DATA COLLECTION AND SPATIAL INTERPOLATION OF BICYCLE AND PEDESTRIAN DATA

FINAL PROJECT REPORT

by

Michael Lowry, Ryan McGrath, Seth Cool, Ryan Cook, and Mitch Skiles
University of Idaho

Zhibin Li, Kristian Henrickson, Matthew Dunlap, and Yinhai Wang
University of Washington

Sponsorship
Washington Department of Transportation

for
Pacific Northwest Transportation Consortium (PacTrans)
USDOT University Transportation Center for Federal Region 10
University of Washington
More Hall 112, Box 352700
Seattle, WA 98195-2700

In cooperation with US Department of Transportation-Research and Innovative Technology Administration (RITA)

Disclaimer

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated under the sponsorship of the U.S. Department of Transportation's University Transportation Centers Program, in the interest of information exchange. The Pacific Northwest Transportation Consortium, the U.S. Government and matching sponsor assume no liability for the contents or use thereof.

Technical Report Documentation Page

<table>
<tr><td>1. Report No.
2013-M-UI-0048</td><td>2. Government Accession No.</td><td colspan="2">3. Recipient's Catalog No.</td></tr>
<tr><td colspan="2" rowspan="2">4. Title and Subtitle
Data Collection and Spatial Interpolation of Bicycle and Pedestrian Data</td><td colspan="2">5. Report Date
7/31/2015</td></tr>
<tr><td colspan="2">6. Performing Organization Code
KLK853; 739436</td></tr>
<tr><td colspan="2">7. Author(s)
Michael Lowry, Ryan McGrath, Seth Cool, Ryan Cook, Mitch Skiles, Zhibin Li, Kristian Henrickson, Matthew Dunlap, and Yinhai Wang</td><td colspan="2">8. Performing Organization Report No.
N15-05</td></tr>
<tr><td colspan="2" rowspan="2">9. Performing Organization Name and Address
National Institute for Advanced Transportation Technology
University of Idaho
875 Perimeter Dr. MS 0901
Moscow, ID 83844-0901</td><td colspan="2">10. Work Unit No. (TRAIS)</td></tr>
<tr><td colspan="2">11. Contract or Grant No.
DTRT12-UTC10</td></tr>
<tr><td colspan="2" rowspan="2">12. Sponsoring Agency Name and Address

Pacific Northwest Transportation Consortium (PACTRANS)
University of Washington
More Hall 112, Box 352700
Seattle, WA 98195-2700</td><td colspan="2">13. Type of Report and Period Covered
Research August 2013 to July 31 2015</td></tr>
<tr><td colspan="2">14. Sponsoring Agency Code</td></tr>
<tr><td colspan="4">15. Supplementary Notes</td></tr>
<tr><td colspan="4">16. Abstract

This report provides guidance for improving manual bicycle and pedestrian count programs and recommendations for using moving Bluetooth sensors to collect travel patterns. The project involved five independent studies. (1) An online questionnaire about manual count programs was sent to transportation specialists across the country, including the Bicycle and Pedestrian Coordinator for each state Department of Transportation. (2) Measurement error from manual counts was assessed through a controlled field experiment with 25 counters at five intersections. (3) A new method was developed to evaluate bicyclist exposure to dangerous situations based on manual count data. (4) GPS enabled Bluetooth sensors were provided to fifty-three people to track their location and intercept Bluetooth devices every seven seconds for a 10 day study period. The data was processed to infer trip purpose and method of travel for both the participants carrying the sensors and those they intercepted. (5) Bluetooth sensors were placed on buses to infer the origin-destination patterns of bus riders by analyzing intercepted devices when boarding and alighting the bus.</td></tr>
<tr><td colspan="2">17. Key Words
Bicycle; pedestrian; Bluetooth;</td><td colspan="2">18. Distribution Statement
No restrictions. Copies available from PACTRANS: www.pactrans.org</td></tr>
<tr><td>19. Security Classification (of this report)

Unclassified</td><td>20. Security Classification (of this page)

Unclassified</td><td>21. No. of Pages
116</td><td>22. Price</td></tr>
</table>

Table of Contents

List of Figures

List of Tables

List of Abbreviations

AADB: Annual Average Daily Bicyclists
AADT: Annual Average Daily Traffic
AFC: Automated Fare Collection
AIC: Aikaike Information Criterion
APE: Average Percent Error
BMP: Bicycle Master Plan
BMT: Bicycle Miles Traveled
CMF: Crash Modification Factors
COD: Class of Device
DOT: Department of Transportation
FIAB: Funf in a Box
GIS: Geographic Information System
GPS: Global Positioning System
MAC: Media Access Control
NBPD: National Bicycle and Pedestrian Documentation Project
PacTrans: Pacific Northwest Transportation Consortium
PercBus: Percent of the trip that occurred on a bus route
PercTrail: Percent of trip that occurred on a trail
SPF: Safety Performance Functions
UI: University of Idaho
UW: University of Washington
UWTS: University of Washington Transportation Services
UWMC: University of Washington Medical Center
WSDOT: Washington State Department of Transportation

Acknowledgments

The authors would like to recognize the important contributions of the late Dr. Michael Dixon. Mike helped craft the proposal for this project, provided insight and enthusiasm during the field experiments, and led an unfinished part of the project that was investigating an innovative new technology to count bicycles using pressure sensitive pneumatic tubes. As our dear friend and colleague, he will be missed.

Executive Summary

Cities and state Departments of Transportation (DOTs) struggle to collect and use bicycle and pedestrian data in effective and meaningful ways. Monitoring non-motorized traffic is still in its infancy compared to the well-developed practice of monitoring motorized traffic. The goal of this project was to provide guidance for improving manual bicycle and pedestrian count programs and investigate the feasibility of new, cutting-edge sensor technology. To this end, five independent studies were pursued to:

1. Evaluate the state-of-the-practice of manual bicycle and pedestrian count programs,
2. Assess the error associated with manual count programs,
3. Develop a process for using manual count data for safety analysis,
4. Explore the feasibility of using moving Bluetooth sensors to collect bicycle and pedestrian data, and
5. Investigate the potential for installing Bluetooth sensors on public buses to gain better understanding of pedestrian interaction with public transportation.

For the first study, an online questionnaire was sent to transportation specialists across the country, including the Bicycle and Pedestrian Coordinator for each state Department of Transportation (DOT). There were 71 responses received from 25 states. Eleven communities were contacted for phone interviews. Information about community-volunteers, scheduling and logistics, data collection techniques, and reasons for conducting manual counts were attained through the questionnaire. The respondents reported many communities are not using automatic counters, not using adjustment factors, counting infrequently, for short two hour periods, and rarely for more than one day. Although occasionally or regularly conducting manual counts has a variety of benefits, it is not clear how the resulting snapshot counts can be used for some of the reasons the survey respondents claimed were motivating their count programs. For example, it is not readily apparent how such counts can be used for analyzing trends year-to-year. On the other hand, some of the stated reasons for conducting manual counts seem compatible with the activities underway. For example, a popular reason cited by survey respondents was to “improve community-wide infrastructure, signage, and paint markings” and another reason given was to “provide opportunities for public engagement and outreach.”

For the second study, measurement error was assessed through a controlled field experiment with 25 counters at five intersections. Lower measurement error rates were observed when using a 4-movement data collection technique compared to a more complicated 12-movement technique; however, the differences were not statistically significant. The overall median absolute percent error for the 12-movement technique was 27% and 7% for bicyclists and pedestrians, respectively. The field experiment showed no consistent increases in measurement error when counters were assigned to collect additional information about passing travelers, such as sex of traveler or whether bicyclists wore helmets.

For the third study, a new method was developed to evaluate bicycle infrastructure plans in terms of exposure to dangerous situations. A typology of 23 dangerous situations for bicyclists was identified through a literature review. Exposure indicators were defined. A geographic information system (GIS) tool was created to analyze some of the dangerous situations and provide a means to compare proposed improvement projects. The tool and process are demonstrated for a case study community. The case study results suggest that the proposed bicycle master plan would decrease exposure for various dangerous situations.

For the fourth study, we explored the feasibility of using moving Bluetooth sensors to collect bicycle and pedestrian data. Fifty three people were provided a smartphone app that tracked their location via global positioning systems (GPS) every seven seconds for 10 days. The app also collected, via Bluetooth, the media access control address (MAC address) of any intercepted device within 100 feet. The intent of the study was twofold, first to explore the potential of using this passive data collection technology to replace traditional travel diaries. The study also provided a means to collect a richer data set of travel patterns by matching intercepted MAC addresses across the city, perhaps as a precursor to the data that might be collected in the future through vehicle-to-vehicle communications. Every night the participants logged into a website to verify information about their travel. They were asked to confirm trip segments, trip purpose, and trip mode, all of which had been inferred through algorithms we developed for the study. Our model predicted mode with 77% accuracy and purpose with 54% accuracy. Over 3,000 devices were intercepted via Bluetooth and recorded over 300,000 GPS points. By matching MAC addresses across town, our data processing algorithms identified over 600 trips, most of which were determined to be pedestrian trips on the University of Idaho campus.

For the fifth study, smartphones with the GPS/Bluetooth app were placed on buses traveling between the University of Washington campus and South Lake Union Neighborhood in Seattle. During the study period 11,041 devices were intercepted. After data processing we determined that 403 bus riders were intercepted. Boarding and alighting locations were determined for the intercepted riders. This study showed that the number of passengers carrying active wireless devices is sufficient to provide a sizable sample of the population, and that the boarding and alighting location of passengers can be inferred from detection time and GPS location.

Chapter 1 Introduction

1.1 Problem Statement

Cities and state Departments of Transportation (DOTs) struggle to collect and use bicycle and pedestrian data in effective and meaningful ways. Monitoring non-motorized traffic is still in its infancy compared to the well-developed practice of monitoring motorized traffic. In the last few years there has been an explosion of citizen-volunteer programs to manually count bicycle and pedestrian volumes. These “grass-roots” efforts to improve bicycle and pedestrian data collection have been motivated by concerns about the environment, community livability, obesity, and personal safety. There are exciting opportunities for using these data for transportation planning purposes in addition to the benefits of increased citizen engagement, but there are also a number of challenges because the count durations are usually very short (typically only two hours), sporadic, and sometimes inconsistent in terms of when, where, and who is collecting the data. In the last few years there have also been significant advances in sensor technology which have greatly increased the possibilities for automated data collection; but these methods also pose a number of challenges that need to be overcome, including counting error due to occlusion, cost, and the limitations of low-fidelity information compared to a human counter that can collect information about direction of travel, sex of traveler, helmet use, etc.

1.2 Project Objectives, Research Approach, and Report Organization

The overall goal of this project was to provide guidance for improving manual bicycle and pedestrian count programs and investigate the feasibility of new, cutting-edge sensor technology. To this end, the following five project objectives were pursued:

- Evaluate the state-of-the-practice of manual bicycle and pedestrian count programs,
- Assess the error associated with manual count programs,
- Develop a process for using count data for safety analysis,
- Explore the feasibility of using moving Bluetooth sensors to collect bicycle and pedestrian data, and
- Investigate the potential for installing Bluetooth sensors on public buses to gain better understanding of pedestrian interaction with public transportation.

This project was led by researchers at the University of Idaho (UI) and the University of Washington (UW). The UI team worked on objectives 1 to 4 and the UW team worked on objective 5.

The chapters of this report address each objective as follows. Chapter 2 describes the evaluation of manual bicycle and pedestrian count programs. An online survey was sent to the Bicycle and Pedestrian Coordinator for all 50 state DOTs. Chapter 2 also presents the results of an assessment of the error associated with manual count programs. Chapters 3 demonstrates how manual counts can be spatially extrapolated to estimate bicycle demand throughout a community. Chapter 4 describes the new process that was developed to use the demand estimation for safety analysis. Chapter 5 presents the UI teams study of moving Bluetooth sensors. The moving sensors were given to 50 study participants for a week in the fall of 2012 and spring of 2013. Chapter 6 describes the UW team's study of moving Bluetooth sensors. Sensors were placed on buses as they traveled around the UW campus and to downtown Seattle.

Chapter 3 Evaluation of Manual Count Programs

2.1 Introduction

Walking and bicycling are increasingly recognized as sustainable modes of transportation (Buehler and Pucher 2011). Throughout the United States there has been a surge in community-volunteer manual count programs to collect bicycle and pedestrian traffic volume data. The volunteers are provided clipboards, clickers, computer tablets, or some other method to tally volumes of non-motorized travelers that pass by a location during a certain time period. Communities are using manual count programs as a low cost and flexible way to gather data while providing an engaging means for public participation.

In the United States, manual count programs are conducted and coordinated by various entities, such as cities, counties, state departments of transportation, and advocacy groups. For example, the Washington State Department of Transportation (WSDOT) has supervised a community-volunteer count program since 2008. The most recent WSDOT count involved 38 jurisdictions, totaling more than 409 observations at intersections (Traffic Monitoring Guide 2013). The California Department of Transportation recently funded an online data-clearinghouse for local communities to upload and share data from their independent count programs (Huff & Brozen, 2014).

The National Bicycle and Pedestrian Documentation Project (NBPD), a collaborative effort by Alta Planning+Design and the Institute of Transportation Engineers, is a pioneer in manual counts (Los Angeles County Bicycle Coalition 2013). The NBPD webpage provides count forms, recommendations for choosing count locations, and a presentation for training community members. The NBPD recommends conducting counts four times per year, once for every season, for two hours in the morning (7:00am-9:00am) and two hours in the evening (4:00pm-6:00pm).

Manual count programs continue to proliferate as communities increasingly turn to non-motorized transportation as a means to establish healthy, livable communities (Forsyth and Oakes 2015; Ryan 2013). Consequently, there is substantial work underway to better understand the issues and challenges related to bicycle and pedestrian traffic counts (Griffin et al. 2014). Much of the work has focused on understanding daily and annual patterns and the impact of temperature and weather (Aultman-Hall et al. 2009), with the goal of developing adjustment factors that can be used to adjust (or inflate) short-duration counts to annual volumes (El Esawey et al. 2013; Figliozzi et al. 2014; Lindsey et al. 2013). Communities with permanent automatic counters can create community-specific adjustment factors from a year's worth of continuous data. Alternatively, communities can use factors developed elsewhere, such as the generic factors provided on the NBPD website (NBPD 2014). However, Nordback et al. (Nordback et al. 2013) concluded that counts obtained over a very short duration (i.e. only a few hours) cannot be adjusted reliably to estimate annual volumes and they warn against using counts from less than one week.

Count error (i.e. measurement error) incurred by volunteer-counters is another possible challenge associated with manual count programs. Diogenes et al. (Diogenes et al. 2007) conducted an error assessment for pedestrian counts collected by contracted staff from a "private consulting firm specializing in data collection" and reported error rates ranging from 8% to 25%. Their study did not investigate if community-volunteers would exhibit a higher error rate or if error increases when volunteers are expected to collect additional information, such as direction of travel, sex of the traveler, or the approximate age of the traveler. The implications that effect measurement error depend largely on the reasons for collecting the count data.

There is little known about why communities are conducting manual counts, nor how volunteers are being trained and instructed. The Federal Highway Administration sponsored a study to assess the state-of-the-practice of bicycle and pedestrian data collection to help craft the new chapter on non-motorized transportation in the new Traffic Monitoring Guide (USDOT 2011). However, the study was not focused on manual count programs and most of the final report deals with automatic counting technologies. The study included a poll of transportation specialists that reported the most common reason for collecting bicycle and pedestrian data was project evaluation (before and after studies) and safety analyses. It is not clear if manual counts that involve community-volunteers are done for the same reasons. Schneider et al. (Schneider et al. 2005) conducted a similar state-of-the-practice questionnaire concerning all aspects of bicycle and pedestrian data collection, including automatic count technology, intercept surveys, and infrastructure inventories. They profiled 29 communities, 6 of which reported conducting manual counts, but it seems they were referring to counts done by staff, not volunteers. In fact, one community spokesperson noted two staff members are typically assigned to count at low volume intersections and three staff members are often assigned to heavy traffic intersections.

Therefore, the intent of this study was twofold; (1) to gain a better understanding of how and why communities are conducting manual counts and (2) to assess the error that might occur for different manual count techniques. An online questionnaire was emailed to transportation specialists across the USA and received 71 responses from 25 states. Eleven communities were contacted for phone interviews. The error assessment involved a controlled field experiment with 25 counters at 5 intersections.

2.2 Background

There are various techniques a community can use to count bicyclists and pedestrians. The most basic approach is to simply tally all non-motorized travelers who pass by a certain point, regardless of method of travel or direction of movement. Another approach is to keep a separate tally for “bicyclists”, “pedestrians”, and “other” for particular travel movements. Figure 2.1 shows various movements that can be captured by volunteer-counters. Screenline counts are conducted mid-segment, while the other techniques collect data at intersections. The NBPD provides forms for a screenline count and a 12-movement count, while WSDOT uses the 4-movement leaving technique.

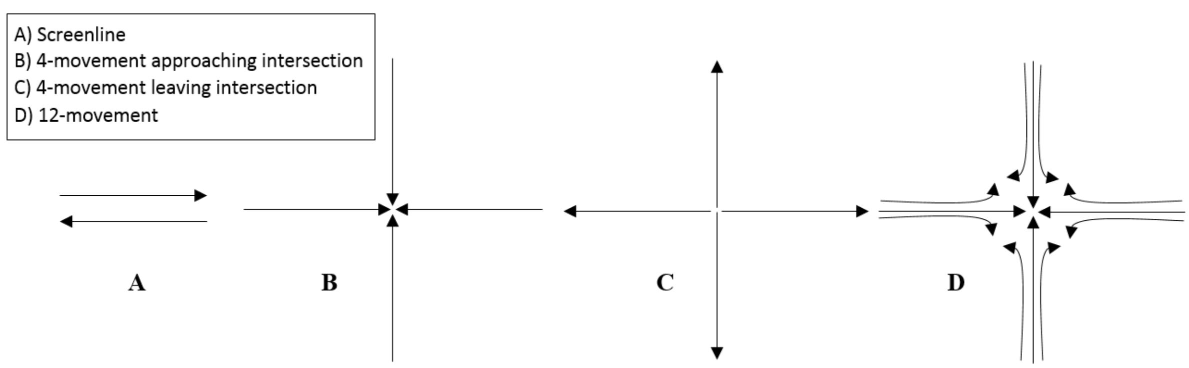


Figure 2.1 Schematic of different count techniques.

Community volunteers are often assigned to collect additional information, such as sex of the traveler, the approximate age, and group size. The additional information can be tallied using unique markings or by providing separate boxes on the count form. The NBPD forms provide separate boxes for tallying male and female pedestrians, and suggest tallying bicyclists on sidewalks. Schneider et al. (2005) reported that one community profiled in the their study has assigned their staff to collect additional information about in-line skaters, scooters, joggers, people walking dogs, people walking with strollers, wheelchair users, bicyclists with helmets, and bicyclists riding in the proper direction.

2.3 Methods

The first part was an online questionnaire sent to practitioners across the country. The second part of this study was a controlled field observation to assess the measurement error exhibited by a group of volunteer counters. Both parts of the study were deemed exempt by the University of Idaho Institutional Review Board.

2.3.1 Practitioner Questionnaire

A 26-item questionnaire was created to gain a better understanding of why communities are conducting manual counts and how the counts are being collected. The questionnaire inquired about a variety of topics germane to manual bicycle and pedestrian count programs, including characteristics of community volunteers, count locations and other logistics, count forms and movements counted, availability of automated count equipment, use of adjustment factors and other resources, purposes of the counts, additional information collected while counting, and challenges and suggestions for improving manual counts. The questions (likert scale, open-ended, check box, and multiple choice) were based on the research team’s experience working with bicycle and pedestrian count programs, and feedback from professional consultants. A web-link to the questionnaire was emailed to the individual(s) that were identified as the Bicycle and Pedestrian Coordinator for every State Department of Transportation (DOT). The email requested the coordinators to forward the questionnaire to anyone working with manual bicycle and pedestrian count programs. Questionnaire respondents were invited to pass the questionnaire on as well (i.e. snowball sampling). A few selected communities were contacted by the researchers for open-ended phone interviews.

2.3.2 Error Assessment

The error assessment was conducted in Moscow, Idaho (population 24,500), a community with a large amount of bicycling and walking due to the presence of the University of Idaho. The community has conducted manual counts every autumn for the past four years (2010-2013), involving approximately 30 community-volunteers annually. The pool of volunteers includes university students, retired community members, adolescents, and advocates for non-motorized transportation.

In the present study, twenty-five university students were recruited to perform the counts. This participant pool was selected in order to replicate a typical group of counters for the community, and to minimize educational and age differences amongst the counters. None of the count participants had participated in any of the community's official count events. Count participants engaged in classroom training sessions using the NBPD training presentation (*11*). The participants also performed a 1 hour practice count at their assigned location prior to the study.

Count participants were randomized into five equal groups and assigned to count at a high volume intersection on two different occasions between 7:00-9:00 AM or 4:00-6:00 PM on a Tuesday, Wednesday, or Thursday over a two week time period during the spring. Count groups were assigned to use either a 4-movement leaving form or a 12-movement form during their manual counts, the order of which was randomized (Figure 2a and Figure 2b).

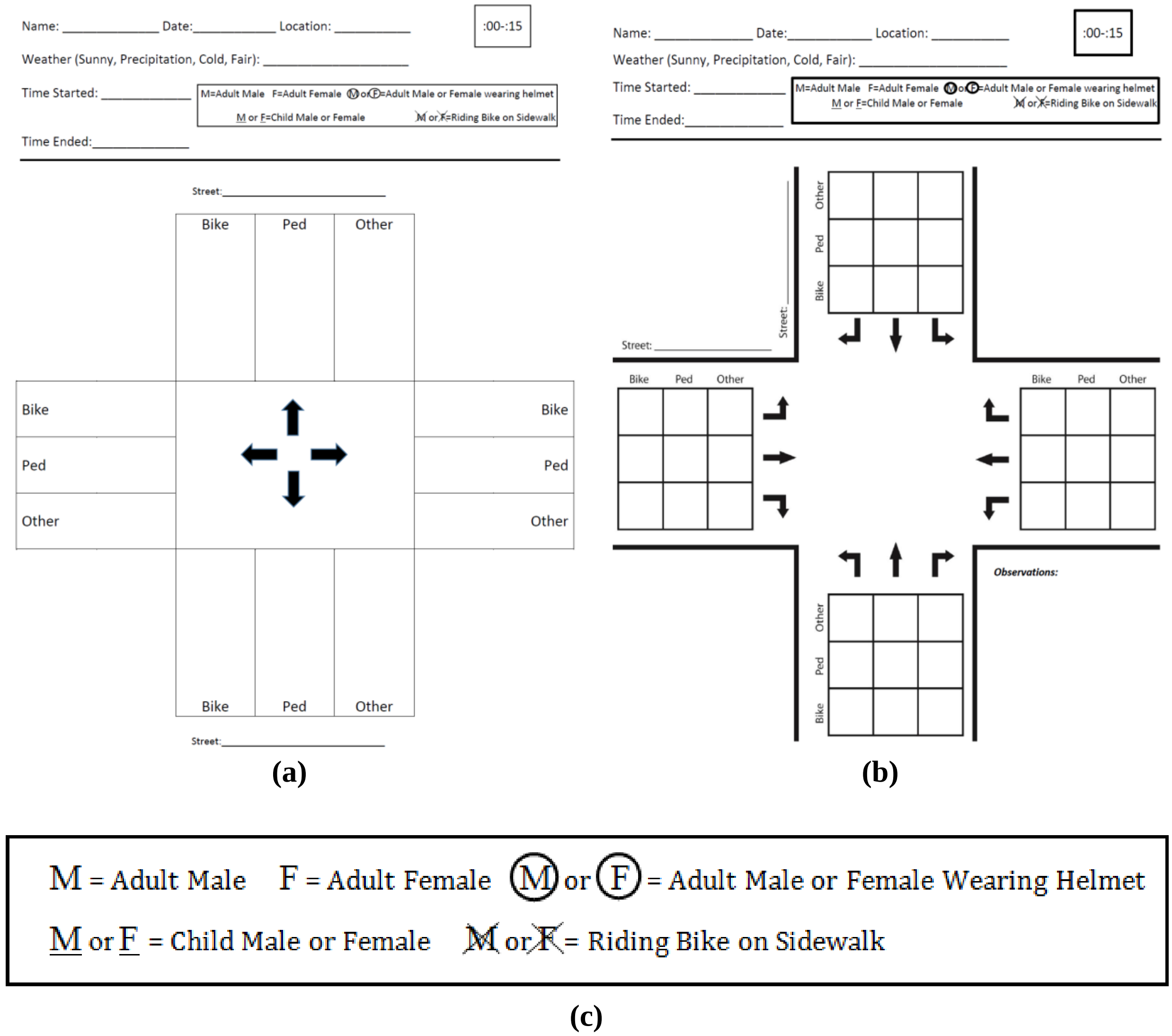


Figure 2.2 Counting forms used for error assessment (a) 4-movement (b) 12-movement (c) enlarged tally key for assignment 5.

If a group was assigned to a particular count form during their first session, they used the other form during their second session, so no group used the same form for both of their counts. Within each group, each counter was randomly assigned to record additional information about the travelers passing through the intersection. The five assignments were:

- Assignment 1: method of travel and direction of travel
- Assignment 2: method of travel, direction of travel, and sex of traveler

- Assignment 3: method of travel, direction of travel, sex of traveler, and bicyclist helmet use
- Assignment 4: method of travel, direction of travel, sex of traveler, bicyclist helmet use, and if a bicyclist was riding on sidewalk or street (legal use)
- Assignment 5: method of travel, direction of travel, sex of traveler, bicyclist helmet use, legal use, and if the person traveling through the intersection was an adult or child.

The assignment information and ordering was determined based on observed frequency of each assignment during pilot testing. A block randomization technique was used for all randomizations (Kang et al. 2008). Figure 2c shows the key for the forms used in this study; count participants only used symbols that were applicable to their assignment.

Two video recorders were placed at intersections to record all travelers. After the field count, two researchers tallied counts based on the video recordings and reconciled any differences between them by returning to the video segments. Absolute percent error (APE) was calculated as follows

$$Absolute\ percent\ error = \left|\frac{T_o - T_v}{T_v}\right| * 100$$

where T_o is the count observed in the field and T_v is the count obtained from the video recording.

Data are reported in the results as median values and interquartile ranges because the raw data were not normally distributed. Wilcoxon-Mann-Whitney tests were then used to compare APE for the bicycle and pedestrian counts for the 4-movement leaving and 12-movement techniques. An α of 0.05 denotes the presence of statistically significant differences. SAS 9.3 software was used for all statistical analyses.

At the conclusion of the study, count participants were asked to complete a short debriefing survey to gather additional feedback about their experience.

2.4 Results

2.4.1 Questionnaire and Interview Findings

The online questionnaire received 71 responses from 25 states (Arizona, California, Connecticut, Georgia, Idaho, Indiana, Kansas, Louisiana, Maine, Massachusetts, Michigan, Minnesota, North Carolina, North Dakota, Nebraska, New Hampshire, New Mexico, Nevada, Ohio, Rhode Island, South Carolina, South Dakota, Vermont, Washington, and West Virginia). A little more than half of the respondents (n=47) said they have been involved with manual count programs in more than one community. These respondents primarily worked for Metropolitan Planning Organizations (32%), State DOT (23%), or consulting firms (11%). Fifty-three percent of them identified their organization as being highly involved with count programs, with the other 47% reporting their organization as having low involvement.

