Multi-Sensor Data Fusion for Traffic Speed and Travel Time Estimation

by

Christian Bachmann

A thesis submitted in conformity with the requirements
for the degree of Master of Applied Science

Department of Civil Engineering
University of Toronto

Your file Votre référence
ISBN: 978-0-494-76428-2

Our file Notre référence
ISBN: 978-0-494-76428-2

Multi-Sensor Data Fusion for Traffic Speed and Travel Time Estimation

Christian Bachmann
Master of Applied Science
Department of Civil Engineering
University of Toronto
2011

Abstract

In this thesis, seven multi-sensor data fusion based estimation techniques are investigated. All methods are compared in terms of their ability to fuse data from loop detectors and Bluetooth tracked probe vehicles to accurately estimate freeway traffic speed. In the first case study, data generated from a microsimulation model are used to assess how data fusion might perform with present day conditions, having few probe vehicles, and what sort of improvement might result from an increased proportion of vehicles carrying Bluetooth-enabled devices in the future. In the second case study, data collected from the real-world Bluetooth traffic monitoring system are fused with corresponding loop detector data and the results are compared against GPS collected probe vehicle data, demonstrating the feasibility of implementing data fusion for real-time traffic monitoring today. This research constitutes the most comprehensive evaluation of data fusion techniques for traffic speed estimation known to the author.

Acknowledgments

Writing acknowledgments is always an enjoyable task; not only because it allows for reflection on all those who have been helpful but also because you can be sure it's one part that someone will actually read. Being the last and final task probably makes it somewhat relieving as well.

First and foremost, I am deeply indebted to my two supervisors, Professor Matthew Roorda and Professor Baher Abdulhai. I cannot express my gratitude enough for all that both of you have done for me over these past few years. Baher: Thank you for taking a chance on me in my undergraduate years and allowing me to get an early start in research; you were the inspiration. Matt: Thank you for your constant presence and enthusiasm since day one; you are a role model for all of us. I hope we are all proud of our joint accomplishments.

Thank you also to Professor Behzad Moshiri from the University of Tehran who was visiting us during some of this work. Your knowledge of data fusion pointed me in the right direction and got things moving along quickly.

I gratefully acknowledge Professor Bruce Hellinga, Pedram Izadpanah, and the team of students from the University of Waterloo who undertook the probe vehicle data collection effort used in the sixth chapter of this thesis.

I feel as though none of this work would have been possible without all of my fellow transportation graduate students. I am very thankful for all of the help I received in the ITS lab throughout the progression of this work. I will never forget that many of you acted as surrogate supervisors to me. I am equally thankful for all of the laughs and lunches we have shared together. I will always look back fondly on my time spent in the ITS lab.

Very special thanks go to all of the students and staff at Chestnut Residence. Being a don throughout my Master's has been an educational experience all in itself. The friends and memories I have made at Chestnut have undoubtedly changed my life in ways I'm not yet sure I fully understand. Thank you to the students on my floor for motivating me and inspiring me with your hard work and dedication. Thank you to my fellow dons for making my life exciting when school was not. The Chestnut Residence community has helped me greatly; I hope I have helped it too.

Of course, I would like to acknowledge my family and friends for their constant love and support. My life has always been blessed with company far better than I deserve and I am always grateful for that.

Funding for this Master's was provided in part by the Social Sciences and Humanities Research Council of Canada (SSHRC) in the form of a Canada Graduate Scholarship (CGS) and by the Ontario Ministry of Training, Colleges, and Universities in the form of an Ontario Graduate Scholarship (OGS).

