UC Berkeley
Dissertations

Title

Design and Development of Novel Routing Methodologies for Dynamic Roadway Navigation Systems

Permalink

https://escholarship.org/uc/item/8d72371n

Author

Zhu, Weihua

Publication Date

2009-09-01

Design and Development of Novel Routing Methodologies for Dynamic Roadway Navigation Systems

Weihua Zhu
University of California, Riverside
2009

UNIVERSITY OF CALIFORNIA
RIVERSIDE

Design and Development of Novel Routing Methodologies for Dynamic
Roadway Navigation Systems

A Dissertation submitted in partial satisfaction
of the requirement for the degree of

Doctor of Philosophy
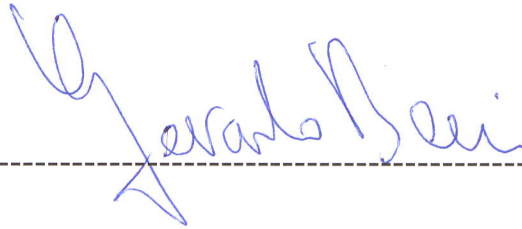
in

Electrical Engineering
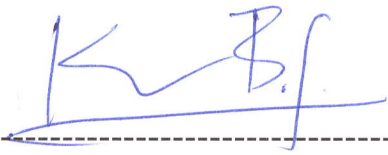
by

Weihua Zhu

March 2009

Dissertation Committee:

    Professor Matthew Barth, Chairperson
    Professor Gerardo Beni
     Dr. Kanok Boriboonsomsin
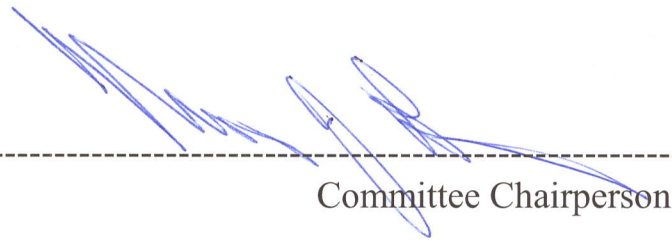
The Dissertation of Weihua Zhu is approved by

Committee Chairperson

University of California, Riverside

Acknowledgment

First of all, I would like to express my deep gratitude to my advisor, Prof. Matthew Barth, for his guidance, advice, support, and encouragement I received throughout my research activity. He introduced me into this field and guided me patiently. Prof. Matthew Barth's sharp insight, extensive knowledge, system-level thinking, and professionalism are and will always be an inspiration to me. Without his help and support, I would not have been here. I will always cherish this great experience.

I would like to thank Dr. Kanok Boriboonsomsin for all the help, discussion, and suggestions I got during my Ph.D. research. I interacted with Dr. Kanok Boriboonsomsin for several years on various research projects and learned much from his professionalism of being an engineer and a researcher.

I am indebted to my other committee members and other faculties who taught me before, including Prof. Gerardo Beni, Prof. Jay Farrell, Prof. Amit K. Roy-Chowdhury, Prof. Ping Liang, Prof. Jie Chen, Prof. Bir Bhanu, Prof. Ilya Dumer, Prof. Yingbo Hua, Prof. Daniel Xu, Prof. Ertem Tuncel, Prof. Tao Jiang, Prof. Eamonn Keogh, Prof. Vassilis Tsotras, Prof. Vana Kalogeraki, Prof. Srikanth Krishnamurthy, Prof. Teodor Przymusinski, and Kris Miller. I have benefited greatly from their inspiring teaching, questions, suggestions, and discussion.

I am also grateful to my coworkers and lab mates at CE-CERT, especially Amanda Gomes, Alex Vu, George Scora, Mike Todd, and Anh Vu. I got a great deal of help from

them during my time at CE-CERT. I would also want to thank CE-CERT and everybody at CE-CERT for providing a fantastic work environment.

My friends at UCR are largely the reason for my happy times in Riverside. They are always there and offering help and fun. I would like to express my gratitude to Dr. Meng Cao, Lili Huang, Yilei Xu, Bi Song, Kezhu Hong, Duo Li, Shan Shan, Anning Chen, Lingfei Zhou, Xiaotao Zou, and Gang Liu among others. All their help is sincerely appreciated.

Finally, thanks to my parents, for their endless and unconditional love. There is no way I could ever repay them for everything they have given to me.

Dedicated

To My Parents

ABSTRACT OF THE DISSERTATION

Design and Development of Novel Routing Methodologies for Dynamic Roadway
Navigation Systems

by

Weihua Zhu

Doctor of Philosophy, Graduate Program in Electrical Engineering
University of California, Riverside, March 2009
Prof. Matthew Barth, Chairperson

To date, traditional navigation systems have embedded algorithms that attempt to
minimize trip distance and/or travel time. However, many drivers are now becoming
increasingly concerned with fuel costs and vehicle emissions that are detrimental to the
environment. Therefore, it is desirable to create new "environmentally-friendly" and
"energy-friendly" navigation algorithms. Taking advantage of the latest navigation
technology, in this dissertation, new navigation techniques have been developed that
focus on minimizing energy consumption and vehicle emissions. These methods combine
sophisticated mobile-source energy and emission models with route minimization
algorithms that are used for navigational purposes. It is also known that different road
types can play a significant role in emissions and fuel consumption. As such, a new
standalone, high-accuracy road type classification methodology has been developed that

only uses a short vehicle velocity trajectory as input, without any external mapping system.

Further, it was found that under chaotic traffic conditions (e.g., those caused by high demand, unexpected road closures, and natural disasters), a shortest-distance route algorithm might suggest a route with unreasonably long travel times, consuming a great deal of energy. On the other hand, under similar chaotic traffic conditions, a shortest-duration routing algorithm might frequently advise a driver to switch routes to avoid congested roadways and maintain reasonable travel time. The number of possible routes varies by the roadway network topology and the location within the network. Thus, it is useful to know how many possible routes exist. Therefore, a new navigational mobility index (NMI) has been developed and justified with an initial focus on freeway networks. NMI can be based on the number of possible routes weighted by shared segments among routes from a source to a destination (referred to as node-to-node NMI). Based on node-to-node NMI, node-NMI and area-NMI are also defined and justified. Different applications of NMI include: 1) measurement of the degree of freedom in which drivers can choose routes from a route choice perspective; 2) determination of the potential effectiveness of navigation systems; 3) determination of the overall connectivity level of an area; and 4) the guidance of the movement of people during an evacuation due to a disaster event.

Based on the proposed NMI concept, a new routing methodology has been developed that is based on maximizing the degree of freedom for re-routing while driving from a known location to a desired destination. Not only is this routing methodology beneficial

for dealing with random incidents, it is also useful during major disaster situations when people in an affected area need to be quickly evacuated and relocated to safer areas. A variety of experiments have been carried out to determine the effectiveness of the proposed concept and routing methodology.

The main contribution of this dissertation are as follows: 1) We prove that a shortest-duration and a shortest-distance route are not necessary the most energy efficient route. We have combined a state-of-art energy/emissions model with navigation technologies to develop an environmentally-friendly navigation methodology, which is unique; 2) Because road type plays an important role in vehicle emission and energy consumption, we have developed a highly accurate, low complexity, and stand-alone road-type classification algorithm that only uses a short vehicle speed trajectory as input without external support such as a map system; 3) We have originally proposed and defined a navigational mobility index (NMI) concept specifically for navigational purposes—compared to other existing similar concepts, it has numerous desirable properties and can be used to evaluate the potential effectiveness of a navigation system; 4) Based on the original NMI concept, node-NMI and area-NMI measures have been further defined that can be used to assess the overall degree of freedom of routing in an area; and 5) For emergency evacuation and navigation under chaotic traffic conditions (e.g., those due to high demand or unexpected road closures), drivers can maximize their degree-of-freedom when re-routing. This is highly desirable under emergency evacuation scenarios, in which drivers are more likely to arrive to the safe area using NMI-based navigation than using the traditional shortest-distance or shortest-duration navigation.

Contents

List of Figures

List of Tables

1. Introduction

1.1 Problem Statement

The total number of vehicle miles traveled (VMT) on roadways continues to increase worldwide, with people driving greater distances and for longer periods of time. This travel demand consumes large quantities of fuel, corresponding to approximately 50% of the United States' energy costs [1]. Extensive roadway systems have already been built over the last century, and in many locations, expanding the network is now less likely due to several issues such as limited land use options and significantly higher building costs. Given this steadily increasing travel demand with limited infrastructure growth, highway congestion continues to get worse. People are driving longer distances and for longer periods of time than necessary. Several studies have been carried out examining roadway congestion in terms of lost productivity and wasted fuel; e.g., in 2003 it has been estimated that 2.3 billion gallons of fuel were wasted due to congestion [2]. There is now a strong need to make the highway travel as efficient as possible. As a result, a variety of intelligent transportation system (ITS) technology is being developed to use the existing roadway network more efficiently.

One example of this ITS technology has been the development of navigation systems (see details in Chapter 2). In recent years, there has been significant advancement of both on-board in-vehicle and off-board internet-based navigation systems that primarily select

routes (i.e. paths, in this dissertation, routes and paths are used interchangeably) corresponding to the shortest-distance between an origin and a destination. It sometimes is also able to calculate a static shortest-duration route using the approximate speed of road segment according to the road type (e.g., freeways have an average free-flow speed of 95 km/h, arterials at 55 km/h, etc.). These systems use a static roadway map and do not use the real time information, and we call these *static roadway navigation systems*. In contrast, we call any navigation systems that utilize the real time traffic information *dynamic roadway navigation systems*.

For example, more recently, new generation systems are now capable of incorporating real time traffic information, providing the ability to find the dynamic shortest-duration path (e.g. [71-72], see details in Chapter 2), in addition to the static shortest-distance path and the static shortest-duration path. These systems can suggest an alternative route if traffic congestion happens and advise the drivers to re-route. "Re-routing" is the point of this kind of dynamic shortest-duration navigation system.

However, there exist some problems of these current navigation systems. First, these navigation systems do not take into account the energy consumption and vehicle emissions. In heavily congested traffic conditions, the static shortest-distance or the static shortest-duration route might result in much higher fuel consumption and vehicle emissions. Second, a dynamic shortest-duration algorithm might direct drivers to a route that actually does not have many routing choices, i.e. a route that has low *degree of freedom* in re-routing. For a route having low degree of freedom in re-routing, the

advantage of the dynamic shortest-duration algorithm is eliminated and the navigation systems will not be of much use.

This dissertation is dedicated to the development of energy/emissions-efficient and better degree-of-freedom navigation algorithms. We first propose and develop novel navigation methodologies that focus on minimizing energy consumption and vehicle emissions. This navigation technique takes advantage of the real time traffic and combines a state-of-art Comprehensive Modal Energy and Emission Model [6-10] with route minimization algorithms. By the algorithm we developed, drivers are given choices of not only the traditional static shortest-distance, static shortest-duration routes, dynamic shortest-duration routes, but also a route consuming less energy and producing less emissions. Therefore, there will be four navigation algorithms available.

As previously mentioned, when a driver gets real time traffic information, (s)he might be advised to switch routes if the current route is not desirable due to real time traffic conditions. With the ability to re-route based on changing network conditions, it is useful to develop an effectiveness index that gives an indication of the number of routing options. For example for an extreme case, if there is only one route between a source-destination pair, choosing an alternative route based on real-time traffic conditions will not be feasible because there is no other available routes. Taken one step further, a network with low index value will not have many re-routing choices; and therefore, choosing an alternative route based on real-time traffic conditions will have limited value. On the other hand, a roadway network with a high index value will be able to provide numerous choices when traffic conditions change. In this dissertation, we then propose an

original navigational mobility index (NMI) concept for a source-destination pair to quantify the routing options between them. We further define a node-NMI and area-NMI based on overall roadway availability.

Normally, drivers want to select any one of the four routing algorithms mentioned above, and would like the traffic conditions to be free-flowing and uninterrupted. However, adverse scenarios often occur; for example, roadway links might unexpectedly shutdown (or become highly congested) due to major incidents (i.e. accidents) or natural disasters (e.g. earthquakes). In these scenarios it is beneficial to re-route, and sometimes it might even be mandatory to re-route. In these cases, it is desirable that a navigation algorithm find an optimal route that offers the highest degree of freedom in re-routing. When considering this kind of optimal route, even if some roadway infrastructure is unusable, the drivers are still likely to find another route to reach their destination without being trapped in a dead end or backtracking too much, and arrive at the destination within a reasonable time. In this dissertation, based on the proposed concept NMI, we further describe a new routing methodology that is based on maximizing the routing choices when driving from a known location to a desired destination. Not only is this routing methodology beneficial for dealing with random incidents, it is also useful during major disaster situations when people in affected areas need to be quickly evacuated and relocated to safer areas.

Therefore, in addition to the traditional shortest-distance/duration algorithms, we provide two possible choices for roadway navigation: environmentally-friendly

navigation algorithm providing a route consuming less energy and an NMI-based navigation algorithm providing a route giving better degree of freedom for re-routing.

1.2 Contribution of the Dissertation

In this dissertation, we consider novel navigation methodologies other than traditional shortest-distance/duration algorithms. Shortest-distance/duration methods have been proven to not necessary consume less energy or produce less pollutants. We first review the related research and give brief background information in Chapter 2. In Chapters 3 and 4, we propose an energy efficient navigation methodology that can guide drivers to routes that consume less fuel and produce lower emissions. With rapid changing traffic conditions, a shortest-distance or shortest-duration routes might not be optimal in any sense (duration/energy consumption). Drivers frequently want to switch routes, especially in metropolitan areas. In the second part of this dissertation (Chapters 5 and 6), we propose a navigation algorithm that gives better degree of freedom for the drivers to switch routes. Specifically, the following lists the main contribution of this dissertation:

- We prove that a shortest-duration and a shortest-distance routes are not necessary the most energy efficient routes. We combine the state-of-art Comprehensive Modal Energy and Emission Model with navigation technologies to develop an environmentally-friendly navigation methodology, which is unique.

- Road type plays an important role in vehicle emission and energy consumption. We have developed a high accuracy, low complexity, and stand-alone road-type

classification algorithm based only on a short vehicle speed trajectory as input without external support such as a map system.

• We originally propose and define a navigational mobility index (NMI) concept specifically for navigational purposes. Compared to other existing similar concepts, it has several desirable properties. NMI can be used to evaluate the potential effectiveness of a navigation system.

• Node-NMI and area-NMI are further defined based on the definition of node-to-node NMI. Node-NMI and area-NMI can be used to assess the overall degree of freedom of an area. Area-NMI is comparable across different areas.

• For emergency evacuation and navigation under chaotic traffic conditions (highly congested or road link closure), drivers frequently want to switch routes. Based on the defined NMI, we originally propose a navigation methodology giving routes in which drivers have the best degree of freedom for re-routing. This is highly desirable under emergency evacuation scenarios, in which drivers are more likely to arrive to the safe area using NMI-based navigation than using the traditional shortest-distance or shortest-duration navigation.

1.3 Organization of the Dissertation

The remaining chapters of this dissertation are organized as follows: in Chapter 2, a literature review and background information is provided. In Chapter 3 we present the development and application of the environmentally-friendly navigation system, including the methodology, justification and experimental results. From the results from

Chapter 3, we know that road type is a very important factor for fuel consumption and emissions. Chapter 4 then describes a methodology for road type classification based on the vehicle speed snippets only. In Chapter 5, we define the concept navigational mobility index (NMI) in detail, give simple illustrative examples, and rigorously prove the NMI properties. We then give a brief description of how to apply the NMI for roadway navigation. Experimental results are also given to prove its effectiveness. Chapter 6 expands the application of NMI on roadway navigation in Chapter 5, and uses an emergency evacuation scenario as a case study. We conclude the dissertation and highlight future work in Chapter 7.

The dissertation is systematically organized and each chapter is closely related to each other. However, we also organize the dissertation in such a way that each chapter can be viewed as a separate and complete entity, and can be understood without reading the previous chapters. Each chapter itself contains enough information for understanding the material. For example, Chapter 6 bases the application of NMI in roadway navigation on Chapter 5, which defines the NMI concept in detail. However, Chapter 6 briefly summarizes the NMI definition in the background section, and if a reader is only interested in NMI-based navigation, he can skip Chapter 5 and the completeness is maintained.

2. Background and Literature Review

In this chapter we introduce various background material for this dissertation, and conduct a literature review of related research. In Chapter 1, we briefly talked about the navigation systems. In Section 2.1, we discuss the current most advanced navigation technologies in detail, and use the TomTom HD traffic service [71-72] as an example. Section 2.2 describes related research focusing on fuel consumption and vehicle emissions. Section 2.3 reviews existing transportation network performance measures and points out the uniqueness of the defined the NMI (described in detail in Chapter 5). The next two sections describe background materials needed for this dissertation: the data sources and an introduction of the wavelet transformation.

2.1 Navigation Technology

One of the major successes in ITS technology lately has been in the area of Advanced Traveler Information Systems (ATIS). In Chapter 1, we introduced the static shortest-distance, static shortest-duration, and dynamic shortest-duration algorithms. Here we give some examples.

There are now several on-line Internet tools (e.g., MapQuest [59] and previous version of Google Map [58]) that provide directions from any origin to any destination in the roadway network. In addition, many vehicle manufacturers (as well as third party companies) now offer on-board navigation systems that use GPS technology combined with sophisticated mapping software to provide driving directions to specified

destinations. These systems do not use dynamic real-time traffic information, and are examples of static navigation systems. As to the price, they are either free of charge (internet-based) or have an affordable price with no subscription fee. For example, the author has a Garmin Nuvi 270 portable GPS navigator [74], which is around $130. Due to the low price of these systems, they are still dominant in the market.

Recently, more advanced navigation systems are emerging that use real-time traffic information on congestion and accidents to determine the best route to take to a desired destination. These advanced navigational tools are emerging on the Internet (e.g. current version of Google Map [58]) and are used where roadway performance data is available (where sensors exist). Further, several vehicle manufacturers and third parties have incorporated real-time traffic information into their on-board navigation systems. These dynamic navigation systems may report to the drivers the length and reason of the delay, travel time, and even alternative routes. These systems generally require a monthly subscription fee, and are slowly growing in the market.

For example, a leading portable navigation solutions provider TomTom is providing a "high definition" (HD) traffic service [72] to users with real time traffic information (mainly in Europe), and a possible suggestions of alternative routes. According to [71], the HD traffic is available to the public in the Netherlands and being prepared for the introduction in a number of other countries in Europe in 2009. There are basically four data sources of TomTom's HD service [71]. The first one and the largest one is data anonymously collected from cell phone networks. The second source of data is provided by users of connected navigation devices. These users provide anonymous, real-time GPS

location and velocity data. The third one is the third-party data providing accidents and roadway closure information, such as the data from a traffic management center. The anonymous historical GPS traces collected from TomTom navigation system users are the fourth data source. These historical data provide information on average speeds by roadway segments by time of the day.

Therefore, it is possible that it covers much larger roadway network if there are enough number of drivers willing to provide such information through cell phones. The manufacture claims to cover up to ten times more roads than conventional traffic system. Figure 2.1 shows such a GPS unit capable of receiving TomTom HD traffic.

Figure 2.1: TomTom one XL HD traffic GPS unit

One point to clarify is that although these routes are calculated based on real time traffic information and they might suggest a desirable travel time, it cannot guarantee that

paths calculated in this way have the shortest duration, due to the fact that roadway incidents (accidents, road segment closure, etc.) cannot be predicted. Having this in mind, however, we still *call* it dynamic shortest-duration algorithm throughout this dissertation.

2.2 Environment / Energy Navigation Research

Navigation technologies described in Section 2.1 certainly helps to travel efficiently. In many cases, a time- or distance-minimizing algorithm will also minimize fuel consumption and emissions. However, there are cases where this is not true, particularly with congestion and in areas with significant road grade. A shortest-distance algorithm may have a driver travel through heavily congested conditions, resulting in higher fuel consumed. On the other hand, there will also be some cases where a shortest-duration (dynamic or static) route will result in longer distance traveled on less congested roadways. Travel at very high speeds for longer distance will generally result in higher fuel consumption (and emissions) compared to a more direct route.

Several studies have shown that selecting different travel routes between the same origin-destination (OD) pair can result in significant differences in the amount of fuel consumed and the amount of emissions emitted [3, 4]. Further, an exploratory study in Sweden found that 46% of the trips were not made on the most fuel-efficient route. These trips could have saved fuel by an average of 8% with the help of a fuel-optimized navigation system [5]. Therefore, there is much to be done in terms of finding fuel efficient routes.

However, only a few researchers have investigated methodologies that provide a

route consuming *less energy* and produce *less emissions* [60][63][64][12]. In this dissertation, we are not necessarily developing new path minimizing algorithms, but rather developing a new set of cost functions based on energy and emissions using the CMEM model. The use of fuel consumption and emissions as a cost function in roadway navigation is unique.

