FUSING NONMOTORIZED TRAFFIC DATA: A DECISION FUSION FRAMEWORK

A Dissertation

by

SIRAJUM MUNIRA

Submitted to the Graduate and Professional School of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Yunlong Zhang
Committee Members,	Mark Burris
	Sue Chrysler
	Xia Ben Hu
	Ipek Nese Sener
Head of Department,	Robin Autenrieth

August 2021

Major Subject: Civil Engineering

ABSTRACT

This dissertation explored an uncharted research territory, fusion of nonmotorized traffic data for estimating reliable nonmotorized demand measures. The research was divided into three sequential stages. The first stage involved developing and applying a guideline to process and homogenize available nonmotorized data sources to estimate demand within a specified scope. Multiple data sources were utilized to develop five bike demand models to estimate annual average daily bike volume at intersections in Austin. Following an in-depth discussion of both nonmotorized traffic data and fusion characteristics, the second stage proposed a decision fusion framework divided into two broad categories: fusion without benchmark data and fusion with benchmark data. Under the first category, four fusion algorithms, including a novel approach, were investigated. Under the second category, a robust state-of-the-art statistical tool, the Dempster Shafer (DST) method, was endorsed. Dempster Shafer with credibility context, proposed by this study, offered a unique way to incorporate subjective judgment of the experts in the mathematical fusion formulation. The third stage was focused on applying the fusion framework on both actual and simulated data to demonstrate the efficacy of the fusion algorithms. The findings illustrated that the novel weighted voting fusion generated a fused estimate of comparable accuracy to the best source estimate when applied to four demand sources. However, when five source fusion was conducted, the accuracy decreased. When applied to simulated data of multiple scenarios, the DST method outperformed the individual source estimate in most cases. Based on the data, categorization and discounts, the change in accuracy varied

from -10% to 7%. Therefore, fusion exhibited the risk of obtaining worse-off results. Moreover, the proposed DST approach outperformed the traditional approach in most cases (above 80%), underscoring the merit of incorporating subjective judgment. The fusion considering knowledge and context is expected to contribute to the field of decision fusion. While the framework offered an additional option of analysis, it is up to the analyst or practitioner to consider and decide the course of option in adopting fusion endeavors given the trade-off between effort and the change in confidence, coverage, and accuracy of the outcomes.

DEDICATION

Mom, Dad, and my Brother- who always believed I could

Abir- who made me believe it

ACKNOWLEDGEMENTS

My Ph.D. journey has been made possible and enjoyable by the contribution of a handful of wonderful people. No word is sufficient to express my gratitude to Dr. Zhang for believing in me and showing me your support throughout my Ph.D. life. You always had the kindest thing to say to lift my spirit, especially when I was going through a personal and academic crisis. I would never forget how strong you stood by me in my Qual and my final Dissertation exam. Thank you also for all the recommendations, which brought many fellowships and awards. It was an absolute privilege to work under your guidance and mentorship throughout this dissertation.

I wish to express my most sincere gratitude and appreciation to Dr. Sener, who was instrumental in getting me started with the research and without whose help, guidance and patience, this dissertation wouldn't have been successful. It was a difficult journey, but no doubt I came out of it with a lot better writing and analytical skill. Thank you for always being kind and supportive and making my Ph.D. journey easier than most.

No amount of Thanks is enough for Dr. Chrysler. Her kindness towards me started when she hired me and then continued throughout this journey. Thank you for standing by me when I was having a mental breakdown and saying the kindest words.

I would like to thank Dr. Burris for his guidance and support in my Qual exam and comments in improving my entire dissertation. Also, thanks to Dr. Hu for his comments and suggestions to improve my dissertation work.

Also, special Thanks to Dr. Subasish Das for referring me to Dr. Chrysler and Safe-D. I would also like to wish my deepest thanks to Russell Henk for trusting me with his project work. I am also sincerely grateful to Dr. Bruce Wang for his guidance and help rendered to me.

A special word of appreciation for my online Statistics and Coding community, who has always been patient when answering my countless questions.

Thanks to my parents Md Haider Ali and Wahida Afsana, and my dearest brother Hassan Ali Askari for always supporting me. I am grateful to the Almighty for giving me the most loving family in this world. Without their love, affection and encouragement, none of my achievements would have been possible. My aim in life has always been to become a great daughter to my parents. I aspire to be the independent, kind and polite human being my parents brought me up to be and the best daughter they deserve.

I would also like to thank my other family- my in-laws for always supporting and encouraging me. Thank you to all of my friends for being the ear to my constant whining.

Finally, Abir, I count my blessings every day to be your partner and best friend. I get to be the whimsical spouse because you are the sensible one. There is nothing you didn’t do, from helping me in complex coding to making comfort foods to lift my spirit. I got into coding only because you pushed me there are then showed all the patience in the world while teaching me. Thank you for actively helping me every step of the way in my life. Every time I stumbled or got frustrated, and God knows there were too many of them, you just waved your magic wand and it (both my life and messy R code) was better again.

CONTRIBUTORS AND FUNDING SOURCES

Contributors

This work was supervised by a dissertation committee consisting of Professor Dr. Yunlong Zhang [Advisor] and Professor Dr. Mark Burris of the Department of Civil & Environmental Engineering, Associate Professor Dr. Xia Ben Hu of Department of Computer Science and Engineering and Dr. Sue Chrysler and Dr. Ipek Nese Sener [Project Supervisor] from the Texas A&M Transportation Institute.

The work conducted for the dissertation was completed by the student independently.

Funding Sources

The research was funded by the Safety through Disruption (Safe-D) National University Transportation Center, a grant from the U.S. Department of Transportation's University Transportation Centers Program (Federal Grant No. 69A3551747115).

NOMENCLATURE

AADT	Annual Average Daily Traffic
AADB	Annual Average Daily Bicycle Volume
ACS	American Community Survey
AIC	Akaike's Information Criterion
BPA	Basic Probability Assignment
CAMPO	Capital Area Metropolitan Council
CBSA	Core Based Statistical Areas
DDM	Direct Demand Model
DOT	Department Of Transportation
DST	Dempster Shafer
DT	Downtown
FHWA	Federal Highway Administration
GIS	Geographic Information System
MPO	Metropolitan Planning Organization
NHTS	National Household Travel Survey
NHTSA	National Highway Traffic Safety Administration
O-D	Origin Destination
TAZ	Traffic Analysis Zone
TTI	Texas A&M Transportation Institute
TxDot	Texas Department Of Transportation

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Page

1. INTRODUCTION AND MOTIVATION

1.1. Background & Motivation

Recognizing the wide range of social, environmental, economic, health, and road safety-related benefits of increased nonmotorized mode share, researchers, policy makers, and health advocates worldwide are working to create bicycling- and walking-friendly communities and reduce automobile dependency. Despite many efforts aimed at promoting bike and walk activity in the United States to create healthy, sustainable, and equitable communities, the quest to reach the envisioned nonmotorized mode share is still an uphill battle in many cities. Statistics based on the 2017 National Household Travel Survey (NHTS) showed that only around 12% of the total daily trips are made by walking and about 1% by biking (Buehler, 2019). Compared with other countries, especially European countries where nonmotorized mode share is as high as 40%, U.S. cities seeking to promote walking and bicycling activity still have a long way to go (Handy, 2019). Nonetheless, pedestrians and bicyclists account for a disproportionate share of the total fatal and serious injury crashes. In the United States, these two modes accounted for around 19% of the total US traffic fatalities in 2019 (The National Highway Traffic Safety Administration [NHTSA], 2020).

Acknowledging the exigency of the issue, safety advocates in multiple areas are persistent in their efforts to develop evidence-based, data-driven strategies to reduce nonmotorized crashes. Although the literature is replete with studies evaluating various aspects of nonmotorized safety, the majority suffer from a major limitation—the absence of nonmotorized demand or exposure data (Turner et al., 2017). The incorporation of robust and reliable exposure estimates is instrumental to the orchestration of an efficacious crash analysis for nonmotorized traffic (Fitzpatrick et al., 2018; Turner et al., 2017). Besides safety analysis, the lack of reliable and robust exposure or demand data also limits the planner's ability to fully comprehend as well as advocate the enormous potential of building/improving infrastructures within the community (Gosse and Clarens, 2014).

The existing approaches of estimating demand or exposure, be it through observed counts, models, or crowdsourced data, exhibit limitations in terms of spatial, temporal, or population representation, or often overall reliability. The observed volume most often falls short in terms of spatial and temporal coverage issues. On the other hand, selecting a modeling approach often warrants a tradeoff consideration in terms of complexity or resource requirement (time, budget, staff, data, etc.) and accuracy or reliability of estimation (Turner et al., 2017). Over the last few years, researchers and transport planners have steered their attention towards nonmotorized activity data from emerging technology-based methods such as GPS-enabled smartphone apps, wearable techs, interactive websites, bike share systems etc., owing to the ubiquitous advancements in technology. These crowdsourced or big data sources, characterized by their volume, velocity, variety, value and veracity (Zhang et al., 2017), offer great potential to understand the detailed spatiotemporal travel pattern of nonmotorized traffic on an unprecedented level of detail (Misra et al., 2014). However, they are often blamed for lacking quality assurance and representativeness (Goodchild & Li, 2012; Jackson et al., 2013) and are considered inadequate without validation from an actual location count (Jestico et al., 2016). Therefore, rather than recognizing these activity data as a reliable exposure measure, researchers have sidelined the sources as a potential proxy solution (Saha et al., 2018) to provide complementary information regarding nonmotorized activity in an area.

The fact that no individual data source or model is sufficiently adequate drives the research question of whether combining or integrating multiple data sources, or in simple terms creating a data fusion, can produce a better estimate of nonmotorized demand or exposure. Researchers in the motorized traffic domain, recognizing that no single data source can provide enough comprehensive and rich information to develop good transport models (Picornell and Willumsen, 2016), are rapidly progressing their utilization of the fusion method. However, despite the suggested efficacy of fusion for motorized traffic flow or demand (Antoniou et al., 2011; Guo et al., 2018), the fusion of nonmotorized traffic data is still uncharted territory. A fusion-based technique, which offers multiple benefits, including increased reliability, robustness, completeness, coverage, and often cost-effectiveness (Bachmann et al., 2013, Ince et al., 2012), when rightly adapted, may prove to be an effective tool for nonmotorized volume/exposure estimation.

1.2. Fusion, Nonmotorized Traffic Data and Context

The concept of data fusion is well-established and researchers across myriad disciplines have acknowledged its advantages to obtain comprehensive and rich information over individual data sources (Wu et al., 2015; Lu et al., 2015, Meng et al., 2017). The application of fusion is still an emergent field in transportation system research. It has been applied in Intelligent Transportation Systems (Dailey et al., 1996), travel time estimation (Martí, 2016), accident analysis (Sohn and Lee 2003), origin-destination estimation (Zhu et al., 2016), etc. The ultimate objective of fusion is to increase the quality and decrease the estimation error rate compared to the individual sources. The above discussion of existing approaches of nonmotorized demand estimation along with their limitations underscored the promising potential of the fusion approach, which, when rightly adapted, can combine strengths and decrease uncertainties associated with individual sources in order to provide a better understanding of the trend and pattern of nonmotorized activity

There is an abundance of research investigating numerous frameworks of data combination or fusion. However, the first and foremost step of selecting an appropriate fusion method is to discern the inherent characteristics and attributes associated with the data and information. Undoubtedly, the selection, customization, and application of fusion methods for combining nonmotorized data sources requires a profound understanding of the inherent characteristics of the data sources and situations. The process also necessitates comprehending the characterization and application of the existing fusion mechanisms. Even though fusion algorithms are being applied in multiple transportation areas, the discussion of aspects or taxonomies of fusion, especially from the transportation data perspective, is yet to be fully addressed. The characteristics of nonmotorized traffic data and commonly used fusion techniques would be carefully examined in the literature review (Chapter 2 and 3) in order to seek a generalizable fusion solution for nonmotorized demand computation. In the following sections, only a brief discussion on the features of nonmotorized data and fusion was outlined to provide a rationale for the research design.

The first issue pertinent to nonmotorized traffic data sources is they are seldom at the same resolution (spatial, temporal, or population) and are most often generated in different data formats. For example, bike-sharing data only represent the bike activity of the system users and are generated in an origin-destination (OD) matrix (station to station) format. Thus, to obtain intersection-level volume from this source, the raw data must go through trip assignment steps in addition to adjustments for population-level representation. This requirement warrants processing of individual sources to reach a homogenized representation before any fusion effort. The processing steps and complexity depend on the format of the raw data and the scope of analysis. Another issue of traditional nonmotorized activity data stems from the limited sample size. Data collection effort (both household survey and sensor-based location data) requires resources (budget and time). It is only possible to collect data for a limited number of locations, unlike motorized traffic data. These key issues, together with some other characteristics (discussed in Chapter 2), are likely to dictate the process of selecting a fusion framework for nonmotorized data. For example, the limited data availability is likely to inhibit the choice of some mechanisms such as artificial intelligence, machine learning, ensemble and deep learning, which are sensitive to sample size. These issues called for eliciting appropriate fusion methods that can adapt and accommodate the characteristics of the nonmotorized traffic data sources considering the constraints and add value to practical application.

The concept of context-aware fusion has been introduced within the perspective of sensor fusion, mainly in the computer science research domain, acknowledging the need for sensing and managing contextual information to improve fusion outcomes (Wu, 2003). A wide range of methods had been applied to handle context in fusion, where the contexts were interpreted from various standpoints. Dey and Abowd defined context as "any information that can be used to characterize the situation of an entity. An entity is a person, place, or object that is considered relevant to the interaction between a user and an application, including the user and application themselves" (Dey and Abowd, 2000).

Dey also noted that “while most people tacitly understand what context is, they find it hard to elucidate” (Dey, 2001). The process of discerning context for traditional context-aware research can be based on automatically acquired information or be done manually. In real-world applications, context detection or sensing mostly relies on manual input (rather than an automated process) and requires a profound understanding of the data dynamics and inherent situations (Dey and Abowd, 2000).

The unique characteristics of nonmotorized activities drive the research question of whether consideration of context, derived from subjective judgment, can further improve the performance of the fusion algorithm for nonmotorized traffic data. For example, walk and bike activity are extremely sensitive to time of day, day of week or month of year, in short temporal differences, and vary significantly with geographic locations. The Traffic Monitoring Guide by the Federal Highway Administration (2016) noted that the spatiotemporal variation of nonmotorized traffic is very location-specific which cannot be generalized. This issue unveiled the potential of incorporating knowledge about the situational context for nonmotorized volume information to enhance and improve the fusion approach. However, to the author's knowledge, no research within the decision fusion domain has incorporated situational knowledge or context (such as place or time), obtained from a subjective judgment, for refining the fusion algorithm.

1.3. Research Scope and Objectives

In light of this uncharted research territory, this dissertation endeavored to develop a fusion-based technique to combine multiple nonmotorized demand data. The objective was to explore, select, and customize fusion mechanisms that can accommodate the distinctive nature of the nonmotorized traffic activity data and/or model output, with an end goal of generating a better-quality demand estimate. To attain the objectives, the research was divided into three sequential stages.

In the first stage, the available data sources representing nonmotorized activity were investigated due to two main issues: i) The spatial and temporal scale of the data sources vary widely ii) the crowdsourced data do not represent the actual nonmotorized activity in an area. Considering these issues, the research acknowledged the need for preprocessing or modeling individual sources to represent the nonmotorized activity in a homogeneous and policy-relevant estimate in terms of spatial, temporal and population scale. As a case study, available and relevant data sources for bicycle traffic for the city of Austin were investigated and processed. Discussing the plausible scope of analysis and recognizing the practical implication, this research had chosen annual average daily volume of nonmotorized traffic at the intersection level as the unit of demand.

The second stage intended to develop a framework to fuse or combine estimates obtained from the first stage. Following an in-depth discussion of both nonmotorized traffic data and fusion characteristics, the dissertation proposed a decision fusion framework that accounted for data availability conditions. The fusion algorithms of the framework were divided into two broad categories: fusion without benchmark data and fusion with benchmark data. Under the first category, four fusion algorithms, including a novel approach, were investigated. The algorithms were unanimity voting, plurality voting, simple majority voting and weighted voting based on interaction. The last algorithm, a proposed approach, considered the pairwise interaction of the data sources to assign weightage on each decision before fusion -which boils down to the framework of weighted majority voting but without the ground truth information

In an endeavor to develop a robust state-of-the-art statistical tool, the Dempster Shafer method was selected as a fusion algorithm when benchmark data is available. The method is a well-known generalization of the Bayesian framework, which can handle ignorance, conflict and inconsistency and was deemed suited for nonmotorized traffic data fusion. In addition to the basic Dempster Shafer method, the dissertation sought to explore if any corrections of extensions can be incorporated to improve the accuracy of the fusion task at hand. In the process of creating a comprehensive understanding of the unique characteristics of the nonmotorized data, the researchers acknowledged that the nonmotorized traffic data are more likely to be highly conflicted and the accuracy of each data source or model may vary with locations. Contemplating the later characteristic as context, the dissertation proposed an approach, Dempster-Shafer fusion with context credibility, which can handle the spatial variability of individual estimates (utilizing grey relation theory) to further improve the fusion outcome. The Dempster-Shafer method was selected for this purpose considering to its ability to incorporate human subjectivity with mathematical probability for fusion.

The third stage illustrated the application of the fusion framework for actual and simulated data. The application of the framework application on the simulated data was deemed specifically relevant as it intended to formulate the answer to the question of under what scenario the fusion methods are appropriate for the application. The numerical outcomes, comparing the performance of different algorithms, were expected to furnish a more intuitive explanation regarding how the fusion algorithm performs under different scenarios or circumstances.

1.4. Dissertation Organization

The dissertation was organized into seven chapters under four broad parts: Introduction, Review of Literature, From Information to Knowledge and Application & Conclusion. The chapters under these parts are

- Introduction,
    - Chapter 1 outlined the motivation and objectives of this dissertation
- Review of Literature,
    - Chapter 2 discussed elements, characteristics, and application of the nonmotorized traffic data along with the characterization of demand estimation process.

    - Chapter 3 explained the aspects and characterization of fusion, leading to the selection of the fusion framework for this dissertation
- From Information to Knowledge
    - Chapter 4 illustrated the design of a guideline for homogenizing multiple data sources under various scopes of estimation and then discussed the mechanism of processing or modeling multiple bike activity-related data to estimate demand in the study area
    - Chapter 5 presented the mathematical formulation of the algorithms utilized in this dissertation
- Application & Conclusion
    - Section 6 demonstrated the application of the fusion framework for both actual and simulated data under different scenarios
    - Chapter 7 discussed the contribution of the research, highlighting the role of the fused estimates in the safety analysis. The limitations and future scopes of the study were also discussed

2. LITERATURE REVIEW: NONMOTORIZED TRAFFIC DATA

The literature review was divided into two main parts. The first part (chapter 2) discussed the elements, characteristics, and application of the nonmotorized traffic data. In order to keep the review in focus, it concentrated on nonmotorized traffic demand/exposure analysis in the later parts (chapter 3).

More specifically, this chapter reviewed literature focusing on four aspects of nonmotorized transportation. The first section discussed the absence of and need for reliable exposure or demand data highlighting the most frequently used exposure measures and their limitations. The second section explored the features and application of different nonmotorized data sources under two broad categorizations. The third section presented the characterization of nonmotorized data sources from a demand analysis perspective. The section highlighted how nonmotorized traffic activity and model and monitoring processes are unique, elucidating the scopes of demand estimation.

2.1. Nonmotorized Demand or Exposure: A Missing Piece of Puzzle

Theoretically, exposure can be defined as potential opportunities for crash occurrence (Turner et al., 2017). According to Greene-Roesel et al. (2007) exposure can be defined as “contact or amount of contact with potentially harmful situation.”. However, in practice, a wide array of exposure measures has been used for crash analysis such as Pedestrian or bicyclist volume, sum of entering nonmotorized flows at intersection, total travel distance, in person-miles of travel etc. (Turner et al., 2017). This research focused on nonmotorized demand at intersection as exposure measure.

Data on bicycle and pedestrian demand or exposure are of utmost value to promote bicycling- and walking-friendly policies, evaluate the related efforts and discern any crash patterns. Without such data, it is difficult to provide a reliable answer to the critical questions of how many people actually use current pedestrian or bicycle facilities and how many will potentially use a new or improved facility if one is built. Safety professionals need location-based volume data to accurately discern the trend in crash rates, identify high-risk locations, and understand the crash causation. The lack of reliable and robust exposure data limits the planners' ability to fully comprehend as well as advocate the enormous potential of building/improving infrastructures and measure pedestrian and bicyclist risks (Gosse and Clarens, 2014). However, information related to demand or exposure is most often missing or available in limited settings.

An ideal exposure measure for crash analysis would be pedestrian or bike activity count at all facility locations across the study area, which is not feasible given the resource constraints. Only a few studies have utilized bicycle and pedestrian count data, obtained from both signalized and unsignalized intersections, as crash exposure measures (Heydari et al., 2017; Strauss et al., 2014). Besides, Wang and Kockelman (2013) used walk-miles traveled as an exposure measure, estimated using household travel survey data and least squares regression. Vehicle miles traveled was another exposure measure used by studies for nonmotorized crash analysis (Nashad et al., 2016; Cai et al., 2017). Since it is often difficult to quantify the number of pedestrian/bicyclist miles of travel and motorized vehicle miles of travel at a zonal level, researchers focusing on the nonmotorized crash have often relied on proxy variables or surrogate measures such as population density (LaScala et al., 2000; Narayanamoorthy et al., 2013), income (Loukaitou-Sideris et al., 2007), activity intensity characteristics (Mitra and Washington, 2012), road network length (Kamel et al., 2019), and so forth. Recent studies have also noted that Strava data (a fitness tracking app monitoring walking and biking activity of individuals) can be a reasonable proxy of the total bicycle volume in certain circumstances (Sanders et al., 2017; Jestico et al., 2016). Most of these studies have emphasized the inadequacy of proxy variables to represent nonmotorized activity, highlighting the need for reliable exposure measures.

A fairly recent study conducted by FHWA (Turner et al., 2018) had discussed some modeling-based methodologies to evaluate or forecast demand at both facility level (intersection or street segment) or zonal (TAZ, Block group, etc.) level. Examples of such models are direct (facility) demand models, regional travel demand models (e.g., trip-based models and activity-based models), trip generation and flow models, geographical information system (GIS)–based models, etc. (Turner et al., 2018). Among these models, the direct (facility) demand model is the most widely utilized modeling approach in the area of pedestrian/bicyclist safety, especially when resources are limited (e.g., input data requirements, technical complexity, budget considerations due to costly model development and maintenance). The particular modeling approach has been used to estimate exposure for all intersections in the study area for both pedestrian and bicycle crash analysis (Munira et al., 2020; Hasani et al., 2019). However, as noted by the researchers, although the approach makes great use of available count data at limited locations and of their surrounding features to estimate volume at locations without counts, it is limited in terms of capturing the underlying behaviors and travel patterns (Munira et al., 2020)

Although not particularly used for developing crash exposure measures, nonmotorized activity models have been developed within the settings of a four-step modeling framework (Clifton et al., 2008; Kuzmyak et al., 2014). However, as indicated by the studies, given the coarse level of spatial analysis structure in the first three steps, the trip-based models exhibit limitations in capturing nonmotorized volume at the fine-grained unit (such as intersection) (Schneider et al. 2009; Griswold et al., 2011). Turner et al. (2017) have also noted that the selection of a modeling approach often demands a tradeoff consideration in terms of complexity or resource requirement (time, budget, staff, data, etc.) and accuracy or reliability of estimation.

2.2. Nonmotorized Data Sources

Methods for estimating nonmotorized demand call for a comprehensive discussion of the data sources, used directly in computing volume or facilitating demand models, including their characteristics, strength and limitation. For ease of discussion, sources were divided into two broad categories: traditional and emerging data sources. While the applications of the data sources were briefly outlined here, more detail on the processing methods for demand analysis was reported in section 2.3.2 under a broader categorization.

2.2.1. Traditional data sources

The discussion on traditional data for nonmotorized traffic focused on on-site data collection efforts, such as short and continuous count, and household surveys, including National Household Travel Survey and American Community Survey.

The most traditional approach for collecting bicycle and pedestrian data is counting traffic at selected locations (intersection, mid-block and trails) in a transportation network. Traditional on-site data collection efforts can be categorized into two types: short count (such as peak/off-peak, 12 hours, 48 hours, 7 days, and so forth) and continuous count. Often, transportation agencies quantify nonmotorized activity through manual short counts performed by staff and volunteers at sites. In recent times, pedestrian and bicyclist data are mostly collected using automated counting technologies such as pneumatic tubes, inductive loops, thermal cameras, infrared sensors (active/passive), magnetometers, piezoelectric devices, radar sensors, and video imaging

(Levinson et al., 2016; Nordback et al., 2016; Ryus et al., 2014). A recent study (Hasani et al., 2019) adopted advanced video processing algorithms to automatically count pedestrians and bicyclists in the San Diego area.

Generally, the short count data collection method is designed to collect data during typical weekdays for a minimum of 24 hours period of time in the selected locations. The duration of short-term observation may vary (such as peak/off peak, 12 hours, 48 hours, 7 days) based on the study objective and their available resources. For example, Tabeshian and Kattan (2014) gathered bicycle and pedestrian counts for three different durations (7 a.m. to 9 a.m., 11 a.m. to 1 p.m., and 4 p.m. to 6 p.m.) in different months. However, the short-period count data are affected by duration, time of week, time of day, month, and season (Hankey et al., 2014; Nordback et al., 2013). Even in the same region, nonmotorized activity may vary dramatically by location, time, and weather conditions. Thus, most often, nonmotorized activity in a location is represented using an annual average estimate which accounts for daily, weekly, and seasonal variation. The temporal variation can be measured using data from the continuous counter installed in a limited number of locations. For example, Munira and Sener (2020) utilized continuous count data at 11 locations for computing adjustment factors which were used to scale 24-hour short count data to estimate annual average daily bicycle (AADB) traffic. Hankey et al. (2017) collected continuous count data for 1 year at four sites and 1-week count data at 97 sites to estimate annual average daily volume.

The Federal Highway Administration’s (2016) *Traffic Monitoring Guide* guides systematic monitoring of nonmotorized traffic using a combination of permanent and short-duration count sites.

The key limitation of the onsite count data is limited temporal and spatial coverages. Given nonmotorized traffic exhibit extensive spatial and temporal variation (discussed in section 2.3.1), count from one location at a given time is not representative of the count at other, even nearby, location. However, it is not cost-effective to install automatic sensors at every location in a city to monitor continuous pedestrian and bicycle volume data. Apart from the limited spatial and temporal coverages, the limitation of the onsite count data also stems from the traffic monitoring technologies devices. Benz et al. (2013) have highlighted the technological challenges associated with the nonmotorized traffic monitoring process. The fact that pedestrians and bicyclists are less confined to fixed lanes or paths affects the accuracy of the sensor. Moreover, when they (pedestrians and bicyclists) travel in closely-spaced groups, sensors are often unable to differentiate the activity type.

Household surveys are another way to obtain nonmotorized activity-related data. The national-level survey such as the National Household Transportation Survey (NHTS) and the American Community Survey (ACS) provide insights into nonmotorized travel activity at the individual and zonal level. While ACS focuses on work-related transportation activity (i.e., walking and biking), NHTS gathers activity data for all types of trips.

The NHTS survey data provides a large national sample to illustrate individual walking and biking travel behavior and variations by demographic and regional factors, including age, gender, race, income, geographical region of the respondents (Santos et al., 2011). It is an excellent source for understanding how, why, and what modes of individual travel activity in the United States. Moreover, the NHTS Add-on Program provides geocoded O-D data for all trips and modes, including transit, walk, and bike. However, the survey involving travel diaries are expensive to administer and often suffers from the issue of relatively small sample sizes (National Household Travel Survey, 2018)

On the other hand, the ACS data provides annual, publicly available information on walk and bike work trips along with social, economic, housing, and demographic characteristics of America's communities at an aggregated geographic level. The survey data is a great source to evaluate changes in work-related walking and biking at various geographic units (Whitfield et al., 2020).

Both NHTS and ACS, representative of national estimates, involve systematic data collection efforts geared to be used in travel demand models for performance measures, project prioritization etc. These data allow demand analysts to understand human travel behavior along with their motivations, choice, and attitudes, where big data sources fall short (Clifton et al., 2004, Whitfield et al., 2020). Data from NHTS can be used in computing various parameters of the demand model such as trip-making characteristics, mode choices, vehicle ownership, trip lengths by purpose etc. The survey data can also be utilized to calibrate and validate travel demand models (National Household Travel Survey, 2018). On the other hand, the census data can be used to support multiple components of travel demand models, such as trip generation, trip distribution, mode choice, traffic assignment, demographic and auto ownership models, etc. (National Academies of Sciences, Engineering, and Medicine, 2007).

However, both of these data collection efforts require significant resources (time, budget, staff). Moreover, although both NHTS and ACS data are useful in understanding active travel trends and facilitating relevant models, the use of these data for a comprehensive evaluation of walk and bike activity is limited, meaning the data cannot be directly and individually used to estimate nonmotorized travel demand at a specific location.

2.2.2. **Emerging data sources**

The limitations of traditional data collection methodologies have steered the attention of the researchers and transport planners towards emerging technology-based methods (or big data) such as GPS-enabled smartphone apps, wearable techs, interactive websites, bike share systems etc. These emerging data sources bring great promises to illuminate human activities and transportation behaviors that were previously poorly understood (Dill, 2009). To take advantage of the massive crowdsourced data generated daily from these sources, researchers endeavored to collect, manage and analyze data to understand patterns and monitor changes in nonmotorized activity. The use of high-resolution crowdsourced data in transportation is expected to create a data-rich environment for nonmotorized researches in the future. The following section briefly discussed the four key sources: fitness apps data focusing on Strava data, bicycle sharing system data, GPS data from cellphones or other sophisticated devices data and Web 2.0 application.

Fitness phone apps, most often connected with wearable devices, record and sync physical activity-related data from which various aspects of the nonmotorized activity, both at individuals or regional level, can be derived. Given the prolific use of the fitness apps, researchers are acknowledging the promising potential of the app gleaned data to analyze walking and bicycle activity. In a study, Nielsen (2014) reported that around one-third of smartphone users use fitness apps, aided in part by wearable technologies such as smart watches and fitness bands. Among many crowdsourced data sources, Strava has been creating a huge dataset, gleaning user data from both pedestrians and bicyclists. Strava is one of the largest cycling fitness apps, having reported distance data of 8 billion miles from 48 million users across 195 countries in 2019 (Strava, 2019). Given the large number of users, a handful of researchers have investigated the use of Strava data sources to examine various issues of nonmotorized travel, including the impact of the built environment and socio-demographic features on cycling behavior (Munira & Sener, 2020; Griffin & Jiao, 2015; Hochmair et al., 2019; Sun et al., 2017), accident risk (Saha et al., 2018; Sanders et al., 2017). A recent study utilized Strava data to estimate AADB in Texas (Dadashova et al., 2020).

Bike share refers to a service that provides short-term access to bicycles for public use, often operating as part of the public transport system of a region (Macioszek et al., 2020). Shared bikes, including e-bikes, and scooters, have become an indispensable part of the transport system in many US cities. In 2019 alone, 136 million trips were made (National Association of City Transportation Officials, 2020). Over the past couple of years, several studies have investigated the bicycle-sharing flow and usage data and explored function, use, issues and capacity (Faghih-Imani et al., 2017; DeMaio, 2009; Frade & Ribeiro, 2015; Lin & Yang, 2011); optimization (Raviv et al., 2013; Raviv, & Kolka, 2013); system prediction (Li & Zheng, 2019); association with demographic and built environment characteristics (Faghih-Imani et al., 2014; Rixey, 2013; Krykewycz et al., 2010; Wang et al., 2012) etc. A recent study (Pogodzinska et al., 2020) estimated the daily volume of bicycles using data from automatic counter loops bike-sharing systems in a city in Poland. The study indicated that the share of bike-sharing users varies significantly among locations. The study concluded that in order to estimate bicycle volume using these sources, it is essential to perform control measurements to verify models and their potential applications in other areas.

Another source of nonmotorized activity data, especially bike, is GPS tracking smartphone apps which often use recruited volunteers for data collection. Such data has been used in an array of nonmotorized researches, including to explore route choice and length (Jackson et al., 2014; Shen et al., 2014; Hood et al., 2011), speed and travel time analysis (Strauss and Miranda-Moreno, 2017), riding experience (Blanc and Figliozzi, 2016), inputs of estimating AADB (Strauss et al., 2015), etc.

Owing to the rise of Location-Based Services (LBS) that logs data from variety sources including GPS, WiFi proximity, aGPS, Bluetooth proximity, and cellular triangulation for smart phones (StreeLight, 2017), different mobility analytic company, such as StreetLight, AirSage, started collecting traffic and travel-related data. Recently, StreetLight started to incorporate biking and walking activity data to get an understanding of the nonmotorized activity on streets, bikeways and other routes (StreetLight, 2019). The activity measure, known as StreetLight Index, represents a sample of bike or walk trips starting in, passing through, or ending in defined zones (Turner, 2020). Although the LBS data is expected to have more coverage than traditional crowdsourced data, including in rural areas (StreeLight, 2017), the strength and applicability of this dataset are yet to be fully discovered. A recent study compared bicyclist counts from 32 locations in eight cities in Texas with StreetLight data and reported promising correlations (Turner et al., 2020)

Web 2.0-based technology is another form of crowdsourced data that allows community-based input and collaboration to make them become a part of the pedestrian and bicycle planning decisions. Three key types of Web 2.0 applications, identified by Nash (2009), are: i) wikis, such as StreetsWiki and the ITE Pedestrian and Bicycle Committee wiki (only for ITE members), ii) personal information sharing such as blogs, YouTube or Twitter, and iii) mash-ups that allows combining information from several sources, such as Walkscore. The data gleaned from Web 2.0 application have been used in multiple areas including to improve area-wide bikeability scoring (Krykewycz et al., 2011); explore people's opinion on nonmotorized project alternatives (Misra et al., 2014), investigate reported collisions, near misses, hazards and thefts related to walk and bike activity (Nelson et al., 2015), increase community awareness regarding biking and walking issue (Town of Blacksburg, 2015) etc.

While the above discussion highlights the strength and use of the big or crowdsourced datasets, it is important to note that the data are not generated specifically for travel demand models or travel behavior analysis. These data have often been blamed for lacking quality assurance and representativeness (Goodchild & Li, 2012; Jackson et al., 2013) and are considered inadequate without validation from an actual location count (Jestico et al., 2016; Romanillos et al., 2016). For example, while the GPS-enabled fitness apps such as Strava mainly represent users who are oriented towards workout and fitness while bike share data provides insights into the overall shared bicycle demand within a city. Despite the extensive coverage, Strava represents a small percentage of nonmotorized users. For example, 3%–9% of bicyclists on trails in Austin were found to use Strava (Griffin and Jiao, 2015). The proportion of Strava users vary by location, land use, nonmotorized facility type, etc. (Jestico et al., 2016; Munira & Sener, 2020).

2.3. The Journey from Data to Demand Estimation

The previous section outlined the existing data sources representing nonmotorized traffic activity, highlighting that the data comes in different forms and for different purposes. It is clear that not all data can be used to compute facility-level demand or exposure. Moreover, some of them have to go through more processing steps than the others. Therefore, this section unfolded the journey from the data to the demand estimation, clarifying the unique characteristics of nonmotorized traffic data.

2.3.1. What Makes Nonmotorized Data Unique?

The literature review regarding the application of fusion in the transportation domain (presented in section 3.7) would reveal numerous existing research utilizing multiple fusion approaches to combine heterogeneous traffic-related data. Hence, before embarking upon the research objectives, the first question was, why can’t we apply the motorized data fusion approaches for nonmotorized traffic? The answer would be that motorized traffic data is different from nonmotorized traffic data, which exhibits unique characteristics. For instance, traffic flow theory, speed and travel time distribution play an important role in motorized traffic demand modeling tasks, which is not the case for nonmotorized traffic. Both walking and biking trips tend to be of short length and most often influenced by the immediate surrounding environment. Moreover, the route choice model or traffic assignment for bicycles is much more complex when compared with motorized vehicles as the bicycle route choice decision is associated with many factors (Ryu et al., 2017; Raith, 2009). The four key characteristics associated with nonmotorized traffic activity and data were discussed below. The first and third characteristics are generally associated with traditional and crowdsourced data, respectively.