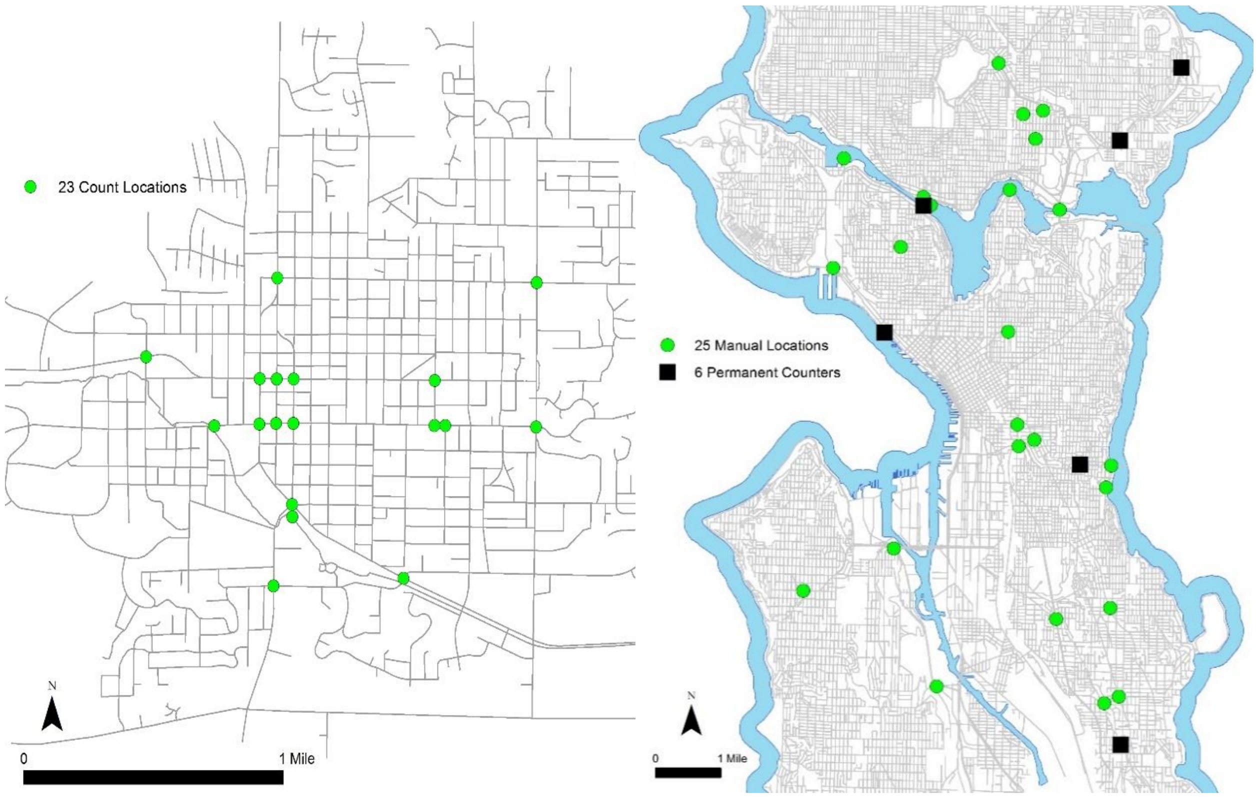
The rest of the respondents (n=24) said they have been involved with manual counts in only one community. The single-community respondents were then asked additional questions; eleven communities were contacted for phone interviews. The single-community respondents came from communities of varying population size: less than 25,000 (25%), 25,000-50,000 (21%), 50,000-100,000 (17%), and greater than 100,000 residents (37%). Sixty-seven percent of single-community respondents identified their involvement as being "highly involved" with manual bicycle and pedestrian counts, while the other 33% reporting their involvement as "low." Most of the single-community respondents said their count programs began recently, with responses such as 2012 (21%), 2009 (17%), 2010 (8%), and 2007 (8%).

An additional 21 responses were excluded from the analyses because they reported not having any involvement with a manual bicycle and pedestrian count program.

2.4.2 Community-Volunteer Characteristics and Count Locations

The single-community respondents (n=24) were asked questions about their community's most recent count event, including questions seeking to better understand the composition of community volunteers. The majority of respondents (83%) suggested that only a few or none of their volunteers were minors. About half of the respondents (46%) said seniors participated in their most recent count and 4% said that seniors comprised the majority of the counters. Most of the respondents (71%) reported the involvement of "first time" counters and 21% said "first time" counters were the majority of their volunteers. Thirty-three percent of the respondents stated the majority of their volunteers were from bicycle and pedestrian advocacy groups. Many of the respondents (63%) reported paid staff was also involved with the count, but only 8% reported their community participants were paid.

The single-community respondents reported counting an average of 15 locations for their most recent count, lasting for 4 hours per day, presumably for 2 hours in the morning and 2 hours in the afternoon. In order to get a sense of geographic spread, a few of the communities contacted by phone were asked for additional information about their count locations. For illustration, Figure 2.3 shows the count locations for the most recent count for a small community and large city. Both communities had about 25 locations. The small community, Moscow, Idaho, exhibits a much higher density of count locations per square mile. The large city, Seattle, Washington, has the advantage of six permanent automatic counters (one permanent counter only detects bicyclists and the other five were installed in 2013).



(a) **(b)**

Figure 2.3 Count locations for (a) Moscow, Idaho and (b) Seattle, Washington.

2.4.3 Count Techniques and Schedule

All respondents were asked about the movements typically counted. It was reported that the following count techniques are used: screenline (38%), 4-movement approaching the intersection (16%), 4-movement leaving the intersection (7%), 12-movement (30%), total count without information about movement direction (16%), and some other movement that was not presented as an option (1%). Eleven percent of respondents did not know what types of forms are used or what movements are typically counted. Most of the respondents said their counters were assigned to collect additional information. Table 2.1 shows the additional information collected for each count technique.

Sixty percent of the respondents reported their most recent count or the typical count lasted for just one day, 9% reported durations of two consecutive days, 17% reported three consecutive days, and 13% did not know. Respondents stated that counts are typically conduct on the following days: Monday (21%), Tuesday (69%), Wednesday (68%), Thursday (69%), Friday (20%), Saturday (45%), Sunday (17%), and 13% of respondents suggested they did not know. Counts occur during the autumn (63%), spring (33%), summer (8%), winter (0%), or on demand (13%). Eight percent of the respondents were unsure in what season their counts typically took place.

Table 2.1 Additional information collected with different count techniques according to survey respondents.

Information collected	Screenline (%, n=27)	4-movement approaching (%, n=11)	4-movement leaving (%, n=5)	12-movement (%, n=21)	Count without movement information, (%, n=11)
Count Only	4	32	27	16	13
Sex	23	16	18	18	33
Helmet	26	16	27	13	20
Legal Use	23	21	18	29	7
Age	21	16	9	18	20
Other	4	0	0	5	7

Note: Some respondents reported using multiple count techniques.

2.4.4 Adjustment Factors, NBPD Resources, and Training

Ninety-two percent of the single-community respondents reported not using automatic counters and none of the remaining 8% who said they use automatic counters said they use them to create temporal adjustment factors. Likewise, only a few respondents (19%) who have worked with multiple communities reported using automatic counters and none said they use the automatic counters to create adjustment factors. Furthermore, only 4 communities reported using the NBPD adjustment factors.

Other NBPD resources are commonly used. Seventy two percent of the respondents reported using NBPD count forms, recommendations about when and where to count, or training material. Respondents said training community-volunteers typically involved providing instructional handouts (51%), giving a PowerPoint presentation (38%), or conducting a live demonstration (28%).

2.4.5 Purpose of the Count

Respondents were asked about the reasons and purpose of conducting manual counts in three different ways. First, they were asked through an opened-ended question to provide reasons for conducting manual counts. This was a required question, so every respondent had to provide an answer to be able to proceed with the questionnaire. A few respondents simply said they collect the data for the purpose of providing it to someone else, such as a State DOT or advocacy group. Many of the respondents provided detailed and thoughtful responses. Responses included, planning for infrastructure improvements, analyzing trends year-to-year, and applying for grants.

Later in the questionnaire, respondents were asked a second time what they use count data for; they were shown a list of reasons for conducting manual counts and asked to identify which reasons from the list they believed to be valid and useful. This was followed by the third method of inquiry whereby respondents were shown the same list again and asked to select one reason as the most important reason. Table 2.2 shows the results from these questions.

The single-community respondents reported submitting their data to advocacy groups (33%) and/or the State DOT (29%). The respondents were not asked the reason for submitting data to other organizations.

Table 2.2 Survey respondent's reasons for collecting manual count data.

Reasons for conducting counts	Selected as Most Important Reason (%)	Identified as a Valid and Useful Reason (%)
Improving community-wide infrastructure, signage, and paint markings	27	66
Analyzing trends year-to-year	23	79
Raising awareness about bicycle and pedestrian activity	18	89
Assessing safety concerns	15	65
Providing public engagement and outreach	5	69
Applying for grants	4	76
Determining the percentage of bicyclists and pedestrians that travel throughout the community	4	69
Validating travel demand models	3	59
Improving location-specific infrastructure, signage, and paint markings	0	78
Measuring physical activity	0	47
Making adjustment to traffic signal timing	0	44
Making comparisons with other communities	0	40
Assessing economic vitality	0	36

2.4.6 Challenges and Suggestions for Improvement

Recruiting volunteers is a problem, as 47% of respondents mentioned needing more people to participate in manual counts when asked with an open-ended question about the challenges facing their manual count programs. Other responses included concerns about count error (22%), lack of funding for manual count programs (18%), the time involved to train volunteers and organize count events (16%), and the limitations of short-duration counts (12%).

When asked about the quality of volunteer training efforts, 38% of single-community respondents thought their current training techniques were sufficient, and 25% stated training should be improved to ensure quality control. This differs from the respondents who have been involved with multiple communities—only 17% thought training was sufficient, while 45% thought more and better training is needed to ensure accurate counts.

2.4.7 Observed Error

Across the five intersections, 348 bicyclists and 1,838 pedestrians were counted (only 8 were children). The overall APE for counting bicyclists and pedestrians with the 4-movement leaving technique was 11% and 4%, respectively; whereas, the overall median absolute error percentage for counting bicyclists and pedestrians with 12-movement technique was 27% and 7% respectively. The APE differences between the 4- and 12-movement techniques were significantly different for bicycles ($P<0.05$), but not pedestrians ($P=0.051$). After removing the data for two participants who admitted to erroneously counting bicyclists that walked their bikes as “pedestrians”, the difference between the 4- and 12-movement techniques was no longer statistically significant for bicycles ($P=0.069$). In other words, despite a greater APE for the 12-movement technique, the apparent differences might be due to random chance. Table 2.3 presents error rates between the 4-movement leaving technique and 12-movement technique for each assignment level.

Table 2.3 Field study median absolute percent error.

Count Assignment	Median Absolute Percent Error (Interquartile Range) (%)							
	Bicyclists					Pedestrians		
	Count	Males	Females	Helmet	Sidewalk	Count	Males	Females
4-Movement Leaving Technique								
1	17 (38)	-	-	-	-	3 (26)	-	-
2	11 (33)	8 (46)	38 (61)	-	-	5 (19)	6 (18)	7 (23)
3	17 (24)	13 (35)	25 (63)	32 (49)	-	5 (19)	5 (16)	10 (21)
4	8 (5)	0 (5)	25 (54)	11 (134)	100 (87)	4 (18)	2 (20)	6 (19)
5	11 (20)	7 (34)	67 (66)	8 (111)	15 (143)	4 (6)	2 (11)	4 (3)
12-Movement Technique								
1	30 (51)	-	-	-	-	8 (9)	-	-
2	27 (57)	15 (62)	73 (105)	-	-	7 (9)	9 (7)	15 (10)
3	28 (29)	14 (12)	91 (201)	9 (26)	-	7 (23)	6 (21)	15 (22)
4	26 (24)	11(13)	100 (119)	15 (30)	18 (106)	7 (20)	8 (11)	9 (26)
5	30 (34)	14 (41)	100 (100)	14 (30)	13 (88)	8 (11)	7 (10)	13 (17)

Note: - = count not assigned.

2.5 Discussion

The intent of this study was to assess the error that might occur for different manual count techniques, and to gain a better understanding of how and why communities are conducting manual counts. The results from the error assessment and findings from the practitioner survey reveal a number of important implications.

Foremost, this study confirms the presence of error in manual bicycle and pedestrian counts. We hypothesized that the 12-movement technique would exhibit more error than the 4-movement leaving technique, but surprisingly, the differences were not statistically significant. A likely explanation for the lack of statistically significant differences between the two techniques is related to the large variability in measurement error between the counters. For example, the absolute percent errors for pedestrians ranged from 0 to 38% and 1 to 28% for the 4- and 12-way techniques, respectively. Therefore, the magnitude of measurement errors made by the counters is highly variable, irrespective of the technique used.

We had assumed the 12-movement technique was more cumbersome since it requires counters to more carefully track each traveler's movement through the intersection. However, in the debriefing survey a participant from the study noted that both techniques require the counter to wait for travelers to complete their entire movement before a tally can be made. Perhaps, error is incurred as counters wait for slow travelers to cross one crosswalk, and then another crosswalk, and then finally complete their movement all in the same time while other bicyclists and pedestrians have traveled through the intersection. For this reason, the 4-movement approaching technique might be more user-friendly and less prone to error since volunteer counters can immediately tally a traveler without needing to track any further movement. Future research should investigate the error associated with the 4-movement approaching technique.

The survey respondents stated the techniques most commonly used are the screenline and 12-movment techniques. Presumably, the advantage of the screenline technique is how user-friendly it can be for volunteers. The results of the survey suggest that communities often combine screenline counts with assignments to collect additional information. This study did not assess the error associated with screenline counts. Future research should fill this gap. The advantage of the 12-movement technique is the extra information gathered about direction of travel. Those extra data can be especially useful for validating travel demand models or making inferences about corridor flow, and origin and destination patterns.

Prior to the error assessment, we suspected error would increase when counters are assigned to collect additional information. Surprisingly, there was no apparent difference in error across the assignment levels for additional information. If a community desires to collect additional information, they might consider using the key presented in Figure 2c. Conversely, participants in our study reported that assignments to collect additional information were onerous and frustrating. For this reason, a community might want to save their volunteers the burden of collecting additional information.

Many of the questionnaire respondents stated that volunteer counters are assigned to collect information about the sex of the traveler. The error assessment appears to show greater error in determining the sex of bicyclists when compared to determining the sex of a pedestrian. This was likely because bicyclists traveled much faster and often wore additional equipment and clothing (e.g., helmet, sunglasses). Our study occurred during the spring; future research could investigate if the error in determining sex is more or less pronounced during cooler or warmer temperatures.

According to the survey, community-volunteers are often “first time counters.” This might be another source of error. Manual count programs seem to be new (most programs in the survey started within the last five years), so there is possibility that communities will begin to have repeat-volunteers. However, the survey respondents cited recruiting and retaining volunteers as the biggest problem they encounter, and most of the participants from the error assessment stated in the debriefing survey that a two-hour count was long and boring. These findings highlight another possible source of error—fatigue. Diogenes et al. (2002) observed greater error at the beginning and end of counting, which they attributed to unfamiliarity and fatigue, respectively. Many count programs involve members of bicycle and pedestrian advocacy groups, who might demonstrate greater stamina and, perhaps, exhibit less error. However, one community leader who was interviewed expressed concern that advocates might inflate count numbers in hopes of sending a stronger message to city council about walking and bicycling prevalence.

The error incurred by counters is possibly overshadowed by unreliability that is inherent to counting programs that are infrequent, cover a very short-duration, and cannot be compared with data from nearby permanent automatic counters. Most respondents reported counting for only one day and many reported counting only once per year. Counting in this manner provides merely a "snapshot" of non-motorized transportation and is highly susceptible to non-recurring variation. For example, it is possible a snapshot has occurred, unbeknownst to the counters, when a mothers' group was walking by or during a large sale at a nearby retail center. Such variation is highly unpredictable. Variation due to weather, on the other hand, can be accounted for by making comparisons with continuous data obtained from automatic counters. However, very few survey respondents reported having or using permanent automatic counters. Nor did many respondents report using the NBPD adjustment factors. It is possible communities don't need to make temporal adjustments; it largely depends on their reasons for collecting manual count data.

Some of the reasons cited by the survey respondents for conducting manual counts seem unrealistic and possibly flawed. Through an opened-ended question we asked respondents to provide their own list of reasons for conducting short-duration manual counts. Next, we provided a list of reasons and asked respondents to identify those they believed to be valid and useful. It turned out that the respondent list and our list were fairly similar. In our list we intentionally included a few reasons we believed to be flawed. For example, it is not clear how short-duration manual counts, from a few intersections, can be used for "determining the percentage of bicyclists and pedestrians that travel throughout the community", yet more than two-thirds of the survey respondents identified that as a valid and useful reason. Likewise, it is not readily apparent how a handful of snapshot counts can be used to "measure physical activity rates", yet nearly half the respondents identified that as a valid and useful reason. Another possibly debatable reason that was cited (to "make comparisons with other communities") is most likely impossible because count-totals depend on when, where, and how many locations were involved. Finally, a reason frequently provided by respondents and selected by many as "the most important reason" is to "analyze trends year-to-year". However, it is doubtful that reliable trends can be identified using short-duration sporadic counts, especially if automatic continuous counters are not nearby and available for making seasonal and temporal adjustments.

Some of the other reasons cited by the survey respondents seem more realistic. For example, the reason that was selected by the most respondents as the "most important reason" to conduct manual counts is "improving community-wide infrastructure, signage, and paint markings". Although this reason is somewhat vague, it is easy to imagine how snapshot data might be useful for comparing the magnitude of non-motorized traffic at different intersections across a community. Likewise, it is conceivable that manual counts can play an important role for "validating travel demand models." Manual counts can also serve an important starting point for further analyses to "assess safety concerns", "make adjustments to traffic signal timing", and "assess economic activity".

Error in count data has little or no bearing for some of the key reasons identified by respondents. For example, "providing an opportunity for public engagement and outreach" is an important reason to establish a manual count program, yet success can be achieved regardless of the data accuracy. Likewise, “raising awareness about bicycle and pedestrian activity” and “applying for grants” are somewhat independent of the actual data that is collected.

Nevertheless, researchers and practitioners should continue to develop strategies to reduce error, resolve concerns about reliability, and find meaningful ways to use the data. It is possible that better training programs for community volunteers could improve error rates. Error rates might be reduced by saturating an intersection with multiple counters and averaging the results; in which case, count program managers would need to balance the tradeoff of improving accuracy at the cost of reducing geographic coverage. A noteworthy characteristic of community-volunteer count programs is the unique opportunity to simultaneously collect data at locations throughout a community. Traditional traffic monitoring programs for vehicle traffic usually do not achieve synchronized counting at so many locations. Researchers should seek ways to take advantage of the wide geographic snapshot obtained during manual count events. Large communities might consider concentrating counters in specific neighborhoods to capture greater count density per mile.

Finally, researchers and practitioners should explore what steps are needed to expand count events for more consecutive days and more periodically throughout the year. Strategies should be developed to determine the optimal schedule to increase reliability. Researchers should continue to look for innovative, low cost alternatives to traditional permanent counters, such as using existing municipal cameras at intersections and other locations, taking advantage of existing inductive loops, and coordinating with vehicle traffic monitoring programs.

2.6 Conclusion

Using community-volunteers to collect manual bicycle and pedestrian counts can be a low cost strategy to collect sustainable transportation data while providing a means for public engagement. The first part of this study involved an online survey to identify how and why communities are conducting manual counts. The respondents were not a random representative sample across the country, nevertheless there was fairly broad response and the findings provide an instructive description of current practices. The respondents reported many communities are not using automatic counters, not using adjustment factors, counting infrequently, for short two-hour periods, and rarely for more than one day. It is not clear how the resulting snapshot counts can be used for some of the reasons the survey respondents claimed were motivating their count programs. For example, it is not readily apparent how such counts can be used for analyzing trends year-to-year. On the other hand, some of the stated reasons for conducting manual counts seem compatible with the activities underway. For example, a popular reason cited by survey respondents was to "improve community-wide infrastructure, signage, and paint markings" and another reason given was to "provide opportunities for public engagement and outreach." Count program managers should conduct a careful self-assessment of their capabilities, limitations, and goals to clearly define their purpose for conducting manual count programs that involve community-volunteers. In particular count program managers should decide if they are striving to have a public involvement activity that produces count data as a byproduct, or vice-versa, or a program that strives to achieve both goals equally.

The second part of this study used a controlled field experiment to assess error exhibited by a group of counters. Statistical inferences are limited to the particulars of the study, yet the essential results are reasonably transferable. The assessment found a slight difference in absolute percent error between counts collected using the 4-movement leaving technique and the more complicated 12-movement technique; however, the differences were not statistically significant. It is worth noting that the accuracy of counts were highly variable between participants, regardless of the technique used. Since the 12-movement technique provides substantially richer data without degrading accuracy, manual count program managers should consider using the 12-movement technique. This study also showed no apparent increase in error when counters were assigned to collect additional information about passing travelers. However, feedback from study participants suggests the task of collecting additional information can be burdensome. Furthermore, a survey of practitioners showed that although many count programs assign their volunteers to collect additional information, the data that are typically collected, such as sex of traveler and helmet use, do not seem to correspond with the reasons given for conducting manual counts. Consequently, manual count program managers should forego collecting additional information during scheduled count events and instead focus on taking advantage of the opportunity to simultaneously collect count data across a wide geographic area. Perhaps, additional information, such as helmet use, could be collected by staff when time permits or as needed.

Chapter 5 Estimating Bicycle Demand

3.1 Introduction

Bicycle volume data is scarce, but there are efforts underway to increase data collection. The National Bicycle and Pedestrian Documentation Project (NBPD) is a coordinated effort that aims to provide a consistent model of bike and walk data collection methods and ongoing data collection (National Bicycle and Pedestrian Documentation Project, 2014). The Project has been successful at increasing what scarce data is available. Data collection methods consist of two basic types: automatic bicycle counters and manual count programs. Automatic counters provide continuous data but these systems require up front capital outlays of $2,000 to $10,000 (Nordback et al., 2013) and thus are rare and concentrated to main routes. Manual data collection is lower cost and more flexible, and an increasingly popular method is citizen-volunteer manual count programs which typically involve providing citizen-volunteers with clipboards to simultaneously collect data at several locations. The counts are typically conducted one or two times a year during peak travel periods to create a “snapshot” of bicycle and pedestrian travel throughout the community. For example, the NBPD recommends conducting counts four times per year, once for every season, for two hours in the morning (7:00am-9:00am) and two hours in the evening (4:00pm-6:00pm).

Count programs are coordinated and conducted by entities at various levels. For example, the Washington State Department of Transportation (WSDOT) has supervised a citizen-volunteer count program since 2008. The most recent WSDOT count involved 38 jurisdictions and more than 409 observations at intersections (Cascade Bicycle Club, 2013). The California Department of Transportation recently funded an online data-clearinghouse for local communities throughout southern California to upload and share their data from their count programs (Huff, 2014).

Our research was conducted with data from Bellingham, WA, where the city administers a volunteer data collection program as part of a statewide effort coordinated by the Washington State Department of Transportation. Bellingham was selected because it has one of the most robust bicycle volume data sets in the Pacific Northwest, and because it is a relatively isolated urban area rather than a large metropolis with complex travel patterns. The City of Bellingham has participated in WSDOT's citizen-volunteer count program every year since 2006, when five locations were counted. Over time the count has expanded and by 2013, 18 locations were counted (Cascade Bicycle Club, 2013). Figure 3.1 displays the distribution of the count locations and relative bicycle volumes observed.

Data collection points are located at street intersections and intersections of streets and trails, in several cases the locations have been moved due to construction. Pedestrians and cyclists are counted as they leave the intersection. Each intersection was counted for two hours in the morning (7-9 AM) and evening (4-6 PM), typically in late September or early October. The 18 locations are spread out geographically and fairly diverse in terms of the functional classification of the intersecting streets, such as Arterial/Local or Trail/Arterial. For each collection point, bicycle data from all available years was averaged.

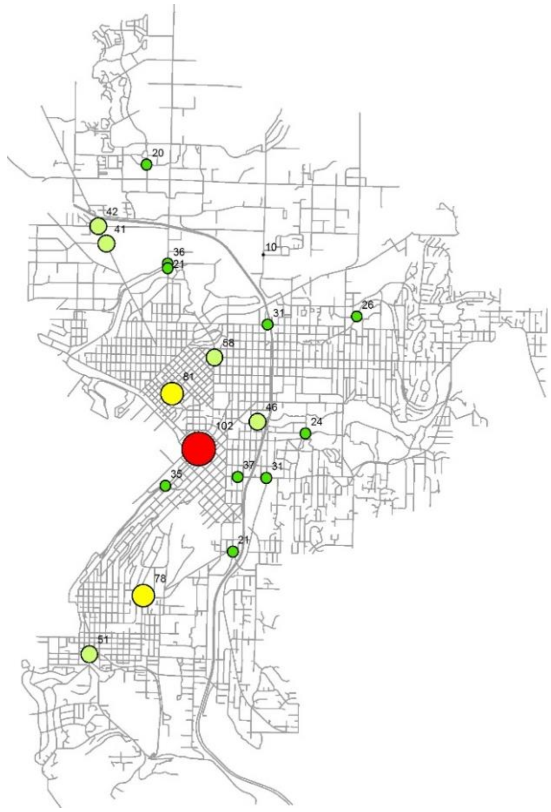


Figure 3.1 Two-hour AM bicycle count average for 2006-2013.

<u>3.2 Spatial Extrapolation of Manual Counts</u>

Traffic volume data collected at specific points can be used to estimate network-wide volumes, this is done by spatially extrapolating the data. We use a method called origin-destination (OD) centrality, which is based on a modified form of stress centrality and was successfully utilized by McDaniel et al. (2014) to estimate bicycle volumes throughout Moscow, ID (population 23,800). The method has three basic steps: 1) the transportation network is characterized in terms of preferred bicycle paths between locations, 2) land use information is used to create origins and destinations and respective multipliers that represent a magnitude of "trip potential" between origin-destination pairs, and 3) regression is used to spatially extrapolate count data from the actual count locations to all links and nodes network-wide.

The use of regression modelling and manual count data to predict bicycle volumes for unobserved locations based on characteristics is not new. Other researchers have used adjacent land use, number of vehicle lanes, vehicle speed limit, and width of the bike lane (Griswold et al., 2011). For example, Jones et al. (2010) developed a regression model ($R^2 = 0.47$) for NBPD data consisting of three explanatory variables: (1) total footage of off-street paths within 0.5 mile, (2) employment density within 0.25 mile, and (3) population density within 0.25 mile.

OD centrality has several features which make it attractive. Compared with other bicycle volume estimation methods it is relatively simple, which allows for the option of modifying the network to create alternative scenarios. It also effectively provides a means to spatially extrapolate count data throughout an entire street and trail network, which provides volume detail on all street and trail segments, as well as individual turn movements, allowing us to analyze specific types of situations.

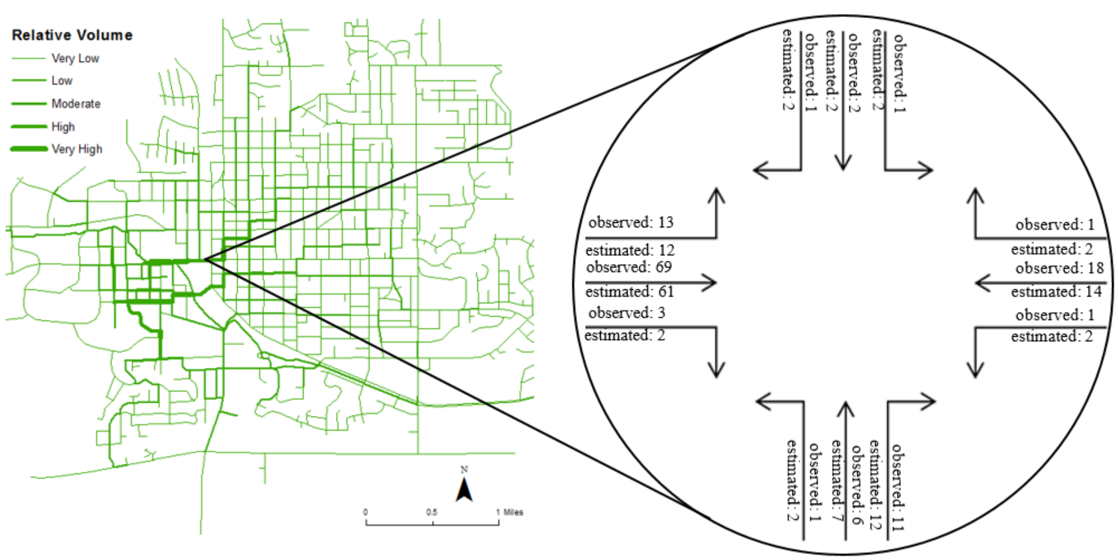


Figure 3.2 Two hour count spatially extrapolated throughout Moscow, ID.