2.3 Transportation Network Performance Measures

As we know from Section 2.1, the dynamic shortest-duration navigation systems have the capability to re-route if the driver wishes to avoid a particular area of the roadway network. With the ability to re-route based on changing network conditions, it is useful to develop an effectiveness index that gives an indication of the number of routing options. A network with a low index value will not have many re-routing choices; and therefore, choosing an alternative route based on real-time traffic conditions will have limited value. On the other hand, a roadway network with a high index value will be able to provide numerous choices when traffic conditions change.

There has been a variety of performance indices developed for different purposes of transportation network evaluation. These indices are generally used to quantify three performance measures, i.e. mobility, accessibility, and reliability. In the context of traffic engineering, mobility has been defined as "the potential for movement, the ability to get from one place to another," according to Handy [42]. Further, Handy [42] also takes the traditional level of service as an example of measure of mobility: higher volume-to-capacity ratio results in lower mobility.

Accessibility index can be interpreted as a measure of how easy it is to reach *opportunities* (e.g. goods, services, activities, educational places). It is more concerned about how easy it is to reach opportunities rather than how easy the travel itself is. Therefore, an accessibility index generally incorporates both an impedance factor indicating the cost (time, distance, fuel consumption, etc.) of reaching an opportunity and an attractiveness factor indicating the quality of that opportunity. A high level of mobility does not necessary result in high accessibility. For example in rural areas, although the traffic conditions are always good, the accessibility level is low because there are few things to reach. Ahmed and Levinson [43] as well as Handy [42] provide a good explanation of the difference and relationship between accessibility and mobility. There are plenty of accessibility indices proposed over the past several decades (see, e.g. [43, 44, 45, 46, 47, 48], and see [43] and [44] for a comprehensive survey of various accessibility indices).

In contrast to the previous two measures, reliability indicates how reliable (how much performance can be maintained) the transportation network is after certain links fail. Examples are connectivity reliability, travel time reliability [49], and capacity reliability [50].

In this dissertation, we propose a new transportation network performance measure called *Navigational Mobility Index* (NMI) to quantify the availability of different roadway facilities from one location to another (with an initial focus on freeways), specifically for navigational purposes. We further define a node-NMI and an area-NMI based on overall roadway availability. For the dynamic routing algorithms, one relevant

question that arises is that if a driver wishes to change his/her route based on real-time traffic information, how many reasonable route choices are possible? The proposed NMI will answer this question fundamentally.

The NMI defined in this dissertation differs itself from any of the three measures previously mentioned. It is specifically defined for navigational purposes. Although we term it as "mobility index", we do not intend to associate the proposed NMI with any previous measures. In the literature, the closest concept that may achieve similar functionality is Path-Size. In the context of route choice modeling, Bekhor et al. [51, 52] proposed the concept of Path-Size to account for the correlation between routes of a source-destination pair and defined it as the weighted number of paths. However, in later chapters we show that the Path-Size concept is not suitable for our navigational purposes. Based on the proposed NMI measure, we develop a routing algorithm giving the best degree of freedom for re-routing. To the best knowledge of the authors, we have not found any similar research to this topic.

The next two sections describe background material useful for later chapters.

2.4 Data Sources

There are basically two types of data sources used for analysis in this dissertation. The first source is *macroscopic* traffic data, which are typically collected by embedded inductive loop detectors, magnetic sensors, microwave radar, video cameras, infrared sensors, etc. [70]. Raw data collected by these sensors can be processed to measure traffic flow, density, and average traffic speed [73].

Among all these kinds of sensors, inductive loop detectors have been used extensively on freeways. Figure 2.2 describes how a typical embedded inductive loop detector looks like. For this dissertation, we utilized California's Freeway Performance Measurement System (PeMS) [14-15] operated by the California Department of Transportation (Caltrans) and the University of California, Berkeley to provide real-time macroscopic traffic parameters. The California PeMS system collects real-time traffic information from inductive loop detectors embedded in California's freeways and makes it available for transportation management, research, and commercial use. The system provides real-time five-minute, per-loop averages of lane occupancy, flow, speed, and congestion for various links in the roadway network. The data are available over the Internet. For more information on PeMS, see http://pems.eecs.berkeley.edu.
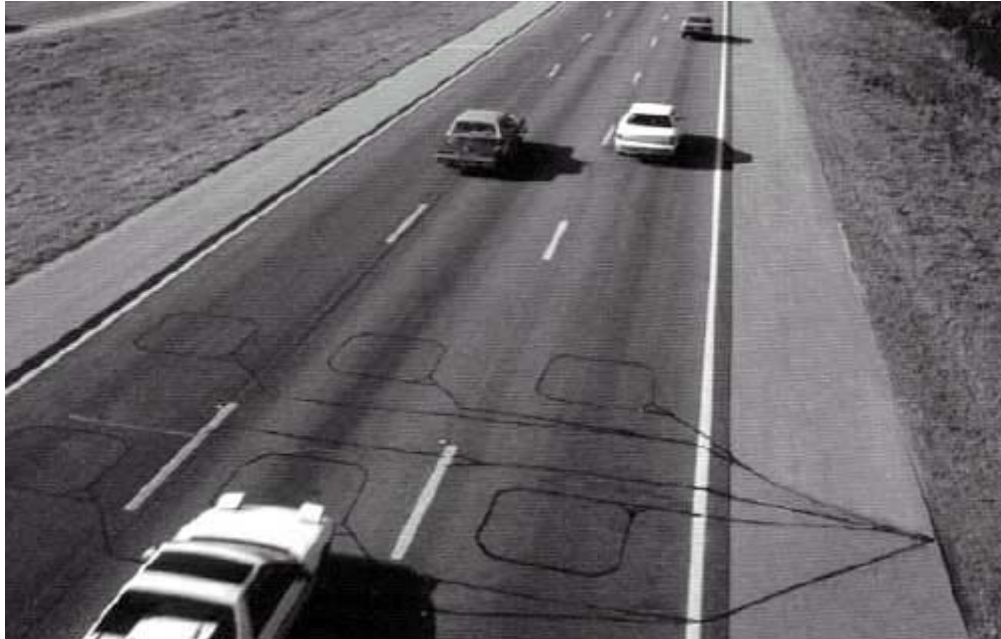
Figure 2.2: A picture of an embedded loop detectors

The second source is *microscopic* vehicle data, which are collected by GPS-equipped probe vehicles and are used to examine the second-by-second velocity trajectory of individual vehicles. At the College of Engineering – Center for Environmental Research and Technology (CE-CERT) at the University of California Riverside, a large database of these vehicle activity data is organized into different layers. In this dissertation, we use this database collected by CE-CERT for the microscopic vehicle data. Figure 2.3 shows an example probe vehicle we used.

A variety of studies have taken place to develop relationships between microscopic vehicle data and macroscopic traffic data. For example, in [69] it was found that the

macroscopic traffic data can be correlated with the velocity trace of the probe vehicle, acting as a representative of the general traffic flow.

Figure 2.3: A picture of a probe vehicle.

2.5 Wavelet Transform Application

Wavelet transforms were developed primarily in the 1980's as a tool to divide data or functions into different frequency components. Each time-series component can then be examined at different scales. Wavelet transforms are well known for their good localization of both time and frequency instead of a frequency-only technique like a

discrete Fourier transform (DFT). There has been a significant amount of activity in recent years applying wavelet techniques to transportation data. For example, Ping designed a wavelet-based linear prediction technique for traffic volumes [24]. Further, Yu investigated wavelet-based determination of aggregation levels for ITS data [25]. Jiang uses a wavelet packet-autocorrelation function method for traffic flow pattern analysis [26].

Wavelet transform is a branch of applied mathematics, and it is a fusion of ideas from several different fields, such as applied mathematics, digital signal processing, and physics. Basically, there are two kinds of wavelet transforms. The first one is continuous wavelet transform (CWT). Let $x(t) \in L^2(R)$, where $L^2(R)$ is the set of all functions f such that the integral of f^2 over the whole real line is finite, and $\phi(t)$ is the mother wavelet function, then the CWT of $x(t)$ is

$$WT_x(a,\tau) = \frac{1}{\sqrt{a}} \int x(t) \phi^* (\frac{t-\tau}{a}) dt =< x(t), \phi_{a\tau}(t) > \qquad (2.1)$$

where a is the scale factor and τ is the translation factor.

CWT is calculated by computing the correlation between a signal and a continuously shifted and continuously scaled version of the mother function. Therefore, much redundancy exists in a CWT. At the same time, for most functions, the CWT has no analytical solutions and therefore needs to be numerically calculated by a computer. Moreover, in many practical applications, the signal is sampled time sequences. Therefore, a discrete wavelet transform (DWT) is more preferable. An important issue in any transform scheme is the question of reconstruction from the transform domain. It

18

turns out that it is possible to reconstruct a signal from its wavelet decomposition. DWT is quite suitable for discrete signal processing, for example, in speech, image, and time sequence processing. There are several ways to present the DWT: from a filter-bank theory approach or from a multiresolution analysis (MRA) point of view. The MRA approach is briefly reviewed (see [33] and [34] for further details).

MRA describe mathematically the process of studying signals at different scales. MRA represents the whole space $L^2(R)$ by a sequence of embedded subspaces for an intelligent choice of appropriate subspaces for an application to get a compromise between accuracy and computation complexity. MRA studies the property of a sequence of subspace V^j and W^j, $j \in Z$, which approximate $L^2(R)$ by satisfying

$$\cdots, V^1 = V^0 \oplus W^0, V^2 = V^1 \oplus W^1, \cdots, V^{j+1} = V^j \oplus W^j \cdots$$

where:

$Union(V^j)_{j \in Z} = L^2(R)$ (union of all V^j is the whole space); $\cap_{i \in Z} V^j = O$

(intersection of all V^j is empty);

$V^j \perp W^j$ \qquad\qquad (V^j is orthogonal to W^j); and

$W^j \perp W^{j'}$, $j \neq j'$ \quad (W^j is orthogonal to $W^{j'}$).

The Haar scaling function is defined as:

$$\phi(t) = \begin{cases} 1 & t \in [0,1) \\ 0 & t \notin [0,1) \end{cases} \qquad\qquad (2.2)$$

and $\phi^j(t) = 2^{j/2} \phi(2^j t - i)$, $j = 0,1,...and \;\; i = 0,1,...,2^j - 1$.

19

The vector space V^j is defined as $V^j = span\{\phi_i^j\}_{i=0,1,...,2^j-1}$. It is straightforward to see

that $V^j \subset V^{j+1}$. The Haar wavelet function is defined as

$$\varphi(t) = \begin{cases} 1 & t \in [0,1/2) \\ -1 & t \in [1/2,1) \\ 0 & otherwise \end{cases}$$ (2.3)

and $\varphi_i^j(t) = 2^{j/2}\phi(2^j t - i)$ $j = 0,1,...$ and $i = 0,1,...,2^j - 1$.

The vector space W^j is defined as $W^j = span\{\varphi_i^j\}_{i=0,1,...2^j-1}$ such that $W^j \subset V^{j+1}$.

By straightforward mathematical computation, we find that the Haar basis has this very

important property:

$$V^{j+1} = V^j \oplus W^j$$

In order to compute a Haar wavelet transform, we let $f \in R^N$, where $N = 2^n$, and

expand this signal into $V^0 \oplus W^1 \oplus \cdots \oplus W^{n-1}$ spanned by the basis vectors. This is

accomplished by first expanding f into $V^{n-1} \oplus W^{n-1}$. The first step can be described as the

following matrix-vector multiplication $f_1 = W_1 f$ where

$$W_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 & & & & & \\ & & 1 & 1 & & & \\ & & & & \ddots & \ddots & \\ & & & & & 1 & 1 \\ 1 & -1 & & & & & \\ & & 1 & -1 & & & \\ & & & & \ddots & \ddots & \\ & & & & & 1 & -1 \end{bmatrix}_{2^2 \times 2^n}$$

It can be seen that the first half rows correspond to the basis $\phi_0^{n-1}, \phi_1^{n-1}, \cdots, \phi_{2^i-1}^{n-1}$ that

span V^{n-1} and the last four rows correspond to the basis vector $\varphi_0^{n-1}, \varphi_1^{n-1}, \cdots, \varphi_{2^i-1}^{n-1}$ that

spans W^{n-1}.

Step two is to expand f into $V^{n-2} \oplus W^{n-2} \oplus W^{n-1}$. The second step can be described by

the matrix-vector multiplication as $f_2 = W_2 f_1$.

$$
W_2 = \frac{1}{\sqrt{2}}
\begin{bmatrix}
1 & 1 & & & & & & & & \\
 & & \ddots & \ddots & & & & & & \\
 & & & & 1 & 1 & & & & \\
1 & -1 & & & & & & & & \\
 & & \ddots & \ddots & & & & & & \\
 & & & & 1 & -1 & & & & \\
 & & & & & & 1 & & & \\
 & & & & & & & 1 & & \\
 & & & & & & & & \ddots & \\
 & & & & & & & & & 1 \\
 & & & & & & & & & & 1
\end{bmatrix}_{2^n \times 2^n}
$$

This can be generalized to:

$$
W_n = \frac{1}{\sqrt{2}}
\begin{bmatrix}
1 & 1 & & & \\
1 & -1 & & & \\
 & & 1 & & \\
 & & & \ddots & \\
 & & & & 1
\end{bmatrix}_{2^n \times 2^n}
$$

and $f_n = W_n W_{n-1} \cdots W_1 f$, where f_n is the wavelet coefficient.

3. Environmentally-Friendly Navigation

3.1 Introduction

This chapter describes an innovative route selection methodology that is designed specifically for minimizing *fuel consumption* and *vehicle emissions*. The key innovation in this research is the integration of current navigation technology with sophisticated vehicle energy and emission models. This research now provides the ability for a driver to not only have a choice of selecting a shortest-distance or shortest-duration route, but also a route that minimizes the fuel consumed and/or pollutant emissions for that particular trip.

As we already know from Chapter 2, although it is true that in many cases, a time- or distance-minimizing algorithm will also minimize fuel consumption and emissions, there are several cases where this is not true, particularly with high levels of congestion and in areas with significant road grade. In this chapter, we have carried out case studies to better understand these cases, estimating the effectiveness of the routing algorithms that minimize fuel consumption and emissions.

In Section 3.2, background information on energy and emissions modeling is provided. Section 3.3 outlines the overall methodology of the research. Several case studies are illustrated in Section 3.4, along with their results.

3.2 Comprehensive Modal Energy and Emission Model

In 1996, CE-CERT at the University of California Riverside began the development of a Comprehensive Modal Emissions Model (CMEM, [6-10]), sponsored by the National Cooperative Highway Research Program and the U.S. Environmental Protection Agency (EPA). The need for this type of microscale model that can predict second-by-second fuel consumption and emissions based on different traffic operations was and remains critical for developing and evaluating transportation policy. In the past, large regional emissions inventory models were being applied for these types of microscale evaluations with little success. The majority of the CMEM modeling effort was completed in 2000 and the model has been updated and maintained since then under sponsorship from the U.S. EPA. CMEM is a public-domain model and has several hundred registered users worldwide.

CMEM was designed so that it can interface with a wide variety of transportation models and/or transportation data sets in order to perform detailed fuel consumption analyses and to produce a localized emissions inventory. CMEM has been developed primarily for microscale transportation models that typically produce second-by-second vehicle trajectories (location, speed, acceleration). These vehicle trajectories can be applied directly to the model, resulting in both individual and aggregate energy/emissions estimates. Over the past several years, CMEM has been integrated into various transportation modeling frameworks, with a focus on corridor-level analysis and intelligent transportation system implementations (e.g., CORSIM, TRANSIMS,

PARAMICS, SHIFT, etc.).

CMEM is comprehensive in the sense that it covers essentially all types of vehicles found on the road today. It consists of nearly 30 vehicle/technology categories from the smallest light-duty vehicles to Class-8 heavy-duty diesel trucks. With CMEM, it is possible to predict energy and emissions from individual vehicles or from an entire fleet of vehicles, operating under a variety of conditions. One of the most important features of CMEM (and other related models) is that it uses a physical, power-demand approach based on a parameterized analytical representation of fuel consumption and emissions production. In this type of model, the entire fuel consumption and emissions process is broken down into components that correspond to physical phenomena associated with vehicle operation and emissions production. Each component is modeled as an analytical representation consisting of various parameters that are characteristic of the process. These parameters vary according to the vehicle type, engine, emission technology, and level of deterioration. One distinct advantage of this physical approach is that it is possible to adjust many of these physical parameters to predict energy consumption and emissions of future vehicle models and applications of new technology (e.g., aftertreatment devices).

CMEM has been rigorously validated [9]. It is also considered to be one of the most detailed and best tested estimates of vehicle exhaust emissions at different speeds and accelerations [11]. Further, CMEM also accounts for road grade effects. It has been shown that road grade has a significant effect on fuel consumption and emissions, e.g., [12] and [13]. For further information on the CMEM effort, please refer to [6-10].

3.3 Methodology

The overall methodology for developing environmentally-friendly navigation can be broken down into two general components: 1) For a given roadway network, *link-based energy and emission factors* have been developed that can be indexed by macroscopic link characteristics such as flow, density, speed, and road grade. Thus, given explanatory variables such as flow, density, and speed, energy and emission functions (calibrated for specific vehicles or fleets) can be used to estimate specific link factors. 2) Given these link-based energy and emission factors, *network-wide routing algorithms* have been developed that minimize fuel consumption and emissions. These routing algorithms utilize real-time data on traffic performance with the developed energy/emission factors for the roadway network under study.

3.3.1 Link-Based Energy/Emissions Cost Factors

The method used to build link-based energy and emission factors is illustrated in Figure 3.1. A large vehicle activity database has been collected from GPS-instrumented probe vehicles for this research, providing time-space indexed speed data of a sample vehicle in the traffic stream. Simultaneously, macroscopic traffic parameters are collected from the California PeMS system, in particular for the embedded vehicle detection station that the probe vehicle is traveling over. The trajectory snippets (e.g., several seconds before and several seconds after the vehicle passes over the sensor location) are then correlated to the macroscopic traffic parameters. We consider the following:

t is time lapse in seconds ($t = 1, 2, \dots, T$). For simplicity, we start $t = 1$ at the first second of the first full minute of time stamp;

p is a 30-second period in which PeMS collects data, i.e.

$$p = (t \setminus 30) + 1;$$

i is an index for vehicle detector station (VDS), i.e.

$$i = 1, 2, \dots, n;$$

j is a lane number ($j = 1, 2, \dots, m(i)$), where lane 1 is the median lane and lane $m(i)$ is the shoulder lane;

c_i is the spatial coverage of VDS i, i.e.

$$c_i = \left[\frac{l_{i-1} + l_i}{2}, \frac{l_i + l_{i+1}}{2} \right);$$

l_i is a centerline distance in km from a starting location (at $t = 1$) to VDS i. For simplicity, we assume $l_0 = 0$ and $l_{n+1} = l_n + 1$.

Then, $v_{c_i,j,t} \in U_{i,j,p}$ where v is second-by-second speeds (i.e. speed snippet) of probe vehicles and U is macroscopic speed of the traffic. The probe vehicle data were collected on the Southern California freeway network at different times of day under different congestion conditions.
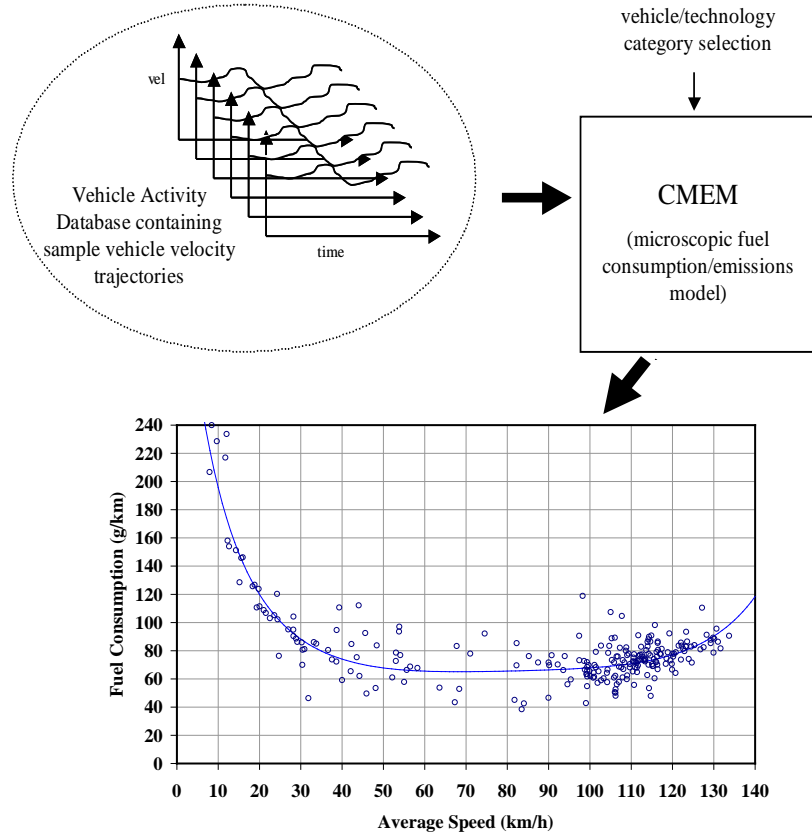
Figure 3.1: Link-level fuel consumption modeling methodology.

