix $\tilde{x}$ for the partially observable link

flows and take the matrix difference. Then replacing the elements of that matrix where the link flows were not observed with zeros and taking the Frobenius norm of that matrix, then squaring and multiplying the result with $1/2$ would have given us the final objective value. Since most of the operations were done on matrices, we could easily parallelize this whole process. These matrix operations also meant that graphics processing unit (GPU) acceleration could be used to speed up the calculation process of the objective value. Considering all these advantages, we decided to solve this optimization problem with a custom algorithm instead of a solver.

3.2.6 Solution algorithm

To solve the optimization problem, we first initialized the OD matrix d^0. Then using an appropriate step size and the gradient of the objective function we found the OD matrix $\mathrm{d}^k_{\mathrm{gd}}$ using the gradient descent method. This update may cause the OD matrix $\mathrm{d}^k_{\mathrm{gd}}$ to leave the feasible set $\mathcal{D}$. Therefore, we needed ot project the OD matrix $\mathrm{d}^k_{\mathrm{gd}}$ back onto the feasible set $\mathcal{D}$ to obtain a feasible OD matrix d^k. We keept repeating this process until the OD matrix converged.

Obtaining final link flows

After, convergance, we calculate the link flows x' using eq. (3.6) and the final OD matrix. We also defined the feasible link flows as the link flows that did not violate the flow conservation and were not less than observed link flows. After calculating the link flows from the OD matrix we project the link flows onto this feasible region. If the link flows were in the feasible region, then this projection operation would not change the link flows. Otherwise, the projection operation ensured a valid of link flows. In the previous OD matrix estimation problem we did not need the flow conservation since logit loading ensured flow conservation there. However, that optimization problem could not ensure that the link flows are not less than the partially observable link flows. That part of the optimization problem was solved here by projecting the link flows onto the feasible link flow region.

To project the link flows onto the feasible link flow region, we solved the following optimization problem.

$$\underset{x}{\text{minimize}} \quad \sum_{(i,j)\in\mathcal{E}} (x_{ij} - x'_{ij})^2 \tag{3.9a}$$

$$\text{subject to} \quad \sum_{j\in\mathcal{N}} x_{ij} - \sum_{h\in\mathcal{N}} x_{hi} = \sum_{s\in\mathcal{N}} d_{is} - \sum_{r\in\mathcal{N}} d_{ri} \qquad \forall i \in \mathcal{N} \tag{3.9b}$$

$$x_{ij} \geq \tilde{x}_{ij} \qquad \forall (i,j) \in \mathcal{E} \tag{3.9c}$$

Here, x′ was the link flow vector obtained from the OD matrix using eq. (3.6). The objective function was the sum of squared differences between the link flows from the OD matrix and the link flows that are the decision variables of this problem. Equation (3.9b) ensured inflow and outflow conservation at each node. Equation (3.9c) ensured that the link flows were not less than the partially observable link flows. Equation (3.9c) was added because partially observable link flows are often undercounts because not everyone used the same services. For example, we used StravaMetro data for the partially observable link flows. However, not everyone used Strava's services. Therefore, trips made by people who were not Strava users would not show up in the partially observable link flows. However, these counts worked well as lower bounds since we knew that at least the recorded number of trips in Strava's dataset were observed. The solution to this problem were the link flows that were the closest to the link flows from the OD matrix and were feasible. The constraints were linear and so the feasible link flow region was convex and the objective function was also convex. Therefore, this was a convex problem which could be easily solved using a convex optimization solver. Finally, all links are sorted by length and estimated flow and only the longest link for each segment is selected to eliminate problems with duplicate links in the dataset.

The solution to this problem was the final estimation of link flows. Some properties about the final link flows are as follows.

1. The final link flows were the closest to the link flows from logit loading using the final OD matrix.
2. The final link flows were all greater than or equal to the partially observable link flows which were historically undercounted.
3. The final link flows satisfied the flow conservation at each node and were consistent with the final OD matrix.

3.3 Implementation

The entire solution was implemented in Python. This implementation made use of several third party libraries. `NumPy` was used for tensor manipulation, `tqdm` was used to visualize the progress of the solution algorithm. All other libraries that were used are part of Python's standard library. The implementation of the solution is divided into several modules. The next subsection discusses each of those modules.

3.3.1 Modules

Network Data

The module `net_data` was written to extract network information from the TNTP network data format. More details about the TNTP data format can be found here. That information is then used to build the network using the next module named "network".

Network

The module `network` contains the `Network` class which stores information about the network like the set of nodes, set of links and associated weights. Several operations on the network were implemented in this module that were used in the final solution algorithm to estimate traffic flows. The operations include adding and removing nodes and edges from the network, obtaining the incoming or outgoing links and shortest path object which implements algorithms related to the shortest paths between nodes in the network. More details can be accessed by using Python's built-in function "help" with the `Network` class as an argument after importing the "network" module in the Python interpreter.

```
>>> from network import *
>>> help(Network)
     ...
```

Shortest Path

The `shortest_paths` module implemented Dijkstra's label setting algorithm for shortest path from one source to all nodes. Other methods that were implemented include methods to find shortest path using the Dijkstra's label setting algorithm, to inspect the cost or disutility of using a path based on the provided costs of links, k-shortest paths algorithm etc. This primary use of this module in the traffic flow estimation solution algorithm was to calculate the node labels from different sources. More details can be accessed

by using Python's built-in function "help" with the `ShortestPath` class as an argument after importing the "shortest_paths" module in the Python interpreter.

```
>>> from shortest_paths import *
>>> help(ShortestPath)
     ...
```

Trip Distribution

The `trip_distribution` module implemented the trip distribution using gravity model with matrix rebalancing. The `TripDistribution` class in that module takes a `Network` object, the production and attraction data as inputs. Then using the ??, it generated the trip table or the OD matrix. More details can be accessed by doing the following.

```
>>> from trip_distribution import *
>>> help(TripDistribution)
     ...
```

Logit Loading

The `network_loader` module was used to assign link flows to different links using logit loading given an OD matrix. The `LogitLoader` class implemented the procedure to create the label L-matrices for different sources. This class also implemented the procedure to create the V-matrices for different sources. These matrices were created using the node labels obtained from Dijkstra's label setting algorithm with different sources. Once these matrices were available, for a given OD matrix using eq. (3.6) the link flows could be estimated. More details can be accessed by doing the following.

```
>>> from network_loader import *
>>> help(LogitLoader)
     ...
```

Flow Estimation

The `link_flow_estimator_tensor_gpu` module implemented the link flow estimator solution algorithm in the class `LinkFlowEstimator`. The inputs to an instance of this solver class were the `LogitLoader` and the `TripDistribution` objects created with the network, trip information. Details can be accessed by doing the following.

```
>>> from link_flow_estimator_tensor_gpu import *
>>> help(LinkFlowEstimator)
     ...
```

3.4 Usage

3.4.1 Requirements

The version of Python used for the Implementation was “3.9.15”. The following third party libraries were used in the implementation of the solution algorithm.

1. NumPy: Version: “1.23.3”
2. PyTorch: Version: “1.13.0” with GPU support.
3. tqdm: Version: “4.64.1”

All of these libraries are open-source and free to use. The implementation of the solution assumes that a CUDA-enabled GPU is available. The links to the websites include a installation guide for the Python interpreter and the libraries mentioned.

3.4.2 Link Flow Estimation

First, a network information needs to be obtained and a `Network` object needs to be created using the “network” module.

```
>>> network_data = GetNetworkData(name=network_name, path=network_path)
>>> links: list[Link] = network_data.links      # Link = tuple[Node, Node]
>>> weights: dict[Link, float] = network_data.net.free_flow_time.to_dict()
>>> network = Network(links, weights)
```

After extracting the production and attraction numbers a `TripDistribution` object can be created and the initial OD matrix can be computed using the ??.

```
>>> # productions: dict[Node, int]
>>> # attractions: dict[Node, int]
>>> td = TripDistribution(productions, attractions, network)
>>> od_matrix: dict[Node, dict[Node, float]] = td.od_matrix()
```

Then logit loader object needs to be created using a value for the θ parameter.

```
>> ll = LogitLoader(
    graph=network,
    od_matrix=od_matrix,
    theta=theta,
    origins=td.origins,
    destination=td.destinations
    )
```

After that the link flow estimation can be done by doing the following.

```
>>> final_od, _ = lf_estimator.steepest_descent(
    init_x=initial_x_partial, epsilon=epsilon, max_iterations=max_iterations
    )
>>> link_flows = lf_estimator.link_flow_adjacency_matrix(final_od)
```

Examples on how to use each of these modules are also available in the source code for the modules after the following line.

```
if __name__ == "__main__":
```

4 Analysis and Visualization

4.1 Introduction

This chapter describes the processing of the Twin Cities Metro Area network data and the modifications made to the network so that it could be used in the developed solution algorithm. After running the algorithm on the networks, the obtained link flows estimates were visualized on a map. The visualization of pedestrian and bicyclist link flows were created using OpenStreetMaps networks. The visualization is done by plotting geometries of the links on a map from OpenStreetMaps and adding associated data to a link as a tooltip. The final visualization can be seen as a webpage. The next section describes the data processing steps, the modifications along with the visualization process and how the visualization can be modified to suit the user's preference.

Processing the Twin Cities network data required multiple operations on the network. Several modifications to the networks were made to make it work with the solution algorithm. We used the Twin Cities area network which contains seven counties: Anoka, Carver, Dakota, Hennepin, Ramsey, Scott, and Washington. This seven county metropolitan area was divided into $3,030$ traffic analysis zones (TAZ) by the Metropolitan Council. The modifications made to the network data are described in the following subsections.

4.2 Data processing

4.2.1 Partial Count Correction

In the available dataset, the partial counts were not available for all days of a year for all links. Therefore, to extract the average daily counts for links where all days of year were not available, we used a filtering approach. First, we created average daily profiles for both bicyclist and pedestrian counts using only the links' counts where all days of year counts were present in the partially available counts dataset. Then, for

the links where more than 50 days counts were missing, we used the average daily profile to first fill the missing data and then used that to compute the average annual daily traffic counts. For links where less than 50 but more than 1 days of counts were missing, we used a Kalman filter using ARMA$(1, 1)$ model to fill the missing data. The ARMA$(1, 1)$ model assumes no observation error which is helpful for our purpose because we do not want to modify the observed data at all. We just use this model to fill in the missing values while keeping the observed data the same. The main idea being, use the ARMA$(1, 1)$ model to fit the available data using a Kalman filter and whenever a missing data is encountered, use the predicted data for the next timestep from Kalman filter and treat that as the observed data. More details about using Kalman filters for general ARMA(p, q) with missing observations can be found in Jones (1980).

4.2.2 Obtaining the network data

We downloaded the entire network data from the OpenStreetMaps (OSM) database using OSMnx (Boeing, 2017) for both the pedestrian and bicyclist networks. Furthermore, OSMnx specifies several network types. Among them the "walk" and "bike" networks contains all streets and paths that pedestrians and cyclists can use respectively. This made obtaining the different types of network straightforward. Using the geometry information (TAZ polygons) we downloaded the $3,030$ smaller networks for both types of network as well. This was done because the entire network was to large to process and use in the solution algorithm in a reasonable amount of time. Therefore, we first divided the network into $3,030$ smaller networks and then we merged the smaller networks to create subnetwork based on neighboring TAZs to perform our analysis.

4.2.3 Modifications to the network

The following description applies for both types of network. From the OSM database we obtained detailed information about coordinates of nodes, geometry information and other properties of links. However, the number of nodes and links were very large in those networks. Networks from OpenStreetMaps are not routable and contains many topological errors (Neis et al., 2012) and OpenStreetMaps nodes are inconsistent (Boeing, 2017). OSMnx can perform a few topological corrections to address these issues. Furthermore, any irregularities in a link's geometry like curves are represented by nodes in OpenStreetMaps networks increasing the total number of nodes and links. Since, those are not graph-theoretic nodes, we use OSMnx to algorithmically remove them and create unified links between the graph-theoretic nodes of the network. This unification of links however resulted in loss of geomtry information for some links. The

new links are also renamed in the corrected network however original node ids are retained for all links. If multiple links were combined then all of those link's OSM ids, names and some other attributes were retained.

4.2.4 Creating subnetworks for analysis

To create subnetworks using which we performed our analysis, we first selected a centering TAZ. We also specified a maximum number of TAZs that can be included in a subnetwork to keep the size of the subnetworks manageable. Then, we found the neighboring TAZs of the centering TAZ. Several steps were taken to achieve this. We describe those steps in the next paragraph.

Selecting neighboring TAZs

Minkowski sum $\oplus$ of two polygons A and B is defined as

$$A \oplus B = \{a + b : a \in A, b \in B\} \tag{4.1}$$

If the polygons are convex and has n and m number of vertices then the Minkowski sum can be done in $\mathcal{O}(n+m)$ time (De Berg, 2000). Minkowski sum of a polygon and a circle inflates the polygon by the radius of that circle. Approximating the circle with a finite number of vertices allow efficient calulation of the inflated polygon. First, we inflated the polygon of the centering TAZ by a very small radius $\epsilon > 0$. After that, we took the intersection of the inflated polygon with all the TAZ polygons which gave us all the neighboring TAZs. A small inflation was sufficient to find the neighboring TAZs which is why the value of ϵ selected to be very small. Then, based on the shortest distance between the centroid of the centering TAZ and the centroid of the neighboring TAZs, we selected the k nearest TAZs where k did not exceed the maximum number of TAZs in a subnetwork. This process gave us the topological neighboring TAZs of the centering TAZ. If the maximum number of TAZs was not exceeded, then we selected another centering TAZ that was a neighbor of the current centering TAZ and repeated the process until the maximum number of TAZs were obtained. Then all of the smaller networks associated with the selected TAZs were merged to create a subnetwork to run our analysis. To merge the smaller networks, we first took the union of the collected polygons and then extracted the subgraph only in that new polygon from the entire network. We used OSMnx to perform this truncation operation on the network to obtain the subnetwork. While performing

this truncation, OSMnx ensured that any edge that has a node inside the subnetwork was not truncated. After obtaining the subnetwork, we had to make further modifications to it.

Node consolidation

To reduce the number of nodes and links further, we also consolidated nodes that were very close to each other. This consolidation process required projecting the the subgraph onto the proper coordinate system and then merging all the nodes that were within 15m of each other. This consolidation of nodes resulted in a significant reduction in the number of total nodes and links in the subnetwork making the solution algorithm more manageable. OSMnx was also used to correct topological errors in the subnetworks by reconnecting the edges and rebuilding the subgraph details of which can be found in Boeing (2017).

Selecting zone nodes

After obtaining the subnetworks, we had to select zone nodes through which the demand would enter and exit the network. We decided to select a maximum of 10 zone nodes for each TAZ. Wherever we noticed a discrepancy between the observed link counts in the adjacent links, we assumed that demand entered or exited through that node. In other words, the node at which flow conservation was violated was selected as a zone node. The reason is that, we assumed that this flow conservation violation would be fixed by the entering or exiting demand through that node if it was selected a zone node. Based on the amount of violation we assigned a weight to each of these zone nodes. If we no flow violations were found then we fell back to the Pagerank algorithm (Brin and Page, 1998) to select important nodes through which demand would enter and exit the network. In that case, we assigned a weight based on the Pagerank value of the node.

Creating virtual zones

We also created one virtual zone for each TAZ and connected that virtual zone to all the zone nodes of that TAZ. This connection was created in both direction to ensure strong connectivity in the subnetwork. Strong connectivity was assumed to ensure any node in the subnetwork can be reached from any other node in the subnetwork. The length of these connector links were determined based on the weights assigned to the zone nodes mentioned in the previous paragraph. Connectors to more important zone nodes were made shorter to increase the likelihood of more people prefering to take those links.

4.3 Analysis

To compare the manually collected dataset with the esimated link flows, first the manual two hour counts need to be converted to daily link flows. To do that, we use the profile given in Figure 4.1. Using this hourly profile and the time of the day when the manual counts were obtained we can convert the manual counts to daily link flows. Using that daily link with the daily profile given in Figure 4.2a and Figure 4.2b we can obtain the annual average daily link flows. It should be noted that the hourly and daily profile is highly dependant on the location of the link. However, here we are using one uniform hourly profile for all links taken from Molino et al. (2012). The daily profile is obtained from the day of year traffic flows from MnDOT's traffic for the year 2019. The daily profile is also just an average of all the available profile. Since, we do not have year long data available for all links we had to use this average profile. However, it should be noted that the daily profile is also highly dependant on the location of the link and therefore, this is source of error that we could not avoid in this analysis.