2.3.1.1. Limited sample size

One of the main differences between nonmotorized and motorized traffic activity monitoring is the scale of data collection (FHWA, 2016). Nonmotorized activity data, particularly the traditional ones, suffer from a limited sample size as agencies most often collect data from a small number of locations due to resource constraints. The temporal resolution of motorized traffic data also tends to be finer compared to nonmotorized traffic. For example, motorized traffic flow or demand analysis models often deal with continuous count data for a large number of locations, updated every few seconds to minutes (Bae, 2017). Whereas, number of locations for which actual on-site data are available for nonmotorized demand analysis is generally between 30 to 100 (Hasani et al. (2019); Munira and Sener (2021); Chen et al. (2017) etc. Continuous count data are available in a more limited setting. A few fairly recent studies in the US have gathered nonmotorized data on a large scale for a region; for example, Griswold et al. (2019) gathered data for 1270 intersections in California and Hankey and Lindsey (2016) utilized data for 471 street segments in Minneapolis.

In addition to on-site data, the sample size of nonmotorized transport users in the household surveys is also limited. For example, while building nonmotorized accessibility model for Arlington (Washington region) using regional travel survey data, researchers acknowledged the limited sample size for bike mode (69 bike trips) (Kuzmyak et al., 2014). The NHTS mode share also confirmed the sample size issue as according to the 2017 National Household Travel Survey, only around 12% of daily trips were made by walking, and about 1% were made by biking (Buehler, 2019). While conducting location-wise analysis, the sample size for biking or walking maybe even smaller than the national statistics. The small sample size inhibits the choice of modeling approaches that demand a large amount of training data. It also puts constraints on the modeling process in multiple ways and negatively affects modeling results and conclusions (Munira and Sener, 2021).

2.3.1.2. **Geographic and Temporal Variation**

Nonmotorized activity is associated with spatial and temporal variability. Unlike motorized traffic, nonmotorized traffic demand depends greatly on rain, temperature, sunshine, wind speed, time of day, day of week and month of the year (Aultman-Hall et al., 2009; Tin et al., 2012). Comprehensive understanding of the variability is often constrained by the limited availability of continuous count data (FHWA, 2016). A study in Seattle (WSDOT, 2014), provided an example of the temporal variation of bike traffic showing how bike traffic varies notably with the time of the day, day of week and season. Other studies have also asserted that the temporal pattern of nonmotorized traffic is somewhat different from car and truck traffic patterns (FHWA, 2016).

Besides, the temporal pattern may not be even similar when bike and pedestrian activity are compared. In Minnesota, a study (Lindsey et al., 2013) comparing pedestrian and bicycle traffic patterns for 43 locations reported the notable variation to suggest that bike and walk traffic data should be separately gathered and modeled.

Similar to temporal variation, the geographic variation of walk and bike activity is also noteworthy. Bike and walking activity vary significantly by location, street and path type (Aultman-Hall et al., 2009). Two adjacent locations in a region may have extremely different nonmotorized activity patterns. One example can be seen from a study by McBain & Caulfield (2018), who illustrated the frequency of trip origin for each bike-sharing station in Dublin showing that even at the closely located stations, the frequency of trips may vary dramatically.

Drawing attention to the difference between motorized and nonmotorized traffic activity, in terms of spatio-temporal pattern, was necessary to underline the fact that the traditional method of spatio-temporal interpolation, for estimating motorized traffic at locations with little or no traffic counts, cannot be directly applied to the nonmotorized context. For example, spatial interpolation for motorized traffic relies on the assumption that annual average daily traffic (AADT) at a location is a function of AADT at the surrounding sites (Bartier and Keller, 1996; Wang and Kockelman, 2009). Typically, the spatio-temporal interpolation method uses two types of neighboring data: temporal-neighboring (data gathered from the same source and in the same time period but neighboring days) and pattern-neighboring (data gathered from the same source and in the same time period but in other days with similar daily flow variation patterns) (Shang et al., 2018). The typical neighboring mechanisms include average or weighted average method, historical average model, the K-Nearest Neighbor (KNN) model etc. (Shang et al., 2018). However, these methods are not practiced for nonmotorized traffic activity prediction, probably due to their geographic and temporal variability across the network in a region. One study (El Esawey, 2018) had addressed the unreliability of long-term counters, subjected to periodic malfunction, and attempted to impute missing data utilizing two approaches: using historical averages and incorporating weather and time-specific variables. The study suggested the inappropriateness of using historical average data to compensate for missing counts and concluded the superiority of location-specific regression count models incorporating weather conditions in estimating bike traffic counts.

2.3.1.3. **Spatial variation of Deviation**

This characteristic is associated with crowdsourced data, which only represent a subpopulation of the total nonmotorized activity. In addition to representing only a part of total activity, crowdsourced data may also exhibit sampling bias by oversampling a certain population group. For example, fitness app-based data, such as Strava, tend to oversample male and fitness riders and undersample female, low incomes, older riders and novice bicyclists (Boss et al. 2018; Jestico et al. 2016; Blanc et al. 2016). In a study in Australia, Lieske et al. (2017) compared app-generated data with Bicycle Journey to Work (JTW) census data. They found that the app-generated data were more representative in urban areas than in rural areas, resulting in a strong urban bias.

Moreover, Munira and Sener (2020) showed that even divergence of Strava volume from the actual volume varied with space. The study compared variations between Strava and actual volume data for 43 intersections in Austin. The comparison was performed using three key statistics—average percent deviation (APD), average of the absolute percent difference (AAPD), and Pearson's correlation coefficient—following guidelines from the Transportation Research Board's Guidebook on Pedestrian and Bicycle Volume Data Collection (Ryus et al. 2014). The APD represents the overall divergence from the actual volume data. The AAPD is a measure of the source's consistency. Pearson's correlation coefficient is a measure of linear correlation between the actual bike volume and Strava cyclist volumes.

The research showed a fairly high APD of −89%. The AAPD was 89%, indicating that the Strava volume was less than the actual volume in all locations. The research also showed that the percentage deviation varied from −43% to −97%, across geographic locations. It was suggested that the differences may be due to the demographic characteristics and trip purpose distribution in different areas.

The research also reported a strong linear association between the volume data from the two sources (Pearson's correlation coefficient $r = 0.63$, $p < 0.0001$). When compared with other studies (Hochmair et al., 2019; Turner et al., 2019), the association was found to be stronger. For instance, comparing Strava and actual volume data in Austin, Turner et al. (2019) reported a Pearson's of $r = 0.59$ and an AAPD of 92.49%.

2.3.1.4. **Traffic Assignment for Nonmotorized Traffic is More Complex than Motorized traffic**

In order to estimate nonmotorized traffic demand at a facility level (street segment or intersection) from an origin-destination matrix (for example, bike sharing data, trip table data from metropolitan planning organizations), it is necessary to perform traffic assignment task. Traffic assignment and route choice for bicyclists and pedestrians are significantly different from that of motorized vehicles. Private vehicles are affected by network congestion which means the travel time is not proportional to road length. But bicyclists are generally not affected by congestion, and they also do not contribute to it. The route choice model or traffic assignment for bicycles is much more complex when compared with the motorized vehicles as the bicycle route choice decision is associated with many factors (Ryu et al., 2017a, 2017b; Raith, 2009) such as travel time or journey length (Stinson and Bhat, 2003; Broach et al., 2011), safety (Akar and Clifton, 2009; Dill and Carr, 2003; Lowry et al., 2012), stress (Mekuria et al. 2012; Sener et al., 2009) etc. Studies have also acknowledged that bicyclist are sensitive to elevation gain, accident-prone route, turn frequency, junction control, noise, pollution, scenery, route with green space and traffic volumes (Ryu et al., 2017b; Broach et al., 2012; Winters et al., 2011)

Often the bicycle forecasting models assign traffic by assuming that cyclists select the minimum-distance path without considering the network attributes. However, nonmotorized traffic is likely to be combinedly affected by multiple attributes. For example, a shared bike path, although the shortest path, with high volume motorized traffic are likely to discourage bicyclists (Klobucar and Fricker, 2007). Hence studies have used multi-criteria approach for traffic assignment (Su et al., 2010; Turverey et al., 2010; Corne et al., 2003; Ryu et al., 2017). For more details about the nonmotorized traffic assignment process and models, readers are referred to Ryu et al., 2017a, 2017b.

2.3.2. **Nonmotorized Data sources from Demand Analysis Perspective**

In light of the discussion in the previous sections, it is clear that several traditional and emerging data sources are available to depict multiple aspects of nonmotorized traffic activity. However, their structure, objectives, scope, or use in the analysis are rarely the same. In order to facilitate understanding of data's value and features, especially within the research context, Table 2.1 tabulated the key characteristics of the commonly available nonmotorized data sources.

The breakdown of the characterization of the individual source will be vital in interpreting and further processing the data to extract volume-related information.

TABLE 2.1. Characterization of Nonmotorized Data Sources

Data Sources	**Temporal Coverage or Frequency of data collection**	**Spatial Coverage**	**Population Resolution**	**Data Output**
Short Duration count (i.e., Manual, Video image etc.)	15 minutes to 24 hours for multiple days or week	limited number of locations (intersection or mid-block)	Total Population	Aggregated volume in a road segment or intersection for a specific time period
Permanent Count (i.e., Inductance Loop, Magnetometer etc.)	Constant count for at least one year	A few locations (intersection or mid-block)	Total Population	Aggregated volume in a road segment or intersection

Data Sources	Temporal Coverage or Frequency of data collection	Spatial Coverage	Population Resolution	Data Output
American Community Survey	1-year, 3-year, and 5-year estimates	Area level estimate for all geographies down to the block group level	Total population	Represent commute, social, economic, demographic, and housing characteristics of the U.S. population
National Household Travel Survey (NHTS)	Every 5 to 7 years	Down to CBSA level in the USA	Total population	Individual travel data for all trips, modes, purposes, trip lengths, etc.
National Household Travel Survey (NHTS) Add on program	Every 5 to 7 years for Add on partners	Smaller and more precise level of geography compared to NHTS data	Total population	Additional data includes origin and destination locations of each trip
Fitness App (i.e., Strava)	Continuous data collection	Entire USA (based on users)	Only Strava users	Origin-destination and node/street level volume
Bike Sharing Data	Continuous data collection	Depends on the coverage of the stations	Only system users	Origin-destination of the trips

Data Sources	Temporal Coverage or Frequency of data collection	Spatial Coverage	Population Resolution	Data Output
StreetLight Data	Continuous data collection	Entire USA (based on users)	A sample of the total trips	Node/street level volume

Table 2.1 depicts that none of the nonmotorized data sources can furnish with activity data for full spatial, temporal or population level resolution. For example, permanent sensor data can provide data for all hours, seasons across the year but only for the location, they have been installed. Data from wearable tech only provides data for the people using the apps.

Therefore, to compute activity-related information, each dataset has to be adjusted, scaled or modeled. The application of the data sources, either as direct input of demand models or as parameters or aids of the model building process, also varies widely. The modeling process to compute demand is divided into three types for ease of discussion. Table 2.2 presented the plausible application of different data sources from a demand anlaysis perspective

TABLE 2.2: Data Processing for Demand Estimation

Data Source Type	Example	Processing Outcome
On-Site location data	Short and Continuous Count data	• Estimating flow patterns and annual average or peak/offpeak hour volume • Computing adjustment factor for scaling short-duration count

		• Developing direct demand model
Survey Data	NHTS, ACS, NHTS-Addon	• Estimating parameters for demand models • Developing mode choice model
Crowdsourced data (at node/zone level)	Strava, Streetlight, Bike-sharing data	• Estimating system user volume • Estimating total volume

2.3.2.1. **Processing On-site Location Data**

The onsite location data refers to the short and continuous counting methods. Processing methods of these data to estimate nonmotorized demand can be categorized into two types:

i. Spatial extrapolation
ii. Temporal extrapolation

The temporal adjustment generally refers to adjustment due to the variation of time of day, day of the week, season etc. Temporal adjustment factors can be calculated for a particular location or computed from knowledge obtained from a different location. For example, due to the unavailability adjustment factors for a study area, Bellingham, WA, Lowry et al. (2015) consulted four sources which were i) National Bicycle and Pedestrian Documentation Project (NBPD) provided with adjustment factors for three climatic zones and two facility types, 2) A previous research conducted in Colorado (Nordback et al.,2013) provided adjustment factors for recreational trails, suburban streets, and urban streets, iii) US DOT's Traffic Monitoring Guide (Federal Highway

Administration, 2013) for Minneapolis, Minnesota and iv) a study by Miranda-Moreno et al., (2013) that included adjustment factors for five North American cities for three facility types.

Lindsey et al. (2013) investigated data from six off-street trails to estimate daily and monthly adjustment factors in Minnesota for estimating the average annual count. The study also highlighted that the findings are not transferable to other geographic locations with different weather conditions. El Esawey (2014) analyzed one-year data in Vancouver to investigate the estimation accuracy of using daily and monthly adjustment factors to estimate AADB and found that the accuracy of AADB is more accurate when monthly adjustment factor is used compared to daily adjustment factors. Ryan et al. (2014) developed a simple method to estimate average daily bicycle volume in the San Diego using average PM peak percentage. The study suggested that PM peak percentage along with the PM peak period manual counts can be used to extrapolate daily bicycle volumes within a network.

On the other hand, nonmotorized traffic volume data, collected at a limited number of locations, can be utilized to estimate network-wide volumes by modeling approach. The direct demand model is an example of such modeling approach that relates bicycle travel demand directly to mode, trip, and traveler attributes using different forms of regression analysis (Munira and Sener, 2017). The resulting models can be used to predict travel activity at similar locations without counts. The models are widely used due to their simplicity in understanding and application. The modeling approach utilizes a wide array of explanatory variables based on locations. A literature review of the determinants of bike and walking activity is beyond the scope of this study. For more details, users are referred to Munira and Sener (2017)

Space Syntax is another example of analyzing nonmotorized movement and predicting volume (Raford and Ragland, 2004). The method generates "movement potentials," for nonmotorized traffic by evaluating the layout and connectivity of urban street grids. The movement potential is compared with sampled volume and land-use indicators at locations. The resulting model can be extrapolated to estimate nonmotorized volumes on street levels for an entire area (Raford and Ragland, 2004).

2.3.2.2. **Processing Survey Data**

The travel-related surveys, including NHTS and ACS are gathered from a representative sample, particularly for travel forecasting models. Hence the data from these surveys serve numerous purposes in demand analysis. The survey data plays a crucial role in formulating various assumptions for travel demand models, including the percentage of mode share of bike and walk trips, temporal variation, trip distribution by purpose, trip distribution by location type etc. For example, Barnes and Krizek (2005) combined NHTS and ACS data to build a relationship between the percentage of the adult population who bike in a day and bicycle commute share percentage for both

Metropolitan Statistical Areas and state level. The basic assumption of their analysis was that a large proportion of the total biking population is contributed by a small proportion of cyclists who bike frequently.

The data can also facilitate the trip generation and trip distribution steps of travel demand models. Although the vehicular trip generation model can be developed from ITE (Institute of Transportation Engineers) trip generation rates, the handbook doesn't include transit or bicycle and pedestrian facilities (Handy et al., 2013). Hence, trip generation rates for nonmotorized traffic can be estimated using factor methods and sketch planning techniques utilizing data from census or surveys (Aoun et al., 2015). Moreover, sometimes agencies build their own trip generation matrices. For example, The Delaware Valley Regional Planning Commission (DVRPC) developed a travel demand model where whey developed the person trip generation rates for nonmotorized trips of the three-person trip purposes (DVRPC, 2011). The nonmotorized trip rates per household, developed by Milwaukee, are segmented by household size, vehicle availability, and area type for all purposes (Singleton and Clifton, 2013). Benz et al., (2013) estimated bicycle demand for the Houston area using H-GAC regional transportation study 2008 household survey. Here, the authors developed stratified trip rates from the estimated number of trips and number of households for each trip purpose by worker stratification and then summed for each traffic analysis zone.

Moreover, with geographic coordinates (such as NHTS add-on), the survey data can also be used to develop a mode choice model for nonmotorized traffic (Kuzmyak et al., 2014). Clifton et al. (2008) developed a pedestrian model using NHTS add-on data where the authors estimated walk trips per household to estimate the total walk trips in the trip generation step and then went through trip distribution and traffic assignment to estimate pedestrian traffic at road segment.

2.3.2.3. **Processing Crowdsourced data**

The process of analyzing crowdsourced data to estimate demand at a traffic network depends on the structure or format of the data. For example, Strava data can be obtained in three formats: streets, origin-destination, and nodes (Munira and Sener, 2020). Hence the activity from crowdsourced data can be modeled to expanded to population-level, at a particular intersection or street, by estimating its relationship with the actual volume and socio-demographic and land use information within the buffer zone. Sanders et al., (2017) estimated the annual average daily traffic from the Strava count where the explanatory variables were the number of bike lanes on the street segment and proximity to a university. Jestico et al., (2016) used Strava data to find its relationship with manual count and other explanatory variables. The built model was then used to predict categories of ridership (low, medium, and high) for all roadways in Victoria with a predictive accuracy of 62%. The significant explanatory variable of the model was crowdsourced cyclist volume, segment slope, posted speed limit, month and if the road had on-street parking. Strauss et al. (2015) utilized GPS records from a regional bike-tracking app, in conjunction with actual observation, to compute annual average daily bike (AADB) volumes. The study reported that the correlation (R-squared values) between GPS data and count data varied from of 0.7 for signalized intersections and 0.58 for non-signalized intersections. Dadashova et al. (2020) modeled AADB from actual observation where Strava observation, reporting 29% prediction error, along with other explanatory variables, were used as predictors. Strava data have also been as proxy exposure in crash models without the validation of actual count where (Saha et al., 2018; Sener et al., 2019).

On the other hand, bike-sharing data generally comes in the origin-destination format with stations as the zones. When the bike-sharing data comes with a GPS trace, the volume at each location can be estimated based on its relationship with actual volume, as done by Pogodzinska et al. (2020) in Poland. However, in the US, the publicly available data doesn't contain GPS trace, making the demand estimation process time-consuming and challenging. No studies have addressed the limited bike-sharing data to estimate actual volume at the location level, to the author's knowledge.

2.3.3. **Scope of Nonmotorized Demand Estimation**

As mentioned in the previous sections, there are multiple modeling approaches, with varying input data requirements, complexity, usability and reliability, to estimate nonmotorized demand. The data sources outlined in the previous sections, either individually or in conjunction with other data, can be utilized to understand nonmotorized activity at varying resolutions. From the perspective of demand or volume analysis, some data need more processing steps than others. Hence, one of the key questions that entail processing individual data for demand analysis is the scope, objective, or resolution of estimation. Such scopes can be divided into three types

i. Spatial scope
ii. Temporal scope
iii. Population-level scope

The description of the scopes is provided below. The detailed understanding of the scopes also facilitates the data alignment or processing steps in section 4.

2.3.3.1. **Spatial Scope**

Under the broad category of geographic scale, the demand models can be categorized into multi-city and single-city models, the latter being the most frequently used approach. Recently, owing to the availability of pedestrian and bicycle count data across 20 U.S. metropolitan statistical areas (MSAs), Le et al. (2018) developed a national-level direct-demand model for bicycle and pedestrian traffic. Dadashova et al., 2020 estimated AADB for 155 locations across 12 cities in Texas.

Even within a single region or city, the geographic coverage of demand estimation varies. For example, Hankey et al. (2017) developed a direct demand model in a smaller community (Blacksburg, Virginia), while Griswold et al. (2019) estimated annual pedestrian crossing volumes at intersections across the whole of California State Highway System (California SHS). However, researchers (Hankey et al., 2017) have also suggested that developing a model for a small community allows gathering count data on a representative number of segments in the network

Both these approaches (single-city and multi-city) can further be categorized into two subgroups:

i. Road segment level

ii. Zone or area level

Road segment level demand analysis is a microscopic analysis where demands are estimated in the intersection (signalized or unsignalized) or street segment. The area or zone level estimation is more aggregated, hence macroscopic, than the street level estimation. Generally, within this category, the volume is estimated for a specific zone such as census block groups, traffic analysis zone (TAZ), core-based statistical area (CBSA) or even a city level. For example, Saha et al. (2018) estimated Strava bike activity at census block group level. The geographic scope mainly depends on the available data condition.

2.3.3.2. **Temporal Scope**

In addition to spatial scale, the temporal resolution for estimating demand or activity also varies. Nonmotorized volumes are generally estimated for hourly, daily, annually during the peak period level. For instance, Hankey et al. (2012) estimated 12-

hour daily volume, Pushkarev and Zupan (1971) estimated hourly flow rates, Schneider et al. (2009) extrapolated collected data to weekly counts, and Nordback (2012) estimated annual average daily bicycles. Annual average daily volume (AADB or AADP) is the most frequently used and policy-relevant unit of demand estimation (Nordback et al., 2013). The volume metrics is important as it can serve in the economic evaluation and project prioritization for nonmotorized traffic. Moreover, spatial distribution and temporal trends of nonmotorized demand, detected through the annual average daily volume, enables better planning and design of infrastructures and policies (El Esawey, 2018). This is also relevant to nonmotorized crash analysis given the fact that bike or pedestrian crashes are most often aggregated for 3 to 5 years (Fitzpatrick et al., 2018)

Although the annual average daily volume is the most plausible unit for a regional demand model, researchers in the direct demand model domain have explored several temporal units for demand estimation. For example, researchers have directly used data for the specific collection period, such as peak-period counts (Fagnant & Kockelman, 2016), often because long-term data to estimate average pedestrian and bicycle traffic were not available. Besides, the hour- or day-specific (such as weekday/weekend) models are also useful to observe the temporal variability of nonmotorized activity and can be utilized to indicate exposures to hazards for a specific period (Lu et al., 2018). Often, the time-specific models add valuable insights into the nonmotorized activity of the study area. For example, Griswold et al. (2011) developed models for 2-hour bicycle counts for weekdays and weekends, noting the variability of demand across the day of the week. Similarly, Miranda-Moreno and Fernandes (2011) developed four models for the a.m. peak period, noon period, p.m. peak period, and entire day to explain pedestrian activity in Montreal, Canada. Lu et al. (2018) developed hour-specific models (one model per hour of the day) for both bicycle and pedestrian traffic. On the other hand, Jestico et al. (2016) utilized a one-minute temporal resolution of Strava data for multiple road segments to estimate hourly, AM and PM peak period total bike ridership. Hence, the temporal resolution also depends on the resolution of collected data from different sources.

2.3.3.3. **Population-level Scope**

Population-level representation has become an important consideration for nonmotorized data, especially because the data collection effort from emerging sources is not specifically designed to be used in demand analysis. Although researchers have used Crowdsourced data, such as Strava, as proxy exposure in crash models without the validation of actual count where (Saha et al., 2018), it is possible to estimate actual volume from crowdsourced data when actual count data is available. The scaling process may or may not involve additional explanatory variables (i.e., socio-demographic and land use information). The population-level modeling has already been discussed in section 2.3.2.3.

2.3.4. **Summary**

The above discussion made it apparent that nonmotorized activity can be computed at different spatial, temporal and even population levels. Generally, it is comparatively easier to transform source estimation to a lower spatial (such as zone) or temporal resolution (such as annual volume) than transforming it into higher resolution (such as hourly volume at intersections), which requires additional assumptions and understanding of the local condition. For example, suppose a source provides a daily estimation of bicycle flow. In that case, the estimation of hourly or peak hour flow using the data would require assumptions regarding the hourly distribution of bicycle volume. Similarly, a network volume can be estimated by aggregating link-specific volume within the network. However, when the data is available in an aggregated format, computation of site-specific volume requires assigning traffic to links, which is a computationally intensive process. Moreover, some of the data sources, especially emerging data sources, requires an understanding of the population bias.

Moreover, given the varying format of different raw data sources, the model building monitoring process are unique for estimating nonmotorized volume are likely to involve multiple steps. It is also imperative to understand the representativeness and distribution of the data before incorporating it in volume or exposure analysis for nonmotorized transport.

3. LITERATURE REVIEW: BACKGROUND AND TYPES OF DATA FUSION

This chapter reviewed the literature on seven aspects of data fusion. First, it provided an overview of the fusion approaches, discussing the definitions, benefits and challenges. The second section explained the characterization schemes, defined from different research perspectives, noting the prevalence of application. In the third section, the review revealed the absence of consensus in defining the fusion terms and presented the accepted definitions and their differences, bridging the terminologies and practices of the diverse fusion research community. Anticipating the need of the research, the focus of the discussion was steered towards the decision fusion approaches in the fourth section. The section discussed the major aspects of data and sources that need to be assessed for selecting a decision fusion framework. The discussion was formulated and adopted, pivoting the transportation data context in focus. Acknowledging that there is no perfect algorithm that is optimal under all conditions, the fifth section provided an overview of the most common decision fusion algorithms highlighting their strength, limitations and uses. The sixth section underlined the need to consider the context in fusion which served as a backdrop for the novel approach proposed by this study. The discussion in the seventh section, on the application of fusion in the transportation domain, was particularly deemed important to underline that fusion mechanisms for motorized traffic data exhibit distinct characteristics and cannot be directly applied in the nonmotorized data context.

3.1. Data Fusion: Background

3.1.1. Concepts and Definitions

In general terms, fusion refers to the process of combining or integrating multiple data sources in order to obtain more accurate and less uncertain information than that provided by any individual source. The concept of data fusion, introduced in the1960s, was first implemented in the 1970s in the area of robotics and defense (Esteban et al., 2005). Since then, a substantial body of literature has focused on the why's and how's of the process of combining – or fusing –multiple data sources to support informed decision making. One of the earliest definitions of fusion, proposed by Llinas (1988), is

"*Fusion can be defined as a process of integrating information from multiple sources to produce the most specific and comprehensive unified data about an entity, activity or event. This definition has some key operative words: specific, comprehensive, and entity. From an information–theoretic point of view, fusion, to be effective as an information processing function, must (at least ideally) increase the specificity and comprehensiveness of the understanding we have about a battlefield entity or else there would be no purpose in performing the function*."

In the later years, the application of fusion spanned a wide range of areas, including robotics (Meurant, 1992), medical decision (Barillot et al., 1993), image processing (Piella, 2003), pattern recognition (Cruz et al., 2018; Awais et al., 2011), mine detection (Gunatilaka et al., 2001), aerospace systems (Schoess & Castore, 1988), law enforcement (Pau, 1990), air traffic control (Olivier et al., 2009), remote sensing (Zhang, 2010), etc. Existing fusion-focused researches have contributed to both human decision-making process, guided by the underlying fusion mechanisms, and automated decision-making where no human intervention was necessary. The disparate application of the concept leads to the development of numerous definitions, some of which are customized to the field of application (Boström et al., 2007). A more generalized definition of fusion, proposed by Gonsalves et al. (2000) is

"*The overall goal of data fusion is to combine data from multiple sources into information that has greater benefit than what would have been derived from each of the contributing parts*."

Nevertheless, the core of the definitions, proposed by numerous researchers, is the notion of improved information in terms of quality and completeness that cannot be achieved with a single sensor or source. In the context of the transportation decision-making process, the fusion or combination should support automated or semi-automated decisions, rather than being the goal or end result (El Faouzi and Klein, 2016).

As noted by Dasarathy (1999), fusion mechanisms can play a vital role when

- Multiple sources of similar data are available
- Each of the data sources represents distinct inherent properties (i.e., specific subpopulation)
- Sources can be combined to produce one or more unified information

Owing to numerous emerging applications of fusion approaches, in both research and commercial environments, the concept is yet to reach an equilibrium of uncontentious terminology and standard tools (Koks & Challa, 2003) and researchers often find it challenging to develop a clear and strict classification for fusion methodologies (Castanedo, 2013). For example, given its interdisciplinarity and growing popularity in various application domains, the terms sensor, multi-sensor, data and information fusion have been used in numerous research articles without much discrimination (Wu, 2003). Over the years, several attempts have been made to denote, interpret and classify various terms and techniques of fusion. Although the terms sensor fusion, evidence fusion, data fusion, information fusion, decision fusion, data combination, data aggregation etc., have been used interchangeably in the existing publications, in some scenarios, they indicate the difference of the fusion framework (Castanedo, 2013). The common terminologies associated with fusion are discussed in detail in section 3.3.

3.1.2. **Benefits and Challenges**

The fusion of multiple source data may offer numerous benefits over single source estimates. A handful of studies have highlighted advantages of fusion-based techniques, including increased reliability, robustness, completeness, coverage, and often cost-effectiveness (Bachmann et al., 2013a, Brooks and Iyengar, 1998; Dailey et al., 1996; Ince et al., 2012; Luo & Kay, 1989). The key advantages of fusion, as discussed in studies, are briefly discussed below.

- *Increased completeness*: if one source is unavailable, denied or lack coverage, other sources can facilitate by providing the missing information.
- *Increased confidence*: When multiple sources measure the same object, incident or activity independently, the redundant information can reduce the overall uncertainty and enhance confidence on the measurement.
- *Reduced ambiguity*: the ambiguity of interpretation may be reduced as the multiple source information reduce the set of hypotheses about the measurement.
- *Enhanced spatial and temporal coverage*: Some sources can detect events and activities at times and places where other sources would fail. When combined, multisource data provide extended special and temporal coverage.

However, along with many successes, there are still some challenges and limitations, both technical and institutional, associated with the fusion concept (Amey et al., 2009; Hall & Steinberg, 2001). The technical challenges arise from data collection, processing and computational issues and the institutional challenges refer to issues related to data ownership, privacy concerns etc. The process of data fusion often adds cost and complexity due to the system requirement for any given application. Moreover, the fusion of multiple sources comes with a risk of actually producing a worse result than the most reliable single source estimate if combined inaccurate data (Ince et al. 2012).

Therefore, both practitioners and researchers are needed to be aware of the problems and challenges associated with the fusion process before implementation. The key issues, discussed by Hall and Garga, 1999; Nahin and Pokoski, 1980 and Hall and Steinberg, 2001, relevant in the research context, are

- *Fusion is not a substitute for a good source*: No amount of fusion of inaccurate sources can replace a single accurate source estimating the activity or phenomena in question. A fusion of multiple inaccurate sources is not likely to provide a more accurate estimate.

- *Fusion of multiple highly accurate sources may not be necessary*: Fusion of multiple highly accurate sources may not necessarily further increase the accuracy of inference

- *The more may not be the merrier*: Adding sources in the fusion process may add to the complexity of the process without adding any value to the process. Nahin and Pokoski (1980) indicated that the largest marginal improvement in sensor fusion could be observed for a moderate number of sensors (less than 7). For a large number of sensors (greater than 8 or 10), the addition of identical sensors may not necessarily increase the inference accuracy

- *There is no magic or golden data fusion* algorithm: There is no perfect or one-size-fits-all algorithm that can be considered optimal under all conditions. The selection of the fusion algorithm must ensure that the input data satisfies the underlying mathematical assumptions upon which the algorithms are formulated.

3.2. Fusion Characterizations

As discussed in the previous section, due to the versatile applications of the fusion concept, there is no one-fits-all framework or fusion mechanisms. Over the years, researchers have proposed several generic as well as customized platforms to accommodate the objective at hand. The task of designing and implementing a fusion framework is complex, associated with several critical issues, including fusion architecture and algorithm selection, software implementation, and validation etc. (Hall, 2000). The development and adaptation of a framework depend on the objectives and the complexity of the target problem.

Therefore, several attempts have been made to characterize and classify multiple aspects of fusion over the years. According to Durrant-Whyte (1990), fusion can be complementary, redundant and cooperative. In the complementary fusion, the input sources representing different aspects or components of a state or object are combined to obtain complete information of the state. In the redundant fusion, multiple input sources providing information about the same state or object are combined to increase confidence and completeness. In the cooperative method, the obtained information is combined into new, more complex information. Several other researchers have also characterized fusion based on steps, level, architecture, algorithms etc. A few notable characterization and classification of fusion are discussed below.

3.2.1. Fusion Steps or Levels

Several studies have divided the fusion process into different steps or levels. One of the earliest, and still most influential, fusion tools is proposed by the US Joint Directors of Laboratories (JDL), in which five levels of data fusion were identified (White, 1991). The key steps of the process were source preprocessing, object refinement, situation assessment, impact assessment and process refinement (Llinas & Hall, 1998). These steps have been further developed or refined by multiple studies in later years (Kong et al., 2007; Ince et al., 2012; Blasch and Plano, 2003). For example, Blasch and Plano (2003) suggested including another level, “Human Refinement” to address the human interface issues. The five steps or levels can also be grouped into two types: low-level and high-level fusion comprising three key components: sources, human-computer interaction and database management system (Castanedo, 2013). The data fusion techniques, explored in transportation-related studies, mainly involve basic steps such as temporal and spatial alignment of input data, data association, and data mining for knowledge extraction (Faouzi and Klein, 2016)

3.2.2. **Fusion architecture & hierarchical levels**

One of the critical issues associated with data fusion is where the data fusion process will be performed. The design choice influences the quality of the fused data, the nature of algorithms and the complexity of processing logic (Ince et al., 2012). Although some studies have categorized the fusion architecture into two parts: centralized and distributed, Castanedo (2013) divided data fusion architecture into four types: centralized, decentralized, distributed and hierarchical. Faouzi & Klein (2016) categorized the architecture as sensor or decision-level, central-level or centralized, and hybrid. In the first category, the source detects, identifies, and produces state estimates of objects which enter the fusion processor. At the central level, data are minimally processed before entering the fusion processor, where features or attributes of the objects are assessed to produce the final output. The hybrid fusion uses both approaches based on the nature of the outputs or resources. Studies focusing on fusion application within the intelligent transportation system context generally utilize the first two architectures (Faouzi & Klein, 2016).

Furthermore, based on the output of the fusion studies (Luo and Kay 1992; Van Lint and Hoogendoorn, 2010) classified the fusion technique into three levels. Level 1 involves the fusion of raw data, for example, speed/density estimation from raw data. Level 2 involves the fusion of attributes extracted from the raw data, for example, identifying queue locations and lengths from the speeds estimated from the raw data.

Level 3 refers to the fusion of objects or events inferred from the attributes of level 2, for example, incident detection.

Dasarathy (1997) categorized fusion into three levels: data fusion, feature fusion, and classifier fusion (also known as decision fusion). They further expanded the classification based on the input/output data types, which are data in-data out (DAI-DAO), data in-feature out (DAI-FEO, feature in-feature out (FEI-FEO), feature in-decision out (FEI-DEO) and Decision In-Decision Out (DEI-DEO). The last category is known as decision fusion which combines input decisions to get better or new decisions (Castanedo, 2013). Similar categorization, although for the application in the biometric systems, was proposed by Ross & Jain (2003), where fusion can be conducted at the i) feature extraction level, ii) matching score level and iii) decision level. The decision fusion techniques, also the focus of this research, uses the classification decisions from different sources to decide the actual class/label of an object or activity to be detected.

3.2.3. **Fusion algorithms**

Researchers have classified fusion algorithms into various types (Luo et al., 2002; El Faouzi et al., 2011). Four broad categories of the existing algorithms are estimation methods, classification methods, inference methods and artificial intelligence methods. Estimation methods include non-recursive methods i.e., weighted average and least squares and recursive methods such as Kalman filtering and extended Kalman filtering. Kalman filtering methods are predominantly preferred for fusing low-level real-time dynamic multi-sensor redundant data. It applies to provide an optimal statistical solution if the system can be described with a linear model. Classification methods are utilized to identify locations or grouping of a feature vector within a multidimensional feature space based on a similarity measure. Examples of this algorithm are parametric templates, cluster analysis, K-means clustering, Kohonen feature map, etc. Inference methods include Bayesian inference and the Dempster-Shafer method. Bayesian inference integrates multi-source information based on the rules of probability theory, where the probabilities of alternative hypotheses are updated depending on observational evidence. Dempster–Shafter evidential reasoning is an extension of the Bayesian inference, which can model the lack of information concerning a proposition's probability. Inference methods, due to their capacity of evidential reasoning, are often used for the high level of inferences, such as incident detection. Examples of artificial intelligence methods include adaptive neural network, expert system, fuzzy logic etc. These techniques are generally used for high-level inferences that need human reasoning, such as pattern recognition, induction, deduction, etc. (Luo & Chang, 2010).

3.3. Explaining Terminologies: Data, Information, Sensor and Decision Fusion

Multiple terminologies such as data fusion, sensor fusion, information fusion, evidence fusion, decision fusion, data combination, data aggregation etc., have been used in the published literature to define the fusion concept. Although often used interchangeably, the terms can also illustrate the distinct difference within the fusion aspects (e.g., theory, systems, frameworks, methods, etc.) (Nakamura et al., 2007). However, due to the lack of standardized categorization and/or consensus, definitions of the terminologies are most often customized to the field of application. For example, military and remote sensing fields, being the early and most prolific area of application of fusion, generally draw the characterization of data fusion from the early definition as "*multilevel, multifaceted process dealing with the automatic detection, association, correlation, estimation, and combination of data and information from multiple sources*" (U.S. Department of Defense, 1991). Later, the term "data fusion" has been suggested to be used as an overall term for fusion (Waltz and Waltz, 2009) which also resonates with the generalized definition by Wald (1999), who stated that "*data fusion is a formal framework in which are expressed means and tools for the alliance of data originating from different sources. It aims at obtaining information of greater quality; the exact definition of 'greater quality' will depend upon the application*."