This velocity snippet database was then run through CMEM (described in Section 3.2), for all vehicle/technology categories represented in the model. As a result, average fuel consumption and emissions factors can be developed for the corresponding macroscopic traffic parameters for each vehicle/technology category in CMEM. The lower part of Figure 3.1 illustrates these factors as a function of average freeway traffic speed using a curve fitting technique for a typical fleet mix in Southern California. The general equation used in the regression analysis is given below:

$$\ln(y) = b_0 + b_1 \cdot x + b_2 \cdot x^2 + b_3 \cdot x^3 + b_4 \cdot x^4$$

where:

y is the fuel consumption or emissions in g/km;

x is average traffic speed in km/h and

b_i represents the set of polynomial coefficients.

It is clear that in the energy/emissions vs. average speed graph in Figure 3.1, a certain amount of data scatter results for a given average speed. This is in part attributable to the fact that the data points in Figure 3.1 are a result of the weighting among the various vehicle/technology categories in CMEM, which have different energy/emissions characteristics associated with the same driving patterns (i.e. snippets).

Even for the same vehicle/technology category, the data points may still remain scattered, caused by different levels of congestion. For example, there are many vehicle trajectories that can have the same average speed, but have different energy/emissions results. Average traffic speed is currently used as the primary explanatory variable in our methodology as it has been shown to be the strongest explanatory variable for predicting energy and emissions. However, other macroscopic freeway traffic parameters (e.g. flow, density) can also be included as explanatory variables in conjunction with average speed to better differentiate congestion conditions and allow for more accurate fuel consumption and emission estimates.

It is important to point out that although energy/emissions variances exist when estimating link-by-link energy/emissions cost factors based on average traffic speed alone, these variances are unlikely to propagate but rather partly cancel each other out when aggregating the energy/emissions of each link in a route. Therefore, the resulting

variance of the total route energy/emissions will likely be lower than the largest variance of the component links. This is especially true if a route is composed of a number of links so that the distribution of individual variances converges to the Gaussian shape.

Further, it is worth mentioning that even if we include additional explanatory variables such as traffic density and flow rates, a certain degree of variance may still remain due to an individuals' driving style (i.e. passive vs. aggressive). Ideally, the energy/emission estimation equation should be calibrated for each individual driver according to his/her personal driving style. This requires collecting driving data (i.e. speed trajectory) for individual drivers along with the other traffic variables, and creating functions based on these customized data sets. This is potentially a subject of future research.

As mentioned previously, we are only considering freeway links at this time. It is clear that driving on surface streets (e.g., arterials, residential streets, etc.) will likely involve multiple stops, which can amount to a significant amount of delay (i.e. stopped delay). This will certainly affect vehicle fuel consumption and emissions [16]. We are currently collecting velocity trajectories on a number of these other road types and are developing additional energy/emissions functions specific for surface street driving. These other roadway link functions will be incorporated into our environmentally-friendly navigation methods at a later date.

Road grade is another critical variable for estimating link-based energy/emission factors. By comparing the measured fuel economy between a flat route and example hilly routes, it has been found that the vehicle fuel economy of the flat route is superior to that

of the hilly routes by approximately 15-20% [17]. Thus, road grade will also be another important factor in the environmentally-friendly navigation methodology. However, to date, digital roadway networks typically do not typically include road grade information; and thus, the incorporation of road grade in our navigation methodology is slated as future research.

3.3.2 Link Fuel Consumption and Emissions Factor Assignment

Most digital roadway networks consist of *nodes* (e.g., intersections, freeway on/off-ramps, point of curvature, etc.) and *links* (i.e., the road sections between nodes). Specific link and node attributes define how the network is connected together and what the general features are of the different links/nodes (e.g., position, length, number of lanes, capacity, speed limit, etc.). Typical navigation algorithms usually consist of finding a particular path between two nodes in the network. This path is usually based on some optimality such as shortest-distance or shortest duration. Dijkstra's algorithm [18] is a prime example of a solution to the macroscale route-planning problem; however, other more efficient algorithms exist (e.g. [19]).

One of the key steps of this research is to assign specific fuel consumption and emissions factors to each link in the roadway network. Because these factors are a function of traffic parameters, ideally each link should have traffic performance data. For the freeway links, these traffic performance data are provided from the PeMS system. An example is given in Figure 3.2. In the figure, different traffic performance values are reported by VDS L through S, as indicated by different colors. These VDS have different

spacing distances where the typical spacing distance between two consecutive VDS is about 1.0-1.6 km. With the knowledge of the actual spacing distances between the VDS, a set of virtual links can be created where each virtual link carries the traffic performance data of the VDS it represents. The spatial coverage of a virtual link generally extends from the mid distance between its VDS and the adjacent VDS on one end to the mid distance between itself and the adjacent VDS on the other end.
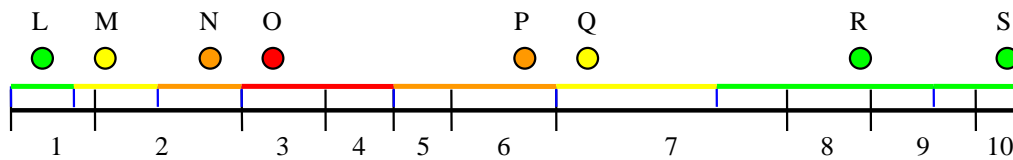
Figure 3.2: Link-based energy/emissions factors assignment methodology.

Each link in the roadway network representation (i.e. Links 1-10 in Figure 3.2) is then assigned a traffic performance value of the overlapping virtual link(s) weighted by the overlapping distance. For the example in Figure 3.2:

$E_1 = (3/4)E_L + (1/4)E_M$

$E_2 = (3/7)E_M + (4/7)E_N$

$E_3 = E_O$

$E_4 = E_O$

$E_5 = E_P$

$E_6 = E_P$

$E_7 = (7/11)E_Q + (4/11)E_R$

31

$E_8 = E_R$

$E_9 = (3/5)E_R + (2/5)E_S$

$E_{10} = E_S$

Similarly, each link in the network should also have a road grade value. We have developed a methodology for determining road grade of each link; however, road grade has not yet been implemented in our algorithms. Once the link-based traffic data (and in the future road grade) are in place, then the fuel consumption and emissions (i.e., carbon dioxide (CO_2), carbon monoxide (CO), hydrocarbons (HC), and oxides of nitrogen (NO_x)) values are calculated for each link, for each vehicle/technology type within CMEM. It is important to note that in addition to calculating environmentally-friendly paths for a specific vehicle/technology type (e.g., heavy-duty trucks), it is also possible to calculate the path for a *fleet* of vehicles. All that is needed is a vehicle fleet composition matrix, providing the percentage of each vehicle/technology group in the fleet for the area of interest.

We have developed our methods using ArcGIS 9.1, a popular Geographical Information System (GIS) software environment that is produced by ESRI [20]. We have imported roadway network data for the southern California region. A specific module was developed to link the roadway network database with the real-time traffic performance data (i.e., PeMS), obtaining the data every five minutes. Further, several other fields were added to the roadway network database, specifically road grade, fuel consumption, CO_2 , CO, HC, and NO_x for the different vehicle/technology categories in CMEM. Using the Network Analyst toolbox, the path minimization feature is utilized,

which can be set to minimize based on any characteristics of the roadway link data. As a result, it is possible to compare minimal routing paths for distance, time, energy, and the various emissions.

3.4 Results and Discussion

As described earlier, a shortest-duration or shortest-distance path will often be the path that also minimizes energy or emissions. This makes sense since the shorter amount of time a vehicle spends on the roadway, the aggregated energy/emissions should be less. However, if there are roadway congestion and other factors, the energy and emissions could be different.

3.4.1 Case Studies

There are several freeway routes that provide access between downtown Los Angeles and the Inland Empire region, which is made up of many suburban communities. The typical travel pattern consists of work commutes from the Inland Empire to downtown Los Angeles in the morning, and return commutes back in the afternoon. Because of the high volume of commute trips, these freeways often suffer from congestion. To illustrate the effect of route choice on energy/emissions under various traffic conditions, we consider example trips from downtown Los Angeles to Chino, California. Two comparable freeway routes are available: 1) taking Interstate 10 followed by State Route 71 (referred to I-10 route), and 2) taking State Route 60 (referred to SR-60 route). They are illustrated in Figure 3.3. These routes have approximately the same distance (around

45.6 km).

In this chapter, we present four case studies that have different levels of congestion on the two alternative routes. For all four case studies, we consider energy/emissions of a typical light-duty truck or sports utility vehicle on the freeways under consideration.

1) Case I: Free-flow Conditions

To demonstrate the case study where the minimal energy/emissions route is the same as the shortest-duration route, we consider an example trip on May 20, 2007 at 1:15 AM. Unsurprisingly, during this nighttime event, all freeways were uncongested and the traffic was at free-flow. The summary of distance, travel time, as well as estimated fuel consumption and emissions are given in Table 3.1 (a). Because of the uncongested conditions, a vehicle could travel at the free-flow speed on both routes. Therefore, the energy and emissions results are similar for either route (i.e., within a few percentage difference).

2) Case II: Free-flow vs. Moderate Congestion

To illustrate the case where the most environmentally-friendly route is not necessarily the same as the shortest-duration route, we consider a trip on June 9, 2007 at 12:20 PM. In this case, the SR-60 route was operating at near free-flow conditions while the alternative I-10 route was experiencing some moderate congestion. In this case, the SR-60 route is definitely the faster route (about 5 minutes faster); however, the I-10 route is

more beneficial environmentally, as summarized in Table 3.1 (b). The reason that fuel consumption and emissions are less is because the vehicles are operating in a moderate congestion regime, where the slightly lower speeds require less load on the vehicle engines, resulting in lower fuel consumption and emissions [21].

3) *Case III: Free-flow vs. Heavy Congestion*

Another case to consider is when one route is operating under free-flow conditions and the other route has heavy congestion. This was the case for I-10 on June 17, 2007 at 5:10 PM. Because the travel time was significantly longer (71%), and the traffic was under stop-and-go conditions for the most part, the fuel use was higher than the free-flow route. This affirms that heavy congestion is typically detrimental to fuel consumption.

We also observe that in Table 3.1 (c), the fuel consumption and emissions are not consistent, i.e. the fuel consumption and emissions of one route do not necessarily have to be consistently better than those of the other route. For example, the I-10 route produces more CO_2, less CO, HC, and NO_x than the SR-60 route. Figure 3.1 explains this phenomenon. In Figure 3.1, we can see that for the free-flow and heavy congestion portion, it is possible for them to have comparatively same amount of fuel consumption and emissions. Therefore, in heavy congestion vs. free-flow with very high speed situations, the comparison result of which route is better in terms of fuel consumption and various emissions is somewhat varied.

35

The last case is when one route is operating under moderate congestion and the other route experiences severe congestion, typically due to an accident. This occurred on I-10 on June 28, 2007 at 2:35 PM. Although the trip on the SR-60 route took longer than it should have been, the trip on the I-10 route was much worse. It took more than twice as much travel time as the SR-60 route. This also resulted in dramatic negative impacts on energy and emissions.

3.4.2 Estimating Overall Potential Energy/Emissions Benefits

As traffic condition on freeways usually varies across different times of day, it is interesting to analyze as to how often each of the cases discussed above actually occurs in a typical time period (e.g. a month). This analysis provides a better idea on the degree to which the energy/environmentally-friendly navigation system can recommend a user to travel on the better route in terms of fuel consumption and emissions. Therefore, the analysis was performed for the I-10 and SR-60 routes used as examples above. Traffic data (in particular, the 5-minute average speed data) on these two routes were acquired from PeMS for the period from May 14 to June 14, 2007. The data were then used to estimate the travel time and fuel consumption of trips from downtown Los Angeles to Chino on both routes, assuming that a new trip pair starts at every 10 minutes throughout the days. This results in 4,601 travel time and fuel consumption estimates for each route. The travel time difference for each trip pair is calculated, which is shown in Figure 3.4. Overall, there is a chance of approximately 45% that this particular trip would require the

same amount of travel time no matter which route is taken. Between the two routes, traveling on I-10 is slightly more likely to reach the destination faster.

The fuel consumption difference for each trip pair is also calculated, and its distribution is shown in Figure 3.5. We can see that about 40% of the trips would need approximately the same amount of fuel for each kilometer the vehicle traveled.

However, the other 60% of the trips would consume different amounts of fuel per kilometer and the differences range from 0.8g/km to 9.2g/km. For an average fuel consumption of this trip (e.g. 82g/km), the saving could then be from 1% to 11%. Between the two routes, traveling on SR-60 is more likely to consume less fuel for each kilometer.

Table 3.1: Results for case studies. (a) Case 1: Free-flow vs. free-flow (05/20/2007 01:15:00; (b)Case 2: Free-flow vs. moderate congestion (06/09/2007 12:20:00); (c) Case 3: Free-flow vs. heavy congestion (06/17/2007 17:10:00); (d) Case 4: Moderate vs. heavy congestion (06/28/2007 14:35:00).

(a):

Measure	SR-60 Route	I-10 Route	% Diff (SR-60 as base)
Distance (km)	44.9	44.0	-2.0%
Travel time (min)	23.6	22.7	-3.8%
Fuel consumption (g)	3,780	3,767	-0.3%
CO_2 (g)	11,600	11,531	-0.6%
CO (g)	133.8	143.1	7.0%
HC (g)	3.90	3.99	2.3%
NO_x (g)	9.80	9.82	0.2%

(b):

Measure	SR-60 Route	I-10 Route	% Diff (SR-60 as base)
Distance (km)	44.9	44.0	-2.0%
Travel time (min)	23.6	29.2	23.7%
Fuel consumption (g)	3,800	3,459	-9.0%
CO_2 (g)	11,645	10,704	-8.1%
CO (g)	141.2	88.0	-37.7%
HC (g)	3.97	3.09	-22.2%
NO_x (g)	9.90	8.77	-11.4%

(c):

Measure	SR-60 Route	I-10 Route	% Diff (SR-60 as base)
Distance (km)	44.9	44.0	-2.0%
Travel time (min)	24.0	41.1	71.3%
Fuel consumption (g)	3,718	4,043	8.7%
CO_2 (g)	11,438	12,480	9.1%
CO (g)	122.4	119.1	-2.7%
HC (g)	3.73	3.58	-4.0%
NO_x (g)	9.59	9.57	-0.2%

(d):

Measure	SR-60 Route	I-10 Route	% Diff (SR-60 as base)
Distance (km)	44.9	44.0	-2.0%
Travel time (min)	41.0	93.4	127.8%
Fuel consumption (g)	3,361	4,777	42.1%
CO_2 (g)	10,505	14,918	42.0%
CO (g)	46.3	63.0	36.1%
HC (g)	2.36	2.83	19.9%
NO_x (g)	8.09	9.24	14.2%

(a):

(b):

(c):

(d):

Figure 3.3: (a): Road network for case study 1: free-flow conditions on both freeways between Los Angeles and Chino; (b): Road network for case study 2: free-flow conditions on SR-60 and moderate congestion on I-10; (c): Road network for case study 3: mostly free-flow condictions on SR-60 and heavy congestion on I-10; (d): Road network for case study 4: moderate condiction on SR-60 and heavy congestion on I-10. Note that roadway links are colored based on their real-time speeds(0-40 km/h" magenta; 40-65 km/h: orange; 65-90 km/h: blue; 90+ km/h: green). This also applies to (b), (c), and (c).
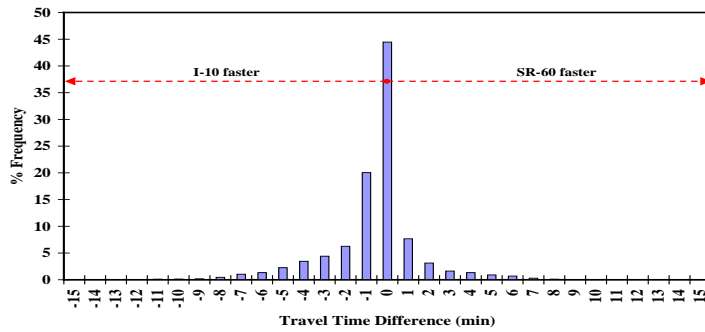
Figure 3.4: Histogram of travel time difference between the two routes during May 14-June
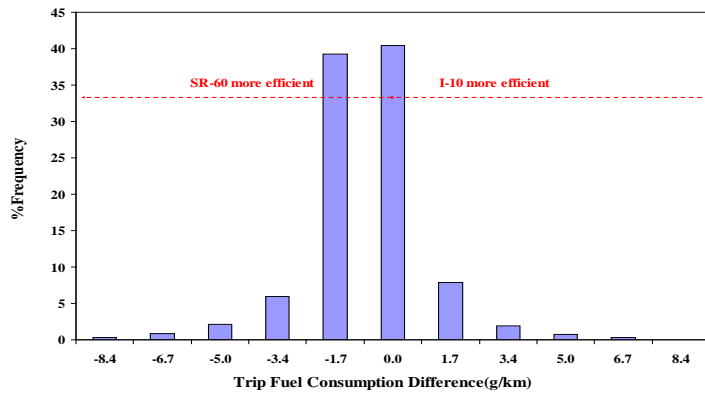
14, 2007.

Figure 3.5: Histogram of trip fuel consumption difference between the two routes during May

14-June 14, 2007.

Trip Fuel Consumption (g/km) on I-10

	76-78	78-80	80-82	82-84	84-86	86-88	88-90	90-up
76-78	0.1	2.4	0.7	0.5	0.5	0	0	0
78-80	0	2.0	6.5	2.4	1.2	0.2	0	0
80-82	0.1	1.3	7.3	13.8	1.9	0.1	0	0
82-84	0	0.1	2.3	23.0	13.5	0.2	0	0
84-86	0	0	1.0	4.4	11.7	0.5	0	0
86-88	0	0	0.4	0.4	0.8	0.4	0	0
88-90	0	0	0.1	0	0	0	0	0
90-up	0	0	0	0	0	0	0	0

Trip Fuel Consumption (g/km) on SR-60 (row axis label)

Figure 3.6: Percentage frequency distribution of the paired trip fuel consumption.

The next step in the analysis is to examine in detail how frequently and how much the fuel consumption of these two routes differs in each bin. In this step, the fuel consumption estimates for each trip pair were allocated to the corresponding cells in the frequency distribution matrix shown in Figure 3.6. The frequency value in each cell was then normalized by the total number of trips (i.e. 4601). The diagonal cells of the matrix in Figure 3.6 refer to those trip pairs that would not have significant fuel consumption difference no matter which route is taken. These trip pairs account for about 44% of the total simulated trips. On the other hand, the remaining 56% of the trips (which fall in the non-diagonal cells of the matrix) would have significantly different fuel consumption depending on which route is taken. For these trips, the energy/environmentally-friendly navigation system could direct the driver to a route that helps save fuel.

Figure 3.7 presents the potential fuel saving due to the use of the energy/environmentally-friendly navigation system. The curves in Figure 3.7 are derived from the fuel consumption vs. speed relationship shown in Figure 3.1. It can be seen that the fuel saving is highly dependent on traffic conditions (i.e. average traffic speed). In the low speed regime (<72 km/h), a fuel savings will be achieved if the average speed on one route is higher than the other. This is opposite for the high speed regime (>72 km/h). Further, an 8 km/h speed difference that occurs at moderate congestion (e.g. 80 km/h) does not give the same amount of fuel saving as an 8 km/h speed difference that occurs at free-flow condition (e.g. 120 km/h). These are all due to the non-linear shape of the fuel vs. speed curve.