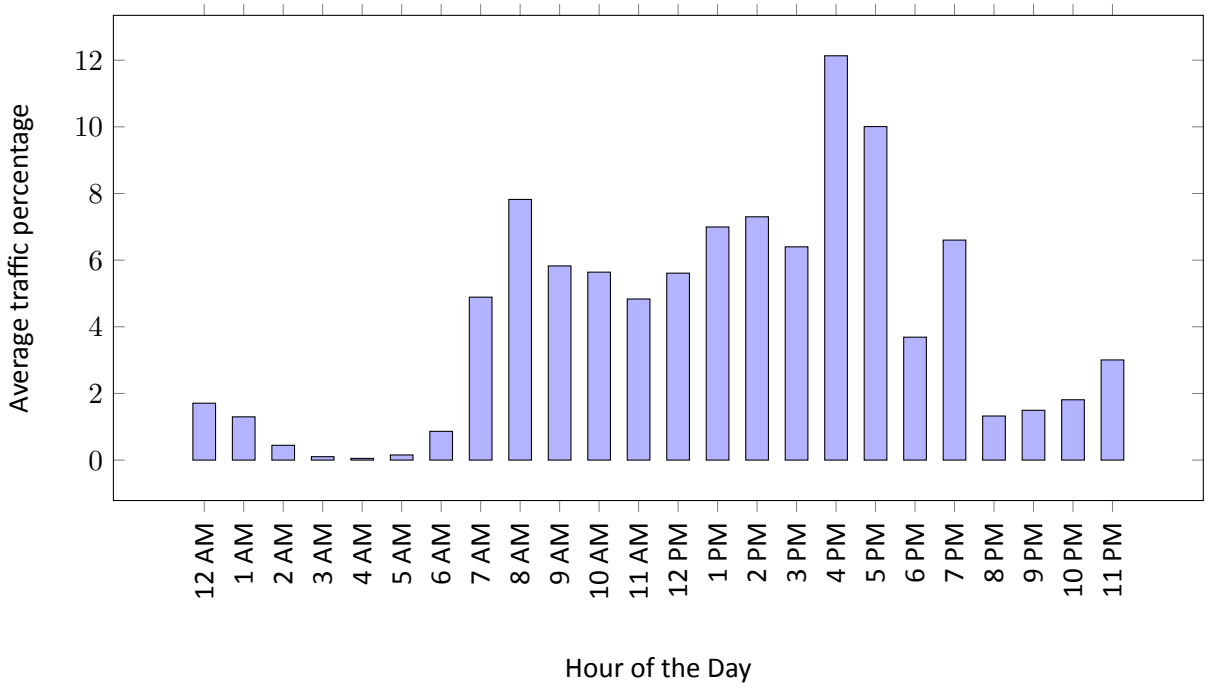
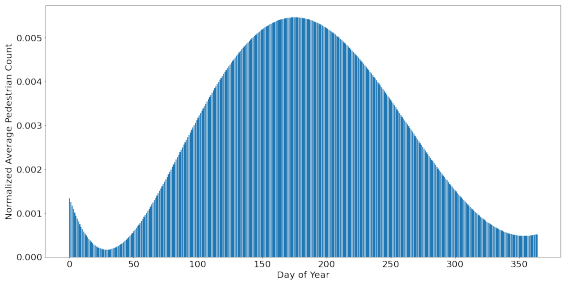
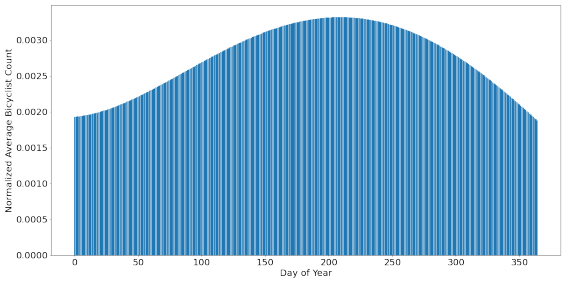


Figure 4.1: Normalized average hourly traffic profile.



(Figure 4.2a:) Normalized average daily traffic profile for pedestrians.



(Figure 4.2b:) Normalized average daily traffic profile for bicyclists.

Figure 4.2: Normalized average daily traffic for pedestrians and bicyclists.

The following tables show the manual two hour counts and the daily count converted from that manaul count using the hourly profile. That daily count is then converted to annual average daily traffic (AADT) using the daily profile for the relevant mode. The name and detailed coordinates about the sites can be found in the appendix of this report. To keep the tables relatively small, only the id of the sites are shown in the following tables. The estimate from the solution proposed is shown and labeled as just the estimate. The adjusted partial counts obtained from StravaMetro is also shown and labeled as Count.

Table 4.1: Bicyclist link flows

Site ID	$2-$hour count	Daily count	AADT (a)	Estimate (e)	Count (s)	$\lvert a-e\rvert/a$ (%)	$\lvert a-s\rvert/a$ (%)
54	7.00	67.03	187.09	40		78.62	
29	76.00	7480.31	1357.53	243	183	82.10	86.52
46	11.50	110.12	106.67	5	4	95.31	96.25
2	55.00	10064.04	4666.86	380	197	91.86	95.78
35	3.00	548.95	1842.55	50	36	97.29	98.05
16	3.00	295.28	893.55	18	18	97.99	97.99
19	1.00	7.93	114.38	3		97.38	
21	0.50	91.49	2118.45	13	13	99.39	99.39
66	4.00	731.93	197.82	51	47	74.22	76.24
69	21.00	10248.90	2638.09	235	184	91.09	93.03
25	253.00	2007.14	81.07	2854	2618	3420.47	3129.36
50	36.00	6587.37	28342.95	200	200	99.29	99.29
47	15.50	602.70	344.86	27	27	92.17	92.17
20	4.00	644.41	530.78	65	65	87.75	87.75
32	5.00	3253.09	3779.34	72	70	98.09	98.15
57	15.50	148.15	225.66	22	16	90.25	92.91
15	2.00	323.69	150.68	1	1	99.34	99.34
22	5.00	430.61	659.50	56	32	91.51	95.15
3	2.00	196.85	154.10	109	109	29.27	29.27
9	58.50	2473.62	281.82	903	883	220.42	213.32
49	16.00	2927.72	1420.33	82	82	94.23	94.23
48	14.50	5401.99	15912.26	58	26	99.64	99.84
27	2.00	19.15	63.92	1		98.44	
70	22.00	1263.06	501.19	330	303	34.16	39.54
43	4.00	1640.91	2698.08	0		0.00	
38	217.00	2077.83	186.49	202	171	8.31	8.31
6	10.00	574.12	175.96	10		94.32	
40	30.00	19518.54	21270.14	299	299	98.59	98.59
12	2.00	43.07	183.31	7		96.18	
11	5.00	263.13	0.00	149	146	inf	inf
8	280.00	2221.34	51.99	217	217	317.38	317.38
28	339.00	220559.53	121354.20	1140		99.06	
18	4.00	393.70	893.11	0		0.00	
67	12.00	114.57	33.58	6		82.13	
37	5.00	39.67	31.89	12		62.37	
56	3.00	1951.85	980.77	60	42	93.88	95.72
58	48.50	636.25	155.37	97	97	37.57	37.57
24	18.00	5734.17	445.17	1		99.78	
55	18.00	171.86	38.31	182	182	375.07	375.07
17	2.00	90.91	158.39	1		99.37	
23	59.00	38386.47	2936.97	30	5	98.98	99.83
63	47.00	372.87	91.50	154	154	68.31	68.31
41	5.00	914.91	11928.20	6		99.95	
42	8.00	3904.34	3915.61	9	9	99.77	99.77
30	22.50	6435.52	2591.15	166	166	93.59	93.59
39	15.00	143.21	897.24	59		93.42	
52	22.00	174.53	901.03	100	90	88.90	90.01
53	22.00	14313.60	43486.42	19		99.96	
64	73.00	698.99	334.07	280	267	16.19	20.08
26	6.00	344.47	58.06	5	4	91.39	93.11
14	3.00	1951.85	986.94	8		99.19	
62	9.00	156.48	52.75	35	31	33.65	41.23
61	24.00	190.40	150.65	18	18	88.05	88.05

Table 4.2: Pedestrian link flows

Site ID	$2-$hour count	Daily count	AADT (a)	Estimate (e)	Count (s)	$\lvert a-e\rvert/a$ (%)	$\lvert a-s\rvert/a$ (%)
54	34.00	325.56	187.09	11	5	94.12	97.33
29	24.00	2362.20	1357.53	286	264	78.93	80.55
46	20.50	196.29	106.67	91	5	14.69	95.31
2	45.00	8234.22	4666.86	796	539	82.94	88.45
35	18.00	3293.69	1842.55	27	24	98.53	98.7
16	17.00	1673.23	893.55	30	23	96.64	97.43
19	6.00	222.65	114.38	277	17	142.18	85.14
21	13.00	4015.50	2118.45	332		84.33	
66	2.00	365.97	197.82	17	8	91.41	95.96
69	10.00	4880.43	2638.09	2	2	99.92	99.92
25	19.00	150.73	81.07	386	386	376.14	376.14
50	288.00	52698.99	28342.95	84	2	99.70	99.99
44	102.00	10039.37	5399.45	56	33	98.96	99.39
47	23.00	674.30	344.86	12		96.52	
20	9.50	1038.14	530.78	120	3	77.39	99.43
32	11.00	7156.80	3779.34	3	2	99.92	99.95
57	46.00	439.78	225.66	5	2	97.78	99.11
15	2.50	288.34	150.68	25	20	83.41	86.73
22	15.00	1271.31	659.50	28	26	95.75	96.06
3	3.00	295.28	154.10	184	52	19.40	66.26
9	21.50	540.72	281.82	575	539	104.03	91.26
49	15.00	2744.74	1420.33	61		95.71	
48	93.50	30799.52	15912.26	2945	2945	81.49	81.49
27	13.00	124.48	63.92	13		79.66	
70	17.00	976.00	501.19	438	437	12.61	12.81
43	13.00	5252.82	2698.08	62		97.70	
38	37.50	359.07	186.49	172	102	7.77	45.31
6	6.00	344.47	175.96	394		123.91	
40	64.00	41639.56	21270.14	45	6	99.79	99.97
12	12.50	358.87	183.31	19	1	89.64	99.45
13	7.00	278.19	140.67	1		99.29	
8	13.00	103.13	51.99	3415	3415	6468.46	6468.46
28	370.00	240728.69	121354.20	3121	3121	97.43	97.43
18	18.00	1771.65	893.11	17	14	98.10	98.43
67	7.00	66.83	33.58	12		64.27	
37	8.00	63.47	31.89	0		0.00	
56	3.00	1951.85	980.77	34	18	96.53	98.16
58	24.00	307.92	155.37	221	221	42.24	42.24
24	4.00	884.46	445.17	4		99.10	
60	2.00	976.09	490.09	1		99.80	
55	8.00	76.38	38.31	34	1	11.25	97.39
17	8.00	315.79	158.39	6	1	96.21	99.37
23	9.00	5855.56	2936.97	178	54	93.94	98.16
63	23.00	182.47	91.50	21	2	77.05	97.81
41	130.00	23787.74	11928.20	2	1	99.98	99.99
42	16.00	7808.69	3915.61	1	1	99.97	99.97
30	14.50	5162.75	2591.15	43	38	98.34	98.53
39	187.00	1785.39	897.24	102	14	88.63	98.44
52	226.00	1792.94	901.03	33	24	96.34	97.34
53	133.00	86532.21	43486.42	45	1	99.90	100.0
64	69.00	660.69	334.07	261	260	21.87	22.17
26	2.00	114.82	58.06	4	1	93.11	98.28
14	3.00	1951.85	986.94	64	1	93.52	99.9
62	6.00	104.32	52.75	65 14	10	73.46	81.04
61	37.00	293.53	150.65	6		96.02	

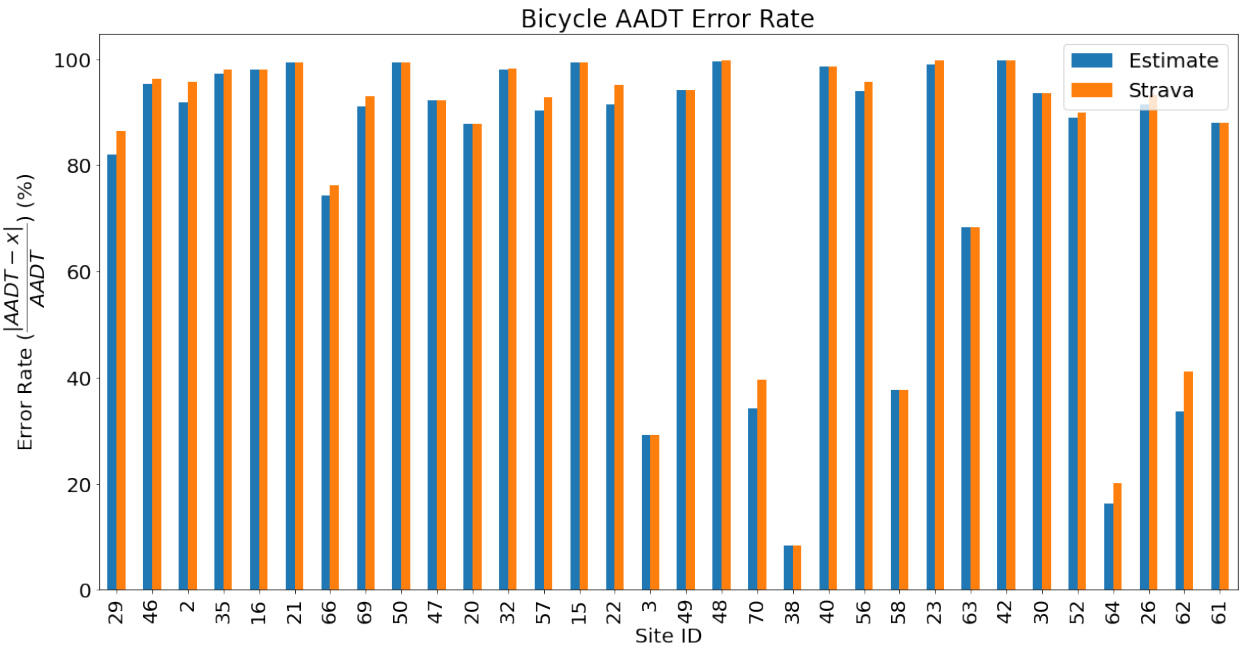


Figure 4.3: Percentage of error in the estimates for bicyclists.

From the table Table 4.2 we can see that some daily count values are very high for similar two hour counts. The reason for this is the time of the collected data. If the two hour count is high for a time of the day where the hourly profile predicts low distribution of traffic then the daily count will be very high after converting the two hour count to the daily count. This high volume is again converted to AADT using the daily profile. The daily profile in reality varies significantly from link to link. However, given the availability of data, we had to use the average daily profile to do that conversion. These are the sources of error that we could not avoid in this analysis. Therefore, we are showing the values of two hour count, daily and AADT so that the effects of these errors can be observed and interpreted correctly by the reader. The estimate from the solution algorithm is also shown in the table which is always greater than or equal to the partially observed count or the initial count. Since, previous literature showed that the partially observed counts are usually underestimates of the actual counts, the solution proposed in this report generates better esimates.