Source: McDaniel et al. (2014)

For our case study we utilized GIS parcel, employment and school attendance data provided by the city of Bellingham. For origins, the inputs of residents per parcel were used. For destinations, employees per parcel, employment data, and parcel size were inputs. A limit of 5 miles was placed on bicycle travel. GIS files provided by the City of Bellingham were modified to create the bicycle transportation network, with GIS attributes comprising bicycle impedance inputs. 10% of the citizen count data was withheld and used for calibration. Spatial extrapolation was completed once for 2-hour AM volumes and again for 2-hour PM volumes. The output from this step was a network-wide bicycle 2-hour volume for AM, and another for PM volume.

3.3 Temporal Extrapolation of Network-Wide Two-Hour Volume

While citizen-volunteer counts are short-duration, bicycle traffic fluctuates depending on the time-of-day, day-of-week, and season. Emerging research has demonstrated the possibility of temporally extrapolating short duration count data to estimate average volumes. This research typically relies on automatic bicycle counters that can collect continuous data; figures 2.3 through 2.5 depict the findings of how bicycle travel fluctuates over time in various situations.

The goal of temporal extrapolation is to estimate Average Annual Daily Bicyclists (AADB), which represents the total annual number of bicyclists passing a point divided by 365 days. This extrapolation is conducted using adjustment factors that convert short term count data into AADB figures. Adjustment factors for short term counts are an emerging phenomenon, and there are few available to cover diverse climatic zones, facility types, and travel patterns in different cities.

For the case study in this project, off the shelf adjustment factors were not available. Adjustment factors were created for Bellingham, WA by adapting factors available from four sources and local knowledge. The first source, the NBPD, provides adjustment factors for three climatic zones called: Long Winter Short Summer, Moderate Climate, and Very Hot Summer/Mild Winter (NBPD, 2009). The second source is a report prepared by Nordback et al. (2013) for the Colorado Department of Transportation. The report provides a variety of adjustment factors for recreational trails, suburban streets, and urban streets throughout Colorado. The third source of adjustment factors is found in the US DOT's Traffic Monitoring Guide (2013) for Minneapolis, Minnesota. The fourth source of adjustment factors is the work by Miranda-Moreno et al., (2013) that included adjustment factors for five North American cities, including Vancouver, B.C. and Portland, OR, but unfortunately did not include a full year. Adjustment factors also vary by facility types (or classes), such as "path" or "pedestrian district" (National Bicycle and Pedestrian Documentation Project, 2009), or "utilitarian", "recreational", or "mixed" (Miranda-Moreno et al., 2013).

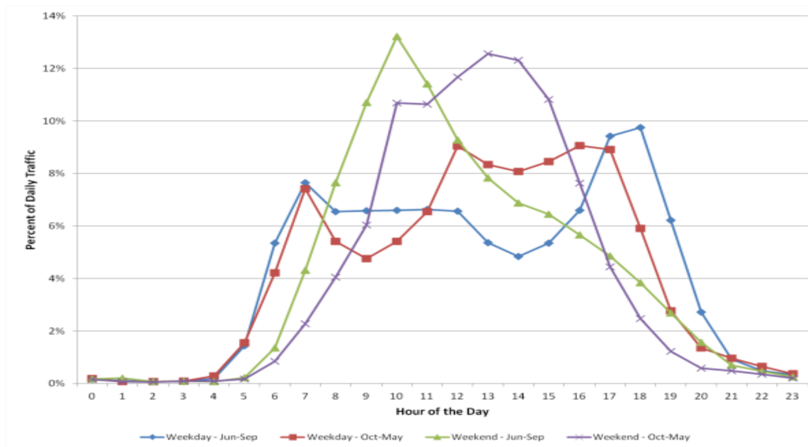


Figure 3.3 Hour of the day patterns on a shared use path in Colorado.

Source: Colorado Department of Transportation, cited in the Traffic Monitoring Guide (2013).

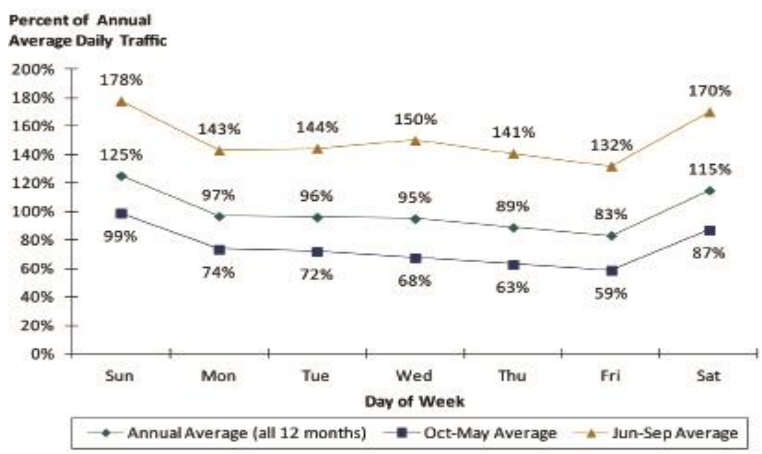


Figure 3.4 Day of the week patterns for a shared use path in Colorado.

Source: Colorado Department of Transportation, cited in the Traffic Monitoring Guide (2013).

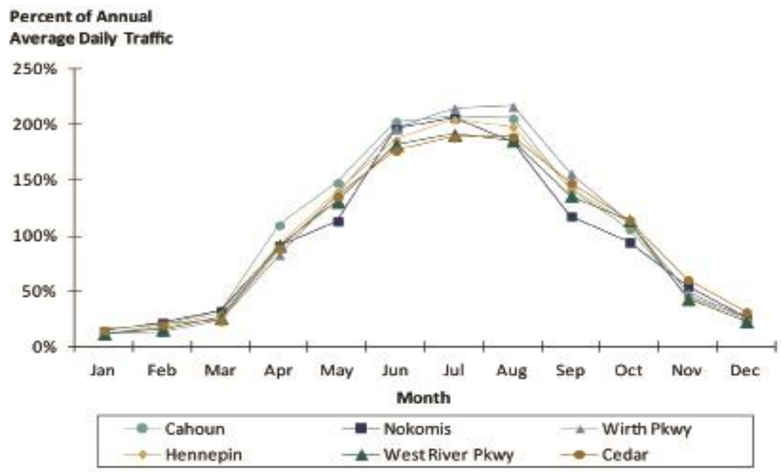


Figure 3.5 Monthly patterns for six shared use paths in Minneapolis, MN.

Source: Greg Lindsey, University of Minnesota, cited in the Traffic Monitoring Guide (2013).

The available adjustment factors outlined above were adapted to reflect Bellingham's climate characterized by a rainy fall, winter, and spring, but a dry and mild summer. Table 3.1 summarizes the sources and adaptations, and the adjustment factors used in the case study are listed in table 2.2. The factors were checked against the NBPD factors created with data from Minneapolis, Vancouver and Portland, as well as by the NBPD. For the case study only certain factors are used; the complete list of adjustment factors is shown for illustration.

Table 3.1 Summary of adaptations to adjustment factors.

Factor	Facility	Source(s)
2 Hour	Street	Adapted from average time of day patterns for bicyclists in Minneapolis MN (Traffic Monitoring Guide p. 4-22, fig. 4-12, 2013) and NBPD count adjustment factors for a pedestrian district 7-9AM and 4-6PM (NBPD 2009)
	Trail	Adapted from hour-of-day factors created for the Colorado's Cherry Creek Trail shared-use path for a weekday October-May (Traffic Monitoring Guide, p. 4-21, fig. 4-11, 2013) and NBPD count adjustment factors for a path 7-9AM and 4-6PM (NBPD 2009)
Day of Week	Street	Adapted from Nordback et al. (p. 107, fig. 56, 2013) commute patterns: low weekend, low monthly variation
	Trail	Adapted from Nordback et al. (p. 107, fig. 56, 2013) front-range non-commute patterns: high weekend, low monthly variation
Month	Street	Adapted from Cherry Creek Trail monthly factors (Traffic Monitoring Guide pp. 4-24 to 4-25, fig. 4-14 and 4-15, 2013). Adjusted similar to monthly trail factors but with increased travel in the winter months and reduced in summer months
	Trail	Adapted from Cherry Creek Trail monthly factors (Traffic Monitoring Guide pp. 4-24 to 4-25, fig. 4-14 and 4-15, 2013). Adjusted to reflect increased winter months due to less snowfall in Bellingham, slower spring increase due to rain, higher peak in late summer due to lack of intense summer heat, rapid drop in fall due to rainy season

Table 3.2 Adjustment factors for AADB used in the case study.

	2-Hour		Day of Week							Month											
Class	AM	PM	S	M	T	W	T	F	S	J	F	M	A	M	J	J	A	S	O	N	D
Street	0.1	0.2	0.1	0.2	0.2	0.2	0.2	0.1	0.1	1.4	1.4	1.3	1.1	0.9	0.8	0.7	0.7	0.8	1.0	1.5	1.5
Trail	0.1	0.2	0.2	0.1	0.1	0.1	0.1	0.1	0.2	1.7	1.7	1.5	1.3	1.0	0.8	0.6	0.6	0.7	1.1	1.7	1.7

The adjustment factors are applied to the two-hour volumes to obtain AADB as follows:

$$AADB = \left[\left(\frac{V_{AM}}{k_{AM,C}} + \frac{V_{PM}}{k_{PM,C}}\right)/2\right] * F_{day,C} * F_{month,C}$$

where

V_{AM} = AM 2-hour bicycle volume,

V_{PM} = PM 2-hour bicycle volume,

$k_{AM,C}$ = AM 2-hour adjustment factor for class C,

$k_{PM,C}$ = PM 2-hour adjustment factor for class C,

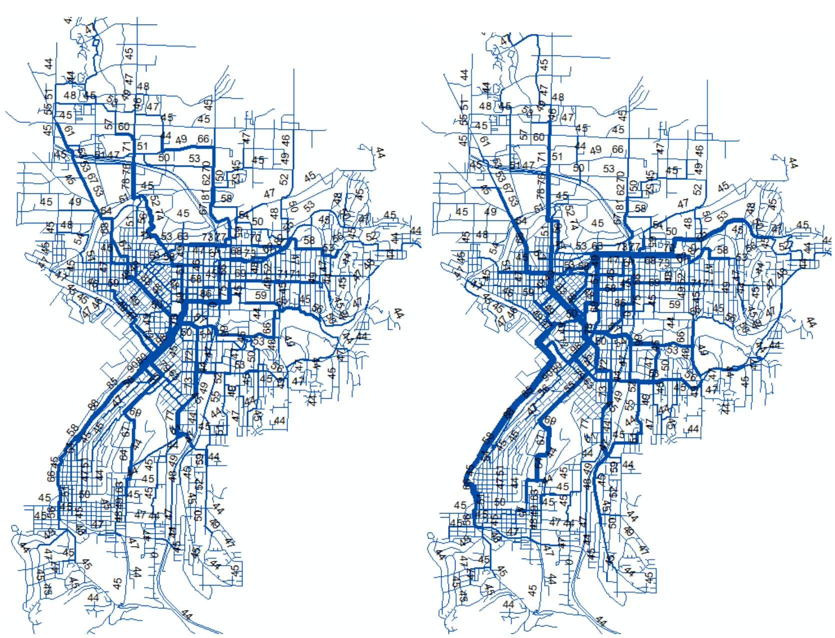
$F_{day,C}$ = day of the week adjustment factor for class C, and

$F_{month,C}$ = month adjustment factor for class C.

3.4 Results

The most important aspect of the OD centrality method is that it produces specific AADB values for every link and every turn movement throughout the network. This can be displayed in map form, as shown in figure 3.6, to depict order-of-magnitude AADB. Practitioners, citizens, and other decision-makers can use maps like these to visualize, discuss, and compare the merits of different proposals.

Origin and destination multipliers were kept fixed, so the total, study-wide bicycle volume is fixed and the change in AADB is merely a shift from travel on certain facilities to other facilities. (Although McDaniel et al. (2014) suggest OD centrality can be used to forecast future growth-scenarios by changing the residential and commercial land use input.)



Scenario 1 **Scenario 2**

Figure 3.6 Estimated AADB for Scenario 1 and Scenario 2.

As an example, Scenario 2 includes extending a shared use path from the south-west into downtown. Not surprisingly, Figure 3.6 shows a decrease in AADB along the parallel arterial because those trips would be shifted to the new shared use path. Practitioners could use maps like these to encourage community dialogue and determine where capital investment funding should be allocated to ensure adequate capacity or improve linkages.

It should be noted that the specific AADB values are rough estimates and the only way to obtain true AADB is by installing and using a continuous counter for the entire year. Consequently, AADB estimation, regardless of the method, can only be roughly validated and the most straight forward method is to simply use local expert knowledge. For the case study, and for scenario planning in general, this level of accuracy is adequate. A potential data collection scenario that could be accomplished within a limited budget would be a manual count program at many locations simultaneously, combined with automated counters that provide continuous data collection at key locations and adjustment factors specific to that locale.

3.5 Conclusion

The growing interest in bicycling is spurring additional count efforts such as volunteer programs and automatic counters. The data from these programs can be extrapolated spatially and temporally to estimate bicycle volumes network-wide. The output is data rich in that it provides volume and turn movement estimates for every link and intersection, respectively. This can provide important information for communities that are making infrastructure improvement decisions, establishing a community vision for future capital improvements, and more. The map outputs can be useful in encouraging public dialogue, and for identifying missing links. The next chapter will demonstrate how these results can be used for quantifying dangerous situation exposure in scenario analysis.

Chapter 6 Analysis of Dangerous Situations for Bicyclists

4.1 Introduction

In the United States, cyclists are 12 times more likely than automobile occupants to be killed per distance traveled (Pucher and Dijkstra, 2003). Arguably, the higher rate is inflated for various reasons, including the fact that most cycling occurs in urban areas where the possibility of accidents is intrinsically higher. Another reason is that cyclists are simply more exposed to danger and do not have the protection that comes from being inside an automobile with seatbelts and airbags (Hurst, 2010). Furthermore, the exposure for bicyclists is prolonged in terms of travel time for distance traveled. Nevertheless, even when compared with other industrialized nations, cycling in the U.S. is relatively more dangerous. Pucher and Dijkstra (2003) report that the U.S. rate of fatalities from cycling is double that of Germany and triple that of the Netherlands, both in terms of number of trips and in distance travelled. For non-fatal accidents, cyclists in the U.S. are eight times more likely to suffer an injury than German cyclists and about 30 times more likely than Dutch cyclists (Pucher and Dijkstra, 2003).

Many U.S. communities are aggressively trying to make cycling safer. They are devising ambitious bicycle improvement plans that will construct new shared-use paths and significantly increase the mileage of on-street bike lanes. One challenge, however, is to determine which projects should be prioritized. For highway and other road projects there are well-established techniques for benefit-cost analysis. Often engineers and planners can use Safety Performance Functions (SPF) to predict the expected number of automobile crashes for a given location based on the location's characteristics. Then, they can compare how different improvement scenarios might reduce accident rates using Crash Modification Factors (CMF) (Highway Safety Manual, 2010). Unfortunately, the lack of robust bicycle accident data makes it difficult to create statistically sound SPFs and CMFs for bicycle infrastructure planning and project prioritization.

Bicycle accident data is sparse for various reasons. Foremost, since few people use their bicycle for daily travel, very few accidents occur compared to automobile accidents (in absolute terms). Bicycle accidents are less likely to be reported since they often do not incur insurance claims or traffic violations. Schimek (2014) suggests that as many as 89% of bicycle accidents go unreported and he further notes that even if reported, the police documentation often lacks the details that are necessary to create SPFs and CMFs. Recently, Nordback et al. (2014) created what is apparently the first and only SPF for bicycles for a U.S. city. They did not create CMFs and their SPF is for a specific city and situation; so while their work is an important step in the right direction, the state-of-the practice still has a long way to go.

This paper presents a method for comparing the safety aspects of proposed improvement scenarios without requiring elaborate accident data or SPFs. We begin by describing two case study scenarios. The first scenario represents the current conditions for a community and the second scenario is a proposed bicycle master plan. Next, we introduce a "typology of dangerous situations for bicyclists." The typology is a first-of-its-kind attempt to enumerate for bicycle infrastructure planning what safety analysts call *situational antecedents*. The typology consists of 23 dangerous situations that we identified through a literature review and serves a valuable contribution to the practice as a framework for future research. Next, we describe how a community can define exposure indicators for these dangerous situations. For illustration, we present example indicators and demonstrate the analysis. The analysis is done using a geographic information system (GIS) tool that we programmed using open-source Python scripting. The GIS tool, instructions, and example data can be obtained from the authors. Engineers and planners can use the results from the GIS analysis to compare projects and communicate the impact of infrastructure investments to elected officials, the general public, and other stakeholders.

4.2 Case Study Scenarios

For demonstration, consider the Bicycle Master Plan (BMP) proposed for Bellingham, WA (population 80,885). The Bellingham BMP includes 184 projects that would create more than 130 miles of new bicycle facilities. The estimated cost to implement the entire plan is $20,452,398 (City of Bellingham, 2013).

Numerous scenarios with different combinations of projects could be analyzed for dangerous situation exposure. For simplicity, we analyzed two scenarios for this paper. Scenario 1 is the existing system and scenario 2 is a modified version of the BMP with new trails, bike lanes, and bicycle boulevards. Bicycle boulevards are high volume bicycle routes on roads with low vehicle speed limits and stops signs favoring the through movement for bicycles. Figure 4.1 shows the bike lanes and trails for the scenarios.

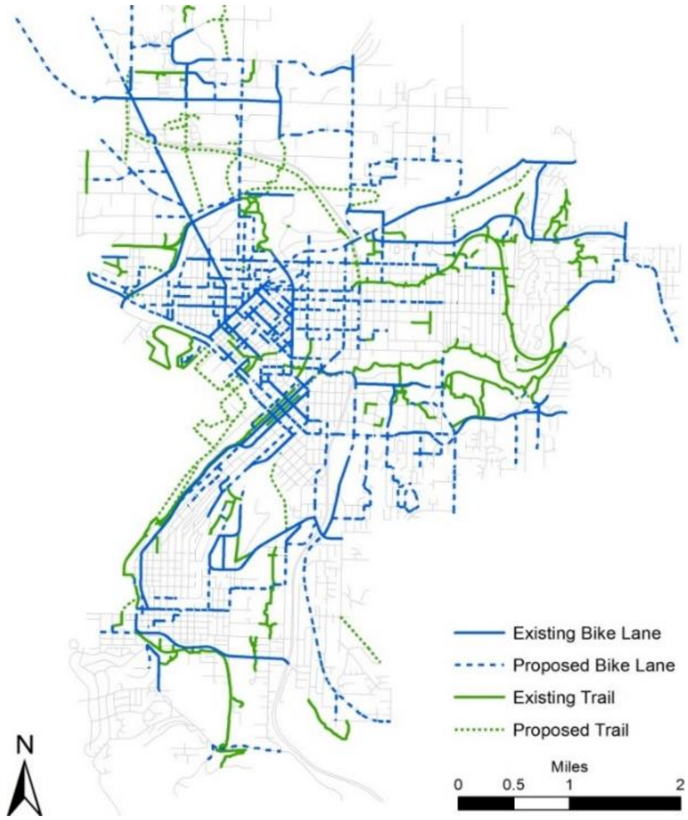
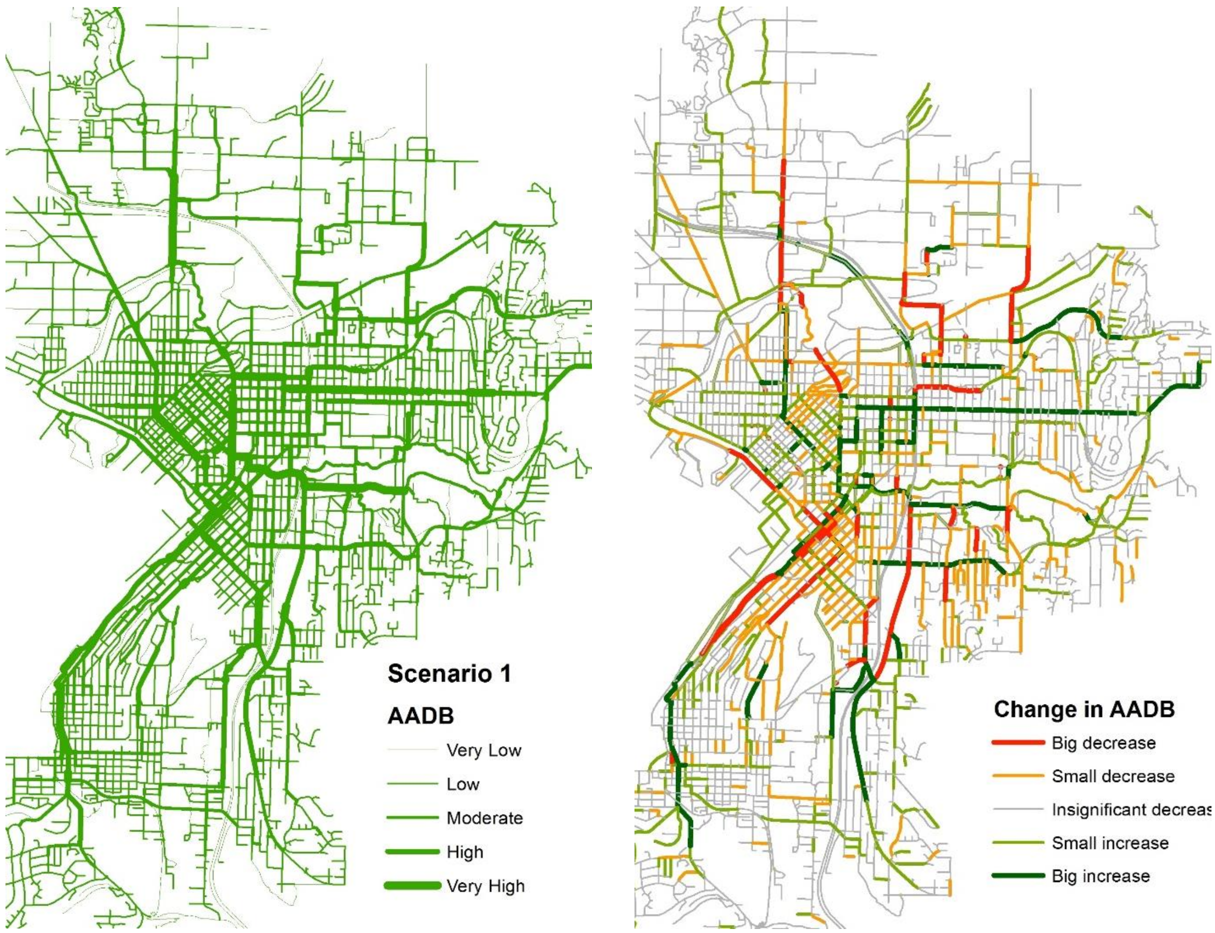


Figure 4.1 Bicycle network for existing scenario 1 and proposed scenario 2.

The analysis presented in this paper requires Annual Average Daily Traffic (AADT) and Annual Average Daily Bicycles (AADB) for every scenario. These can be obtained through any means, including travel demand models (Ortúzar, J. and L. Willumsen, 2011). Techniques for estimating AADB are not as advanced and robust as for automobiles, but in recent years significant strides have been made to improve AADB estimation (Liu. et al., 2012). For this paper, AADB was estimated using the approach developed by McDaniel et al. (McDaniel et al., 2014), but any method could be used. Figure 4.2 shows the estimated AADB for scenario 1 and the change in AADB expected under scenario 2. AADB is expected to increase for some locations and decrease for other locations. The travel demand model predicted a shift in AADB toward new bike lanes and bicycle boulevards.



(a) **(b)**

Figure 4.2 Case study (a) scenario 1 AADB (b) scenario 2 change in AADB.

<u>4.3 Dangerous Situation Typology</u>

An important contribution from this paper is the dangerous situation typology outlined in this section. We created the typology with the intent to enumerate common, distinct, and definitive situational antecedents for bicycle accidents. The list of 23 dangerous situations that we identified is based on a scan of the literature and careful deliberation. Some of the situations are dangerous for overlapping reasons and can be resolved mutually through a single mitigation strategy; for others, mitigation would require choosing between trade-offs.

All of the dangerous situations are subject to personal interpretation. Different bicyclists, with different levels of skill and experience, would certainly have their own definition of "dangerous" for each situation. For a few situations there is substantial debate and disagreement on what is safe and what is not safe. For example, some bicyclists might have enormous tolerance for vehicle traffic while others have little or no tolerance for certain situations. Nevertheless, our goal was to identify a set of situations that a community could define in a meaningful way.

Table 4.1 lists the dangerous situations we identified. Many are well-known and regularly discussed, such as the "right hook" situation or "dooring". We tried to identify names that are concise and descriptive for easy conversation. The accompanying references are intended as a general starting point for further investigation. The remainder of this section summarizes key aspects for each situation, highlighting defining thresholds where possible. The dangerous situations can be organized into five categories:

- Traffic characteristics along street segments (Dangerous Situations 1-2),
- Physical conditions along street segments (Dangerous Situations 3-10),
- Intersection and network movement (Dangerous Situations 11-16),
- Cyclist behavior (Dangerous Situations 17-21), and
- Environmental factors (Dangerous Situations 22-23).

4.3.1 Traffic Characteristics Along Street Segments

The first dangerous situation is **hazardous mixed cycling**. Mixed cycling is commonly defined as cycling without a painted or physical separation from motorized vehicles; "hazardous", on the other hand, is not as easy to define. In fact, for the past thirty years there has been ongoing, and often polarizing, debate about the hazards of mixed cycling (Mapes, 2009). Nevertheless, both sides of the debate would agree that most people have a particular, albeit elusive, tolerance for (1) vehicle volume, (2) vehicle speed, and (3) percent of heavy vehicles.

Considerable research has investigated the dangers associated with these three traffic characteristics.

Traffic volume is often considered the single most important correlate of accidents and accident severity (Allen-Munley et al., 2004). Schepers et al. (2011) designate 8,000 AADT as a cutoff threshold for what one might consider hazardous. The Dutch Design Manual for Bicycle Traffic (CROW Manual) has various thresholds, including 4,000 AADT to demarcate an undesirable level of traffic (CROW, 2007).

Vehicle speed also shows correlation with injury severity and fatalities (Klop and Khattak, 1999). Kim et al., (2007) reports that if vehicle speeds exceed 50 mph (80 km/h), the death rate for bicyclists increases sixteen-fold. They suggest mixed cycling should be avoided if the speed limit is 30 mph (50 km/h) or more.

Percent of heavy vehicles (trucks, buses, and large vans) is another defining characteristic of hazardous traffic because they have less maneuverability and more blind spots. Harkey and Stewart (Harkey and Stewart, 1997) recommend communities avoid having mixed cycling when there is a high percentage of heavy vehicles, but they do not provide a defining threshold.

Hazardous separated cycling, in the form of bike lanes or buffered bike lanes is intended to eliminate, or at least diminish, danger. Essentially separated cycling raises the thresholds that define "hazardous" for vehicle volume, vehicle speed, and percent of heavy vehicles. Reynolds et al. (2009) report for their study that bike lanes "consistently [reduced] injury rate, collision frequency or crash rates by about 50%." Nevertheless, there is some evidence that separated cycling can introduce new dangers. One concern is that the separation, especially if it is merely a painted line, gives a false sense of protection to both driver and cyclist, thus discouraging attentive travel for both parties (Van Houten and Seiderman, 2005). Parkin and

Meyers (2010) showed that drivers give less lateral space to cyclists in a bike lane compared to mixed cycling. The false sense of security is particularly problematic at intersections (Lusk et al., 2011).

Table 4.1 Dangerous situations for bicyclists.