Wald (1999) also noted that data taken from the same source at different instants could be considered as distinct sources.

Highlighting the discordance in the terminology, Dasarathy (2001) adopted the term "Information Fusion" defining it as "*in the context of its usage in the society, it encompasses the theory, techniques and tools created and applied to exploit the synergy* *in the information acquired from multiple sources (sensor, databases, information gathered by humans, etc.) in such a way that the resulting decision or action is in some sense better (qualitatively or quantitatively, in terms of accuracy, robustness, etc.) than would be possible if any of these sources were used individually without such synergy exploitation*." Hence, as noted by Nakamura et al. (2007), both the terms data fusion and information fusion have equivalent meanings.

On the other hand, the terms Sensor or Multisensor Fusion, subsets of information fusion, are generally used to denote the fusion of raw sensory data (Nakamura et al., 2007). Sensor fusion generally involves objects with measurable properties, therefore, corresponds to low-levels data fusion which is numerically based (Bellenger. 2013). Although often the term "data fusion" has been used to denote fusion of raw data, as emphasized by Elmenreich (2002) and Nakamura et al. (2007), fusion involving raw (or low level) data should explicitly use the term raw or low-level data fusion for avoiding confusion.

Finally, decision fusion (often referred to as classifier fusion) involves making a decision based on the given knowledge of the perceived situation using inference methods (Nakamura et al., 2007). Decision fusion refers to high-level fusion, often taking symbolic representations of events and activities, which reasons to obtain a more accurate decision accounting for the uncertainties and constraints (Castanedo, 2013). Examples of classical inference methods are Bayesian inference and Dempster-Shafer

Belief theory. The concept has been receiving constant research attention since the last two decades (Xu et al., 1992; Kittler et al., 1998; Kuncheva et al., 2001; Kuncheva, 2014) and has been a very active field of research in machine learning and pattern recognition (Cruz et al., 2018; Bashbaghi et al., 2017; Zhang et al., 2012). As the focus of this research is decision fusion, the subsequent discussions are focused on the characterization of decision fusion algorithms.

3.4. Decision Fusion Framework: Aspects and Related Research

As discussed in the previous section, there are myriad approaches of fusion, none of which are optimal for all conditions and situations. Hence, contemplating the requirement of this research, the focus of the discussion is kept on the decision fusion approaches. The domain of decision fusion is also vast, necessitating a further understanding of inherent characteristics of the data and sources (Kuncheva, 2014; Ren et al., 2016). The following subsections describe the major aspects of data and sources that need to be assessed for selecting a decision fusion framework. The discussion was formulated and adopted pivoting the transportation data context in focus.

3.4.1. Types of outputs

The first aspect that needs to be assessed for selecting a decision or classifier fusion approach is the type of output from each data source. The fusion framework largely depends on the type of output of the classifier. Xu et al. (1992) categorized the outputs into three types:

i. Abstract level: The source or classifier provides a predicted class level without allocating any confidence measure. The abstract level is universal as any classifier can produce a level for x.

ii. Rank Level: The classifier provides a ranking of the class labels. The output of each source is ranked in order of plausibility. This type output is particularly suitable for problem data with a large number of classes, such as character, face, speaker recognition etc. (Kuncheva, 2014).

iii. Measurement level (soft level): The classifier allocates a confidence measure/score that a specific input belongs to a given class.

Here, each source or classifier D_i ($i = 1, \dots, S$) produces a n-dimensional vector [$d_{i,1}, \dots, d_{i,n}$], where $d_{i,j}$ is the support for the hypothesis that the vector x belongs to class ω_j $(j=1 \dots n)$

Other categories of output level are also available. For example, Bezdek et al. (1999) categorized output levels into three types: crisp, possibilistic and probabilistic. Here, the crisp label is equivalent to the abstract level.

Among the above-mentioned levels, the amount of information in the output increases from the abstract to the measurement level. Intuitively, the measurement level classification algorithms may obtain the best results. However, in the real application, abstract level classifiers are more common. From the nonmotorized volume perspective, the outputs are expected to be available at the abstract level. While the transformation of an output with high-level information to lower level, for example, measurement level to abstract level, is always possible, the opposite transformation (lower to higher) is only possible when adequate ground truth data is available such as by performing an empirical probability distribution over a set of training data (Ruta & Gabrys, 2000) or developing confusion matrix for the classifiers (Deng et al., 2016).

3.4.2. Type of Combiners

The strategy of combining outputs of multiple classifiers, including for abstract level fusion, can be separated into two types: trainable and non-trainable (Tulyakov et al., 2008; Duin, 2002; Xu, 2014). The trainable approach uses the outputs of each classifier to train a new model. One example of this approach is when outputs of each classifier are concatenated into a feature vector to use machine learning techniques for model training. The appeal of this approach stems from the fact that the outputs of each classifier can be directly (or without processing) used as an input for model training. However, the major limitation of this approach is every time a new source is added, the whole combiner has to be trained again.

On the other hand, the non-trainable approaches use pre defined combination rules for fusing or combining the outputs of multiple classifiers. In most cases, the outputs of the initial classifiers need to be preprocessed to be made comparable, for example, as class label probabilities. The process of transforming the outputs into probabilities is one of the most complex and crucial components of these approaches.

The majority voting method is non-trainable, while the Bayes Belief Integration and Behavior Knowledge Space methods are examples of the trainable model (Rothe et al., 2019; Battiti and Colla, 1994; Xu et al., 1992, Ji and Ma, 1997). Although the non-trainable approach may lead to sub-optimal results when compared with trainable approaches (Xu, 2014), the non-trainable approach offers the flexibility of adding a new classifier when it becomes available. Studies have explored the pros and cons of these methods in the context of fusion. Raudys (2006) had argued that complex trained rules are appropriate only when a large enough data set is available. Duin (2001) and Roli et al. (2002) asserted that trained rules are a worthwhile approach while dealing with classifiers with unbalanced performances, especially when compared with a nontrainable approach like majority voting.

3.4.3. Benchmark/Ground truth data or validation set

Ground truth data refers to the actual information collected on fields or sites which can be used to verify extracted information from processed data. For example, in the nonmotorized demand context, the ground truth data is the walk or bike volume data, collected using video images in the field.

The combination methods can also be categorized into three types, based on the use of priori knowledge or ground truth data (Di Lecce et al., 2000). The first category does not require any priori information, for example, majority voting (Suen et al., 1990; Ho et al., 1994; Hull et al., 1992). The second category uses information/performance at the level of individual classifiers as a priori knowledge, for example, the Dempster-Shafer Method (DS) (Lu and Yamaoka, 1994). The third category utilizes information at the level of the entire set of combined classifiers, for example, the BKS method (Huang, 1993, 1994)

The quality and size of the benchmark/ground truth data or validation set are also among the fundamental aspects of designing the fusion framework. Although often the same training set is used for designing trainable rules and validation, it is recommended to have an independent validation set to avoid the bias of the fusion rule (Raudys, 2006).

While the trainable models are asymptotically optimal, they require an adequate size of validation sets. Complex trainable models such as BKS demand very large validation sets, while the belief functions can be developed using a limited size of validation sets (Roli et al, 2002).

3.4.4. Correlation and Conflict

A combination of information needs consideration of other aspects of information quality. In addition to the reliability of sources (obtained by comparing with ground truth), another important aspect of information quality is its credibility (Florea et al., 2010). The credibility of a source can be quantified based on its relation/interaction with other sources. Both reliability and credibility measures influence the trust in the pieces of evidence.

The issue of credibility is related to the issue of conflict, the existence of which indicates the presence of at least one unreliable source. On the other hand, the absence of the conflict does not necessarily confirm the reliability of sources as multiple sources may incorrectly agree on a decision. (Rogova & Nimier, 2004). Therefore, it is often necessary to quantify the interactions among the information from multiple sources. Studies have proposed different methods to quantify the interactions or conflict between the pieces of information to further refine the fusion process (Yong et al., 2004; Liu, 2006; Martin et al., 2008)

Moreover, when faced with many classifiers and data sources, it is often necessary to choose some sources, among many, to increase the performance of the fusion method. The selection of a set of classifiers is often an issue before the implementation of the final fusion framework. It is because combining similar classifiers may increase the computational complexity without increasing the efficiency of the fusion. An ideal set of classifiers should be robust as well as optimal so that it can provide the best solution with the minimum amount of calculation time and data memory (Niu et al., 2007)

3.5. Decision Fusion Algorithms

As the main objective of this paper is to combine nonmotorized traffic activity data which are generally available as the abstract (crisp) level output, relevant aspects and algorithms for abstract level fusion methods serve as the focal point of the following discussions. A handful of fusion algorithms are available to handle and combine abstract-level data. A few of the most frequently used methods are discussed here.

3.5.1. **Voting Based Rules**

Voting Based methods are perhaps the oldest and simplest strategies for decision making (Day, 1988) which involves consensus patterns among the sources to make a decision. The voting methods can be applied to abstract types of outputs and have been utilized by several studies in different area of research (Lee, 1993; Ruta and Gabrys, 2002; Lin et al., 2003; Burduk, 2017; Alotaibi and Elleithy, 2016; Atallah and Al-Mousa, 2019; Lam and Suen, 1997). There are multiple voting rules based on the same underlying reasoning, equal weight to the decisions of the sources, but different consensus patterns. The most common approaches are unanimity voting, simple majority voting, plurality voting. Another voting approach is threshold voting which considers a threshold to make a decision whether an object belongs to a specific class. For example, while simple majority voting needs to take three out of five votes to decide an outcome, a threshold of two as the minimum necessary number can be selected to assign a label (Raol, 2015).

Another variant of the voting approach is weighted voting which relaxes the assumption of equal weightage by assigning different weights to each decision, commonly estimated based on the reliability of the sources. Generally, the weight on the decision of the sources that perform better than others should be high. Hence, it requires a training sample to compute the reliability or accuracy of the sources. The approach has also been utilized by several studies (Kim et al., 2011; Li and Ngan 2017; Kuncheva and Rodríguez, 2014; Lopez-Otero et al., 2017) for fusion. The success of the approach depends on the method of computing weights. Nagi and Bhattacharyya (2013) suggested that the weighted voting method performs well only when the weights of the classifiers are precisely allocated. While some researchers have followed the conventional approach of computing weightage based on classification accuracy (Moslem et al., 2011), several studies have utilized some other performance measures, such as Area under the ROC Curve (AUC), F-score etc., for weight estimation (Texier et al., 2019; Lopez-Otero et al., 2017). For example, Lopez-Otero et al. (2017) evaluated reliability based on precision, recall, accuracy, F-score and mutual information. On the other hand, in the method proposed by Kim et al. (2011), larger weights are assigned on classifiers that perform better on hard-to-classify instances. Moreover, viewing Weigh computation as an optimization problem, researchers have also utilized artificial intelligence techniques for improving the classification performance (Ekbal and Saha, 2011)

Despite the simplicity, the approach along with its variants, have been found to be effective in the decision fusion process. Lee (1993) had utilized and compared this method with several other complex combination methods such as Bayesian, neural network etc. and indicated that the simple majority rule is just as effective as the other complex methods for improving the recognition rate. Roli et al. (2002) suggested that majority voting generally performs well for classifiers with a similar level of accuracy. Lin et al. (2003) illustrated the effectiveness of plurality voting in achieving the tradeoff between rejection rate and error rate. The study concluded that the combination of independent classifiers, using the voting method, can result in significant improvement in accuracy.

The mathematical formulation of the voting approach is discussed in detail in section 5.1.1.

3.5.2. **Behavior-Knowledge Space (BKS)**

Behavior-Knowledge Space (BKS) is another abstract level fusion method where a knowledge space is created to record the behavior of all the classifiers on a set of samples. In general terms, it creates a lookup table containing every classifier output combination (Huang and Suen 1993). The table is created using a training set where each of the combinations of the class labels is used to match the ground truth values to estimate their relative frequencies. For an unknown test sample, the output class label of each classifier is used to index a unit of the table to assign it to the class label with the most training samples. The table below is an example BKS lookup table with a two label, two source problem (Butler, 2012):

Table 3.1: Example of BKS lookup table

Source 1	Source 2	Ground Truth	% Occurrence
1	1	1	90%
0	0	0	80%
1	0	0	50%
0	1	1	75%

According to the table above, source 2 is good at identifying label 1 compared to source 1. Moreover, if both sources assign label 1, it is highly likely to be true.

The strength of the BKS method stems from the fact that it doesn't require the decisions of the classifiers to be independent. However, it demands large training samples to build the knowledge space, hence not suitable for problems with low sample size. It also takes a lot of memory when the number of sources and possible outcomes is large. Moreover, the combiner sometimes suffers from the overtraining problem that means it may work well on the training data but poorly on testing data (Cozzolino, 2013; Kuncheva, 2014). The method has been used in multiple research areas, such as forgery detection (Ferreira et al., 2016), signal processing (Yang et al., 2019) etc.

3.5.3. **Bayesian Inference**

Bayesian fusion uses Bayes probability theory to combine evidence from multiple sources. The method is formulated on the assumption of mutual independence of the sources. The statistical inference method represents uncertainty in terms of conditional probabilities describing the belief of a hypothesis. The belief can take values in the interval [0,1], 1 being the absolute belief and 0 being no belief (Nakamura et al., 2007). Bayes rule is used to obtain the posterior probability density function which is based on the prior belief of the hypothesis and conditional probability (likelihood) of the observation.

It is also capable of incorporating a wide range of prior knowledge, hence applicable for the analysis when small sample size is available (Antal et al., 2008). When more data is available, the approach can be used to update the posterior probability of based on the additional observations. The popularity of the Bayesian approach stems from its rigorous statistical underpinning and ability to quantify and reduce uncertainties in complex problems (Kuncheva, 2014; Nagel, 2019). The performance of the method has been reported as quite robust, even when the assumption of independence is violated (Kuncheva, 2014). However, the Bayesian inference has some major shortcomings, as highlighted by Hall and Llinas (1997). In addition to the fact that the approach demands the hypothesis to be mutually exclusive, it also needs a well-defined prior and conditional probabilities of the hypothesis. Moreover, the method cannot account for the general uncertainty of the data (Wu, 2003).

The Bayesian fusion method has been applied in various areas including to detect events (Atrey et al., 2006); tracking (Stolkin et al., 2012); image processing (Wei et al., 2015; Khademi et al., 2018) etc.

3.5.4. **Dempster Shafer Evidential Reasoning**

The Dempster Shafer theory (DST), first introduced by Dempster (Dempster 1967) and then formalized by Shafer (Shafer 1976), is one of the most popular methods that characterize data using probability mass to combine multiple sources. The framework can address both probabilistic and epistemic uncertainty and process uncertain, imprecise, and incomplete information from multiple data sources (Wang, 2016). Compared with the Bayesian model, the DST theory can explicitly model the absence of information and allows a finer representation of uncertainty in a comprehensive framework. The framework is recognized due to its ability to simultaneously handle conflict and incompleteness. In contrast to the Bayesian method, the DST relaxes the mutual exclusivity assumption and doesn't demand a well assignment of the prior probability; rather the probability assignment depends upon the availability of supporting information (Wang, 2016).

The theory has been utilized in a wide area of research, including transportation such as to estimate crash analysis (Leung et al., 2017; Sohn & Lee, 2001), incident detection (Klein, 2000); evaluating sustainable transport solutions (Awasthi & Chauhan, 2011); public transportation planning (Kronprasert & Kikuchi, 2011); travel time (El Faouzi et al., 2009); traffic flow prediction (Soua et al., 2016); traffic state estimation (Kong & Liu, 2009; De Donato et al., 2014).

A comprehensive discussion along with the mathematical formulation of the Dempster Shafer was presented in section 5.2.

3.6. Fusion with Context Consideration

The concept of context-aware computing has been introduced within the perspective of sensor fusion, mainly in the computer science research domain, acknowledging the need of sensing and managing context information (Wu, 2003). The formal definition of context-aware computing can be found from Dey and Abowd (2000), who suggested that "*Context is any information that can be used to characterize the situation of an entity. An entity is a person, place, or object that is considered relevant to the interaction between a user and an application, including the user and application themselves"*

Therefore, the term context can cover an extremely broad range, such as spatial, temporal and environmental information, social and physiological situation etc. For example, ultrasonic sensors are affected by the change in temperature (Hillel et al., 2014) or certain monitoring sensors in airplanes exhibit different fault patterns during ascent/take of (Sarkar et al., 2013)

The context can be categorized as intrinsic context, which directly influences source measurement and extrinsic context, where the interpretation of the collected data is affected (Virani, 2017). Dey and Abowd (2000) identified the primary categories of context which are: identity, activity, location, and time. The first two are high-level attributes while the last two categories low-level attributes. Contexts are generally task-specific in nature, varying across the application domain. For example, for object recognition by image processing, the context can be the visual scene while for a mobile tour planning application, the user's availability and preference would be considered as context (Wu, 2003).

Researchers focusing on context-aware computing have relied on both automatically and manually acquired information. Although, ideally, automatically obtained context would nullify the need for manual acquisition, in real-world applications, context sensing mostly relies on manual input (Dey and Abowd, 2001). A wide range of methods has been applied to handle context in fusion, where the methods interpret context from varying perspectives.

Despite a growing emphasis by several researchers, context consideration is often avoided in the application as there is no general way to extract and handle context. Liu et al. (2017a) proposed fusion with contextual reliability evaluation which relies on inner reliability and relative reliability concepts. From a different perspective, researchers have pursued context extraction for sensor fusion using clustering techniques (Frigui et al., 2008; Virani et al., 2016) and mixture modeling methods (Ratto, 2012). However, most of these methods are generally steered for data rich sensor fusion environment. Nonetheless, researchers have emphasized the need of incorporating context for better accuracy of fusion.

Although within the same domain, the context has been deemed from a different perspective, particularly in the application of the Dempster Shafer algorithms. For sensor fusion, researchers have computed contextual discounting of the belief function that considers the reliability of a source of information (Mercier et al., 2008). For instance, a sensor may be more efficient in recognizing a specific type of target than other types. Another study (Mercier et al., 2012) focused on generating contextual discounting, hypothesizing that some sensors may have better reliability in recognizing some situations or objects. However, to the author's knowledge, no research within the decision fusion domain has incorporated the situation or context of sources (such as place or time), obtained from a subjective judgment, for refining the fusion algorithm.

3.7. Data Fusion in Transportation Research

Owing to the advancement of technology followed by the overwhelming volume of data from heterogeneous sources along with the demand for highly accurate estimates, the transportation area has been recognized as an ideal candidate for the application of the data fusion approach. The tool was first utilized by El Faouzi and Lesort (1995) and Sethi et al. (1995) in the traffic and transportation area. Since then, numerous efforts have been made in this area. The users of the fusion approach were categorized into three types: transportation system users, transportation service providers, and transportation planners (Amey et al., 2009). Notable examples of the area of application include advanced transportation management systems, automatic incident detection, network-wide control, advanced traveler information systems, advanced driver assistance, traffic demand estimation, traffic forecasting and traffic monitoring, traffic flow variables estimation for motorized traffic etc. (Faouzi and Klein, 2016; He et al., 2016).

The main sources of motorized traffic-related data are GPS, blue tooth, video, automatic vehicle identification (AVI), plate scanning, loop detector, probe vehicle, Cellular handoff probe system (CHPS), microwave sensors, CCTVs, toll collection stations data, historical data, taxi trajectories, etc. (Wu et al.,2015; Zhu et al., 2016; Zhu et al., 2016; Alibabai and Mahmassani,2009); Meng et al., 2017). A substantial number of studies have highlighted the need and superiority of fusion to provide a complete depiction of traffic state on roads, given that the individual data sources are often associated with sparseness and noise (Treiber et al., 2011). The fusion approach has been used for estimating speed, flow, density, travel time, queue length, etc., (Peng et al., 2009; Han, 2012; Van Lint and Hoogendoorn, 2010; Li et al., 2013; Nanthawichit et al., 2003; Zhou et al., 2011). The fusion approach has also been utilized in computing vehicle trajectories (Mehran and Kuwahara, 2013), OD flow estimation (Lu et al., 2015); Dynamic Origin-Destination Demand Matrices (Zhu et al., 2016; Alibabai and Mahmassani (2009); City-wide traffic volume (Meng et al., 2017).

The characteristics of the data from multiple sensors and the flexibility of quantifying traffic parameters using the dynamics of relationship have made it possible to explore and compare a wide array of approach or algorithms, such as Kalman filters, Bayes methods and Artificial Neural Networks, Simple Convex Combination, Dempster-Shaefer theory or Fuzzy Logic, etc. (Ou, 2011). The choice of approach is determined by domain specific constraints, input parameters, operational constraints and the objective and application of the fusion (Klein 2012; El Faouzi and Klein (2016). For example, fusion approaches to estimating traffic states, where traffic flow variables (speed, position, travel time, density, etc.) are computable using established measurement equations (such as state-space model) from raw sensor data, generally utilize model-based approach, such as Kalman filter (Zhou et al., 2011; Peng et al., 2009). Studies have also used data-driven approaches, incorporating statistical models such as artificial neural networks (ANN) or regression, for fusing data representing traffic state variables (Zhou, 2015).

Although a majority of the studies have reported the superior performance of fusion approaches over individual estimates, multiple studies have acknowledged the fact that it is difficult to conclude which technique performs the best. A study by Bachmann et al. (2013a), comparing seven fusion-based estimation process to combine data from loop detectors and probe vehicles to compute highway traffic speed, concluded that most techniques tend to improve the accuracy of estimation when compared to estimation based on single sources. However, in addition to the technique, the improvement depends on the number of sources (such as probe vehicles) as well as the situation (i.e., traffic conditions).

A more focused literature review, conducted by this research, also revealed the use of decision fusion algorithms, including Bayesian Estimation and Dempster-Shafer evidence theory (Liu et al., 2016; Zhu et al., 2016; Shi et al., 2017; He et al., 2016) for estimating travel time, traffic state, OD matrices etc. He et al. (2016) noted that the effectiveness of Dempster-Shafer fusion generally relies on generating the liability of measurements from the sources. Shi et al. (2017) found the Dempster-Shafer approach efficient in fusing travel time data from various sources under different information conflict situations, including highly inconsistent, slightly inconsistent, and highly consistent situations.

Section II Summary

The review of literature, in chapters 2 and 3, underlined the dire need for robust and accurate nonmotorized activity-related information revealing potential avenues of research of combining-or fusing- multiple nonmotorized data sources. The review asserted that both traditional and emerging data sources representing nonmotorized traffic activity come with great potential to understand the trend, attitudes and behaviors of bicyclists and pedestrians. However, each of the sources exhibits some limitations in terms of spatial, temporal and population representation which subsequently influence the accuracy and reliability of associated demand models or estimation process. Moreover, the individual sources or models are often incomplete, conflicting, or scattered. Hence, a fusion mechanism that can offer completeness along with improved reliability and accuracy may prove to be a powerful tool to overcome the lack of knowledge or limitations associated with the individual sources.

Outlining the commonly used algorithms and tools of fusion, the discussion also emphasized that the nonmotorized traffic activity data exhibits unique characteristics, in contrast to motorized traffic data, and methods and steps of estimating demand or states for motorized traffic cannot be directly adopted in nonmotorized traffic areas. The review revealed that the decision fusion domain is an appropriate fusion mechanism that can accommodate the characteristics of the nonmotorized traffic data sources. Finally, the discussion of various decision fusion algorithms, along with their strength and limitations, served a crucial role in selecting potential fusion algorithms for this research. The discussion on context-aware fusion studies, coupled with the unique features of nonmotorized data, motivated a new research question as to whether the incorporation of context credibility can refine the fusion process.

4. NONMOTORIZED DATA TO DEMAND MODELING

The review of literature in the previous two chapters, outlined the generalized options of decision fusion mechanisms given the inherent characteristics of nonmotorized activity data and aspects related to demand analysis.

One of the crucial issues associated with nonmotorized data, considering the fusion perspective, stems from the fact that different data sources come in different formats and structures. The spatial, temporal and population-level representation also vary across sources. Hence, the first step of the fusion endeavor is to process the raw sources to obtain a homogeneous representation. Therefore, chapter 4 elaborated a conceptual framework outlining the steps and aspects of processing and homogenizing the sources, in the light of the discussion delineated in section 2.3. This chapter also intended to create a better understanding of the sources and models along with their coverage, representativeness and skewness. At the end of the chapter, the potential context of nonmotorized traffic activity estimates, required for the betterment of the fusion mechanisms, was discussed

In an endeavor to select and customize fusion mechanisms to accommodate the processed/modeled data, the research considered the practical applicability under different circumstances of benchmark data availability. In addition to traditional approaches, the novel approaches proposed by this dissertation also contemplated the scenarios. Chapter 5 was devoted to illustrating the mathematical formalism of the algorithms utilized in this research.

4.1. The Framework of Processing Nonmotorized Traffic Data

As discussed in section 2.3, nonmotorized data, both traditional and crowdsourced, are generated in different structures and formats. While some of the sources can be processed to estimate activity, some, such as Web 2.0 Tech, are not meant for demand representation. Moreover, some of the data sources need minimal processing while others have to go through an array of steps to compute demand, given the scope of estimation.

However, there is no clear guidance of how different data sources can be processed and brought together to compute nonmotorized volume or exposure within a facility or area. Motivated by the ardent need for a comprehensive guideline, this research developed a conceptual framework that outlines the steps and aspects of processing and homogenizing different sources. The outcome of data processing, illustrated in this section, will be used in the fusion approach, discussed in the next chapter. The framework consisting of processing data and applying fusion mechanisms can be applied to all nonmotorized traffic data such as bicycle and pedestrian. However, for demonstration purposes, as a case study, only bicycle-related data were collected and processed in the following chapters.

The key steps of gathering and processing nonmotorized traffic data are

Step 1: Select the study area

Step 2: Gather available data sources and identify the data structure

Step 3: Determine the scope of estimation

Step 4: Analyze/model data sources to obtain a homogenized representation

The following sections explained the steps in detail and present a successful case study as an example of the application.

4.2. Select Study Area

The first step of the framework is to select a study area for gathering, processing and analyzing nonmotorized traffic data. The study area can be a city, county, multiple census tracts or Traffic Analysis Zone (TAZ) etc. The selection of a study area depends on the ultimate objectives of the researchers (such as to analyze pedestrian crashes in a city or design bike lanes in an urbanized area), data availability etc. The size of the study area should also be a consideration as some nonmotorized models (such as the direct demand model) exhibit scalability and transferability issues (Munira et al., 2021). For example, while building a bike direct demand model for the City of Austin, Munira et al., 2021 noted that the model prediction was drawn for locations within the limits of the city's bicycle route map and did not include areas farther into the suburban regions.

For this study, the city of Austin was selected as a case study. With an area of 326 square miles, the city of Austin accommodates a population of over 996,369 (City of Austin Planning and Zoning, 2020). Downtown Austin, which is located on the north bank of the Colorado River, is the central business district of the city. The University of Texas (UT) at Austin, accommodating over 50,000 students, is located north of the downtown area. Although the eastern part of the city is flat, the western part contains some hilly terrain. The city is also home to several natural and man-made lakes. By its very nature, the city is diverse in terms of age, culture, income, and built environment characteristics, and it has experienced a steep rise in the degree of socioeconomic spatial separation over the last few decades. Despite being heavily car dependent, especially in suburban neighborhoods, the city has observed a significant increase in bicycle commuters in the last few years.

Austin makes an excellent case study for this research for multiple reasons. Austin has ranked 27th among the US cities in terms of high bicycling and walking levels in 2012 (Swanson, 2012). The same report also indicated that Austin stood 15th and 37th for bicycle and walk commute share respectively when compared with other US cities. Being known as a bike-friendly city, it is endeavoring to adopt a holistic approach to increase safety and mobility for pedestrians and bicyclists of all ages. Austin City Council adopted the 2014 Austin Bicycle Master Plan to develop a connected and protected walking and biking network. The city accommodates a total of 267.5 miles of bicycle facilities, including protected and buffered bicycle lanes and urban trails (City of Austin, 2018a). Approximately, 36% of the region's arterial streets have traditional painted bicycle lanes (City of Austin. 2015). The planning and implementation of various projects since 2009 have observed a significant increase of bicycle commuters in the city. The citywide mode share of bicycles doubled in 2011 (around 2%) compared to 2009 (City of Austin, 2015). Moreover, within the 32 square miles area of central Austin, 5.5% mode share was observed for bicycles in 2012. Some census tracts in the area even reported as high as 13% mode share for bicycle commute.

The city also adopted the Vision Zero initiative to reduce traffic related death and injury to zero by the year 2025. To promote safe walking and biking among the residents, the city has also taken various programs including nonmotorized friendly street design, pedestrian safety action plan etc. Given the strong commitment to its Vision Zero goals and long-term planning to improve nonmotorized infrastructure, the city needs reliable and robust nonmotorized activity data and tools to facilitate strategic data-informed decisions. Therefore, a comprehensive data fusion framework for reliable nonmotorized demand/exposure estimation, applied for this study, would be of great value to policymakers, and practitioners along with the scholars.

4.3. Gather Available Data Sources and Identify the Data Structure

The second step after selecting the study area is to identify available data sources relevant to nonmotorized volume estimates in the region. Table 2.1 depicted the characterization of the nonmotorized data sources and Table 2.2 presented their plausible application from a demand analysis perspective. Hence, it is important to gather all available data sources and recognize the attributes associated with each source, such as the frequency of data collection, geographic unit, spatial, temporal and population-level resolution, etc.

For the City of Austin, five primary data sources, relevant to bike activity were gathered. These were:

i. Actual bicycle volume counts (Permanent and Short count)
ii. Bicycle sharing data
iii. NHTS add on data
iv. Strava data
v. StreetLight data

The ACS and other land-use related data were also extracted for the study area, which served as auxiliary data sources for modeling purposes.

The five key datasets represented bike activity at various spatial, temporal and sub-population representations. The structure of the datasets was presented in Table 4.1

Table 4.1: Structure of the Bicycle Data Sources in Austin

Data Sources	Temporal Coverage or Year of Data Collection	Spatial Coverage	Population Resolution	Source
Video-based Short count	24-hour count in 2017	44 intersection	Yes	City of Austin Transportation Department
Inductive Loop based permanent count	Constant count from 2012 to 2017	11 locations	Yes	City of Austin Transportation Department
NHTS-Addon Survey	Gathered in 2017	Households in the City of Austin	Yes	TxDOT
Strava	Trips in 2017	Intersections in City of Austin	No	TxDOT
Bike Sharing Data	Trips in 2017	63 stations in Downtown Austin	No	Public Website
StreetLight Data	Trips in 2018	Zone/Intersections in City of Austin	No	StreetLight Inc.
ACS data	2017	Census tract or block group level in City of Austin	N/A	Public Website
Land Use Related Data	2014 to 2017	Point or Area-based data in the City of Austin	N/A	Internal communication and Public Website

Table 4.1 confirmed that the data sources exhibit heterogeneous structures. In order to obtain a homogeneous estimate, which is meaningful and relevant to policy planning, each of the data sources needs to be cleaned, processed and modeled given a scope of estimation as explained in the next section.

4.4. Determine the Scope of Estimation

The processing of individual data sources inevitably entails the question of scope for demand analysis, outlined in section 2.3.3. The scopes of estimation are Spatial scope, Temporal scope and Population-level scope. As the details of the scopes and their applicability have already been discussed, the following section explains the premise and reasoning behind selecting scope/unit of estimation for the case study.

4.4.1. Temporal Unit of Volume/Demand Estimation

It is imperative to select a temporal unit of analysis for representing nonmotorized traffic demand from all sources. Intuitively, estimate at the finer unit (such as hourly volume) requires data at finer temporal resolution compared to a larger unit (such as annual average volume). However, data at finer resolution may not be available from all sources. For example, the four-step demand modeling approach generally estimates traffic as annual average daily volume. Transforming estimations at a finer unit requires the involvement of multiple assumptions and an understanding of the local condition.

The temporal unit of analysis for this study was selected as annual average daily bicycle (AADB) volume. The rationale behind the selection was also motivated by the fact that the particular metric is a widely accepted policy-relevant unit of demand representation for policy planning and safety analysis.

4.4.2. **Geographic Scale of Volume/Demand Estimation**

The task of estimating nonmotorized demand requires the analyst to select a geographic or spatial unit for which the volume will be estimated. The plausible spatial units are facility level (such as intersections, mid-block) or zone level (such as TAZ, Block etc.) (Turner et al., 2017).

For this research, intersections are selected as the spatial unit of volume analysis. The rationale behind the selection stemmed from the fact that a large proportion of crashes occur at intersections in the Texas region (Texas Department of Transportation, 2016). A microscopic level (intersection) analysis is expected to facilitate focused policy efforts and effective safety implementation plans. Moreover, the actual count data for the study area was available at the intersection level, which is expected to serve as the ground truth and benchmark data for models.

To identify the intersections for the study area, a bicycle network, developed and managed by the City of Austin Transportation Department, was obtained. The bike route network is generally different from the regular street maps as it identifies various bicycle related facility segments (off-street, on-street bicycle facilities, special facilities etc.) in addition to regular roadways that allow bicyclists. The process of extracting intersections from the bike network was explained in section 4.5.1.1

4.4.3. Population Scale

Given that some of the data sources represent a different subpopulation of the actual bike activity, it is only logical to expand or scale such activity from individual sources to the total population level. While the analyst may not have to choose a particular population scale for homogenization, it is imperative to understand the representativeness and skewness of the sources.

4.5. Analyze/Model Data Sources to Obtain Homogenize Representation

Given the selected scopes of bike activity estimation, AADB at intersections, the data sources of different structures have to go through multiple analysis steps. Using the data sources, tabulated in table 4.1, this research developed five models, which were:

i. Direct Demand Model
ii. Four-Step Model
iii. Bike Sharing Model
iv. Strava Model &
v. StreetLight Model

A tabulated summary of the key features of the modeling steps was presented in Table 4.2, followed by an in-depth discussion of individual processes.

Table 4.2: Key Steps of Volume Estimation Models

Model	Main Input	Use of Explanatory variables	Model Type or steps	Adjust Population Scale?
Direct Demand Model	Short and permanent Counter data	Yes	Regression	No
Four-Step Model	Demographics and Employment data,	Yes	Trip generation, distribution, mode choice, assignment	No
Bike-Sharing Model	Bike-sharing volume	No	Trip assignment	Yes
Strava Model	Node/road level app volume	Yes	Regression	Yes
StreetLight Model	Zone level volume	Yes	Regression	Yes

The description of the process was divided into two parts. As each of the data sources had to undergo some processing and modeling tasks, the common steps that were utilized multiple times, such as gathering auxiliary data, regression model building process, traffic assignment and actual volume process were explained, in the first part. In the second part, the models developed from the induvial sources were explained.

4.5.1. **General Steps**

Under this section, some of the steps that are relevant to multiple or all models, developed by this study, are explained first to avoid repetition when discussing individual models. The discussion included processing of bike network data to obtain intersections for the study area, processing of explanatory variables for models and creating bikeability index measures, developed for modeling and simulation purposes.

The section also explained the generalized variable and model building process and traffic assignment task, required for bike-sharing and four-step models.

4.5.1.1. **Processing bike network**

The bike network of the study area is one of the crucial and essential information or input required for activity or demand estimation of nonmotorized traffic. It is required to locate the intersections or street segment locations for which the demand has to be computed. The network map is also necessary for route choice analysis in traffic assignment task. A bike or pedestrian route network is generally different from the regular street maps as it identifies various bicycle or pedestrian specific facilities and excludes the access-controlled routes on which nonmotorized activities are not allowed.

For the study area, a bicycle network map, developed and managed by the City of Austin Transportation Department, was obtained. The map included information regarding both bicycle facility type (on-street/off-street etc.) and bicycle comfort level for each segment. The comfort level was computed based on a study by Geller (2009), taking into account a number of factors, including traffic speeds and volumes, roadway widths, bicycle facility type, and other readily available metrics, to determine how comfortable a segment is for people of all ages and abilities. The comfort level was mainly categorized into four types: high comfort sections, medium comfort sections, low comfort sections, and extremely low comfort sections. The bike network excluded road segments that are without any bike facility or of low comfort but with an alternative route option. The residents of Austin are also allowed to suggest new routes or to update the comfort levels of the existing routes. Hence, the mapping follows a robust process that takes continuous advantage of public feedback to avoid the routes where bicyclists never or very seldom ride. Therefore, the network is expected to provide a good representation of the actual route choices by the bicyclists of the study area.

As the network was processed to be used in intersection identification and route choice analysis, it was imperative to perform a comprehensive quality check to check its completeness and accuracy. It was observed that the route map didn't include access-controlled roads (i.e., Interstate highways) where bicyclists are not allowed, as expected. Moreover, acknowledging the fact that even the most carefully developed networks are bound to have errors, the obtained network file was checked for two common digitizing errors: overshoots and undershoots. Overshoot and undershoot errors happen when a line is not connected with the neighboring line with which it should intersect (Krizek et al., 2009). Following an exhaustive manual investigation of the network, a network correction task was performed (in ArcGIS) to avoid error in the traffic assignment process.