From the tables it can be seen that the percentage of error is almost always smaller for the estimate from the proposed solution compared to the partially observed counts extracted from StravaMetro. The difference between the magnitudes of the AADT and our esimates are high. However, the AADT values were obtained by performing two separate conversions that have introduced errors in the final values. Furthermore, the estimates are obtained using network, production, attraction and partial count dataset from the year 2019 but the manual counts were collected in 2022. One potential simple approach to reduce

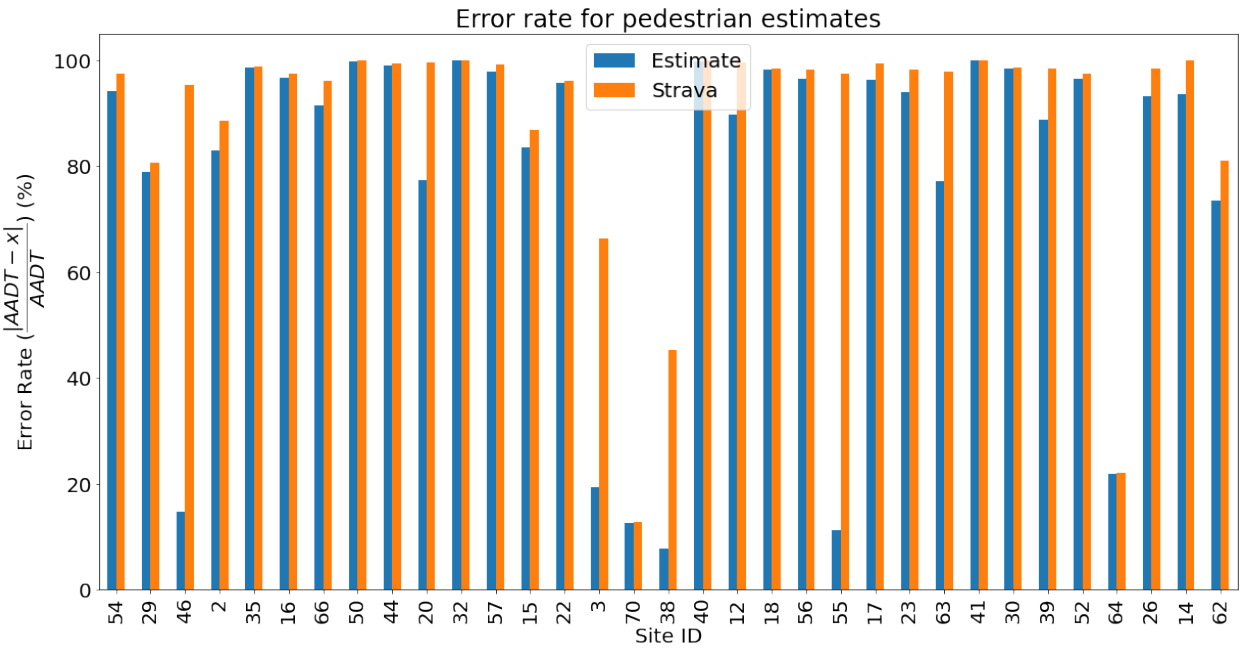


Figure 4.4: Percentage of error in the estimates for pedestrians.

the difference between the magnitudes of the AADT and the estimates is to scale the estimates by some factor dependent on the features of the link. Better network, production and attraction data and more accurate hourly and yearly profiles that are dependant on the locations and different features of the links can also lead to more accurate estimates.

4.4 Visualizing the maps

After running the solution algorithm on the networks, we obtained the link flow estimates for both the bicyclists and pedestrians. All the links where geometry information was available were visualized. To create the visualization, we used an open source tool called kepler.gl which is designed for geospatial data analysis, developed by Uber. The reason for choosing this tool is that it is open source, allows non-technical audiences to visualize data easily and a large amount of data can be visualized using this in a browser. This data can also be exported directly to a GIS file. We had the link geometries for most of the links and the link flow estimates for all the links. Using the link geometries which are linestrings, we first plotted the links on the map. Then, the link flow esimates were added to each link plotted on the map as a tooltip. The tooltip pops up when a user hovers over a link as shown in Figure 4.5.

By clicking the right arrow (shown in Figure 4.6) on the top left corner of the map, more options to

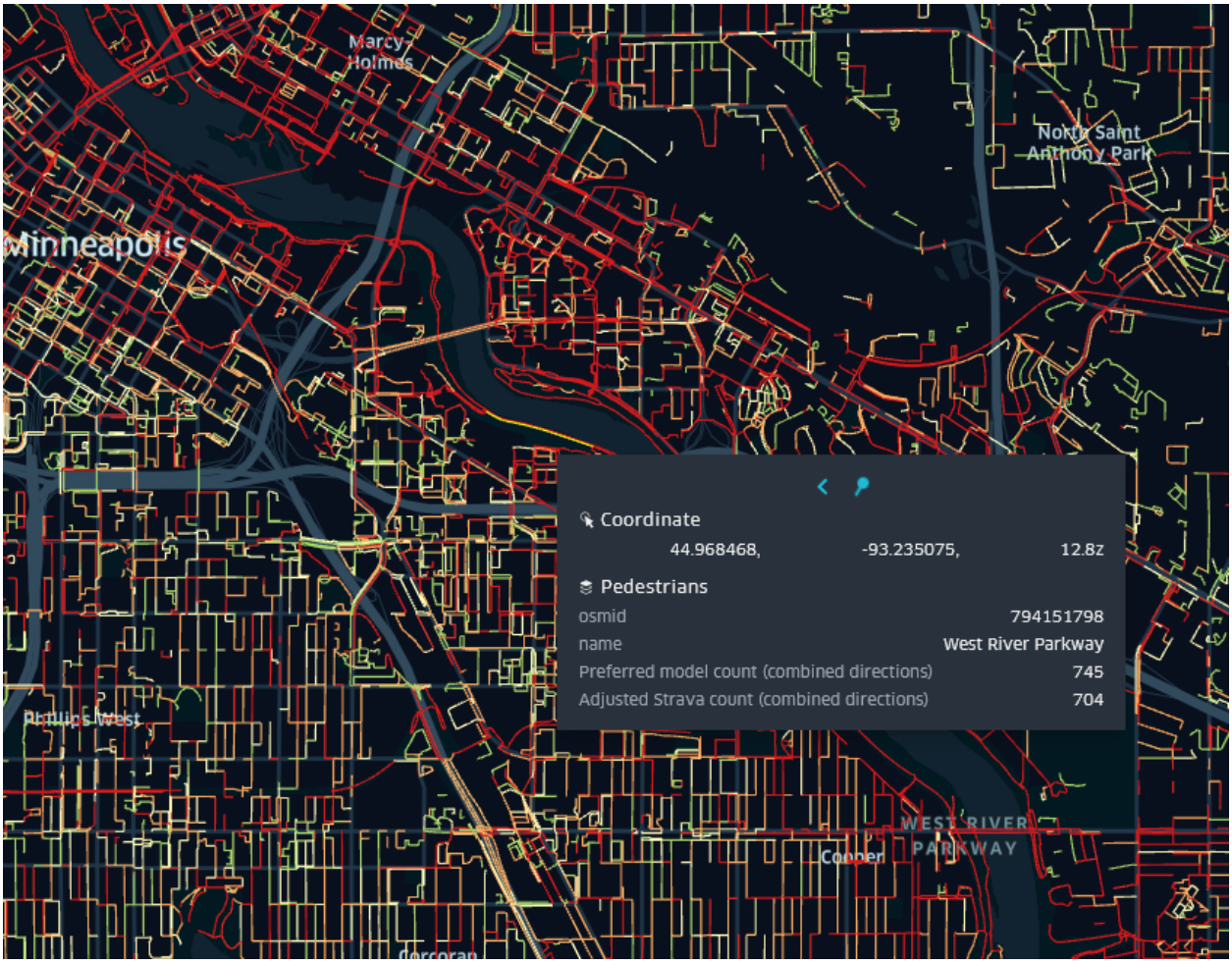


Figure 4.5: Link flow estimates for the pedestrian network

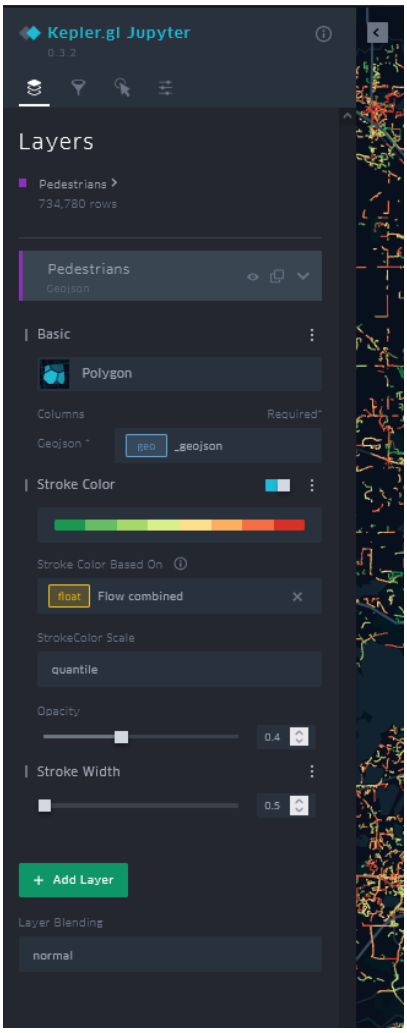
Figure 4.6: Access to more options to modify the map

modify the map can be accessed. On clicking the right arrow, a menu pops up on the left side of the map as shown in Figure 4.7. According to the user's preference, the heatmap can be modified based on the link flow estimates or the StravaMetro flows or the any other features of the links that are provided in the map. To access what features are available, the user can click on the layers option then the three vertical dots next to the text Stroke Color after enabling it. Then, the drop down menu accessed by clicking the Stroke Color Based On field will show all the features that are available. The color sequence can also be modified to recolor the links based on the user's preference. The interactions option in the expanded menu can be used to modify the interactions with the map. The user can activate the tooltip option and select different attributes of the links to be shown as a tooltip when the user hovers over a link.

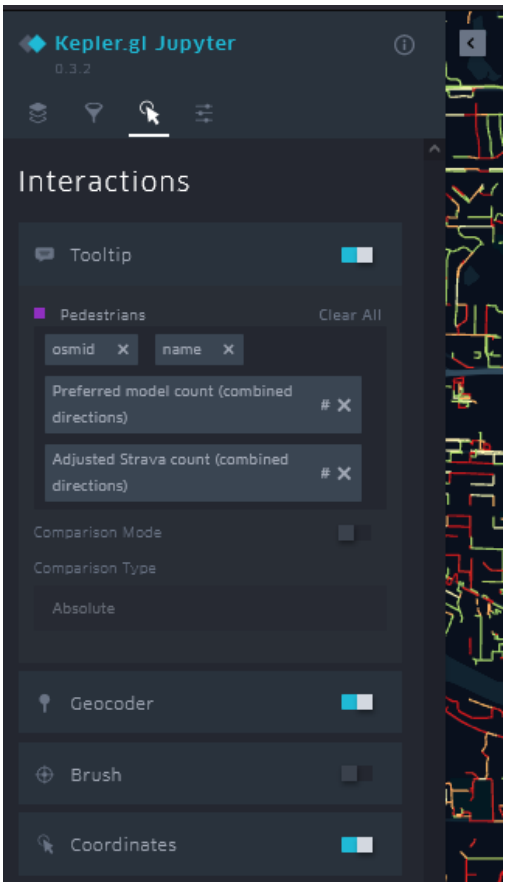
Other basic map interactions like zooming in and out, displaying legends, 3D view, searching for an address or coordinate etc can be done using the buttons and search bar provided on the top right corner of the map. Details about other features of the map can be found in the following urls: filters, map styles, interactions, map settings.

4.5 Summary

This chapter described how the Twin Cities Metro Area network data was processed and modified to make it suitable for the developed solution algorithm. The report also described how the link flow estimates were visualized on a map using the open source tool called kepler.gl. The data processing and visualization steps described in this report was done for both the pedestrian and bicyclist networks. Furthermore, how the maps can be modified to suit the user's preference was also discussed. The steps to modify the maps were discussed along with visual examples. Sources where detailed description and documentation about the map and the tool used to create the visualization were also provided.



(Figure 4.7a:) Menu to modify the map layers



(Figure 4.7b:) Menu to modify the map interactions

Figure 4.7: Menu to modify different aspects of the map.

Link to the map files

The GIS shape files are available at `https://z.umn.edu/bike_ped_volume_estimation` and the following maps are available at `https://z.umn.edu/bike_ped_maps`:

1. Peds_custom_buckets_1.html
2. Peds_custom_buckets_2.html
3. Peds_20_1.html
4. Peds_20_2.html
5. Peds_25_1.html
6. Peds_25_2.html
7. Bikes_custom_buckets.html
8. Bikes_20.html
9. Bikes_25.html
10. Book1.xlsx

The html files with the string "custom_buckets" in their names have the scale set to break at 25, 50, 100, 200, 500, 1000, $2000+$. For the other two html files, the scale is set to break every 20 and 25 percentile. A table showing the number of segments and miles of facilities that fall within each group is also given in the Book1.xlsx file.

5 Conclusions and Next Steps

5.1 Lessons Learned

Over the course of the research project, several key lessons were learned that will help future research activities. Some important lessons are identified in bulleted form below:

- There are inherent limitations associated with relying on incomplete data sources such as OpenStreetMaps. Specifically, OpenStreetMaps was missing many links in the bicyclist and pedestrian networks. This made it challenging to account for all possible paths.
- Relying on model-based data such as StreetLight pedestrian data has inherent limitations. For example, if input data are based purely on calibrated models and not on actual bicyclist or pedestrian counts, the downstream methods for estimating link-level volumes may be inaccurate as behavior shifts and calibrated models no longer are accurate. Therefore, additional work will be needed to ensure that the input data are current and coming from count-augmented trip estimates.
- While OpenStreetMaps provides a valuable, open-source resource for mapping, starting with a more complete bicycle and pedestrian oriented map would have saved considerable time, and likely yielded a more complete final product.
- Knowing the use cases for the resulting data product in advance would have been beneficial. Specifically, future efforts should convene more extensive conversations and include involvement with MnDOT GIS product users (e.g., as TAP members) to help the researchers understand potential use cases for the data and understand what data format is most useful for the end-users.
- As a result of the OpenStreetMaps used and data format, some links in the network do not have AADP or AADB estimates, and others have multiple estimates that are overlapping. When using the map,

we recommend the following:

- Overlapping estimates: Use the highest estimate on overlapping estimates. This should be the correct estimate that the algorithm converged to.
- Missing estimates: On links with missing AADP or AADB estimates, we recommend interpolating by averaging between the nearest two links on the same road with complete AADP or AADB information. To calculate missing AADB:

$$AADB_i = \frac{AADB_{i-1} + AADB_{i+1}}{2},$$

where $AADB_i$ is the missing AADB value on link i, and $AADB_{i-1}$ and $AADB_{i+1}$ are the nearest two available AADB measurements on the upstream and downstream links $i-1$ and $i+1$, respectively. Similarly, for AADP, the calculation

$$AADP_i = \frac{AADP_{i-1} + AADP_{i+1}}{2}$$

should be used, where $AADP_i$ is the missing AADP value on link i, and $AADP_{i-1}$ and $AADP_{i+1}$ are the nearest two available AADP measurements on the upstream and downstream links $i-1$ and $i+1$, respectively.

5.2 Next Steps

While the research project accomplished the goals identified at the onset, there is still ample room for expansion and continuation of this work before it is fully integrated into the decision process. Thoughts on potential next steps are provided in bulleted form below:

- The current analysis is only applicable to the Minneapolis-St. Paul Metro area. Additional work would be needed to expand the analysis to the entire state.
- In its current form, the output is not used to inform funding decisions. However, with additional accuracy, this could be expanded to provide data for funding allocation when considering infrastructure investments.

- As mentioned in the previous bullet above, this work currently relies on simulated input data. However, additional work on data collection could be conducted to expand this work to use mobile phone data directly, thus eliminating the need for additional third-party data vendors.

- Additional analysis could be conducted to identify where there are gaps in the bicyclist and pedestrian network that, if filled, would improve access. Furthermore, quantifying the improvement in access would be of substantial value.

- The current project only considers data from bicyclists and pedestrians. However, increasingly, micromobility (e.g., e-scooters) is becoming prevalent in cities. This work could be expanded to include data from micromobility vendors in addition to bicyclists and pedestrians.

- Moreover, additional data exists such as bus ridership data. How can such data be integrated to improve accuracy of the bicyclist and pedestrian count estimates?

- Additionally, many interesting questions around route choice arise when considering how infrastructure investments influence bicyclist and pedestrian design. These types of questions can potentially be investigated with the data and models produced in this project.

- There is potential bias that is introduced when considering data from only a subset of users. Additional work is needed to study how to generalize about mobility trends by only considering a subset of users without introducing inherent bias.