Dangerous Situation		Description	References
1	Hazardous mixed cycling	Cycling in vehicle travel lane with high vehicle volume, speed, and/or percent heavy vehicle.	Schepers et al., 2011; CROW 2007; Kim et al., 2007; Allen-Munley et al., 2004; Klop and Khattak, 1999.
2	Hazardous separated cycling	Cycling in bike lane or cycle track with high vehicle volume, speed, and/or percent heavy vehicle.	Parkin and Meyers, 2010; Reynolds et al., 2009; Lusk et al., 2011; Van Houten and Seiderman, 2005.
3	Cramped space	Narrow roads without a bike lane or shoulder.	Vandenbulcke 2011; Harkey and Stewart, 1997.
4	Excessive space	Wide roads without a bike lane.	Allen-Munley et al., 2004; Hunter et al., 1999.
5	Dooring	Street segments with on-street parking and high parking turnover.	Tilahun et al., 2007.
6	Driveways	Street segments with frequent or unexpected access points.	Räsänen and Summala, 1998.
7	Railroad tracks	Crossing or riding alongside railroad tracks.	Teschke et al. 2012.
8	Poor pavement	Pot holes and abrupt uneven surfaces.	-
9	Winding road	Frequent and/or sudden sharp curves.	Kim et al., 2007.
10	Steep hills	Hilly terrain and/or steep grades.	Teschke et al., 2012, Klop and Khattak, 1999.
11	Hazardous crossing	Crossing a road with high vehicle volume, speed, and/or percent heavy vehicle.	CROW, 2007; Summala et al., 1996.
12	Oncoming cross	Oncoming left-turning vehicles cut off through movement bicyclists.	Shepers et al., 2014; Summala et al., 1996.
13	Right hook	Right-turning vehicle conflicts with through movement cyclist.	Furth et al., 2014; Schimek, 2014; Weigand, 2008; McCarthy and Gilbert, 1996.
14	Left sneak	For left turn, sneaking across travel lanes, waiting for a gap in oncoming traffic, and sneaking in front of oncoming.	Hunter et al., 1999; Wachtel and Lewiston, 1994.
15	Complicated intersection	Navigating for example, five point intersections or roundabouts.	Daniels et al., 2009.
16	Bikeway gap	Discontinuity in bicycle network.	Mekuria et al., 2012; Krizek and Roland, 2005.
17	Wrong-way riding	Cycling the wrong-way.	Hunter et al., 1999; Summala et al., 1996.
18	Sidewalk riding	Cycling on sidewalks.	Schimek, 2014.
19	Safety in numbers	Low cyclist volume.	Nordback et al., 2014, Jacobsen, 2003.
20	Crowded path	High volume shared use paths.	Teschke et al., 2012; CROW 2007.
21	Reckless riding	Riding behavior that is unsafe.	Minikel 2012; Kim et al., 2007.
22	Bad weather	Inclement weather that decreases visibility and/or cyclist control.	Kim et al., 2007.
23	Darkness	Insufficient lighting.	Schimek, 2014; Reynolds et al., 2009.

4.3.2 Physical Conditions Along Street Segments

Cramped space is dangerous when motorists are inclined to overtake cyclists. Vandenbulcke et al., (2013) suggest cramped space is the primary factor for increased accidents on bridges. Harkey and Stewart (1997) found drivers typically exhibit 6 ft (1.83m) of separation when overtaking bicyclists and recommend providing cyclists at least 4 ft (1.2m. On the other hand, **excessive space** can be unsafe also. Allen-Munley et al. (2004) found injuries were more severe on wide streets, perhaps because excess space encourages higher vehicle speeds and can give cyclists a false sense of security. Hunter et al. (1999) suggest “Lanes wider than 4.6 m [15 ft] sometimes result in the [dangerous situation] of two motor vehicles side by side”.

Dooring is the term cyclists use to describe being hit by a car door when riding next to on-street parking. Tilahun et al. (2007) found that cyclists are willing to add more than 9 minutes to a 20 minute commute to avoid on-street parking. Cycling next to parking can also be dangerous because of decreased sight distances. **Driveways** are dangerous because entering or exiting motorists often do not expect to see cyclists, or vice versa (Räsänen and Summala, 1998). The danger is exacerbated when combined with wrong-way riding and sidewalk riding. Allen-Munley et al. (2004) found that 20% of collisions in their study occurred at driveways.

Crossing railroad tracks can result in tire slippage or cause riders to swerve into traffic in an attempt to cross at safe angles **(**Teschke et al., 2012). **Poor pavement quality** is often included in bicycle suitability calculations; however we were not able to find any studies relating pavement quality to accidents. **Winding roads** can reduce sight distance and maneuverability. Kim et al. (2007) found accidents involving curving roads tend to be more severe. **Steep hills** can affect braking efficiency, sight distance, and speed differential between vehicles and cyclists.

Various studies have shown increased accident severity for grades greater than 4% *(*Allen-Munley et al., 2004).

4.3.3 Intersection and Network Movement

Intersections are the greatest point of vehicle-bicycle interaction (Wang and Nihan, 2004). Wachtel and Lewiston (1994) reported that 74% of collisions, in their study, occurred at intersections. The study by Schimek (2014) found that 81% of collisions occurred at intersections.

Hazardous crossings can be mentally and physically demanding (Summala et al., 1996). Once again, “hazardous” is subjective and can be defined based on vehicle volume, vehicle speed, and percent heavy vehicle. The CROW manual (2007) defines the following thresholds: up to 800 passenger car equivalent per hour (pcu/h), crossability is considered “reasonable without a central traffic island”; from 800 to 1,600 pcu/h, crossability is “reasonable provided crossings can be made in two stages” with a central refuge island; from 1,600 to 2,000 pcu/h, crossability is “moderate to poor”; and above 2,000 pcu/h, crossability is “poor”. The CROW manual (2007) also recommends cross street speed limits should not exceed 20 mph (30 km/h) for uncontrolled intersections with a high volume of bicyclists.

The **oncoming cross** is when oncoming vehicles turning left cut in front of bicyclists. In a study in Finland, it was found that 12% of vehicle-bicycle collisions involved an oncoming cross (Räsänen and Summala, 1998). In the U.S., Schimek (2014) found that 9% of vehicle-bicycle collisions involved a oncoming cross, the highest of any specific intersection movement. One remedy offered by Shepers et al. (2011) is to provide a left-turn lane in the oncoming direction, even if not warranted for vehicle needs, to give oncoming drivers “extra time to slow down and notice cyclists”.

The infamous **right hook** is the situation where a right turning vehicle collides with a straight moving bicyclist. In a London study, one-third of fatal collisions occurred because of this situation (vehicles in Great Britain drive on the left side of the road, so it is a “left hook”) (McCarthy and Gilbert, 1996). In an US study, 6% of bicycle collisions were right hook collisions (Schimek, 2014). There are several strategies to reduce right hook collisions, including advanced stop lines and special traffic signals (Weigard, 2008). For pedestrians, Furth et al. (2014) note that some states have thresholds (e.g. 250 right-turning veh/hr in Massachusetts) that trigger recommendations for special infrastructure, but no state has similar thresholds for bicyclists. In the Netherlands special accommodations are provided for bicyclists when there are more than 150 right turning vehicles per hour (Furth et al., 2014).

The **left sneak** is dangerous for three reasons: first, there is the danger of sneaking across travel lanes in order to make a left turn; second, there is the risk of being rear-ended while waiting to turn left; and third, is the danger from sneaking through a gap in oncoming traffic to complete the turn. Hunter et al. (1999) found that 6% of bicycle collisions in Austin, TX involved cyclists attempting to turn left. At some locations, cyclists crossed multiple lanes at once. Their study summarized observed left turns as follows: 44% “did so like a motor vehicle with proper lane destination positioning,” 8% “with improper lane destination positioning”, 17% “made pedestrian style left turns, where the cyclist would ride all the way to the intersection and then use the crosswalk to get across the street like a pedestrian” (Hunter et al., 1999).

Complicated intersections, which include 5-leg intersections, one-way/two-way intersections, and roundabouts, can be dangerous to navigate (Daniels et al., 2009). **Bikeway gaps** often occur at intersections where bike lanes terminate and do not continue on the other side. Although little research has investigated safety issues associated with gaps, Krizek et al.

(2005) and Mekuria et al. (2012) suggest discontinuities introduce high levels of stress and discomfort.

4.3.4 Cyclist Behavior

Wrong-way riding is dangerous because it increases the possibility of head-on collisions, which are usually more severe. It also increases the possibility of broadside collisions because motorists entering or exiting a street often do not look in the direction of wrong-way riders (Summala et al., 1996). Note that wrong-way riding is not the same as “contraflow” or “facing traffic cycling”, where cycling against vehicle traffic is permitted and intentional. Contraflow riding on low volume, low speed roads is relatively safe due to increased opportunity for cyclists and motorists to see each other and react (Kim et al., 2007). In the Netherlands, contraflow bike lanes are frequently installed to preempt wrong-way riding (CROW, 2007). Alta Planning + Design (2004) found that “sharrows” can reduce wrong-way riding by as much as 80%. Hunter et al. (1999) found that wrong-way riding was more prevalent on streets without a bike lane, suggesting bike lanes encourage proper riding. **Sidewalk riding** increases risk of collision with motorists at crosswalks and driveways, as well as with pedestrians. Schimek (2014) found 77% of injured sidewalk riders were going the “wrong way”.

Safety in numbers is the term for the empirically observed phenomena that accident risk decreases as the number of cyclists increases (Elvik et al., 2009). Jacobsen (2003) found the non-linear relationship can be fit with a power exponent of 0.4, which translates to about a one-third increase in injuries for a doubling of cyclists. Nordback et al. (2014) found that signalized intersections with fewer than 200 AADB have substantially higher collisions per cyclist, while intersections with greater than 600 AADB have very low risk. On the other hand, **crowded shared-use paths** are dangerous due to potential conflict with other bicyclists or pedestrians.

Teschke et al. (2012) suggest shared-use paths are more dangerous than bike-only paths and cycle tracks. The CROW manual (2007) recommends separating bicyclists with paint markings if the pedestrian flow rate is greater than 100 pedestrians per hour and separating with a physical barrier if the pedestrian flow rate is greater than 200 pedestrians per hour.

Reckless riding includes alcohol consumption, aggressive behavior, distracted cycling, and other unsafe practices. Minikel (2012) reports for a Berkley dataset, police found 41% cyclist at fault, 54% motorist at fault, and the remaining 5% no-fault. Bicyclist intoxication increases severity of head injuries (Kim et al., 2007).

4.3.5 Environmental Factors

Kim et al. (2007) found that **bad weather**, i.e. “rain, snow, fog, etc.,” increase the probability of fatality by 128% and suggest reduced visibility and traction are to blame. They also found **darkness** increases the probability of fatality by 110%. Schimek (2014) estimates that crashes in low-light conditions account for 20% of injuries and 47% of fatalities in the U.S. Reynolds et al. (2009) report that street lighting on rural roads reduced the rate of cyclists' injuries by half.

<u>4.4 Analysis and Results</u>

4.4.1 Defining and Assessing Dangerous Situation Thresholds

The first step for a community is to define thresholds for the dangerous situations they would like to analyze. For example, the threshold for “hazardous mixed cycling” might be: a street without a bike lane, with AADT > 3,000 AADT, speed limit > 30 mph, and percent heavy vehicle > 5%. A community should define their own community-specific thresholds for their needs, circumstances, and desires, taking into consideration the proportion of older adults, children, and novice riders in their community. A small rural community might define “hazardous traffic” considerably different than a large urban community with a long tradition of bicycling. Furthermore, a community might decide to have more than one threshold for a particular dangerous situation. Geller (2007) suggested there are four types of bicyclists: (1) “Strong and Fearless”, (2) “Enthused and Confident”, (3) “Interested but Concerned”, (4) “No Way No How.” He further postulates the majority (60%) of bicyclists are Interested but Concerned. A community might want to define their thresholds for this type of bicyclist. The thresholds should be defined based on public input, local experience, and findings from research.

The next step for the community is to use our GIS tool to asses the dangerous situation thresholds for various bicycle infrastructure improvement scenarios. The tool requires two input files. The first is a textfile with threshold values. The tool is fully modifiable, so a community could add and remove thresholds or change how thresholds are interpreted by the tool. The second input is a GIS street network with AADT, AADB and any other attributes that the user may need for their thresholds. AADT and AADB can be obtained through any travel demand forecasting model or some other process. The tool calculates an *exposure indicator* for every dangerous situation by checking threshold conditions for every street segment and intersection. The calculation takes 3 seconds on a standard laptop computer for the case study network, which has 4,379 street segments and 3,271 intersections.

For the case study, we defined thresholds for 12 dangerous situations. Some of the dangerous situations not analyzed would most likely require special GIS data. For example, analysis of “sidewalk riding” would most likely involve GIS data about the presence of sidewalks. The case study used commonly available street attributes, including lane width, speed limit, presence of bike lane, percent heavy vehicle, slope, access point frequency, and on-street parking turnover rate (The last two were derived from functional classification and adjacent land use. Slope was calculated using a digital elevation map.). It might not be possible to analyze some dangerous situations using GIS, such as “bad weather” and “reckless riding”; however, there is no obvious reason why these situations would need to be analyzed when comparing infrastructure scenarios, since they are primarily mitigated through education and enforcement programs.

4.4.2 Dangerous Situation Exposure

The results from the case study are shown Table 4.2 and Table 4.3. Table 4.2 shows exposure along street segments in terms of bicycle miles traveled (BMT). The GIS tool calculates BMT by multiplying street segment length times AADB. Column 2 shows the thresholds used for the analysis. The resulting indicator values (columns 3 and 4) represent dangerous situation exposure and provide a means for comparing scenarios (columns 5 and 6). The indicator values are not a prediction of accidents and should not be construed to represent actual conflict numbers. For example, “hazardous mixed cycling” is calculated by summing BMT across every street segment that meets the threshold conditions, but since some cycling would occur when traffic volumes are not hazardous the value cannot be expected to represent the true BMT of hazardous mixed cycling. Furthermore, the exposure indicators rely heavily on the output from travel demand models which only produce order-of-magnitude estimates.

Table 4.2 Dangerous situation exposure along street segments.

Dangerous Situation	Conditions and Thresholds	Scenario 1: Existing Conditions (BMT)	Scenario 2: w/Proposed Improvements (BMT)	Change (BMT)	Percent Change (%)
Hazardous mixed cycling	mixed cycling, >3,000 AADT, >30 mph, >5% heavy vehicle	11,437	5,138	-6,299	-55
Hazardous separated cycling	separated cycling, >8,000 AADT, >50 mph, >10% heavy vehicle	4,860	9,953	+5,093	+105
Cramped space	mixed cycling, lane and shoulder width < 12 ft, >1,000 AADT, >20 mph,	1,349	1,059	-290	-21
Excessive space	mixed cycling, lane and shoulder width > 15 ft	8,684	3,232	-5,452	-63
Dooring	on street parking, turnover > 4 per hour	13,545	13,186	-359	-3
Driveways	access points > 30 per mile	16,592	17,324	+732	+4
Steep hills	grade > 4%	9,680	9,832	+152	+2
Safety in numbers	< 200 AADB	40,503	41,003	+500	+1
Wrong-way riding	wrong-way riding occurrence	252	236	-16	-6

Table 4.2 shows a 55% decrease to “mixed cycling in hazardous traffic”. However, this is accompanied by a much larger (105%) increase in “separated cycling in hazardous traffic”. This peculiar result is because some roadways simply changed categories with the provision of new bike lanes. In other words, in scenario 1 the BMT was counted as mixed cycling, while in scenario 2 the BMT was counted as separated cycling, and despite higher thresholds, the BMT was still deemed hazardous. Furthermore, the travel demand model may have shifted some cyclists to roadways with new bike lanes despite the hazardous conditions. Nevertheless, overall cycling in hazardous traffic (the sum of both situations) decreased by 1,206 BMT, which can be

attributed to bike lanes alleviating danger and new bike boulevards and trails shifting cyclists away from hazardous conditions.

Cycling in “cramped space” decreased by 21%. This is because new bicycle facilities attracted cyclists away from narrow streets. Cycling in “excessive space” showed an even greater decrease (63%). Some of the decrease can also be attributed to route shifts incurred by new bike lanes and trails. However, much of the decrease is simply from providing bike lanes since “excessive space” is only tallied for mixed cycling according to the case study thresholds. A community could choose to include additional thresholds for when a bike lane is present.

“Dooring” showed a slight 3% decrease because new bicycle provisions elsewhere shifted cyclists away from bustling on-street parking. On the other hand, BMT for “driveways” increased by 4%. This is because scenario 2 provides numerous new bicycle boulevards that travel through residential neighborhoods, where access points are more frequent. BMT for “steep hills” increased by 2%, presumably because new bicycle facilities would persuade cyclists to ride up and down hills. Scenarios 2 concentrates cyclists along certain corridors. However, as a consequence, some streets would exhibit less cyclists, explaining the unfavorable 1% change for “safety in numbers”. “Wrong-way riding” decreased by 6% because, according to the travel demand model, new bicycle accommodations would discourage shortcuts.

Table 4.3 summarizes exposure at intersections in terms of AADB. There is an increase for “hazardous crossings” of 5%. This is probably because cyclists were shifted to arterials with new bike lanes. The other situations showed modest decreases. An analyst could explore how the results might change under various threshold values. Likewise, it is possible to change how the indicators are calculated. For example, in the case study “hazardous crossings” was tallied if any of the three thresholds were exceeded, but an analyst might choose to only tally if all three of the thresholds are exceeded. Through multiple executions of the tool, the analyst might discover new insights about the peculiars of certain locations in their community. Furthermore, they can make maps for specific dangerous situations like the one shown in Figure 4.3. The map shows locations where there is a high number of bicyclists going straight and vehicles turning right, i.e. the right hook dangerous situation. These locations can be visited and inspected in the field to gain a better understanding of the actual dangers.

Table 4.3 Dangerous situation exposure at intersections.

Dangerous Situation	Conditions and Thresholds	Scenario 1: Existing Conditions (AADB)	Scenario 2: w/Proposed Improvements (AADB)	Change (AADB)	Percent Change (%)
Hazardous crossing	bicyclist traveling straight, cross street: > 8,000 AADT, > 50 mph, > 10% heavy vehicle	31,595	33,297	+1,702	+5
Oncoming cross	bicyclist traveling straight, oncoming left-turning AADT > 2,000	45,577	42,516	-3,061	-7
Right hook	bicyclist traveling straight, right turning vehicles > 2,000 AADT	51,603	47,737	-3,866	-7
Left sneak	bicyclist turning left, adjacent vehicles > 8,000 AADT oncoming vehicles > 8,000 AADT	9,015	8,798	-217	-2

Figure 4.3 Map highlighting dangerous situation locations for the right hook.

4.5 Conclusion

This paper introduced a novel method for engineers and planners to evaluate projects for sketch-level scenario planning in terms of bicyclist exposure to dangerous situations. The results can be presented through tables, figures, and maps to elected officials, the general public, and other stakeholders at town hall meetings or other forums for deliberative decision-making. The method and tools are intuitive and easy to use.

The dangerous situation typology provides a framework for future research. Perhaps SPFs could be developed based on exposure indicators by investigating correlation with bicycle accident data. Research should seek to develop thresholds for dangerous situations not included in this study. For example, one possible approach to analyze “sidewalk riding” would be to assume that streets passing through areas of vibrant commercial activity, such as a lively Main Street or “Urban Village”, would have a percentage of sidewalk riding if there is not a bike lane. “Winding roads” might be analyzed through some sophisticated spatial GIS procedure. “Crowded path” could be analyzed with GIS data about painted center lines and pedestrian volumes. Finally, the tool could be enhanced through an optimization procedure that could identify high priority projects by minimizing dangerous situation exposure.

Chapter 8 Passive Travel Data Collection

5.1 Introduction

Mobile devices such as cell phones, PDAs, and hands free audio devices are becoming pervasive in society as most individuals carry at least one of these with them everywhere they go. Many companies and researchers are recognizing the vast potential of this device network. For example, some companies are using the GPS information from smart phones to learn and report real time traffic conditions. Other researchers are using the GPS data from volunteers to replace travel diaries.

One area that has not received as much attention is using the communicative power of Bluetooth devices. Most mobile devices have the ability to communicate with other devices via Bluetooth. In addition they are constantly scanning and noting all other Bluetooth devices within communication range (anywhere from 10 to 100 meters depending on the device). There are applications that allow devices to keep a record of all intercepted Bluetooth signals, via the intercepted devices unique 6-hex digit MAC address, as well as the time and location that the device was detected. If many phones across a city were collecting this data and uploading it to a central server there would be a wealth of information available for researchers and planners. The goal of this paper is to present a proof of concept for this data collection method.

One possible application of this is learning what parts of the city have the most travelers. While there are many ways to get this information for car traffic, this Bluetooth network would also pick up pedestrians and bicyclists. Another potential would be to use the unique MAC address of intercepted devices to track travelers across space and time as they travel around the city. This could provide information on route selection, origin/destination demand, and travel speeds (which can be used to infer travel mode).

In addition to providing a wealth of information, this data collection system would be very cheap and easy to distribute. Rather than hiring individuals to go out into the city and count travelers or conduct surveys, this method crowd sources the data collection. Citizens throughout a city simply need to download an app on their smart phone, and then go about their lives as normal. The device would be scanning, recording, and uploading with no input from the owner.

5.2 Background

5.2.1 Passive GPS Travel Diaries

Travel diaries have been used for decades by researchers to learn where, how, when, and why individuals are traveling. A travel diary allows for detailed reports on trips. However they are completely reliant on the effort and thoroughness of the volunteer. A common problem is individuals forgetting to record short trips, especially those part of a trip chain, or for volunteers to put off filling out the journal for a few days, by which time they have forgotten the details of the trip.

Because travel diaries are reliant on volunteers to provide all the information, they are considered active feedback. Individuals carrying a Global Positioning System (GPS) enabled device that records their location, is a passive form of data collection that is much easier on the participant. All they have to do is remember to carry the GPS device with them. With this passive data collection there are no recall issues, as all the data is stored electronically. Also it takes much less effort from the participant, so researchers can conduct studies for longer time-periods without having to worry about participant fatigue. For this reason GPS data collection has become quite popular in both movement science and travel engineering research.

One drawback to GPS data collection is the data processing requirements. Thousands of data points per participant are recorded each day. Before this data can be used the trips have to be identified, a challenge into which much research has been done. Once the trips are identified researchers often want to know the mode and purpose of the trip. This study tests predicting trip mode and purpose with multinomial logit models based on various parameters gathered from the GPS data, city road/path and bus networks, and city land use data.

In 2014 Broach et al. performed a similar study that used a multinomial logit model to predict trip mode. However one difficulty they had was in gathering the ground truth data form individuals. Individuals weren't asked information about their trips until the end of the study, by which time many had difficulty recalling their travel mode. This study aims to pick up where Broach et al. left off. Individuals will be asked for information about their trips within one to three days to try to avoid recall problems. Also this study will test trip purpose prediction in addition to mode prediction.

5.2.2 Moving Bluetooth Detection

The most common application of Bluetooth detection in traffic research is to gather travel time data along major corridors. Static Bluetooth sensors are placed at either end of a corridor, and record the MAC address of all cars with Bluetooth devices that drive by. If the same MAC address is then seen down road by the second sensor the travel time can be calculated. This use of static Bluetooth sensors has been researched and utilized extensively. Where much less research has been done, is the use of moving Bluetooth sensors.

Wang at the University of Washington performed a small study where four individuals walked around campus for an hour with Bluetooth sensors. This gave a small scale look into what this type of data collection could provide. The results were promising. The busy parts of campus were clearly identified, as well as detecting a spike in traffic during the class change. In addition many devices were detected more than once allowing for travel information to be obtained.

Versichele et al used a moving Bluetooth sensor to predict crowd sizes at a bicycling race in Belgium. This paper used the Class of Device (COD) code, one of the pieces of information that sensors learn about detected devices similar to the MAC address. The COD tells the type of device that was detected. In this study, only phone detections were used as all others were assumed to be either redundant or static devices such as desktops.

5.3 Method

5.3.1 Field Experiment Overview

For this experiment, 53 volunteers were given apps for android devices (either installed on their own phone or installed on a phone provided them for the experiment). This app records the phone's location through use of Global Positioning Satellites (GPS) as well as any Bluetooth signals it detects. See Figure 5.1. This data is recorded once a second for three seconds every ten seconds. All this data is stored on the android phone until it connects to a WiFi signal when it uploads the data to a central server. The participants were instructed to simply carry these devices with them for 10-12 days going about their life as normal.

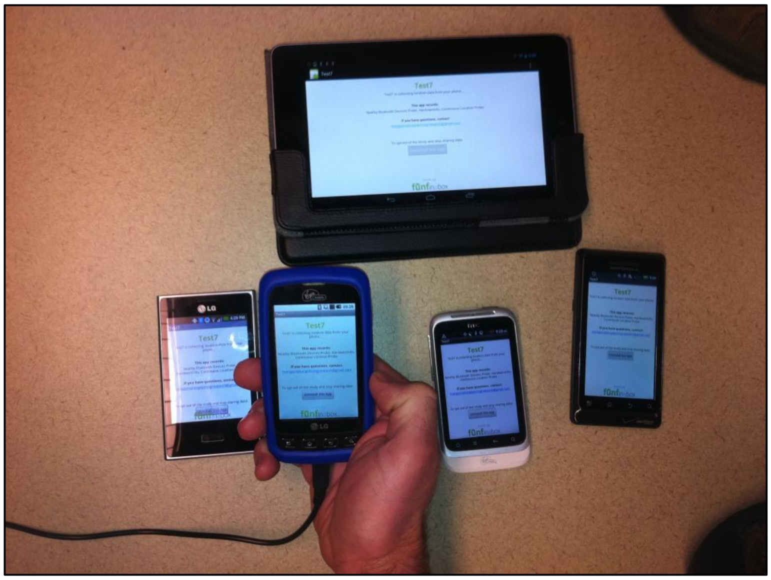


Figure 5.1 Android devices and Funf travel app.

This data collection took place in and around Moscow, Idaho. Moscow is a small town with a population of 24,500 and contains the University of Idaho. Eight miles to the West is Pullman, Washington which is slightly larger and contains the University of Washington. A large portion of our volunteers were students of the University of Idaho (36 of the 42 from whom we received data).

An important aspect of the analysis was the development of an algorithm to process a stream of GPS points into individual trips or trip segments. The trip isolation algorithm was based on the density of point within a specified dwell radius defined by distance and time. Figure 5.2 shows a schematic of the process.

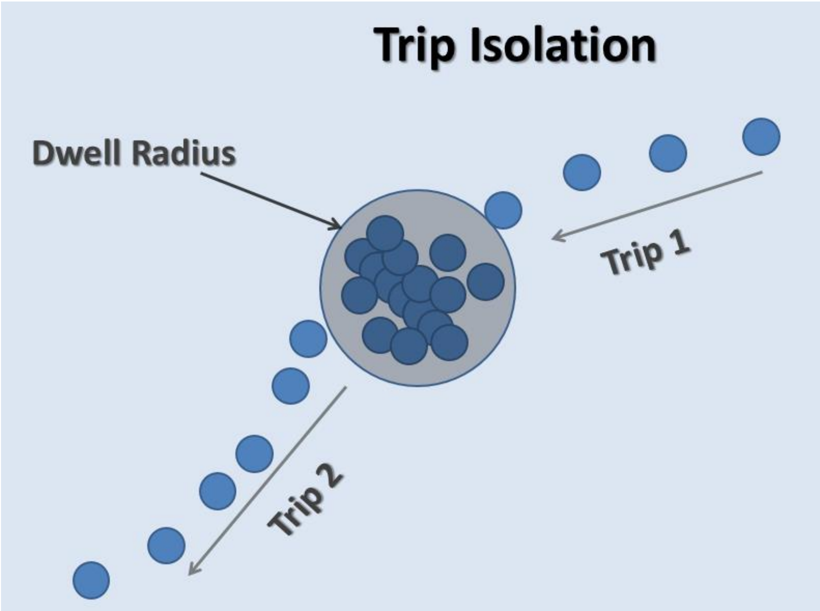


Figure 5.2 Trip isolation to distinguish trip segments.

5.3.2 GPS Data Collection

Every seven seconds this app recorded the phones GPS location once a second for three seconds. The app then uploaded this data to servers whenever it was connected to WiFi. This GPS data was then analyzed and trips were identified. At the end of each day trips were uploaded to a website. Each volunteer was then asked to log in to the website and answer some questions about the trips they made that day. Figure 5.3 shows a screen shot of this website. Volunteers were asked to confirm that they did make the shown trip, if there were any inaccuracies with the displayed trip, what their mode was, and where they were coming from and traveling to. This was repeated every day for 10 days. We received usable data and confirmation from 37 of the volunteers giving a total of 294 trips; Table 5.1 shows a summary of the data collected.