In Figure 3.1, the bottom of the curve seems somehow flat, a question might be asked: will the fuel consumption corresponding to average speeds between 48 km/h and 112 km/h be different? The answer is true. This speed range corresponds to the final speed values (x axis in Figure 2.7) of between 64km/h and 96km/h. The red solid line in Figure 3.7 suggests that the fuel saving in this speed regime could be from 2% to 10%. For the non-flat part of the curve in Figure 3.1, a significant amount of fuel saving (larger than 10%) could happen if, for example, we lower the speed higher than 112 km/h down by 16 km/h, or increase the speed lower than 44 km/h up by 16 km/h.
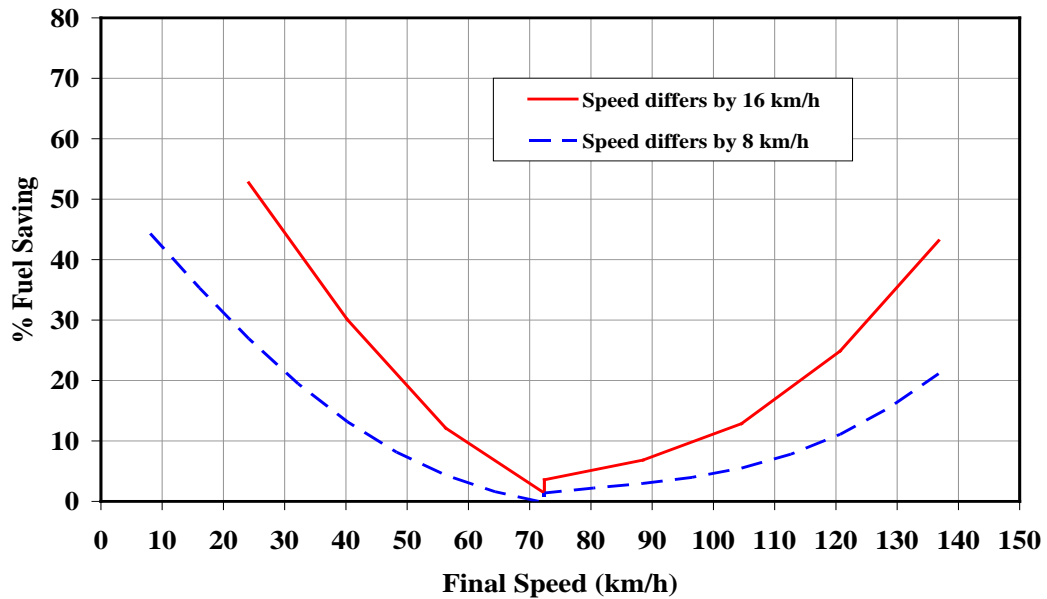
Figure 3.7: Potential fuel saving due to the use of energy/environmental-friendly navigation.

By combining the results in Figure 3.4, 3.5, 3.6 & 3.7, it is estimated that for this particular example the energy/environmentally-friendly navigation system could potentially help save fuel by an average of 1.4 g/km, if the more fuel efficient route for each trip pair is always selected for all the 4601 trip pairs. It is important to note that this estimate is based on the assumption that the probability of making a trip is equal across the days. This assumption unintentionally causes an underestimation of the fuel saving benefit. If only peak periods (morning and afternoon commutes) were used in the analysis, the potential fuel saving would have been much higher because the traffic conditions of these two routes are likely to be different more frequently.

4. Vehicle Trajectory-Based Road Type and Congestion Recognition using Wavelet Analysis

4.1 Introduction

From Section 2.4 we know that there are basically two kinds of data sources for measuring roadway-traffic performance. One is macroscopic traffic data, which can be collected from fixed embedded inductive loop detectors. Data from these detectors are very useful for transportation officials who are interested in historical and current roadway performance, in order to better manage traffic. In addition, these systems are beneficial for drivers who want to see where congestion is occurring in real-time and want to estimate travel time for interested link segments via ATIS technology, i.e. navigation tools.

There are still many roadways that do not have embedded loop detectors, or they haven't been hooked up yet to a traffic management center. As a duality and complement to fixed embedded loop sensor systems, *probe vehicles* can also be used to gather information on the roadway. In general, probe vehicles carry on-board instrumentation to measure their own velocity trajectories and/or travel times as they traverse the roadways; in many applications, the probe vehicles' information can be used as a surrogate for average traffic data, such as average speed or travel time on a section of roadway. The

measurements can either be relayed back to a management center in real-time via some type of communication, or be stored locally for later analysis.

In addition to simply estimating average traffic speed, probe vehicle trajectory data can be interpreted to determine further characteristics of traffic and the roadway. Rather than transmitting velocity and position trajectory information to a centralized location for analysis (requiring high communications bandwidth), it is possible to perform on-board trajectory interpretation for purposes of determining roadway type and localized congestion conditions. This information can be useful in several applications:

1) *centralized traffic management systems:* sending locally-processed data greatly reduces the bandwidth required when probe vehicle relay information back to a centralized traffic management center;

2) *de-centralized traffic management systems*: there has been a good deal of interest lately in eliminating the need for a centralized traffic management system; instead it has been proposed to have vehicles communicate among themselves to determine traffic conditions (see, e.g. [22] and [23]).

3) *in-situ emissions and/or energy management control systems:* based on the type of driving the vehicle is experiencing, intelligent control systems can modify their internal emissions control system parameters to minimize pollutants and energy consumption. Further, for hybrid vehicles, parameterized driving information can be used to optimize their energy management systems.

4) *energy consumption and emissions estimation:* understanding fleet-based fuel consumption and developing an emissions inventory as a function of traffic conditions is important in many congested cities.

5) *traffic simulation model calibration:* the understanding of real-world second-by-second velocity patterns associated with different roadtype/congestion levels can potentially be used to calibrate microscopic traffic models implemented for the same network. Default vehicle behavior parameters (e.g., car-following logic) in these models typically need to be calibrated for local traffic conditions.

In this chapter, a method is described to recognize roadway types and congestion conditions simply by analyzing short-term vehicle velocity trajectories (i.e., 128-second snippets). Normally, a location system can be used to determine roadway type and congestion levels when tied in with a traffic information system (e.g., based on embedded loop detector data). However, we propose that roadway type and congestion conditions can be determined without these external systems, making the task simpler and more robust by analyzing snippets of local velocity patterns.

The implementation hardware for this application should be made small, economic, and easy to use. Therefore, the associated algorithms should have low complexity. In our implementation, Haar waveform transformations are computed and principal component analysis (PCA) is used on the leading wavelet coefficients, meeting the criteria for low complexity. The hardware can consist of a simple embedded system with sufficient memory in order to extract the first several wavelet coefficients and then using these coefficients for further processing. After evaluating this initial approach, it was found that

46

better recognition could be achieved by using the Adaboost algorithm [27-28] to intelligently select the wavelet coefficients that are important for classification, rather than ignoring the tail wavelet coefficients in the PCA approach.

An introduction of PCA and the reason why Harr wavelet is suitable for this application is given in Section 4.2. Section 4.3 that describes the proposed methodology including PCA and the Adaboost algorithm to recognize roadway facility type and congestion levels. Section 4.4 describes overall results. Section 4.5 of the chapter presents the conclusions and future work.

4.2 Background

One method to deal with the problem of high dimensionality is to reduce the dimensionality by combining features. Linear combinations are preferable because they are simple to compute. Principal components analysis (PCA) is one kind of such linear combination methods and it seeks a projection that best represents the data in a least square sense. PCA finds components that are useful for representing data. The main application of PCA is to reduce the dimensionality of data while keep as much information as possible. Since an eigenvalue decomposition of the covariance matrix is necessary, the computation complexity can be high. Further details on the technique can be found in [35].

In the case of using vehicle velocity trajectories as the time-series data, there are two key properties: (1) a vertical shifting variance; and (2) a trend variance. It is assumed that average speed of the short vehicle velocity trajectory is an important indicator of the road

type. Therefore, roadtype classification is highly sensitive to vertical shifting. Tilting the sequence also affects roadtype classification greatly. We find that the Haar wavelet transform is good for roadtype classification since its first coefficient is the average of the data. Haar coefficients are also sensitive to the general trend because tilting the sequence affects all the Haar coefficients [36]. In this research, we first use discrete wavelet transform for feature extraction. The resulting feature vectors have many useful properties. They give us a good representation of the natural features of the original time series, not only at small scale, such as sharps peaks, but also at large scale, such as wide mountains. After the feature extraction of wavelet transform, initially PCA is applied to wavelet subbands to extract features further and therefore reduce the dimensionality. Later, it was found that if we use the Adaboost algorithm to intelligently select features (wavelet coefficients), better classification results can be achieved.

4.3 Methodology

4.3.1 Method using DWT and PCA

Figure 4.1 illustrates the overall process of calculating the wavelet transform and PCA on the fly for vehicle velocity trajectory snippets (128 seconds). Data are continuously acquired, typically through an interface to the vehicle's on-board bus that monitors the vehicle velocity as one of many parameters. This time-series data are pre-processed, applying noise filtering. The discrete wavelet transform is first applied, which extracts a set of features from each time series. Based on these feature vectors, we can do

further analysis for classification. This is indeed a problem of dimension reduction. The combination of discrete wavelet transform combined with principal components analysis is applied to achieve this dimensionality reduction. After the feature vectors are extracted (dimension reduction), the classification step follows. There are several classification methods, such as multiclass LDA, nearest-neighbor matching, and neural networks. A multiclass LDA approach tries to project the data in such directions that are efficient for discrimination. For LDA, the data have to be linear discriminable. A K-nearest-neighbor classifies a time series by assigning the label most frequently represented among the k nearest samples in the training set. In other words, a decision is made by looking at the labels on the k nearest neighbors and choosing the majority. For a neural network approach, many layers might be required and a large amount of neurons might be needed. Further, the convergence might be very slow and the computation complexity is high. In this application, a nearest-neighbor algorithm is applied. In the last step, it is desired to estimate the roadway type and congestion level. Note that training data are required for this recognition.
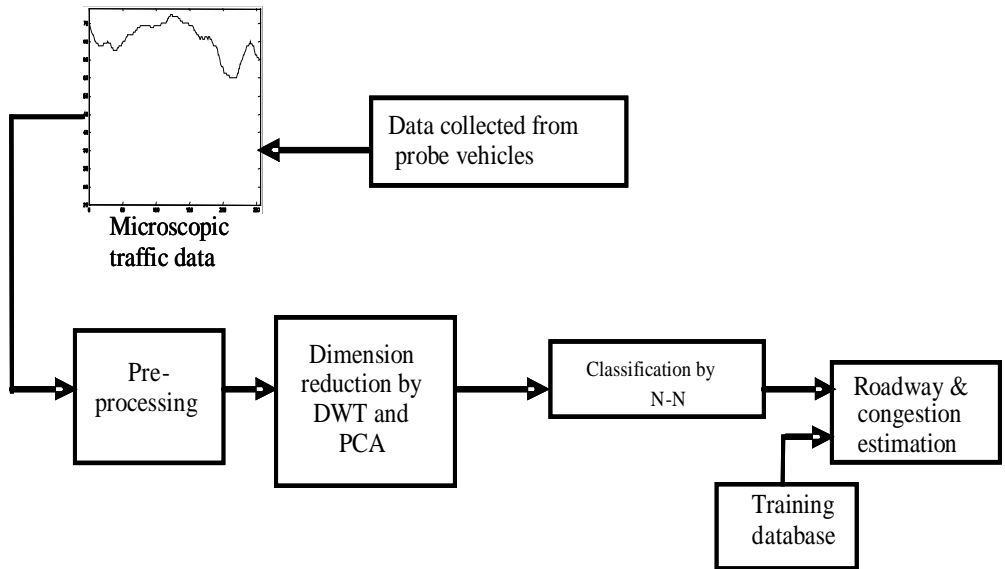
Figure 4.1: Process diagram of the project.

In the proposed methodology, the Haar DWT transform is used as the feature extractor for the roadtype and congestion classification. Subsequently, PCA is applied to further reduce the dimensionality. In the road type and congestion classification problem, the probability distribution of the velocity speed is unknown due to many uncertain factors affecting the velocity speed, such as general traffic interaction, time of weekday or weekend, weather conditions, and personal driving behavior. We do not have prior information about the probability density function. Therefore, nearest neighbor is chosen as the classification rule in the final step, which is a nonparametric technique and doesn't assume a probability distribution.

Adopting the same metric as in [31], [32], we use the square root of the sum of squared difference as distance function to measure the similarity between two sequences. The distance between two time sequences $\mathbf{x}, \mathbf{y}$ of equal length n is

$$D(\mathbf{x},\mathbf{y}) = \left(\sum_{i=0}^{n-1} (y_i - x_i)^2 \right)^{1/2} \tag{4.1}$$

A nearest neighbor method is used instead of setting a hard threshold ε to determine "similar" sequences. It should be noted that for other applications, such as range query in database, hard thresholds for decision of similar sequences are typically used. In [32], it is shown that the Euclidean distance is preserved under Haar wavelet transforms and using the first several coefficients of Haar transform will not cause false dismissal for a range query. In this chapter, based on the multi-resolution analysis and the Haar transformation matrix described in the previous section, we show a similar idea in a much simpler way. Although we are not using a hard decision rule, we still include the proof of no false dismissal here for completeness.

Lemma 1. Euclidean distances of both coefficients in time domain and in transformed wavelet domain are the same for Haar DWT transformation. In other words, Euclidean distance is preserved under Haar DWT transformation.

Proof: from last section, we have $f_n = W_n W_{n-1} \cdots W_1 f$, where W_i is unitary matrix.

$$\| WT(f_n^1) - WT(f_n^2) \|^2$$
$$= \| W_n W_{n-1} \cdots W_1 f^1 - W_n W_{n-1} \cdots W_1 f^2 \|^2$$
$$= \| Q_n f^1 - Q_n f^2 \|^2 = \| Q_n (f^1 - f^2) \|^2$$
$$= (f^1 - f^2)^H Q_n^H Q_n (f^1 - f^2)$$
$$= \| (f^1 - f^2) \|^2$$

Lemma 2. If only the first several dimensions of the Haar transformation coefficients are selected, no false dismissal will occur.

Proof: Considering the inequality $D(\mathbf{x},\mathbf{y}) \leq \varepsilon$

51

If we only select the first h dimensions of the transformation coefficients,

$$D(\mathbf{x}, \mathbf{y}) = D(WT(\mathbf{x}), WT(\mathbf{y}))$$
$$= D(WT(\mathbf{x})(1:h), WT(\mathbf{y})(1:h)) + \Delta$$
$$\leq \varepsilon$$

where Δ is a nonnegative number. We employ the MATLAB syntax where $WT(\mathbf{x})(1:h)$ means the 1 to h rows of $WT(\mathbf{x})$. Therefore, we have $D(WT(\mathbf{x})(1:h), WT(\mathbf{y})(1:h)) \leq \varepsilon$, which guarantees no false dismissal.

We use Haar wavelet transform before PCA considering the computation complexity. The computation complexity of PCA is very high. The computation of computing the eigenvector from a covariance matrix of a dataset is $o(N^3)$, where N is the dimension of the dataset. If the number of training vectors M is less than N, the computation complexity is reduced to $o(M^3)$. Therefore, the computation complexity can be expressed as $o(p^3)$, where $p = \min(M, N)$.

4.3.2 Method using DWT and Adaboost

The method in the previous section assumes that the tail wavelet coefficients of the signal are not important for the classification. However, we know that this assumption is not strong, or at least not guaranteed.

In terms of the feature selection process, boosting algorithms were originally proposed by Schapire and Freund [27], [28]. The most recent version, referred to as Adaboost, has been shown to be very effective. These boosting algorithms have very broad applications, including face recognition [29], medical imaging [30], etc. These methods are effective for improving the accuracy of a learning algorithm. It produces a

very accurate classification rule by combining weak or moderately accurate rules with accuracy only slightly better than random guessing. It repeatedly reweights the examples in the training set and reruns the weak learning algorithms on these reweighted training examples.

In this section, the Adaboost algorithm is used instead to intelligently select the wavelet coefficients that are important for classification, i.e. as part of the feature selection process. The value of this type of boosting algorithm is that it effectively forces the weak learning algorithm to concentrate on the hardest examples. The final combined hypothesis is a weighted vote of the weak hypotheses. If we assume $h_1, h_2, ..., h_T$ is a set of weak hypothesis, then

$$f(x) = \sum_{t=1}^{T} \alpha_t h_t(x)$$

where α_t is the weight of each weak hypothesis within the set, $f(x)$ is the final classification hypothesis.

In real world, many problems are multiclass classification problems, such as roadtype/congestion classification in this chapter, handwritten digit recognition, vehicle classification, etc. Multiclass learning problems can be generally solved by reducing them into binary classification problems. Several ways are typically used to achieve this. The most straightforward reduction is to use a binary learner to learn each individual class. Therefore, this approach learns k individual binary functions $f_1, \cdots, f_k$, one for each class. Finally we will have a collection of hypotheses, each of which tries to predict whether an instance belongs to one particular class. To assign a new item, X, to one of the

53

classes, all f_i will be evaluated for *X*, and *X* will be assigned class *j* with highest $f_i(X)$. A separate rule will be used to break ties. This is called one-per-class approach, following the terminology of [37].

Another approach, first proposed by Sejnowski [38], is to associate with each class a unique binary string of length *n*, which is referred to as "codewords". One hypothesis is learned for each of the *n* bit positions. When we train for an example from class *k*, we specify the desired outputs of these *n* hypotheses by the *n*-bit binary string associated with class *k*. A test instance is then classified to belong to the class whose codewords (*n*-bit string) is closest in hamming distance to the string of predictions generated by the *n* hypothesis.

Based on this approach, Dietterich's algorithm [39] tries to pick the codewords associated with the classes to belong to error-correcting codes which have some special properties in communication theory. This error-correcting code approach suggests that we can view the machine learning as a kind of communications problem. The true identity that is the correct output class for a new instance is transmitted over a channel. The channel generally can be understood to consist of the features selection process, the collection of training examples, and the learning algorithms. The class information is corrupted due to the errors introduced by the imperfect channel. If we encode the class into error-correcting codes and transmit each bit sequentially, we might be able to recover errors introduced by channel.

Motivated by the Adaboost and error-correcting algorithms, Schapire [37] developed one algorithm called Adaboost.OC, having the performance advantages of boosting while

relying only on binary weak learner. Guruswami [40] proposed a variant of this algorithm called Adaboost.ECC that uses a better weighting of the weak hypothesis, and in fact chooses the weights purely as a function of the error of the binary weak hypothesis, and therefore represents a more direct reduction of multiclass learning to binary learning.

Table 4.1 and Table 4.2 below contain a formal description of the two algorithms with application to the roadtype/congestion classification problems. In this chapter, we follow the terminology of [37] and use the notation $\|\Psi\|$ that is defined to be 1 if proposition Ψ holds, and 0 otherwise. Basically these two algorithms are given m training examples in the form (x_i, y_i) where x_i is from some space X, and the corresponding known label y_i is from a set Y with $|Y| = k$, where k is the cardinality. On each round t a distribution D_t is computed over the training examples, and a mapping function $\mu_t : Y \to \{0,1\}$, which is referred to as coloring and which divides the label set Y into two parts. The data is then relabeled according to μ_t, the k classification problem will then be reduced to binary problem, and the weak learner will be trained on this relabeled data weighted according to D_t. For example, if there are four classes, we can randomly compute the function μ_t at round t as $\mu_t : \{1 \to 0, 2 \to 1, 3 \to 0, 4 \to 1\}$. Therefore, for training example $(x,1)$, it will be relabeled as $(x,0)$, for $(x,2)$, it will be relabeled as $(x,1)$, and for $(x,4)$, it will be relabeled as $(x,1)$, Please refer to [39] for how to select the coloring function through error-correcting codes and [37], [40] for the details and theoretical analysis of these two algorithms.

Table 4.1: Schapire's algorithm.

Adaboost.OC
Given the training set: $(x_1, y_1), \cdots, (x_m, y_m) : x_i \in X, y_i \in Y$
Initialization: $\breve{D}_1(i,l) = \dfrac{1}{m(k-1)}$ if $l \neq y_i$,   otherwise $\breve{D}_1(i,l) = 0$
For t =1:T   Compute coloring $\mu_t : Y \to \{0,1\}$;   Let $D_t(i) = \dfrac{1}{U_t} \sum_{l \in Y} \breve{D}_t(i,l) \|\mu_t(y_i) \neq \mu_t(l)\|$, and U_t is a scaling factor.   Get hypothesis $h_t : X \to \{0,1\}$ from the weaker learner for distribution D_t Let $\breve{h}_t(x) = \{l \in Y : h_t(x) = \mu_t(l)\}$.   Let $\breve{\varepsilon}_t = \frac{1}{2} \sum_{i=1}^{m} \sum_{l \in Y} \breve{D}_t(i,l) \cdot (\|y_i \notin \breve{h}_t(x_i)\| + \|l \in \breve{h}_t(x_i)\|)$.   Let $\alpha_t = \frac{1}{2} \ln(\dfrac{1-\breve{\varepsilon}_t}{\breve{\varepsilon}_t})$.   Update $\breve{D}_{t+1}(i,l) = \frac{1}{Z_t} \cdot \breve{D}_t(i,l) \exp\{\alpha_t (\|y_i \notin \breve{h}_t(x_i)\| + \|l \in \breve{h}_t(x_i)\|)\}$ and Z_t is a scaling factor .
Output the final hypothesis: $H_f(x) = \arg\max_{l \in Y} \sum_{t=1}^{T} \alpha_t \|l \in \breve{h}_t(x)\|$

Table 4.2: Guruswami's algorithm.