Table of Contents

List of Tables

List of Figures

List of Appendices

List of Acronyms

ADAS = Advanced Driver Assistance

ADVANCE = Advanced Driver and Vehicle Advisory Navigation Concept

AGV = Autonomous Guided Vehicles

AID = Automatic Incident Detection

ATIS = Advanced Traveler Information Systems

AVL = Automatic Vehicle Location

DSER = Dempster-Schafer Evidential Reasoning

EOBR = Electronic On Board Recorder

FEFM = Federated Evidence Fusion Model

GEP = Generalized Evidence Processing

GPS = Global Positioning System

ILD = Inductive Loop Detector

ITS = Intelligent Transportation Systems

KF = Kalman Filter

LS = Least Square

MAC = Media Access Control

MAE = Mean Absolute Error

MAPE = Mean Absolute Percentage Error

MARE = Mean Absolute Relative Error

ME = Mean Error

ML = Maximum Likelihood

MRE = Mean Relative Error

MSDE = Mean State Decision Error

MSE	= Mean Squared Error
MTO	= Ministry of Transportation Ontario
OWA	= Ordered Weighted Averaging
RME	= Relative Mean Error
RMSE	= Root Mean Squared Error
SCAAT	= Single-Constraint-At-A-Time
TMC	= Traffic Management Center
VDS	= Vehicle Detector Station

Chapter 1
Introduction

"The beginning is the most important part of the work."

– Plato

1.1 Bluetooth Traffic Monitoring

Travel time measurement in real-time is a major function of Intelligent Transportation Systems. Recently, there has been an interest in developing an anonymous probe vehicle monitoring system to measure travel times on highways and arterials based on wireless signals available from technologies such as Bluetooth. The majority of consumer electronic devices produced today come equipped with Bluetooth wireless capability to communicate with other devices in close proximity. For example, it is the primary means to enable hands-free use of cell phones. Bluetooth enabled devices can communicate with other Bluetooth enabled devices anywhere from 1 meter (class 3) to 100 meters (class 1), depending on the power rating of the Bluetooth in the devices. The Bluetooth protocol uses an electronic identifier in each device called a Media Access Control (MAC) address. By mounting a simple antenna adjacent to the roadway (Figure 1-1), MAC addresses for visible devices can be easily logged and time-stamped. If these MAC addresses are consecutively logged at multiple stations, the unique MAC addresses can be matched, and the difference in time stamps can be used to estimate travel times (Figure 1-2). Studies in Maryland (Young, 2008) and Indiana (Wasson, Sturdevant, & Bullock, 2008) show the feasibility of using such a system to estimate travel time.

A real-time traffic monitoring system is currently under development in Toronto, Canada, which detects Bluetooth-enabled devices travelling past roadside receivers, allowing for the aforementioned method of travel time estimation. The system also makes use of RouteTrackers, which are currently installed in over 20,000 trucks in over 250 firms. A RouteTracker is a Global Positioning System (GPS) tracking device that is connected directly to the vehicles engine computer. Engine diagnostic reporting combined with GPS data is gathered to help fleets optimize their operations and automate regulatory compliance (Xata Turnpike, 2010). These RouteTrackers download GPS data wirelessly to roadside receivers in *pseudo* real time along freeways throughout Toronto. It is noted that the GPS data obtained by this system are only "real-time" in the sense that closely spaced stations provide very frequent updates of recent truck location and speed data. A preliminary analysis of this system by Roorda et al. (2009) showed that travel time estimates can be obtained by observing either RouteTracker-enabled trucks or other vehicles carrying a Bluetooth device at consecutive locations on the highway.

Figure 1-1: Bluetooth station installation (Roorda et al., 2009)