- One potential next step would be to use a geographic information system (GIS) platform such as QGIS or Esri to produce GIS feature classes to share with MnDOT and ensure the data are as maximally usable (i.e., no overlapping line segments) as possible.

References

Boeing, G. (2017). Osmnx: New methods for acquiring, constructing, analyzing, and visualizing complex street networks. Computers, Environment and Urban Systems, 65:126–139.

Brin, S. and Page, L. (1998). The anatomy of a large-scale hypertextual web search engine. Computer Networks and ISDN Systems, 30(1-7):107–117.

Camacho-Torregrosa, F. J., Llopis-Castelló, D., López-Maldonado, G., and García, A. (2021). An examination of the strava usage rate—a parameter to estimate average annual daily bicycle volumes on rural roadways. Safety, 7(1):8.

Dadashova, B. and Griffin, G. P. (2020). Random parameter models for estimating statewide daily bicycle counts using crowdsourced data. Transportation Research Part D: Transport and Environment, 84:102368.

Dadashova, B., Griffin, G. P., Das, S., Turner, S., and Sherman, B. (2020). Estimation of average annual daily bicycle counts using crowdsourced strava data. Transportation Research Record, 2674(11):390–402.

De Berg, M. (2000). Computational geometry: Algorithms and applications. Springer Science & Business Media, Heidelberg, Germany.

Estellés-Arolas, E. and González-Ladrón-de Guevara, F. (2012). Towards an integrated crowdsourcing definition. Journal of Information Science, 38(2):189–200.

Federal Highway Administration (2001). Traffic monitoring guide. Washington, DC: Federal Highway Administration.

Fratar, T. J. (1954). Vehicular trip distribution by successive approximations. Traffic Quarterly, 8(1).

Jestico, B., Nelson, T., and Winters, M. (2016). Mapping ridership using crowdsourced cycling data. Journal of Transport Geography, 52:90–97.

Jones, R. H. (1980). Maximum likelihood fitting of arma models to time series with missing observations. Technometrics, 22(3):389–395.

Kwigizile, V., Oh, J.-S., Kwayu, K., et al. (2019). Integrating crowdsourced data with traditionally collected data to enhance estimation of bicycle exposure measure. Technical report.

Lee, K. and Sener, I. N. (2020). Emerging data for pedestrian and bicycle monitoring: Sources and applications. Transportation research interdisciplinary perspectives, 4:100095.

Lee, K. and Sener, I. N. (2021). Strava metro data for bicycle monitoring: A literature review. Transport Reviews, 41(1):27–47.

Lin, Z. and Fan, W. D. (2020). Modeling bicycle volume using crowdsourced data from strava smartphone application. International journal of transportation science and technology, 9(4):334–343.

Lißner, S., Francke, A., and Becker, T. (2018). Modeling cyclists traffic volume–can bicycle planning benefit from smartphone based data? Proceedings of the 7th Transport Research Arena TRA.

McGuckin, N. and Fucci, A. (2018). Summary of travel trends: 2017 national household travel survey. Technical report.

Molino, J. A., Kennedy, J. F., Inge, P., Bertola, M. A., Beuse, P. A., Fowler, N. L., Emo, A. K., and Do, A. (2012). A distance-based method to estimate annual pedestrian and bicyclist exposure in an urban environment. Technical report.

Neis, P., Zielstra, D., and Zipf, A. (2012). The street network evolution of crowdsourced maps: Openstreetmap in germany 2007–2011. Future Internet, 4(1):1–21.

Nelson, T., Roy, A., Ferster, C., Fischer, J., Brum-Bastos, V., Laberee, K., Yu, H., and Winters, M. (2021). Generalized model for mapping bicycle ridership with crowdsourced data. Transportation Research Part C: Emerging Technologies, 125:102981.

Nishi, K., Tsubouchi, K., and Shimosaka, M. (2014). Hourly pedestrian population trends estimation using location data from smartphones dealing with temporal and spatial sparsity. In Proceedings of the

22nd ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, pages 281–290.

Ohlms, P. B., Dougald, L. E., and MacKnight, H. E. (2019). Bicycle and pedestrian count programs: Scan of current US practice. Transportation Research Record, 2673(3):74–85.

Roll, J. (2018). Bicycle count data : What is it good for? a study of bicycle travel activity in central lane metropolitan planning organization. Oregon. Dept. of Transportation.

Roy, A., Nelson, T. A., Fotheringham, A. S., and Winters, M. (2019). Correcting bias in crowdsourced data to map bicycle ridership of all bicyclists. Urban Science, 3(2):62.

Saad, M., Abdel-Aty, M., Lee, J., and Cai, Q. (2019). Bicycle safety analysis at intersections from crowdsourced data. Transportation Research Record, 2673(4):1–14.

Sanders, R. L., Frackelton, A., Gardner, S., Schneider, R., and Hintze, M. (2017). Ballpark method for estimating pedestrian and bicyclist exposure in seattle, washington: Potential option for resource-constrained cities in an age of big data. Transportation Research Record, 2605(1):32–44.

Strauss, J., Miranda-Moreno, L. F., and Morency, P. (2015). Mapping cyclist activity and injury risk in a network combining smartphone gps data and bicycle counts. Accident Analysis & Prevention, 83:132–142.

Turner, S., Sener, I. N., Martin, M. E., Das, S., Hampshire, R. C., Fitzpatrick, K., Molnar, L. J., Colety, M., Robinson, S., and Shipp, E. (2017). Synthesis of methods for estimating pedestrian and bicyclist exposure to risk at areawide levels and on specific transportation facilities. Technical report.

Turner, S. M., Martin, M. W., Griffin, G. P., Le, M., Das, S., Wang, R., Dadashova, B., and Li, X. (2020). Exploring crowdsourced monitoring data for safety. Technical report.

Xiao, Y. and Watson, M. (2019). Guidance on conducting a systematic literature review. Journal of Planning Education and Research, 39(1):93–112.

Appendix A: Site Selection

The locations were sorted in a descending order based on the Strava counts. Then the top and bottom count locations were selected. StreetLight provides the counts for different regions so the sites for StreetLight were selected based on whether those links had higher count in Strava dataset. Some other criteria like proximity of the sites was also considered to make sure a lot of the sites that were very close was not selected.

Site ID: 1

Site Location: Cedar Avenue South, Hale, Nokomis, Minneapolis, Hennepin County, Minnesota, 55407 (Map)

Nearest Intersection/Site: Cedar Ave S. at W Lake Nokomis Pkwy and Cedar Ave S. at W Lake Nokomis Pkwy at E 52nd St

(Figure 5.1a:) Site View

(Figure 5.1b:) Satellite View

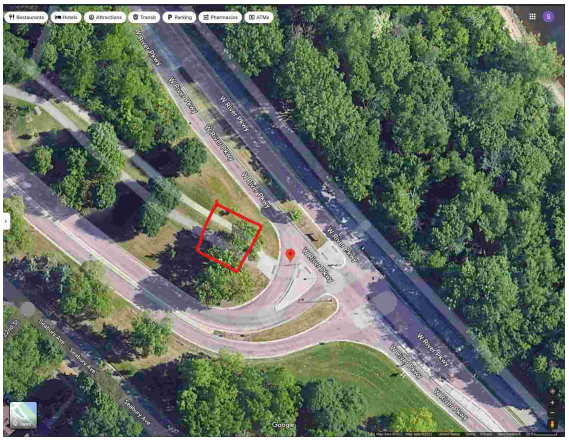
Description: One of the top Strava pedestrian and StreetLight bicyclists count locations. Near the Lake Nokomis Park, Triangle Park and the Grand Rounds National Scenic Byway which makes it a scenic route. Sidewalk present. Close to Hale Elementary School.

Site ID: 2

Site Location: West River Parkway, Seward, Minneapolis, Hennepin County, Minnesota, 55455-0500 (Map)

Nearest Intersection/Site: E Franklin Ave at W River Pkwy

(Figure 5.2a:) Site View



(Figure 5.2b:) Satellite View

Description: One of the top Strava and StreetLight count locations for pedestrian and bicyclists. Link is part of West River Parkway Trail. Proximity to Mississippi river and the west river parkway trail makes it Scenic and attractive to tourists. Walking path and bicycle lane present. Separate walking and bicyclist lane present.

Site ID: 3

Site Location: Fox Street, Orono, Hennepin County, Minnesota, 55391 (Map)

Nearest Intersection/Site: Orono Orchard Rd S at Fox St

(Figure 5.3a:) Site View

(Figure 5.3b:) Satellite View

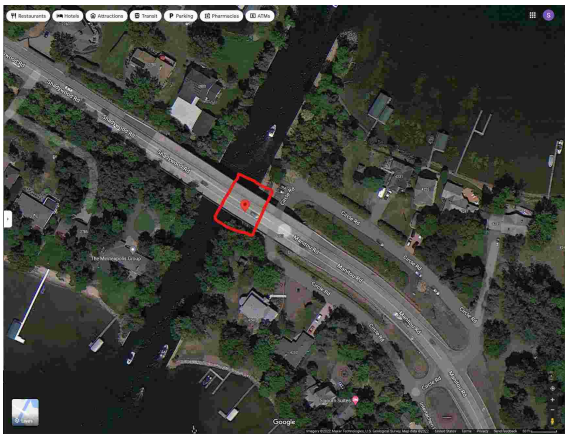
Description: One of the top Strava count locations. High Strava count observed for pedestrians but sidewalk or walking facilities or even shoulder not present. For this location the count from Strava was collected during a weekday.

Site ID: 4

Site Location: Narrows Bridge, Shadywood Road, Navarre, Orono, Hennepin County, Minnesota, 55391 (Map)

Nearest Intersection/Site: Manitou Rd at Circle Rd

(Figure 5.4a:) Site View



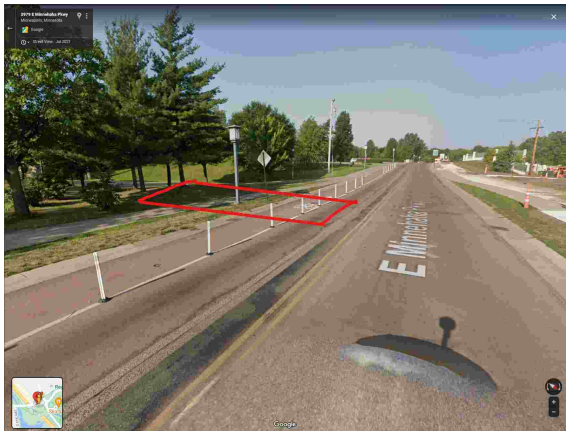
(Figure 5.4b:) Satellite View

Description: One of the top Strava count locations. High Strava count observed for pedestrians. No sidewalk present but wide shoulder is present. Lafayette bay and Carman bay can be seen from the site which makes it a scenic route.

Site ID: 5

Site Location: 4003-4063 E Minnehaha Pkwy, Minneapolis, MN 55406 (Map)

Nearest Intersection/Site: Godfrey Circle



(Figure 5.5a:) Site View

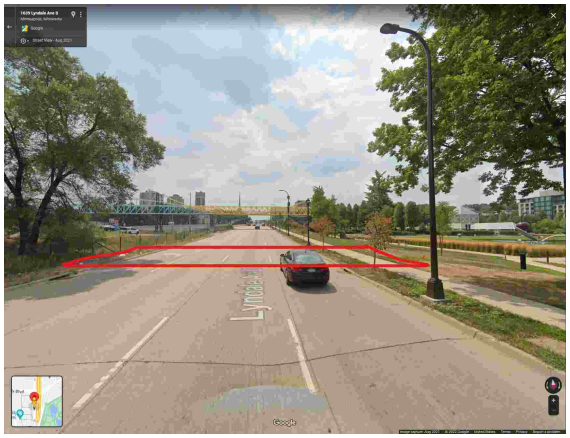
(Figure 5.5b:) Satellite View

Description: One of the top Strava count locations. Connects to the West River Parkway Trail. Close to Minnehaha Regional Park, Minnehaha Falls and Longfellow Gardens which makes it a scenic and attractive to tourists route. Higher pedestrian count observed.

Site ID: 6

Site Location: Lyndale Ave S, Loring Park, Minneapolis, Hennepin County, Minnesota, 55403 (Map)

Nearest Intersection/Site: Lyndale Avs S at Dunwoody Blvd



(Figure 5.6a:) Site View

(Figure 5.6b:) Satellite View

Description: One of the top Strava count locations. High Strava count observed for pedestrians. Connects to the West River Parkway Trail. Close proximity the Spoonbridge and Cherry, Minneapolis Sculpture Garden and Walker Art Center makes it a scenic site and attractive to tourists.

Site ID: 7

Site Location: Dean Parkway, Cedar - Isles - Dean, Bde Maka Ska - Isles, Minneapolis, Hennepin County, Minnesota, 55416 (Map)

Nearest Intersection/Site: Dean Pkwy at W 28th St

(Figure 5.7a:) Site View

(Figure 5.7b:) Satellite View

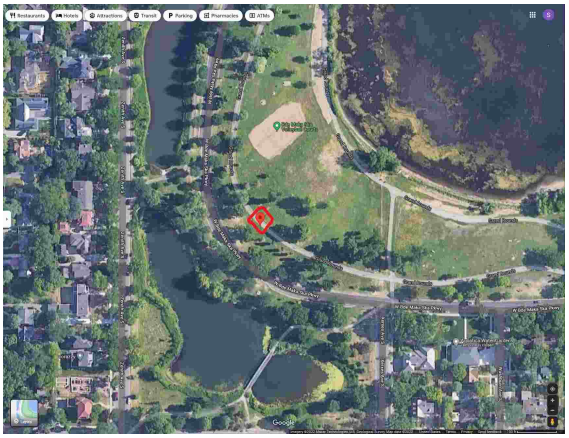
Description: One of the top Strava count locations. High Strava count observed for pedestrians. Near the Cedar Lake, Lake of the isles park and Bde Maka Ska which makes it a Scenic and touristy route. Part of the Dean Parkway Trail.

Site ID: 8

Site Location: West Bde Maka Ska Parkway, West Maka Ska, Bde Maka Ska - Isles, Minneapolis, Hennepin County, Minnesota, 55416 (Map)

Nearest Intersection/Site: Pleasant St SE at Pillsbury Dr SE

(Figure 5.8a:) Site View



(Figure 5.8b:) Satellite View

Description: One of the top Strava and StreetLight count locations. High Strava count observed for pedestrians and bike. Part of the Bde Maka Ska trai Near Bde Maka Ska lake, Thomas Beach which makes it scenic. Part of the Bde Maka Ska trail.

Site ID: 9

Site Location: Theodore Wirth Pkwy, Minneapolis, Hennepin County, Minnesota, 55411-1124 (Map)

Nearest Intersection/Site: Theodore Wirth Parkway at Plymouth Ave N

(Figure 5.9a:) Site View

(Figure 5.9b:) Satellite View

Description: One of the top Strava count locations. High Strava count observed for pedestrians. Connects to Theodore Wirth Wildflower Trail which makes it a Scenic route and a Trail. Higher pedestrian count observed. Bike lane and walkway present. Potential automated counter deployment location.

Site ID: 11

Site Location: Star Loop, Bloomington, Hennepin County, Minnesota, 55438 (Map)

Nearest Intersection/Site: Hyland Lake Park Reserve & Play Area and Star Loop

(Figure 5.10a:) Site View

(Figure 5.10b:) Satellite View

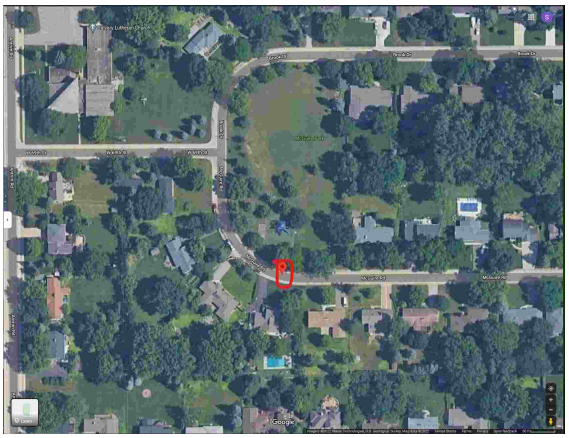
Description: Low Strava count location for both pedestrian and bicyclists. Low Strava count observed for the pedestrians. Near the Hyland Lake Park which makes it scenic and touristy.