ABOOST.ECE
Given the training set: $(x_1, y_1), \cdots, (x_m, y_m) : x_i \in X, y_i \in Y$
Initialization: $\breve{D}_1(i,l) = \dfrac{1}{m(k-1)}$ if $l \neq y_i$, otherwise $\breve{D}_1(i,l) = 0$
For t =1:T   Compute coloring $\mu_t : Y \to \{0,1\}$;   Let $D_t(i) = \dfrac{1}{U_t} \sum_{l \in Y} \breve{D}_t(i,l) \| \mu_t(y_i) \neq \mu_t(l) \|$, and U_t is a scaling   factor.   Compute hypothesis $h_t : X \to \{0,1\}$ from the weaker learner   Compute the weight of positive and negative votes α_t and β_t respectively: $\alpha_t = \beta_t = \frac{1}{2} \ln \left(\dfrac{\sum\limits_{i:h_t(x_i)=\mu_t(y_i)} D_t(i)}{\sum\limits_{i:h_t(x_i)\neq\mu_t(y_i)} D_t(i)} \right)$ Define $g_t(x) = \alpha_t$ if $h_t(x) = +1$ $g_t(x) = -\beta_t$ if $h_t(x) = 0$ Update $\breve{D}_{t+1}(i,l) = \frac{1}{Z_t} \cdot \breve{D}_t(i,l) \exp\{(g_t(x_i) \cdot (2\mu_t(l) - 1)$ $- g_t(x_i) \cdot (2\mu_t(y_i) - 1))/2\}$ and Z_t is a scaling factor.
Output the final hypothesis: $H_f(x) = \arg\max_{l \in Y} \sum\limits_{t=1}^{T} g_t(x)\mu_t(l)$

4.4 Experiments and Results

In order to determine the recognition capability of our proposed method, probe vehicle data were necessary for establishing a training set. As part of another vehicle activity study, probe vehicle velocity data were collected for approximately 40 days in Southern California for many different roadway types (e.g., freeways, arterial roads, and residential streets) under different congestion conditions. The data were collected using a GPS receiver on a set of vehicles, with a sampling frequency of 2 Hz. where both

57

position and velocity information were acquired. The data collection was conducted between 8AM and 8PM and consisted of a variety of congestion conditions while excluded extreme conditions such as gridlocked traffic. The datasets were subsequently post-processed, identifying the roadway type and congestion level based on map-matching and an existing traffic performance system (see, e.g., [15] and [41]).

As part of the experimentation, three things were investigated for the wavelet/PCA algorithm. In the previous section, it was claimed that the Haar transform of the velocity trajectory data concentrates the energy of the signal in the first few coefficients. Next, it was of interest to determine how the recognition precision varies with the PCA level. Finally, it is desired to determine how the recognition rate depends on the number of the sequences in the training dataset.

It is well known that the magnitudes of wavelet coefficients of a signal at each resolution level are proportional to the corresponding energy in the signal. In Figure 4.2, the energy concentration is plotted as a function of resolution. Fifty-four sequences for each class were used in the training set for this experiment. Each sequence has a dimension of 256. Resolution "0" corresponds to the low frequency component, and resolutions 1 through 8 correspond to the high frequency components. The low frequency component and resolution 1 through 8 together make the Haar wavelet transform. For this analysis, concentration is defined as follows:

$$concentration = \frac{Sum_{amp\_in\_a\_resol}}{Sum_{all} \times number_{coef\_in\_a\_resol}} \tag{4.2}$$

where $Sum_{amp\_in\_a\_resol}$ is the sum of all the values of amplitudes with a resolution, Sum_{all} is the sum of all the amplitudes of all resolutions, and $number_{coef\_in\_a\_resol}$ is the number of coefficients within a resolution.

We can see from Figure 4.2 that for three roadway types with mixed congestion levels, they all have their energy concentrated in the first few scales. This verifies our original hypothesis. It is interesting to note that within the three types of roads, freeway roadways concentrate their energy in the first few coefficients much more than arterial and residential streets. There is no significant difference between the arterial and residential streets regarding energy concentration.
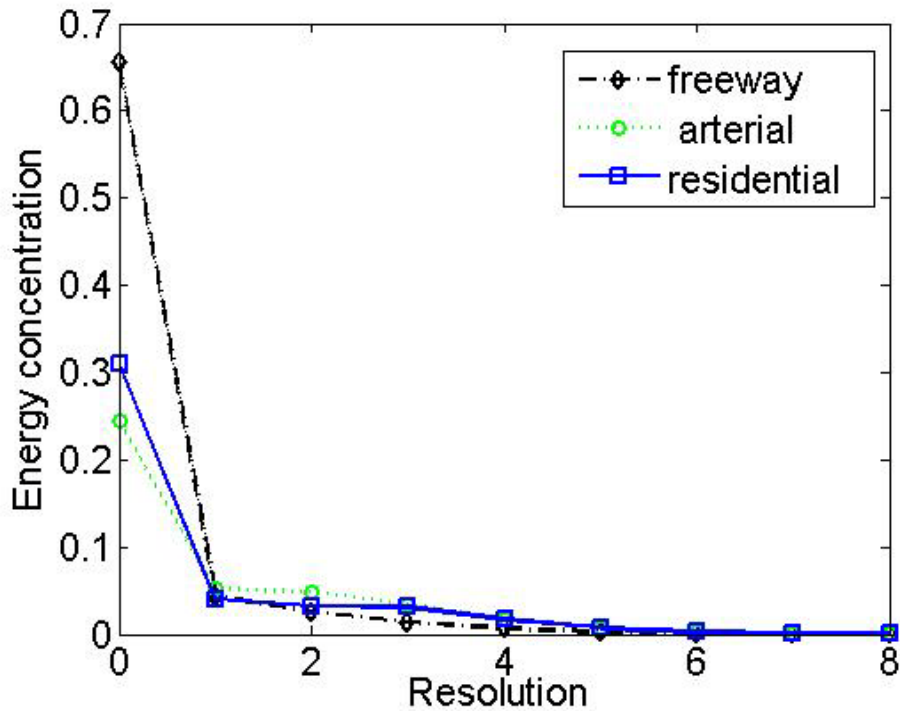

Figure 4.2: Energy concentration per coefficient vs. resolution.

In the following experiments, the cut off resolution is chosen at 4 since the coefficients after resolution 4 are negligible. We use the overall recognition rate as a major performance measurement for this application of roadtype classification, which is defined as

$$recognition/congestion = \frac{Num_{correct}}{Num_{total}} \qquad (4.3)$$

where $Num_{correct}$ is the number of sequences that are correctly classified, and the Num_{total} is the total number of testing sequences.

In this experiment, 300 sequences are used for training purposes, 100 of each class and 150 sequences for testing, 50 of each class. Each sequence has a dimension of 256. The results are shown in Table 4.3. From Table 4.3, it can be seen that using PCA levels 6 can provide sufficient performance, and further reduce the dimension from 16 to 6.

Table 4.3: Recognition rate vs. PCA levels, 300 sequences for training, 150 for testing.

PCAlevel	2	4	6	8	10	12
Reco.Rate	0.7933	0.8600	0.9133	0.9000	0.9000	0.9267

Next, it was desired to see how the number of training sequences affects the overall recognition performance. The PCA level is set to 6. Each sequence has a length of 256, and the cut off resolution is chosen to be 4. The same testing set is used, consisting of 150 testing sequences, 50 of each class. For the training sequences, a multiple of 60 was chosen, and each class has the same testing number of sequences. It can be seen in Figure 4.3 that by a general trend, the recognition rate increases as the number of training sequences increases.
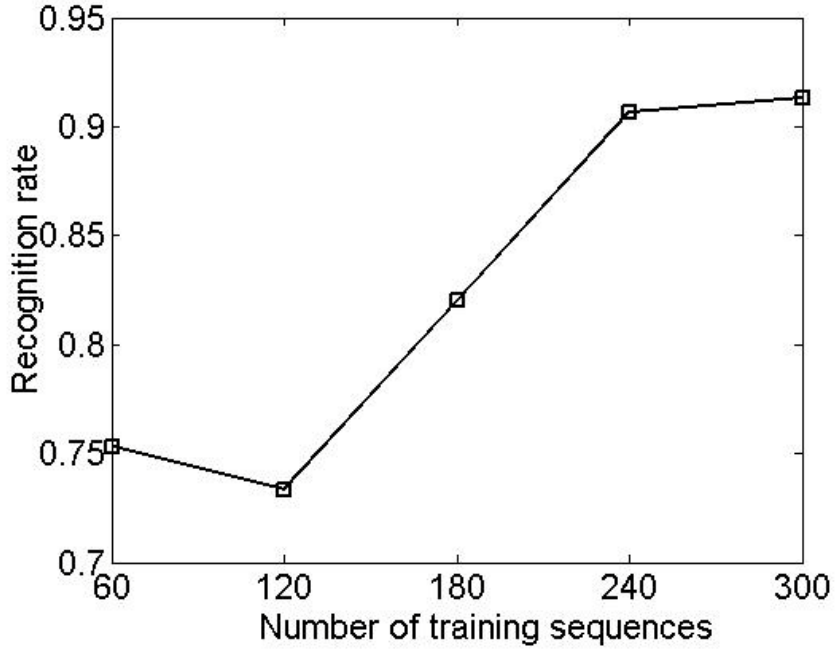
60

Figure 4.3: General trend of recognition rate vs. number of training sequences, PCA level 6.

Simulations were also performed for the Adaboost algorithm. Both Schapire's and Guruswami's algorithms were investigated. Since there are three classes, the coloring functions are randomly (with equal probability) selected as:

$$\mu_t = \begin{cases} \{1 \to 1,\ 2 \to 0, 3 \to 0\} \\ \{1 \to 0,\ 2 \to 1, 3 \to 0\} \\ \{1 \to 0,\ 2 \to 0, 3 \to 1\} \end{cases} \cdot$$

In this experiment, 562 instances are used for training, 135 in freeway, 230 in arterial, and 197 in residential road. We also used 60 items for testing, 20 of each class. The overall testing error is illustrated in Figure 4.4. It can be seen that generally the Guruswami's algorithm is more accurate than Schapire's algorithm for the roadtype/congestion classification problem. When is the number of iterations is more

than 600 hundred, the Guruswami's algorithm is stable and can achieve an accuracy of 0.9167. It can be expected that with an increased number of training items, the accuracy will continue to increase. It can also be inferred that 600 iterations are enough to achieve good performance.
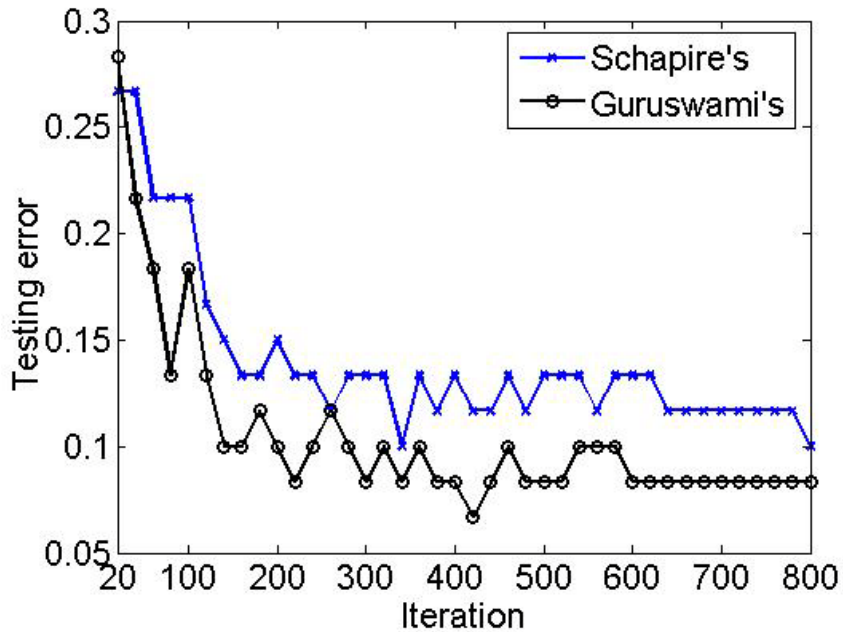

Figure 4.4: A comparison of Schapire's Adaboost.OC and Guruswami's Adaboost.ECC algorithm: overall testing error vs. the number of iterations.

Although it is observed that there is not an apparent advantage of the wavelet/Adaboost algorithm over the wavelet/PCA algorithm in terms of the performance, the Adaboost algorithm is more suitable for real-time processing. For the PCA algorithm, the nearest-neighbor algorithm is necessary for the final classification, i.e., when each test instance is classified, it is necessary to compute the Euclidean distance between this test instance's final computed features and all the training instances'. For each test instance,

it is necessary to process the entire dataset, even though the training step was previously completed. The complexity will grow rapidly with the size of the training set. For the Adaboost algorithm, once the trained simple final hypothesis is computed, the computation will no longer depend on the training dataset. For each test instance, it is not necessary to go over the whole training dataset. Therefore, we conclude that if the training dataset is small, wavelet/PCA and wavelet/Adaboost will have similar performance and complexity, however, when the training dataset is large, wavelet/Adaboost will have a large advantage over the wavelet/PCA algorithm in terms of computational complexity.

5. Defining a Freeway Mobility Index for Roadway Navigation

5.1 Introduction

As described in previous chapters, today's latest navigation systems are able to monitor real time traffic conditions, and have the capability to suggest an alternative path for re-routing if the traffic conditions are bad in a particular area. With the ability to re-route, it is useful to develop an effectiveness index that gives an indication of the possible number of routing options for an origin-destination pair. With a larger the number of the routing options, drivers have a better freedom to re-route. Further, a network with a low index value will not have many re-routing choices; and therefore, dynamic shortest-duration algorithm based on real-time traffic conditions will have limited value in this area. On the other hand, dynamic shortest-duration algorithm will have larger value in a roadway network with a high index value. In this chapter, we will develop and justify such an effectiveness index.

The remainder of this chapter is organized as follows. Section 5.2 describes the overall methodology, including the definition of node-to-node NMI, node-NMI, and area-NMI. Also, a justification is given for the proposed concepts, along with the search algorithm. Section 5.3 presents experimental results and discussions using freeway networks in California as case studies.

5.2 Methodology

In this section, we: 1) explain the network structure and the concept of critical node used, 2) propose a new mobility index for navigational purposes based on the availability of freeway facilities, 3) justify the effectiveness of the new index, 4) discuss various forms of the new index, and 5) present an algorithm to implement the index in navigation applications.

5.2.1 Network Representation

In this research, we initially consider the network connectivity of a freeway system without loss of generality when it is extended to other roadway facility types (e.g., arterials, local streets, etc.). In the experiment, we use the latest census-based TIGER/Line map roadway shapefiles as our data source [53], but any roadway network representation will also be sufficient. In the roadway shapefiles, a freeway segment is a polyline connecting one node to another. An actual freeway is typically represented by a series of segments in the network file to account for its curvature and connectivity. We define a "location" in a freeway network as a critical node, i.e. an intersection of freeways or a start point/end point of a physical roadway segment. There are a number of non-critical nodes that are used to represent only curvature in the roadway shapefiles. However, we do not consider these nodes due to the following reason. Assume one driver is traveling over a pre-selected route; according to the real-time traveler information system, the driver finds that an accident has occurred and the traffic ahead is jammed.

The driver will then want to switch to another route that has reasonable travel time and distance. However, the driver can only switch to another route at a critical node, i.e. freeway intersection (again, if not considering arterial roads).

Because of this reason, we only consider the connectivity between critical nodes of the freeway network and combine all map lines representing physical roadway segments into "edges" connecting critical nodes. An edge is defined as a roadway segment with a specific direction. A simple example is given in Figure 5.1 to illustrate the concept above. Figure 5.1 is a map representing a simple roadway network. Piecewise lines are used to represent the curvatures of roadway segment. Each line consists of a start node and an end node represented by white dots in the figure. It can be seen that one physical freeway segment consists of a series of connected lines. For example segment *2-5* of freeway *a* consists of line *2-3*, *3-4*, and *4-5*. In our modified network, these lines are combined into one segment represented by *2-5*. After combining the lines, the remaining critical nodes in Figure 5.1 are *2, 5, 6, 7, 10* and *14*.
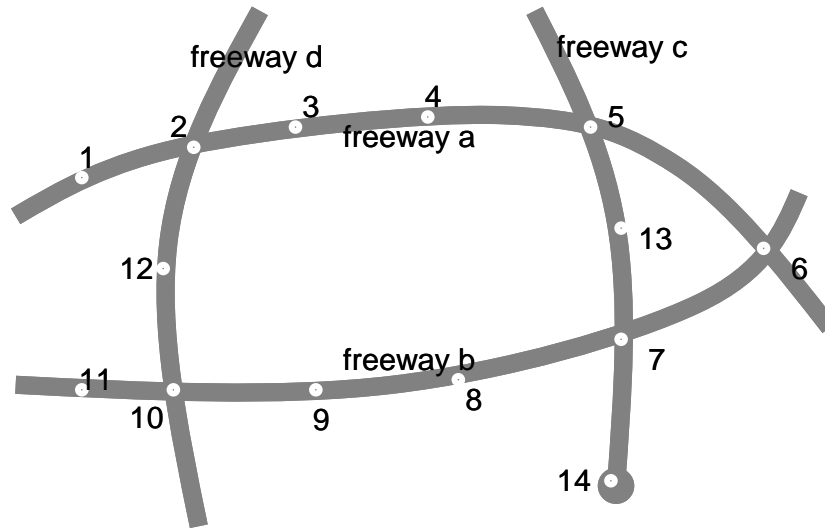
Figure 5.1: A simple example of roadway infrastructure.

In the following discussions, all of the "nodes" denote the "critical nodes" in our abstract graph, and the terms are used interchangeably. The edges denote the combined segments in the abstract graph. A graph can be represented either by an adjacent matrix or a linked list [18]. Further, a route is defined as a loopless path for a source-destination pair in the constructed network, and will be used interchangeably with the term (loopless) path. There exist several possible routes between a source destination pair, and each route normally consists of several segments.

5.2.2 Navigational Mobility Index Based on the Availability of Freeway Facilities

In order to find *K* shortest routes from node *i* to node *j,* Yen's *K*-shortest path algorithm can be used for the computation. Yen's algorithm is a deviation algorithm

where only loopless paths are determined. For a network consisting of *N* nodes, we first summarize the notations in Table 5.1. A brief summary of Yen's algorithm is then given in Table 5.2. Please refer to [54] for the details and analysis of this algorithm. The computation complexity of this algorithm is $O(Kn(e + n\log n))$ if the Dijkstra algorithm is used to find the shortest path, where *e* is the number of edges and *n* is the number of nodes in the network. Surprisingly, there has been no significant improvement in terms of worst computation complexity since its first development in the 1970's. Please refer to [55] for a slightly faster implementation of Yen's algorithm using new data structures.

Table 5.1: Selected notations of Yen's K-shortest path algorithm.

$(i), i = 1, 2, \cdots, N.$ is used to denote the node of a path, where (1) is the origin and (N) is the destination, i.e. (1) is the first node of a path and (N) is the last node of a path;
$(1) - (i) - \cdots - (j), i \neq j \neq \cdots \neq 1$ is used to represent a path from (1) to (j) passing through (i) and the following nodes;
$d_{ij} \geq 0, i \neq j$ is the distance of the direct arc from (i) to (j) if the arc exists, otherwise it is infinite;
$A^k = (1) - (2^k) - (3^k) - \cdots (Q_k^k) - (N), k = 1,2,\cdots, K$ be the *k*th shortest path from (1) to (N) where $(2^k),(3^k),\cdots,(Q_k^k)$ are the 2nd 3rd,…, $Q_k th$ node of the *k*th shortest path and $Q_k = N - 1$.

Table 5.2: The procedure of Yen's algorithm.

Iteration 1. Determining A^1 using any shortest path algorithm, such as Dijkstra or A* algorithm [23].

Iteration k $(k=2, 3, \cdots, K)$. Determining A^k. This is an iterative method, therefore in order to find A^k, all previous routes $(A^1, A^2, \cdots A^{k-1})$ must have been computed. A^k may be computed by the following steps:

1) For each of $i = 1, 2, \cdots, Q_{k-1}$ do the following:

    a. Updating the network. Check if the subpath consisting of the first i nodes of A^{k-1} coincide with the subpath consisting of the first i nodes of A^j, $j = 1, 2, \cdots k-1$. If it coincides any of those subpath, set d_{iq} of those $A^j s$ to infinity, where *(q)* is the *(i+1)th* node of A^j, otherwise make no changes.

    b. Applying a shortest-path algorithm to find the shortest path from *(i)* to *(N),* allowing it to pass only those nodes that are not yet included in the path to avoid loops. The subpath from *(1)* to *(i)* is called R_i^k and the subpath from *(i)* to *(N)* is called S_i^k.

    c. Constructing A_i^k by joining R_i^k and S_i^k and add A_i^k to list B.