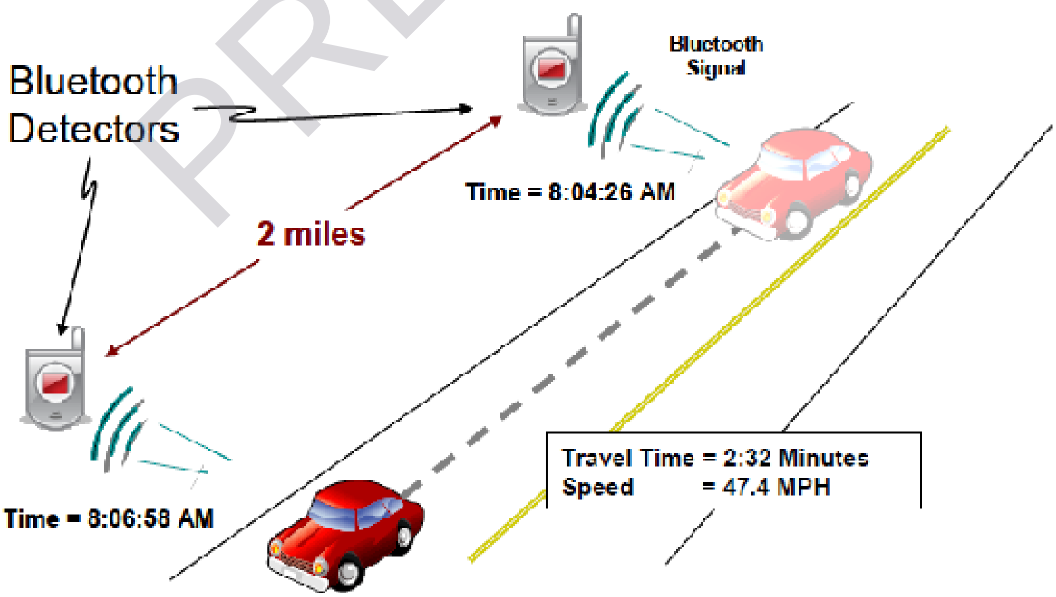

Figure 1-2: Bluetooth traffic monitoring operation concept (Young, 2008)

1.2 Loop Detectors

More traditionally, vehicle detector stations (VDS) are the major elements of a freeway traffic management system. Inductance loops are the most widely used detectors in freeway traffic management systems because of their reliability in data measurements and flexibility in design (Ministry of Transportation, 2010).

As the name suggests, the main function of loop detectors is to detect the passage and presence of vehicles on the freeway. Data collected at vehicle detector stations are initially processed by a micro processor located at the side of the highway. The processed data contain traffic volumes, vehicle speeds, road occupancy, and vehicle length information. This information is then transmitted at regular intervals to a Traffic Management Center (TMC) via a communications system. The computer system uses the data to monitor traffic patterns and also attempts to identify traffic incidents as they occur.

An inductance loop detector system basically consists of three components: a loop embedded in the pavement consisting of multiple turns of wire; a lead-in cable which connects the loop wire to the input of the loop detector amplifier; and a detector amplifier that intensifies the electrical energy produced by the detector loop (Ministry of Transportation, 2010)..

Many vehicle detector stations in the Greater Toronto Area have a double-loop arrangement and are specifically designed to measure vehicle speeds and lengths in addition to the traffic volumes and occupancy information. Stations with one loop per lane are capable of directly measuring traffic volumes and occupancy information only, from which speed can be estimated (Ministry of Transportation, 2010).

1.3 Research Questions and Objectives

Clearly, both Bluetooth device tracking and loop detectors provide a means of traffic speed and travel time estimation. The co-existence of these systems begs a number of research questions. Most obviously, how do these systems compare in terms of their ability to estimate freeway traffic speeds and travel times? Furthermore, is one system superior to the other? Bluetooth device monitoring is a relatively new method for traffic monitoring, and accordingly, its performance is not well known. Moreover, there has been no attempt thus far to compare it with monitoring via loop detectors.

An even more interesting question is how to fuse data from these systems together. The rationale for fusion is strong when you consider how the sensors complement one another. The data from loop detectors cover almost all the vehicles that have travelled on the road section, resulting in excellent temporal sampling and resolution. However, these measurements can be imprecise and the spatial sampling depends on the sensor spacing. Moreover, such measurements typically only represent traffic speed at the location of the sensor and not over the entire link. On the other hand, probe vehicles can be more accurate, although of variable quality, and with good spatial coverage. They describe the state of traffic on the entire road link, but are not exhaustive as they are only a small portion of the vehicles that make up all of traffic in the network. With this complementarity in mind, one can imagine how fusing data from these sources together might enhance a traffic monitoring system.