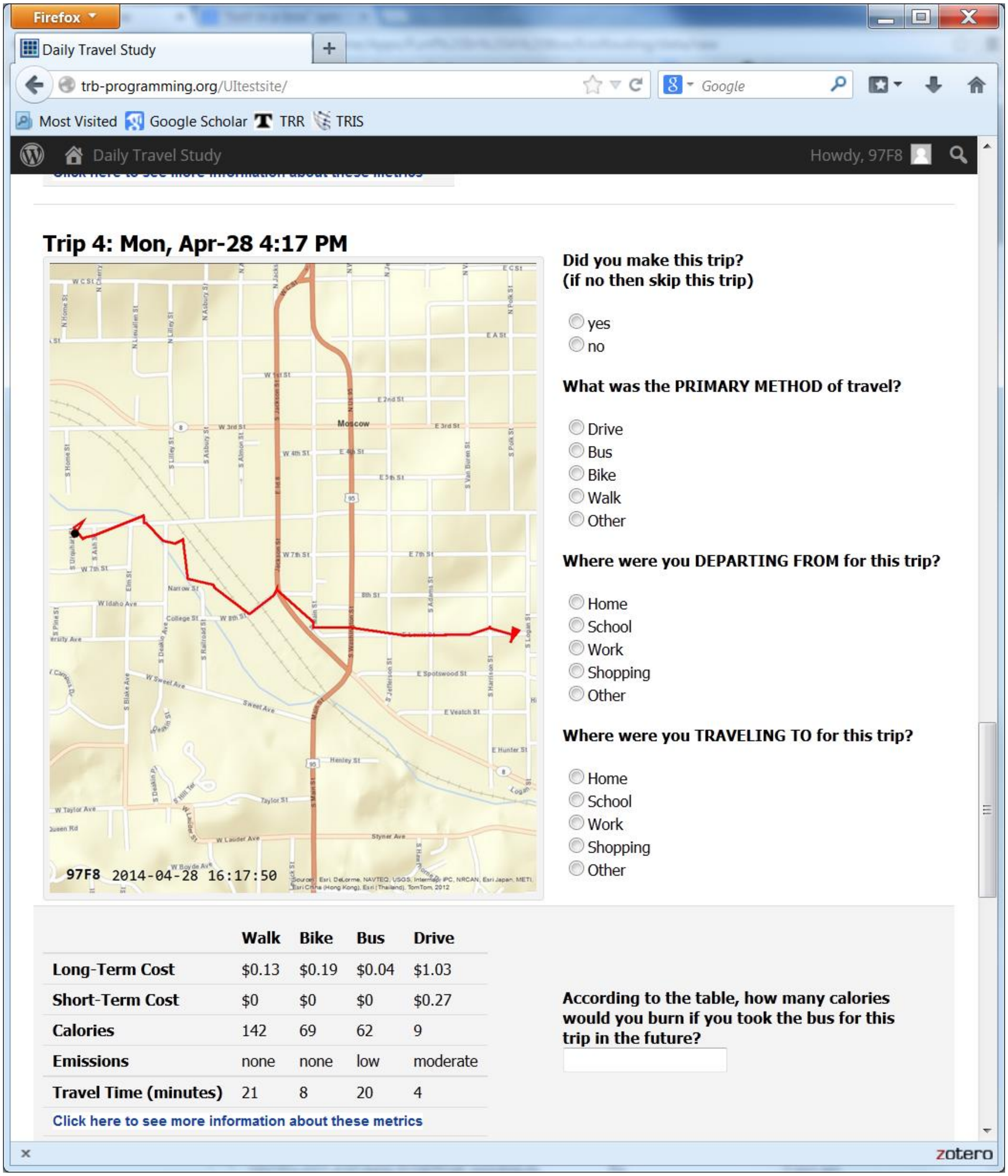


Figure 5.3 Screenshot of data collection website.

Table 5.1 Summary of collected data.

Mode	Number of Trips	Destination Purpose	Number of Trips
Bike	35	Home	91
Bus	11	Other	71
Drive	118	School	82
Other	14	Shopping	13
Walk	116	Work	37
All modes	294	All Purposes	294

5.3.3 Bluetooth Detection

After collecting data for twelve days over three thousand devices were intercepted by the phones of the volunteers, recording over three hundred thousand data points. Next the reported COD codes of the intercepted devices were decoded to report the class of device. A large number of the intercepted devices were static devices like desktop computers. Static devices are unhelpful as they are not being carried by an individual, and therefore cannot be used to track and individuals movement. After keeping only the mobile devices, this still leaves over fifteen hundred unique devices and over thirty-six thousand data points. Table 5.2 below summarizes the collected data.

Table 5.2 Data summarized by device class.

Device Class	Unique Devices	Data Points	Mobile Device
Computer	**1,278**	**226,281**	
Laptop	902	164,522	
Desktop workstation	317	58,702	
Handheld PC / PDA (Clam Shell)	39	2,503	Yes
Palm sized PC / PDA	16	145	Yes
Server-class computer	1	288	
Uncategorized	3	121	
Audio / Video	**1,037**	**44,483**	
Hands-free device	660	4,657	Yes
Wearable headset device	203	2,503	Yes
Set-top box	98	35,409	
Video Display and Loudspeaker	31	324	
Loudspeaker	20	365	
Portable audio	13	107	Yes
HiFi Audio Device	3	15	
Uncategorized	9	1,103	
Phone	**574**	**27,125**	
Cellular	362	14,659	Yes
Smart phone	206	12,149	Yes
Wired modem or voice gateway	3	308	
Cordless	2	8	
Uncategorized	1	1	
Imaging	**76**	**1,816**	
Printer	70	1,728	
Scanner	6	88	
Peripheral	**64**	**362**	
Keyboard-Uncategorized device	13	67	
Pointing device - Uncategorized device	17	71	
Not Keyboard / Not Pointing Device - Gamepad	8	139	
Not Keyboard / Not Pointing Device - Joystick	8	48	
Pointing device - Digitizer tablet	3	14	
Not Keyboard / Not Pointing Device - Uncategorized device	15	23	
Toy - Robot	**3**	**8**	
Wearable - Wrist Watch	**2**	**39**	**Yes**
Miscellaneous	**86**	**2,514**	
Uncategorized	**154**	**4,135**	
Grand Total	**3,274**	**306,763**	

5.4 Results

5.4.1 Predicting Mode and Trip Purpose from GPS data

The predictive models used were multinomial logit models. This is an application of logistic regression used for categorical response data. Logistic regression is a generalized linear model used for binary response variables and takes the following form.

$$log\left[\frac{P(Y_i = 1)}{P(Y_i = 0)}\right] = \beta_0 + \beta_1 x_{i_1} + \beta_2 x_{i_2} + \cdots + \beta_p x_{i_p}$$

where Y is the response variable which equals either 1 or 0, the x's are the explanatory variables, and the β's are the coefficients predicted by regression analysis. To define multinomial logit models: suppose Y_i=1, 2 … K denotes K categories, and let π_{i1}, π_{i2} … π_{iK} be the probabilities of each category. The multinomial logit model is then defined by the following equations.

$$log\left[\frac{\pi_{i_2}}{\pi_{i_1}}\right] = \beta^{(2)}{}_0 + \beta^{(2)}{}_1 x_{i_1} + \beta^{(2)}{}_2 x_{i_2} + \cdots + \beta^{(2)}{}_p x_{i_p}$$

$$log\left[\frac{\pi_{i_3}}{\pi_{i_1}}\right] = \beta^{(3)}{}_0 + \beta^{(3)}{}_1 x_{i_1} + \beta^{(3)}{}_2 x_{i_2} + \cdots + \beta^{(3)}{}_p x_{i_p}$$

$$\vdots$$

$$log\left[\frac{\pi_{i_K}}{\pi_{i_1}}\right] = \beta^{(K)}{}_0 + \beta^{(K)}{}_1 x_{i_1} + \beta^{(K)}{}_2 x_{i_2} + \cdots + \beta^{(K)}{}_p x_{i_p}$$

This is a series of logistic regression models where each one compares the probability of category 1, the reference category, to the probability of a different category. Some algebra to solve for the probabilities gives the following.

$$\pi_{i_c} = \frac{e^{\eta_{i_c}}}{1 + \sum_{t=2}^{K} e^{\eta_{i_t}}}$$

where $\eta_{i_c} = \beta^{(c)}{}_0 + \beta^{(c)}{}_1 x_{i_1} + \beta^{(c)}{}_2 x_{i_2} + \cdots + \beta^{(c)}{}_p x_{i_p}$

To create the models numerous variables were considered: day of the week, time of day in minutes, time period of day (morning, afternoon, evening, or night), trip duration, average speed, standard deviation of speed, number of stops, start latitude, start longitude, end latitude, end longitude, dwell time (time spent without moving) before trip, dwell time after trip, start land use, end land use, gps distance, percent of the trip that occurred on the street and path network, percent of the trip that occurred on a bus route (PercBus), percent of trip that occurred on a trail (PercTrail), distance.

To predict mode, a multinomial logit model was used with the volunteers' stated mode as the response variable. Forward selection based on the Aikaike information criterion (AIC) selected the following explanatory variables: average speed in miles per hour, standard deviation of speed, PercTrail, PercBus, and distance in miles. The model coefficients are shown in Table 5.3; note that Drive is the reference category. Table 5.4 shows the number of trips whose mode this model correctly predicts as well as percentages correctly predicted. To account for model optimization, ten random iterations of 10-fold cross validation were performed. The column Cross-Percentage gives the prediction accuracy after accounting for model optimization. The model ended up predicting mode with a 77% accuracy.

Table 5.3 Mode prediction regression coefficients.

Mode	(Intercept)	Average Speed	Standard Deviation of Speed	PercTrail	PercBus	Distance
Bike	1.408	-0.117	-8%	781%	0.361	-0.223
Bus	-4.515	-0.201	19%	447%	3.081	0.121
Other	0.338	-0.23	12%	275%	0.73	-0.199
Walk	4.741	-0.521	21%	657%	0.575	-0.516

Table 5.4 Mode prediction accuracy.

Mode	Correctly Predicted	Incorrectly Predicted	Percent Correct	Cross-Percent
Drive	107	11	91%	89%
Bike	14	21	40%	35%
Bus	4	7	36%	29%
Other	0	14	0%	0%
Walk	107	9	92%	90%
Total	232	62	79%	77%

To predict purpose, a multinomial logit model was used with the volunteers' stated destination as the response variable. Forward selection based on the Aikaike information criterion (AIC) selected the following explanatory variables: destination land use (EndLU) (commercial, residential, or recreational), time of day in minutes since midnight, start land use (StartLU) (commercial, mixed, residential, or recreational), destination longitude (EndLon), duration in minutes, number of stops, starting longitude (StartLon), and dwell time after the trip

(EndDwell). The model coefficients are shown in Table 2.7; note that Home is the reference category. Table 5 shows the number of trips whose purpose this model correctly predicts as well as percentages correctly predicted. To account for model optimization, ten random iterations of 10-fold cross validation were performed. The column Cross-Percentage gives the prediction accuracy after accounting for model optimization. The model ended up predicting purpose with a 54% accuracy.

Table 5.5 Purpose prediction regression coefficients.

Purpose	(Intercept)	EndLU Commercial	EndLU Residential	Minutes	StartLU Mixed	StartLU Recreational
Other	-17.233	-0.657	-2.083	2.47E-05	53.113	0.208
School	465.422	317.192	313.949	-3.58E-03	-16.885	-56.137
Shopping	-7.033	237.577	233.978	4.65E-04	-6.633	-84.410
Work	-628.828	-0.930	-3.868	-4.49E-03	-8.239	-149.469

Table 5.6 Purpose prediction regression coefficients (continued).

Purpose	StartLU Residential	EndLon	Duration	NumStops	StartLon	EndDwell
Other	2.166	10.268	-3.29E-03	0.037	-10.423	-8.29E-04
School	1.981	31.762	4.89E-03	-0.205	-25.108	1.60E-04
Shopping	0.922	9.293	-5.71E-04	0.032	-7.313	-2.58E-04
Work	2.112	-6.875	-8.78E-03	-0.036	1.465	3.21E-04

Table 2.7 Purpose prediction accuracy.

Purpose	Correctly Predicted	Incorrectly Predicted	Percent Correct	Cross-Percent
Home	69	22	76%	70%
Other	34	37	48%	38%
School	62	20	76%	72%
Shopping	0	13	0%	0%
Work	13	24	35%	23%
Total	178	116	61%	54%

5.4.2 Processing Bluetooth Data

Below in Figure 5.4 you can see the location of each intercepted device during the study. (Note that this only shows Moscow, there were also data points going out to Pullman, Seattle, Lewiston, Sandpoint, and Spokane). Given that almost all our volunteers were students, it's not surprising that most of the data points are on the University of Idaho campus (the South Western Quadrant of Moscow), however there are quite a few points from around the city. This figure gives a good idea of the high traffic areas in Moscow.

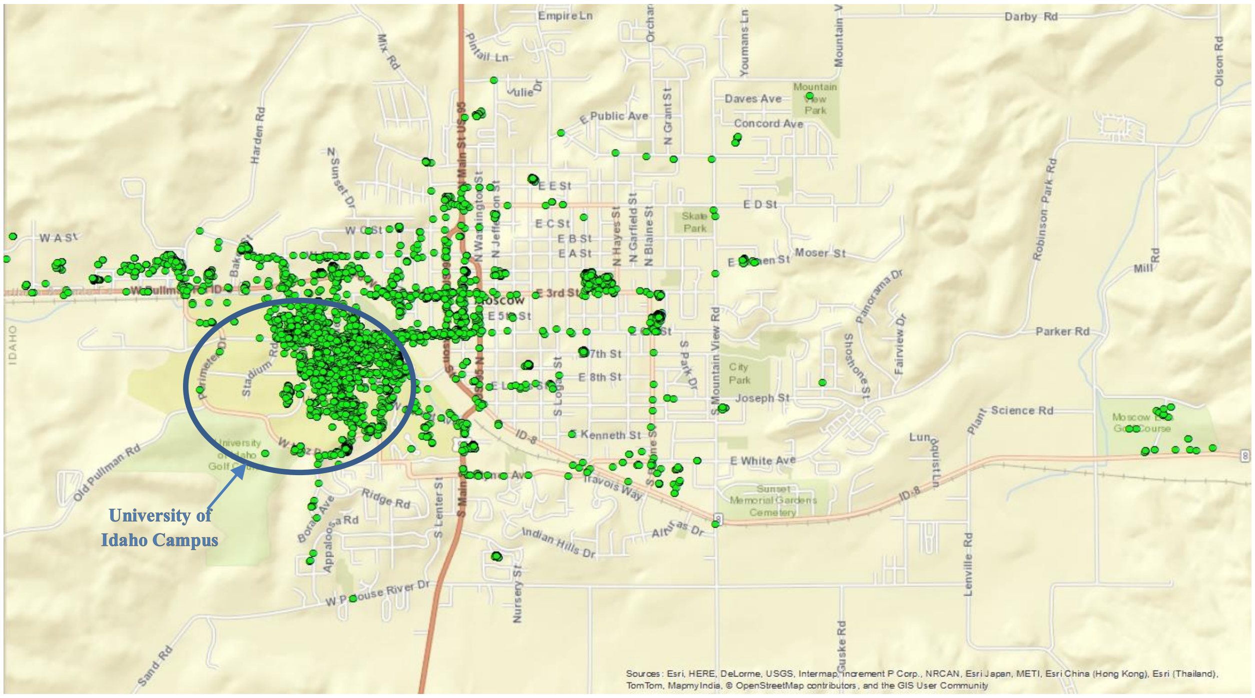


Figure 5.4 All intercepted devices.

Another way to view the data is on a daily basis. Figure 5.5 shows all the intercepted devices on Friday May 2nd, 2014, while Figure 5.6 shows Saturday May 3rd, 2014. Note that on Friday most of the activity is on campus; this is as expected since students have classes on Friday. On Saturday, however, there is much less activity on campus and more in other parts of town, especially a couple noted recreational areas.

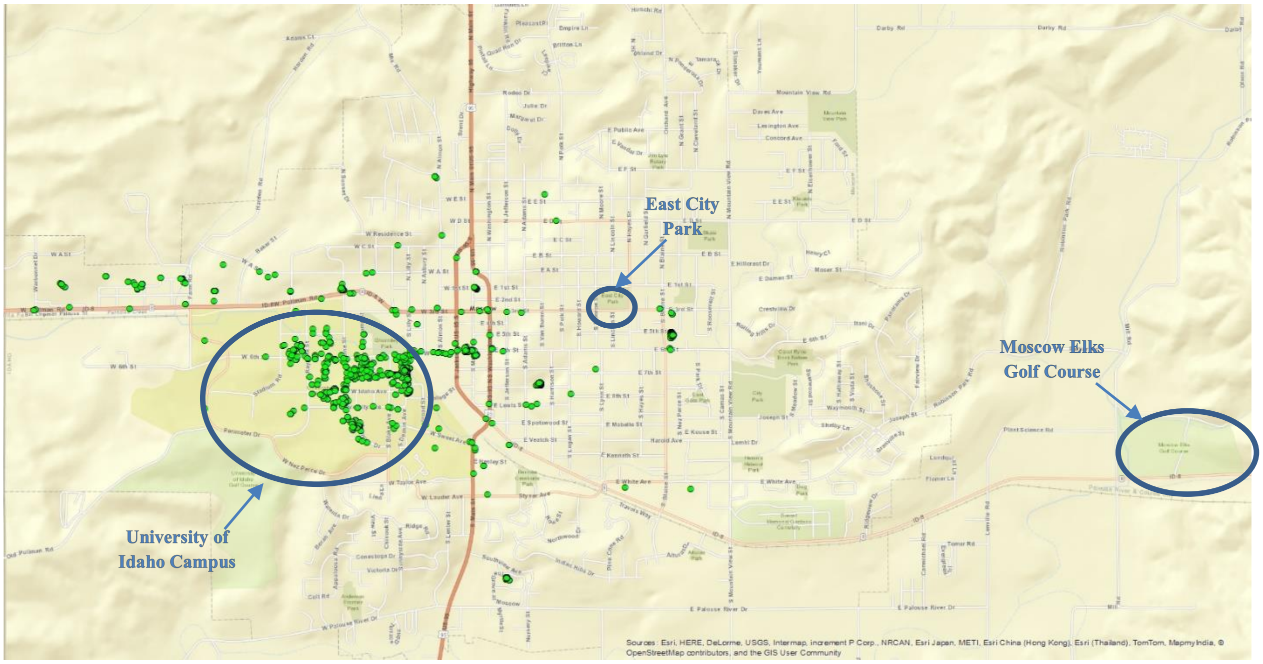


Figure 5.5 Bluetooth interception Friday May 2nd 2014.

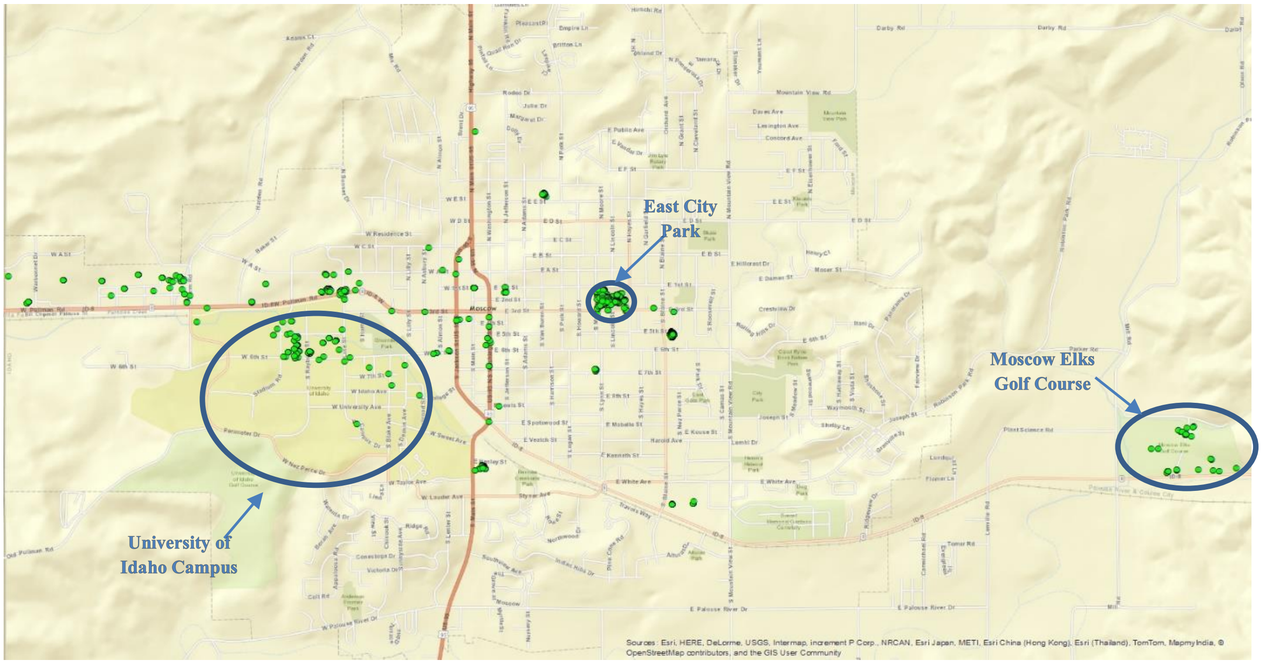


Figure 5.6 Bluetooth interception Saturday May 3rd 2014.

The data can also be viewed hourly. Figure 5.7 shows the number of intercepted mobile devices each hour. Note that the data peaks around 11 AM and then has a minimum around 2 AM. This has more activity in the middle of the day than is usually seen because college students are constantly moving around between classes.

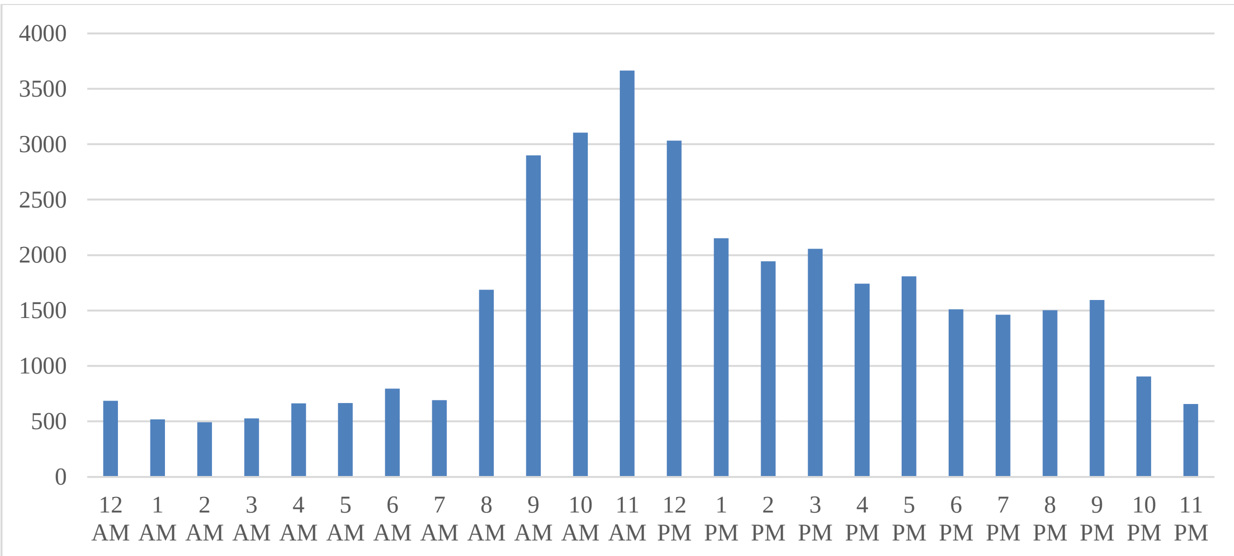


Figure 5.7 Bluetooth detections by hour.

If the MAC address of the intercepted devices is considered, as well as time and location, then movement of individual devices can be tracked. About half of the intercepted devices were seen more than once. This means that for these 721 devices we could detect movement by matching the two points where the device was detected. Figure 5.8 shows all thirty-five thousand tracks for these devices. Figure 5.9 shows the time and distance between these matched points (Figure 5.9 only shows times up to 1 hour and distances up to 3 miles).

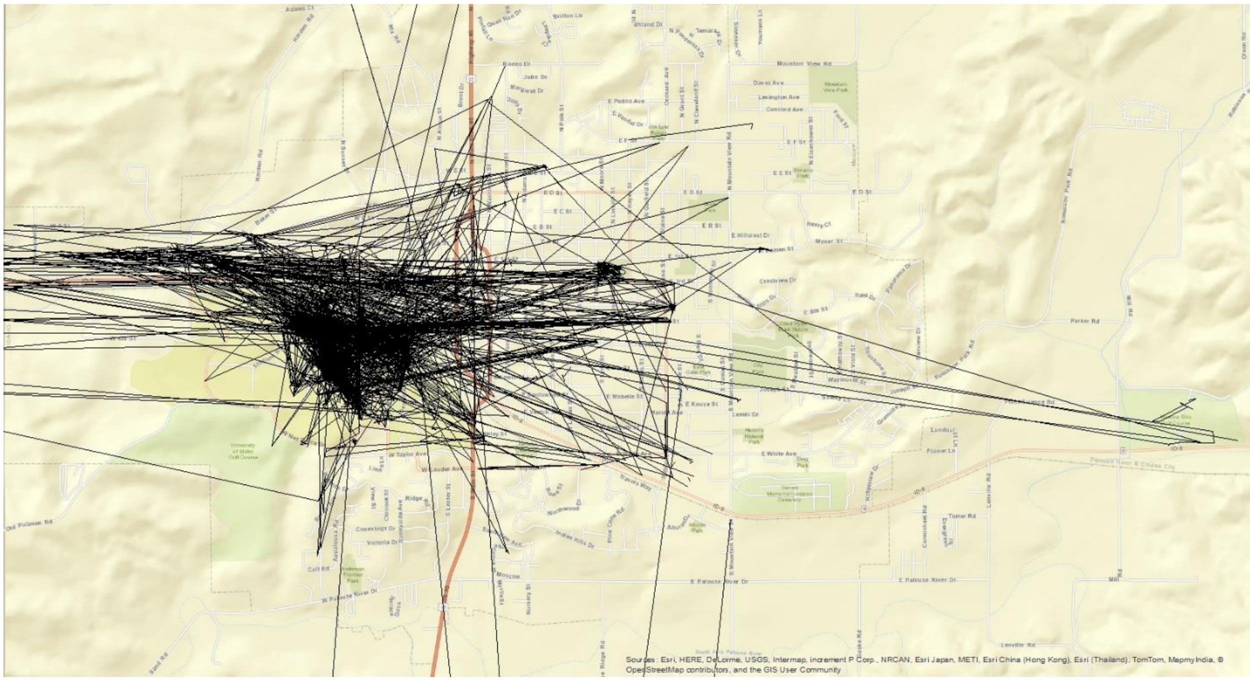


Figure 1.8 Matched Bluetooth points.