2) Finding from List B the path(s) that have the minimum length.
If the path(s) found plus the path(s) already in List A exceed *K*, we are done. Otherwise denote this path by A^k and move it from List B to List A. Keep the rest of the paths in List B still in List B. Go to iteration *k+1*.

Dial [56] proposed a concept of "efficient path" assuming that drivers are supposed to select efficient paths. A path is efficient "if every link in it has its initial node closer to the origin than is its final node, and has its final node closer to the destination than its

initial node" [56]. Paths selected by the *K* shortest path algorithms might not be efficient

paths. However, in this chapter we consider the situation where severe traffic conditions

occur, and we believe it is a common practice nowadays that a driver is willing to

backtrack slightly to avoid severe traffic conditions.

5.2.2.1 Definition and Formulation of NMI

Motivated by Yen's *K*-shortest path algorithm [54], the NMI *for a source-destination*

pair is defined in a recursive way as follows. The NMI is the cumulative weighted

number of paths of all possible routes for a source-destination pair. First, we assume that

the shortest path is the only freeway facility available for this node pair, and we next

consider the contribution of the next-shortest path if it is added to the existing facility

considering the possible overlapping segments with the previous shortest path, and so on.

We first use Yen's algorithm to find *K* shortest paths (the shortest, the second shortest,

and so on), and store them in a set **setKPaths** in an increasing order (according to the

path length, i.e. A^1 for the shortest one, and A^2 for the second shortest one, and so on).

Each path normally consists of multiple segments. We proceed in the order from

A^1 to A^K. Next, we have another set called **segmentsIncluded** to collect all road

segments whose contributions to NMI have been counted. The set **segmentsIncluded** is

initially empty. The detailed steps are given below.

1. We begin from the shortest route A^1: the contribution (weighted number of path)

   of A^1 to the NMI is calculated as 1 since none of its segments is in set

segmentsIncluded. We then add all the segments of this route to the set **segmentsIncluded**.

2. For the second shortest route A^2, its contribution to the NMI is calculated as

$$\sum_{l_a \in L_2} \frac{|l_a|}{|L_2|} G_{l_a} \tag{5.1}$$

where L_2 is the collection of all segments of route A^2, $|L_2|$ is the total length of route A^2, and $|l_a|$ is the length of segment l_a. G_{l_a} is a general function depending on l_a and the set **segmentsIncluded**. In this chapter we specifically choose δ_{l_a} for G_{l_a}, where δ_{l_a} is a function whose value is 1 if segment l_a is not already in the set **segmentsIncluded,** and is 0 if otherwise. We then add all $L_2's$ segments into the set **segmentsIncluded**.

3. Repeat step 2 for all the remaining shortest paths.

4. Compute the total NMI of the freeway facility available for this source-destination pair as

$$NMI = \sum_{k=1}^{K} \sum_{l_a \in L_k} \frac{|l_a|}{|L_k|} G_{l_a} \tag{5.2}$$

We specifically choose δ_{l_a} for G_{l_a} because it is simple and intuitive. Using δ_{l_a}, we do not consider the contribution to NMI of a new route's segments that has already been counted in previous routes. However, in order to grant some small weighted contribution

71

of these segments, G_{l_a} can be selected as a general function inversely proportional to the number of occurrences of l_a in the set **segmentsIncluded.** For example, we can construct:

$$G_{l_a} = \beta^{-O(l_a)}$$

where $O(l_a)$ is the number of occurrence of l_a in the set **segmentsIncluded**, and β is a constant larger than 1 (e.g. 2).

When drivers are selecting routes, unreasonably long routes are usually not considered. Therefore, in the calculation of K shortest paths, another criterion is implemented: whenever an mth ($m < K$) shortest route whose length is longer than $ratio * Length_{shortest}$ is found, the K-shortest path algorithm stops. For a moderately complex freeway network in an urban area, a large value of K is usually selected and the *ratio* parameter is normally set at around 1.5. Therefore, the K-shortest path algorithm usually stops *before* the Kth shortest path is found.

Even for the routes whose lengths are shorter than $ratio * Length_{shortest}$, their attractiveness to drivers could still be different due to having different lengths; i.e., a shorter route might be more attractive than a longer one. We define an *appealing function* to quantify the attractiveness of a route to drivers in terms of the route length. Therefore, the appealing function should have a maximum value of 1 for the shortest route and a minimum value of 0 when the route length is longer than a threshold (i.e. $ratio * Length_{shortest}$). For example, we could have the following three simple forms of this function:

$$P_0(x) = \begin{cases} 1 & x \le ratio \\ 0 & otherwise \end{cases},$$

$$P_1(x) = \frac{1}{(ratio - 1)^2}(x - ratio)^2,$$

or, $P_2(x) = 1 - \dfrac{1}{(ratio - 1)^2}(x - 1)^2$

where $P_j(x)$, j = 0, 1, 2 is the appealing function, and x is the path length divided by the length of the shortest path. For $P_0(x)$, it is assumed that routes with different lengths are equally appealing if their lengths are shorter than $ratio * Length_{shortest}$. For $P_1(x)$ and $P_2(x)$, shorter routes are generally considered more attractive than longer routes. We could also apply other appealing functions by using the route choice model. There is a large body of literature on discrete route choice model, e.g. [51, 52]. However, we do not encourage incorporating the complex route choice model into the NMI definition. First, the NMI definition is simple and intuitive; incorporating the route choice model will incur difficulty in understanding the concept as well as introducing greater complexity in computation. Second, in the route choice model, a similar concept of "path-size" was proposed. We think it is problematic to use some existing concepts to construct a new concept achieving similar functionalities.

In the definition of NMI described earlier in this section, we implicitly used the $P_0(x)$ as the appealing function. However, we could also use other appealing functions, such as $P_1(x)$, $P_2(x)$, or other functions by modifying (5.2). That is,

$$NMI = \sum_{k=1}^{K} P(x_k) \sum_{l_a \in L_k} \frac{|l_a|}{|L_k|} G_{l_a}$$

where $x_k = \frac{|L_k|}{|L_1|}$ and $|L_1|$ is the length of the shortest route. Throughout this chapter, we use $P_0(x)$ as the appealing function for illustration purpose due to its simplicity. Therefore, the appealing function will not be explicitly included in calculation due to the property of $P_0(x)$ (i.e. being 1 if within a range).
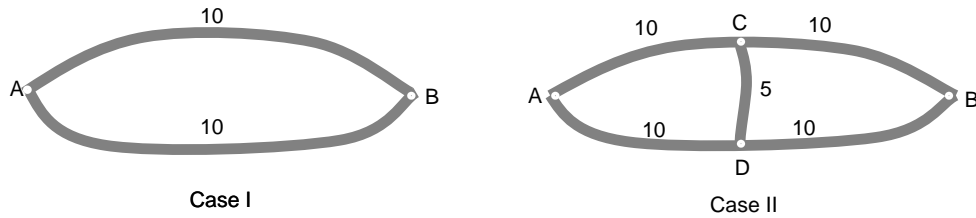
5.2.2.2 Numerical Example

Figure 5.2: Counter intuitive cases.

Here, we give a simple illustrative example of how NMI is calculated. In case II in Figure 5.2, there are four possible routes from node A to node B. Assuming 1.5 for *ratio,* all the four routes are included and we have **setKPaths** = {A-C-B; A-D-B; A-C-D-B; A-D-C-B}, where A-C-B is used to denote a path from A to to C to B. We proceed from route A^1 = A-C-B (shortest) to A^4 = A-D-C-B (longest) according to their length. Set **segmentsIncluded** is used to collect all road segments whose contributions to NMI have been counted and is initially empty.

1. We begin from the shortest route A^1 : the contribution (weighted number of path) of A^1 to the NMI is calculated as 1 since none of its segments is in the set **segmentsIncluded**. We then put all the segments (with accumulative number of occurrences) of this route into the set **segmentsIncluded**, i.e.

segmentsIncluded = {AC(1); CB(1)}

2. For the second shortest route A^2 , its contribution to the NMI is calculated as

$$\sum_{l_a \in L_2} \frac{|l_a|}{|L_2|} G_{l_a} = \frac{10}{20} G_{AD} + \frac{10}{20} G_{DB}$$
$$= 0.5 \cdot 1 + 0.5 \cdot 1 = 1$$

if G_{l_a} is selected as δ_{l_a} . We then add all L_2's segments into the set **segmentsIncluded**, i.e.

segmentsIncluded = {AC(1); CB(1); AD(1); DB(1)}

3. For the third shortest route A^3 = A-C-D-B, its contribution to the NMI is calculated as

$$\sum_{l_a \in L_3} \frac{|l_a|}{|L_3|} G_{l_a} = \frac{10}{25} G_{AC} + \frac{5}{25} G_{CD} + \frac{10}{25} G_{DB}$$
$$= 0.4 \cdot 0 + 0.2 \cdot 1 + 0.4 \cdot 0 = 0.2$$

if G_{l_a} is selected as δ_{l_a} . We then add all L_3's segments into the set **segmentsIncluded**, i.e.

segmentsIncluded = {AC(2); CB(1); AD(1); DB(2);CD(1) }

4. For the fourth shortest route A^4 = A-D-C-B, its contribution to the NMI is calculated as

$$\sum_{l_a \in L_4} \frac{|l_a|}{|L_4|} G_{l_a} = \frac{10}{25} G_{AD} + \frac{5}{25} G_{DC} + \frac{10}{25} G_{CB}$$
$$= 0.4 \cdot 0 + 0.2 \cdot 1 + 0.4 \cdot 0 = 0.2$$

if G_{l_a} is selected as δ_{l_a}. We then add all $L_4\text{'}s$ segments into the set

segmentsIncluded, i.e.

segmentsIncluded = {AC(2); CB(2); AD(2); DB(2);CD(1); DC(1) }

5. Compute the total NMI of the freeway facilities available for the source-destination (A, B) pair as

$$NMI = \sum_{k=1}^{K} \sum_{l_a \in L_k} \frac{|l_a|}{|L_k|} G_{l_a}$$
$$= 1 + 1 + 0.2 + 0.2 = 2.4$$

5.2.2.3 Comparison with Other Approaches

To the best knowledge of the authors, the concept that could achieve the most similar functionality (calculating the cumulative weighted number of paths between a source-destination pair) is the Path-Size concept [52]. Originally, the Path-Size concept is proposed for computing the probability of a driver selecting a specific route and it works very well for that purpose. Bekhor et al. [52] used the Path-Size to account for the correlation (overlap) between routes and defined it as the weighted number of paths for one path between a source destination pair. Mathematically, for path *k (k=1, 2, ..., n)*, PS_{kn} can be written as:

$$PS_{kn} = \sum_{a \in \Gamma_k} (\frac{|l_a|}{|L_k|}) \frac{1}{\sum_{j \in C_n} \left(\frac{|L_k|}{|L_j|} \right)^{\gamma} \delta_{aj}} \qquad (5.3)$$

where $|l_a|$ is the length of link a, $|L_k|$ is the length of path k; δ_{aj} is 1 if link a is within path j and 0 otherwise, Γ_k is the set of all links of path k, and γ is a positive parameter.

We could use this definition for our calculation of the total number of weighted routes between a pair of source and destination by calculating all the feasible PS_{kn} (k=1, 2, …, n) of each route between this source-destination pair and adding them together to formulate the total weighted number of paths between the given source-destination pair. However, if calculated in this way, the resulting number of paths tends to be underestimated. A comparison of the total cumulative weighted number of paths between a source-destination pair using the proposed NMI and the Path-Size for Figure 5.2 is shown in Table 5.3. From Figure 5.2 and Table 5.3, we can see that adding additional infrastructure (from C to D) only increases the weighted number of paths from 2 to 2.04 if using the Path-Size concept. Therefore, we could not use the Path-Size concept in this scenario.

Table 5.3: Comparisons of weighted number of paths using NMI and Path-Size of the two special cases.

Case	Weighted # using NMI	Weighted # using Path-Size ($\gamma=0$)	Weighted # using Path-Size ($\gamma=1$)	Weighted # using Path-Size ($\gamma=2$)
I	2	2	2	2
II	2.4	2	2.02	2.04

5.2.3 Properties of the Navigational Mobility Index

The proposed definition of NMI has several attractive properties. We already observe that the K shortest paths have some overlapping segments among them and these overlapping segments have to be taken care of. Thus, a *dependency matrix* is used to represent the correlation (overlapping segments) among all the K shortest routes (or m smaller than K if the algorithm stops earlier), and to prove important properties of the proposed definition of NMI. The dependency matrix (having size m by m) is constructed as shown in (5.4), and is divided into blocks with each block representing the correlation of one route pair (e.g. $D_{i,j}$ represents the correlation between route i and route j). Blocks are arranged from short to long in terms of their lengths, e.g., the block $D_{1,1}$ represents the correlation between the shortest path and itself.

$$
\mathbf{D} =
\begin{bmatrix}
D_{1,1} & D_{1,2} & \cdots & D_{1,m} \\
D_{2,1} & D_{2,2} & \cdots & D_{2,m} \\
\vdots & \vdots & \ddots & \vdots \\
D_{m,1} & D_{m,2} & \cdots & D_{m,m}
\end{bmatrix}
\tag{5.4}
$$

where

$$
D_{i,j} =
\begin{bmatrix}
r_{i,1}^{j,1} & r_{i,1}^{j,2} & \cdots & r_{i,1}^{j,Q_{r_j}} \\
r_{i,2}^{j,1} & r_{i,2}^{j,2} & \cdots & r_{i,2}^{j,Q_{r_j}} \\
\vdots & \vdots & \ddots & \vdots \\
r_{i,Q_{r_i}}^{j,1} & r_{iQ_{r_i}}^{j,2} & \cdots & r_{i,Q_{r_i}}^{j,Q_{r_j}}
\end{bmatrix}
$$

with Q_{r_j} being the number of segments of route j. For the block $D_{i,j}$ representing the correlation between route i and route j, the row within this block represents the segments of route i, and the column within this block represents the segments of route j. If the

78

p^{th} segment of route i overlap with the qth segment of route j, the value of element (p, q),

i.e. $r_{i,p}^{j,q}$, in this block is the length of the overlapping segment. In other words, the

element $r_{i,p}^{j,q}$ represents the overlapping length of the pth segment of the ith path with the

q^{th} segment of the jth path. Note that for two segments, either they completely overlap

or the overlap is 0 due to the way the matrix is constructed.

Again, we give a simple illustrative example of how to construct a dependency matrix

using case II in Figure 5.2. There are 4 routes all together, and the dependency matrix is

constructed as

$$\mathbf{D} = \begin{bmatrix} D_{1,1} & D_{1,2} & D_{1,3} & D_{1,4} \\ D_{2,1} & D_{2,2} & D_{2,3} & D_{2,4} \\ D_{3,1} & D_{3,2} & D_{3,3} & D_{3,4} \\ D_{4,1} & D_{4,2} & D_{4,3} & D_{4,4} \end{bmatrix}$$

Next, we use block $D_{i,j}$ to represent the correlation between route i and route j. For

example, for route A^2 = A-D-B and route A^3 = A-C-D-B,

$$D_{2,3} = \begin{array}{c} \quad \; AC \; \; CD \; \; DB \\ \begin{array}{c} AD \\ DB \end{array} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 10 \end{bmatrix} \end{array}$$

Therefore, the dependency matrix is:

$$
\mathbf{D} =
\begin{array}{c}
 \\
AC \\ CB \\ AD \\ DB \\ AC \\ CD \\ DB \\ AD \\ DC \\ CB
\end{array}
\begin{array}{c}
\begin{array}{cccccccccc}
AC & CB & AD & DB & AC & CD & DB & AD & DC & CB
\end{array} \\
\left[
\begin{array}{cccccccccc}
10 & 0 & 0 & 0 & 10 & 0 & 0 & 0 & 0 & 0 \\
0 & 10 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 10 \\
0 & 0 & 10 & 0 & 0 & 0 & 0 & 10 & 0 & 0 \\
0 & 0 & 0 & 10 & 0 & 0 & 10 & 0 & 0 & 0 \\
10 & 0 & 0 & 0 & 10 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 5 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 10 & 0 & 0 & 10 & 0 & 0 & 0 \\
0 & 0 & 10 & 0 & 0 & 0 & 0 & 10 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 5 & 0 \\
0 & 10 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 10
\end{array}
\right]
\end{array}
$$

The proposed definition of NMI has some attractive properties as follows:

1. The dependency matrix is symmetric with $D_{i,j} = D_{j,i}$;

2. All diagonal elements of the dependency matrix are nonzero;

3. For each row of each block, at most one element is nonzero. For each column of each block, at most one element is nonzero;

4. In any row of the dependency matrix that has other nonzero elements besides the diagonal elements, the value of all those nonzero elements is equal to the value of the diagonal element; and

5. Most importantly, adding additional infrastructures will NOT decrease the NMI value.

Properties 1 and 2 are relatively straightforward. Property 3 is due to the fact that the route is loopless, i.e. each route can only pass one segment and one node at most once. Property 4 is due to the fact that each row of the dependency matrix is representing if one

segment overlaps with the segments of other routes. If it overlaps, the corresponding value will be the length of that segment, otherwise 0. Property 5 is important and needs a proof.
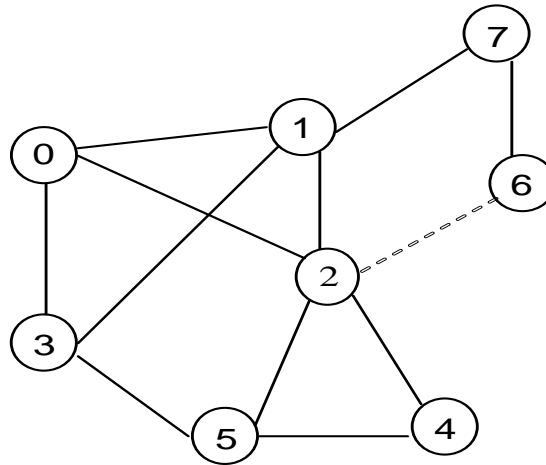
Figure 5.3: A sample roadway network with added facility.

Proof:

We have an existing freeway network structure represented by solid lines in Figure 5.3. Assume there is an additional freeway facility (i.e. the freeway segment between node 2 and node 6) added to the network. If using function δ_{l_a} in the calculation of NMI, then the NMI for any source-destination pairs is not supposed to decrease. For one specific source-destination pair, assume a large value of K is selected, and the K-shortest path algorithm stops early when the mth route ($Length_{m+1} > ratio * Length_{shortest}$) is found. The dependency matrix for that specific source-destination pair after the additional freeway segment is added, assuming that only one additional route is added, can be written as:

$$D = \begin{bmatrix} D_{1,1} & D_{1,2} & \cdots & D_{1,j} & \cdots & D_{1,m+1} \\ D_{2,1} & D_{2,2} & \cdots & D_{2,j} & \cdots & D_{2,m+1} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ D_{j,1} & D_{j,2} & \cdots & D_{j,j} & \cdots & D_{j,m+1} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ D_{m+1,1} & D_{m+1,2} & \cdots & D_{m+1.j} & \vdots & D_{m+1,m+1} \end{bmatrix} \tag{5.5}$$

Without loss of generality, assume that only one additional route is added. Thus, in (5.5) we have one additional column and one additional row of blocks added. Suppose the additional route is the jth shortest route. Therefore, blocks $D_{j,k}, k = 1, \cdots, m+1$ and $D_{k,j}, k = 1, \cdots, m+1$ are added. All the values of other existing blocks remain unchanged. To calculate the NMI, we proceed from the top left of the dependency matrix to the bottom right. We first compute the weighted contribution of the first route by considering block $D_{1,1}$. The weighted contribution for the first path is 1 according to the definition. We then compute the weighted contribution of the second shortest path by looking at the blocks $D_{1:2,1:2}$, where $D_{1:u,1:v}$ is the submatrix consisting of blocks $D_{p,q}$, where $1 \le p \le u$ and $1 \le q \le v$. We check all diagonal elements of $D_{2,2}$ by starting from the element $r_{2,1}^{2,1}$. We examine all the elements that are at the same column as $r_{2,1}^{2,1}$ in the dependency matrix and whose row index in the dependency matrix is smaller than the row index of $r_{2,1}^{2,1}$. If any such element is nonzero, the contribution to the second path of this element is 0 since the contribution of this segment to the NMI has already been counted when calculating the previous route. We do the same thing with all other diagonal elements of block $D_{2,2}$. We then repeat the same procedure for all other diagonal blocks.

When we process block $D_{j,j}$ (which is due to the added infrastructure), we find that all the elements of path j are shared with other routes except the newly added segment (it is possible that the new route j will utilize other new roadway facilities that have not been included before, but in that scenario, the NMI will be increased even more). Segments could be shared by route j with other routes either longer than, equal to, or shorter than route j.