However, considering data fusion adds further questions. First, is there a best way of fusing data from these systems together? Once the data have been fused, is there an improvement in accuracy? If there is no improvement in accuracy, should we bother with data fusion? Another interesting point lies in the number of probe vehicles captured by the Bluetooth traffic monitoring system. More specifically, how does the number of probe vehicles captured affect the accuracy of the system and of the fusion result? Finally, with all of the other questions answered, how feasible is implementing a data fusion based system today?

With these questions in mind, the research objectives of this thesis are as follows:

- Compare the accuracies of loop detectors and Bluetooth traffic monitoring.
- Identify all of the data fusion techniques that could be used to fuse loop detectors and probe vehicle estimates.
- Compare the applicable data fusion methods in their ability to fuse loop detector data and probe vehicle data to accurately estimate freeway traffic speeds.
- Using microsimulation scenarios, investigate how the number of probe vehicles captured by the Bluetooth traffic monitoring system affects its accuracy and the subsequent fusion estimate.
- Fuse real-world data coming from the Bluetooth traffic monitoring system and corresponding loop detector data and compare against GPS collected data to determine if these techniques are “practice-ready”.

This thesis is the first attempt to fuse a Bluetooth traffic monitoring system with loop detectors. Furthermore, it constitutes the most comprehensive evaluation of data fusion techniques for traffic speed estimation known to the author.

1.4 Thesis Structure

Chapter 1 introduced the underlying motivations and research objectives. Chapter 2 provides the relevant background information to familiarize the reader with the field of multi-sensor data fusion with the intent of making the remainder of the thesis comprehensible. Chapter 3 provides an introduction to data fusion applications in transportation engineering and a comprehensive literature review of data fusion research conducted in traffic speed and travel time estimation. Chapter 4 presents the mathematical details of the data fusion techniques utilized in this research; every effort was made to make this chapter as self contained as possible, but the reader may find they require additional texts to aid in understanding the complex portions of this material (which can be found in the cited references). Chapter 5 provides the results of a microsimulation case study of Highway 400 in Toronto, Canada. Chapter 6 provides the results of fusing real-world data from the Bluetooth traffic monitoring system with corresponding loop detector data on Highway 401, also in Toronto. Chapter 7 concludes with a summary of key findings and directions for future work.

Chapter 2
Background

"Nature provides the main inspiration in designing intelligent systems."

– G.W Ng

2.1 What is Data Fusion?

The general concept of multi-sensor data fusion is analogous to the manner in which humans and animals use a combination of multiple senses, experience, and the ability to reason to improve their chances of survival (Mitchell, 2007). In particular, the brain fuses information from our surrounding environment and attempts to derive knowledge, draw conclusions or inferences from the fused information (Ng, 2003). For example, consider how many sensors are used by a human being when eating. Assessing the quality of an edible substance may not be possible using only the sense of vision; the combination of sight, touch, smell, and taste is far more effective (Hall & Llinas, 2001).

While there is not one commonly referenced definition of data fusion, there is a general consensus of what fusing data means. Mitchell (2007) suggests that multi-sensor data fusion is "the theory, techniques and tools which are used for combining sensor data, or data derived from sensory data, into a common representational format…in performing sensor fusion our aim is to improve the quality of the information, so that it is, in some sense, *better* than would be possible if the data sources were used individually." Hall & Llinas (2001) propose "data fusion techniques combine data from multiple sensors and related information to achieve more specific inferences than could be achieved by using a single, independent sensor." Ng (2003) provides the simplest definition, stating that "fusion involves the combination of data and information from more than once source." As can be seen from these three definitions, there is a common understanding that data fusion encompasses a wide variety of activities that involve using multiple data sources. Unfortunately, the universality of data fusion has engendered a profusion of overlapping research and development in many applications. A jumble of confusing terminology (Figure 2-1) and ad hoc methods in a variety of scientific, engineering, management, and educational disciplines obscures the fact that the same ground has been covered repeatedly (Hall & Llinas, 2001).