The final task was to identify intersections from the route network. Both three and four-legged intersections were identified in the process. The study area consists of a total of 2,518 intersections as shown in figure 4.1.

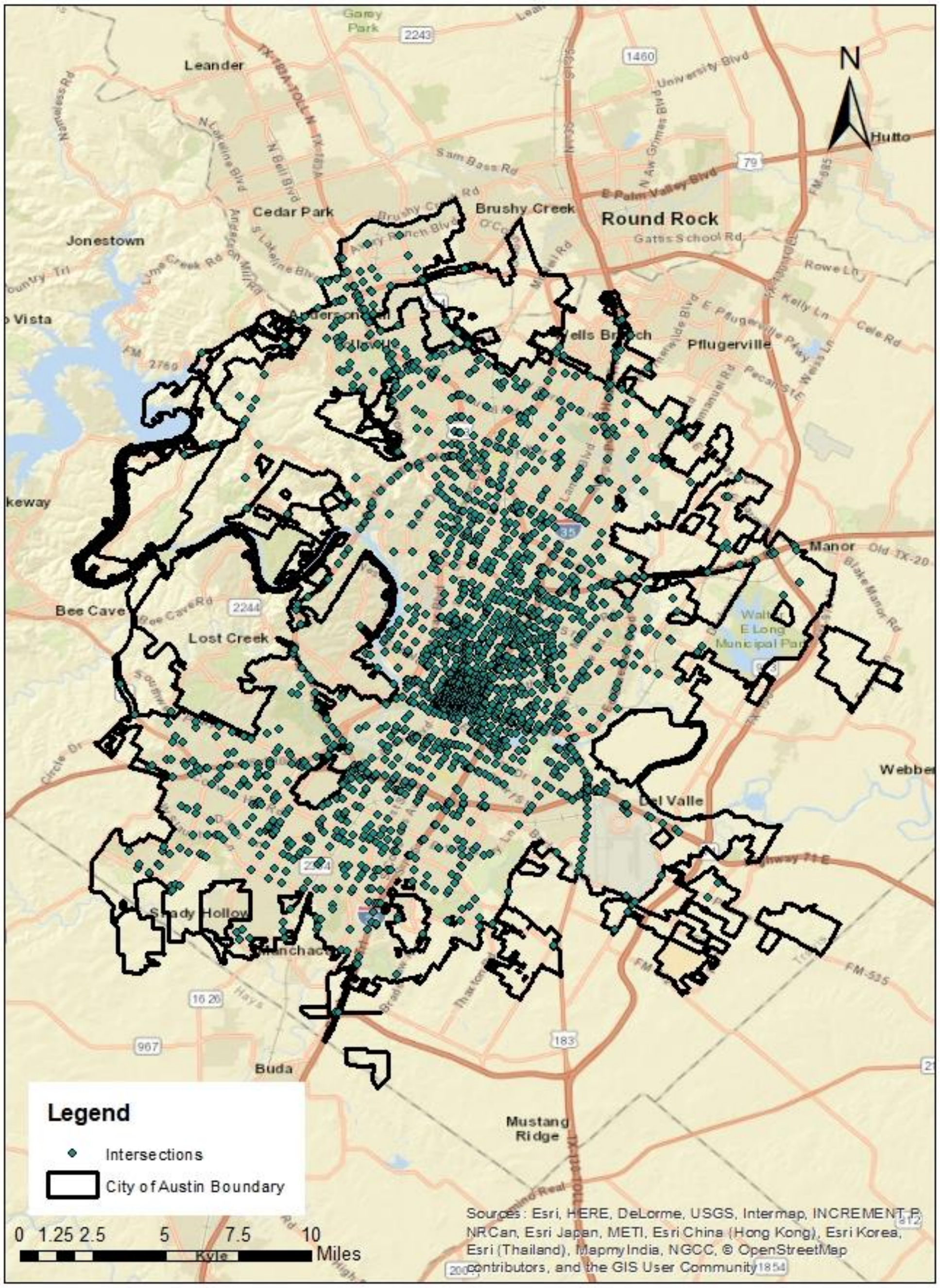


Figure 4.1: Intersections within the study area

4.5.1.2. **Processing explanatory variables**

In an effort to assemble explanatory variables for the study, prior studies focusing on direct demand models and bike determinants were reviewed (Chen et al., 2017a; Dill, 2009; Hankey et al., 2017; Hasani et al., 2019; Tabeshian & Kattan, 2014). This study sought to build a rich set of explanatory variables using insights from the earlier studies as well as the data available for the study area. To develop the explanatory variables for the model, first, a comprehensive search was performed to see which data were publicly available. Data were gathered from the City of Austin data portal (https://data.austintexas.gov/), 2017 American Community Survey (https://data.census.gov/cedsci/), City of Austin Planning and Development Review Department, Texas Education Agency (http://schoolsdata2-tea-texas.opendata.arcgis.com/datasets/059432fd0dcb4a208974c235e837c94f_0), Austin Transportation Department Arterial Management Division, Capital Metro (https://data.texas.gov/Transportation/CapMetro-Shapefiles-JUNE-2018/rwce-6ann), and BCycle (Austin bike-sharing agency) data portal (https://data.austintexas.gov/Transportation-and-Mobility/Austin-MetroBike-Trips/tyfh-5r8s). After identifying the gap between the required and available datasets, relevant authorities were contacted and asked to provide additional data for research purposes. Additionally, a review of factors included in past studies identified a need to explore new variables related to the features that have the potential to drive the bicycle demand at an intersection. Therefore, this study examined some additional variables—including the presence of bike-sharing stations or bike signals around an intersection, bike-accessible bridges, and so forth—anticipating their possible impacts on bicycle volume in the area.

Since all of the raw datasets obtained from different sources were at different spatial scales, the datasets were cleaned and processed to bring them to homogenous spatial scales (buffer level). Over 400 variables for three buffer zones—0.1 miles, 0.5 miles, and 1 mile—were created. The variables were categorized into seven groups following the categorization suggested by Munira and Sener (2017): demographics, socioeconomics, network/interaction with vehicle traffic, pedestrian- or bicycle-specific infrastructure, transit facilities, major generators, and land use.

The demographic and socioeconomic variables included age, gender, education, race, household size and occupancy status, income, and commute mode and time of the surrounding population. The network and bicycle-specific infrastructure-related variables included different types of bicycle infrastructure based on the conditions and comfort level, which was developed by the City of Austin (2017), as well as bike signal, intersection density, and bike-sharing stations. Various transit-facility-related variables were compiled, including frequency of transit stops, transit route length, and distance from hub locations. Major generators and land use variables, such as the number of schools, offices, industries, open areas, mixed-use developments, water areas, and bicycle-accessible bridges, were also gathered based on available data.

4.5.1.3. Creating Bikeability Index

The research developed a Bikeability Index- a composite measure to quantify bike friendliness of the network, mainly to handle the issue of small sample size of the study area. As noted by Munira et al. (2021), the small sample puts constraints on the modeling process including the direct demand models in multiple ways such as limiting the number of predictors that can be added to the regression model to avoid the overfitting issue (Howell, 1997). The issue may result in the omission of important variables.

The bikeability index was mainly inspired from the approach of Winters et al. (2013). The process of developing the bikeability index was detailed in the Munira et al. (2021). Five attributes were used in the process which were: bicycle route length, high comfort bicycle route length, connectivity of bicycle-friendly streets, destination density and transit coverage, within 1 mile of the intersection. Each of the five attributes were scored in a scale of 1 to 10 based on the quantile value, where 1 = *least bikeable environment* and 10 = *most bikeable environment*. The final index value was created by summing the scale value for each attribute. The highest index value could range from 5 to 50. Munira et al. (2021) showed that the high bikeability intersections (above 40) were located near the central region.

4.5.1.4. **Variable selection and model building**

For variable selection, an extensive three-stage procedure was followed. First, the relative strengths of relationships between each of the explanatory variables and the dependent variable was assessed through a simple OLS model. At this stage, variables that had significant association (at a 90% confidence level) with the dependent variable (AADB at intersections) were identified. Second, correlation between variables were examined through Pearson's correlation coefficients. In this stage, a large number of significantly correlated variable pairs (at 0.7) were recognized. Finally, by iterating numerous combinations of variables which are not highly correlated, a final model was identified examining its predictive accuracy such as mean absolute error (MAE), root mean square error (RMSE), misclassification error and fitness (adjusted R^2). In addition, while selecting the variables for the best model, the statistical significance of individual variables and intuitive interpretation were also considered.

The performance of the models was examined using cross-validation, a resampling technique that helps identify a parameter value, in order to ensure a proper balance between bias and variance (Chan-Lau, 2017). For this process, a 10-fold, cross-validation method was used to evaluate and compare the performance of the developed models. The performance evaluation criterion was the average accuracy.

It's worth noting here that the process of variable selection could also be done through various state-of-the-art machine learning approaches, such as Lasso or Random

Forest which are often referred to as the black box approach with limited interpretability (Chen et al. 2017a). The approach followed in this study intended to promote a profound understanding of the variables along with the dynamics of their relationship.

4.5.1.5. **Traffic Assignment**

A traffic assignment task is needed to allocate traffic to the facilities of the transport network. This is the last step of the four-step modeling process. The step is also required in bike-sharing traffic allocation as the raw data generally comes in a zone (station)-to-zone(station) trip format.

For conducting the traffic assignment task, several software packages are available such as PTV VISUM from PTV Group, EMME by INRO, CUBE Voyager by Citilabs, TransCAD by Citilabs, which require a commercial license. In order to be used in the bicycle trip assignment task for this dissertation, an open-source software was preferred. It was found that “Tranus” is an open-source software that can be used in developing land use and transportation model at an urban or regional scale. First developed in 1982 by Modelistica, the software has gone through several versions incorporating theoretical developments and practical requirements (Parsons Brinckerhoff Quade & Douglas, Inc., 1999). This project used Tranus version 12.10.1.

Tranus has been utilized in a large number of studies, for regions of varying socioeconomic and cultural contexts, such as Latin America, USA, Europe and Japan.

Oregon Department of Transportation had used the Tranus to develop an integrated land use-transport model at the statewide level (Parsons Brinckerhoff Quade & Douglas, 1999). A research project promoted by the French National Research Agency (ANR) utilized Tranus platform to develop an integrated land use and transport model for the urban region of Grenoble, France (Feudo et al., 2017). The researchers concluded that they were able to achieve their goal of building a meaningful and policy-relevant model using the Tranus system. Wegener (2004) reviewed twenty contemporary urban land-use transport models and highlighted that the Tranus stands out as a particularly advanced and well-documented demand modeling platform with its attractive user interface. Center for International Intelligent Transportation Research of the Texas Transportation Institute (TTI) and Modelistica developed a binational travel demand model for El Paso, Texas, and Ciudad Juarez, Mexico, using Tranus to help transportation agencies of both regions anticipate traffic flow and needs (Collins et al., 2009).

The traffic assignment task in the Tranus platform, for this study, went through an exhaustive process utilizing an array of assumptions and hypotheses. The section also highlights the process and underlying algorithms that have been utilized for the dissertation purpose. The models were customized based on the characteristics and requirements of the study area.

4.5.1.6. **Processing actual bicycle volume counts**

This study obtained and analyzed two types of bicycle count data: the short-count (24-hour) data from the City of Austin Transportation Department, and continuous-count data from Eco-Counter (Eco-Counter, 2019). The continuous-count data were required to calculate adjustment factors to estimate the annual average daily bicycle (AADB) volume for the specific locations (Nordback et al, 2013) using the short count data.

As shown in Munira et al. (2021), the 24-hour bicycle count data were available for 44 locations. According to city officials, the locations were identified using the City of Austin bicycle route map (City of Austin, 2017) and based on the professional judgment of local planners. Standard procedures were followed for data collection where the data were collected using a video recorder in each of the intersections on typical weekdays distributed over 5 months (April, May, June, August, and October) in 2017. The permanent location count data, since 2012, were obtained from Eco-Counter for 11 locations. The continuous data were used to estimate the daily and monthly factors in order to calculate the AADB volume for each location where the short-count data were available.

As shown by Munira et al. (2021), with a minimum of 43 and a maximum of 1,282 riders, the intersections exhibited notable variation in terms of AADB volume,.

4.5.2. Model Building from Individual Sources

This section describes the demand models developed utilizing multiple sources. In doing so, the characteristics and features of different datasets are explained briefly.

While explaining the model building process, the general steps mentioned in the previous section are referred if/when necessary.

4.5.2.1. Direct Demand Model

Among several approaches to estimate and predict the demand of pedestrian and bicycle travel, the direct (facility) demand model is the most frequently used modeling approach in the area of pedestrian/bicyclist safety (Turner et al., 2017). The model utilizes actual count observations to estimate or forecast demand in a specific geographic area (such as intersection) by directly associating the observations with mode, trip, and traveler attributes, generally utilizing a form of regression (Ortuzar & Willumsen, 2011). This research utilized the actual AADB estimate (section 4.5.1.6), bikeability index (in section 4.5.1.3) and explanatory variables (outlined in section 4.5.1.2) to develop a direct demand model that can estimate AADB for all intersections of the study area.

As the characteristics of the dependent variables (Actual AADB at 44 intersections) was over-dispersed , a negative binomial model was utilized. The variable selection process is outlined in section 4.5.1.4. In addition to the bikeability index, the direct demand model, contained five demographic and land-use variables which are total population of age under 14, Black or African American population, population with no or some academic degree, bike signals, and presence of bicycle accessible bridge, as shown in Table 4.3.

Table 4.3: Negative Binomial Regression Model.

Variable (buffer width)	Estimates	T-Stat

(Intercept)	4.96	8.82
Bikeability index (1 mile)	0.02	1.65
Black or African American population (1 mile) (in 100s)	−0.02	−3.23
Population with no or some academic degree (0.1 mile)	0.03	2.98
Total population of age 14 and under (0.5 mile) (in 100s)	−0.12	−2.99
Bike signal (0.1 mile)	0.3	2.46
Presence of bicycle accessible bridge (0.1 mile)	0.54	2.25
Model Statistics		
N (sample size) : 44		
Adjusted R^2 : 0.7		
RMSE : 171		
MAE : 132		

The adjusted R^2 for the prediction model was 0.7, and RMSE was 171. Finally, the model was utilized to estimate AADB for all intersections of the study area. The predicted AADB ranged from 15 (mainly at the areas away from the downtown core) to 1,398 (at the downtown core).

The model also provided interesting insights into how different attributed influence bicycle demand within the Austin region. The AADB in intersections is characterized by sociodemographic (total population of age younger than 15, Black or African American population, and adult population of age 25 or older with no or some academic degree), bicycle infrastructure (bike signals), and built environment (bikeability index and presence of bicycle-accessible bridge) variables. The varying buffer scales of the variables emphasized the importance of using various buffers in developing the best model, being consistent with previous studies (Miranda-Moreno and Fernandes 2011). Moreover, while some variables exhibited the effectiveness of some of the policies taken by the city for improving safety and mobility of bike traffic, some variables provided insights into the scope of future improvements. For more details on the interpretation of the variables of the model, readers are referred to Munira et al. (2021).

4.5.2.2. **Bicycle Sharing Model**

Bicycle sharing data was obtained from the data portal of the BCycle-Austin bike-sharing agency. the bike-sharing system in Austin mainly covers the downtown region, with only 63 stations as of 2018 (BCycle, 2018). The research extracted bicycle trip data, from all stations, for 2017 when a total of 193492 trips were logged. The dataset contained limited information regarding each trip including trip start and end time and station.

Initial investigation on the temporal distribution of the trips revealed that the highest numbers of trips were observed during March through April, followed by September. This distribution might be attributed to the pleasant and mild weather during these months. The Texan summer, when temperatures climb into the mid to high 90s with high humidity, might be the reason for low bike-sharing trip activity from June through August. It might also be attributed to the fact that during the summer time, University of Texas Austin students are less likely to reside in the campus. Intuitively, Friday, Saturday, and Sunday observed a higher number of trips compared to other days of the week. Moreover, the majority of the trips were made by the users with a 24-hour pass (compared to annual pass, monthly pass, and 3-day pass users). This finding explains the high number of plausible recreational purpose trips that started and ended at the same station. Despite having a wide temporal coverage, the spatial coverage of the data was limited because the sharing stations mainly cover the downtown region. Trip starting and ending volumes were concentrated near the Lady Bird Lake (downtown) area. Moreover, the bike-sharing activity only accounted for a proportion of the total bike activity in that area, and thus needed population-level scaling.

The process of estimating annual average bicycle volume from this specific data went through three key stages: (a) examination and cleaning to obtain data in a meaningful format, (b) trip assignment, and (c) population-level scaling. In the first step, the distribution of the data, in terms of spatial unit, temporal unit, membership level, etc., was investigated. The step also included data cleaning; for example, trips starting and ending at the same station were removed because they would not add value to the assignment process (i.e., no trips would be assigned to the routes). Then, an OD matrix was developed using the station-level data. For this purpose, trips for the months April, May, and June, which represented the typical month of bike-sharing activity, were extracted. Then an OD table was built using the average trips for each station (OD) pair for the above-mentioned three months. For example, if there were six trips in three days (during the three months) for an OD pair, the average trip for that particular OD pair was taken as 2. Trips were rounded up for the decimal numbers. In the second step, a trip assignment process was conducted (as explained previously) to assign the trips at the intersection level. The output of this step was intersection-level bike-sharing volume. Finally, the estimate was scaled to population level using actual AADB (presented in section 4.5.1.6) and utilizing a negative binomial model.

4.5.2.3. **Strava Model**

Strava Metro is a data service that produces activity data from users of the Strava app. The app allows cyclists and runners to track their activities (such as rides, runs, and walks) on a smartphone or other GPS device. Strava allows access (in a subscription-based format) of the anonymized and aggregated activity data to the transportation agencies, city governments, and corporations. For this study, bike activity data were collected from Strava Metro through TxDOT.

The obtained dataset contains three subsets in three formats: streets, origin-destination, and nodes. To meet the objective of this study, node-level data (street intersections) were gathered. The researchers processed the total bicycle volume count for all nodes for the year 2017 to estimate the daily average estimate. Strava activity data for 2,303 intersections were processed. However, the data only represents a subpopulation. In order to scale the volume to the population level, the relationship between the actual AADB (presented in section 4.4.1.6) and Strava volume was built utilizing a negative binomial model, with the intersection density as an explanatory variable.

4.5.2.4. **Four-step Model**

This dissertation developed a four-step demand model to estimate intersection-level bicycle volume in the Austin area. The general steps of the models were: trip generation, trip distribution, modal split and trip assignment.

The first two steps, trip generation and trip distribution at the TAZ level, were performed using the traffic forecasting model tools from the local transportation planning agency, known as the Capital Area Metropolitan Planning Organization (CAMPO). The trip generation step computed the number of trip ends (daily) produced in and/or attracted to each TAZ of the study area, based on sociodemographic and land-use information, for multiple trip purposes (i.e., home-based work, non-home-based work, and non-home-based other). The step utilized TripCAL6 in TexPACK v3.0 beta version, which was created by TTI travel forecasting program for the demographic preparation and trip generation process. For the trip distribution step, the outputs from TripCAL6 in TexPACK v3.0 were converted to conform to the format of the CAMPO 2010 model. The CAMPO model uses a standard gravity model equation and applies friction factors to represent the effects of impedance (i.e., travel time, spatial separation) between zones. The output of the trip distribution step is an OD matrix that, for each trip purpose, indicates the travel flow between each pair of TAZs.

As the trip distribution output for multiple trip purposes, including all modes of transportation, it is necessary to develop a mode choice model to obtain an OD matrix for bicycle traffic. To do so, the general framework of the nonmotorized demand model proposed by the National Cooperative Highway Research Program (Kuzmyak et al., 2014) was followed. The main dataset utilized for this analysis was 2017 NHTS add-on data for the Austin region. To facilitate the model building process, TAZ-level socioeconomic and land-use variables were also used.

For the study area, 27,950 trips were extracted, of which 353 were bicycle trips. Given that the bicycle trips only occurred for a limited number of purposes (i.e., no bike trips for pick-up/drop-off purposes), only trips of purpose (home based non work, home based work and none home based) that had at least one bike trip were selected. Then, the TAZ location of each trip's origin and destination (based on coordinates) was identified. As the final step of data generation for the model building process, the land-use variables for each TAZ were matched to each trip's origin and destination TAZ. Thus, one trip was associated with two features of each land-use variable, denoted as origin land use and destination land use.

Three mode choice models were developed: home-based non-work trip model, home-based work trip model, and non-home-based work trip model. For all three trip purposes, the binary logit model was used to estimate an OD score (based on utility) based on the variables of distance between origin and destination (skim), origin land use, and destination land use. The main rationale was that people's decision to bike depends on the characteristics of both origin and destination. Table 4.4 presented the mode choice model results for three trip purposes.

Table 4.4: Binary logit model for three trip purposes

Variable	Home Based Non Work Model		None Home Based Model		Home Based Work Model	
	Estimate	Pr(>|z|)	Estimate	Pr(>|z|)	Estimate	Pr(>|z|)
(Intercept)	0.77	0.06	2.11	0.01	2.08	0.00
Skim (Distance)	-0.30	0.00	-0.18	0.11	-0.19	0.00
Origin-Freq of transit Stop	0.06	0.20	-	-	-	-
Origin-Mixed land Use	0.31	0.15	-	-	-1.65	0.17
Dest- Freq of transit Stop	0.06	0.18	-	-	-	-
Dest-Mixed land Use	0.50	0.13	-	-	-	-
Dest-High comfort bike facility	0.00	0.14	0.00	0.24		
Origin- Low comfort bike facility	-	-	0.00	0.20	0.00	0.07
Dest- Low comfort bike facility	-	-	-	-	0.00	0.30
Origin-Commercial land use	-	-	-0.06	0.30	-	-
Dest-Commercial land use	-	-	-0.09	0.10	-	-
Misclassification error	0.24		0.19		0.25	
ROC:	0.76		0.82		0.74	

The final model variables were then used to calculate the OD score for each TAZ pair and for each trip purpose. The computed OD score was categorized into several bins to estimate the rates of mode split for each bin by trip purpose, as outlined by Kuzmyak et al. (2014). The graph and equation developed from this step exhibited a clear pattern of higher rates of bike mode share for TAZ pairs of higher OD scores. The OD score and mode split relationship were then used to estimate bicycle trips for the study area. The output of the process was a trip distribution table, at the TAZ level, for bicycle traffic.

Since bike trips are generally short, TAZs that were more than 20 car minutes away from each other were removed from the trip distribution table.

Finally, Tranus was used for the trip assignment step to allocate the trips at the intersections of the study area. The model outcome was found to underestimate intersection-level volume when compared with the actual AADB. This underestimation was probably due to the model being developed using the demographic and trip characteristics of the 2010 CAMPO model. Thus, to scale the volume to 2017, actual AADB was used in a negative binomial model to estimate AADB from the four-step model.

4.5.2.5. **StreetLight Model**

StreetLight generates data representing walking and biking activity metrics that are derived from three main sources: general location-based services data, mode-tagged location-based services data, and validated bicycle and pedestrian counts (StreetLight, 2018). Additional sources, such as GPS-enabled travel diaries and traditional surveys about active mode behavior, are also used during the algorithmic development of the metrics.

The raw datasets go through a series of data processing steps to measure the active mode trips for an area. The platform utilized a probabilistic approach for mode inference (car, bike, walk) based on machine learning models and using multiple trip-related features. StreetLight (StreetLight, 2018) has noted that due to the relatively low sample, pedestrian and bike activity are not adjusted for population biases. To represent activity metrics, StreetLight generates index values that reflect a sample of bicycle trips starting in, passing through, or ending in defined zones.

This research gathered StreetLight Index data for bike traffic in terms of annual average daily volume in 2018 for 950 intersections of the study area. In order to scale the volume to the population level, the relationship between the actual AADB and StreetLight Index was built utilizing a negative binomial model with the population under age 14 (within a 1-mile buffer) as an explanatory variable.

4.5.3. **Discussion on the Model Outcomes**

Because the output from the five models was to be used as the input for fusion, it was imperative to discern the distribution, skewness, and coverage of the AADB estimates. The study area consisted of 2,518 intersections, both signalized and unsignalized. Overall, the AADB estimates from the models were found to be in the range from 0 to around 1,400. The intersections with high bike volume were generally concentrated in the downtown area. The estimates from the five demand models were deemed as reasonable and intuitive given that the actual count data also illustrated a high concentration (maximum of 1,282 riders) of bike ridership in downtown Austin. It is important to note that since the locations of counts were identified based on the city's bicycle route map, the counts are drawn from locations within the limits of the map and do not include areas farther into the suburban regions.

The DDM generated estimates for the maximum number of intersections (2,518), where the minimum volume was 15 and the maximum was 1,398. The bicycle-sharing model estimated bike activity in 793 intersections located near the bike-sharing stations in the central regions. The four-step model provided estimates for 2,397 intersections. Strava and StreetLight data were available for 2,303 and 950 intersections, respectively.

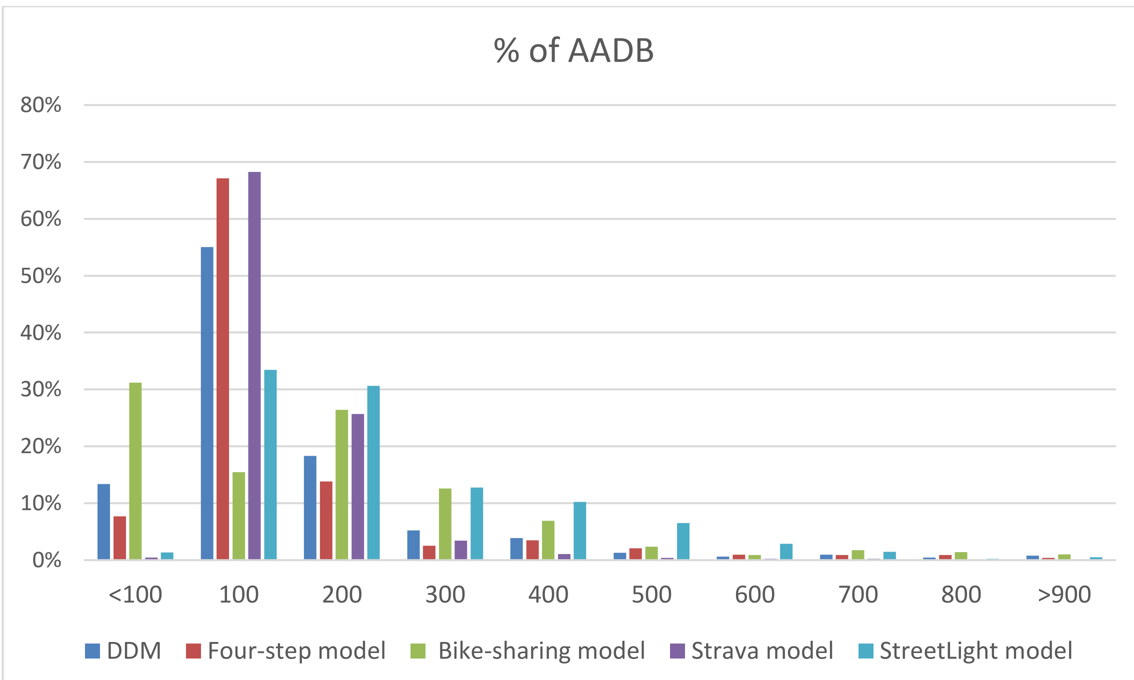


Figure 4.2: Frequency Distribution of AADB Estimates

Figure 4.2 presented the distribution of the estimated AADB from the five models. The figure showed that the volume distribution was right skewed. The majority of the estimates were found to be under 300 AADB. The mode of distribution for the bike-sharing model estimates was between 0 to 100 AADB. For the other four models, the peak was at the 100 to 200 AADB bins. Only a few observations were reported for AADB exceeding 800.

When compared with the actual AADB at 44 intersections, it also showed similar distribution. For the actual AADB, the mode of distribution was also between 100 to 200 AADB.

4.6. Context for Nonmotorized traffic data

As discussed in section 3.6, the process of discerning context for the traditional context-aware researches can be based on automatically acquired information or be done manually. Automatic detection requires a large-scale training database which is most often not an issue for researchers dealing with sensors (such as image sensor) data. However, in real-world applications, context sensing mostly relies on manual input (Dey and Abowd, 2001), which requires a profound understanding of the dynamics of the data and inherent situations.

The discussions in the previous chapters, along with this one, drew the implication that for nonmotorized demand data, the context in question can be derived from its spatial and temporal circumstances. For example, volume data obtained from bicycle sharing data is expected to be more reliable at the locations near the bike stations compared to intersections that are far from the stations. Similarly, the reliability of volume data may vary based on time of day, when data are available at finer temporal resolution.

Although this research didn't have adequate data to investigate the influence of temporal features as context, the spatial variability of reliability of the data sources was put forward for investigation. For example, the examination of the spatial variability of the deviation of the Strava data (as outlined in section 2.3.1.3) indicated the fact that the reliability of the volume estimates, obtained from different sources, are expected to have a spatial variation. Therefore, the spatial feature of the estimates, from each source, can serve as a "context" in the fusion algorithm.

5. MATHEMATICAL FRAMEWORK FOR FUSION OF NONMOTORIZED TRAFFIC DATA

5.1. Fusion Algorithms for Nonmotorized Traffic Data

The in-depth discussions on the characteristics and dynamics of the nonmotorized data sources and models shed light on multiple aspects that entail the choice and formulation of plausible fusion mechanisms for combining knowledge on nonmotorized demand. Indeed, the main objective of designing a fusion framework to combine multiple sources is to obtain an estimate which has better accuracy than the individual source. While there is no doubt on the need to develop a robust state-of-the-art statistical approach, it is also imperative to bring attention to practical constraints and limitations of nonmotorized information.

As mentioned in Chapter 3, a multitude of mechanisms for decision fusions have been applied in different areas of research based on multiple criteria and aspects of data. Among various elements and categories of fusion algorithms, the two key aspects deemed crucial to nonmotorized activity-related knowledge were

i) The type of output provided by the sources/models and

ii) The use of ground truth data.

Nonmotorized models are likely to generate knowledge or information as the abstract (crisp) level output. Moreover, given the resource requirements of gathering on-site nonmotorized activity data, which should serve as a ground truth or benchmark information for generating some fusion models, some practical scenarios may call for an option of fusion algorithms that don't warrant ground truth information. Therefore, considering the characteristics of the output of the nonmotorized activity information, availability of ground truth information, and sample size, this research adopted and illustrated multiple classifier/decision fusion mechanisms. Moreover, in addition to the traditional approaches, this research acknowledged the need for novel approaches accommodating the unique characteristics of nonmotorized demand models.

The fusion algorithms developed and applied for this research were categorized into types

i. Fusion without benchmark data

ii. Fusion with benchmark data

Under the first category, three traditional fusion algorithms, under the rationale of the voting method, were illustrated. Then a novel approach was proposed that uses the rationale behind the traditional weighted majority approach but generates decision weights without using the ground truth data

The fusion framework endorsed a state-of-the-art statistical approach—the Dempster Shafer (DST) method—which requires benchmark data. The method is a well-known generalization of the Bayesian framework, which can handle the uncertainty or ignorance of the data. The literature review had already outlined that the Dempster-

Shafer theory is particularly adequate for nonmotorized data fusion and application. Finally, in an endeavor to develop a robust statistical tool considering context, the Dempster Shafer with context credibility was proposed due to its ability to incorporate human subjectivity with mathematical probability to combine information (Wu, 2003)

The following sections of this chapter were devoted to illustrating the mathematical formalism of the algorithms utilized in this research.

5.2. Fusion Methodology without Benchmark Data

Fusion mechanisms under this category were described in two sections. The first section identified the traditional voting fusion approaches, applicable without ground truth data. The second section proposed a weighted majority voting approach that considers the pairwise interaction of the data to assign weightage on each decision and obtain a fused estimate.

5.2.1. Voting rules

The Majority voting is the simplest method to combine abstract-level classifiers. It doesn't require any prior training from the sources (Suen and Lam, 2000). Hence it can function without the ground truth data. In the simple majority voting method, the decision from each classifier or source is assigned an equal weight.

In general, there are three forms of voting rules (Raol, 2015; Kuncheva, 2014)

i. Unanimity: When combining the decisions of k sources or classifiers, a decision is made when all k sources agree on the decision/class label. Otherwise, the sample is rejected.

ii. Simple majority: When combining the decisions of k sources or classifiers, a decision is made when at least m of the sources agree on the decision/class label. Otherwise, the sample is rejected

$$m = \begin{cases} \frac{k}{2} + 1, if \text{ k } is\ even \\ \frac{k+1}{2}, if\ k\ is\ odd \end{cases} \quad \text{.............................(1)}$$

iii. Plurality: When combining the decisions of k sources or classifiers, a decision is made based on the most voted label.

Let us assume that class labels are obtained from each classifier or source. The decision of the kth classifier is defined $d_{k,j}$ € $0,1$, where k$= 1 \ldots . s$ (s classifiers) and $j = 1 \ldots . n$ (n class pattern). Each input x is classified into one of the n classes. $d_{k,j} = 1$ if k labels x as w_j, $d_{k,j} = 0$ otherwise. In this case, plurality will result in a decision for

w_j, if $\sum_{k=1}^{s} d_{k,j} = \max_{j=1}^{n} \sum_{k=1}^{s} d_{k,j}$ ....(2)

5.2.2. **Novel Weighted Majority Vote**

Weighted majority voting is another framework of combining decision-level data where each classifier is given a certain weight/importance in the decision-making.

Intuitively, the stronger classifier has more weightage in the decision compared to the weaker classifier. Typically, the weight corresponds to the accuracy of the classifier (Polikar, 2006); hence it requires ground truth data or a training set.

The label outputs can be represented as the degree of support for the class using the following equation

$$d_{k,j} = \begin{cases} 1, if\ D_k\ assignes\ x\ as\ w_j \\ 0,\ otherwise \end{cases} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots..(3)$$

If b_k is the weightage assigned to the classifier D_k, the class support for w_j can be obtained from

$$\mu_j(x) = \sum_{k=1}^{s} b_k\, d_{k,j} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots..(4)$$

The classifiers, in interest, would choose class w_m , given the following condition hold true

$$\sum_{k=1}^{s} b_k d_{k,m} = \max_{j=1}^{n} \sum_{k=1}^{s} b_k d_{k,j} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots..(5)$$

To estimate weights, this dissertation proposed to use a dissimilarity measure that can compute the pairwise interaction of the datasets. The underlying assumption was that if the decisions of two sources are far from each other, they are considered to be dissimilar. Therefore, if a data source has high dissimilarity with the other data sources, the piece of data is regarded as less important and will have a small weight in the final decision. To measure the pairwise dissimilarity, the concept of Euclidian distance was utilized. If x_r and y_r are the r^th observation of data source x and y, Euclidian distance between two data source is

$$\mathbf{d}(x,y) = \sqrt{\sum_{\mathrm{r}=1}^{n}(x_r - y_r)^2} \text{ ....(6)}$$

Here,

$$\mathbf{d_r}(x,y) = \begin{cases} 0, \text{if } x_r = y_r \\ 1, \text{if } x_r \neq y_r \end{cases} \text{..........................(7)}$$

Based on the pairwise distance measure, the weight of each source can be quantified based on its dissimilarity with the rest of the sources. For *k* information sources (k=1,…s), the dissimilarity and similarity of k^{th} source can be defined as

$$\mathrm{D}(k,s) = \frac{1}{\mathrm{s}-1}\sum_{j=1,\mathrm{k}\neq j}^{s} d(k,j) \text{ ....(8)}$$

$$Sim_k = 1/\mathrm{D}(k,\mathrm{s}) \text{..........................(9)}$$

Finally, the weight of k^{th} source, to be assigned in the decision process is

$$\mathrm{b}_k = Sim_k / \sum_{\mathrm{k}=1}^{S} Sim_k \text{..........................(10)}$$

5.3. Fusion Methodology with Benchmark Data

Given the characteristics of the nonmotorized activity data, which are often scarce, vague, incomplete, conflicting, or scattered, a robust fusion algorithm is needed to represent the knowledge and handle uncertainty. In this regard, Dempster Shafer (DST) had been deemed as a suitable mathematical framework that elicits decisions from multiple sources and measures uncertainty involving ambiguity (or ignorance) and conflict (Kronprasert, 2012), where traditional probability theory falls short. The following sections illustrated the principle, terminology and axioms of the mechanisms along with a brief literature review when deemed necessary. The section also described the proposed DST with context credibility.

5.3.1. **Dempster Shafer Combination: Overview**

Dempster Shafer method (DST), the core mathematical framework in this study, is a generalization of traditional probability theory. The DST framework utilizes belief measures to allocate degrees of support to one or multiple hypotheses instead of one mutually exclusive outcome as in the probability theory (Kronprasert, 2012). The theory can handle evidence with different levels of precision and ambiguity.