Site ID: 12

Site Location: West 70th Street, Eden Prairie, Hennepin County, Minnesota, 55344 (Map)

Nearest Intersection/Site: W 69th St at McGuire Rd and McGuire Park

(Figure 5.11a:) Site View



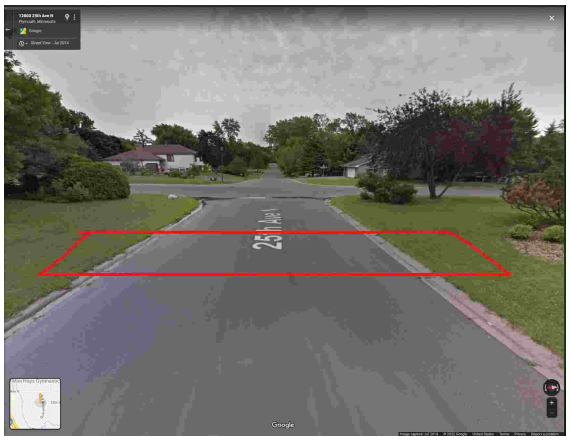
(Figure 5.11b:) Satellite View

Description: Low Strava count observed for the pedestrians. Close to Edina High School, Valley View Middle School and Calvary Lutheran Church. Very close to the McGuire Park which makes it touristy.

Site ID: 13

Site Location: 12800, 25th Avenue North, Plymouth, Hennepin County, Minnesota, 55441 (Map)

Nearest Intersection/Site: 25th Ave N at Quinwood Ln N



(Figure 5.12a:) Site View

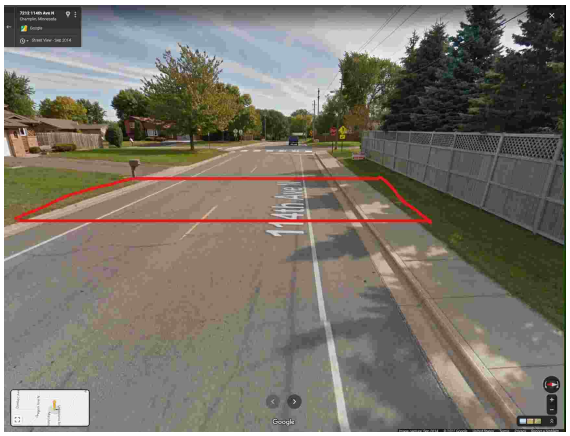
(Figure 5.12b:) Satellite View

Description: Low Strava count observed for the pedestrians. Sidewalk present. Close to Plymouth Creek, West Medicine Lake park and Heritage Park.

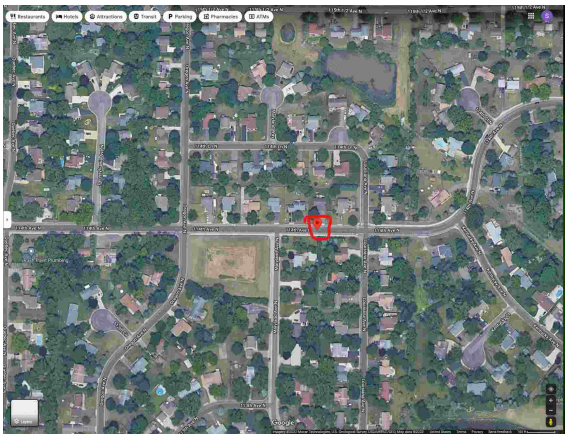
Site ID: 14

Site Location: 7120, 114th Avenue North, Champlin, Hennepin County, Minnesota, 55316 (Map)

Nearest Intersection/Site: 114th Ave N at Louisiaina Ave N



(Figure 5.13a:) Site View



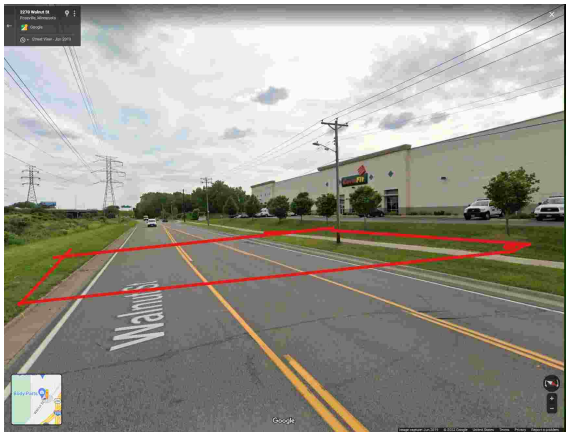
(Figure 5.13b:) Satellite View

Description: Low Strava and StreetLight count observed for the pedestrians. Sidewalk present. Close to Andrews Park, Brittany Park and Oxbow Creek Park. Close to Jackson Middle School and Chamlin Park High School.

Site ID: 15

Site Location: Walnut Street, Roseville, Ramsey County, Minnesota, 55113 (Map)

Nearest Intersection/Site: Terminal Rd at Walnut St. near CertiFit Auto Body Parts.



(Figure 5.14a:) Site View

(Figure 5.14b:) Satellite View

Description: Low Strava pedestrian and StreetLight bicyclist counts were observed. Sidewalk Present.

Site ID: 16

Site Location: 2600, Idaho Avenue South, Bronx Park, Saint Louis Park, Hennepin County, Minnesota, 55426 (Map)

Nearest Intersection/Site: W 26th St at Idaho Ave S

(Figure 5.15a:) Site View

(Figure 5.15b:) Satellite View

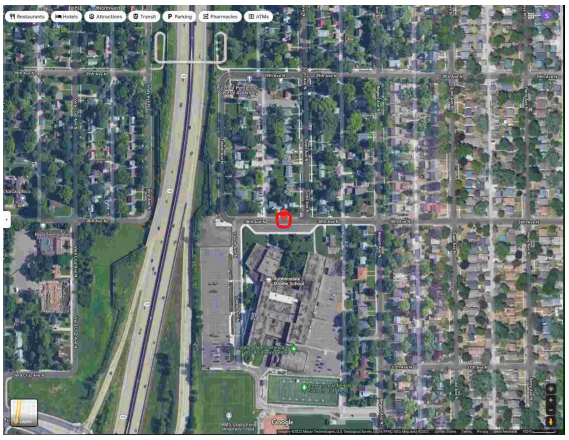
Description: Low Strava count observed for the pedestrians. Sidewalk Present. Near Nelson Park, Dakota Park, Carlson Field and Cedar Knoll Park. Proximity to Peter Hobart Primary Center makes is close to an educational institute.

Site ID: 17

Site Location: 38th Avenue North, Robbinsdale, Hennepin County, Minnesota, 55422 (Map)

Nearest Intersection/Site: 38th Ave N at Scott Ave N and 38th Ave N at Toledo Ave N

(Figure 5.16a:) Site View



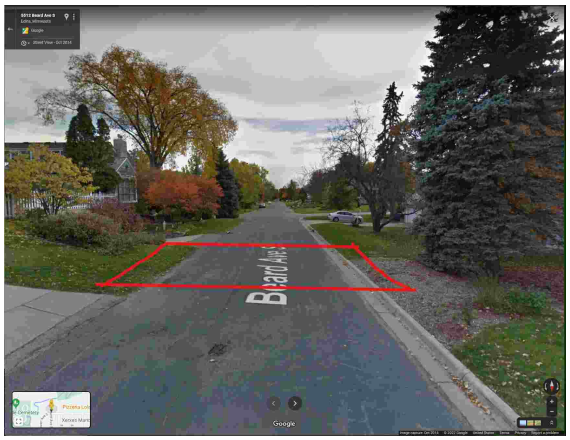
(Figure 5.16b:) Satellite View

Description: Low Strava count observed for the pedestrians. Close Robbinsdale Middle School, RiverTree School and Church of Pentecost. Sidewalk Present.

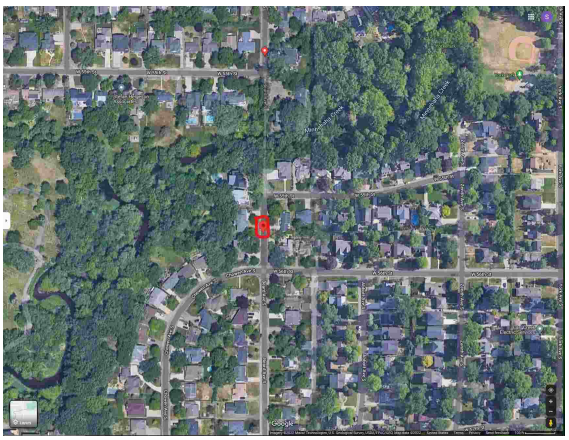
Site ID: 18

Site Location: Beard Avenue South, Edina, Hennepin County, Minnesota, 55410 (Map)

Nearest Intersection/Site: Beard Ave S at W 55th St and Beard Ave S at Chowen Ave S



(Figure 5.17a:) Site View



(Figure 5.17b:) Satellite View

Description: Low Strava count observed for the pedestrians. Close to Minnehaha Creek and York Park which makes it a link to a scenic route.

Site ID: 19

Site Location: Eagle Lake Drive, Maple Grove, Hennepin County, Minnesota, 55369 (Map)

Nearest Intersection/Site: Eagle Lake Dr at 65th Ave N and Eagle Lake Dr at 64th Ave N

(Figure 5.18a:) Site View

(Figure 5.18b:) Satellite View

Description: Low Strava and StreetLight count observed for the pedestrians. Close to the Eagle Lake, Woodcrest Park and Thoresen Park.

Site ID: 20

Site Location: 19 Av S, East 42nd Street, Standish, Powderhorn, Minneapolis, Hennepin County, Minnesota, 55406:55407 (Map)

Nearest Intersection/Site: S 19th Ave at E 42nd St and at E 43rd St

(Figure 5.19a:) Site View

(Figure 5.19b:) Satellite View

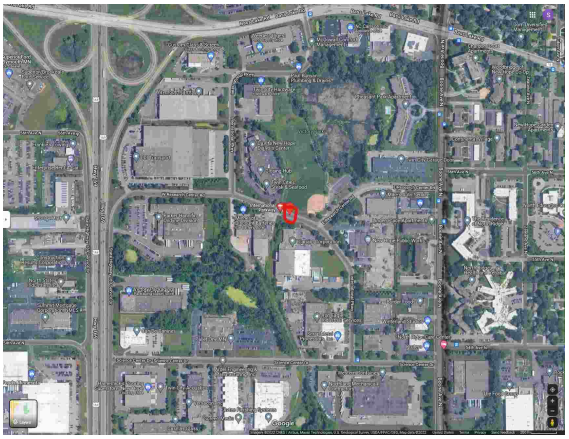
Description: Low Strava and StreetLight count observed for the pedestrians. Sidewalk present. Close to lake Hiawatha.

Site ID: 21

Site Location: International Pkwy New Hope, MN 55428 (Map)

Nearest Intersection/Site: International Pkwy at Science Center Drive and International Pkwy at W Research Center Rd.

(Figure 5.20a:) Site View



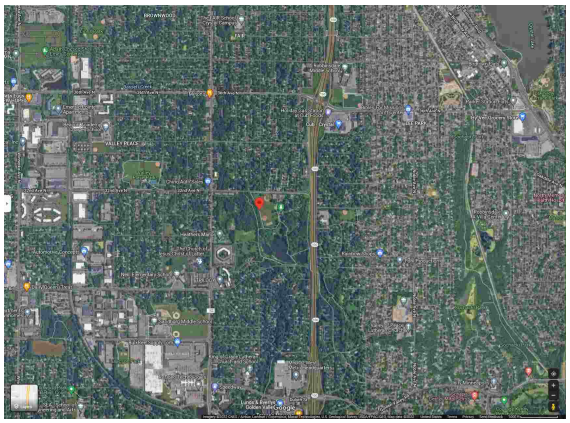
(Figure 5.20b:) Satellite View

Description: High Strava count observed for the bikes. Close to Victory Park. Close to North Education Center.

Site ID: 22

Site Location: Bassett Creek, Crystal, MN 55422 (Map)

Nearest Intersection/Site: Bassett Creek at 32nd Ave N



(Figure 5.21a:) Site View

(Figure 5.21b:) Satellite View

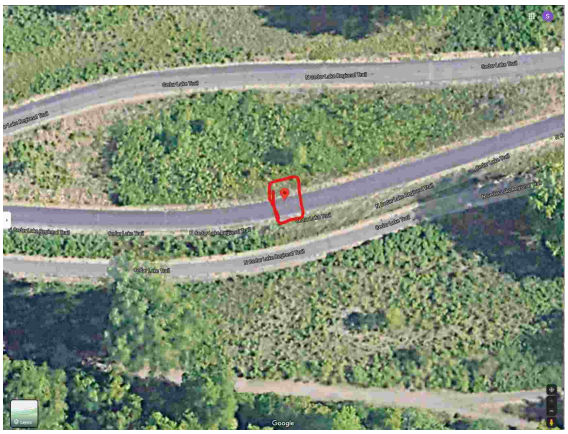
Description: High Strava count observed for the bikes. Proximity to Bassset Creek Park makes it a scenic and touristy route. Part of the Bassett Creek Trail. Close to Neill Elementary School, Beacon Academy Charter School and Valley Community Presbyterian Church.

Site ID: 23

Site Location: N Cedar Lake Regional Trail, Minneapolis, MN 55405 (Map)

Nearest Intersection/Site: Cedar Lake Rd at Grand Rounds Trail

(Figure 5.22a:) Site View



(Figure 5.22b:) Satellite View

Description: High Strava count observed for the bikes. Part of the N Cedar Lake Regional Trail. Close to Eloise Butler Wildflower Garden and the Cedar Lake which makes it a scenic and touristy route.

Site ID: 24

Site Location: Upland Ln N, Maple Grove, MN 55369 (Map)

Nearest Intersection/Site: Upland Ln N at 93rd Ave N

(Figure 5.23a:) Site View

(Figure 5.23b:) Satellite View

Description: High Strava count observed for the bikes. Near Medicine Lake Trail and Scenic. Sidewalk Present. Near Radiant Montessori School and Maple Grove Covenant Church.

Site ID: 25

Site Location: Midtown Greenway, Minneapolis, MN 55407 (Map)

Nearest Intersection/Site: Midtown Greenway at Cedar Ave S and Midtown Greenway at E 28th St

(Figure 5.24a:) Site View

(Figure 5.24b:) Satellite View

Description: High Strava and StreetLight count observed for the bikes. Separate lanes for bikes and pedestrians present. Potential automatic counter deployment location. Part of the Midtown Greenway Trail. Close to South High School.

Site ID: 26

Site Location: 103rd Ave N, Brooklyn Park, MN 55445 (Map)

Nearest Intersection/Site: Midtown Greenway at Cedar Ave S and Midtown Greenway at E 28th St

(Figure 5.25a:) Site View

(Figure 5.25b:) Satellite View

Description: High Strava count observed for the bikes. Sidewalk present. Suburban location. Close to the Rush Creek Regional Trail.

Site ID: 27

Site Location: NE 2nd St, Minneapolis, MN 55418 (Map)

Nearest Intersection/Site: NE Lowry Ave at NE 2nd St

(Figure 5.26a:) Site View

(Figure 5.26b:) Satellite View

Description: High Strava count observed for the bikes. Sidewalk Present.

Site ID: 28

Site Location: Lake Harriet Bike Trail, Lynnhurst, Minneapolis, MN 55419 (Map)

Nearest Intersection/Site: Lake Harriet Pkwy at Morgan Ave S and Lake Harriet Pkwy at Oliver Ave S

(Figure 5.27a:) Site View

(Figure 5.27b:) Satellite View

Description: High Strava and StreetLight count observed for the bikes. Sidewalk present. View of the lake Harriet makes it a scenic route.