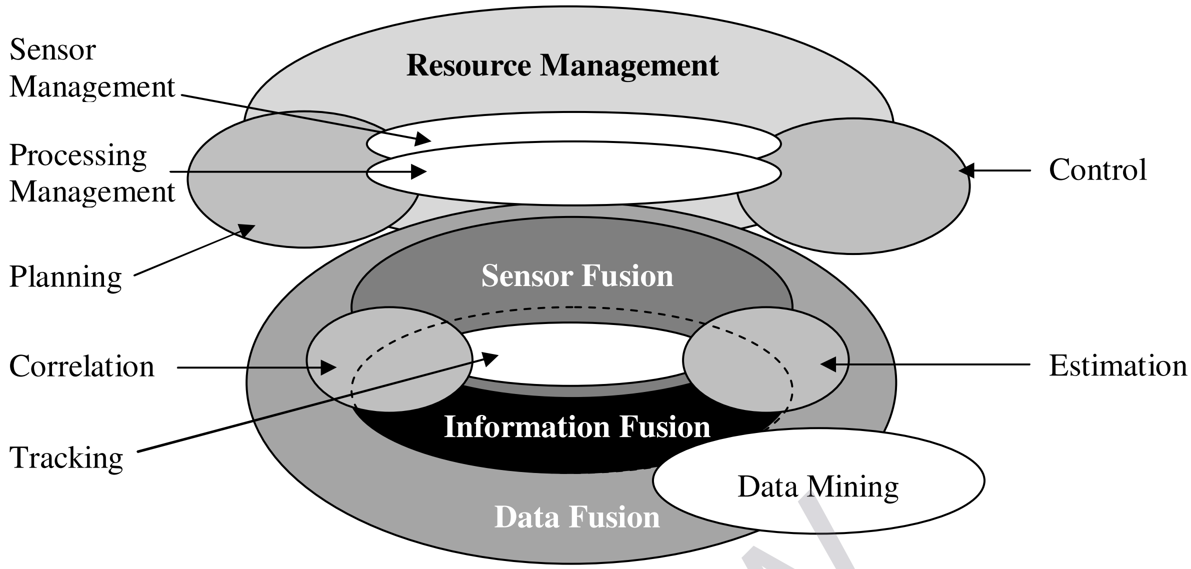

Figure 2-1: (Con)fusion of terminology – adapted from (Hall & Llinas, 2001)

2.2 Importance of Data Fusion – Why Fuse Data?

There are many reasons why we need data and information fusion systems as noted by Mitchell (2007), Ng (2003), Hall & Llinas (2001), Brooks & Iyengar (1998) and Luo & Kay (1989):

- **Reliability/Robustness/Redundancy**: A system that depends on a single source of input is not robust in the sense that if the single source fails to function properly, the whole system operation will fail. However, the system fusing several sources of data has a higher fault-tolerance since multiple sensors providing redundant information serve to increase reliability in the case of sensor error or failure.
- **Accuracy/Certainty**: Combining readings from several different kinds of sensors can give a system more accurate information. Combining several readings from the same sensor makes a system less sensitive to noise and temporary glitches. Therefore, multiple independent sources of data can not only help improve accuracy, but can also add certainty by removing ambiguity in the data.
- **Completeness/Coverage/Complementarity**: More data sources will provide extended coverage of information on an observed object or state. Extended coverage is particularly relevant in spatial and temporal environments for the sake of completeness. Sometimes information from multiple sensors is complementary and allows features in the environment to be perceived that are impossible to perceive using just the information from each individual sensor operating separately (see section 2.3.3).

- **Cost effectiveness**: To build a single sensor that can perform multiple functions is often more *expensive* than to integrate several simple and cheap sensors with specific functions.
- **Representation**: Another problem that sensor fusion attempts to address is information overload. The amount of time required to make a decision increases rapidly as the amount of information available increases. Sensor fusion is necessary to combine information and clearly present the best interpretation of the sensor data to allow for a well informed and timely decision.
- **Timeliness**: More timely information may be provided by multiple sensors due to either the actual speed of operation of each sensor, or the processing parallelism that may be possible, as compared to the speed at which it could be provided by a single sensor.

Of course, all of these benefits hinge on the assumption that there is no single perfect source of information. This assumption is well made since all sensors have a few things in common: every sensor device has a limited accuracy, limited coverage, is subject to the effect of some type of noise, and will under some conditions function incorrectly. Hence, there is no single perfect source of information.