It should be mentioned here that The DST and Bayesian methods generate identical results when all the hypotheses are singletons and mutually exclusive (Wu, 2003). However, the framework for fusing nonmotorized data sources, when benchmark data is available, was built under the DST formulation, over Bayesian, for two reasons. First, the DST framework is applicable where information or decision of different granularity are needed to be dealt with at the same time. For example, two different sources or models may estimate a volume using different class boundaries, such as one as 100-200 AADB while the other data as 75-250 AADB. The second reason is the DST can accommodate imprecision and inaccuracy, i.e. missing knowledge, which is a common phenomenon for nonmotorized traffic data. The DST framework allows the modeling of the ignorance and missing information by generating explicit estimation of inaccuracy and the conflict between the information of various data sources (Chust et al., 2004)

5.3.2. Dempster Shafer: Basic Concept, Principles and Formulation

Outlining the basic elements and formulation of the DST, a brief literature review regarding the process of generating Basic Probability Assignments (BPA) was provided to illustrate how the BPA generation mechanism was adopted by this study. This section also discusses the need to incorporate discount factors in the DST framework.

5.3.2.1. Frame of discernment, Focal elements and Mass function

The frame of discernment (Ω) is the fundamental concept in the DST. A frame of discernment denotes a problem domain $\Omega = \{\omega_1, \omega_2, \ldots \omega_N\}$ which consists of an exhaustive set of mutually exclusive hypotheses. The power set of Ω is 2^{Ω} which is the set of all possible subsets of Ω (including the empty set $\emptyset$). That means $2^{\Omega} = \{\emptyset, \{\omega_1\}, \{\omega_2\}, ., \{\omega_N\}, \{\omega_1, \omega_2\}, \ldots, \{\omega_1, \omega_2 \ldots, \omega_N\}\}$.

Any subset of Ω may represents a proposition or hypothesis regarding the state of the object or system. The subsets containing only one element are called singletons. Here, each subset is a focal element where $\emptyset$ denotes empty set and the last subset $\{\omega_1, \omega_2 \ldots, \omega_N\}$ denotes complete ignorance as it doesn't provide any specific information.

In contrast to Bayesian theory, the DST framework assigns all missing evidence to ignorance.

In the DST framework, a notion of basic probability assignment (BPA), a function of mass function (m), is used which represents the degree of belief that supports a particular focal element. The range of a mass function is between 0 and 1 where 0 represents no belief and 1 represents a complete belief. The mass function has the following properties:

$m(\emptyset) = 0$ and $\sum_{A \in 2^{\Omega}} m(A) = 1$ ……………………………… *(11)*

Here, A denotes a focal element where $m(A) > 0$. Moreover, $m(A)$ denotes a degree of belief associated with the hypothesis that $'\omega \in A'$. The above equation corresponds to a closed world assumption, which means the frame of discernment is considered exhaustive. When evidence relevant to any focal element cannot be obtained, the remainder BPA is allocated to ignorance.

5.3.2.2. **Dempster's Rule of Belief Combination**

The DST fusion framework offers a mathematical architecture that allows each of the data sources to contribute to reaching a decision regarding the state of an object or incident. If m_1 and m_2 is two mass function obtained from two independent sources of information, the new mass function based on Dempster's combination rule can be denoted as $m = m_1 \oplus m_2$.

It can be represented as

$$m_1 \oplus m_2 (A) = \frac{1}{1-q} \sum_{B \cap C = A} m_1(B) m_2(C) \quad \text{..........................(12)}$$

$$m_1 \oplus m_2 (\emptyset) = 0 \quad \text{..........................(13)}$$

$$q = \sum_{B \cap C = \emptyset} m_1(B) m_2(C) \quad \text{..........................(14)}$$

Here, A, B & $C \in 2^{\Omega}$. q is the degree of conflict, alternatively called the degree of normalization or global conflict, between m_1 and m_2. Large q indicates more conflict between sources. The combination rule is valid only when $q < 1$.

Dempster's rule of combination is associative and commutative. That means the fused mass function does not depend on the order of aggregation. For s information sources, the combined mass function after the DST fusion process is

$$m_{combined} = m_1 \oplus m_2 \oplus \dots . m_s \quad \text{..........................(15)}$$

5.3.2.3. **Pignistic Probability Transformation**

The final step of the fusion is to make a decision regarding the label or class from the frame of discernment. The decision to reach a singleton can be made using various theories, including maximum belief (pessimistic strategy), maximum plausibility (optimistic strategy) or pignistic probability transform (Xu, 2014; Barnett, 1991; Denoeux, 1997).

Pignistic probability transform is a popular method that offers to transform a mass function m into a probability function for decision-making. The pignistic probability transformation of the final mass values was introduced in the concept of the Transferable belief model (Smets and Kennes, 1994). It transforms the mass function into a probability measure utilizing the principle of insufficient reason. It can be defined as

$$BetP_m(\omega_n) = \sum_{A \subseteq \Omega, \omega_n \in \boldsymbol{A}} \frac{m(A)}{|\mathrm{A}|} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots..(16)$$

Here |A| is the cardinality (number of elements) of focal element A. $BetP_m$ transfers the positive belief mass of each nonspecific element to the singletons in that element, based on the cardinal number of the proposition.

5.3.2.4. **Generating Basic Probability Assignments- A look at the literature**

The issue of generating the BPA has been a focus of research attention for decades as it plays a crucial role in the success of the DST (Parikh et al., 2001). One of the prominent methods to estimate BPA is utilizing information obtained from the confusion matrices. In a groundbreaking paper, published in 1992, Xu, Krzyzak and Suen combined multiple abstract level classifiers where BPA function for each of the K classifiers was defined using recognition rate, substitution rate and rejection rate (Xu et al., 1992). The methodology considered an additional class representing unknown classes or ignorance. However, in the method, each class of a classifier shared the same BPA, which couldn’t identify the source/classifier’s ability to recognize different classes.

Rogova (1994) applied the DST to combine fuzzy outputs where BPAs were generated for each class ‘n’ for each source ‘S’. The method was interesting as the construction of the first per-class-per-source BPAs utilized information regarding the recognition ability of each source across the classes. A similar approach was later utilized by Kuncheva et al. (2001), who proposed decision templates (DT) based fusion technique. In order to improve the approach proposed by Rogova (1994), Al-Ani and Deriche (2002) developed BPA based on the distance between the classification of an input vector and a reference vector, estimated as the mean classification output on training data for the class. Parikh et al., (2001) proposed to modify the approach proposed by Xu et al., (1992), by constructing BPA for each class, utilizing a precision rate estimated from the confusion matrix. Thiel et al., (2005) utilized DST concept to combine the normalized outputs of multiple source or classifiers where samples were rejected if the information were highly conflicting. Other methods of BPA generation include an approach based on expert knowledge about the domain of application (Milisavljevic and Bloch, 2003; Hégarat-Mascle et al., 2003) which can assign beliefs to compound hypotheses in addition to singletons.

In a relatively recent publication, Deng et al. (2016) proposed to construct BPA using both precision and recall rate for each class and demonstrated the effectiveness of the approach. A similar approach was later used by Schmitz et al. (2020) for land cover classification. This research also adapted the approach proposed by Deng et al.

5.3.2.5. **Discounting Approach in the Dempster–Shafer theory**

The traditional DST rule of combination has some limitations. Studies have shown that the theory poses the risk of leading to counterintuitive results while dealing with highly conflicting evidence. This limitation was first demonstrated by Zadeh (1984) and has been discussed by several studies, including Wilson (1993), Haenni (2005), Yager (1987) etc. The issue has been addressed from two different perspectives: i) preprocessing the evidence before combination, without changing the DST combination rule (Murphy, 2000; Yong et al., 2005; Martin et al., 2008; Chen et al., 2018) and ii) modifying the combined rule (Yager, 1987; Dubois and Prade, 1988). This research opted towards the first option to modify or discount the evidence before the DST combination process.

To manage conflict between different disagreeing sources within the DST framework, researchers have introduced the idea of discounting to reduce the strength of a mass function. As noted by Noble and Smith (2015) and Florea et al., (2010), two key factors influence the degree of trustworthiness and justify the discounting of a mass/belief function prior combination (1) either the data source is not found fully reliable; (2) or evidence from source is not fully credible. Both cases warrant for weakening the belief as argued by Shafer (1976) "*discounting at higher rates those belief functions one particularly distrusts and whose influence one wants to reduce*".

In the context of DST, the discounting factor based on reliability may reduce the strength of a mass function by allocating some masses to the ignorance state. The discounted mass function m_α can be written as

$m_\alpha(A) = \alpha \mathrm{m(A)}$ &

$$m_{\alpha}(\Omega) = 1 - \alpha\big(1 - \mathrm{m}(\Omega)\big),\ \mathrm{A} \subset \Omega \qquad \text{.............................}(17)$$

Here, $\alpha \in [0,1]$ is the discount factor of source S.

The issue of estimating discounting rates has been addressed by numerous researchers. Studies have computed discount rates based on reliability degrees of the source, calculated using ground truth information (Mercier et al., 2008; Smet, 1993). The discounting rate can also be based on credibility which can be computed, without any reference to ground truth data, based on a degree of disagreement between evidence (Noble and Smith, 2015; Florea et al., 2010). The rationale behind this estimation is the more one source is in disagreement with the others, the more the source is unreliable. Thus, the underlying assumption implies that a majority of experts are reliable (Martin et al., 2008).

Discounting rates can also be computed based on consensus between belief functions using evidence distance as a measure of the difference (Yong et al., 2004; Guo et al., 2006; Martin et al., 2008; Liu et al., 2017a). Several definitions of distance, in the evidence theory, have been proposed, such as Jousselme distance (Jousselme et al., 2001), cosine similarity (Wen et al., 2008), belief function distance metric (Sunberg and Rogers, 2013) etc. A more detailed discussion of the discounting rates (including dissimilarity measures between belief functions and membership degrees based on a distance measure) can be found in Florea et al., 2010, Chen et al., 2018.

This research adopted the Jousselme distance measure for estimating conflict.

5.3.3. **Dempster Shafer Combination: Adaptation**

The previous section laid out the background that lead to the adaptation of the DST framework for this dissertation. After illustrating the mathematical formulation of the BPA generation process, this section outlined the distance measure adopted to handle the conflicting evidence for this study

5.3.3.1. **BPA Generation**

The first key step of the DST approach is to estimate BPA functions from the different information sources. This research followed the approach proposed by Deng et al. (2016) to construct the BPA functions, recognizing its ability to capture source or classifier's performance based on the classes.

In this framework, let $m_i^{\varphi_k}$is the BPA function for each class ω_i $(i = 1 \dots N)$ in every data source φ_k $(k = 1 \dots S)$. The BPA for each class was estimated using recall and precision rates, obtained from the Confusion matrix C_φ, developed for each source.

The confusion matrix for each source can be represented as follows

$$C_\varphi = \begin{bmatrix} n_{11} & n_{12} & \cdots & n_{1N} & n_{1(N+1)} \\ n_{12} & n_{12} & \dots & n_{2N} & n_{2(N+1)} \\ \vdots & \dots & \ddots & \dots & \vdots \\ n_{N1} & n_{N2} & \cdots & n_{NN} & n_{N(N+1)} \end{bmatrix} \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots (18)$$

Here, n_{ij} denotes the number of samples belonging to class ω_i but identified as ω_j by source φ_k. The last column represents the rejection class that means not identified as any of the N classes. From the confusion matrix, precision rate e_{ij}^p and recall rate e_{ij}^r can be obtained from the following equations.

$e_{ij}^{r}=\frac{n_{ij}}{\sum_{j=1}^{N+1} n_{ij}}$ and $e_{ij}^{p}=\frac{n_{ij}}{\sum_{i=1}^{N+1} n_{ij}}$……………………………..(19)

Both e_{ij}^{r} and e_{ij}^{p} can be used to develop a corresponding Recall matrix C_{φ}^{r} and Precision matrix C_{φ}^{p} as follows

$$C_{\varphi}^{r}=\begin{bmatrix} e_{11}^{r} & e_{12}^{r} & \cdots & e_{1N}^{r} & e_{1(N+1)}^{r} \\ e_{21}^{r} & e_{22}^{r} & \dots & e_{2N}^{r} & e_{2(N+1)}^{r} \\ \vdots & \dots & \ddots & \dots & \vdots \\ e_{N1}^{r} & e_{N2}^{r} & \cdots & e_{NN}^{r} & e_{N(N+1)}^{r} \end{bmatrix} \text{ and } C_{\varphi}^{p}=\begin{bmatrix} e_{11}^{p} & e_{12}^{p} & \cdots & e_{1N}^{p} & e_{1(N+1)}^{p} \\ e_{21}^{p} & e_{22}^{p} & \dots & e_{2N}^{p} & e_{2(N+1)}^{p} \\ \vdots & \dots & \ddots & \dots & \vdots \\ e_{N1}^{p} & e_{N2}^{p} & \cdots & e_{NN}^{p} & e_{N(N+1)}^{p} \end{bmatrix} \ldots(20)$$

BPA m_{i}^{r} and m_{i}^{p} for each source φ_{k} can be derived using the following equations

$m_{i}^{r}(\{\omega_{i}\}) = \frac{e_{ii}^{r}}{\sum_{j=1}^{N} e_{ji}^{r}}$ and $m_{i}^{p}(\{\omega_{i}\}) = \frac{e_{ii}^{p}}{\sum_{j=1}^{N+1} e_{ij}^{r}}$ $where, \omega_{i} \epsilon\, \Omega$ ……………..(21)

$m_{i}^{r}(\Omega)$= 1-$m_{i}^{r}(\{\omega_{i}\})$ and $m_{i}^{p}(\Omega)$= 1-$m_{i}^{p}(\{\omega_{i}\})$ $where\ \omega_{i} \epsilon\, \Omega$……………..(22)

$m_{i}^{r}(A) = 0\ and\ m_{i}^{p}(A) = 0\ where\ \forall A \epsilon\ 2^{\Omega}\ \setminus \{\{\omega_{i}\}, \Omega\}$……………..(23)

The final BPA $m_{i}^{\varphi_{k}}$ for dataset φ_{k} and class ω_{i} can be obtained by combining both $m_{\varphi_{k},i}^{r}\ and\ m_{\varphi_{k},i}^{p}$

$m_{i}^{\varphi_{k}} = m_{\varphi_{k},i}^{r} \oplus m_{\varphi_{k},i}^{p}$……………..(24)

If a sample had been classified as a rejection class, the BPA is $m_{i}^{\varphi_{k}}(\Omega) = 1$.

This BPA generated from the training data can be utilized to fuse independent test data. For each sample in the test area, every data source allocates a class $\omega_{p}^{\varphi_{k}}$. According to the predictions, the corresponding BPA function $m_{i}^{\varphi_{k}}$ is selected. For the unclassified samples, $m^{\varphi_{k}}(\Omega)$=1. If $f\left(\varphi_{\mathrm{k}}, \omega_{p}^{\varphi_{k}}\right)$ represents the BPA $m_{i}^{\varphi_{k}}$ associated with the predicted class $\omega_p^{\varphi_k}$, for a sample, the combined BPA for the sample can be obtained from the Dempster Shafer's combination rule.

$$m_p^{combined} = f\left(\varphi_1, \omega_p^{\varphi_1}\right) \oplus f\left(\varphi_2, \omega_p^{\varphi_2}\right) \dots . \oplus f\left(\varphi_S, c_p^{\varphi_S}\right) \text{ ……………..(25)}$$

The last step is the final decision-making step where a final class is assigned to each sample based on the combined BPA, applying the pignistic transformation method.

5.3.3.2. **Conflict Measure Based on Distance**

To handle the conflict between sources, this research utilized the relative reliability measure, proposed by Martin et al., (2008), to define the distance between the BPAs. The underlying assumption is that if the evidence of two sources is far from each other, they are considered to be in conflict. Therefore, if a piece of evidence is highly conflicting with other evidence, this piece of evidence should be regarded as less important and will have a small influence on the final combination results.

The research utilized the Jousselme distance (Jousselme et al., 2001) measure in which the distance between two BPA m_1 and m_2 can be defined as

$$d_{12} = \sqrt{\frac{1}{2}\left(\overrightarrow{m_1} - \overrightarrow{m_2}\right)^T D \left(\overrightarrow{m_1} - \overrightarrow{m_2}\right)} \text{……………..(26)}$$

Here, m_1, m_2 are two BPAs under the frame of discernment Ω. The power set of the frame of discernment 2^Ω is regarded as a 2^N-linear space. D is a $2^N \times 2^N$ matrix, whose elements are

$$D(A, B) = \frac{|A \cap B|}{|A \cup B|} \; \forall A, B \in 2^\Omega.$$

Therefore, the conflict measure between two sources can be defined by

$Conf(1,2) = d_{12}$……………..(27)

The weight on each source can be quantified based on its conflict with the rest of the sources. For k information sources (k=1…S), the conflict measure between k^{th} and the other S-1 sources can be defined as

$Conf(k, \mathrm{S}) = \frac{1}{S-1}\sum_{j=1,k\neq j}^{S} Conf(k,j)$ ……………..(28)

Therefore, the reliability of each source can be estimated as

$\alpha_k = f(Conf(k, \mathrm{S}))$……………..(28)

Here, f denotes a decreasing function.

$\alpha_k = \left(1 - Conf(k,S)^{\lambda}\right)^{\frac{1}{\lambda}}$, where λ > 0. ……………..(29)

Here, λ is the degree of conflict discount. The conflict measure can be used to adjust the belief assignment before the DST combination, as illustrated in equation 17.

5.3.4. **Dempster Shafer with Context Credibility**

Acknowledging the potential of incorporating context to improve fusion (as discussed in section 3.6) coupled with the fact that the nonmotorized activity exhibits contextual properties (section 2.3.1), this research proposed an approach within the DST framework where the belief function of each source can be modified based on the contextual situation of an entity or object. The DST framework was particularly chosen for this purpose, given its ability to synthesize human subjectivity with mathematical probability (Wu, 2003).

While the relative reliability or conflict-based discount computed the disagreement among the evidence, the proposed method estimated a discounting factor based on contextual situation (such as spatial location, temporal scale etc.). This was deemed particularly important for the fusion of nonmotorized data, which exhibited significant variation based on spatial location. Moreover, as previously discussed, the deviation of the source estimate from the actual observation also varied with space, probably due to differences in demographic characteristics and trip purpose distribution in different areas (Munira and Sener, 2020). That means the reliability of the volume estimates, obtained from various sources, may exhibit a spatial variation. For example, a volume estimate from bicycle sharing data may be more reliable on spatial zones near the kiosk stations. Identification of such context would require subjective judgment based on knowledge of the local conditions.

In the proposed mechanism, the discounting measure based on contextual situation (such as location) was estimated based on the Grey theory. Grey theory allows exploring the intrinsic information of the system without requiring a specific relationship as an assumption (Song et al. 2005). This study had particularly chosen the theory as it is an effective approach to analyze the relationship between cases with small samples and inadequate information, overcoming the limitation of the traditional statistical methods

(Chang et al., 2003; Liu and Forrest, 2010). The theory was also used to measure sensor credibility (Ye et al., 2016), Rotor's fault detection (Zhang et al., 2012) etc. The proposed framework consisting of DST with context credibility, based on Grey theory, was illustrated in figure 5.1.

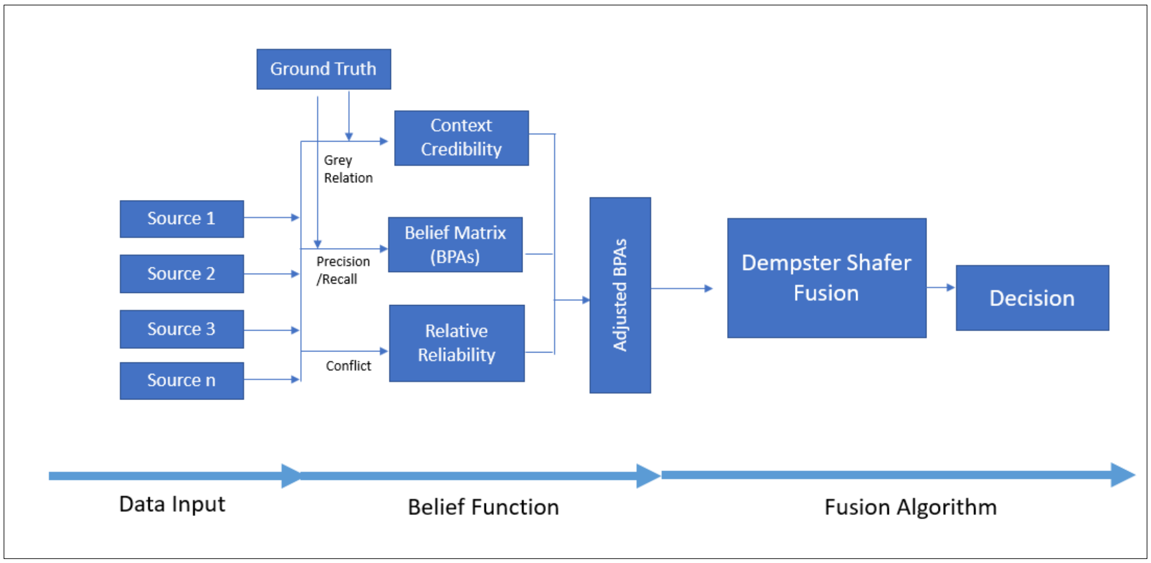


Figure 5.1: Proposed Framework of DST with Context Credibility

Grey correlation analysis (GCA) is one of the main components of the Grey System Theory proposed by Deng (1987). The measure can quantify the correlation between the reference sequence and the comparison sequence (Yunlong et al., 2019).

Suppose that the x_o= $[x_{o(1),}x_{o(2),} \dots x_{o(n)}]$ is the reference sequence and x_k =$[x_{1(1),}x_{1(2),} \dots x_{1(n)}]$ is the comparison sequence, $k = 1 \dots S$.

$\gamma\big(x_o(r), x_k(r)\big)$ denotes the comparison measurement of x_o and x_k at the r^th (r = 1….n) point in the grey relation factor space. If $\gamma(x_o, x_k)$ is the average value of $\gamma\big(x_o(r), x_k(r)\big)$ at all points, the relationship can be represented as

$$\gamma(x_o, x_k) = \frac{1}{n}\sum_{r=1}^{n} \gamma\big(x_o(r), x_k(r)\big) \quad \text{……………..(30)}$$

$$\text{Where, } \gamma\big(x_o(r), x_k(r)\big) = \frac{\min_k \min_r \Delta_{ok}(r) + \zeta \max_k \max_r \Delta_{ok}(r)}{\Delta_{ok}(r) + \zeta \max_k \max_r \Delta_{ok}(r)} \quad \text{………….….(31)}$$

$\zeta \in [0,1]$ is the distinguishing coefficient. Usually, ζ =0.5 is used as it has moderate discrimination and good stability (Yunlong et al., 2019)

$$\Delta_{ok}(r) = \begin{cases} |\ x_o(r) - x_k(r)\ |, \text{ when } x_o(r) \text{ and } x_k(r) \text{ are numerical} \\ 1, \text{when } x_o(r) \text{ and } x_k(r) \text{ are categorical and } x_o(r) \neq x_k(r) \\ 0, \text{when } x_o(r) \text{ and } x_k(r) \text{ are categorical and } x_o(r) = x_k(r) \end{cases} \quad \text{…………(32)}$$

Based on the GCA, the proposed spatial correction σ_k, for k^th source is

$$\sigma_k = \gamma(x_o, x_k)^{\beta}, \text{ where } \beta > 0. \quad \text{……………..(33)}$$

Here, β is the degree of discount due to context credibility. Therefore, the new discount function and adjusted belief function are

$\Theta_k = \sigma_k * \alpha_k$, $\Theta_k > 0$

$m_{adj}(A) = \Theta_k m(A),\ where\ \forall A \epsilon\ 2^{\Omega} \setminus \{\Omega\}$……………..(34)

$m_{adj}(\Omega) = 1 - \Theta_k(1 - m(\Omega))$ ……………..(35)

5.4. Summary

The main objective of the chapter was to design a fusion framework for consolidating nonmotorized data. The framework was devised keeping both statistical robustness and practical application in mind. The voting approach, despite its simplicity, is expexted to add great value for the fusion of multiple sources, especially when adequate ground truth data is not available. The fusion framework also endorsed a state-of-the-art statistical approach—the Dempster Shafer (DST) method—which requires benchmark data but is adequate to accommodate the often incomplete and conflicting nonmotorized demand output with a mechanism of handling uncertainty involving ambiguity (or ignorance) and conflict. Moreover, the basic DST was reinforced, incorporating a robust belief assignment method based on both precision and recall matrix and a conflict discounting mechanism based on Jousselme distance. Besides, the DST fusion approach with credibility context, as proposed by this study, offered a unique way to incorporate the subjective judgment of experts in mathematical fusion formulation. The novel DST approach that considers contexts is also expected to add value to other areas of research, in addition to nonmotorized traffic, especially when the sample size is a constraint and the analyst's judgment senses the potential variability of reliability of the sources across one or multiple contexts of an entity or object.

While theoretically, both the traditional and novel approaches were deemed well-founded and promising, it is imperative to observe the performance of the approaches when they are applied to real-world data. The next Chapter consequently focused on the numerical outcomes of the application.

6. EXPERIMENTATIONS AND EVALUATIONS OF FUSION FRAMEWORK

Based on the formalism and interpretations of the fusion algorithms, illustrated in the previous chapter, Chapter 6 intended to elucidate the application of the fusion framework to address the critical question of when or under what scenario the fusion methods are suitable for the application. In explaining the application and efficaciousness of the framework, the section utilized both actual and artificially generated data to exhibit the performance of different algorithms. The findings from the numerical application were expected to furnish a more intuitive explanation regarding the potential and limitations of the fusion algorithms.

Application of the fusion algorithms necessitates an experimental setup with mechanisms for demonstrating the validation process. The application and validation process was divided into two parts. First, actual bike volume estimates from the five bike demand models, as described in chapter 4, were utilized for fusion application. Due to the unavailability of adequate ground truth data, voting algorithms were applied for fusing the estimates. However, for explaining the application and interpretation of the voting algorithms, the actual volume information (from 44 locations) was used.

For demonstrating the efficacy of the DST framework, the size of the ground truth data for the study area was not deemed adequate. However, the framework was expected to be useful for areas where adequate ground truth data is available. Hence, the research generated artificial datasets, conforming to the actual data condition and context as much as possible, to evaluate the performance of the DST mechanism. The purpose was to empirically test the performance of the DST approaches, both traditional and novel, by applying them on a set of artificially generated datasets with different intrinsic characteristics. The goal was to gain an understanding of which strategy, and under what condition, performs the best in order to bolster the conclusion of this research. The artificial datasets also allowed for the creation of multiple scenarios to examine the range of outcomes and understand the potential factors that contribute to the performance of the algorithm.

This chapter is divided into three parts. First, the validation process to demonstrate the effectiveness of the algorithms was explained. The second and third sections illustrated the findings of the numerical experiment using actual (voting fusion) and simulated data (DST fusion), respectively. A chapter summary encapsulated the key lessons learned from the analysis.

6.1. Experimental Setup

The ultimate goal of any fusion endeavor is to produce estimates that are superior to the individual source estimates. The key operative word here is ‘superior,’ which may entail completeness, accuracy, or a mix of both. While the completeness can be evaluated by comparing the coverages (for example spatial coverage), the superiority of the accuracy can be demonstrated by showing that the accuracy of the fused estimate is higher than the best individual source.

The first step of the experimental setup was to select a performance measure to evaluate the performance of the modeling framework. Although there are several measures for assessing the performance of classification algorithms (Reich and Barai, 1999), the focus of this research was given to the evaluation of the overall classification accuracy, meaning the number of correctly predicted samples in a test fold. In doing so, an exhaustive set of mutually exclusive hypotheses, or a frame of discernment for DST, had to be created with the AADB estimates. While the DST framework can accommodate different class labels from different sources, this research used the same class label (categorization) for all sources.

While the overall accuracy is the performance measure, this research adopted k-fold cross-validation for validating the DST framework. Consulting previous studies (Kohavi, 1995; Rodriguez et al., 2009), k=5 was used to maintain a trade-off between bias and variance while minimizing computational effort. To compare the performances, the partitioning of every data source was kept exactly the same for all of the sources and fusion estimates. The final accuracy of the fused algorithm was computed by averaging the estimated accuracy of each fold.

6.2. Application of Voting Fusion on Actual Data

The application of the voting fusion algorithms was demonstrated utilizing the outputs from five bike demand models. The first task in the process was to categorize or assign class labels to the estimates. Four class labels, reflecting the distribution of the datasets (Figure 4.2) and ensuring adequate observation in each of the categories, were assigned to the model outputs: < 100 AADB, 100 to 250 AADB, 251 to 400 AADB, > 400 AADB.

Table 6.1 presents the deviation of the individual model outputs from the actual bicycle volume count data (termed as accuracy) along with their coverage. The direct demand model exhibited the highest accuracy, while the bike-sharing model had the lowest accuracy for the study area.

Table 6.1 Accuracy and Spatial Coverage of the Sources

Model	Number of Intersection	Overall Accuracy (%)
Direct demand Model	2518	0.59
Strava Model	2303	0.50
Bike-Sharing Model	793	0.27
Fourstep Model	2397	0.41
StreetLight Model	950	0.52

Four voting fusion algorithoms, as outlined in Chapter 5, were applied to combine the bike estimates. Intuitively, a fusion algorithm can be considered effective when the accuracy is higher than that of the direct demand model (best individual source). In addition to the accuracy, it is also essential to examine the detection rate (or coverage) of each fusion method because the goal is to obtain high-accuracy estimates for the maximum number of intersections. Thus, it is essential to evaluate both the detection accuracy and overall accuracy of the algorithms.

Figure 6.1 presents the detection rate and accuracy among the detected AADBs of each voting fusion approach. For evaluating the latter, the actual volume counts were utilized.

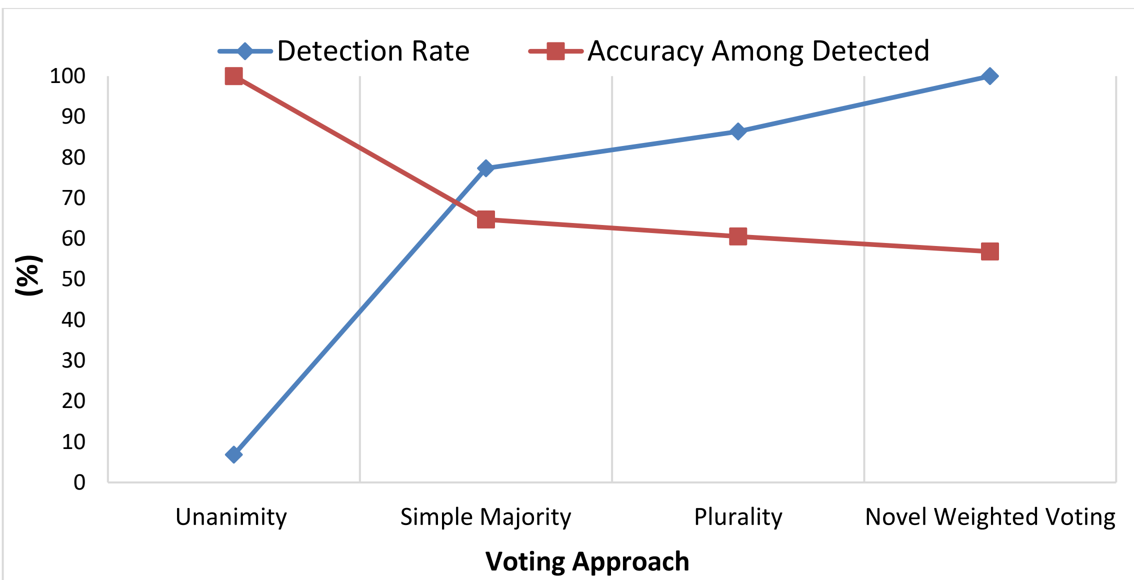


Figure 6.1 Detection Rate and Accuracy (Among Detected) of the Voting Approaches

The results indicated that the unanimity voting approach had the lowest detection rate (for the entire study area) yet the highest accuracy among the detected estimates. This finding is intuitive because to reach a decision for unanimity, each of the sources has to agree. The trade-off between detection rate and accuracy for the first four voting approaches is also made apparent in the figure. The novel weighted voting approach allocated AADB at all the intersections of the study area with a detection accuracy of 57%. Therefore, the evaluation suggests that the novel weighted approach exhibits decent performance considering the detection rate and the detected accuracy. However, if the analyst prefers accuracy over coverage, the unanimity and simple majority voting approaches may be regarded as potential options.

Based on evaluation of the overall performance of the approaches and comparison of each with the best individual source (as shown in **Error! Reference source not found.**6.2), it was observed that the novel weighted voting approach exhibited slightly lower accuracy compared to the best individual model—the direct demand model.

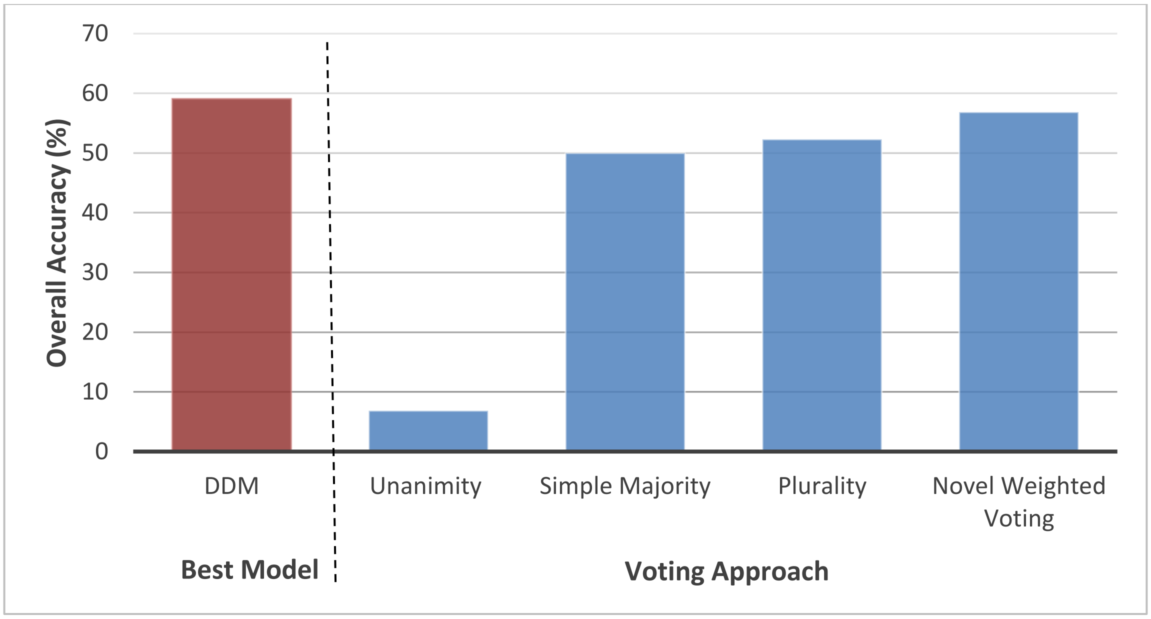


Figure 6.2: Overall Accuracy of the Voting Fusion Approaches

Another weighted voting fusion algorithm was run with four data sources, removing the StreetLight data. It was observed that the novel weighted voting approach with four data sources (direct demand model, Strava model, bike-sharing model & fourstep model) exhibited almost the same accuracy (0.59%) as the best individual model.

In light of the limited sample of the validation data for the current case study, it can be noted that the weighted voting fusion algorithms may not result in a better outcome than the best individual source. However, when rightly adapted, it may at least exhibit similar accuracy compared to the best individual model/source outcome.

The weighted voting approach was developed to apply when the analyst does not have any knowledge of the accuracy of the individual sources. It was expected that the application of the novel weighted method, although with no significant increase in accuracy, may instill confidence in the estimate, which is another objective of the fusion endeavor. The approach also had the potential to provide better coverage of the estimates when the individual sources have a significant number of missing cases. However, the findings indicated that the weighted voting approach actually poses the risk of obtaining worse off results compared to individual sources. It was also noted that adding the maximum number of sources may not add value to the process. A minimal understanding of the individual sources is desired to adopt an effective voting fusion approach.

6.3. Application of DST Fusion on Artificial Data

The purpose of this experiment was to empirically test the performance of the DST approaches applying them on a set of artificially generated datasets of different characteristics. As the DST approach requires adequate ground truth data, which wasn't available for the study area, meticulous attention has been given to ensure that the artificial data conform to the real-world condition. In this section, the experimental procedures, including data and scenario generation, were explained before discussing the findings from the numerical analysis.