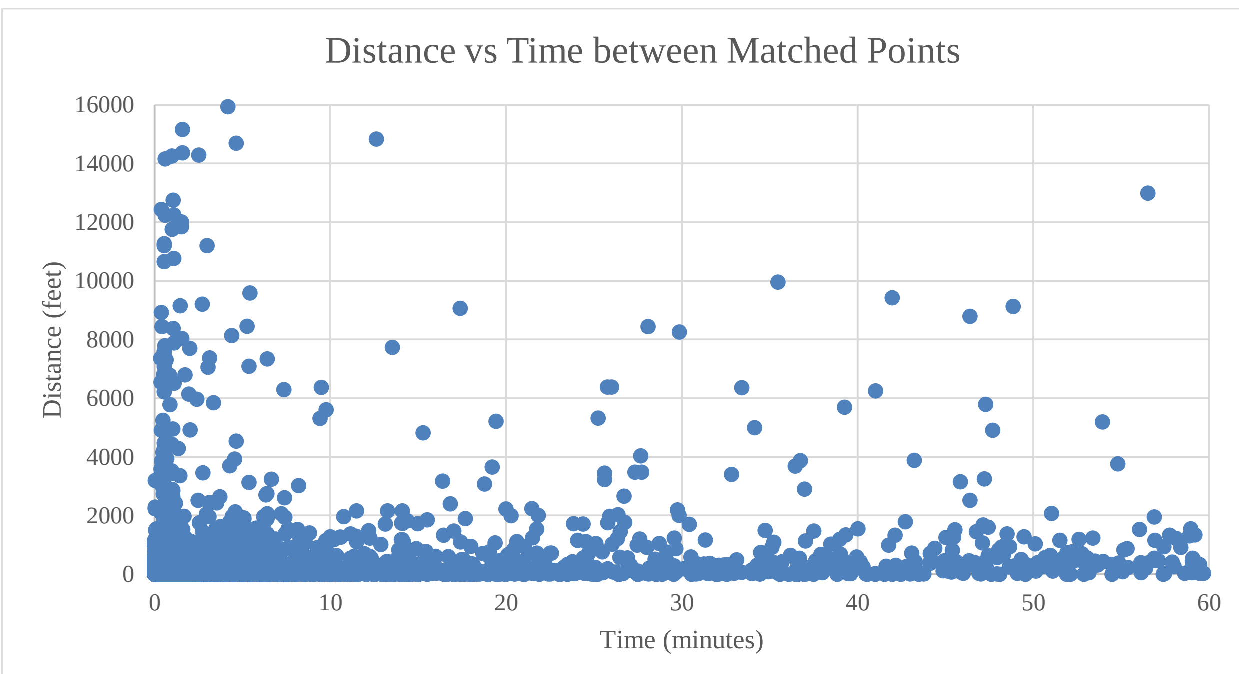


Figure 5.9 Distance vs time between matched points.

One goal with these matched points could be to identify trips. However, most of these matched points had long gaps of time between multiple sightings which makes it hard to determine that the device traveled directly from the first point to the second (they could have made significant detours that weren't detected). Also, there were many times were a device was detected multiple times in the same location, so these matched points don't create trips. A trip can be identified if a device is detected in at least two different places over a short period of time. This experiment detected almost 600 trips, all of which are shown in Figure 5.10.

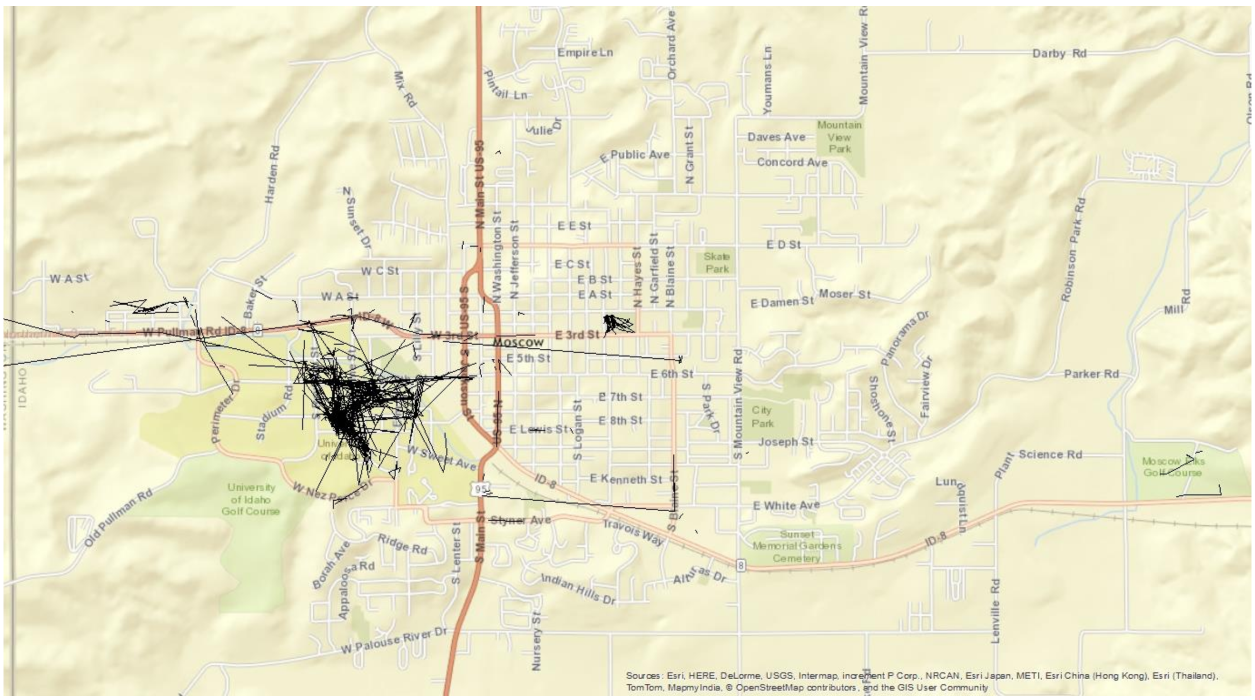


Figure 5.10 Identified trips.

5.5 Conclusion

As a proof of concept this experiment was successful. With only 53 participants, there was quite a bit of useful information to be pulled from the Bluetooth detection data. High traffic areas (as well as how this changes from one day to the next), peak travel hours, and trips can all be identified.

It's worth noting that with only 53 participants (the majority of which were students on campus most of the day) the coverage area of the network is quite limited. For example, on Friday May 2nd there could have been quite a bit of traffic in Mountain View Park due to a parade, but if none of the participants were there then this would go undetected. In addition, while some trips were identified, it is difficult to tell what route was used as the individuals were only detected at the beginning and end of their trip. However, if there were more participants and a wider geographic distribution of participants both of these shortcomings would be reduced.

In conclusion, this data collection method works and can provide a considerable amount of data for very little cost.

Chapter 9 Bluetooth Data Collection on Transit

6.1 Introduction

The synergistic relationship between non-motorized travel and public transit is both intuitive and well documented in research. Public transit use has been shown to be correlated with higher rates of non-motorized travel, and is recognized as a key determinant of the accessibility of an urban area by biking and walking modes (Besser & Dannenberg, 2005; Chanam & Moudon, 2006; Bachand-Marleau et al., 2011). A number of cities in the US have identified the role of transit in increasing the number and distance of trips accessible by active travel modes, and considered this interaction in planning transit and cycling facilities. For example, the city-wide Capitol Bikeshare program in Washington D.C. has placed a great deal of emphasis on transit connectivity, and the most recent user survey for the program indicates that approximately 64% of all trips began or ended at a metro station (LDA Consultants, 2015; Pucher & Buehler, 2008). However, despite the increasing emphasis on transit and non-motorized mode integration over the last few years, enhancing this trend has not been prioritized to a large enough extent in policy making (Wang & Liu, 2013).

A number of previous studies have suggested that better integration between non-motorized modes and transit is a fundamental step toward increasing walking and cycling rates. For example, Lachapelle & Noland (2012) showed a statistically significant relationship between transit use and walking rates, and conclude that access to transit via non-motorized modes can result in more non-motorized trip chains and higher walking rates overall. Likewise, the analysis described in Martin & Shaheen (2014) suggests that coordination between transit and public bike sharing programs has the potential to both augment existing transit service and to reduce the reliance on transit for many users. Cheng & Liu (2012) evaluated factors that correspond to a cyclist's perceived convenience, and their findings indicate that the presence of "intra-transit" cyclist facilities is a key factor in a traveler's decision to make cycling part of their commute. Some studies have addressed the need to identify the most cost effective improvements in terms of transit/cycling integration (e.g. Krizek & Stonebraker, 2011), but it is clear that such efforts can benefit from more detailed and accurate data describing local transit use patterns.

The majority of past work in data acquisition for transit planning has been based on traveler surveys, which are known to be susceptible to stated preference bias and error (Wardman, 1988). In recent years, there has been a strong and growing interest in using passenger smart card and other automated fare collection (AFC) system data for data collection (Pelletier et al., 2011). However, smart card and other AFC systems are designed for a very specific purpose, and as such are not typically optimized for traveler data collection. Though some methods have been developed to extract meaningful information from AFC systems, this process generally involves the fusion of multiple data sources and can be quite complicated (e.g. Zhao et al., 2007; Ma et al., 2012). Because of this, though the majority of major transit agencies have already implemented some form of automated payment, many do not have a mature framework for extracting mobility patterns form the resulting data. Further, even with the growing popularity of AFC, many smaller and specialized transit agencies do not have the resources or financial incentive to adopt this technology.

There has been significant recent interest in onboard mobile communication device sensing for transit data collection, which promises to address many of the shortcomings of survey or AFC based methods. In principle, multiple sensors are used to record the unique Bluetooth or Wi-Fi media access control (MAC) address for each wireless communication device at points along their travel route. By re-identifying a traveler at multiple locations over time, it is possible to reconstruct entire trips and even compute travel speed. This technology has been used to some extent in a range of transportation applications, including corridor travel time estimation, pedestrian travel analysis, and vehicle origin-destination data collection (Bachmann et al., 2013; Malinovskiy et al. 2012; Blogg et al., 2010). Mobile MAC address sensing has also been applied to some extent in transit data collection. For example, Vassilis et al. (2013) showed that bus-mounted Bluetooth MAC address sensors can be used to infer traveler boarding and alighting locations and produce relatively detailed information about the temporal and spatial distribution of demand. Similarly, Canon-Lozano et al. (2013) demonstrated a web-service platform for origin-destination estimation based on static Bluetooth sensor data.

Though previous work has addressed some of the difficulties inherent to mobile device MAC address sensing, a number of challenges remain. As pointed out in several studies (Canon-Lozano et al., 2013; Malinovskiy et al., 2012), many individuals carry more than one wireless communications device, and a greater number choose to disable the Bluetooth and Wi-Fi capabilities of their device(s) when not in use. Additionally, GPS errors and extraneous detections are a constant concern, and a substantial amount of preprocessing work must be completed to insure that such observations are corrected or discarded. Thus, the focus of this project is demonstrating a set of methodologies to collect, process, and analyze wireless sensor-based transit traveler data.

6.2 Methods

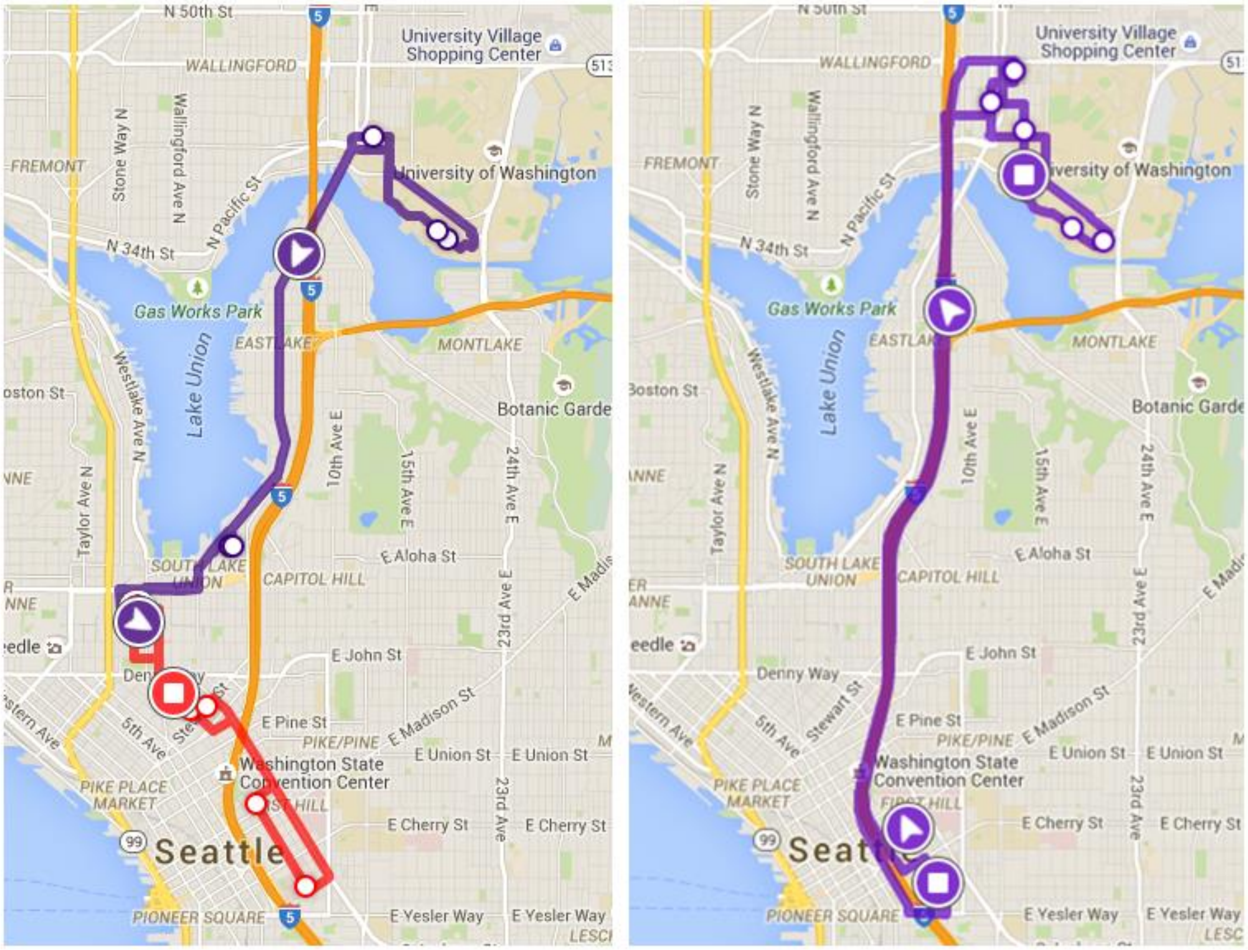
6.2.1 Field Experiment Overview

Field experiments were conducted on transit vehicles owned and operated by the University of Washington Transportation Services (UWTS). The University of Washington operates a small fleet of busses serving traffic between the UW Medical Center and Harborview Medical Center and various locations near the UW campus. A dedicated shuttle service also serves traffic between the South Lake Union Neighborhood and UW Medical Center. These services provide free and accessible travel for customers of the UW hospital facilities, many of whom rely on this service to access the care and medical attention they need. The data collected in this experiment will aid UWTS in assessing the efficiency, reliability and customer satisfaction of their services, and may be used for future service planning and funding applications. UWTS expressed a strong interest in improving passenger origin/destination estimates, and was eager to support data collection efforts.

The Health Sciences Express (see Figure 6.1 (a)) provides transit services between the UW Medical Center and Harborview Medical Center, with stops at the UW Roosevelt Clinic and the UW Tower. UW faculty, staff, students, and medical center patients all have access to this service. Health Sciences Express buses depart UWMC every 15 minutes from 6 a.m. to 5:45 p.m., Monday through Friday, excluding University holidays. The evening service is sponsored by Graduate Medical Education, and leaves UWMC at 6:15 p.m. and HMC at 7 p.m. The UW/Fred Hutch South Lake Union Shuttle (see Figure 6.1 (b)) consists of two routes, the UWMC to South Lake Union via Fred Hutch, and South Lake Union to Harborview. Both routes include stops at several additional locations, including medical research facilities.

(a) Health Sciences Express bus **(b) Shuttle**
Figure 6.1 Two types of transit services offered by UWTS.

Two types of transit services offered by UWTS, South Lake Union Shuttle and Health Sciences Express, are shown above in Figure 6.1 ((a) and (b) respectively). This map was obtained from the UWTS real-time bus location mapping application, so the arrow and square icons in Figure 6.2 indicate the location and travel direction of in-service transit vehicles. Both of the routes shown below were equipped with sensors in this study, but the analysis and results are only presented for the Health Science Express route shown on the right in Figure 6.2.



(a) South Lake Union Shuttle **(b) Health Science Express**

Figure 6.2 UWTS transit routes (UWTS, 2015).

The stop information for the Health Science Express route can be found in Table 6.1, which lists the location and building of each stop. There are a total of eight stops, including those that are used only for passenger drop off.

Table 6.1 UWTS bus and shuttle routes (UWTS, 2015).

University of Washington Medical Center (UWMC) – Harborview Medical Center (HMC)

Building	Location	Stops
Medical Facilities Stops		
UWMC - Cascade Tower	NE Columbia Rd outside the Cascade Tower on the southeast corner of UWMC.	Stop 1
UWMC - D-Wing	NE Columbia Rd at the D-Wing shelter.	Stop 2
UW Tower	West side of 12th Ave. NE, halfway between NE 43rd St. and NE 45th St.	Stop 4 - to HMC Stop 7 - to UWMC
Roosevelt Clinic	West side of Roosevelt Way NE, south of NE 43rd Street.	Stop 5 - to HMC Stop 8 - to UWMC
HMC	Rider Alert: Between Dec 1 and 5 this stop will be located outside the Maleng Building on 9th Ave due to a lane closure in front of the Research & Training Building. East side of 9th Avenue in front of the Harborview Research & Training Building.	Stop 6
Flag Stops		
Brooklyn Avenue Flag Stop	Northeast corner of Brooklyn Avenue NE and NE Campus Parkway at the Metro bus stop. Bus stops outbound to Roosevelt Clinic and HMC.	Stop 3
Inbound Drop-Off Only (No Pick-Up)		
Lander Hall	NE Campus Parkway, between 12th Ave NE and Brooklyn Ave NE. Bus stops inbound to UWMC from HMC.	N/A
Hitchcock Hall	At the passenger load zone in front of the Rotunda on NE Pacific Street. Bus stops inbound to UWMC from HMC.	N/A

Field data collection was conducted in two separate events during the time periods from March to May, 2015, with each events lasting about two weeks. At any given time on a typical day of operation, there are 4 buses and 3 shuttles in service. All the buses and shuttles instrumented with Bluetooth/Wi-Fi detection devices which will be introduced in the following section. The detection devices are supplied with power through the bus electronic system in order to insure that the data collection period is not limited by battery life.

6.2.2 Data Collection Hardware and Software

Sensing hardware used in this experiment consists of a set of mobile phones running the Android operating system. For the mobile sensors, power adapters were connected directly to the transit vehicle power supply, and connected to the phones with a standard USB charging cable. Each stationary sensor was connected to an auxiliary battery to extend the deployment time, and placed inside a weather-proof box.

A mobile phone Application (app) was used to detect the Bluetooth and Wi-Fi devices within nearby spaces. The app used in this study was developed using the Funf in a Box (FIAB) platform, a simple application development tool which provides access to mobile device internal sensors. FIAB is part of the Funf Open Sensing Framework, an open source API designed to give application developers and researchers a set of tools for acquiring, processing, and managing Android mobile device sensor data (Aharony et al. 2011). To use FIAB, the user specifies the data collection parameters (e.g. sensor types, observation frequency, etc.) and a Dropbox folder, and an app is automatically generated to collect and upload data to the specified location. The application can then be downloaded directly or distributed through the Android Market. In this study, sensor data was collected and stored in device memory, and uploaded to Dropbox whenever the UW Wi-Fi network is available. Otherwise, data can be obtained data manually by connecting the mobile device to the laptop. The app creation interface is shown in Figure 6-3.

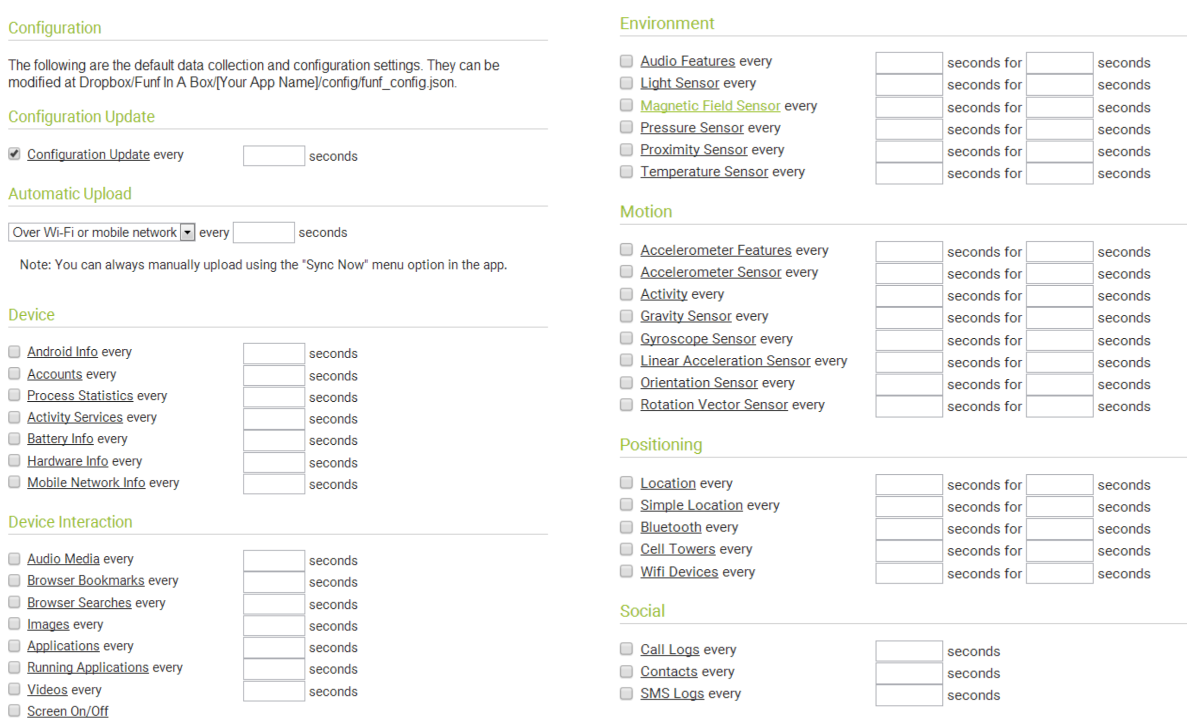


Figure 6.3 Interface of the app creation process (FOSF, 2015).

In this study, the Bluetooth and Wi-Fi MAC addresses for all devices within sensing range are collected, as well as the GPS location of the data collection device. For Bluetooth detection, data is recorded for 10 seconds on 20 second intervals. For the Wi-Fi detection, data is recorded for 5 seconds on 20 second intervals. GPS location is recorded every 20 sections for a duration of 4 seconds. Battery status is also collected every 2 hours, though the devices were supplied with power through the bus electronic system.

6.2.3 Sensor Placement

As described previously, both mobile (bus-mounted) and static sensors were used in this data collection effort. Mobile sensors were positioned in near the driver's seat in the bus and wired into the bus power supply, and so did not rely on the device battery for power. Static sensors were placed in weather-proof cases near bus stop locations, and so were equipped with auxiliary batteries to increase the operating time and reduce the need to recharge. Figure 6.4 (below) shows the placement of the mobile sensors, which were plugged into a power adapter wired into the 24V power panel in the busses. Figure 6.5 shows the wiring and placement of the static sensor, in this case in a tree adjacent to a bus stop.

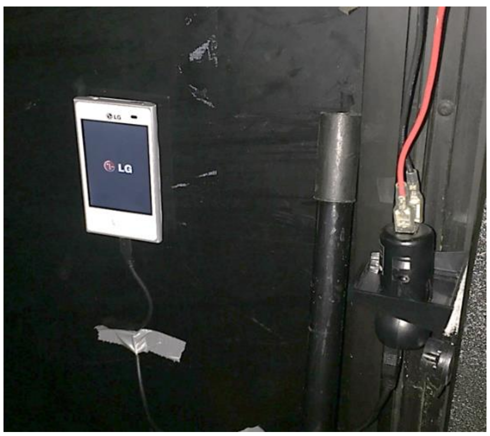


Figure 6.4 Mobile sensor placement.

(a) Sensor with auxiliary battery **(b) Sensor placement**

Figure 6.5 Static sensor setup and placement.

6.2.4 Manual Data Collection

For validation purposes, a number of volunteers were asked to ride the instrumented busses at various intervals throughout the data collection process. Volunteers were given instructions to board a bus at the UWMC stop, and record the number of boarding and alighting passengers at each stop over a ~ 1 hour period. Over the course of the data collection period, approximately 17 hours of manual count data was collected during the mid-day and peak time periods.

Manual boarding counts were also collected by bus drivers on the South Lake Union shuttles, and made available to the research team by UWTS. These counts are collected as a regular course of operation, and provided a useful additional source of data for this project.

6.2.5 GPS Data Collection

The format of the GPS data recorded in the field experiment is illustrated in Table 6.2. This raw data contains a great deal of supporting information including detection id, accuracy, timestamp, and others, not all of which is useful in our analysis. Only the three pieces of information that are used in subsequent analysis are shown in Table 6.3. The table shows the GPS location of the bus, represented by the latitude and longitude coordinates, and the time of observation. This data can be combined with the MAC address sensing information to determine the approximate location of each device at the time when it was detected.

Table 6.2 Raw GPS data from mobile sensors.

id	device	timestamp	mAccuracy	mAltitude	mBearing	mExtras_r	mExtras_r	mHasAccur	mHasAltit	mHasBeari	mHasSpeed	mLatitude	mLongitud	mProvider	mSpeed	mTime	timestamp
1cc4274a-	64d5dd97-	1.43E+09	24	6.5	0			TRUE	TRUE	FALSE	TRUE	47.65198	-122.316	gps	0	1.43E+12	1.43E+09
1cc4274a-	64d5dd97-	1.43E+09	33	0	0	cached	wifi	TRUE	FALSE	FALSE	FALSE	47.65242	-122.316	network	0	1.43E+12	1.43E+09
1cc4274a-	64d5dd97-	1.43E+09	36	0	0	cached	wifi	TRUE	FALSE	FALSE	FALSE	47.65243	-122.316	network	0	1.43E+12	1.43E+09
1cc4274a-	64d5dd97-	1.43E+09	36	0	0	cached	wifi	TRUE	FALSE	FALSE	FALSE	47.65243	-122.316	network	0	1.43E+12	1.43E+09
1cc4274a-	64d5dd97-	1.43E+09	24	6.5	0			TRUE	TRUE	FALSE	TRUE	47.65198	-122.316	gps	0	1.43E+12	1.43E+09
1cc4274a-	64d5dd97-	1.43E+09	36	0	0	cached	wifi	TRUE	FALSE	FALSE	FALSE	47.65243	-122.316	network	0	1.43E+12	1.43E+09
1cc4274a-	64d5dd97-	1.43E+09	39	0	0	cached	wifi	TRUE	FALSE	FALSE	FALSE	47.6524	-122.316	network	0	1.43E+12	1.43E+09
1cc4274a-	64d5dd97-	1.43E+09	39	0	0	cached	wifi	TRUE	FALSE	FALSE	FALSE	47.6524	-122.316	network	0	1.43E+12	1.43E+09
1cc4274a-	64d5dd97-	1.43E+09	25	-2.6	0			TRUE	TRUE	FALSE	TRUE	47.65191	-122.316	gps	0	1.43E+12	1.43E+09
1cc4274a-	64d5dd97-	1.43E+09	25	-2.6	0			TRUE	TRUE	FALSE	TRUE	47.65191	-122.316	gps	0	1.43E+12	1.43E+09
1cc4274a-	64d5dd97-	1.43E+09	25	-2.6	0			TRUE	TRUE	FALSE	TRUE	47.65191	-122.316	gps	0	1.43E+12	1.43E+09
1cc4274a-	64d5dd97-	1.43E+09	33	0	0	cached	wifi	TRUE	FALSE	FALSE	FALSE	47.65243	-122.316	network	0	1.43E+12	1.43E+09
1cc4274a-	64d5dd97-	1.43E+09	37	0	0	cached	wifi	TRUE	FALSE	FALSE	FALSE	47.65243	-122.316	network	0	1.43E+12	1.43E+09
1cc4274a-	64d5dd97-	1.43E+09	37	0	0	cached	wifi	TRUE	FALSE	FALSE	FALSE	47.65243	-122.316	network	0	1.43E+12	1.43E+09
1cc4274a-	64d5dd97-	1.43E+09	25	-2.6	0			TRUE	TRUE	FALSE	TRUE	47.65191	-122.316	gps	0	1.43E+12	1.43E+09
1cc4274a-	64d5dd97-	1.43E+09	37	0	0	cached	wifi	TRUE	FALSE	FALSE	FALSE	47.65243	-122.316	network	0	1.43E+12	1.43E+09
1cc4274a-	64d5dd97-	1.43E+09	37	0	0	cached	wifi	TRUE	FALSE	FALSE	FALSE	47.65243	-122.316	network	0	1.43E+12	1.43E+09
1cc4274a-	64d5dd97-	1.43E+09	37	0	0	cached	wifi	TRUE	FALSE	FALSE	FALSE	47.65243	-122.316	network	0	1.43E+12	1.43E+09
1cc4274a-	64d5dd97-	1.43E+09	25	-2.6	0			TRUE	TRUE	FALSE	TRUE	47.65191	-122.316	gps	0	1.43E+12	1.43E+09
1cc4274a-	64d5dd97-	1.43E+09	37	0	0	cached	wifi	TRUE	FALSE	FALSE	FALSE	47.65243	-122.316	network	0	1.43E+12	1.43E+09
1cc4274a-	64d5dd97-	1.43E+09	43	0	0	cached	wifi	TRUE	FALSE	FALSE	FALSE	47.65253	-122.316	network	0	1.43E+12	1.43E+09
1cc4274a-	64d5dd97-	1.43E+09	43	0	0	cached	wifi	TRUE	FALSE	FALSE	FALSE	47.65253	-122.316	network	0	1.43E+12	1.43E+09
1cc4274a-	64d5dd97-	1.43E+09	25	-2.6	0			TRUE	TRUE	FALSE	TRUE	47.65191	-122.316	gps	0	1.43E+12	1.43E+09
1cc4274a-	64d5dd97-	1.43E+09	43	0	0	cached	wifi	TRUE	FALSE	FALSE	FALSE	47.65253	-122.316	network	0	1.43E+12	1.43E+09
1cc4274a-	64d5dd97-	1.43E+09	40	0	0	cached	wifi	TRUE	FALSE	FALSE	FALSE	47.65244	-122.316	network	0	1.43E+12	1.43E+09
1cc4274a-	64d5dd97-	1.43E+09	40	0	0	cached	wifi	TRUE	FALSE	FALSE	FALSE	47.65244	-122.316	network	0	1.43E+12	1.43E+09
1cc4274a-	64d5dd97-	1.43E+09	25	-2.6	0			TRUE	TRUE	FALSE	TRUE	47.65191	-122.316	gps	0	1.43E+12	1.43E+09
1cc4274a-	64d5dd97-	1.43E+09	40	0	0	cached	wifi	TRUE	FALSE	FALSE	FALSE	47.65244	-122.316	network	0	1.43E+12	1.43E+09
1cc4274a-	64d5dd97-	1.43E+09	40	0	0	cached	wifi	TRUE	FALSE	FALSE	FALSE	47.65244	-122.316	network	0	1.43E+12	1.43E+09
1cc4274a-	64d5dd97-	1.43E+09	40	0	0	cached	wifi	TRUE	FALSE	FALSE	FALSE	47.65244	-122.316	network	0	1.43E+12	1.43E+09
1cc4274a-	64d5dd97-	1.43E+09	30	-3.6	0			TRUE	TRUE	FALSE	TRUE	47.6519	-122.316	gps	0	1.43E+12	1.43E+09
1cc4274a-	64d5dd97-	1.43E+09	30	-3.6	0			TRUE	TRUE	FALSE	TRUE	47.6519	-122.316	gps	0	1.43E+12	1.43E+09
1cc4274a-	64d5dd97-	1.43E+09	30	-3.6	0			TRUE	TRUE	FALSE	TRUE	47.6519	-122.316	gps	0	1.43E+12	1.43E+09
1cc4274a-	64d5dd97-	1.43E+09	40	0	0	cached	wifi	TRUE	FALSE	FALSE	FALSE	47.65244	-122.316	network	0	1.43E+12	1.43E+09

Table 6.3 GPS data retained for analysis.