For segments shared by route j with routes shorter than it, the contribution to NMI of these segments under route j is zero since it has already been counted when calculating the previous routes. For segments shared by route j with routes longer than it, the contribution to NMI of these segments should be calculated now under route j instead of under the later routes. Since the length of route j is smaller than the length of the later routes, the contribution to NMI of these segments will be larger according to the definition. The contribution of the newly added facility is also computed under route j, which increases the value of NMI even more. Similarly, we can follow the same process to prove property (5) for any G_{l_a} other than δ_{l_a}. This concludes the proof.

5.2.4 Node-NMI and Area-NMI

Thus far, we have discussed the definition of a node-to-node NMI, which provides the foundation for the definition of a *node-NMI* and an *area-NMI*.

In general, a node-to-node index can be extended to measure the performance of one location and the performance of one area. Assume there is an area consisting of *N*

locations and one specific kind of index from location i to location j is defined as a_{ij}. An *integral index* was first introduced in [48] for each location i in an area consisting of N locations and an extension was proposed in [45] as

$$A_i = \frac{1}{N-1} \sum_{\substack{j=1 \\ j \neq i}}^{N} a_{ij}, i = 1,2,\cdots,N \tag{5.6}$$

This integral accessibility index can be used to compare the performance among different locations within an area. However, it does not capture the overall performance over an entire area and is not comparable among locations belonging to different areas. Bruce et al. [45] further extended the integral to an *overall index* as:

$$A = \frac{1}{N(N-1)} \sum_{i=1}^{N} \sum_{\substack{j=1 \\ j \neq i}}^{N} a_{ij} \tag{5.7}$$

The resulting overall index quantifies the performance of an entire area. Since it is normalized with respect to the number of location pairs in an area, it can be compared across different areas.

In this chapter, the node-NMI is defined by adopting the definition of the integral accessibility index in (5.6), and the node-NMI for node i is defined as:

$$A_i^{NMI} = \frac{1}{(N-1)} \sum_{\substack{j=1 \\ j \neq i}}^{N} (NMI_{ij}), i = 1,2,\ldots,N \tag{5.8}$$

One important application of the node-NMI is its potential usage in emergency evacuation. Natural disasters (e.g. flood, earthquake, hurricane, and forest fire) and terrorism attacks always pose a potential threat to the security of any cities. During these events, people need to relocate from dangerous locations to safer locations as soon as

possible. During this movement, the routes to be taken should be chosen in such a way that also provides the best possible re-routing options due to uncertain conditions of the roadway network. That is, we should evacuate along the roads (intersections) that have high node-NMI value and avoid those nodes that have low node-NMI value.

For an area-wide measure of navigational mobility, we adopt the definition of the overall accessibility index presented in (5.7), which is averaged over all node pairs, i.e.

$$A^{NMI} = \frac{1}{N(N-1)} \sum_{i=1}^{N} \sum_{\substack{j=1 \\ j \neq i}}^{N} (NMI_{ij}) \tag{5.9}$$

where N is the total number of critical nodes in an area.

The value of A^{NMI} gives an estimate of the average number of effective routes between two randomly selected nodes in an area. It is a good measure of the degree of travel freedom within an area. Therefore, the area-NMI can be used to evaluate the vulnerability of roadway network layout in an area. Transportation agencies can use the area-NMI to systematically identify (and even rank) roadway links that are critical to the mobility of road users. For example, if a roadway link was to be close for maintenance or due to structural failure (e.g., I-35W bridge collapse in Minnesota in December 2007), it is important to know how much the area-NMI would decrease (i.e., how many alternative routes remain available to road users without that particular roadway link). On the other hand, constructing the same length of roadway segments can result in different area NMI values. Thus, the area-NMI can serve as another planning tool in determining how to design the layout of roadway network efficiently. Since A^{NMI} is normalized with respect to the number of nodes in an area, it can be compared across different areas.

5.2.5 NMI-based Navigation Algorithm

As previously mentioned, there has been a significant proliferation of both on-board and off-board navigation systems that primarily provide routing advice across a roadway network (e.g. Google map [57] and MapQuest [58]). Some of these systems are now capable of incorporating real-time congestion delays, providing the ability to find the shortest-duration path and minimum fuel consumption path [60] in addition to the more standard shortest-distance path. However, in a network that is prone to incidents or high congestion, all the above navigation algorithms may not work well. With the ability to re-route based on changing network conditions, NMI is useful in giving an indication of the number of routing options. A network with a low NMI value will not have many re-routing choices; and therefore, the ability to choose an alternative route based on real-time traffic conditions will have limited value.

Normally, drivers want to select either a shortest-duration or shortest-distance path, and would like the traffic conditions to be free-flow and uninterrupted. However, various scenarios sometimes evolve where links are unexpectedly shutdown/highly congested due to major incidents (i.e., accidents) or natural disasters (e.g., earthquakes). In these scenarios the driver must mandatorily re-route. In this section, we briefly present a new navigational methodology based on NMI that could be used for mandatory re-routing scenarios. Please refer to Chapter 5 and [61] for more comprehensive presentation and discussion.

In circumstances when a "good and feasible" route needs to be found for a pair of source and destination nodes, node-to-node NMI values can be applied as part of the navigational methodology. We first calculate all the node-to-node NMI values of the pairs from each node in the network to the destination node. We then attach the corresponding node-to-node NMI value to each node, respectively. In order to incorporate the NMI value into the generalized cost, we have to convert the NMI value to a connectivity cost using the following function:

$$NMICost = \left(\frac{\min * \max}{NMI} \right)^{\omega} \tag{5.10}$$

where *min* is the minimum nonzero node-to-node NMI value in the network, *max* is the maximum node-to-node NMI value in the network, ω is a positive parameter, and *NMI* is the actual node-to-node NMI value.

Currently, we have an *NMICost* value attached to each critical node. For a link going from Node A to Node B, the *NMICost* value of this link is defined as the *NMICost* of the end Node, i.e. Node B. The reason for this is that when a driver is on the link, (s)he has no other routing choices until (s)he reaches the end node where (s)he can change routes. Finally, we define the generalized cost incorporating both distance and NMI value as

$$Cost = f(length) * g(NMICost) \tag{5.11}$$

where *f(length)* is a function depending on how much weight needs to be put on the distance. In this chapter, it is chosen as

$$f(length) = (length)^{k}$$

where k is a parameter for the weight. The term $g(NMICost)$ is a function depending on how much weight needs to be put on the $NMICost$. For example, it can be chosen as

$$g(NMICost) = (NMICost)^{\omega}$$

where ω is a parameter for the weight. In Chapter 5 and [61] different forms of function f and g are also investigated.

Since we have to calculate all the node-to-node NMI values of the pairs from each node in the network to the destination node, the computation complexity will be $O(Kn^2(e + n \log n))$ if the Dijkstra algorithm is used to find the shortest path, where e is the number of edges and n is the number of nodes in the network. Thus, the calculation of the generalized cost depends on the size and complexity of the networks (e.g. for the experiment in section 5.3.4, it takes a couple minutes). This calculation and the subsequent routing determination may be performed off-board on a fast computer server before the result is transmitted to an on-board device for display. Alternatively, a methodology may be developed to significantly reduce the computation time, thus allowing for on-board calculation. The methodology may involve establishing a meaningful search area (i.e. using a small searching area instead of the entire map) in order to eliminate unnecessary calculations. This is to be addressed in future work.

5.3 Experimental Results and Discussion

5.3.1 Experimental Setup

The freeway networks of the Los Angeles (LA) County in Southern California and the San Francisco (SF) Bay Area in Northern California were selected as case studies due to having the moderate number of nodes and multiple parallel corridors. Census 2000 TIGER/Line shapefiles [53] were used for the experiment. Using ArcGIS 9.1, we extracted all the roadway segments whose Census Features Class Codes (CFCC) belongs to categories A1, A2, and A3 corresponding to the freeway categories. Abstract graphs were then built using the corresponding database file to describe the roadway connectivity. Any nodes connected to two other nodes are considered non-critical nodes and the adjacent lines represented by non-critical nodes are combined. Any nodes connected to more than two nodes are intersections and considered critical nodes. Any nodes connected to only one other node are end nodes and are also considered critical nodes. In the end, the LA County has 542 critical nodes and the SF Bay Area has 400 critical nodes.

5.3.2 Comparison of Area-NMI between the LA County and the SF Bay Area

In this section, we compare the area-NMI of the LA County network and the SF Bay Area network. We suspect that the area-NMI results may be sensitive to the coverage of

the area. Therefore, instead of using Equation (5.9), we use the following equation to

calculate the area-NMI of the networks:

$$A^r = \frac{1}{N} \sum_{i=1}^{N} \frac{1}{\left| S_i^r \right|} \sum_{j \in s_i^r} NMI_{ij} \tag{5.12}$$

where, S_i^r is the set of all nodes that fall in the circle whose radius is r and is centered at

node i, and $\left| S_i^r \right|$ is the number of nodes in the set S_i^r. In essence, this area-NMI provides a

general indication of the number of possible routes between any node pairs within an area.

Figure 5.4 shows the comparison of the area-NMI between the LA County and the SF

Bay Area. It implies that the LA County network is better than the SF Bay Area network

in terms of the connectivity of freeways at all radius values. This is primarily due to the

fact that the Bay Area network is geographically separated by the bay, which reduces the
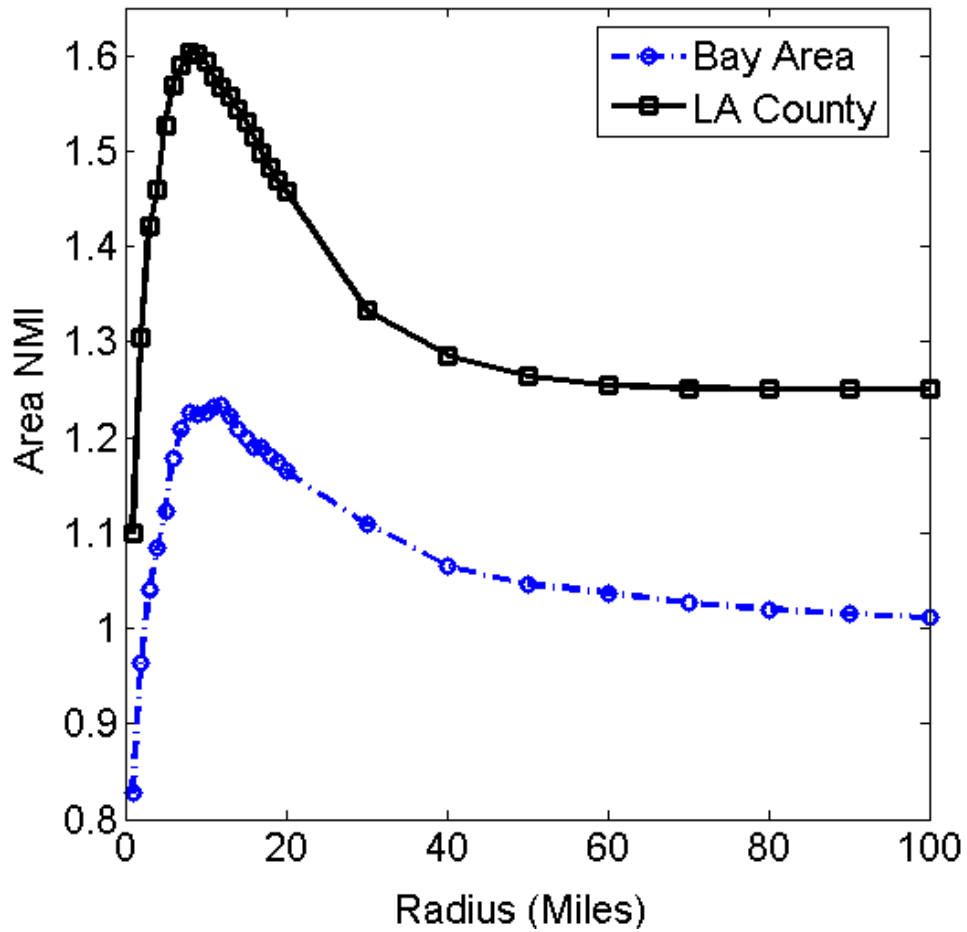
overall connectivity of the network.

Figure 5.4: Area NMI of the LA County and the San Francisco Bay area with different radius values.

5.3.3 Sensitivity Analysis of the Area Coverage

Similar to Section 5.3.2, instead of using Equation (5.8), we use the following equation to calculate the NMI of node i in the network:

$$A_i^r = \frac{1}{\left| S_i^r \right|} \sum_{j \in S_i^r} NMI_{ij}$$

In this experiment, we select K (the number of shortest paths) as 10. Whenever a path m where $Length_m > 1.5 * Length_{shortest}$ is found, the iteration stops at m. The value of r is selected as 10 miles, 50 miles, and 1000 miles (i.e. the coverage is the whole area). Figures 5.5(a) through 5.7(a) show the comparisons of freeway connectivity between the LA County network and Bay Area network. We use Google Earth [59] for the display. Each link is assigned a "link NMI" whose value is the average between the NMI values of the start node and end node of this link. The link is colored in a way that green represents the highest NMI (2.6) and red represents the lowest value (0).

We can see that the maximum value of link NMI is comparable for the two areas. However, the LA County network has many more links that have high NMI values. Further, the NMI value is sensitive to the range of coverage. When the coverage is small, the number of nodes included is also small, resulting in small values of the average NMI. When the coverage is very large, the number of nodes included is also very large. However, these nodes include many isolated nodes in distant areas, inversely causing the average NMI values to be small. From the comparisons in Figures 5.5(b) through 5.7(b), it can be seen that in general the LA County network provides more freedom of route selection. The number of nodes in the LA County having high NMI values is larger than that in the Bay Area. It also can be observed that the distributions of NMI values loosely follow a Gaussian distribution. The outliers with very low NMI values are mostly those isolated nodes and road segments.

Figure 5.5(a): NMI of each link in LA County (upper) and in Bay area (bottom), radius 10 miles.
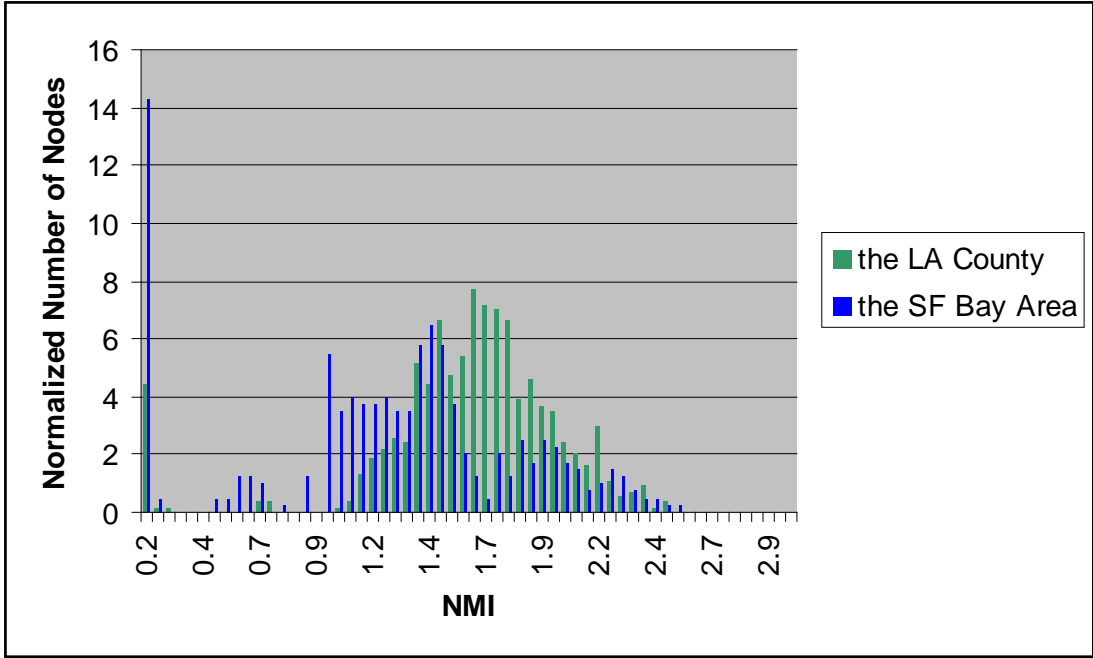
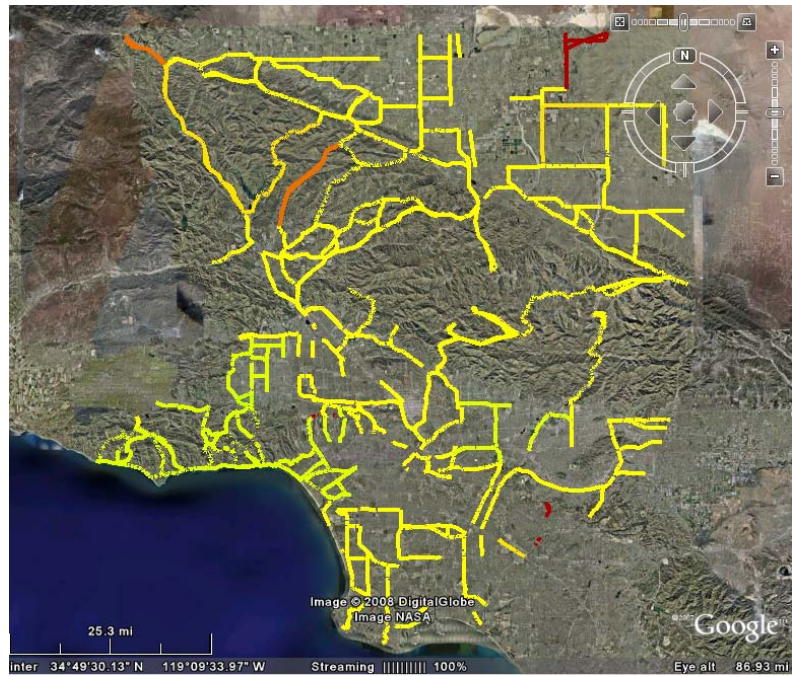Figure 5.5(b): The distribution of NMI values, radius 10 miles.

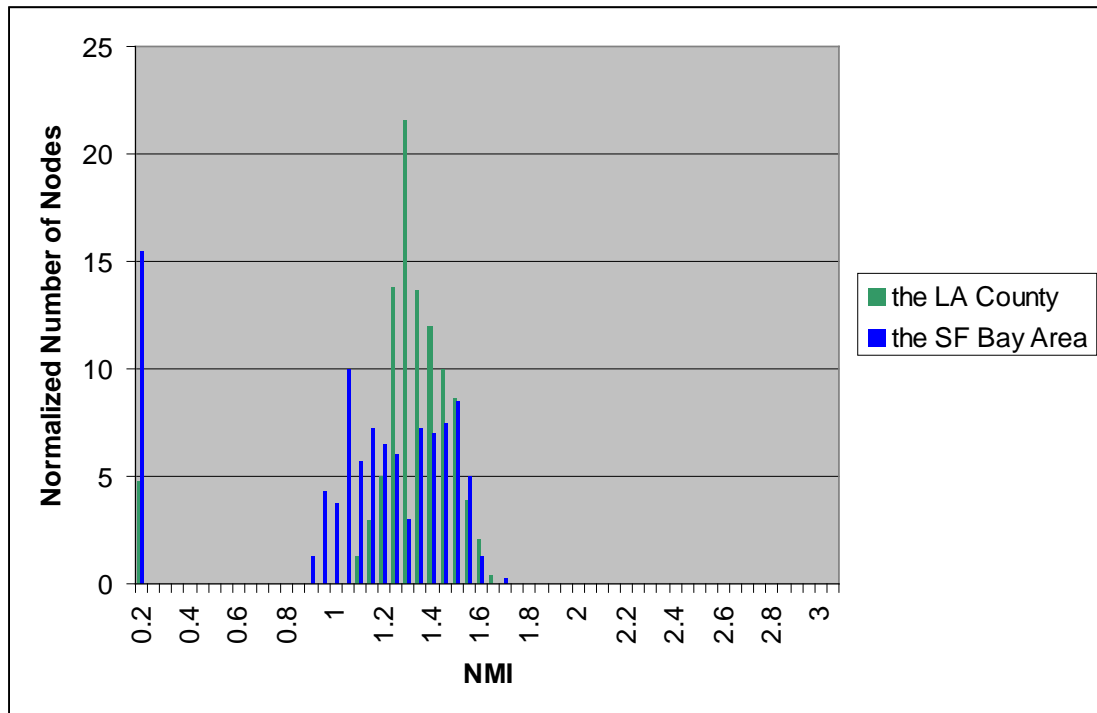Figure 5.6(a): NMI of each link in LA County (upper) and in Bay area (bottom), radius 50 miles.