6.3.1. Design of Simulated Experiments

In order to ensure that the demonstration of the effectiveness of the fusion framework is objective and convincing, this dissertation followed a systematic approach to design the simulated experiment.

For doing the experiment, a ground truth dataset and multiple other datasets which have some degree of relationship with the ground truth data were needed. While the ground truth dataset was required to train and validate the model, the other data sets served as input sources for fusion, as in DDM or four-step model output for nonmotorized traffic. While creating the datasets, the following criteria were deemed essential for consideration.

i. The data should conform to the actual nonmotorized model outputs for the study area. However, the artificially generated data should have a larger, yet realistic range of nonmotorized volume

ii. To demonstrate the effectiveness of the contextual discount, proposed by this study, the artificial dataset should exhibit a contextual difference in reliability, in some form, which conforms to the actual scenario

In order to create the ground truth data, representing a realistic nonmotorized volume, the relationship between the bikeability index and bike volume gleaned from the direct demand model, was taken as a reference point. First, a random set of 2500 data points (as in locations or intersections) of bikeability index were created which was of uniform distribution. The random bikeability index values were run through the equation

E[Y] =exp(0.14 * bikeability index). Finally, the Ground Truth was simulated from each Poisson (E[Y]) variable.

The other datasets, denoted as the individual sources, were created using Normal distribution noise with the ground truth data. The key equation for building the source-datasets was

$$Source\ K1 = Groundtruth + Y,\ Where, Y \sim N(\mu, \sigma^2) \quad \text{………………..6.1}$$

Here, μ denotes mean and σ denotes standard deviation

As the normal distribution noise results in both positive and negative values, the negative estimates were assigned as missing values. Additional missing values were also imputed in the datasets, randomly, for building the missing case scenarios.

While creating the datasets, careful attention was given to instill contextual accuracy differences in the sources. In doing so, the reliability variation of actual model estimates (as shown in Table 6.2), based on their location as of downtown (DT) or out of downtown, were consulted.

Table 6.2: Overall Accuracy of the models within and outside Downtown (DT)

Model	in DT	Outside of DT
Direct demand Model	0.6	0.59
Strava Model	0.6	0.47
Bike-Sharing Model	0.6	0.17
Fourstep model	0.5	0.38
StreetLight	0.6	0.5

The table showed that some of the models have a notable variation in terms of accuracy based on the locations. Hence, the location of the intersection could serve as a context for the actual models. In order to replicate the real data, each of the artificial datasets was divided into two sets, context 1 and 2, where the accuracy varied. An example is shown in Table 6.3, where for each scenario, the accuracy varied across two contexts.

Table 6.3: Example of Context Variation of the Artificial Datasets

Source	Context 1 accuracy	Context 2 accuracy
Source 1	0.75	0.63
Source 2	0.63	0.56
Source 3	0.52	0.55

6.3.2. **Scenario Design**

Given that it is possible to generate an infinite number of simulated data sets to validate an experiment, it is often easy to lose sight of essential questions that need be addressed. This research formulated four distinct questions for the evaluation process. These questions were relevant to nonmotorized data characteristics and local situations to which the simulated scenarios were expected to respond:

i. Does the fusion algorithm always provide a better estimate?

ii. Is the more (number of sources) always better (for fusion estimates)?

iii. Is the fusion estimate sensitive to degree of conflict and context credibility discount?

iv. Is the fusion estimate sensitive to categorization or class labels?

To address these questions, three main scenarios, each containing two sub-scenarios, were created. Each scenario was developed with three datasets of varying accuracy and missing case situations.

The three key scenarios were

Scenarios 1: No missing observations, fairly similar accuracy across the sources

Scenarios 2: No missing observations, a varying level of accuracy across the sources

Scenarios 3: Each of the sources has missing observations, a varying level of accuracy across the sources

For each sub-scenarios, models were built for two categorizations or class labels, mainly to examine if the algorithms are sensitive to categorization type.

Two categorizations denoted as trial 1 and 2 are

Trial 1: 4 categories

<100", 100-250, 251-400, >400

Trial 2: 12 categories

< 50, 50–100, 101–200, 201–300, 301–400, 401–500, 501–600, 601–800, 801–1,000, 1,001–1,200, 1,201–1,400, > 1,400.

The scenario also assessed the sensitivity of the DST fusion based on the degree of conflict (λ) and context credibility (β). The lower the value of λ, the more heavily the mass function is discounted due to conflict. On the other hand, higher β indicates greater discounting of the mass function. The value of λ and $\beta > 0$ mean discounts are applied based on conflict and context credibility.

Finally, to address whether the more (number of sources) is always better, the lowest accuracy source was removed for each of the six scenarios (three key and two sub scenarios) and the model performance was evaluated.

6.3.3. Scenario Analysis and Comparison

This section evaluated the performance of the DST when applied on the three scenarios explained in the previous section. As noted before, the objective of fusion is to obtain a superior estimate (better accuracy and/or coverage) than the best individual source. This research presented the analysis outcome in terms of change in accuracy, representing the difference of accuracy between the fusion estimate and the best individual source. Hence, a positive change in accuracy indicates that the fusion estimate is better than the best individual source estimate. For each scenario, varying values (0, 0.5, 1) of λ and β were tested for the two categorizations or trials. While illustrating the relationship between λ and change in accuracy, the value of β was kept as 0 and vice versa.

Before delving into the specifics of the fusion outcomes, a descriptive statistic of the scenarios was presented in Table 6.4, which reported the coverage, in terms of complete cases, and accuracy of the individual sources for both categories. Here, the complete case indicated the number of data points (as in intersection for the nonmotorized model) for which all three sources have estimates. As explained in the previous section, the first two scenarios had no missing cases.

Table 6.4: Scenario Description and Accuracy of Individual Sources

Scenario	Complete cases	Source 1		Source 2		Source 3	
		Cat 1	Cat 2	Cat 1	Cat 2	Cat 1	Cat 2
1	2500	0.69	0.43	0.70	0.47	0.71*	0.48*
2	2500	0.57	0.32	0.68	0.43	0.78*	0.55*
3	1974	0.54	0.29	0.60	0.40	0.70*	0.50*

*indicates the best source to compute change in accuracy (for each categorization)

Table 6.4 showed that the accuracy of the sources varied with the categorization type. Overall, trial 1 (4 labels) has higher accuracy than trial 2 (12 labels) for each source. The change in accuracy of the fusion method was computed comparing with the best individual source, marked with asterisks, for each category.

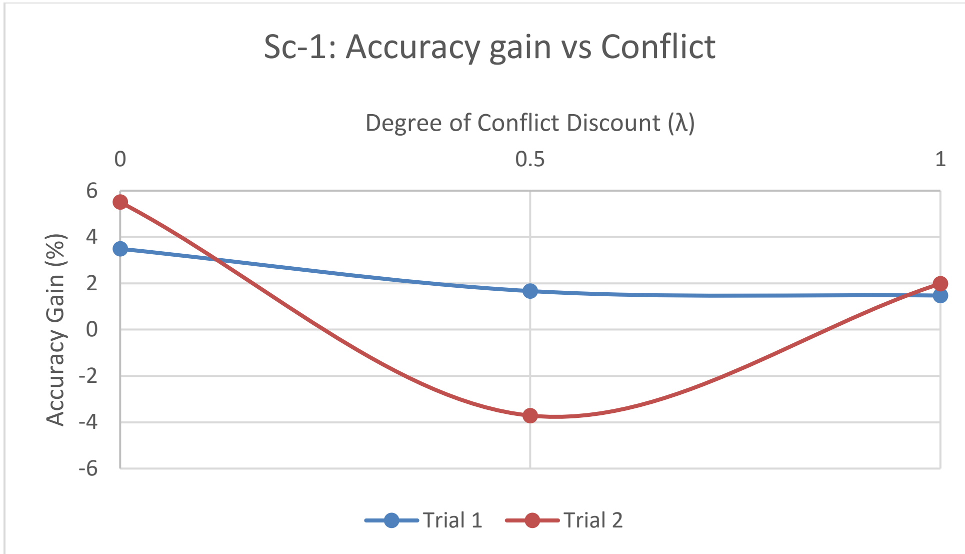
6.3.3.1. Scenario 1 Analysis

Scenario 1 was built with no missing observations where the sources have fairly similar accuracy. It intended to explore whether fusion of similar sources, without any missing cases, may add computational value to the system. Figure 6.3 reported the analysis outcome for scenario 1.

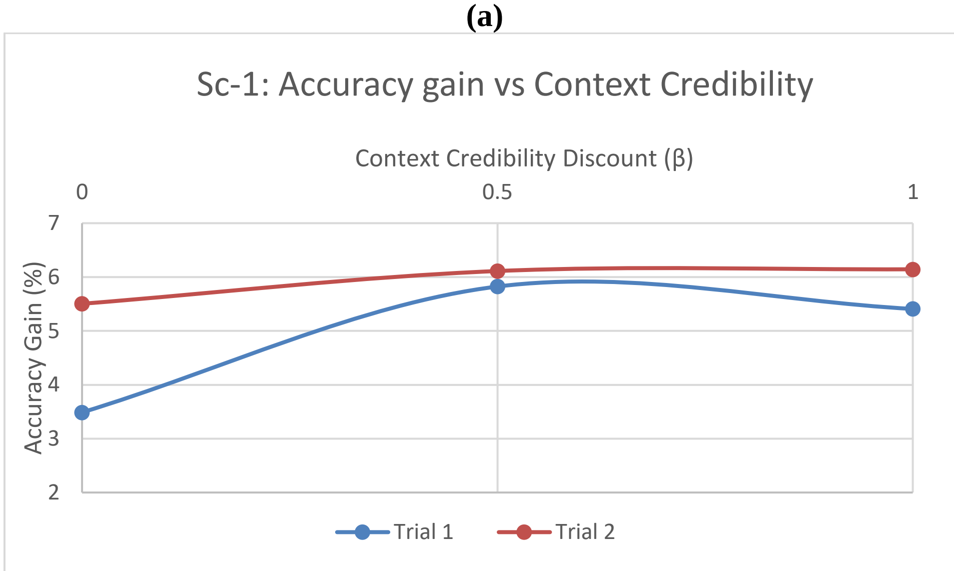
As shown in Figure 6.3, a positive change in accuracy, meaning the fusion estimate is better than the best individual source, was observed in most cases. However, the change in accuracy varied with both categorization type and values of λ and β. Even the optimal value of λ and β varied with the categorization.

For trial 1, the optimal value of λ and β were 0 and 0.5, respectively, and for trial 2, the optimal values were 0 and 1, respectively. While trial 1 exhibited positive change in accuracy across the values of λ and β, label 2 showed accuracy loss (fusion is not better than the best individual estimate) when λ was 0.5 (for β=0). Overall, the positive change in accuracy indicated that even for no missing case scenarios, the fusion algorithm may prove to be of great value to obtain better estimates, especially when the individual sources are of similar accuracy. The optimal value of β>0 asserted the superiority of the application of the contextual discount, as proposed by the study.

Across all (six) tested values of λ and β, the highest change in accuracy for trial 1 and trial 2 were 5.8% and 6.1% respectively. The average change in accuracy for trial 1 and trial 2 was 3.55% and 3.58% respectively.



(a)

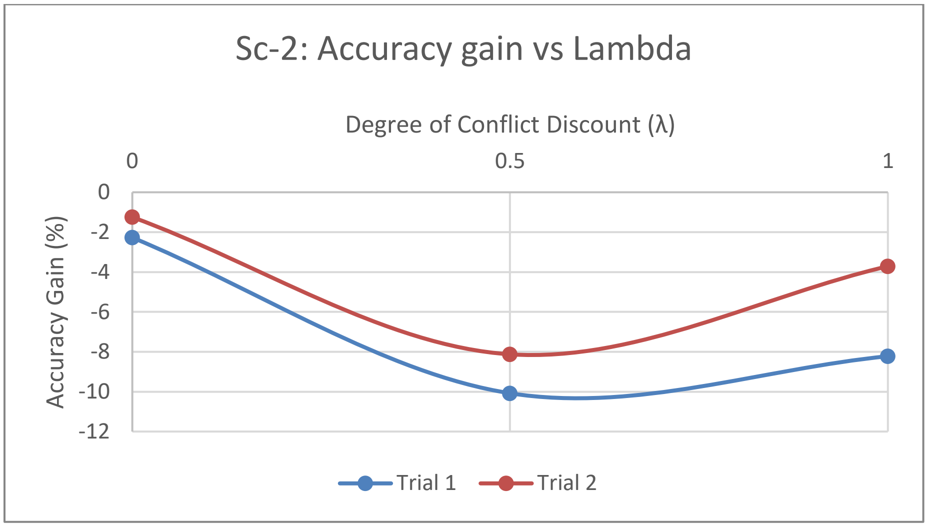


(b)

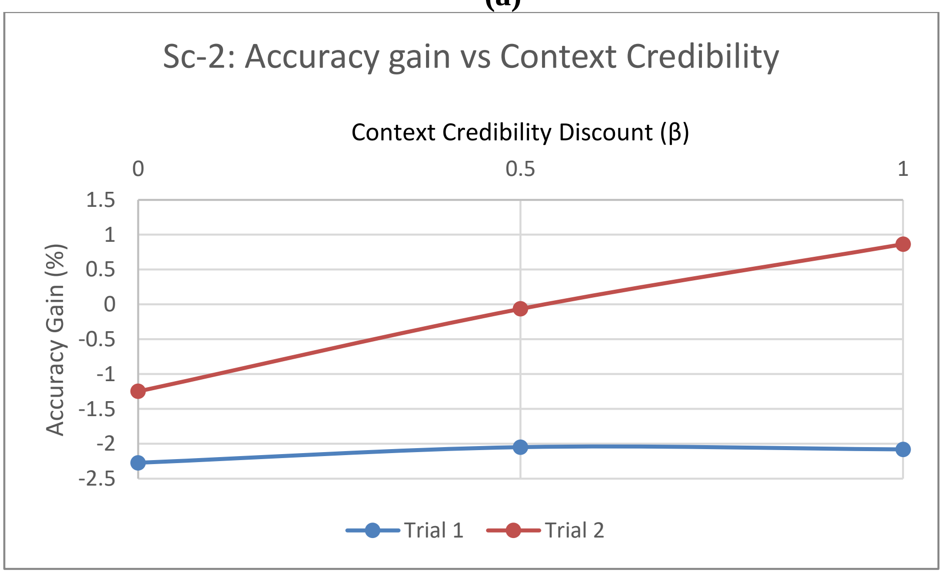
Figure 6.3: Change in accuracy based on conflict (a) and context credibility (b) for Scenario 1

6.3.3.2. **Scenario 2 Analysis**

Scenario 2 was built with no missing observations, similar to scenario 1. However, the accuracy of the sources, in this scenario, varied notably. Figure 6.4 reported the analysis outcome for scenario 2.



(a)



(b)

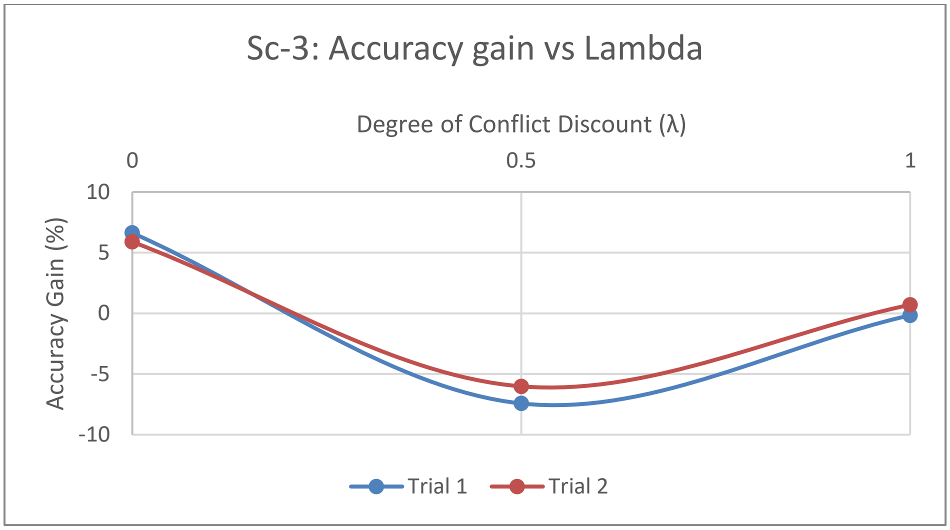
Figure 6.4: Change in accuracy based on conflict (a) and context credibility (b) for Scenario 2

Assessment of Figure 6.4 illustrated that for label 1, the fusion algorithm couldn't yield positive change in accuracy across the tested values of λ and β, meaning fusion effort was not effective. However, for label 2, a positive gain was observed for the values of β=1 and λ=0. This indicated that even for the same dataset, the fusion algorithm might be effective for one categorization or label but not for the other one. It also indicated that fusing low and high accuracy sources may not yield the desired outcome for no missing case scenarios.

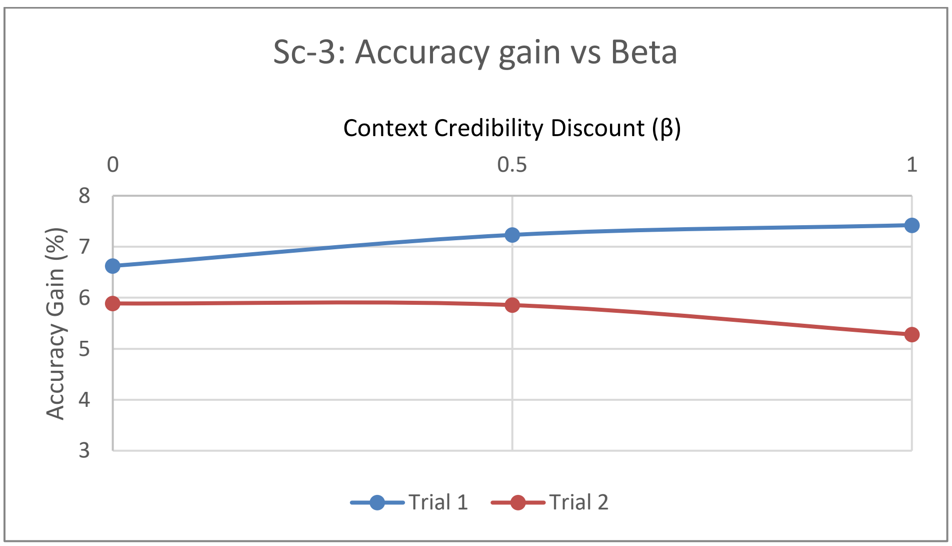
For this scenario, the average accuracy loss across all (six) tested values of λ and β, for trial 1 and trial 2 were -4.49% and -2.25% respectively. The lowest change in accuracy (loss) was -10% for trial 1.

6.3.3.3. **Scenario 3 Analysis**

For Scenarios 3, each of the sources had missing observations and their accuracy also varied notably. Figure 6.5 reported the analysis outcome for scenario 3.



(a)



(b)

Figure 6.5: Change in accuracy based on conflict (a) and context credibility (b) for Scenario 3

As reported in Figure 6.5, the application of the fusion algorithm was fruitful for both categorizations across most of the values of λ and β. Overall, the change in accuracy was higher compared to scenario 1 and 2. For label 1, the optimal value of λ and β were 0 and 1, respectively, and for label 2, the optimal values were 0 for both cases.

Overall, across all (six) tested values of λ and β, the highest change in accuracy for trial 1 and trial 2 were 7.42% and 5.89% respectively. The average change in accuracy for trial 1 and trial 2 was 3.38% and 2.93% respectively.

The result indicated that when there is a significant number of missing observations in each source, a fusion endeavor may complement each other (sources), especially by providing a decision for a case where an estimate is not available from all the sources.

Analysis from three scenarios indicated that for the first trial, the average change in accuracy over 18 cases was 0.82%. For the second trial, the average change in accuracy was 1.42%. .

In this regard, it is worth mentioning here that each scenario was built with an intention to contributes to the conclusion of this research. Although the magnitude of change in accuracy or loss, depends on data characteristics, trial or categorization type and values of λ and β, when findings from all scenario were considered, it was seen that the average change in accuracy was below 1% for trial 1. The findings suggested that along with the risk of resulting in worse off result that the individual sources, the fusion endeavor may also generate an outcome that is non-noteworthy.

6.3.3.4. **Scenario Analysis for Fewer Data Sources**

While the analysis of the three scenarios, explained in the previous section, undoubtedly suggested the efficacy of the fusion algorithms to obtain a better, the magnitude may vary, estimate in most cases, it also indicated that fusion may not always yield the desired outcomes. In light of the findings, this research intended to investigate further if fusion with fewer sources can yield better estimates for the same scenarios. To examine this, the lowest accuracy source (identified from table 6.4) was removed for each of the scenarios and fusion was conducted utilizing the remaining two sources. As for the two source fusion, the λ (conflict discount) would not add any value to the analysis; the optimal value of β, for each scenario, was taken to conduct the fusion process for both trials.

As shown in Figure 6.6, accuracy difference referred to the difference between the accuracy of two source fusion and the best-reported accuracy of the three source fusion (Figure 6.3, 6.4 and 6.5) for each scenario and categorization. The result showed that for scenario 1 and 3, three source fusion scenarios were better than the two-source fusion. However, for scenario 2, removing the source that had considerably lower accuracy than the other two sources improved the accuracy of the fused estimate, meaning two source fusion is better than both individual estimate and three source fusion. For this scenario, the fusion of two highest accuracy sources (table 6.4) resulted in change in accuracy of 0.8% and 1.7% for trial 1 and 2 respectively.

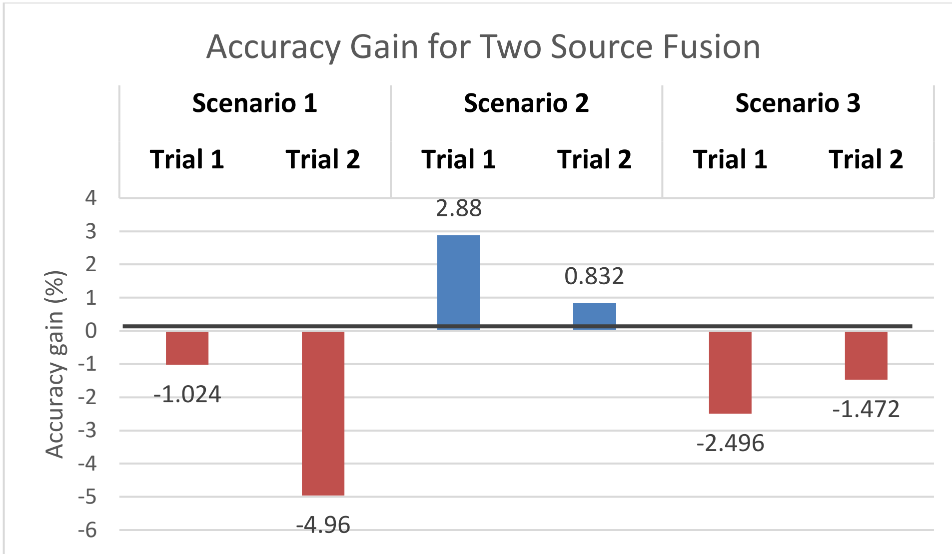


Figure 6.6: Accuracy Comparison by Removing One Source

The finding emphasized that adding more sources in the fusion process may add to the computational complexity without adding any value in terms of accuracy. Hence, a low reliable source in the fusion pool may add risks of obtaining a worse outcome.

6.3.3.5. **Key Lessons Learned**

Overall, the scenario analysis results, presented in figured 6.4 to figure 6.6, suggest the following key takeaways

- Fusion accuracy is sensitive to categorization (trial) type, meaning for a dataset, fusion effort may yield change in accuracy for one label but not for the others
- The change in accuracy varies with the degree of conflict (λ) and context credibility (β)

- Even for the same dataset, the optimal values of λ and β may vary with categorization label (trial)
- Adding more sources may not necessarily yield improvements in the fusion—in other words, more is not always better.
- However, if there are missing values in the individual sources, even a comparatively lower accuracy source may add value to the fusion process.

6.4. Concluding Remarks

The main objective of this chapter was to demonstrate the effectiveness of the fusion framework formulated for consolidating nonmotorized data. The novel weighted voting fusion algorithm may be considered for fusion when the analyst has no prior knowledge about the reliability of individual sources. In the scenario where local agencies do not have adequate actual count data but have access to various model outputs and crowdsourced datasets with enhanced coverage, application of voting algorithms, both the traditional and novel weighted approaches, may be advantageous to obtain better coverage in some cases. As shown by this study, the novel weighted fusion of four sources generated an estimate that is comparable to the best individual source. However, when five sources were considered, the accuracy of fused estimate was lower than the individual source. Hence, the method should be applied with caution as it poses the risk of obtaining a worse off result.

This chapter also demonstrated the application of the DST mechanisms on the simulated data that conformed to real-world nonmotorized data characteristics. The most important finding is that in the majority of cases, the use of fusion methods outperforms the maximum individual source performance where magnitude of the change in accuracy varied. The optimal value of $\beta > 0$ (Figure 6.4 to 6.6) asserted the superiority of incorporating contextual discount in the DST, as proposed by the study. In addition to nonmotorized fusion, the concept of context credibility can be used in other areas of fusion when deemed necessary.

Overall, the results of this numerical experiment led to the conclusion that the performance of fusion firmly depends on the fusion method coupled with the data and situation characteristics. The findings presented in this section addressed the critical question of which fusion method, and under what condition, can outperform individual estimates. There is no optimal categorization or lambda/beta value that can accommodate all data situations. The future application of the fusion approach requires a profound understanding of the data, situation, and sensitivity of the fusion models for varying beta to obtain the optimal combination for a given categorization. Thus, the findings of this research are expected to help analysts derive the best course of action regarding nonmotorized fusion based on the available knowledge and local context.

7. SUMMARY & CONCLUSIONS

This chapter summarized the contribution of the dissertation from both research and application perspectives. The implication of the research findings in safety analysis was highlighted. The concluding remarks discussed the limitations and suggestions for further research.

7.1. Summary and Contributions

The dissertation contributed to both research and professional practice by tapping the uncharted research territory of fusion for nonmotorized traffic. A fastidious approach was undertaken to seek a generalizable fusion solution for nonmotorized demand computation. A major effort in this project was expended to understand the intrinsic characteristics of nonmotorized data and models and explore relevant and accommodative fusion mechanisms to facilitate safety-focused decision-making and infrastructure planning. In doing so, an extensive review of the literature elucidating the aspects of nonmotorized data and demand models and characterization and application of the existing fusion practices was carried out. The literature review led to the recognition of three fundamental issues: (i) Nonmotorized data has unique characteristics, especially when compared with motorized traffic; (ii) The availability and size of actual observation (ground truth), needed to validate crowdsourced data as well as to train models, is often a constraint; (ii) Different nonmotorized data sources are generated in different form and structure which necessitates additional processing steps to obtain a demand or exposure estimate at a given spatial and temporal scope. The understanding of the nonmotorized data coupled with the insights into the fusion mechanism and practice steered the research attention to the decision fusion algorithm, which was deemed as an effective strategy to consolidate information from various nonmotorized traffic data sources.

Because there was no clear guidance of how different data sources—including both traditional and crowdsourced—can be processed and brought together to compute nonmotorized exposure within a facility or area, a conceptual framework was developed for analysis. The bike demand models developed for this project not only illustrated the use of different datasets of varying forms and resolutions to bring into a homogeneous estimate at the micro-level (intersection), they also shed light on the characteristics and aspects of the nonmotorized activity. For example, the direct demand model developed for this research took the traditional approach to the next level by incorporating a bikeability index that can facilitate the modeling approach even when only limited count data are available. Moreover, some of the variables provided unique insights into bike travel behavior within the city, such as the significant and positive influence of the presence of bike signals and bike-accessible bridges. Additionally, Strava and StreetLight data were examined, providing insights into the potential use of crowdsourced data in transportation studies, especially when resources are limited. In summary, the findings benefit stakeholders by explaining the determinants of bicycle activity within the region, thus providing guidance to formulate effective strategies, training, and educational programs geared toward creating a friendlier environment for bicyclists.

From a theoretical perspective, the DST fusion approach with credibility context, as proposed by this study, offered a unique way to incorporate the subjective judgment of experts in mathematical fusion formulation. The experiment on the simulated data, where the proposed approach outperformed the traditional approach, in varying magnitude in above 80% of the cases, underscored the merit of the mechanism, not only for nonmotorized data analysis but also for application in other areas where an analyst's subjective judgment calls for contemplating context for belief refinement in the DST algorithm. The novel weighted approach is also expected to add value in fusion endeavors when adequate ground truth data are unavailable. To summarize, the dissertation developed fusion mechanisms that are pertinent and adaptive to accommodate the characteristics of the nonmotorized activity data. It also formulated guidance to select and customize fusion approaches based on the availability and sample size of the actual demand observation. Therefore, the application of the framework has the adaptability to fit into the need of the local condition.

While there is no doubt that the fusion endeavors may offer increased confidence and, in some cases, coverage, the effort poses the risk of resulting in both worse results or negligible improvement in outcomes. While the framework offered an additional option of data analysis, it is up to the analyst or practitioner to consider and decide the course of option in adopting fusion endeavors given the trade-off between effort and the change in confidence, coverage, and accuracy of the outcomes. Nevertheless, the proposed fusion framework promotes data-driven safety analysis and informed planning, while enhancing the strategic use of available information.

7.2. Generalizability and Transferability

For effective future application, it is imperative to discuss the generalizability and transferability of the individual modeling approaches as well as the fusion methods. This section discussed the scalability and transferability of the models, including the direct demand model and Strava and StreetLight models. The applicability of the fusion approaches was also discussed here.

First, the research followed a structured framework to search and gather a rich set of explanatory variables for building models. While exploring the variables, it was made sure that final variables cover all seven groups (demographics, socioeconomics, network/interaction with vehicle traffic, pedestrian- or bicycle-specific infrastructure, transit facilities, major generators, and land use) as suggested by Munira and Sener (2017). The model outcomes asserted that the bike volume in the Austin area is characterized by sociodemographic (such as the total population of age younger than 15, Black or African American population, and adult population of age 25 or older with no or some academic degree), bicycle infrastructure (such as bike signals), and built environment (such as bikeability index and presence of bicycle-accessible bridge, intersection density) variables. While findings for some variables conformed to previous studies, some other model variables provided unique insights into bike travel behavior within the Austin region. For example, the contribution of transit coverage to the bikeability index highlighted the need to include such a variable to quantify bike friendliness, especially for cities such as Austin, where buses and trains offer bike-friendly services. Moreover, the positive association between the presence of bicycle-accessible bridges, located over lakes or creeks, and bicycle volume can be interpreted in two ways. First, people might be more likely to bike near locations on or around water bodies, and second, they might tend to appreciate the greater accessibility provided by the crossing facilities that the city is building (City of Austin Transportation Department, 2016). Therefore, future model-building processes need to include both traditional variables along with the variables that may reflect the unique characteristics of the study area.

Moreover, the proposed framework provides an important step toward the development of a more generalized direct-demand modeling approach. The bikeability index, a composite measure encompassing multiple pertinent built environment features, emerged as a useful tool for improving the performance of the models as well as quantifying the bicycle friendliness of the entire street network for the region, which is often needed for evaluating development scenarios and policies.

While the final estimation of the model still must be unique to the region because the variable values in the model equation are specific to each region; however, the underlying theory and its application, including the formulation of the bikeability index and integration of it into direct demand modeling, provides a generic approach that can be used in any region. Moreover, the bike friendliness measure, which could be termed in different forms, such as level of service, rating, and score, is often readily available from local urban planning and/or transportation agencies. Therefore, by using the bikeability index, analysts can circumvent the step of seeking a wide range of built-environment-related variables for developing direct-demand models. As a result, the scalability and transferability of the models are improved. Therefore, this index serves as an excellent basic variable to be included in direct demand models, in combination with additional variables, regardless of the study area.

On the other hand, the future adoption and application of the fusion framework would also depend on the data availability as well as local condition. While theoretically, the voting fusion algorithm wouldn’t require ground truth data, there is no doubt of the need of some actual observation data to understand the performance of the individual source and validate the fusion approach. Utilizing all available data sources in the fusion, without any understanding or knowledge of their performance poses the risk of actually obtaining a worse-off result than the best individual source. Hence the application of the voting fusion algorithm requires careful consideration and attention to every analytical step.

The contribution of the DST fusion framework stemmed from its adaptability to fit into the need of the local condition. Future application would need to find the optimal lambda, beta and most importantly categorization labels by performing sensitivity test. Analysts also need to understand the missing case scenario as well as individual accuracy of the models to derive the best course of action.

7.3. Contribution in Safety Analysis

Given that this research's main motivation stemmed from seeking a reliable and robust exposure measure, this section was devoted to discussing the application and benefits of the research findings in terms of safety analysis.

First, it is worth mentioning that the spatial unit of demand analysis, intersection, was selected keeping the safety application in mind. Although nonmotorized crashes occur on various road facilities, such as intersections, driveways, and midblock locations, safety planners often focus on intersection-related crashes because a large proportion of crashes are observed in or near intersections (Choi, 2010). The Texas Strategic Highway Safety Plan reported that more than one-third of fatal and incapacitating injury crashes in Texas in 2013 were identified as intersection related (Texas Department of Transportation [TxDOT], 2016a). A report analyzing crashes in the Capital Area Metropolitan Planning Organization (CAMPO) region estimated that the total cost of intersection crashes was around $3.3 billion from 2010 to 2014 (Texas Department of Transportation (TxDOT, 2016a). The same report also revealed that more than one of every seven severe crashes (fatal and suspected serious injury) at intersections in the CAMPO region involved pedestrians or bicyclists. Hence, the location-based exposure data is expected to contribute to the efforts of city officials to discern the trend in crash rates, identify high-risk locations, and understand the crash causation.

Moreover, the direct demand model, developed for this study (Munira et al., 2021) itself, was reported as a reliable source of volume estimate for the entire study area, exhibiting high performance in terms of accuracy and goodness of fit. The predicted intersection bike volume for the entire region is expected to be useful to safety analysts, supporting the City's goal of incorporating data for informed decision-making for nonmotorized transport. A similar approach had also been undertaken in a study (Munira et al., 2020) exploring pedestrian crash severities at intersections through a Bayesian multivariate spatial Poisson-lognormal model, where a direct demand model was utilized to estimate the annual average daily pedestrian volume-the exposure measure. The direct demand model also offered unique insights into the effectiveness of some of the policies taken by the City's transportation department to improve the safety and mobility of bicyclists. For example, the positive association between the presence of bike signal and bicycle volume indicated that the number of bike signals positively influences bike volume around intersections. The City of Austin Transportation Department installed bicycle signals in several locations in 2017 to improve safety for both bicyclists and other road users (City of Austin 2018b). Some of the intersections have a leading pedestrian and bicycle interval to ensure that pedestrians and bicyclists have extra time to start crossing. The model results might be an implication of bicyclists' perception of feeling safe at such intersections. The city might evaluate the feasibility and the need to install additional bike signals at intersections where bicycle volume is high.

Finally, the multi-source fusion endeavor, offering a superior (better coverage and accuracy) estimate than the individual sources, supports safety planner's goal of incorporating robust and reliable exposure estimates, which is instrumental to the orchestration of an efficacious crash analysis for nonmotorized traffic. The fused estimate can be applied in both micro and macro-level models. For example, a macro-level model (such as for block group) may aggregate the intersection AADB volumes into a specific zone (using the mid-value of each category) to compute the average zonal exposure measure. These estimates can benefit various zone-level crash models, including hotspot analysis which often relied on vehicular volume or population measure due to the unavailability of nonmotorized volume-related exposure data (Wang et al., 2016). An area- or macro-level safety analysis is pivotal to recognizing safety problems in a larger area to facilitate long-term policy planning to reduce crashes (Wang et al., 2016).

Moreover, the fused estimates can also be used as categorical variables in crash models. For example, in exploring individual's safety-related perceptions (micro model) regarding biking, the maximum bike volume at a buffer zone of the one's household location can be used as an explanatory variable. Such models may help recognize if the biking activity in the neighborhood formulates an individual's subjective perception. Nonetheless, a robust exposure measure obtained from a fusion endeavor would support focused policy efforts and effective safety implementation plans for the city.

7.4. Limitations and Future Scope of Research

The dissertation is not without limitations and calls for further research in this area. The major limitations stemmed from the limited sample size of the actual observation, which may not be fully representative of the entire region. To better capture the spatial variation of bike volume, a robust site collection approach involving sampling strategy and a larger data-collection program are warranted for selecting a representative sample of the study area.

Moreover, the actual bicycle volume count data (from 44 locations) were utilized as the ground truth data for validating the voting fusion algorithms. Due to the limited sample size, it would not be meaningful to consider this process as a validation exercise with reasonable confidence. Moreover, in the ideal case (when the analysts have the option to separate the validation dataset from the training dataset), a separate validation dataset is desirable to conduct a true evaluation. Thus, this process was instead used to draw insights and gain understanding of how each of the models and the voting algorithms were performing.