Latitude	Longitude	timestamp
47.65241647	-122.304183	1425429196
47.6525856	-122.305113	1425429196
47.65241647	-122.304183	1425429216
47.6499316	-122.309938	1425429216
47.6498885	-122.309717	1425429219
47.6498885	-122.309717	1425429219
47.65241647	-122.304183	1425429236

6.2.6 Bluetooth and Wi-Fi detection

The raw Bluetooth detection data is shown in the Table 6-4 below, with identifying information censored for privacy reasons. The Wi-Fi dataset is very similar to the Bluetooth, and so is not shown for brevity. This table includes the MAC address specific to each device, the name of the device given by the passenger, the timestamp of the detection, and others, not all of which was useful in our analysis. Some information such as “name” could be considered invasive, and because it is not useful for our study was deleted. Only the two most important information are retained: the timestamp of detection and the MAC address, as shown in Table 6-5 (again, censored for privacy reasons).

Table 6.4 Raw Bluetooth detection data.

id	device	timestamp	android.b	android.bluetooth.de	android.bluetooth.de	android.blu	timestamp
d2f0f866-b53c-4571-bl	64d5dd	1430254854	5898756	1C:62:B8: : :		-100	1430254854
d2f0f866-b53c-4571-bl	64d5dd	1430254860	1060	CC:6D:A0: : :		-97	1430254860
d2f0f866-b53c-4571-bl	64d5dd	1430254863	1060	CC:6D:A0: : :		-97	1430254863
d2f0f866-b53c-4571-bl	64d5dd	1430254961	2360324	00:1D:86: : :		-87	1430254961
d2f0f866-b53c-4571-bl	64d5dd	1430254982	15205380	00:08:E0: : :		-99	1430254982
d2f0f866-b53c-4571-bl	64d5dd	1430255013	2098184	00:05:4F: : :		-81	1430255013
d2f0f866-b53c-4571-bl	64d5dd	1430255074	2098184	10:C6:FC: : :		-92	1430255074
d2f0f866-b53c-4571-bl	64d5dd	1430255093	66816	00:A0:96: : :		-75	1430255093
d2f0f866-b53c-4571-bl	64d5dd	1430255158	9472	00:17:E9: : :		-98	1430255158
d2f0f866-b53c-4571-bl	64d5dd	1430255177	2098184	10:C6:FC: : :		-90	1430255177
d2f0f866-b53c-4571-bl	64d5dd	1430255197	5898756	00:1D:29: : :		-93	1430255197
d2f0f866-b53c-4571-bl	64d5dd	1430255274	5898764	E4:12:1D: : :		-95	1430255274
d2f0f866-b53c-4571-bl	64d5dd	1430255274	5898764	E4:12:1D: : :		-92	1430255274
d2f0f866-b53c-4571-bl	64d5dd	1430255276	5374468	74:45:8A: : :		-93	1430255276
d2f0f866-b53c-4571-bl	64d5dd	1430255402	4063500	30:10:B3: : :		-93	1430255402
d2f0f866-b53c-4571-bl	64d5dd	1430255437	12452108	24:FD:52: : :		-93	1430255437
d2f0f866-b53c-4571-bl	64d5dd	1430255441	3670276	00:25:00: : :		-90	1430255441
d2f0f866-b53c-4571-bl	64d5dd	1430255503	2098180	00:13:8A: : :		-90	1430255503
d2f0f866-b53c-4571-bl	64d5dd	1430255682	7936	00:1A:5D: : :		-92	1430255682
d2f0f866-b53c-4571-bl	64d5dd	1430255693	7936	00:1A:5D: : :		-97	1430255693
d2f0f866-b53c-4571-bl	64d5dd	1430255717	3408904	90:03:B7: : :		-90	1430255717
d2f0f866-b53c-4571-bl	64d5dd	1430255719	3670276	EC:35:86: : :		-96	1430255719
d2f0f866-b53c-4571-bl	64d5dd	1430255758	3670284	28:CF:E9: : :		-97	1430255758
d2f0f866-b53c-4571-bl	64d5dd	1430255761	2098184	10:C6:FC: : :		-97	1430255761
d2f0f866-b53c-4571-bl	64d5dd	1430255797	2098184	10:C6:FC: : :		-95	1430255797
d2f0f866-b53c-4571-bl	64d5dd	1430255799	2098184	10:C6:FC: : :		-91	1430255799
d2f0f866-b53c-4571-bl	64d5dd	1430255801	2098184	10:C6:FC: : :		-94	1430255801
d2f0f866-b53c-4571-bl	64d5dd	1430255815	2098184	10:C6:FC: : :		-91	1430255815
d2f0f866-b53c-4571-bl	64d5dd	1430255817	2098184	00:1E:B2: : :		-91	1430255817
d2f0f866-b53c-4571-bl	64d5dd	1430255817	2098184	10:C6:FC: : :		-92	1430255817

Table 6.5 Bluetooth data retained for analysis.

MAC_ID	timestamp
B8:76:3F:XX: XX:XX	1425429217
04:0C:CE: XX: XX: XX	1425429238
18:83:31: XX: XX:XX	1425429239
28:CF:E9: XX: XX: XX	1425429244
18:83:31: XX:XX:XX	1425429257
00:22:41: XX: XX: XX	1425429318
00:1B:63: XX:XX:XX	1425429322
00:1D:4F: XX:XX: XX	1425429325

6.2.7 Combination of GPS, Bluetooth and Wi-Fi Data

The GPS, Bluetooth and Wi-Fi data are combined by matching the timestamp information in each data element. In many cases, an exact match between the timestamps will not be found due to differences in the observation intervals. In this case, each Bluetooth or Wi-Fi detection will be matched to the closest available GPS observation. An example of combined Bluetooth and GPS data is shown in Table 6.6. A unique proxy id code is used instead of the actual device MAC address for privacy reasons.

Table 6.6 Combined Bluetooth and GPS data.

BTtime	GPStime	Latitude	Longitude	ID
1425430857	1425430856	47.65853	-122.318	1213
1425431044	1425431036	47.65853	-122.318	1053
1425431484	1425431476	47.60392	-122.323	1012
1425431959	1425431956	47.60392	-122.323	639
1425432606	1425432599	47.6529	-122.312	1273
1425433343	1425433336	47.65579	-122.315	22
1425434219	1425434219	47.64398	-122.322	39

With the combined information, the location of each Bluetooth or Wi-Fi detection can be visualized in an ArcGIS map. Figure 6.6 shows the location of Bluetooth Detections onboard a single vehicle, bus #592. Figure 6.7 shows the location of Wi-Fi Detections onboard Bus #592. Most of the detections are along the bus routes, but several appear well outside the bus operating area. This indicates that some data pre-processing will be required before the data can be used for any analysis.

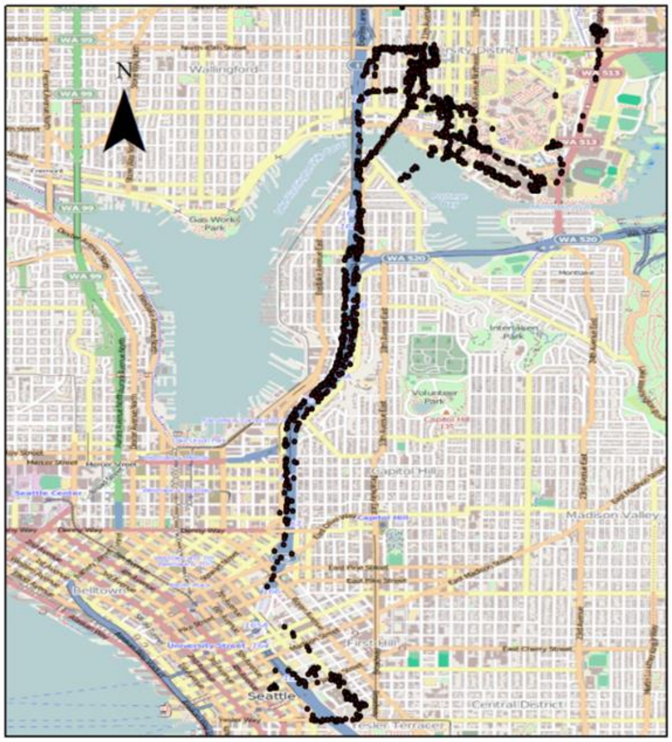


Figure 6.6 Bluetooth detections onboard bus #592.

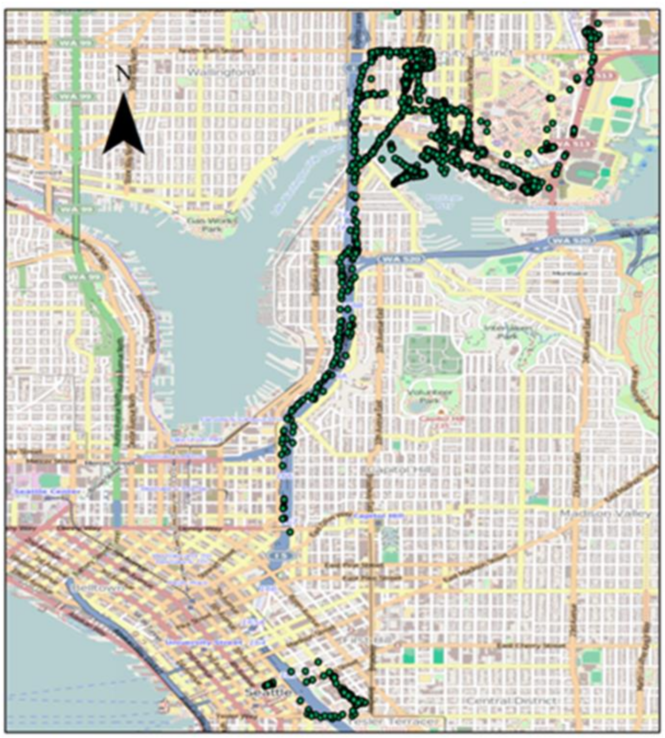


Figure 6.7 Wi-Fi detections onboard bus #592.

6.3 Results

6.3.1 Data Processing

The Bluetooth and Wi-Fi data collected in the field experiment contain a great number of errors and inconsistencies, and a certain amount of extraneous data was collected. For example, when a bus stops, the Bluetooth and Wi-Fi from both the passengers boarding the bus and the non-passengers who are near the bus will be detected. Clearly, this non-passenger data should be excluded from further analysis. In addition, some people may turn on or off their Bluetooth or Wi-Fi signal during the trip. In such cases, it is difficult to infer the boarding or alighting bus stop.

Data processing was conducted to exclude confounding data from further analysis. The data cleaning process in this project contains three principle steps. The first step is to delete random detections. To do this, we rely on the fact that a passenger should be detected repeatedly during a bus trip. Thus, the MAC address with fewer than 3 detections were excluded from the final dataset. In this way, most of the random non-passenger data is removed be removed.

In the second step, the data is processed to separate the trip rides from the detection records. Based on historical bus operating records, the running time between two bus stops is always above 60 seconds and the time from the initial stop to the last stop on a route is less than 1.5 hours. Thus, in this project, if no two detections for a single device are more than 60 s apart, it is not considered a trip ride. Also, if the time difference between two detections is over 5400 s, they are considered to belong to two separate rides. All detections for a given device within the 60 s and 5400 s windows are considered to belong to the same trip ride. After this step, the origin timestamp and destination timestamp for each trip, together with the origin location and destination location, can be inferred. The resulting origin and destination are shown in Figure 6-8.

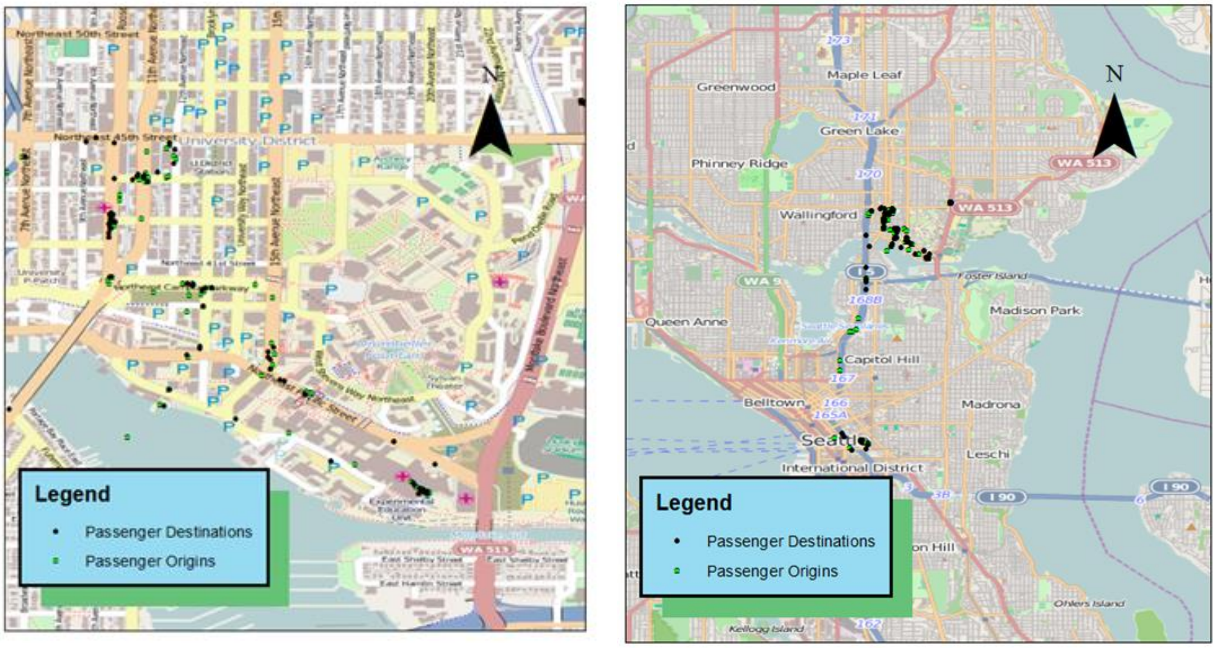


Figure 6.8 Passenger origins and destinations from Bluetooth detection data.

The third step matches the origin and destination of riders to transit stop locations. In each trip, the initial and final timestamp and GPS location are recorded. The trips are selected only if the origins and destinations of rides are within 300 ft of known transit stop location. This step can delete the device trips which do not start and end at bus stops. To perform the data cleaning in this step, a bus stop location layer is first created in an ArcGIS map, as shown in Figure 6-9. The exact latitude and longitude coordinates of each bus stop are specified in the layer. Next, a spatial analysis is conducted using common GIS tools to select the data within bus stop areas. The results of data selection after step 3 are shown in Figure 6-10. The detected origins and destinations are now clustered around the bus stop location.

Figure 6.9 Bus stop locations.

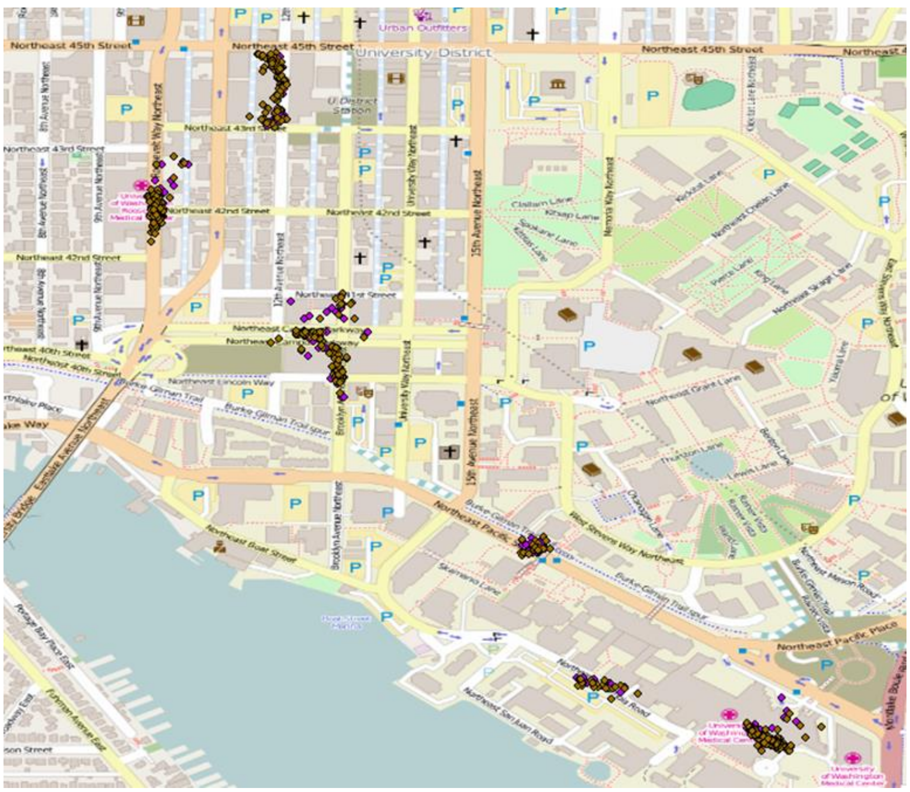


Figure 6.10 Distribution of detections near bus stop locations.

The number of samples remaining after each data cleaning step is shown in Table 6.7. The raw data contains 11041 devices and 58668 rides. After step 1, only 6476 devices and 14126 rides are left. The number of devices is reduced by about 50% while the number of rides are reduced by 75%. This indicates that a great number of devices and are detected fewer than three times. Such devices are most likely other road users near the sensor on the bus, rather than the bus passengers. After filter step 2, only 1220 devices and 5426 rides remain. After filter step 3, only 403 rides are left. Those rides can be used for obtaining the origin and destination information. Note that, even though a great deal of data was discarded in due to the strict cleaning rules, the remaining dataset should be very reliable for inferring passenger origins and destinations.

Table 6.7 Detection count after each data cleaning step.

	Number of devices	Number of Rides
Raw Data	11041	58688
After Filter 1	6476	14126
After Filter 2	1220	5426
After Filter 3	NA	403

6.3.2 Data Analysis Results

With the processed Bluetooth detection data, we plotted the number of detections at different time a single week. The results are shown in Figure 6.11. It shows that data were detected during day time when bus is in operation, with minimal variation from day to day.

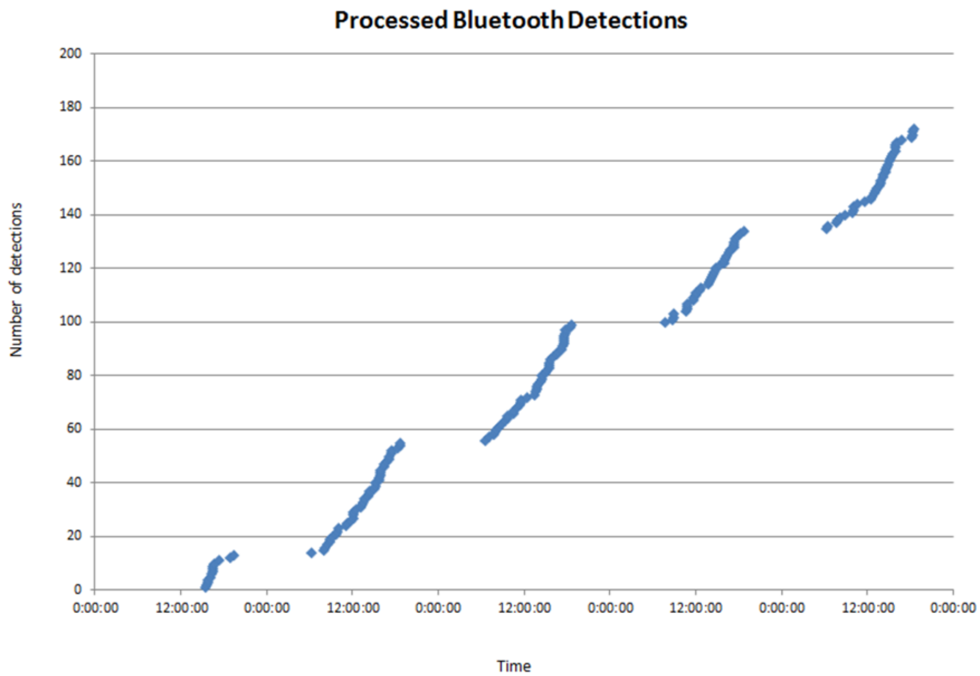


Figure 6.11 Number of detections over time.

The distribution of trip duration for the processed Bluetooth is shown in Figure 6.12. It is clear that the comparatively few longer duration trips were recorded, and that most of the trips are within 10 minutes.

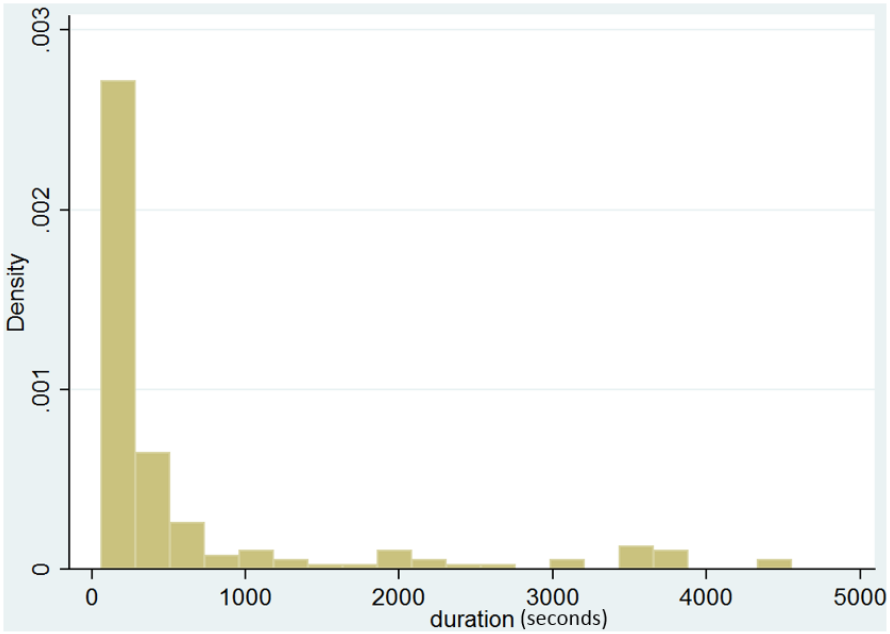


Figure 6.12 Distribution of trip time duration.

The origin-destination matrix for the processed Bluetooth data is shown in Table 6.8. The left-most column represents the origin bus stop, and the first row represents the destination. The color in the table represents the number of rides for each combination of origin and destination. The most popular trip is from the Roosevelt stop to the HVC stop, with 92 passengers boarding at the Roosevelt stop and alighting at the HVC stop. Other popular trips include HVC to Roosevelt, Hitchcock Hall to Cascade Tower, Lander Hall to Brooklyn, and UW Tower to Roosevelt. The percentage of trip for each origin-destination pair is estimated as shown in Table 6-9.

Table 6.8 Estimated origin-destination matrix (bus stops).

O/D	HMC	D-Wing	Hitchcock Hall	Brooklyn and Lincoln Way	Lander Hall	Roosevelt	UW Tower	Cascade Tower
HMC	0	0	0	2	0	47	13	0
D-Wing	0	0	0	3	1	2	0	8
Hitchcock Hall	0	0	0	0	0	0	0	33
Brooklyn and Lincoln Way	1	2	2	0	12	10	1	12
Lander Hall	0	2	0	16	0	4	1	5
Roosevelt	92	3	1	3	4	0	12	2
UW Tower	14	0	0	1	0	16	0	1
Cascade Tower	1	11	13	6	1	2	0	0

Table 6.9 Percentages of origin-destination trips (bus stops).

O/D	HMC	D-Wing	Hitchcock Hall	Brooklyn and Lincoln Way	Lander Hall	Roosevelt	UW Tower	Cascade Tower
HMC	0.00%	0.00%	0.00%	0.56%	0.00%	13.06%	3.61%	0.00%
D-Wing	0.00%	0.00%	0.00%	0.83%	0.28%	0.56%	0.00%	2.22%
Hitchcock Hall	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	9.17%
Brooklyn and Lincoln Way	0.28%	0.56%	0.56%	0.00%	3.33%	2.78%	0.28%	3.33%
Lander Hall	0.00%	0.56%	0.00%	4.44%	0.00%	1.11%	0.28%	1.39%
Roosevelt	25.56%	0.83%	0.28%	0.83%	1.11%	0.00%	3.33%	0.56%
UW Tower	3.89%	0.00%	0.00%	0.28%	0.00%	4.44%	0.00%	0.28%
Cascade Tower	0.28%	3.06%	3.61%	1.67%	0.28%	0.56%	0.00%	0.00%

The research team processed the Wi-Fi data on Bus #592 from April 27 to May 1. The same cleaning procedure was followed to process the raw data, and the resulting origin and destination points are shown in Figure 6.13. This shows that, similar to the Bluetooth data, the Wi-Fi trip starts/ends are clustered around the bus stops after cleaning is completed. The origin-destination matrix is calculated for the Wi-Fi data, shown in Table 6.10 and Table 6.11. In comparing the results from Bluetooth and Wi-Fi data, the origin-destination matrices appear somewhat similar. Based on Wi-Fi data, the most popular trip is from the Roosevelt stop to the HMC stop. Other popular routes include UW tower to HMC, HMC to Roosevelt and Cascade Tower, Lander Hall to Cascade Tower, Cascade Tower to UW Tower, and UW Tower to Roosevelt.