Site ID: 29

Site Location: Minnetonka Blvd, Minnetonka, MN 55305 (Map)

Nearest Intersection/Site: Minnetonka Boulevard at Baker Rd

(Figure 5.28a:) Site View

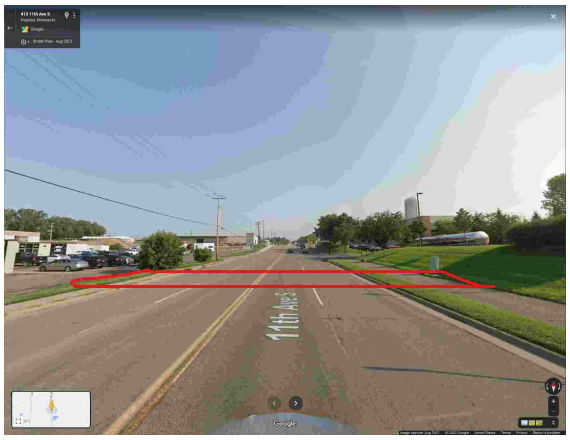
(Figure 5.28b:) Satellite View

Description: Low Strava count observed for the bikes. Separate bicycle lane and sidewalk present. Potential automated counter deployment location. Close to the Lake Minnetonka LRT Regional Trail.

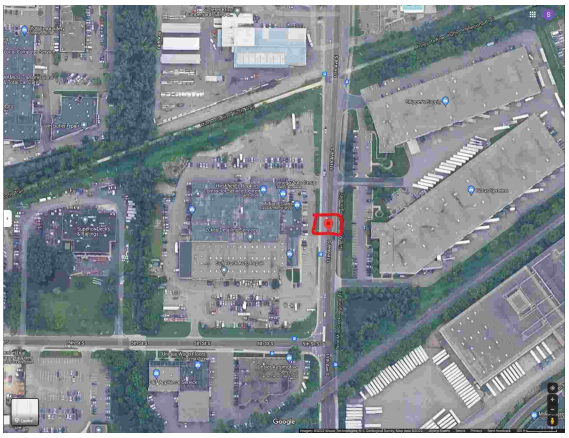
Site ID: 30

Site Location: 11th Ave S, Hopkins, MN 55343 (Map)

Nearest Intersection/Site: 11th Ave S at 5th St S



(Figure 5.29a:) Site View



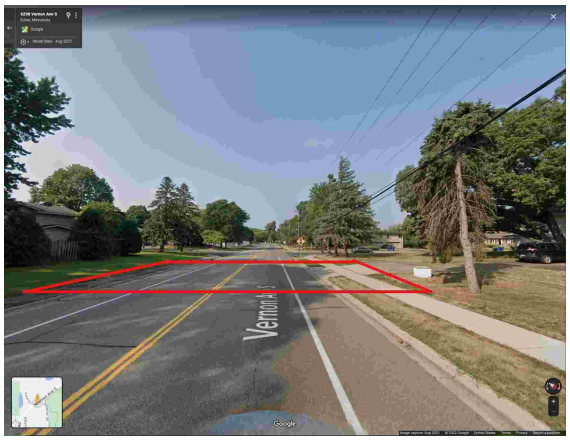
(Figure 5.29b:) Satellite View

Description: Low Strava count observed for the bikes. Sidewalk present.

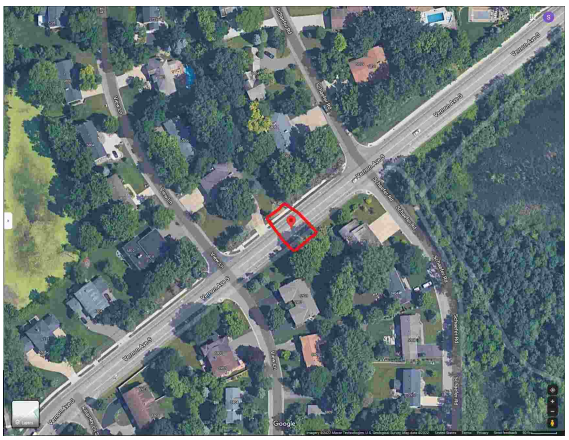
Site ID: 31

Site Location: 6299-6291 Vernon Ave S, Edina, MN 55436 (Map)

Nearest Intersection/Site: Vernon Ave S at Schaefer Rd and Vernon Ave S at View Ln



(Figure 5.30a:) Site View



(Figure 5.30b:) Satellite View

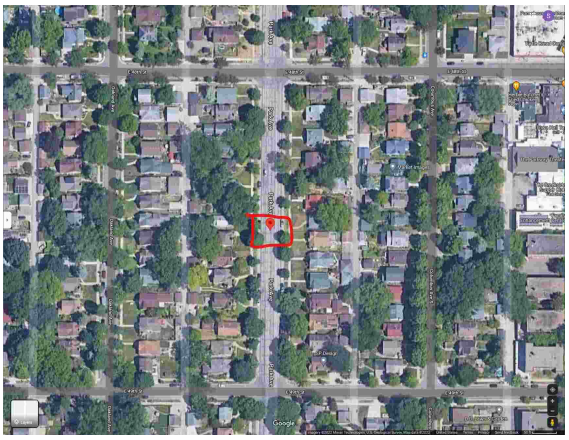
Description: Low Strava count observed for the bikes. Sidewalk present and separate bike lanes exist. Potential automatic counter deployment location.

Site ID: 32

Site Location: 64899-4801 Park Ave, Minneapolis, MN 55417 (Map)

Nearest Intersection/Site: Park Ave at E 49th St and Park Ave at E 48th St

(Figure 5.31a:) Site View



(Figure 5.31b:) Satellite View

Description: Low Strava count observed for the bikes. Sidewalk present and separate bike lanes exist. Potential automatic counter deployment location.

Site ID: 34

Site Location: 8127-8071 12th Ave S, Bloomington, MN 55425 (Map)

Nearest Intersection/Site: 12th Ave S at E 82nd St and 12th Ave S at E 80th St

(Figure 5.32a:) Site View

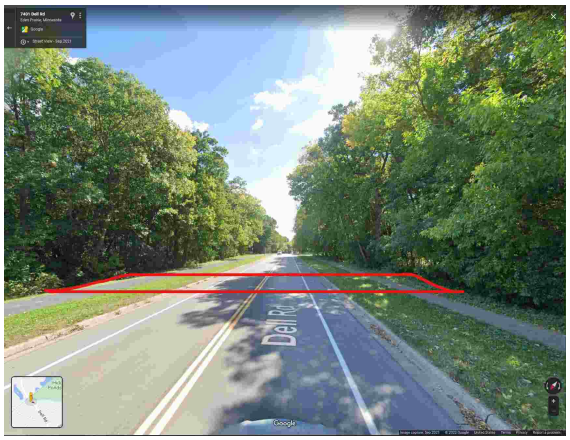
(Figure 5.32b:) Satellite View

Description: Low Strava count observed for the bikes. Sidewalk and bike lane present. Potential automatic counter deployment location.

Site ID: 35

Site Location: Dell Rd, Eden Prairie, MN 55346 (Map)

Nearest Intersection/Site: Dell Rd at Twilight Trail



(Figure 5.33a:) Site View

(Figure 5.33b:) Satellite View

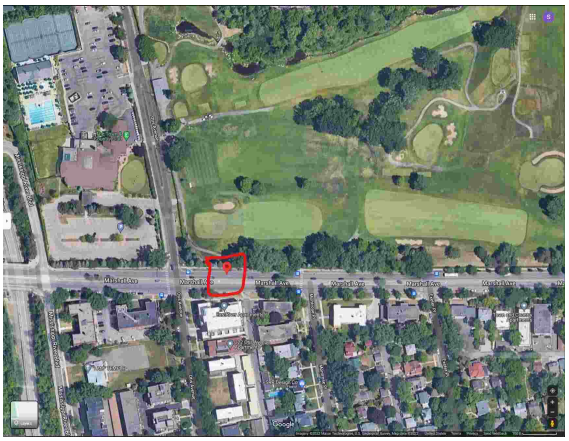
Description: Low Strava count observed for the bikes. Sidewalk and bike lane present. Potential automatic counter deployment location. Twilight trail is accesible from this route and close to Hidden Ponds Park make it a scenic route.

Site ID: 36

Site Location: D300 N Mississippi River Blvd, St Paul, MN 55104 (Map)

Nearest Intersection/Site: Marshall Ave at Otis Avenue

(Figure 5.34a:) Site View



(Figure 5.34b:) Satellite View

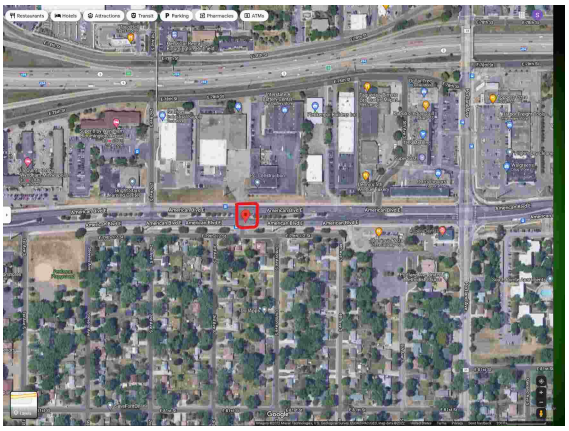
Description: Low Strava count observed for the bikes. Sidewalk and bike lane present. Potential automatic counter deployment location. Close to the Mississippi River Bridge makes it a scenic route.

Site ID: 37

Site Location: American Blvd, Bloomington, MN 55420 (Map)

Nearest Intersection/Site: American Blvd at Nicollet Ave and American Blvd at Portland Ave

(Figure 5.35a:) Site View



(Figure 5.35b:) Satellite View

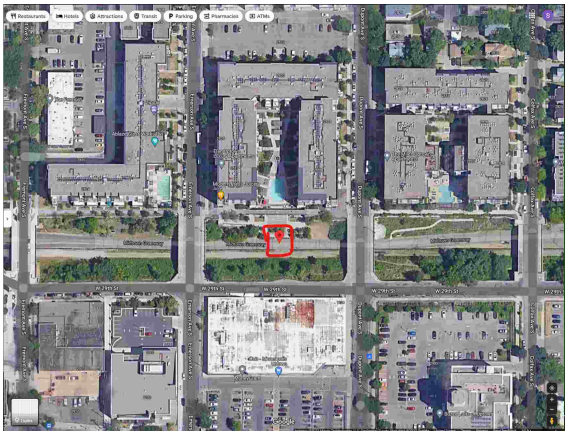
Description: High StreetLight count observed for the pedestrians. Sidewalk present.

Site ID: 38

Site Location: Midtown Greenway, Minneapolis, MN 55408 (Map)

Nearest Intersection/Site: Can be accessed from Midtown Greenway at Humboldt Ave S or Midtown Greenway at Nicollet Ave

(Figure 5.36a:) Site View



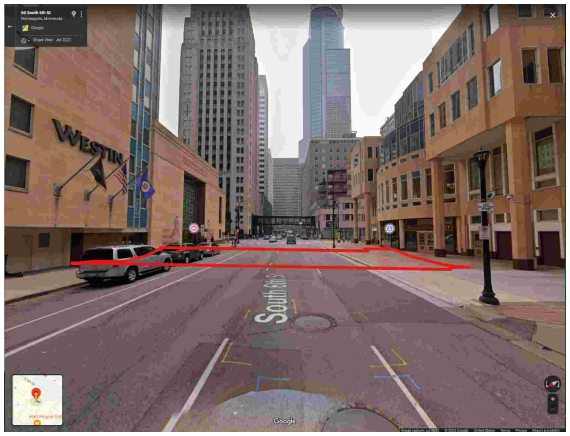
(Figure 5.36b:) Satellite View

Description: High StreetLight count observed for the pedestrians. Potential automatic counter deployment location.

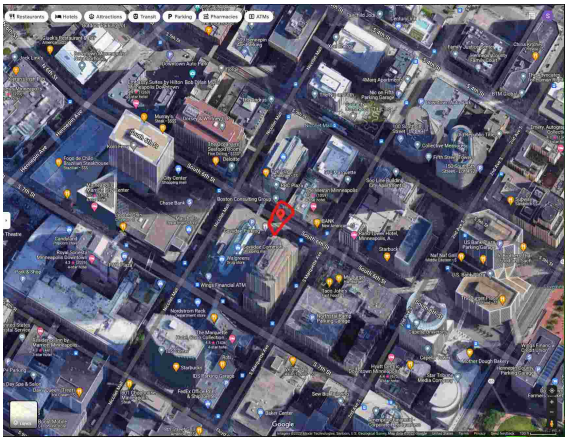
Site ID: 39

Site Location: 69-91 South 6th St, Minneapolis, MN 55402 (Map)

Nearest Intersection/Site: South 6th St at Nicollet Mall and South 6th St at S Marquette Ave



(Figure 5.37a:) Site View



(Figure 5.37b:) Satellite View

Description: High StreetLight count observed for the pedestrians. Separate bike lane and sidewalk present.

Site ID: 40

Site Location: 59-103 Willow St, Minneapolis, MN 55403 (Map)

Nearest Intersection/Site: Willow St at Yale Pl and Willow St at W Grant St

(Figure 5.38a:) Site View

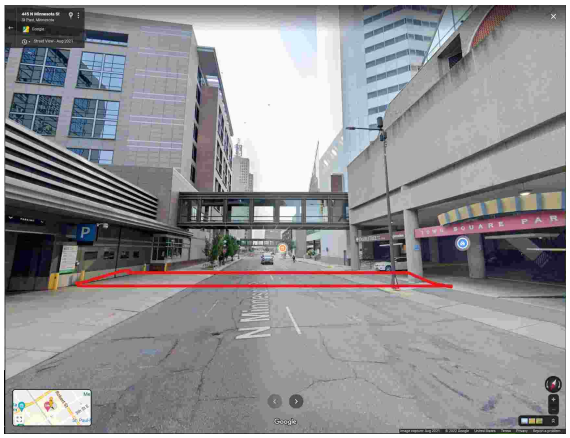
(Figure 5.38b:) Satellite View

Description: High StreetLight count observed for the pedestrians. Close to the Loring Park and Berger Fountain which makes it a scenic route. Sidewalk present. Close to the Minneapolis College.

Site ID: 41

Site Location: 396-426 Minnesota St, St Paul, MN 55101 (Map)

Nearest Intersection/Site: Minnesota St at 7th Pl E and Minnesota St at 6th St E



(Figure 5.39a:) Site View

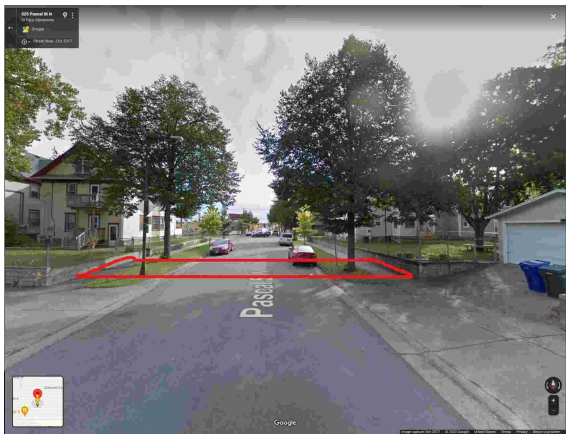
(Figure 5.39b:) Satellite View

Description: High StreetLight count observed for the pedestrians. Sidewalk present.

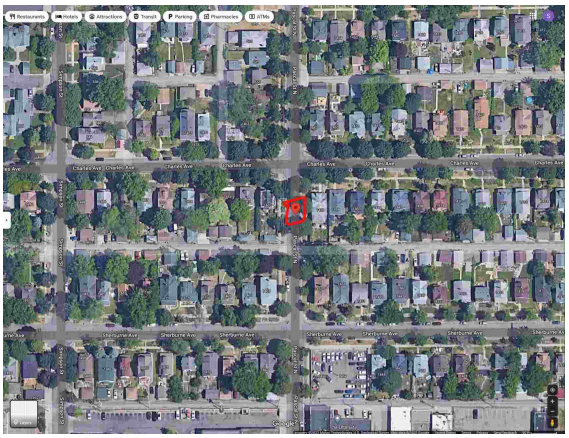
Site ID: 42

Site Location: 526-540 Pascal St N, St Paul, MN 55104 (Map)

Nearest Intersection/Site: Pascal ST N at Charles Ave and Pascal ST N at Sherburne Ave



(Figure 5.40a:) Site View



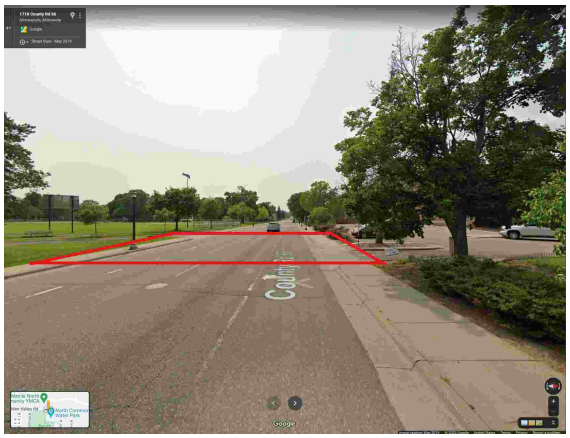
(Figure 5.40b:) Satellite View

Description: High StreetLight count observed for the pedestrians. Sidewalk present. Suburban location. Near LEAP High School.