The nonmotorized activity and land use data were generated in different years, ranging from 2010 to 2018. Although the estimates were scaled using the actual count data, the process still poses a risk of bias, leading to inaccurate outputs. Future efforts should seek to collect data for the same year to obtain more accurate demand estimates.

The traffic assignment model for this research only considered distance and route comfort as factors of the route choice model. The route choice model for Austin's bicyclists may be more complex, warranting consideration of multiple factors such as safety, stress, pollution, scenery etc. Future research may consider developing multi-factor traffic assignment model to obtain a more accurate facility-level volume.

Besides, the research considered only one context in refining the fusion algorithms. Given that context has a very broad definition, some additional information about a state or object may also serve as a meaningful context. For example, temporal variability is also essential in nonmotorized demand data but was beyond the scope of the current research due to the unavailability of relevant data to investigate the influence of temporal features as context. The process of investigating and exploring multiple contexts may be an interesting topic for further research. Moreover, future research may explore other fusion algorithms, such as ensemble learning-based fusion, when more readily available location-based data (such as crowdsourced data) are available.

REFERENCES

1. Akar, G., & Clifton, K. (2009). Influence of individual perceptions and bicycle infrastructure on decision to bike. *Transportation Research Record: Journal of the Transportation Research Board*, (2140), 165-172
2. Al-Ani, A., & Deriche, M. (2002). A new technique for combining multiple classifiers using the Dempster-Shafer theory of evidence. *Journal of Artificial Intelligence Research*, *17*, 333-361.
3. Alibabai H, Mahmassani H (2009) Dynamic Origin-Destination Demand Estimation Using Turning Movement Counts. Transportation Research Record: Journal of the Transportation Research Board 2085:39–48.
4. Alotaibi, B., & Elleithy, K. (2016, April). A majority voting technique for wireless intrusion detection systems. In *2016 IEEE Long Island Systems, Applications and Technology Conference (LISAT)* (pp. 1-6). IEEE.
5. Amey, A., Liu, L., Pereira, F., Zegras, C., Veloso, M., Bento, C., and Biderman, A. (2009). *State of The Practice Overview of Transportation Data Fusion: Technical and Institutional Considerations*. Transportation Systems-Working Paper Series
6. Antal, P., Millinghoffer, A., Hullám, G., Szalai, C., & Falus, A. (2008, September). A Bayesian view of challenges in feature selection: feature aggregation, multiple targets, redundancy and interaction. In *New Challenges for Feature Selection in Data Mining and Knowledge Discovery* (pp. 74-89).
7. Antoniou, C., Balakrishna, R., & Koutsopoulos, H. N. (2011). A synthesis of emerging data collection technologies and their impact on traffic management applications. *European Transport Research Review*, 3(3), 139-148
8. Aoun, A., Bjornstad, J., DuBose, B., Mitman, M., Pelon, M., and Fehr and Peers (2015). *Bicycle and Pedestrian Forecasting Tools: State of the Practice*. DTFHGI-11-H-00024. Federal Highway Administration, Washington, DC.
9. Atallah, R., & Al-Mousa, A. (2019, October). Heart Disease Detection Using Machine Learning Majority Voting Ensemble Method. In *2019 2nd International Conference on new Trends in Computing Sciences (ICTCS)* (pp. 1-6). IEEE.
10. Atrey, P. K., Kankanhalli, M. S., & Jain, R. (2006). Information assimilation framework for event detection in multimedia surveillance systems. *Multimedia systems*, *12*(3), 239-253.
11. Aultman-Hall, L., Lane, D., & Lambert, R. R. (2009). Assessing impact of weather and season on pedestrian traffic volumes. *Transportation research record*, *2140*(1), 35-43.
12. Awais, M., Yan, F., Mikolajczyk, K., & Kittler, J. (2011, September). Novel fusion methods for pattern recognition. In *Joint European Conference on Machine Learning and Knowledge Discovery in Databases* (pp. 140-155). Springer, Berlin, Heidelberg.

13. Awasthi, A., & Chauhan, S. S. (2011). Using AHP and Dempster–Shafer theory for evaluating sustainable transport solutions. *Environmental Modelling & Software*, *26*(6), 787-796.
14. Bachmann, C., Abdulhai, B., Roorda, M. J., & Moshiri, B. (2013). A comparative assessment of multi-sensor data fusion techniques for freeway traffic speed estimation using microsimulation modeling. *Transportation research part C: emerging technologies*, *26*, 33-48.
15. Bae, B. (2017). Real-time Traffic Flow Detection and Prediction Algorithm: Data-Driven Analyses on Spatio-Temporal Traffic Dynamics.
16. Barillot, C., Lemoine, D., Le Briquer, L., Lachmann, F., and Gibaud, B. (1993). Data Fusion in Medical Imaging: Merging Multimodal and Multipatient Images, Identification of Structures and 3D Display Aspects. *European Journal of Radiology*, Vol. 17, No. 1, pp 22-27.
17. Barnes, G., & Krizek, K. (2005). Estimating bicycling demand. *Transportation Research Record: Journal of the Transportation Research Board*, (1939), 45-51.
18. Bartier, P.M.; Keller, C.P. (1996). Multivariate interpolation to incorporate thematic surface data using inverse distance weighting (IDW). *Comput. Geosci. 22*, 795–799
19. Bashbaghi, S., Granger, E., Sabourin, R., & Bilodeau, G. A. (2017). Dynamic Selection of Exemplar-SVMs for Watch-list Screening through Domain Adaptation. In *ICPRAM* (pp. 738-745).
20. Battiti, R., & Colla, A. M. (1994). Democracy in neural nets: Voting schemes for classification. *Neural Networks*, *7*(4), 691-707.
21. BCycle (2018). *Austin B-cycle Expands Service to UT Campus and West Campus Neighborhoods.* Retrieved from https://www.bcycle.com/news/2018/02/14/austin-b-cycle-expands-service-to-ut-campus-and-west-campus-neighborhoods
22. Bellenger, A. (2013). *Semantic Decision Support for Information Fusion Applications* (Doctoral dissertation, Rouen, INSA).
23. Benz, R. J., Turner, S., & Qu, T. (2013). Pedestrian and bicyclist counts and demand estimation study. Texas A&M Transportation Institute
24. Benz, R. J., Turner, S., & Qu, T. (2013). Pedestrian and bicyclist counts and demand estimation study. Texas A&M Transportation Institute
25. Bezdek, J. C., Keller, J., Krisnapuram, R., & Pal, N. (1999). *Fuzzy models and algorithms for pattern recognition and image processing* (Vol. 4). Springer Science & Business Media.
26. Blanc, B., and Figliozzi, M. (2016). Modeling the Impacts of Facility Type, Trip Characteristics, and Trip Stressors on Cyclists' Comfort Levels Utilizing Crowdsourced

Data. *Transportation Research Record: Journal of the Transportation Research Board,* No. 2587, pp. 100–108

27. Blanc, B., Figliozzi, M., Clifton, K. (2016) How Representative of Bicycling Populations Are Smartphone Application Surveys of Travel Behavior? *Transportation Research Record: Journal of the Transportation Research Board* 2587, 78–89. doi:10.3141/2587-10.
28. Blasch, E., & Plano, S. (2002, April). JDL Level 5 Fusion Model 'User Refinement' Issues and Applications in Group Tracking. *In Proc. SPIE*, Vol. 4729, pp. 270-279.
29. Boström, H., Andler, S. F., Brohede, M., Johansson, R., Karlsson, A., Van Laere, J., ... & Ziemke, T. (2007). On the definition of information fusion as a field of research. Informatics Research Centre, University of Skövde, Tech. Rep. HS-IKI-TR-07-006, 2007. http://urn.kb.se/resolve?urn=urn:nbn:se:his:diva-1256
30. Broach, J., Dill, J., and Gliebe, J. (2012). Where Do Cyclists Ride? A Route Choice Model Developed with Revealed Preference GPS Data. *Transportation Research Part A: Policy and Practice,* Vol. *46, No.* 10, pp. 1730-1740.
31. Broach, J., Gliebe, J. P., & Dill, J. (2011). *Bicycle route choice model developed from revealed-preference GPS data* (No. 11-3901).
32. Brooks, R. R., and Iyengar, S. S. (1998). *Multi-Sensor Fusion: Fundamentals and Applications with Software*. Prentice-Hall, Inc.
33. Buehler, R. (2019). *A first look: Trends in walking and cycling in the United States 2001–2017*. Presented at the National Household Travel Survey (NHTS) Data for Transportation Applications Workshop. Retrieved from http://onlinepubs.trb.org/onlinepubs/Conferences/2018/NHTS/BuehlerTrendsinWalkingandCycling.pdf
34. Burduk, R. (2017, June). Integration base classifiers based on their decision boundary. In *International Conference on Artificial Intelligence and Soft Computing* (pp. 13-20). Springer, Cham.
35. Cai, Q., Abdel-Aty, M., & Lee, J. (2017). Macro-level vulnerable road users crash analysis: A Bayesian joint modeling approach of frequency and proportion. *Accident Analysis & Prevention*, *107*, 11–19.
36. Castanedo, F. (2013). A review of data fusion techniques. *The Scientific World Journal*, *2013*.
37. Centers for Disease Control and Prevention (2016). Bicycling and Walking in the United States 2016, Benchmarking Report. Alliance for Biking and Walking, Washington, DC.
38. Chang, C. L., Tsai, C. H., & Chen, L. (2003). Applying grey relational analysis to the decathlon evaluation model. *International Journal of the Computer, the Internet and Management*, *11*(3), 54–62.

39. Chan-Lau, M. J. A. (2017). *Lasso regressions and forecasting models in applied stress testing*. IMF Working Paper, Institute for Capacity and Development. doi:10.5089/9781475599022.001
40. Chen, C., Anderson, J. C., Wang, H., Wang, Y., Vogt, R., & Hernandez, S. (2017b). How bicycle level of traffic stress correlate with reported cyclist accidents injury severities: A geospatial and mixed logit analysis. *Accident Analysis & Prevention*, *108*, 234-244.
41. Chen, L., Diao, L., & Sang, J. (2018). Weighted evidence combination rule based on evidence distance and uncertainty measure: An application in fault diagnosis. *Mathematical Problems in Engineering*, *2018*.
42. Chen, P., Zhou, J., & Sun, F. (2017a). Built environment determinants of bicycle volume: A longitudinal analysis. *Journal of transport and land use*, *10*(1), 655-674.
43. Chen, P., Zhou, J., & Sun, F. (2017a). Built environment determinants of bicycle volume: A longitudinal analysis. *Journal of transport and land use*, *10*(1), 655-674.
44. Choi, E. H. (2010). *Crash factors in intersection-related crashes: An on-scene perspective* (No. HS-811 366).
45. Choi, K., & Chung, Y. (2002). A data fusion algorithm for estimating link travel time. *ITS journal*, *7*(3-4), 235-260.
46. Chust, G., Ducrot, D., & Pretus, J. L. (2004). Land cover discrimination potential of radar multitemporal series and optical multispectral images in a Mediterranean cultural landscape. *International Journal of Remote Sensing*, *25*(17), 3513-3528.
47. City of Austin Planning and Zoning. (2019). *Demographics*. Retrieved from http://www.austintexas.gov/demographics
48. City of Austin. (2015). *2014 Austin bicycle plan*. Retrieved from https://www.austintexas.gov/sites/default/files/files/2014_Austin_Bicycle_Master_Plan__Reduced_Size_.pdf
49. City of Austin. (2017). *Austin Texas Bike Map*. Retrieved from https://austintexas.gov/sites/default/files/files/Transportation/2017_Austin_Bike_Map_-_Side_2.pdf
50. City of Austin. (2018). *The City of Austin Transportation Department 2018 annual report*. Retrieved from https://1d0d7cb9-bfde-4667-8eeb-20a5ee919bd6.filesusr.com/ugd/956239_455271f084ea4a4d9caf7bb098fb18ab.pdf
51. City of Austin. 2018b. "*Austin mobility news*." Accessed May 4, 2021. https://mailchi.mp/austintexas/mobilitynews_09172018.
52. Clifton, J. K., Burnier V. C., Schneider, R., Huang, S., Kang, W.M. (2008). Pedestrian Demand Model for Evaluating Pedestrian Risk Exposure. The National Center for Smart Growth Research and Education University of Maryland, College Park. Maryland State Highway Administration, Office of Traffic and Safety. Available http://kellyjclifton.com/MoPeD/Summary_Final.pdf. Accessed 10th September 2018

53. Clifton, K., & Krizek, K. J. (2004). The utility of the NHTS in understanding bicycle and pedestrian travel. In *National Household Travel Survey Conference: understanding our nation's travel* (pp. 1-2).
54. Collins, T. W., Grineski, S. E., & Aguilar, M. D. L. R. (2009). Vulnerability to environmental hazards in the Ciudad Juárez (Mexico)–El Paso (USA) metropolis: a model for spatial risk assessment in transnational context. *Applied Geography*, *29*(3), 448-461
55. Corne, D. W., Deb, K., Fleming, P. J., & Knowles, J. D. (2003). The good of the many outweighs the good of the one: evolutionary multi-objective optimization. *IEEE Connections Newsletter*, *1*(1), 9-13.
56. Cruz, R. M., Sabourin, R., & Cavalcanti, G. D. (2018). Dynamic classifier selection: Recent advances and perspectives. *Information Fusion*, *41*, 195-216.
57. Dadashova, B., Griffin, G. P., Das, S., Turner, S., & Sherman, B. (2020). Estimation of Average Annual Daily Bicycle Counts using Crowdsourced Strava Data. *Transportation Research Record*, 0361198120946016.
58. Dailey, D. J., Harn, P., and Lin, P.-J. (1996). *ITS data fusion.* Technical Report Research Project T9903, Task 9, Washington State Transportation Commission, Department of Transportation and the U.S. Department of Transportation, Federal Highway Administration.
59. Dasarathy, B. V. (1997). Sensor fusion potential exploitation-innovative architectures and illustrative applications. *Proceedings of the IEEE*, *85*(1), 24-38.
60. Dasarathy, B. V. (1999). Pattern Recognition-a Pillar in the Sensor Fusion Edifice or Vice Versa?. In *International Conference on Advances in Pattern Recognition* (pp. 445-454). Springer, London.
61. Dasarathy, B. V. (2001). Information fusion-what, where, why, when, and how?. *Information Fusion*, *2*(2), 75-76.
62. Day, W. H. (1988) Consensus methods as tools for data analysis. In: H. H. Bock, editor, *Classification and Related Methods for Data Analysis*, pp. 317–324. Elsevier Science Publishers B.V. (North Holland)
63. De Donato, W., Pescapé, A., & Dainotti, A. (2014). Traffic identification engine: an open platform for traffic classification. *IEEE Network*, *28*(2), 56-64.
64. DeMaio, P. (2009). Bike-sharing: History, impacts, models of provision, and future. *Journal of public transportation*, *12*(4), 3.
65. Dempster, A. P. "Upper and lower probabilities induced by a multivalued mapping," *Ann. Math. Stat.*, vol. 38, no. 2, pp. 325–339, 1967.
66. Deng, J. L. (1987). Basic method of grey system. *Huangzhong Polytechnical Institute Press, Wuhan, China*.
67. Deng, X., Liu, Q., Deng, Y., & Mahadevan, S. (2016). An improved method to construct basic probability assignment based on the confusion matrix for classification problem. *Information Sciences*, *340*, 250-261.

68. Denoeux, T. (1997). Analysis of evidence-theoretic decision rules for pattern classification. *Pattern recognition*, *30*(7), 1095-1107.
69. Dey, A. K. and Abowd, G. D (2000). Towards a better understanding of context and contextawareness. In the Workshop on the What, Who, Where, When and How of Context-Awareness, affiliated with the 2000 ACM Conference on Human Factors in Computer Systems (CHI 2000), The Hague, Netherlands. April 1-6, 2000. Available at: ftp://ftp.cc.gatech.edu/pub/gvu/tr/1999/99-22.pdf
70. Dey, A. K., Abowd, G. D., & Salber, D. (2001). A conceptual framework and a toolkit for supporting the rapid prototyping of context-aware applications. *Human–Computer Interaction*, *16*(2-4), 97-166.
71. Di Lecce, V., Dimauro, G., Guerriero, A., Impedovo, S., Pirlo, G., & Salzo, A. (2000, September). Classifier combination: the role of a-priori knowledge. In *IWFHR* (Vol. 7, pp. 143-152).
72. Dill, J. (2009). Bicycling for transportation and health: the role of infrastructure. *Journal of public health policy*, *30*(1), S95-S110.
73. Dill, J., & Carr, T. (2003). Bicycle commuting and facilities in major US cities: if you build them, commuters will use them. *Transportation Research Record: Journal of the Transportation Research Board*, (1828), 116-123
74. Dubois, D., & Prade, H. (1988). Representation and combination of uncertainty with belief functions and possibility measures. *Computational intelligence*, *4*(3), 244-264.
75. Duin, R. P. (2002, August). The combining classifier: to train or not to train?. In *Object recognition supported by user interaction for service robots* (Vol. 2, pp. 765-770). IEEE.
76. Duin, R. P. W. (2001) *A Discussion on Trained Combining of Classifiers*. concept paper
77. Durrant-Whyte, H. F. (1990). Sensor models and multisensor integration. In *Autonomous robot vehicles* (pp. 73-89). Springer, New York, NY.
78. DVRPC. 2011. DVRPC Travel Demand Model Upgrade -Travel Improvement Model (TIM) 1.0. Available https://www.dvrpc.org/reports/TR10006.pdf
79. Eco-Counter. (2019). *Products*. Retrieved from https://www.eco-compteur.com/en/produits/multi-range/urban-multi/
80. Ekbal, A., & Saha, S. (2011). Weighted vote-based classifier ensemble for named entity recognition: a genetic algorithm-based approach. *ACM Transactions on Asian Language Information Processing (TALIP)*, *10*(2), 1-37.
81. El Esawey, M. (2014). Estimation of annual average daily bicycle traffic with adjustment factors. *Transportation Research Record*, *2443*(1), 106-114.
82. El Esawey, M. (2018). Daily bicycle traffic volume estimation: comparison of historical average and count models. *Journal of Urban Planning and Development*, *144*(2), 04018011.

83. El Faouzi, N. E., & Klein, L. A. (2016). Data fusion for ITS: techniques and research needs. *Transportation Research Procedia*, *15*, 495-512
84. El Faouzi, N. E., & Lesort, J. B. (1995). *Travel time estimation on urban networks from traffic data and on-board trip characteristics* (No. Volume 1).
85. El Faouzi, N. E., Klein, L. A., & De Mouzon, O. (2009). Improving travel time estimates from inductive loop and toll collection data with Dempster–Shafer data fusion. *Transportation research record*, *2129*(1), 73-80.
86. El Faouzi, N. E., Leung, H., & Kurian, A. (2011). Data Fusion in Intelligent Transportation Systems: Progress and Challenges–A Survey. *Information Fusion*, Vol. 12, No. 1,pp 4-10.
87. Elmenreich, W. (2002). An introduction to sensor fusion. *Vienna University of Technology, Austria*, *502*, 1-28.
88. Esteban, J., Starr, A., Willetts, R., Hannah, P., & Bryanston-Cross, P. (2005). A review of data fusion models and architectures: towards engineering guidelines. *Neural Computing & Applications*, *14*(4), 273-281.
89. Ewing, R., & Cervero, R. (2001). Travel and the built environment: a synthesis. *Transportation research record*, *1780*(1), 87-114.
90. Faghih-Imani, A., Eluru, N., El-Geneidy, A. M., Rabbat, M., & Haq, U. (2014). How land-use and urban form impact bicycle flows: Evidence from the Bicycle-Sharing System (BIXI) in Montreal. *Journal of Transport Geography*, Vol. 41, pp. 306–314.
91. Faghih-Imani, A.; Hampshire, R.; Marla, L.; Eluru, N. (2017) An empirical analysis of bike sharing usage and rebalancing: Evidence from Barcelona and Seville. *Transportation Research Part A: Policy and Practice*, 97, 177–191.
92. Fagnant, D. J., and Kockelman, K. (2016). A Direct-Demand Model for Bicycle Counts: The Impacts of Level of Service and Other Factors. Environment and Planning B: Planning and Design, Vol. 43, No. 1, pp. 93–107.
93. Federal Highway Administration. 2013. *Traffic Monitoring Guide*. Washington, DC, Report FHWA-PL-13-015. https://www.fhwa.dot.gov/policyinformation/tmguide/tmg_fhwa_pl_17_003.pdf
94. Ferreira, A., Felipussi, S. C., Alfaro, C., Fonseca, P., Vargas-Munoz, J. E., dos Santos, J. A., & Rocha, A. (2016). Behavior knowledge space-based fusion for copy–move forgery detection. *IEEE Transactions on Image Processing*, *25*(10), 4729-4742.
95. Feudo, F. L., Morton, B., Capelle, T., & Prados, E. (2017). Modeling Urban Dynamics using LUTI Models: Calibration Methodology for the Tranus-Based Model of the Grenoble Urban Region. *In HKSTS 2017-22nd International Conference of Hong Kong Society for Transportation Studies*.

96. Figliozzi, M. A., Unnikrishnan, A., Kothuri, S., Caviedes, A., & Soto Padín, D. R. (2018). *Compliance and Surrogate Safety Measures for Uncontrolled Crosswalk Crossings in Oregon* (No. FHWA-OR-RD-19-02). Oregon. Dept. of Transportation. Research Section.
97. Fitzpatrick, K., Avelar, R., & Turner, S. M. (2018). *Guidebook on Identification of High Pedestrian Crash Locations* (No. FHWA-HRT-17-106). United States. Federal Highway Administration. Office of Safety Research and Development.
98. Florea, M. C., Jousselme, A. L., & Bossé, É. (2010). Dynamic estimation of evidence discounting rates based on information credibility. *RAIRO-Operations Research*, *44*(4), 285-306.
99. Frade, I. & Ribeiro, (2015)/ A. Bike-sharing stations: A maximal covering location approach. *Transportation Research Part A: Policy and Practice*, 82, 216–227.
100. Frigui, H., Gader, P. D., & Abdallah, A. C. B. (2008, April). A generic framework for context-dependent fusion with application to landmine detection. In *Detection and Sensing of Mines, Explosive Objects, and Obscured Targets XIII* (Vol. 6953, p. 69531F). International Society for Optics and Photonics.
101. Geller, R. (2009). *Four types of cyclists*. Retrieved from https://www.portlandoregon.gov/transportation/44597?a=237507
102. Gonsalves, P., Cunningham, R., Ton, N., & Okon, D. (2000, July). Intelligent threat assessment processor (ITAP) using genetic algorithms and fuzzy logic. In *Proceedings of the Third International Conference on Information Fusion* (Vol. 2, pp. THB1-18). IEEE.
103. Goodchild, M. F., & Li, L. (2012). Assuring the quality of volunteered geographic information. *Spatial Statistics*, *1*, 110–120. doi:10.1016/j.spasta.2012.03.002
104. Gosse, C., and Clarens, A. (2014). Estimating Spatially and Temporally Continuous Bicycle Volumes by Using Sparse Data. Transportation Research Record, No. 2443, pp. 115–122.
105. Greene-Roesel, R., Diogenes, M.C., Ragland, D.R. (2007). *Estimating Pedestrian Accident Exposure: Protocol Report*. (In Estimating Pedestrian Accident Exposure: Final Report, California PATH Research Report UCB-ITS-PRR-2010-32, Final Report for Task Orders 5211/6211).
106. Griffin, G. P., & Jiao, J. (2015). Where does bicycling for health happen? Analysing volunteered geographic information through place and plexus. *Journal of Transport and Health*, *2*(2), 238–247. doi:10.31235/osf.io/5gy3u
107. Griswold, J., Medury, A., and Schneider, R. (2011). Pilot Models for Estimating Bicycle Intersection Volumes. Transportation Research Record 2247. pp. 1–7.

108. Gunatilaka, A. H., & Baertlein, B. A. (2001). Feature-level and decision-level fusion of noncoincidently sampled sensors for land mine detection. *IEEE transactions on pattern analysis and machine intelligence*, *23*(6), 577-589.

109. Guo, F., Polak, J. W., & Krishnan, R. (2018). Predictor fusion for short-term traffic forecasting. *Transp. Res. C, Emerg. Technol.*, *92*, 90-100.

110. Guo, H., Shi, W., & Deng, Y. (2006). Evaluating sensor reliability in classification problems based on evidence theory. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, *36*(5), 970-981.

111. Haenni, R. (2005, July). Shedding new light on Zadeh's criticism of Dempster's rule of combination. In *2005 7th International conference on information fusion* (Vol. 2, pp. 6-pp). IEEE.

112. Hall, D. L., & Garga, A. K. (1999, July). Pitfalls in data fusion (and how to avoid them). In *Proceedings of the Second International Conference on Information Fusion (Fusion'99)* (Vol. 1, pp. 429-436).

113. Hall, D. L., & Steinberg, A. (2001). *Dirty secrets in multisensor data fusion*. Pennsylvania State Univ University Park Applied Research Lab.

114. Hall, D., & Llinas, J. (Eds.). (2001). *Multisensor data fusion*. CRC press

115. Hall, D.L. (1992) Mathematical Techniques in Multisensor Data Fusion. Norwood, MA: Artech House,

116. Han, J. (2012). Multi-sensor data fusion for travel time estimation. Doctoral Dissertation. Imperial College London, United Kingdom

117. Handy, S., Shafizadeh, K. R., & Schneider, R. (2013). *California smart-growth trip generation rates study* (No. CA13-1940). University of California, Davis Urban Land Institute and Transportation Center.

118. Hankey, S., Lindsey, G., & Marshall, J. (2014). Day-of-year scaling factors and design considerations for nonmotorized traffic monitoring programs. *Transportation Research Record*, *2468*(1), 64-73.

119. Hankey, S., Lindsey, G., Wang, X., Borah, J., Hoff, K., Utecht, B., & Xu, Z. (2012). Estimating use of nonmotorized infrastructure: Models of bicycle and pedestrian traffic in Minneapolis, MN. *Landscape and Urban Planning, 107*, 307–316. doi:10.1016/j.landurbplan.2012.06.005

120. Hankey, S., Lu, T., Mondschein, A., and Buehler, R. (2017). Merging Traffic Monitoring and Direct-Demand Modeling to Assess Spatial Patterns of Annual Average Daily Bicycle and Pedestrian Traffic. Transportation Research Board 2017 Annual Meeting.

121. Hasani, M., Jahangiri, A., Sener, I. N., Munira, S., Owens, J. M., Appleyard, B., ... & Ghanipoor Machiani, S. (2019). Identifying high-risk intersections for walking and bicycling using multiple data sources in the city of San Diego. Journal of advanced transportation, 2019.

122. He, S., Zhang, J., Cheng, Y., Wan, X., and Ran, B. (2016). Freeway Multisensor Data Fusion Approach Integrating Data from Cellphone Probes and Fixed Sensors. *Journal of Sensors,* Vol.2016, Article ID 7269382, 2016.

123. Hégarat-Mascle, L., Richard, D., & Ottlé, C. (2003). Multi-scale data fusion using Dempster-Shafer evidence theory. *Integrated Computer-Aided Engineering*, *10*(1), 9-22.

124. Heydari, S., Fu, L., Miranda-Moreno, L. F., & Joseph, L. (2017). Using a flexible multivariate latent class approach to model correlated outcomes: A joint analysis of pedestrian and cyclist injuries. *Analytic Methods in Accident Research, 13,* 16–27.

125. Hillel, A. B., Lerner, R., Levi, D., & Raz, G. (2014). Recent progress in road and lane detection: a survey. *Machine vision and applications*, *25*(3), 727-745.

126. Ho, T. K., Hull, J. J., & Srihari, S. N. (1994). Decision combination in multiple classifier systems. *IEEE transactions on pattern analysis and machine intelligence*, *16*(1), 66-75.

127. Hochmair, H. H., Bardin, E., & Ahmouda, A. (2019). Estimating bicycle trip volume for Miami-Dade County from Strava tracking data. *Journal of Transport Geography, 75,* 58–69. doi:10.1016/j.jtrangeo.2019.01.013

128. Hood, J., Sall, E., & Charlton, B. (2011). A GPS-based bicycle route choice model for San Francisco, California. *Transportation Letters*, Vol. *3, No.* 1, pp. 63-75.

129. Howell, D. C. (1997). *Statistical methods for psychology* (4th ed.). Belmont, CA: Duxbury Press.

130. Huang, Y. S. (1993). An optimal method of combining multipe classifiers for unconstrained handwritten numeral recognition. In *Proc. 3rd International Workshop on Frontiers in Handwriting Recognition* (pp. 11-20).

131. Huang, Y. S. (1994). A neural network approach for multi-classifier recognition systems. In *Proc. 4th International Workshop on Frontiers in Handwriting Recognition, Dec. 1994* (pp. 235-244).

132. Huang, Y. S., & Suen, C. Y. (1993). The behavior-knowledge space method for combination of multiple classifiers. In *IEEE computer society conference on computer vision and pattern recognition* (pp. 347-347). Institute of Electrical Engineers Inc (IEEE).

133. Huang, Y., Suen, C.: (1995) A Method of Combining Multiple Experts for Recognition of Unconstrained Handwritten Numerals. IEEE Trans. on Pattern Analysis and Machine Intelligence (TPAMI) 17(1) (1995) 90-94

134. Huang, Z., Ling, X., Wang, P., Zhang, F., Mao, Y., Lin, T., & Wang, F. Y. (2018). Modeling real-time human mobility based on mobile phone and transportation data fusion. *Transportation research part C: emerging technologies*, *96*, 251-269.

135. Hull, J. J., Ho, T. K., Favata, J., Govindaraju, V., & Srihari, S. (1992). Combination of segmentation-based and wholistic handwritten word recognition algorithms. *From Pixels to Features III: Frontiers in Handwriting Recognition*, 261-272

136. Ince, A. N., Topuz, E., Panayirci, E., & Isik, C. (2012). *Principles of Integrated Maritime Surveillance Systems.* Vol. 527. Springer Science & Business Media.

137. Jackson, S., Miranda-Moreno, L. F., Rothfels, C., & Roy, Y. (2014). *Adaptation and implementation of a system for collecting and analyzing cyclist route data using smartphones*(No. 14-4637).

138. Jackson, S., Mullen, W., Agouris, P., Crooks, A., Croitoru, A., & Stefanidis, A. (2013). Assessing completeness and spatial error of features in volunteered geographic information. *ISPRS International Journal of Geo-Information*, *2*(2), 507–530. doi:10.3390/ijgi2020507

139. Jestico, B., Nelson, T., & Winters, M. (2016). mapping ridership using crowdsourced cycling data. *Journal of Transport Geography*, *52*, 90–97. doi:10.1016/j.jtrangeo.2016.03.006

140. Ji, C., & Ma, S. (1997). Combinations of weak classifiers. In *Advances in Neural Information Processing Systems* (pp. 494-500).

141. Ji, Y., Ma, X., Yang, M., Jin, Y., & Gao, L. (2018). Exploring spatially varying influences on metro-bikeshare transfer: A geographically weighted Poisson regression approach. *Sustainability, 10(5),* 1526.

142. Jousselme, A. L., Grenier, D., & Bossé, É. (2001). A new distance between two bodies of evidence. *Information fusion*, *2*(2), 91-101.

143. Kamel, M. B., Sayed, T., & Osama, A. (2019). Accounting for mediation in cyclist-vehicle crash models: a Bayesian mediation analysis approach. *Accident Analysis & Prevention*, *131*, 122-130.

144. Khademi, G., & Ghassemian, H. (2018). Incorporating an adaptive image prior model into Bayesian fusion of multispectral and panchromatic images. *IEEE Geoscience and Remote Sensing Letters*, *15*(6), 917-921.

145. Kim, H., Kim, H., Moon, H., & Ahn, H. (2011). A weight-adjusted voting algorithm for ensembles of classifiers. *Journal of the Korean Statistical Society*, *40*(4), 437-449.

146. Kittler, J., Hatef, M., Duin, R. P., & Matas, J. (1998). On combining classifiers. *IEEE transactions on pattern analysis and machine intelligence*, *20*(3), 226-239.

147. Klein L. A. (2012). Sensor and data fusion: a tool for information assessment and decision making. 2nd Edition. SPIE press

148. Klein, L. (2000). Dempster–Shafer data fusion at the traffic management centre, Transportation Research Board, in: 79th Annual Meeting, Paper No. 00-1211, Washington, DC, Transportation Research Board, 2000

149. Klobucar, M., & Fricker, J. (2007). Network evaluation tool to improve real and perceived bicycle safety. *Transportation Research Record: Journal of the Transportation Research Board*, (2031), 25-33.

150. Kohavi, R. (1995). *A study of cross-validation and bootstrap for accuracy estimation and model selection.* In Proceedings of the 14th International Joint Conference on Artificial Intelligence IJCAI'95, Montreal, QC, Canada, 20–25 August 1995 (Vol. 14, No. 2, pp. 1137-1145).
151. Koks, D., & Challa, S. (2003). *An introduction to Bayesian and Dempster-Shafer data fusion*. Defence Science and Technology Organisation Salisbury (Australia) Systems Sciences Lab.
152. Kong, Q. J., & Liu, Y. (2009). A Model of Federated Evidence Fusion for Real-Time Traffic State Estimation. In *Sensor and Data Fusion*. IntechOpen.
153. Kong, Q. J., Chen, Y., & Liu, Y. (2007). *An Improved Evidential Fusion Approach for Real-Time Urban Link Speed Estimation*. Intelligent Transportation Systems Conference, 2007. ITSC 2007. IEEE (pp. 562-567). IEEE
154. Krizek, K. J., Iacono, M., El-Geneidy, A. M., Liao, C. F., & Johns, R. (2009). *Access to destinations: Application of accessibility measures for non-auto travel modes* (No. MN/RC 2009-24). Minnesota. Dept. of Transportation. Research Services Section.
155. Kronprasert, N. (2012). *Reasoning for public transportation systems planning: Use of Dempster-Shafer theory of evidence* (Doctoral dissertation, Virginia Tech).
156. Kronprasert, N., & Kikuchi, S. (2011). Measuring validity of reasoning process for transportation planning using Bayesian inference and Dempster-Shafer theory. In *Vulnerability, Uncertainty, and Risk: Analysis, Modeling, and Management* (pp. 121-128).
157. Krykewycz, G., Pollard, C., Canzoneri, N., & He, E. (2011). Web-Based" Crowdsourcing" Approach to Improve Areawide" Bikeability" Scoring. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2245, pp. 1-7.
158. Krykewycz, G., Puchalsky, C., Rocks, J., Bonnette, B., Jaskiewicz, F., 2010. Defining a Primary Market and Estimating Demand for Major Bicycle-Sharing Program in Philadelphia, Pennsylvania. *Transportation Research Record: Journal of the Transportation Research Board,* No. 2143, pp. 117-124
159. Kuncheva, L. I. (2014). *Combining pattern classifiers: methods and algorithms*. John Wiley & Sons.
160. Kuncheva, L. I., & Rodríguez, J. J. (2014). A weighted voting framework for classifiers ensembles. *Knowledge and Information Systems*, *38*(2), 259-275.
161. Kuncheva, L., J. C. Bezdek, and R. P. W. Duin. (2001). "Decision Templates for Multiple Classifier Fusion: An Experimental Comparison." *Pattern Recognition* 34 (2): 299–314.
162. Kuzmyak, J. R., Walters, J., Bradley, M., & Kockelman, K. M. (2014). *Estimating bicycling and walking for planning and project development: A guidebook* (No. Project 08-78).

163. Lam, L., & Suen, S. Y. (1997). Application of majority voting to pattern recognition: an analysis of its behavior and performance. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, *27*(5), 553-568.

164. LaScala, E. A., Gerber, D., & Gruenewald, P. J. (2000). Demographic and environmental correlates of pedestrian injury collisions: A spatial analysis. *Accident Analysis & Prevention*, *32*, 651–658.

165. Lee, D. S. (1993). Handprinted digit recognition: A comparison of algorithms. In *Proceedings of the Third International Workshop on Frontiers in Handwriting Recognition* (pp. 153-164).

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

167. Leung, Y., Li, R., & Ji, N. (2017). Application of extended Dempster–Shafer theory of evidence in accident probability estimation for dangerous goods transportation. *Journal of Geographical Systems*, *19*(3), 249-271.

168. Levinson, D., Boies, A., Cao, J., and Fan, Y. (2016). The Transportation Futures Project: Planning for Technology Change. Minnesota Department of Transportation, St. Paul, MN.

169. Li, J. Q., Zhou, K., Shladover, S., and Skabardonis, A., (2013). *Estimating Queue Length Under the Connected Vehicle Technology: Using Probe Vehicle, Loop Detector, and Fused Data*. Transportation Research Board 92nd Annual Meeting, Washington, DC,

170. Li, L., & Ngan, C. K. (2017). A weight-adjusted-voting framework on an ensemble of classifiers for improving sensitivity. *Intelligent Data Analysis*, *21*(6), 1339-1350.