2.3 On the Use of Multiple Sensors

Durrant-Whyte (1988) first classified a multi-sensor data fusion system according to its sensor configuration. The typology proposed by this author gained popularity and is now widely used in the data fusion research community. The three basic types of configurations are: complementary, competitive, and cooperative. While these divisions are defined by the functionality of the sensor network, they are not necessarily mutually exclusive.

2.3.1 Complementary

A sensor configuration is called complementary if the sensors do not directly depend on each other, but can be combined in order to give a more complete image of the phenomenon under study. Complementary sensors help resolve the problem of incompleteness. As a simple example, Figure 2-2 shows a temperature monitoring system that consists of several thermometers each covering a different region. This configuration is complementary because each thermometer provides the same type of data but for a different geographic region. In general, fusing complementary data is intuitive and easy.

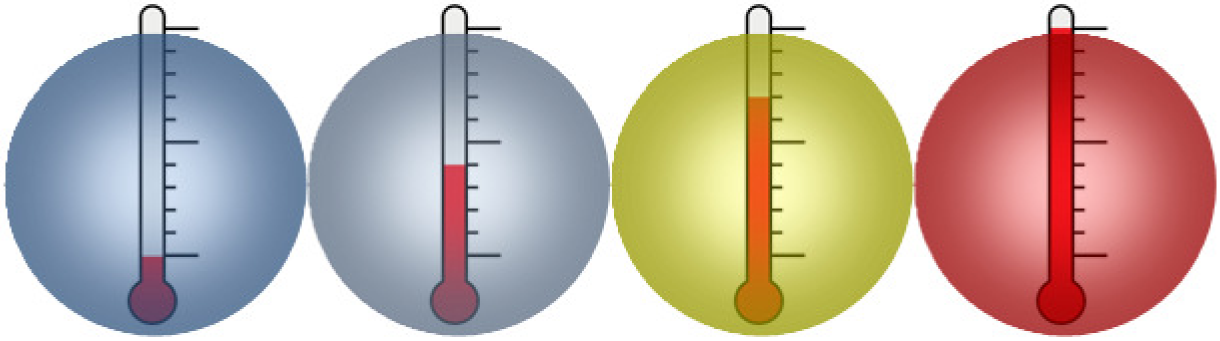

Figure 2-2: A complementary sensor network may consist of several thermometers, each covering a different geographical region (note there is no overlap in coverage)

2.3.2 Competitive

A sensor configuration is competitive if each sensor delivers an independent measurement of the same property. Since they provide what should be identical data, the sensors are in competition as to which reading should be believed by the system in the case of discrepancies. Competing sensors can be identical or they can use different methods of measuring the same attribute. The aim of competitive fusion is reduce the effect of uncertain and erroneous measurements, provide greater reliability, and/or add fault tolerance to a system. Figure 2-3 shows three thermometers partially surveying the same region (shaded darker). Note that this type of configuration would still be able to function for the joined region if one of the thermometers were to cease functioning.

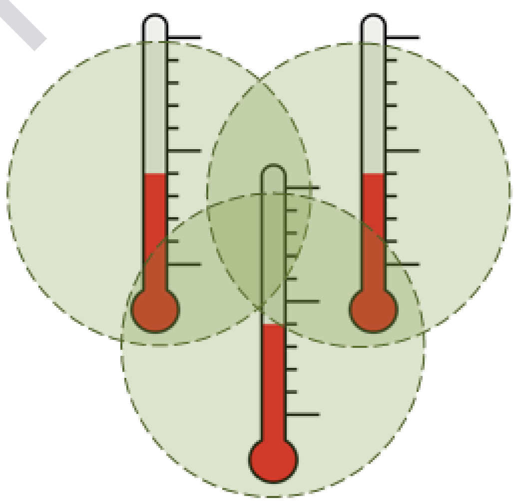

Figure 2-3: Competitive thermometers would all return information regarding the same region (note the overlap in coverage)