Site ID: 43

Site Location: 1718-1734 County Rd 66, Minneapolis, MN 55411 (Map)

Nearest Intersection/Site: Golden Valley Rd at Logan Ave N and Golden Valley Rd at N Knox Avenue North



(Figure 5.41a:) Site View

(Figure 5.41b:) Satellite View

Description: High StreetLight count observed for the pedestrians and low StreetLight count for bicyclists. Sidewalk present. Suburban location. Close to the North Community High School. Close to the North Commons Park.

Site ID: 44

Site Location: Powderhorn Lake, Powderhorn, Minneapolis, MN 55407 (Map)

Nearest Intersection/Site: 14th Ave S at E 33rd St

(Figure 5.42a:) Site View

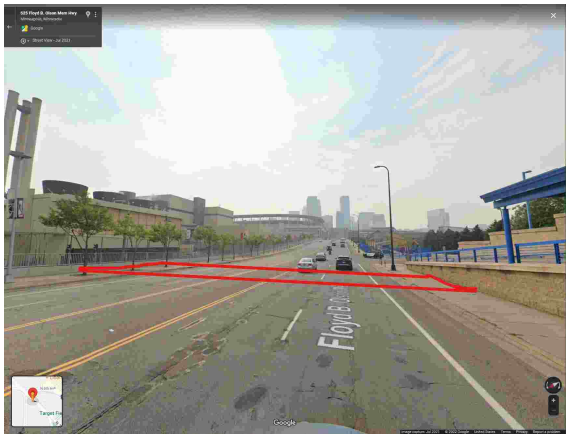
(Figure 5.42b:) Satellite View

Description: High StreetLight count observed for the pedestrians. Sidewalk present. Close to the Powderhorn Lake and park. Close to the Laura Ingalls Wilder School.

Site ID: 45

Site Location: 507-599 N 7th St, Minneapolis, MN 55405 (Map)

Nearest Intersection/Site: N 7th St at N 5th Ave and N 7th St at N 6th Ave



(Figure 5.43a:) Site View

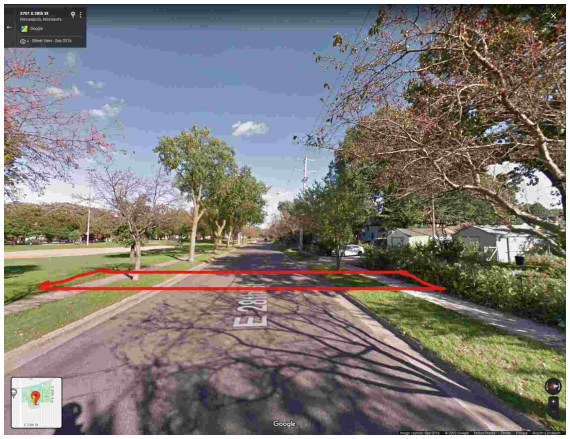
(Figure 5.43b:) Satellite View

Description: High StreetLight count observed for the pedestrians. Sidewalk and bike lane present.

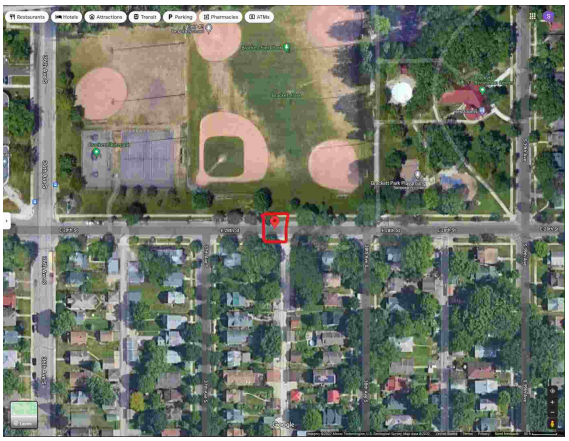
Site ID: 46

Site Location: Brackett Park, E 28th Ave, Minneapolis, MN 55406 (Map)

Nearest Intersection/Site: E 28th St at 37th Ave S and E 28th St at 38th Ave S



(Figure 5.44a:) Site View



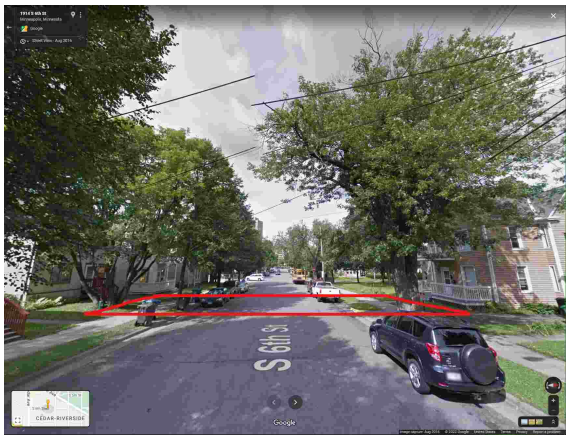
(Figure 5.44b:) Satellite View

Description: High StreetLight count observed for the bicyclists. Sidewalk present. Close to the Hiawatha Collegiate High School. Proximity to Brackett Park makes it a scenic route.

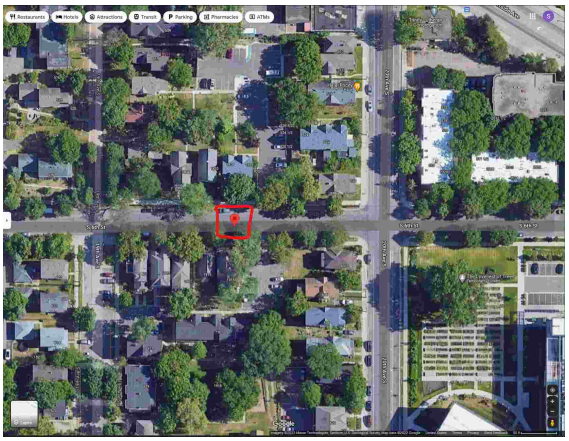
Site ID: 47

Site Location: 1901-1957 S 6th St, Minneapolis, MN 55454 (Map)

Nearest Intersection/Site: S 6th St at 19th Ave S and S 6th St at 20th Ave S



(Figure 5.45a:) Site View



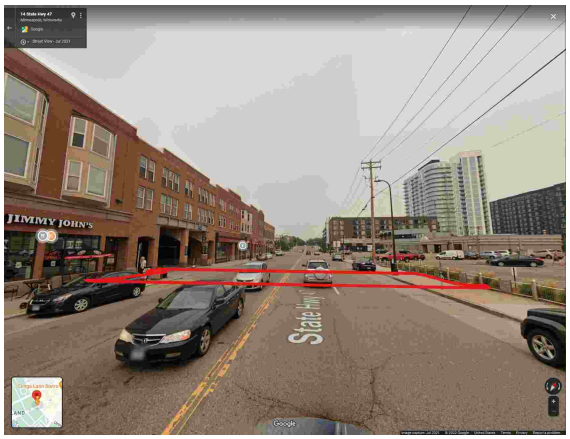
(Figure 5.45b:) Satellite View

Description: High StreetLight count observed for the bicyclists. Sidewalk present. Close to Augsburg University and Carlson School of Management.

Site ID: 48

Site Location: 1-33 University Ave NE, Minneapolis, MN 55413 (Map)

Nearest Intersection/Site: University Ave NE at E Hennepin Ave and University Ave NE at 1st Ave NE



(Figure 5.46a:) Site View

(Figure 5.46b:) Satellite View

Description: High StreetLight count observed for the bicyclists. Sidewalk present. Close to the Nicollet Island. Close to DeLaSalle High School and Marcy Arts and Magnet Elementary School.

Site ID: 49

Site Location: 1498-1414 Fillmore St NE, Minneapolis, MN 55413 (Map)

Nearest Intersection/Site: Fillmore St NE at 14th Ave NE and Fillmore St NE at NE 18th Ave

(Figure 5.47a:) Site View

(Figure 5.47b:) Satellite View

Description: High StreetLight count observed for the bicyclists. Sidewalk present. Close to the Northeast Park. Close to the Yinghua Academy.

Site ID: 50

Site Location: 3127 21st Ave S, Minneapolis, MN 55407 (Map)

Nearest Intersection/Site: 21st Ave S at E 31st St and 21st Ave S at E 32nd St

(Figure 5.48a:) Site View

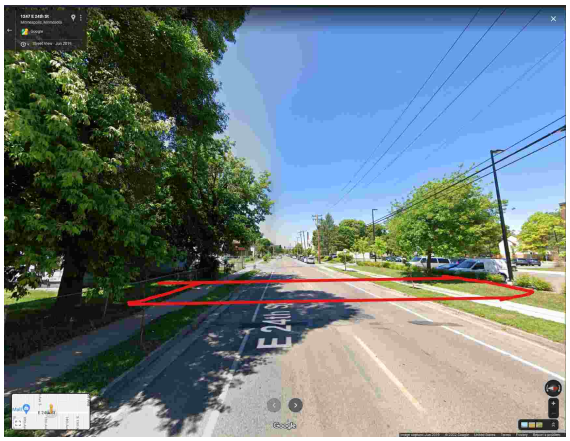
(Figure 5.48b:) Satellite View

Description: High StreetLight count observed for the bicyclists. Close to the South High School. Sidewalk present.

Site ID: 51

Site Location: 1201-1247 E 24th St, Minneapolis, MN 55404 (Map)

Nearest Intersection/Site: E 24th St at 12th Ave S and E 24th St at 13th Ave S



(Figure 5.49a:) Site View

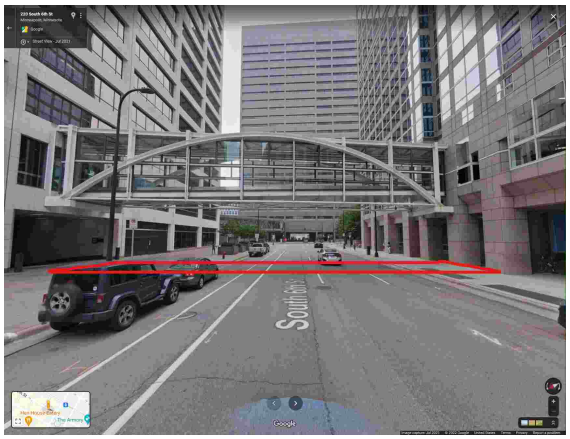
(Figure 5.49b:) Satellite View

Description: High StreetLight count observed for the bicyclists. Sidewalk and bike lane present.

Site ID: 52

Site Location: 220-200 South 6th St, Minneapolis, MN 55402 (Map)

Nearest Intersection/Site: South 6th St at 2nd Ave S and South 6th St at 3rd Ave S



(Figure 5.50a:) Site View

(Figure 5.50b:) Satellite View

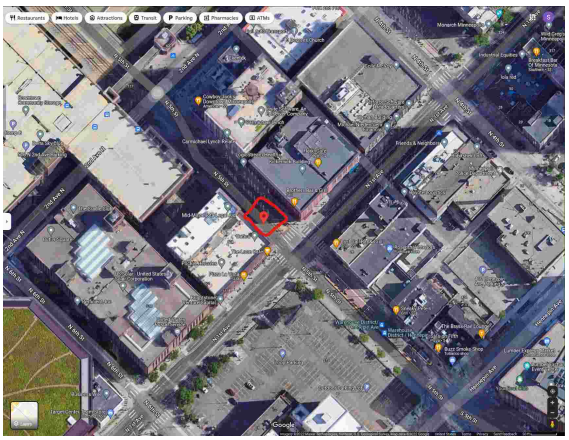
Description: High StreetLight count observed for the bicyclists. Sidewalk and bike lane present. Next to Capella University.

Site ID: 53

Site Location: 512 N 1st Ave, Minneapolis, MN 55403 (Map)

Nearest Intersection/Site: N 5th St at N 1st Ave and N 5th St at 2nd Ave N

(Figure 5.51a:) Site View



(Figure 5.51b:) Satellite View

Description: High StreetLight count observed for the bicyclists. Sidewalk present.

Site ID: 54

Site Location: 310-398 SE 4th St, Minneapolis, MN 55414 (Map)

Nearest Intersection/Site: SE 4th St at 4th Ave SE and SE 4th St at SE 3rd Ave

(Figure 5.52a:) Site View

(Figure 5.52b:) Satellite View

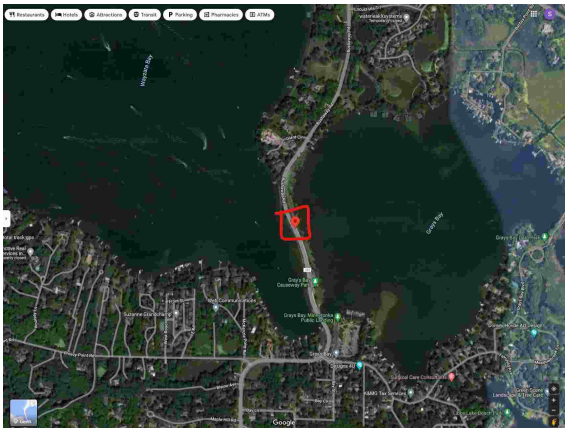
Description: High StreetLight count observed for the bicyclists. Sidewalk and bike lane present. Close to Marcy Arts Magnet Elementary School. Close to the Stone Arch Bridge and Father Hennepin Bluff Park.

Site ID: 55

Site Location: 700 Bushaway Rd, Wayzata, MN 55391 (Map)

Nearest Intersection/Site: Bushaway Rd at Grays Bay Blvd and Bushaway Rd at Giant Cir

(Figure 5.53a:) Site View



(Figure 5.53b:) Satellite View

Description: High Income location. High Pedestrian and Bike from Strava was observed. Close to Gray's Bay and Gray's Bay Causeway Park. Scenic route and close to green space.

Site ID: 56

Site Location: 1750 Covington Ln, Eagan, MN 55122 (Map)

Nearest Intersection/Site: Johnny Cake Ridge Rd at Covington Ln and Johnny Cake Ridge Rd at Sherwood Way

(Figure 5.54a:) Site View

(Figure 5.54b:) Satellite View

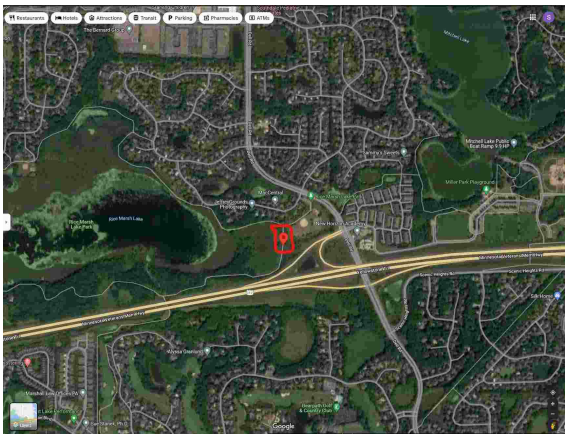
Description: High Income location. High Pedestrian and Bike from Strava was observed. Close to Lebanon Hills Regional Park and RidgeCliff Park. Close to green space and scenic route. Sidewalk present and link is classified as bicycle friendly.