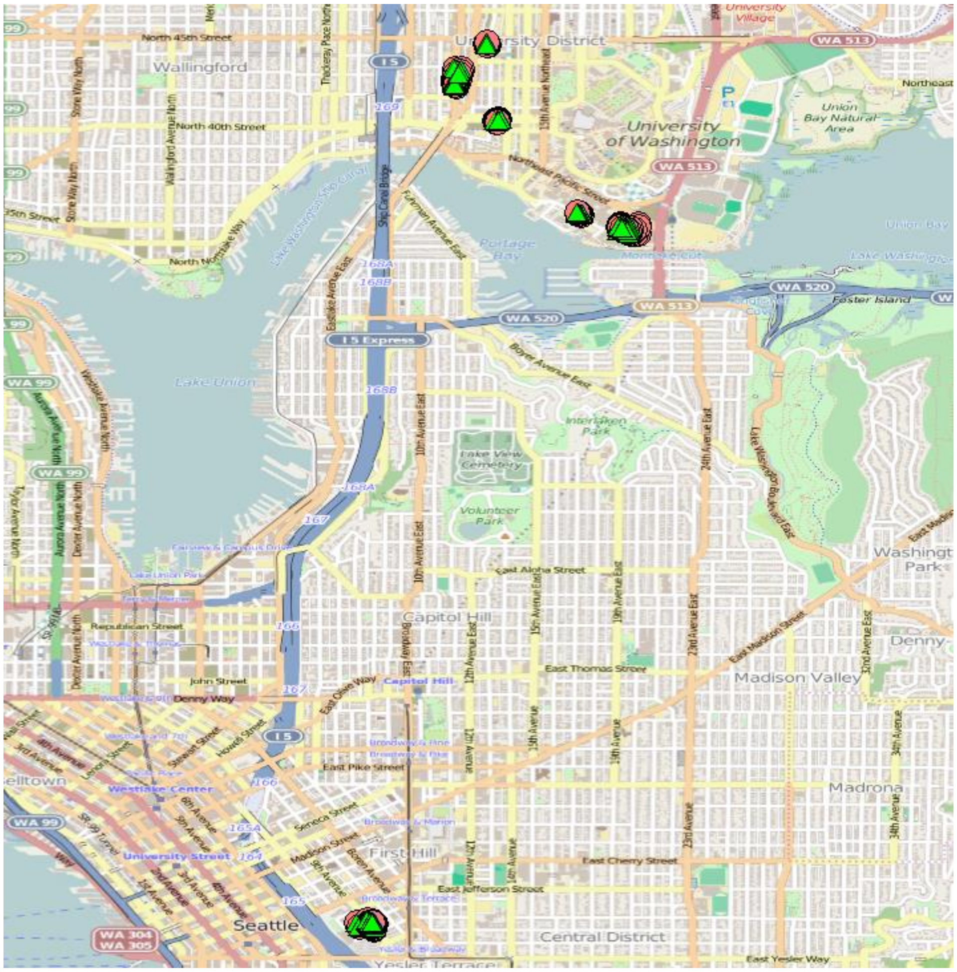


Figure 6.13 Origins and destinations from Wi-Fi data.

Table 6.10 Estimated origin-destination matrix from Wi-Fi (bus stops).

O/D	HMC	D-Wing	Hitchcock Hall	Brooklyn and Lincoln Way	Lander Hall	Roosevelt	UW Tower	Cascade Tower
HMC	0	2	0	0	3	20	16	23
D-Wing	11	0	0	0	1	2	3	20
Hitchcock Hall	0	0	0	0	0	0	0	0
Brooklyn and Lincoln Way	0	0	0	0	0	0	0	0
Lander Hall	0	0	0	0	0	0	0	26
Roosevelt	40	4	0	0	2	0	0	7
UW Tower	33	0	0	0	0	20	0	6
Cascade Tower	6	16	0	0	4	9	20	0

Table 6.11 Percentages of origin-destination trips from Wi-Fi (bus stops).

O/D	HMC	D-Wing	Hitchcock Hall	Brooklyn and Lincoln Way	Lander Hall	Roosevelt	UW Tower	Cascade Tower
HMC	0.00%	0.68%	0.00%	0.00%	1.02%	6.80%	5.44%	7.82%
D-Wing	3.74%	0.00%	0.00%	0.00%	0.34%	0.68%	1.02%	6.80%
Hitchcock Hall	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Brooklyn and Lincoln Way	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Lander Hall	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	8.84%
Roosevelt	13.61%	1.36%	0.00%	0.00%	0.68%	0.00%	0.00%	2.38%
UW Tower	11.22%	0.00%	0.00%	0.00%	0.00%	6.80%	0.00%	2.04%
Cascade Tower	2.04%	5.44%	0.00%	0.00%	1.36%	3.06%	6.80%	0.00%

6.4 Discussion

The results presented here are based on a subset of the total data collected, as collection efforts are ongoing. Further, though a sizable quantity of manual count data was obtained for verification and validation, these efforts are ongoing and so no results are presented here. That said, the methods and results described in this report provide a solid illustration of what is possible through wireless MAC address sensing. This study shows that the number of passengers carrying active wireless devices is sufficient to provide a sizable sample of the population, and that the boarding and alighting location of passengers can be inferred from detection time and GPS location. In communications with the UWTS, it is clear that they have a great deal of interest in obtaining low-cost origin destination data and time distribution of ridership. As noted previously, certain UWTS buses have manual data collection equipment which allow the driver to tabulate the boardings at each stop. Compared to this approach, MAC address sensing requires no effort on the part of the driver and provides both boarding and alighting location for a sample of passengers.

There are some obvious technological and functional differences between Wi-Fi and Bluetooth sensing which must be considered in developing a data collection and analysis methodology. Most importantly, Wi-Fi has a longer detection range, and so a greater area is covered by each sensor. Though this results in a greater number of detections, it adds a significant amount of spatial uncertainty to the data. As a result, Wi-Fi data is much noisier than

Bluetooth data and contains a greater number of extraneous detections. This characteristic of longer detection range can be helpful in cases where a large area must be covered by a single sensor, for example, if a single device needs to cover a courtyard or parking lot. For on-board transit passenger detection, Bluetooth may provide more useable data. In any case, more work is needed to determine the optimal hardware type and configuration for different detection scenarios.

One of the challenges that must be addressed in future work arises from the fact that extraneous detections are more likely to occur when there are a large number of non-passenger individuals near bus stop locations. Intuitively, it is likely that pedestrians and other non-passenger travelers tend to appear in greater numbers near popular bus stops where the majority of passengers board and alight. Because of this, great care must be taken in the data cleaning step to avoid discarding useful data. This issue may be resolved through the use of more sophisticated data filtering methods, a topic that will be examined in greater detail as verification work is completed.

6.5 Conclusion

The work described in the section demonstrates a possible application of wireless MAC address sensing in transit data acquisition. A field experiment was designed and executed, and a set of data cleaning and processing methods were developed and applied to the field data. Based on the research team's experiences, it is thought that the methods applied here can be expanded temporally and spatially to provide a robust and detailed source of data for transit planning and operations analysis. With high temporal resolution, low initial investment, and near real-time monitoring, it is clear that wireless sensing has a number of potential advantages over conventional data collection methods. Additional advantages include:

- Low hardware, software, and installation costs
- Minimal maintenance of and access to hardware is required once it has been installed
- Automated, wireless data acquisition and upload
- Observed traveler behavior, as opposed to survey or other stated preference data

One additional point that should be made with regard to the work described here is that wireless MAC address sensing is not limited to any one travel mode in particular, making integrated motor vehicle, non-motorized, and transit travel data collection possible. One can imagine a network of low cost sensors providing rich, real-time information about inter- and intra-mode travel patterns on a variety of on-road and active transportation facilities.

Chapter 11 Conclusions and Recommendations

The goal of this project was to provide guidance for improving manual bicycle and pedestrian count programs and investigate the feasibility of new, cutting-edge sensor technology. To this end, five independent studies were pursued to:

1. Evaluate the state-of-the-practice of manual bicycle and pedestrian count programs,
2. Assess the error associated with manual count programs,
3. Develop a process for using manual count data for safety analysis,
4. Explore the feasibility of using moving Bluetooth sensors to collect bicycle and pedestrian data, and
5. Investigate the potential for installing Bluetooth sensors on public buses to gain better understanding of pedestrian interaction with public transportation.

Through an online survey of transportation specialists, including Bicycle and Pedestrian Coordinator for various state DOTSs, it is clear that many communities are using community-volunteers to collect manual bicycle and pedestrian counts as a low cost strategy. Furthermore, the respondents reported many communities are not using automatic counters, not using adjustment factors, counting infrequently, for short 2 hour periods, and rarely for more than one day. Although occasionally or regularly conducting manual counts has a variety of benefits, it is not clear how the resulting snapshot counts can be used for some of the reasons the survey respondents claimed were motivating their count programs. Count program managers should conduct a careful self-assessment of their capabilities, limitations, and goals to clearly define their purpose for conducting manual count programs that involve community-volunteers. In particular count program managers should decide if they are striving to have a public involvement activity that produces count data as a byproduct, or vice-versa, or a program that strives to achieve both goals equally. A community that lacks resources for permanent counters might conclude that manual counts with error are better than no counts at all.

For the second part of this project a controlled field experiment found a slight difference in absolute percent error between different count techniques; however, the differences were not statistically significant. Nevertheless, since the 12-movement technique provides substantially richer data without degrading accuracy, manual count program managers should consider using the 12-movement technique. This study also showed no apparent increase in measurement error when counters were assigned to collect additional information about passing travelers. The survey of practitioners showed that many count programs assign their volunteers to collect additional information, and a few noted that sex of traveler and helmet use are used to measure the success of programs to encourage helmet use and increase bicycling among women. However, feedback from the participants in our count study suggests the task of collecting additional information can be burdensome. Manual count program managers might consider focusing count events that involve community-volunteers as an opportunity to simultaneously collect count data across a wide geographic area and, instead, collect additional information such as helmet use. Also, the data could be used by city staff on a different occasion when this information is needed to assess specific programs or for grant applications.

For the third part of this project a new method was developed for engineers and planners to evaluate projects for sketch-level scenario planning in terms of bicyclist exposure to dangerous situations. The results from the new method can be presented through tables, figures, and maps to elected officials, the general public, and other stakeholders at town hall meetings or other forums for deliberative decision-making. The method and tools are intuitive and easy to use. Central to the new method is a typology of 23 dangerous situations for bicyclists that were identified through a literature review. The dangerous situation typology provides a framework for future research. Perhaps SPFs could be developed based on exposure indicators by investigating correlation with bicycle accident data. Research should seek to develop thresholds for dangerous situations not included in this study. For example, one possible approach to analyze “sidewalk riding” would be to assume that streets passing through areas of vibrant commercial activity, such as a lively Main Street or “Urban Village”, would have a percentage of sidewalk riding if there is not a bike lane. “Winding roads” might be analyzed through some sophisticated spatial GIS procedure. “Crowded path” could be analyzed with GIS data about painted center lines and pedestrian volumes. Finally, the tool could be enhanced through an optimization procedure that could identify high priority projects by minimizing dangerous situation exposure.

For the fourth study, we explored the feasibility of using moving Bluetooth sensors to collect bicycle and pedestrian data. Fifty three people were provided a smartphone app that tracked their location via global positioning systems (GPS) every seven seconds for 10 days. The app also collected, via Bluetooth, the media access control address (MAC address) of any intercepted device within 100 feet. The intent of the study was twofold, first to explore the potential of using this passive data collection technology to replace traditional travel diaries. The study also provided a means to collect a richer data set of travel patterns by matching intercepted MAC addresses across the city, perhaps as a precursor to the data that might be collected in the future through vehicle-to-vehicle communications. Every night the participants logged into a website to verify information about their travel. They were asked to confirm trip segments, trip purpose, and trip mode, all of which had been inferred through algorithms we developed for the study. Our model predicted mode with 77% accuracy and purpose with 54% accuracy. Over

3,000 devices were intercepted via Bluetooth and recorded over 300,000 GPS points. By matching MAC addresses across town, our data processing algorithms identified over 600 trips, most of which were determined to be pedestrian trips on the University of Idaho campus.

For the fifth study, smartphones with the GPS/Bluetooth app were placed on buses traveling between the University of Washington campus and South Lake Union Neighborhood in Seattle. During the study period 11,041 devices were intercepted. After data processing we determined that 403 bus riders were intercepted. Boarding and alighting locations were determined for the intercepted riders. This study showed that the number of passengers carrying active wireless devices is sufficient to provide a sizable sample of the population, and that the boarding and alighting location of passengers can be inferred from detection time and GPS location. Based on the research team's experiences, it is thought that the methods applied here can be expanded temporally and spatially to provide a robust and detailed source of data for transit planning and operations analysis. With high temporal resolution, low initial investment, and near real-time monitoring, it is clear that wireless sensing has a number of potential advantages over conventional data collection methods. Additional advantages include:

- Low hardware, software, and installation costs
- Minimal maintenance of and access to hardware is required once it has been installed
- Automated, wireless data acquisition and upload
- Observed traveler behavior, as opposed to survey or other stated preference data

References

Aharony, Nadav, Wei Pan, Cory Ip, Inas Khayal, and Alex Pentland. "Social fMRI: Investigating and shaping social mechanisms in the real world." Pervasive and Mobile Computing 7, no. 6 (2011): 643-659.

Allen-Munley, C., Daniel, J., and Dhar, S. (2004). Logistic Model for Rating Urban Bicycle Route Safety. Transportation Research Record, 1878(1), 107–115.

Alta Planning + Design. (2004) San Francisco's Shared lane Pavement Markings: Improving Bicycle Safety. Prepared for San Francisco Department of Parking and Traffic.

Aultman-Hall L, Lane D, Lambert R. 2009. Assessing impact of weather and season on pedestrian traffic volumes. Transportation Research Record: Journal of the Transportation Research Board 2140:35-43.

Bachand-Marleau, Julie, Jacob Larsen, and Ahmed M. El-Geneidy. "Much-anticipated marriage of cycling and transit." Transportation Research Record: Journal of the Transportation Research Board 2247, no. 1 (2011): 109-117.

Bachmann, Chris, Matthew J. Roorda, Baher Abdulhai, and Behzad Moshiri. "Fusing a bluetooth traffic monitoring system with loop detector data for improved freeway traffic speed estimation." Journal of Intelligent Transportation Systems 17, no. 2 (2013): 152-164.

Besser, Lilah M., and Andrew L. Dannenberg. "Walking to public transit: steps to help meet physical activity recommendations." American journal of preventive medicine 29, no. 4 (2005): 273-280.

Blogg, Miranda, Conor Semler, Manu Hingorani, and Rod Troutbeck. "Travel time and origin-destination data collection using Bluetooth MAC address readers." In Australasian Transport Research Forum, pp. 1-15. 2010.

Buehler, R., and J. Pucher. 2011. Sustainable Transport in Freiburg: Lessons from Germany's Environmental Capital. International Journal of Sustainable Transportation, Vol. 5, No. 1, Jan. 2011, pp. 43–70.

Canon-Lozano, Y., Angie Melo-Castillo, Cesar Augusto Gomez-Perilla, Klaus Banse, and L. F. Herrera-Quintero. "Web service platform for automatic generation of O/D matrix for mass transportation systems." In ITS Telecommunications (ITST), 2013 13th International Conference on, pp. 462-467. IEEE, 2013.

Cheng, Yung-Hsiang, and Kuo-Chu Liu. "Evaluating bicycle-transit users' perceptions of intermodal inconvenience." Transportation Research Part A: Policy and Practice 46, no. 10 (2012): 1690-1706.

City of Bellingham. (2013). Bellingham Bicycle Master Plan, Community Survey and Interactive Map Findings, Interactive Maps: Current and Desired Routes.

CROW. (2007) Design Manual for Bicycle Traffic. Record No. 25. Centre for Research and Contract Standardization in Civil and Traffic Engineering CROW, Ede.

Daniels, S., Brijs, T., Nuyts, E., and Wets, G. (2009). Injury crashes with bicyclists at roundabouts: influence of some location characteristics and the design of cycle facilities. Journal of Safety Research, 40(2):141-8.

Diogenes M, Greene-Roesel R, Arnold L, Ragland D. 2002. Pedestrian counting methods at in intersections: a comparative study. Transportation Research Record: Journal of the Transportation Research Board 2002:26-30.

El Esawey M. Lim C. Sayed T. Mosa A. 2013. Development of daily adjustment factors for Bicycle traffic. Journal of Transportation Engineering 139:859-871.

Elvik, R., Hoye, A., Vaa, T. (2009). The Handbook of Road Safety Measures. Emerald.

Federal Highway Administration. 2013. Traffic Monitoring Guide. Washington, DC, Report FHWA-PL-13-015.

Figliozzi M, Johnson P, Monsere C, Nordback K. 2014. Methodology to characterize ideal short-term counting conditions and improve AADT estimation accuracy using a regression-based correcting function. Journal of Transportation Engineering 140:04014014.

Forsyth, A., and J. M. Oakes. Cycling, the Built Environment, and Health: Results of a Midwestern Study. *International Journal of Sustainable Transportation*, Vol. 9, No. 1, Mar. 2013, pp. 49–58.

Funf Open Sensing Framework (FOSF). Funf in a box website. Last modified 3/25/2015. Accessed 5/10/2015.

Furth, P., Koonce, P., Yu, P., Peng, F., and Littman, M. (2014). Mitigating the Right Turn Conflict Using Protected-Yet-Concurrent Phasing for Cycle Track and Pedestrian Crossings. Submitted to the Transportation Research Board, August 1, 2013: 1–13.

Geller, R. (2007). Four Types of Cyclists. Portland, OR: City of Portland Office of Transportation, undated, circa 2007, http://www.portlandoregon.gov/transportation/44597?a=237507 (accessed April 23, 2014).

Griffin G, Nordback K, Götschi T, Stolz E, Kothuri S. 2014. Monitoring bicyclists and pedestrian travel and behavior: current research and practice. Transportation Research E-Circular E-C183.

Harkey, D., and Stewart, J. (1997). Evaluation of Shared-Use Facilities for Bicycles and Motor Vehicles. Transportation Research Record 1578

Highway Safety Manual. AASHTO, 2010.

Hurst, R. The Cyclist's Manifesto. Guilford, CT: Falcon Guides, Globe Pequot Press, 2009.

Huff H, Brozen M. 2014. Creating the bicycle count data clearinghouse for Los Angeles County, California. Presented at the 93rd Annual Meeting of the Transportation Research Board, 12-16 January 2014, Washington DC, USA.

Hunter, W., Stewart, J., Stutts, J., Huang, H., and Pein, W. (1999). A Comparative Analysis of Bicycle Lanes Versus Wide Curb Lanes : Final Report. FHWA-RD-99–034.

Jacobsen, P. (2003). Safety in numbers: more walkers and bicyclists, safer walking and bicycling. Injury Prevention, 9(3), 205–209.

Kang M, Ragan B, Park J. 2008. Issues in outcomes research: an overview of randomization Techniques for clinical trials. Journal of Athletic Training 43:215-221.

Kim, J., Kim, S., Ulfarsson, G., and Porrello, L. (2007). Bicyclist injury severities in bicycle-motor vehicle accidents. Accident Analysis and Prevention, 39(2), 238–51.

Klop, J., and Khattak, A. (1999). Factors influencing bicycle crash severity on two-lane, undivided roadways in North Carolina. Transportation Research Record, 1674(1), 78–85.

Krizek, K. and Roland, R. (2005). “What is at the end of the road? Understanding discontinuities of on-street bicycle lanes in urban settings.” Transportation Research Part D: Transport and Environment, 10(1), 55–68.

Krizek, Kevin J., and Eric W. Stonebraker. "Assessing options to enhance bicycle and transit integration." Transportation Research Record: Journal of the Transportation Research Board 2217, no. 1 (2011): 162-167.

Lachapelle, Ugo, and Robert B. Noland. "Does the commute mode affect the frequency of walking behavior? The public transit link." Transport policy 21 (2012): 26-36.

LDA Consulting. “2014 Capital Bikeshare Member Survey Report”. (2015).

Lee, Chanam, and Anne Vernez Moudon. "Correlates of walking for transportation or recreation purposes." Journal of Physical Activity & Health 3 (2006): S77.

Lindsey G. Chen J. Hankey S. 2013. Adjustment factors for estimating miles traveled by nonmotorized traffic. Presented at the 92nd Annual Meeting of the Transportation Research Board, 13-17 January 2013, Washington DC, USA.

Liu, F., J. Evans, and T. Rossi. (2012) Recent Practices in Regional Modeling of Nonmotorized Travel. Transportation Research Record, 2303(1) 1–8.

Los Angeles County Bicycle Coalition. 2013. Instructions for 2013 Bike + Ped Counters. Available at http://la-bike.org/countvolunteers [Accessed 31 July, 2014].

Lusk, A., Furth, P., Morency, P., Miranda-Moreno, L., Willett, W., and Dennerlein, J. (2011). Risk of injury for bicycling on cycle tracks versus in the street. Injury Prevention, 17(2), 131–5.

Ma, Xiao-lei, Yin-hai Wang, Feng Chen, and Jian-feng Liu. "Transit smart card data mining for passenger origin information extraction." Journal of Zhejiang University Science C 13, no. 10 (2012): 750-760.

Malinovskiy, Yegor, Nicolas Saunier, and Yinhai Wang. "Analysis of pedestrian travel with static bluetooth sensors." Transportation Research Record: Journal of the Transportation Research Board 2299, no. 1 (2012): 137-149.

Mapes, J. (2009) Pedaling Revolution: How Cyclists Are Changing American Cities. Oregon State University Press: Corvallis, OR.

Martin, Elliot W., and Susan A. Shaheen. "Evaluating public transit modal shift dynamics in response to bikesharing: a tale of two US cities." Journal of Transport Geography 41 (2014): 315-324.

McDaniel, S. and Lowry, M., and Dixon, M. (2014). Using Origin-Destination Centrality to Estimate Directional Bicycle Volumes. Transportation Research Record, in press.

McCarthy, M., and Gilbert, K. (1996). Cyclist road deaths in London 1985-1992: drivers, vehicles, manoeuvres and injuries. Accident Analysis and Prevention, 28(2), 275–9.

Mekuria, M., Furth, P., and Nixon, H. (2012). Low-Stress Bicycling and Network Connectivity.

Minikel, Eric. (2012). “Cyclist Safety on Bicycle Boulevards and Parallel Arterial Routes in Berkeley, California.” Accident Analysis and Prevention 45, 241–7.

Munizaga, Marcela A., and Carolina Palma. "Estimation of a disaggregate multimodal public transport Origin–Destination matrix from passive smartcard data from Santiago, Chile." Transportation Research Part C: Emerging Technologies 24 (2012): 9-18.

National Bicycle and Pedestrian Documentation Project. Available at http://bikepeddocumentation.org/ [Accessed 31 July, 2014].

Nordback K, Marshall W, Janson B, Stolz E. 2013. Estimating annual average daily bicyclists. Transportation Research Record: Journal of the Transportation Research Board 2339: 90-97.

Nordback, K., Marshall, W., Janson, B. (2014). Bicyclist Safety Performance Functions for a U.S. City. Accident Analysis and Prevention 65(C) 114–122.

Ortúzar, J. and L. Willumsen. Modelling Transport. Wiley, West Sussex, UK, 2011.

Parkin, J., and C. Meyers. The effect of cycle lanes on the proximity between motor traffic and cycle traffic. Accident Analysis & Prevention, Vol. 42, No. 1, Jan. 2010, pp. 159–165.

Pelletier, Marie-Pier, Martin Trépanier, and Catherine Morency. "Smart card data use in public transit: A literature review." Transportation Research Part C: Emerging Technologies 19, no. 4 (2011): 557-568.

Pucher, J., and Dijkstra, L. (2003). Promoting safe walking and cycling to improve public health: lessons from The Netherlands and Germany. American Journal of Public Health, 93(9), 1509–16.

Pucher, John, and Ralph Buehler. "Making cycling irresistible: lessons from the Netherlands, Denmark and Germany." Transport Reviews 28, no. 4 (2008): 495-528.

Räsänen, M., and Summala, H. (1998). Attention and expectation problems in bicycle-car collisions: an in-depth study. Accident Analysis and Prevention, 30(5), 657–66.

Reynolds, C., Harris, M, Teschke, K., Cripton, P., and Winters, M. (2009). The impact of transportation infrastructure on bicycling injuries and crashes: a review of the literature. Environmental Health, 8:47.

Ryan S. 2013. Establishing an automated regional non-motorized transportation data collection system to support active transportation performance monitoring. Presented at the 92nd Annual Meeting of the Transportation Research Board, 13-17 January 2013, Washington, DC, USA.

Schneider R, Patten R, Toole J. 2005. Case study analysis of pedestrian and bicycle data collection in U.S. communities. Transportation Research Record: Journal of the Transportation Research Board 1939:77-90.

Schimek, P. (2014). Circumstances of bicyclist injuries. Paper prepared for presentation at the 93rd Annual Meeting of the Transportation Research Board, January 2014.

Schepers, J., Kroeze, P., Sweers, W., and Wüst, J. (2011). Road factors and bicycle-motor vehicle crashes at unsignalized priority intersections. Accident Analysis and Prevention, 43(3), 853–61.

Summala, H., Pasanen, E., Räsänen, M., and Sievänen, J. (1996). Bicycle accidents and drivers' visual search at left and right turns. Accident Analysis and Prevention, 28(2).

Teschke, K., Harris, M., Reynolds, C., Winters, M., Babul, S., Chipman, M., Cripton, P., Cusimano, M., Brubacher, J., Hunte, G., Friedman, S., Monro, M., Shen, H., Vernich, L., and Cripton, P., (2012). Route Infrastructure and the Risk of Injuries to Bicyclists: A Case-Crossover Study. American Journal of Public Health, 102(12) 2336–2343.

Tilahun, N., D. Levinson, and k. Krizek. (2007). "Trails, Lanes, or Traffic: Valuing Bicycle Facilities with an Adaptive Stated Preference Survey." TR Part A: 41(4) 287–301.

United States Department of Transportation. 2011. Pedestrian and Bicycle Data Collection.

University of Washington Transportation Services (UWTS). UW Shuttle Facility Services Webpage. Last modified 4/23/2015. Accessed 5/10/2015.

Van Houten, R., and C. Seiderman. How Pavement Markings Influence Bicycle and Motor Vehicle Positioning: Case Study in Cambridge, Massachusetts. Transportation Research Record: Journal of the Transportation Research Board, Vol. 1939, 2005, pp. 1–14.

Vandenbulcke, G., Thomas, I., and Int Panis, L. (2013). Predicting cycling accident risk in Brussels: A spatial case-control approach. Accident Analysis and Prevention 62, 341-57.

Wang, Rui, and Chen Liu. "Bicycle-Transit Integration in the United States, 2001-2009." Journal of Public Transportation 16, no. 3 (2013).

Wang, Y., and Nihan, N. (2004). Estimating the Risk of Collisions between Bicycles and Motor Vehicles at Signalized Intersections. Accident Analysis and Prevention 36(3) 313–21.

Wachtel, B., and Lewiston, D. (1994). Risk Factors for Bicycle-Motor Vehicle Collisions at Intersections. Institute Transportation Engineers Journal, September, 30–35.

Wardman, Mark. "A comparison of revealed preference and stated preference models of travel behaviour." Journal of Transport Economics and Policy (1988): 71-91.

Weigand, L. (2008). A Review of Literature: Intersection Treatments to Improve Bicycle Access and Safety. Initiative for Bicycle and Pedestrian Innovation, PSU, Portland, Oregon.

Zhao, Jinhua, Adam Rahbee, and Nigel HM Wilson. "Estimating a Rail Passenger Trip Origin‐Destination Matrix Using Automatic Data Collection Systems." Computer‐Aided Civil and Infrastructure Engineering 22, no. 5 (2007): 376-387.