171. Li, Y., & Zheng, Y. (2019). Citywide bike usage prediction in a bike-sharing system. *IEEE Transactions on Knowledge and Data Engineering*, *32*(6), 1079-1091.

172. Lieske, S. N., Leao, S. Z., & Conrow, L. (2017) Validating Mobile Phone Generated Bicycle Route Data in Support of Active Transportation. *presented at SOAC 2017 – State of the Australian Cities Conference, 28-30 November 2017, Adelaide, Australia*

173. Lin, J.R. & Yang, T.H. (2011) Strategic design of public bicycle sharing systems with service level constrains. Transp. Res. Part E 2011, 47, 284–294.

174. Lin, X., Yacoub, S., Burns, J., & Simske, S. (2003). Performance analysis of pattern classifier combination by plurality voting. *Pattern Recognition Letters*, 24(12), 1959-1969.

175. Lindsey, G., Hankey, S., Wang, X., & Chen, J. (2013). The Minnesota bicycle and pedestrian counting initiative: Methodologies for non-motorized traffic monitoring.

176. Liu, K., Cui, M. Y., Cao, P., & Wang, J. B. (2016). Iterative Bayesian Estimation of Travel Times on Urban Arterials: Fusing Loop Detector and Probe Vehicle Data. PloS ONE, Vol. 11, No. 6, e0158123.

177. Liu, S., & Forrest, J. Y. L. (2010). *Grey systems: theory and applications*. Springer Science & Business Media.

178. Liu, W. (2006). Analyzing the degree of conflict among belief functions. *Artificial Intelligence*, *170*(11), 909-924.
179. Liu, Z., Pan, Q., Dezert, J., Han, J. W., & He, Y. (2017a). Classifier fusion with contextual reliability evaluation. *IEEE transactions on cybernetics*, *48*(5), 1605-1618.
180. Llinas, J. (1988). Toward the utilization of certain elements of ai technology for multi sensor data fusion. *Application of artificial intelligence to command and control systems, Peter Peregrinus Ltd*.
181. Llinas, J., & Hall, D. L. (1998, May). An introduction to multi-sensor data fusion. In *ISCAS'98. Proceedings of the 1998 IEEE International Symposium on Circuits and Systems (Cat. No. 98CH36187)* (Vol. 6, pp. 537-540). IEEE.
182. Lopez-Otero, P., Docio-Fernandez, L., & Garcia-Mateo, C. (2017). Ensemble audio segmentation for radio and television programmes. *Multimedia Tools and Applications*, *76*(5), 7421-7444.
183. Loukaitou-Sideris, A., Liggett, R., & Sung, H. G. (2007). Death on the crosswalk: A study of pedestrian-automobile collisions in Los Angeles. *Journal of Planning Education and Research*, *26*(3), 338-351.
184. Lowry, M., Callister, D., Gresham, M., Moore, B. 2012. Assessment of community-wide bikeability with bicycle level of service. Transportation Research Record 2314: 41-48
185. Lowry, M., McGrath, R., Cool, S., Cook, R., Skiles, M., Li, Z., ... & Hall, M. (2015). *Data Collection and Spatial Interpolation of Bicycle and Pedestrian Data* (No. 2013-M-UI-0048).
186. Lu, Z., Rao, W., Wu, Y. J., Guo, L., & Xia, J. (2015). A Kalman filter approach to dynamic OD flow estimation for urban road networks using multi-sensor data. *Journal of Advanced Transportation*, *49*(2), 210-227
187. Luo, R. and Kay, M. (1992). Data fusion and sensor integration: State-of-the-art 1990s. Data Fusion in Robotics and Machine Intelligence, pages 7-135. 9
188. Luo, R. C., & Chang, C. C. (2010). Multisensor fusion and integration aspects of mechatronics. *IEEE Industrial Electronics Magazine*, *4*(2), 20-27.
189. Luo, R. C., and Kay, M. G. (1989). Multisensor Integration and Fusion In Intelligent Systems. *IEEE Transactions on Systems, Man, and Cybernetics*, Vol.19, No. 5, pp 901-931.
190. Luo, R. C., Yih, C. C., and Su, K. L. (2002). Multisensor Fusion and Integration: Approaches, Applications, and Future Research Directions. IEEE Sensors journal, Vol. 2, No. 2, pp. 107-119.
191. Macioszek, E., Świerk, P., & Kurek, A. (2020). The bike-sharing system as an element of enhancing sustainable mobility—A case study based on a city in Poland. *Sustainability*, *12*(8), 3285.
192. Martí, F. S. (2015). *Highway travel time estimation with data fusion* (Vol. 11). Springer.

193. Martin, A., Jousselme, A. L., & Osswald, C. (2008, June). Conflict measure for the discounting operation on belief functions. In *2008 11th International conference on information fusion* (pp. 1-8). IEEE.

194. McBain, C., & Caulfield, B. (2018). An analysis of the factors influencing journey time variation in the cork public bike system. *Sustainable cities and society*, *42*, 641-649.

195. Mehran, B., & Kuwahara, M. (2013). Fusion of probe and fixed sensor data for short-term traffic prediction in urban signalized arterials. *International Journal of Urban Sciences*, *17*(2), 163-183.

196. Mekuria, M., Furth, P., Nixon, H. 2012. Low-stress bicycling and network connectivity. Mineta Transportation Institute, San José State University

197. Meng, C., Yi, X., Su, L., Gao, J., & Zheng, Y. (2017). City-wide traffic volume inference with loop detector data and taxi trajectories. In *Proceedings of the 25th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems* (p. 1). ACM.

198. Mercier, D., Lefèvre, É., & Delmotte, F. (2012). Belief functions contextual discounting and canonical decompositions. *International Journal of Approximate Reasoning*, *53*(2), 146-158.

199. Mercier, D., Quost, B., & Denœux, T. (2008). Refined modeling of sensor reliability in the belief function framework using contextual discounting. *Information fusion*, *9*(2), 246-258.

200. Meurant, G. (1992). *Data Fusion in Robotics and Machine Intelligence*. Academic Press. San Diego, CA

201. Milisavljevic, N., & Bloch, I. (2003). Sensor fusion in anti-personnel mine detection using a two-level belief function model. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, *33*(2), 269-283.

202. Miranda-Moreno, L. F., Nosal, T., Schneider, R. J., & Proulx, F. (2013). Classification of bicycle traffic patterns in five North American Cities. *Transportation research record*, *2339*(1), 68-79.

203. Misra, A., Gooze, A., Watkins, K., Asad, M., & Le Dantec, C. (2014). Crowdsourcing and its application to transportation data collection and management. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2414, pp. 1-8.

204. Mitra, S., & Washington, S. (2012). On the significance of omitted variables in intersection crash modeling. *Accident Analysis & Prevention*, *49*, 439-448.

205. Moslem, B., Khalil, M., Diab, M. O., Chkeir, A., & Marque, C. (2011, December). A multisensor data fusion approach for improving the classification accuracy of uterine EMG signals. In *2011 18th IEEE International Conference on Electronics, Circuits, and Systems* (pp. 93-96). IEEE.

206. Munira, S., & Sener, I. N. (2017). Use of the Direct-Demand Modeling in Estimating Nonmotorized Activity: A Meta-Analysis. *Safety through Disruption (Safe-D) National University Transportation Center (UTC) Program*.

207. Munira, S., & Sener, I. N. (2020). A geographically weighted regression model to examine the spatial variation of the socioeconomic and land-use factors associated with Strava bike activity in Austin, Texas. *Journal of Transport Geography*, 88, 102865.

208. Munira, S., Sener, I. N., & Dai, B. (2020). A Bayesian spatial Poisson-lognormal model to examine pedestrian crash severity at signalized intersections. *Accident Analysis & Prevention*, *144*, 105679.

209. Munira, S., Sener, I., Zhang.Y (2021). Using Bikeability Index to Improve Direct Demand Model for Estimating Bicycle Demand in The Austin Area. In press. *Journal of Urban Planning and Development*

210. Murphy, C. K. (2000). Combining belief functions when evidence conflicts. *Decision support systems*, *29*(1), 1-9

211. Nagel, J. B. (2019). *Bayesian techniques for inverse uncertainty quantification* (Vol. 504, pp. 1-205). ETH Zurich.

212. Nagi, S., & Bhattacharyya, D. K. (2013). Classification of microarray cancer data using ensemble approach. *Network Modeling Analysis in Health Informatics and Bioinformatics*, *2*(3), 159-173.

213. Nahin, P. J., & Pokoski, J. L. (1980). NCTR Plus Sensor Fusion Equals IFFN or Can Two Plus Two Equal Five?. *IEEE Transactions on Aerospace and Electronic Systems*, Vol. 3, pp. 320-337.

214. Nakamura, E. F., Loureiro, A. A., & Frery, A. C. (2007). Information fusion for wireless sensor networks: Methods, models, and classifications. *ACM Computing Surveys (CSUR)*, *39*(3), 9-es.

215. Nanthawichit, C., Nakatsuji, T., & Suzuki, H. (2003). Application of probe-vehicle data for real-time traffic-state estimation and short-term travel-time prediction on a freeway. *Transportation research record*, *1855*(1), 49-59.

216. Narayanamoorthy, S., Paleti, R., & Bhat, C. R. (2013). On accommodating spatial dependence in bicycle and pedestrian injury counts by severity level. *Transportation research part B: methodological*, *55*, 245-264.

217. Nash, A. (2009). Web 2.0 Applications for Improving Public Participation in Transport Planning. Transportation Research Board 89th Annual Meeting.

218. Nashad, T., Yasmin, S., Eluru, N., Lee, J., & Abdel-Aty, M. A. (2016). Joint modeling of pedestrian and bicycle crashes: Copula-based approach. *Transportation Research Record, 2601*, 119–127.

219. National Academies of Sciences, Engineering, and Medicine 2007. A Guidebook for Using American Community Survey Data for Transportation Planning. Washington, DC: The National Academies Press. https://doi.org/10.17226/13895

220. National Association of City Transportation Officials. 2020. *Shared Micromobility In The Us: 2019*. https://nacto.org/wp-content/uploads/2020/08/2020bikesharesnapshot.pdf

221. National Highway Traffic Safety Administration (2017). 2016 Motor Vehicle Crashes: Overview. U.S. Department of Transportation, Washington, DC.

222. National Household Travel Survey. 2018. *2018 National Household Travel Survey Workshop*. Transportation Research Board. Washington, D.C https://nhts.ornl.gov/assets/2018_NHTS_Workshop_E-Circular_238.pdf

223. Nelson, T. A., Denouden, T., Jestico, B., Laberee, K., & Winters, M. (2015). Bikemaps. Org: A Global Tool for Collision and Near Miss Mapping. Frontiers in Public Health, 3.

224. Nielsen. (2014). Hacking health: How consumers use smartphones and wearable tech to track their health. Nielsen. Retrieved 17th August 2017 from http://www.nielsen.com/us/en/insights/news/2014/hacking-healthhow-consumers-use-smartphones-and-wearable-tech-to-track-their-health.html

225. Niu, G., Han, T., Yang, B. S., & Tan, A. C. C. (2007). Multi-agent decision fusion for motor fault diagnosis. *Mechanical Systems and Signal Processing*, *21*(3), 1285-1299.

226. Noble, H., & Smith, J. (2015). Issues of validity and reliability in qualitative research. *Evidence-based nursing*, *18*(2), 34-35.

227. Nordback, K. (2012). *Estimating annual average daily bicyclists and analyzing cyclist safety at urban intersections*. Doctoral dissertation, University of Colorado at Denver.

228. Nordback, K., Kothuri, S., Figliozzi, M., Phillips, T., Gorecki, C., and Schrope, A. (2016). Investigation of Bicycle and Pedestrian Continuous and Short Duration Count Technologies in Oregon. Oregon Department of Transportation, Salem, OR.

229. Nordback, K., Marshall, W. E., Janson, B. N., & Stolz, E. (2013). Estimating annual average daily bicyclists: Error and accuracy. *Transportation research record*, *2339*(1), 90-97.

230. Olivier, B., Pierre, G., Nicolas, H., Loïc, O., Olivier, T., and Philippe, T. (2009). Multi sensor data fusion architectures for Air Traffic Control applications. Sensor and Data Fusion, In-Tech.

231. Ortuzar, J. D. D., & Willumsen, L. G. (2011). *Modelling transport* (4th ed.). Chichester, UK: John Wiley & Sons.

232. Ou, Q., Bertini, R. L., Van Lint, J. W. C., & Hoogendoorn, S. P. (2011). A theoretical framework for traffic speed estimation by fusing low-resolution probe vehicle data. *IEEE Transactions on Intelligent Transportation Systems*, *12*(3), 747-756.

233. Parikh, C. R., Pont, M. J., & Jones, N. B. (2001). Application of Dempster–Shafer theory in condition monitoring applications: a case study. *Pattern Recognition Letters*, 22(6-7), 777-785.

234. Parsons Brinckerhoff Quade & Douglas, (1999). *Development and Calibration of the Statewide Land Use-Transport Model.* Transportation and Land Use Model Integration Program. Oregon Department of Transportation

235. Pau, L. F. (1990). Behavioral knowledge in sensor/data fusion systems. *Journal of Robotic Systems*, *7*(3), 295-308.

236. Peng, D., Zuo, X., Wu, J., Wang, C., & Zhang, T. (2009). A Kalman Filter Based Information Fusion Method for Traffic Speed Estimation. In Power Electronics and Intelligent Transportation System (PEITS), 2009 2nd International Conference on (Vol. 1, pp. 374-379). IEEE.

237. Picornell, M., and Willumsen, L. (2016). *Transport Models and Big Data Fusion: Lessons from Experience*. In European Transport Conference 2016 Association for European Transport (AET).

238. Piella,G. (2003). A General Framework for Multiresolution Image Fusion: from Pixels to Regions, *Information Fusion*, Vol. 4, No. 4, pp. 259–80.

239. Pogodzinska. S., Kiec. M., D'Agostinob. C (2020). Bicycle Traffic Volume Estimation Based on GPS Data. *Transportation Research Procedia*. Vol 45, 2020, Pp 874-881

240. Polikar, R. (2006). Ensemble based systems in decision making. *IEEE Circuits and systems magazine*, *6*(3), 21-45.

241. Pushkarev, B., & Zupan, J. (1971). Pedestrian travel demand. *Transportation Research Record*, *377*, 37–53.

242. Raford, N., & Ragland, D. (2004). Space syntax: Innovative pedestrian volume modeling tool for pedestrian safety. Transportation Research Record: Journal of the Transportation Research Board, *(1878), 66-74.*

243. Raith, A. (2009). *Multiobjective routing and transportation problems* (Doctoral dissertation, ResearchSpace@ Auckland).

244. Raol, J. R. (2015). *Data fusion mathematics: theory and practice*. CRC Press.

245. Ratto, C. R., (2012) *Nonparametric Bayesian Context Learning for Buried Threat Detection*. PhD thesis, Duke University, .

246. Raudys, Š. (2006). Trainable fusion rules. II. Small sample-size effects. *Neural Networks*, *19*(10), 1517-1527.

247. Raviv, T., & Kolka, O. (2013). Optimal inventory management of a bike-sharing station. *Iie Transactions*, *45*(10), 1077-1093.

248. Raviv, T., Tzur, M., & Forma, I. A. (2013). Static repositioning in a bike-sharing system: models and solution approaches. *EURO Journal on Transportation and Logistics*, *2*(3), 187-229.
249. Reich, Y., & Barai, S. V. (1999). Evaluating machine learning models for engineering problems. *Artificial Intelligence in Engineering*, *13*(3), 257-272.
250. Ren, Y., Zhang, L., & Suganthan, P. N. (2016). Ensemble classification and regression-recent developments, applications and future directions. *IEEE Computational intelligence magazine*, *11*(1), 41-53.
251. Rixey, R., (2013). Station-Level Forecasting of Bike Sharing Ridership: Station Network Effects in Three U.S. Systems. Paper presented at the 92nd Transportation Research Board Annual Meeting 2013, Washington, DC.
252. Rodriguez, J. D., Perez, A., & Lozano, J. A. (2009). Sensitivity analysis of k-fold cross validation in prediction error estimation. *IEEE transactions on pattern analysis and machine intelligence*, *32*(3), 569-575.
253. Rogova, G. (1994). Combining the results of several neural network classifiers. *Neural Networks 7*. 777–781
254. Rogova, G. L., & Nimier, V. (2004, June). Reliability in information fusion: literature survey. In *Proceedings of the seventh international conference on information fusion* (Vol. 2, pp. 1158-1165).
255. Roli, F., Raudys, Š., & Marcialis, G. L. (2002, June). An experimental comparison of fixed and trained fusion rules for crisp classifier outputs. In *International Workshop on Multiple Classifier Systems* (pp. 232-241). Springer, Berlin, Heidelberg.
256. Romanillos, G., Zaltz Austwick, M., Ettema, D., & De Kruijf, J. (2016). Big Data and Cycling. *Transport Reviews*, Vol. 36, No. 1, pp 114-133.
257. Ross, A., & Jain, A. (2003). Information fusion in biometrics. *Pattern recognition letters*, *24*(13), 2115-2125.
258. Rothe, S., Kudszus, B., & Söffker, D. (2019). Does classifier fusion improve the overall performance? Numerical analysis of data and fusion method characteristics influencing classifier fusion performance. *Entropy*, *21*(9), 866.
259. Ruta, D., & Gabrys, B. (2000). An overview of classifier fusion methods. *Computing and Information systems*, *7*(1), 1-10.
260. Ryan, S., Appleyard, B., Schroeder, C., & Prescott, A. (2014). *Estimating Daily Bicycle Volumes Using Manual Short Duration and Automated Continuous Counts* (No. 14-5395).
261. Ryu, S., Chen, A., Christensen, K., & Choi, K. (2017a). *Development of multi-class, multi-criteria bicycle traffic assignment models and solution algorithms* (No. TRCLC 15-10). Western Michigan University. Transportation Research Center for Livable Communities.

262. Ryu, S., Chen, A., Su, J., & Choi, K. (2017b). Two-Stage Bicycle Traffic Assignment Model. *Journal of Transportation Engineering, Part A: Systems*, *144*(2), 04017079.
263. Ryus, P., Ferguson, E., Laustsen, K. M., Prouix, F. R., Schneider, R. J., Hull, T., & Miranda-Moreno, L. (2014). *Methods and technologies for pedestrian and bicycle volume data collection*. Transportation Research Board.
264. Saha, D., Alluri, P., Gan, A., & Wu, W. (2018). Spatial analysis of macro-level bicycle crashes using the class of conditional autoregressive models. *Accident Analysis & Prevention*, *118*, 166–177. doi:10.1016/j.aap.2018.02.014
265. Sanders, R. L., Frackelton, A., Gardner, S., Schneider, R., & Hintze, M. (2017). Ballpark Method for Estimating Pedestrian and Bicyclist Exposure in Seattle, Washington: Potential Option for Resource-Constrained Cities in an Age of Big Data. *Transportation Research Record: Journal of the Transportation Research Board*, (2605), 32-44.
266. Santos, A., McGuckin, N., Nakamoto, H. Y., Gray, D., & Liss, S. (2011). Summary of travel trends: 2009 national household travel survey (No. FHWA-PL-11-022). United States. Federal Highway Administration.
267. Sarkar, S., Mukherjee, K., Sarkar, S., & Ray, A. (2013). Symbolic dynamic analysis of transient time series for fault detection in gas turbine engines. *Journal of Dynamic Systems, Measurement, and Control*, *135*(1).
268. Schmitz, S., Weinmann, M., Weidner, U., Hammer, H., & Thiele, A. (2020). Automatic Generation of Training Data for Land Use and Land Cover Classification by Fusing Heterogeneous Data Sets. Publ. Der Dtsch. Ges. Für Photogramm. Fernerkund. Und Geoinf. EV, 29, 73-86.
269. Schneider, R., Arnold, L., & Ragland, D. (2009). Pilot Model for Estimating Pedestrian Intersection Crossing Volumes. Transportation Research Record 2140. pp. 13–26
270. Schoess, J., & Castore, G. (1988, August). A distributed sensor architecture for advanced aerospace systems. In *Sensor fusion* (Vol. 931, pp. 74-87). International Society for Optics and Photonics.
271. Sener, I. N., Eluru, N., & Bhat, C. R. (2009). An analysis of bicycle route choice preferences in Texas, US. *Transportation*, *36*(5), 511-539.
272. Sethi, V., Bhandari, N., Koppelman, F. S., & Schofer, J. L. (1995). Arterial incident detection using fixed detector and probe vehicle data. *Transportation Research Part C: Emerging Technologies*, *3*(2), 99-112.
273. Shafer, G. *A Mathematical Theory of Evidence*. Princeton, NJ, USA: Princeton Univ. Press, 1976.
274. Shang, Q., Yang, Z., Gao, S., & Tan, D. (2018). An imputation method for missing traffic data based on FCM optimized by PSO-SVR. *Journal of Advanced Transportation*, *2018*.

275. Shen, Q., Chen, P., Schmiedeskamp, P., Bassok, A., & Childress, S. (2014). *Bicycle Route Choice: GPS Data Collection and Travel Model Development* (No. 2012-S-UW-0019).
276. Shen, Y., Zhang, X., & Zhao, J. (2018). Understanding the usage of dockless bike sharing in Singapore. *International Journal of Sustainable Transportation, 12(9),* 686-700
277. Shi, C., Chen, B. Y., Lam, W. H., & Li, Q. (2017). Heterogeneous data fusion method to estimate travel time distributions in congested road networks. *Sensors*, *17*(12), 2822.
278. Singleton, P. A., & Clifton, K. J. (2013). Pedestrians in regional travel demand forecasting models: state-of-the-practice. In *92nd Annual Meeting of the Transportation Research Board, Washington, DC* (pp. 13-4857).
279. Smets, P. (1993). Belief functions: the disjunctive rule of combination and the generalized Bayesian theorem. *International Journal of approximate reasoning*, *9*(1), 1-35.
280. Smets, P., & Kennes, R. (1994). The transferable belief model. *Artificial intelligence*, *66*(2), 191-234.
281. Sohn, S. Y., & Lee, S. H. (2003). Data fusion, ensemble and clustering to improve the classification accuracy for the severity of road traffic accidents in Korea. *Safety Science*, *41*(1), 1-14.
282. Song, Q., Shepperd, M., & Mair, C. (2005, September). Using grey relational analysis to predict software effort with small data sets. In *11th IEEE International Software Metrics Symposium (METRICS'05)* (pp. 10-pp). IEEE.
283. Soua, R., Koesdwiady, A., & Karray, F. (2016). Big-data-generated traffic flow prediction using deep learning and dempster-shafer theory. In *2016 International joint conference on neural networks (IJCNN)* (pp. 3195-3202). IEEE.
284. Steinberg, A.N., Bowman, C. L. and White, F. E. (1999). "Revisions to the JDL Data Fusion model," NSSDF, pp. 235 - 251.
285. Stinson, M., & Bhat, C. (2003). Commuter bicyclist route choice: Analysis using a stated preference survey. *Transportation research record: journal of the transportation research board*, (1828), 107-115.
286. Stolkin, R, Rees, D., Talha, M., and Florescu, I. (2012). Bayesian fusion of thermal and visible spectra camera data for region based tracking with rapid background adaptation. In *Multisensor Fusion and Integration for Intelligent Systems (MFI), 2012 IEEE Conference on,* pages 192-199. IEEE.
287. Strauss, J., & Miranda-Moreno, L. F. (2017). Speed, travel time and delay for intersections and road segments in the Montreal network using cyclist Smartphone GPS data. *Transportation research part D: transport and environment*, *57*, 155-171.
288. Strauss, J., Miranda-Moreno, L. F., & Morency, P. (2014). Multimodal injury risk analysis of road users at signalized and non-signalized intersections. *Accident Analysis & Prevention*, 71, 201-209.

289. Strauss, J., Miranda-Moreno, L. F., and Morency, P. (2015). Mapping Cyclist Activity and Injury Risk in a Network Combining Smartphone GPS Data and Bicycle Counts. *Accident Analysis and Prevention,* Vol. 83, pp. 132–142.

290. StreetLight. 2017. *Location-Based Services Data is Better than Cellular for Transportation – and Especially for Rural Areas.* https://www.streetlightdata.com/location-based-services-data-transportation-rural-studies/?type=blog

291. StreetLight. 2018. *StreetLight Active Mode Methodology, Data Sources and Validation Version 1.0.* Technical Report. https://www.streetlightdata.com/wp-content/uploads/StreetLight_Active-Mode_Methodology-and-Validation_190108.pdf

292. StreetLight. 2019. This glowing map shows the most-biked streets in San Francisco. https://www.streetlightdata.com/fast-company-virtual-bike-and-pedestrian-sensors-data/

293. Su, J. G., Winters, M., Nunes, M., & Brauer, M. (2010). Designing a route planner to facilitate and promote cycling in Metro Vancouver, Canada. *Transportation research part A: policy and practice*, *44*(7), 495-505.

294. Suen, C. Y. (1990). Recognition of totally unconstrained handwritten numerals based on the concept of multiple experts. In *Proc. 1st International Workshop on Frontiers in Handwriting Recognition, April 1990* (pp. 131-143).

295. Suen, C. Y., & Lam, L. (2000, June). Multiple classifier combination methodologies for different output levels. In *International workshop on multiple classifier systems* (pp. 52-66). Springer, Berlin, Heidelberg.

296. Sun, Y., Du, Y., Wang, Y., & Zhuang, L. (2017). Examining associations of environmental characteristics with recreational cycling behaviour by street-level Strava data. *International Journal of Environmental Research and Public Health, 14*(6), 644. doi:10.3390/ijerph14060644

297. Sunberg, Z., & Rogers, J. (2013). A belief function distance metric for orderable sets. *Information Fusion*, *14*(4), 361-373.

298. Swanson, K. (2012). *Bicycling and Walking in the United States Benchmarking Report. Alliance for Biking and Walking*, Washington, D.C.

299. Tabeshian, M., & Kattan, L. (2014). Modeling nonmotorized travel demand at intersections in Calgary, Canada: Use of traffic counts and geographic information system data. *Transportation Research Record*, *2430*, 38–46. doi:10.3141/2430-05

300. Texas Department of Transportation (TxDOT) (2016). State of Texas Classification of Motor Vehicle Traffic Crashes in Texas: 2016 Edition Retrieved from (2016) https://ftp.dot.state.tx.us/pub/txdot-info/library/forms/cit/crash102_0616.pdf

301. Texas Department of Transportation (TxDOT). (2016a). *Texas Intersection Safety Implementation Plan: Preliminary Findings for Texas's Capital Area Metropolitan Planning Organization*. Retrieved from https://www.texasshsp.com/wp-content/uploads/2017/02/Preliminary-Findings_CAMPO_2016-06-15.pdf

302. Texier, G., Allodji, R. S., Diop, L., Meynard, J. B., Pellegrin, L., & Chaudet, H. (2019). Using decision fusion methods to improve outbreak detection in disease surveillance. *BMC medical informatics and decision making*, *19*(1), 1-11.

303. The National Highway Traffic Safety Administration [NHTSA]. (2020). *Preview of Motor Vehicle Traffic Fatalities In 2019*. TRAFFIC SAFETY FACTS. Research Note. https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/813021

304. Thiel, C., Schwenker, F., & Palm, G. (2005, June). Using dempster-shafer theory in mcf systems to reject samples. *In International Workshop on Multiple Classifier Systems (pp. 118-127).* Springer, Berlin, Heidelberg.

305. Tin, S. T., Woodward, A., Robinson, E., & Ameratunga, S. (2012). Temporal, seasonal and weather effects on cycle volume: an ecological study. *Environmental Health*, *11*(1), 1-9.

306. Town of Blacksburg. 2015. BLACKSBURG BICYCLE MASTER PLAN 2015. https://www.blacksburg.gov/home/showpublisheddocument?id=5413

307. Treiber, Martin, Arne Kesting, and R. Eddie Wilson. (2011) "Reconstructing the traffic state by fusion of heterogeneous data." *Computer-Aided Civil and Infrastructure Engineering* 26, no. 6: 408-419.

308. Tulyakov, S., Jaeger, S., Govindaraju, V., & Doermann, D. (2008). Review of classifier combination methods. In *Machine learning in document analysis and recognition* (pp. 361-386). Springer, Berlin, Heidelberg.

309. Turner, S. M., Sener, I. N., Martin, M. E., White, L. D., Das, S., Hampshire, R. C., ... & Wijesundera, R. K. (2018). *Guide for scalable risk assessment methods for pedestrians and bicyclists* (No. FHWA-SA-18-032). United States. Federal Highway Administration. Office of Safety.

310. Turner, S., Martin, M., Griffin, G., Le, M., Das, S., Wang, R., ... & Li, X. (2020). Exploring Crowdsourced Monitoring Data for Safety.

311. Turner, S., Sener, I. N., Martin, M., Das, S., Shipp, E., Hampshire, R., Fitzpatrick, K., Molnar, L., Wijesundera, R., Colety, M., and Robinson, S. (2017). *Synthesis of Methods for Estimating Pedestrian and Bicyclist Exposure to Risk at Areawide Levels and on Specific Transportation Facilities.* U.S. Department of Transportation, Federal Highway Administration, Office of Safety, Washington, DC.

312. Turverey, R. J., Cheng, D. D., Blair, O. N., Roth, J. T., Lamp, G. M., & Cogill, R. (2010, April). Charlottesville bike route planner. In *Systems and Information Engineering Design Symposium (SIEDS), 2010 IEEE* (pp. 68-72). IEEE

313. U.S. Department of Defence (1991). Data Fusion Subpanel of the Joint Directors of Laboratories, Technical Panel for C3, "Data fusion lexicon,"

314. Van Lint, J. W. C., and Hoogendoorn, S. P. (2010). A Robust and Efficient Method for Fusing Heterogeneous Data from Traffic Sensors On Freeways. *Computer-Aided Civil and Infrastructure Engineering*, Vol. *25, No.* 8, pp 596-612.

315. Virani, N. (2017). Learning Data-Driven Models for Decision-Making in Intelligent Physical Systems.

316. Virani, N., Sarkar, S., Lee, J. W., Phoha, S., & Ray, A. (2016). Algorithms for context learning and information representation for multi-sensor teams. In *Context-Enhanced Information Fusion* (pp. 403-427). Springer, Cham.

317. WALD, L. (1999). Some terms of reference in data fusion. *IEEE Trans. Geosci. Remote Sens. 13,* 3 (May), 1190–1193.

318. Waltz, E., & Waltz, T. (2009). The principles and practice of image and spatial data fusion. In *Proc. 8th Natl. Data Fusion Conf*.

319. Wang, S. (2016). *Surveillance video data fusion* (Doctoral dissertation, Kingston University).

320. Wang, X., & Kockelman, K. M. (2009). Forecasting network data: Spatial interpolation of traffic counts from texas data. *Transportation Research Record*, *2105*(1), 100-108.

321. Wang, X., Lindsey, G., Schoner, J., Harrison, A., (2012). Modeling Bike Share Station Activity: The Effects of Nearby Businesses and Jobs on Trips to and from Stations. Paper presented at the 92nd Transportation Research Board Annual Meeting 2012, Washington, DC.

322. Wang, X., Yang, J., Lee, C., Ji, Z., & You, S. (2016). Macro-level safety analysis of pedestrian crashes in Shanghai, China. *Accident Analysis & Prevention, 96*, 12–21.

323. Wang, Y., Kockelman, K., & Jamali, A. (2017). A synthesis of spatial models for multivariate count responses. In R. Jackson & P. Schaeffer (Eds.), *Regional research frontiers* (Vol. 2, pp. 221–237). Cham, Switzerland: Springer

324. Wegener, M. (2004). Overview of land-use transport models. *Handbook of transport geography and spatial systems*, *5*, 127-146.

325. Wei, Q., Dobigeon, N., & Tourneret, J. Y. (2015). Bayesian fusion of multi-band images. *IEEE Journal of Selected Topics in Signal Processing*, *9*(6), 1117-1127.

326. Wen, C., Wang, Y., & Xu, X. (2008, September). Fuzzy information fusion algorithm of fault diagnosis based on similarity measure of evidence. In *International Symposium on Neural Networks* (pp. 506-515). Springer, Berlin, Heidelberg.

327. White, F. E. (1991). *Data fusion lexicon*. Joint Directors of Labs Washington DC.

328. Whitfield, G. P., McKenzie, B., Graff, K. A., & Carlson, S. A. (2020). Peer Reviewed: Monitoring State-Level Changes in Walking, Biking, and Taking Public Transit to Work—American Community Survey, 2006 and 2017. *Preventing chronic disease*, *17*.

329. Wilson, N. (1993, January). The assumptions behind Dempster's rule. In *Uncertainty in Artificial Intelligence* (pp. 527-534). Morgan Kaufmann.

330. Winters, M., Brauer, M., Setton, E. M., & Teschke, K. (2013). Mapping bikeability: A spatial tool to support sustainable travel. *Environment and Planning B: Planning and Design*, *40*, 865–883. doi:10.1068/b38185

331. Winters, M., Davidson, G., Kao, D., & Teschke, K. (2011). Motivators and deterrents of bicycling: comparing influences on decisions to ride. *Transportation*, *38*(1), 153-168.

332. WSDOT, 2014. *Methods for Estimating Bicycling and Walking In Washington State.* Oregon Transportation Research and Education Consortium. *https://www.wsdot.wa.gov/research/reports/fullreports/828.1.pdf*

333. Wu, C., Thai, J., Yadlowsky, S., Pozdnoukhov, A., & Bayen, A. (2015). Cellpath: Fusion of Cellular and Traffic Sensor Data for Route Flow Estimation Via Convex Optimization. *Transportation Research Procedia*, Vol. 7, pp. 212-232.

334. Wu, H. (2003). *Sensor data fusion for context-aware computing using dempster-shafer theory* (Doctoral dissertation, Carnegie Mellon University, the Robotics Institute).

335. Xu, L., Krzyzak, A., & Suen, C. Y. (1992). Methods of combining multiple classifiers and their applications to handwriting recognition. *IEEE transactions on systems, man, and cybernetics*, *22*(3), 418-435.

336. Xu, P. (2014). *Information fusion for scene understanding* (Doctoral dissertation).

337. Yager, R. R. (1987). On the Dempster-Shafer framework and new combination rules. *Information sciences*, 41(2), 93-137.

338. Yang, H., Xia, K., Anqi, B., & Qian, P. (2019, December). New Applications of Late Fusion Methods for EEG Signal Processing. In *2019 International Conference on Computational Science and Computational Intelligence (CSCI)* (pp. 617-621). IEEE.

339. Ye, F., Chen, J., Li, Y., & Kang, J. (2016). Decision-making algorithm for multisensor fusion based on grey relation and DS evidence theory. *Journal of Sensors*, *2016*.

340. Yong, D., WenKang, S., ZhenFu, Z., & Qi, L. (2004). Combining belief functions based on distance of evidence. *Decision support systems*, *38*(3), 489-493.

341. Yunlong, W., Kai, L., Guan, G., Yanyun, Y., & Fei, L. (2019). Evaluation method for Green jack-up drilling platform design scheme based on improved grey correlation analysis. *Applied Ocean Research*, *85*, 119-127.

342. Zhang, J. (2010). Multi-Source Remote Sensing Data Fusion: Status and Trends. International *Journal of Image and Data Fusion*, Vol. 1, No. 1, pp 5-24.

343. Zhang, L. (2009). Context-dependent fusion with application to landmine detection.

344. Zhang, W., Su, Y., Min, J., Teng, R., & Zhou, Y. (2012, August). Application of grey relation degree in rotor's fault identification. In *2012 International Symposium on*

Instrumentation & Measurement, Sensor Network and Automation (IMSNA) (Vol. 1, pp. 183-185). IEEE

345. Zhang, X., Pérez-Stable, E. J., Bourne, P. E., Peprah, E., Duru, O. K., Breen, N., ... & Denny, J. (2017). Big data science: opportunities and challenges to address minority health and health disparities in the 21st century. *Ethnicity & disease*, *27*(2), 95.

346. Zhou, X., Sharma, S., & Peeta, S. (2011). Development of a Mobile Probe-Based Traffic Data Fusion and Flow Management Platform for Innovative Public-Private Information-Based Partnerships. USDOT Region V Regional University Transportation Center Final Report.

347. Zhou, Z. (2015). A Multi-Sensor Data Fusion Approach for Real-Time Lane-Based Traffic Estimation. Doctoral dissertation. Arizona State University.

348. Zhu, S., Guo, Y., Zheng, H., Peeta, S., Ramadurai, G., & Wang, J. (2016). *Field Data Based Data Fusion Methodologies to Estimate Dynamic Origin-Destination Demand Matrices from Multiple Sensing and Tracking Technologies* (No. NEXTRANS Project No. 109PUY2. 1).