Site ID: 57

Site Location: Eden Prairie, MN 55347 (Map)

Nearest Intersection/Site: Close to Rice Marsh Lake Park

(Figure 5.55a:) Site View



(Figure 5.55b:) Satellite View

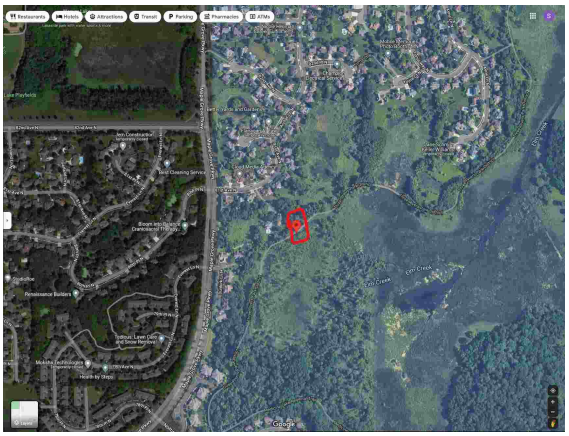
Description: High Income location. High Pedestrian and Bike from Strava was observed. Close to Rice Marsh Lake Park, Bearpath Golf and Country Club and Miller Park Playground. Close to Rice Marsh Lake, Lake Riley and Mitchell Lake.

Site ID: 58

Site Location: Three Rivers Park District, Maple Grove, MN 55311 (Map)

Nearest Intersection/Site: Close to Elm Creek

(Figure 5.56a:) Site View



(Figure 5.56b:) Satellite View

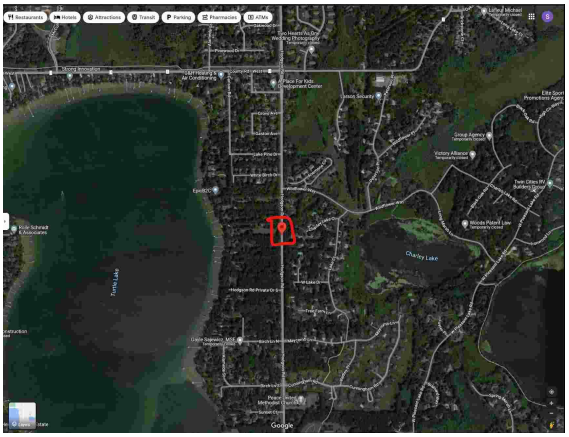
Description: High Income location. High Pedestrian and Bike from Strava was observed. Close to Elm Creek, Weaver Lake and Weaver Lake Community Park.

Site ID: 59

Site Location: 5305-5221 Hodgson Rd, North Oaks, MN 55126 (Map)

Nearest Intersection/Site: Hodgson Rd at Hodgson Rd Private Dr S and Hodgson Rd at Wildflower Way

(Figure 5.57a:) Site View



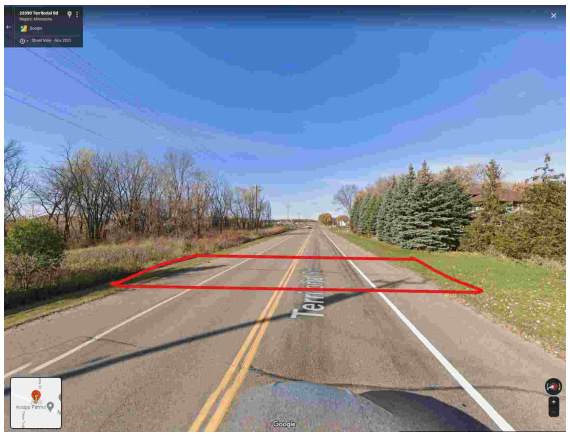
(Figure 5.57b:) Satellite View

Description: High Income location. High Pedestrian and Bike from Strava was observed. No sidewalk or bike lanes present. Close to the Turtle Lake and Charley Lake. Close to McCullough Park and Turtle Creek Open Space. Chippewa Middle School, Peace United Methodist Church are close. Sidewalk present and link is classified as bicycle friendly.

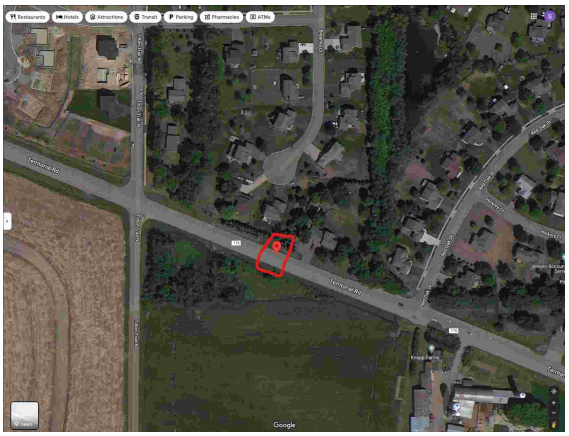
Site ID: 60

Site Location: Rogers, MN 55374 (Map)

Nearest Intersection/Site: Territorial Rd at Tilton Train N and Territorial Rd at Red Oak Dr



(Figure 5.58a:) Site View



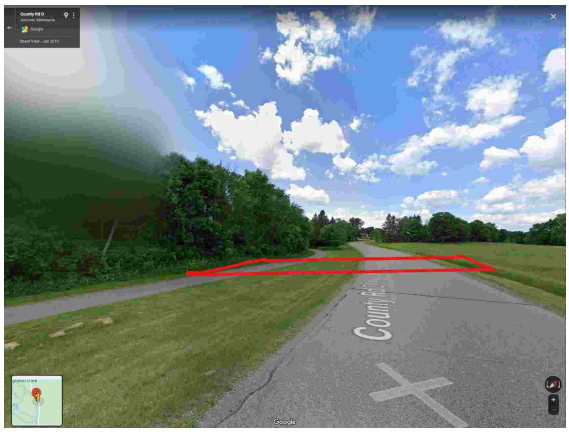
(Figure 5.58b:) Satellite View

Description: High Income location. High Pedestrian and Bike from Strava was observed. No sidewalk or bike lanes present, however large shoulder present and link is bicycle friendly.

Site ID: 61

Site Location: Bunker Hills, Andover, MN 55304 (Map)

Nearest Intersection/Site: Bunker Hills Regional Park



(Figure 5.59a:) Site View

(Figure 5.59b:) Satellite View

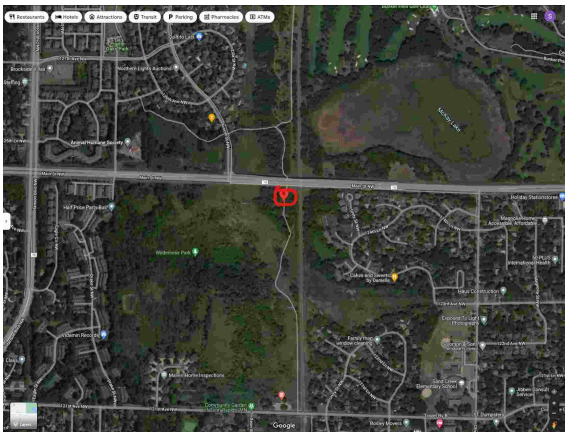
Description: High Income location. High Pedestrian and Bike from Strava was observed. Close to the Bunker Hills Regional Park, Bunker Lake which makes it a scenic route. Close to green space.

Site ID: 62

Site Location: Wilderness Trail, Coon Rapids, Minnesota (Map)

Nearest Intersection/Site: Wilderness Trail at 121st Ave NW and Wilderness Trail at Avocet St Nw

(Figure 5.60a:) Site View



(Figure 5.60b:) Satellite View

Description: High Income location. High Pedestrian and Bike from Strava was observed. Close to the Wilderness Park, McKay Lake and Bunker Hills Golf Club. Close to Sand Creek Elementary School and Hanson Blvd KinderCare.

Site ID: 63

Site Location: South St Paul, MN 55075 (Map)

Nearest Intersection/Site: Close to the Wildflower Levee Park

(Figure 5.61a:) Site View

(Figure 5.61b:) Satellite View

Description: Low Income location. High Pedestrian and Bike from Strava was observed. Close to the Wildflower Levee Park, Pigs Eye Islands and Mississippi River. Close to green space and route is scenic or touristy.

Site ID: 64

Site Location: Coon Rapids Dam (Map)

Nearest Intersection/Site: Coon Rapids Dam

(Figure 5.62a:) Site View

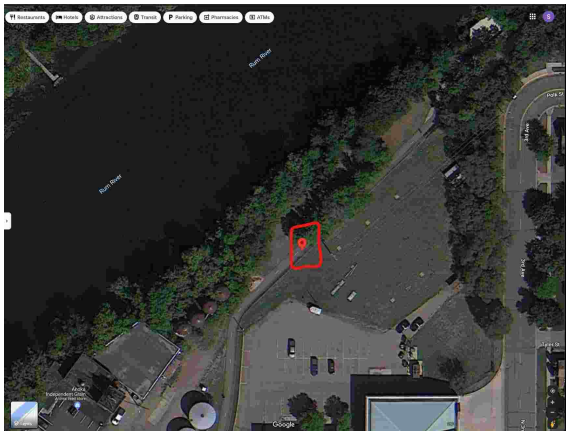
(Figure 5.62b:) Satellite View

Description: Low Income location. High Pedestrian and Bike from Strava was observed. Close to the Close to the Coon Rapids Dam Regional Park, Pavillion No. 3, Rush Creek Regional Trail-Trailhead and Mississippi Gateway Regional Park. Close to green space and site is scenic and touristy.

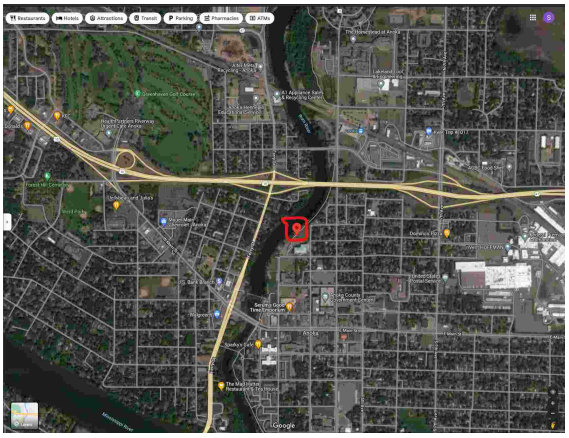
Site ID: 65

Site Location: Anoka, MN 55303 (Map)

Nearest Intersection/Site: Close to Anoka Police Department



(Figure 5.63a:) Site View



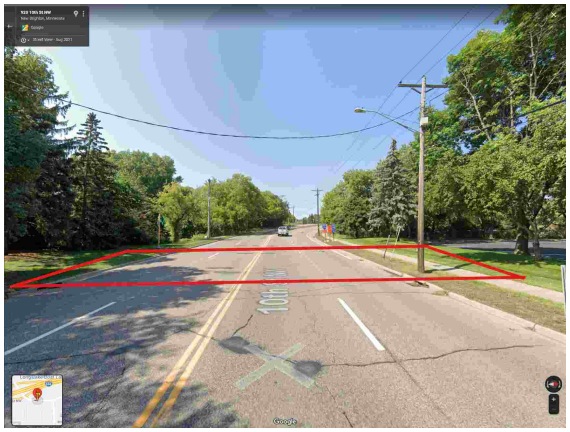
(Figure 5.63b:) Satellite View

Description: Low Income location. High Pedestrian and Bike from Strava was observed. Close to the Riverfront Memorial Park, Rudy Johnson Park, Greenhaven Gold Course and Rum River.

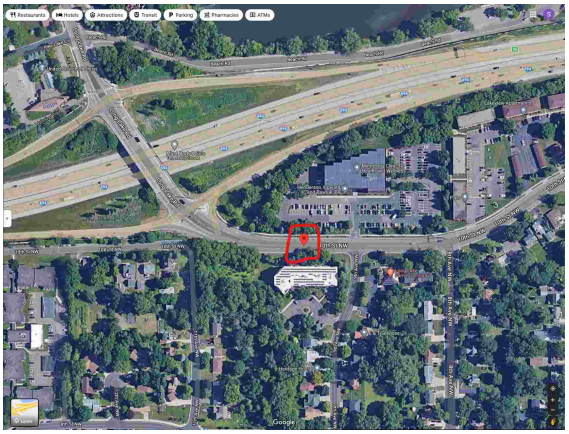
Site ID: 66

Site Location: 922 10th St NW, New Brighton, MN 55112 (Map)

Nearest Intersection/Site: 10th St NW at Long Lake Rd and 10th St NW at at 9th Ave NW



(Figure 5.64a:) Site View



(Figure 5.64b:) Satellite View

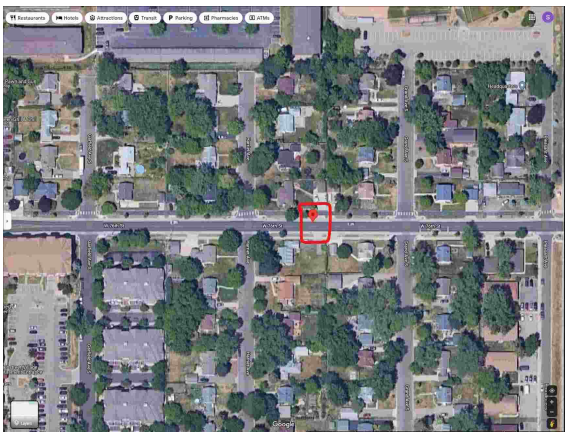
Description: Low Income location. High Pedestrian and Bike from Strava was observed. Close to the Close to the Long Lake, Pike Lake and Hansen Park.

Site ID: 67

Site Location: 499-401 W 76th St, Richfield, MN 55423 (Map)

Nearest Intersection/Site: W 76th St at Harriet Ave and W 76th St at Grand Ave S

(Figure 5.65a:) Site View



(Figure 5.65b:) Satellite View

Description: Low Income location. High Pedestrian and Bike from Strava was observed. Close to the Lincoln Field, Fremont Park and Wood Lake. Close to Central Education Center and Richfield Community Education.

Site ID: 68

Site Location: 10514-10560 Lyndale Ave S, Bloomington, MN 55420 (Map)

Nearest Intersection/Site: Lyndale Ave S at W 106th St and Lyndale Ave S at Hopkins Rd.

(Figure 5.66a:) Site View

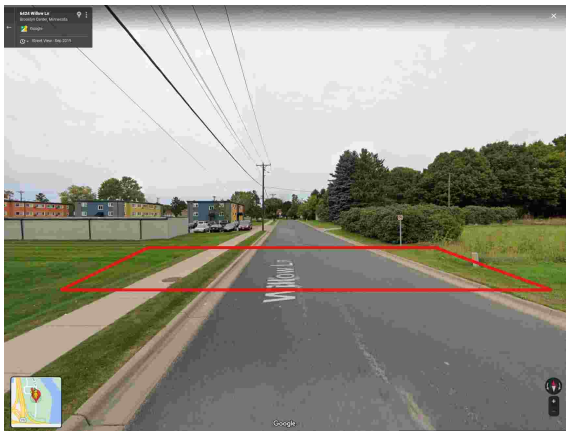
(Figure 5.66b:) Satellite View

Description: Low Income location. High Pedestrian and Bike from Strava was observed. Close to the Bloomington Cemetary. Close to the Oak Grove Middle School and Oak Grove Elementary School.

Site ID: 69

Site Location: 6416-6422 Willow Ln, Brooklyn Center, MN 55430 (Map)

Nearest Intersection/Site: Close to Willow Ln at 65th Ave N



(Figure 5.67a:) Site View

(Figure 5.67b:) Satellite View

Description: Low Income location. High Pedestrian and Bike from Strava was observed. Close to the Firehouse Park, Islands of Peace Regional Park and Mississippi River.

Site ID: 70

Site Location: 2684-2798 N 2nd St, Minneapolis, MN 55411 (Map)

Nearest Intersection/Site: N 2nd St at 28th Ave N and N 2nd St at N 26th Ave

(Figure 5.68a:) Site View

(Figure 5.68b:) Satellite View

Description: Low Income location. High Pedestrian and Bike from Strava was observed. Close to the Farview Recreation Center, Farview Park, Ole Olson Park.

Site ID: 78

Site Location: 5341 Hanrehan Lake Blvd, Savage, MN 55378 (Map)

Nearest Intersection/Site: Hanrehan Lake Blvd at Edgewater Dr and Hanrehan Lake Blvd at Murphy Lake Blvd

(Figure 5.69a:) Site View

(Figure 5.69b:) Satellite View

Description: Low Income location. High Pedestrian and Bike from Strava was observed. Close to Wildflower Early Leaningn Center. Close to the Hanrehan Lake, Murphy Hanrehan Park Reserver and MTB Trailhead. Close to green space and route is scenic or touristy. Potential automated counter installation site.