Figure 5.6 (b): The distribution of NMI values, radius 50 miles.
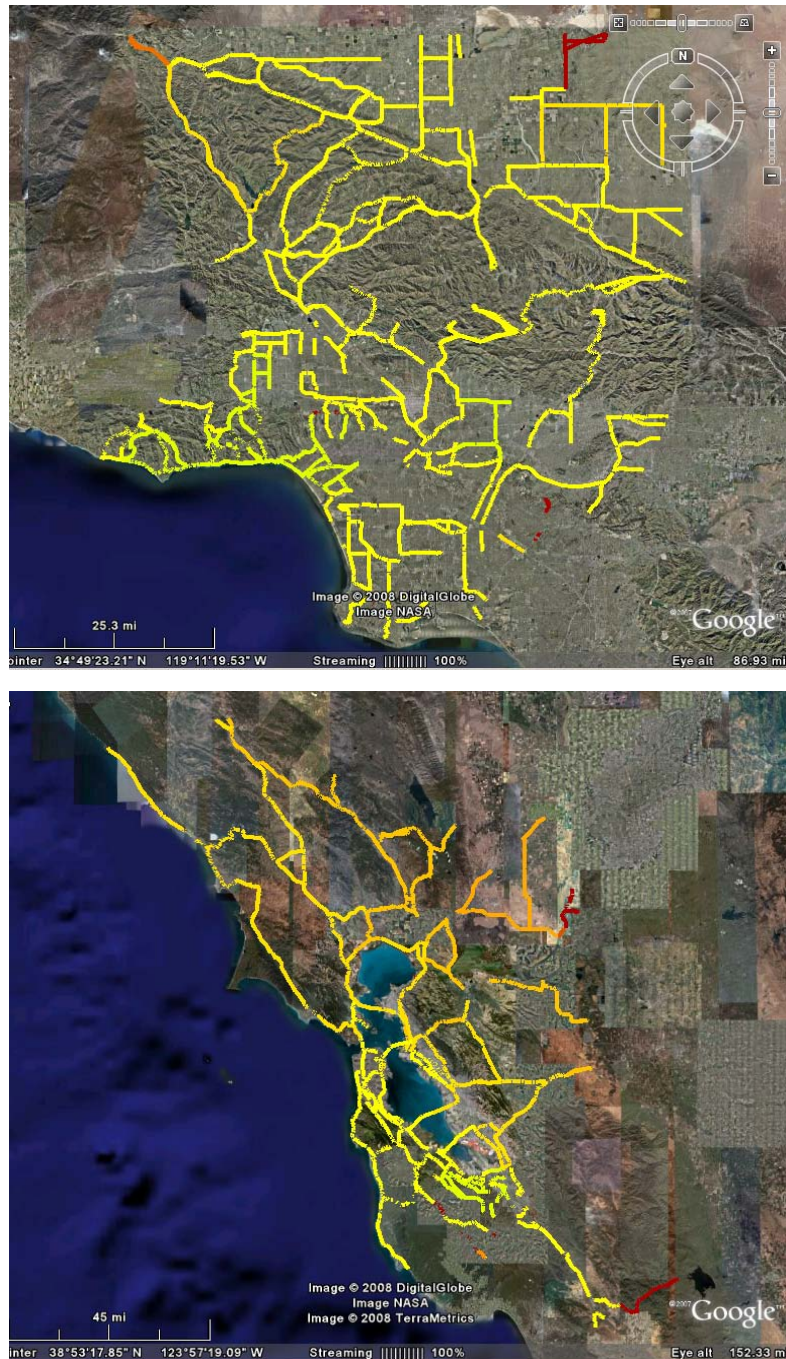
Figure 5.7(a): NMI of each link in LA County (upper) and in Bay area (bottom), radius 1000 miles.
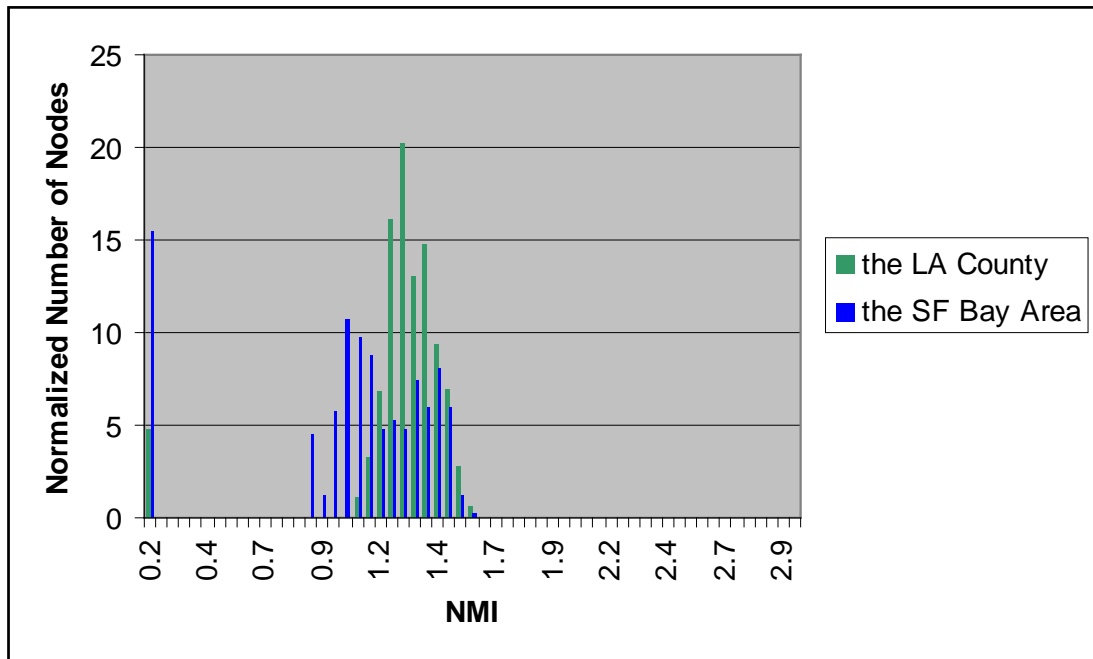
Figure 5.7(b): The distribution of NMI values, radius 1000 miles.

5.3.4 NMI-Based Navigation

In this section, we compare the shortest-distance route and the least generalized cost route. For this analysis, we use the network of the South Bay Area in Northern California as a case study. The generalized cost is constructed as in (5.11). The selection of ω depends on how much weight is to be placed on the NMI. If ω is 0, then it is simply a shortest-distance routing algorithm. The higher ω is, the more weight we put on the NMI. In this experiment, we select a moderate value 3 for ω.

Figure 5.8 shows the results of the recommended routes according to different criteria. The purple-colored route is based on the shortest-distance, and the blue-colored route is according to the least generalized cost. We can see from the figure that these two routes

98

have a significant number (about 50%) of non-overlapping segments. The least generalized cost route does not include those nodes along the SF Bay because those nodes generally have low values of NMI to the destination.
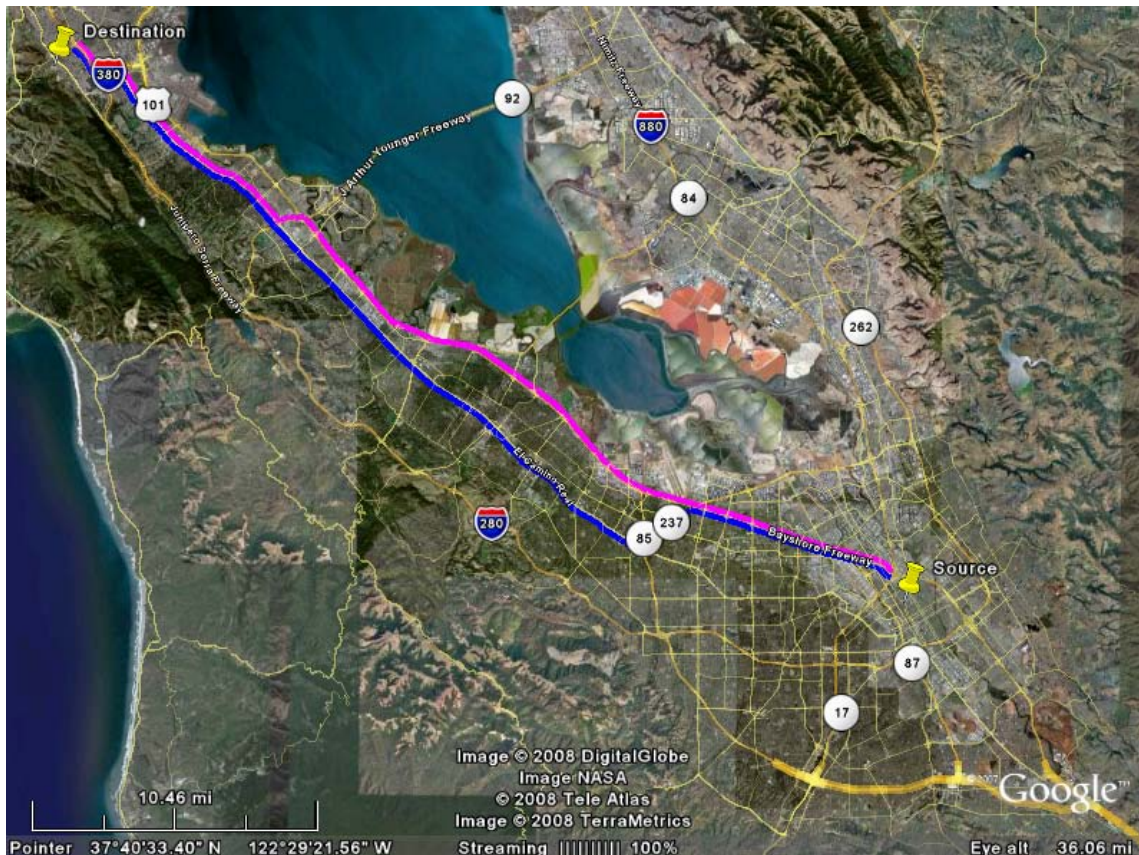
Figure 5.8: The routes selected by the shortest-distance (purple), and the least generalized cost (blue).

A generalized cost incorporating both distance and NMI can be particularly useful when traffic conditions change frequently and unexpectedly, for example, during rush hour, when multiple accidents occur. It can also be useful for disaster evacuation purposes (see Chapter 5 and [61]). Ultimately, the NMI can serve as another meaningful

99

metric when performing route optimization in addition to the standard metrics of distance, time, and fuel consumption.

6. Mobility Index-Based Navigation for Mandatory Re-Routing Scenarios

6.1 Introduction

Current dynamic navigators have the capability to re-route if the driver wishes to avoid a particular area of roadway network due to traffic conditions. In Chapter 5, we have developed a *navigational mobility index* (specifically for navigational purpose) that quantifies the number of possible routing options, e.g. the weighted number of possible routes between an origin-destination pair. Basically, the NMI was used to quantify connectivity of facility from one location to another. Furthermore, corresponding node-NMI and area-NMI were also defined.

In this chapter, we develop a new navigational methodology based on NMI for dealing with mandatory re-routing scenarios. Normally, drivers want to select either a shortest-duration or shortest-distance path, and would like the traffic conditions to be free-flow and uninterrupted. A mandatory re-routing situation occurs when one or more links in a roadway network are closed (or highly congested) and it is preferable (sometimes it is even mandatory) to choose a different path. There are basically two scenarios where mandatory re-routing might be need: one is disaster evacuation, and the other is navigation under chaotic conditions.

Disasters (earthquakes, hurricanes, tornados, forest fires, floods, terrorist attacks, etc.)

can happen anywhere anytime and can have severe consequences. For example, in the recent Wenchuan earthquake (May 12, 2008) in China, over 69,000 people were killed and a significant amount of roadway infrastructure was destroyed. During these disasters, roadway facilities can be damaged but their conditions may not be known immediately to the navigation systems. Nevertheless, it is extremely important to relocate drivers in affected areas to safer locations as soon as possible. In these scenarios, drivers might have to re-route frequently due to the failure of roadway segments. It is very important that drivers are able to re-route. In another word, the degree of freedom of re-routing should be as high as possible to avoid reaching a dead end. Both shortest-distance and shortest-duration algorithms may recommend routes that are not high in degree of freedom for re-routing, causing the drivers less likely to reach their destinations.

Nowadays, traffic conditions are often unpredictable. Roadway incidents happen frequently and cause traffic interruptions. Dynamic shortest-duration algorithms typically use roadway real time speed data, as well as historical data (long term and short term) [66, 71] when calculating a shortest-duration route. This works well when there are no roadway incidents happening that causes severe congestion. However, when an incident happens and the traffic becomes congested, the navigator might advice the drivers to re-route. This is another scenario of mandatory re-routing. Although this may not really be "mandatory", and re-routing is just an option, we still *call* it mandatory re-routing because staying in the old route is no longer desirable. In this scenario, new shortest-duration route will not exist if there are no route choices available (for example, if the driver is on a bridge with no other connections). This emphasizes that the degree of

freedom in re-routing may be a desirable property for determining an optimal route provided by navigation system.

In these cases, it is desirable that a navigational algorithm find an optimal route that offers the highest degree of freedom in re-routing. When considering this kind of optimal route, even if some roadway infrastructure is unusable, the drivers are still more likely to find another route to reach their destination in a reasonable travel time without being trapped in a dead end or backtracking too much, compared to other navigation algorithms.

The remainder of this chapter is organized as follows. Section 6.2 summarizes the overall methodology including the definition of NMI and the routing algorithm. Section 6.3 presents the experimental results of two mandatory re-routing scenarios and provides discussions.

6.2 Methodology

6.2.1 Review of Navigation Mobility Index

NMI was proposed in Chapter 5 and [65] to quantify the connectivity of roadway facilities from one location to another indicating *the weighted number of available routes*. Corresponding node-NMI and area-NMI were proposed to indicate the connectivity of a node and connectivity of an area, respectively. The definition of node-to-node NMI is briefly summarized below.

Yen's algorithm [54] is first used to find a total of K shortest paths (i.e. the shortest, the

second shortest, and so on), and they are stored in set **setKPaths** in an increasing order (according to the route length, i.e. A^1 for the shortest one, A^2 for the second shortest one, and so on). Each path consists of multiple segments (components of a route). We proceed in the order from A^1 to A^K. Next we have another set called **segmentsIncluded** to collect all road segments whose contribution to NMI has been counted. The set **segmentsIncluded** is initially set to empty. The detailed steps are given below:

1. We begin from the shortest route A^1: the contribution (weighted number of path) of A^1 to the NMI is calculated as *1* since none of the segments is in set **segmentsIncluded**. We then put all the segments of this route to the set **segmentsIncluded**.

2. For the second shortest route A^2, its contribution to the NMI is calculated as:

$$\sum_{l_a \in L_2} \frac{|l_a|}{|L_2|} G_{l_a} \qquad (6.1)$$

where L_2 is the collection of all segments of route A^2, $|L_2|$ is the total length of route A^2, and $|l_a|$ is the length of segment l_a. G_{l_a} is a general function depending on l_a and set **segmentsIncluded**. In this chapter we specifically choose δ_{l_a} for G_{l_a}, where δ_{l_a} is a function whose value is *1* if segment l_a is not within the set **segmentsIncluded**, and is *0* if otherwise. We then add all $L_2's$ segments into the set **segmentsIncluded**.

3. Repeat step 2 for all the remaining shortest paths.

4. Compute the total NMI of the roadway facilities available for this origin-destination pair as

104

$$NMI = \sum_{k=1}^{K} \sum_{l_a \in L_k} \frac{|l_a|}{|L_k|} G_{l_a} \qquad (6.2)$$

In this chapter, we specifically choose δ_{l_a} for G_{l_a} because it is simple and intuitive. Using δ_{l_a}, we do not consider the contribution to NMI of a new route's segments that has already been counted in previous routes. In order to grant some small weighted contribution of these segments, G_{l_a} can be selected as a general function inversely proportional to the number of occurrences of l_a being within set **segmentsIncluded**.

When drivers are selecting routes, unreasonably long routes are usually not considered. Therefore, in the calculation of K shortest paths, another criterion is implemented: whenever an mth ($m < K$) shortest route whose length is longer than $ratio * Length_{shortest}$ is found (where $Length_{shortest}$ is the length of the shortest path), the K-shortest path algorithm stops. For a moderately complex freeway network in an urban area, a large value of K is usually selected and the ratio parameter is normally set at around 1.5.

This definition of NMI described above has some attractive properties. One of them is that adding additional infrastructure will not decrease the value of the index. Please refer to Chapter 5 and [65] for other properties and their proofs. Due to limited space, please also refer to Chapter 5 and [65] for simple illustrative examples of how NMI is computed.

The corresponding node-NMI for node i is defined as:

$$A_i = \frac{1}{(N-1)} \sum_{\substack{j=1 \\ j \neq i}}^{N} (NMI_{ij}), i = 1, 2, \ldots, N \tag{6.3}$$

and the area-NMI are NMI averaged over all node pairs:

$$A = \frac{1}{N(N-1)} \sum_{i=1}^{N} \sum_{\substack{j=1 \\ j \neq i}}^{N} (NMI_{ij}) \tag{6.4}$$

where N is the total number of nodes in the area.

6.2.2 NMI-based Navigation Algorithm

As previously mentioned, there has been a significant proliferation of on-board and off-board navigation systems that primarily provide routing advice across a roadway network. With the ability to re-route based on changing network conditions, NMI is useful in giving an indication of the number of routing options. A network with low NMI value will not have many re-routing choices and therefore choosing an alternative based on real-time traffic conditions will have limited value.

In contrast, it may be desirable to choose a preferred route also based on the number of available routing options. This is particularly valuable if the roadway network changes frequently due to the freeway incidents, congestion, etc. Instead of using duration or distance only, a generalized cost considering multiple measurements (distance, routing options, emission/energy, etc.) can be used. In essence, a "good" route can be chosen that provides a smallest generalized cost that incorporate both distance and freedom of route selection, if the network characteristics (i.e. traffic conditions) change.

Such kinds of routing methods are even more important for disaster evacuation. During an evacuation, people need to be relocated to safer areas as soon as possible.

Some roadway infrastructure is likely to be damaged and vehicles might not be able to get through. For disaster evacuation, there are basically two requirements: the first one is that the probability that vehicle is trapped in dead ends should be minimized; the second one is that the distance traveled should also not be long.

Node-to-node NMI can be used for mandatory re-routing. A route providing the highest degree of freedom in re-routing needs to be found for an origin-destination pair. We first calculate the node-to-node NMI value of all the pairs from each node in the network to destination nodes. We then attach the corresponding node-to-node NMI value to each node, respectively. In order to incorporate the NMI value into the generalized cost, we have to convert the NMI value to a connectivity cost first using a mapping function. The mapping function has to be inversely proportional to the NMI value. An example is given below:

$$NMICost = \left(\frac{\min * \max}{NMI} \right)^{\omega} \tag{6.5}$$

where *min* is the minimum nonzero node-to-node NMI in the network, *max* is the maximum node-to-node NMI value. So far, we have an NMI value for each node. We define the NMI value of a link as follows. For a link going from Node A to Node B, the NMI value of this link is defined as the NMI of the end node, i.e. Node B. The reason for this is that when a driver is on the link, (s)he has no other routing choice until (s)he reaches the end node before (s)he can change routes. Finally we define the generalized cost incorporating both distance and NMI value as

$$Cost = f(length) * g(NMICost) \tag{6.6}$$

where *f(length)* is a function depending on how much weight need to be put on the distance, and in this chapter it is chosen as

$$f(length) = (length)^k \qquad\qquad (6.7)$$

where *k* is a parameter for the weight. The term *g(NMICost)* is a function depending on how much weight needs to be put on the *NMICost*, and in this chapter it is chosen as

$$g(NMICost) = e^{NMICost} \qquad\qquad (6.8)$$

After we have the generalized cost for each link, we are able to do the NMI-based navigation for mandatory routing. Disaster evacuation is selected for the illustration of the navigational method, but routing under severe congestion should be very similar.

In the disaster evacuation scenario, we only have the original "unaffected" map, which is the map with all roadway infrastructures in good condition before a disaster. We also assume that when a driver is at a node, he is able to observe if a segment he is moving onto next is in serviceable condition or not by watching the general traffic flow in that segment or traffic signs/bulletins in that segment .

The steps of the mandatory re-routing based on least generalized cost are:

1. the generalized cost for each link;

2. Find the least generalized cost path between the origin-destination pair using any shortest-path finding algorithms (e.g. Dijkstra [18], A*[19] [62], etc.);

3. Proceed from the origin node;

4. Decide if the next link to move onto is in serviceable condition;

5. If the link to move onto for step 4 is not serviceable, go to step 6; otherwise move

one link forward, and if the next node is destination, stop at destination, but if the next node is not destination, go to step 4;

6. Stay at the node; disable the unserviceable segment in the map, and find another least generalized cost path (going backwards is allowed). If such a route does not exist, this vehicle is trapped in a dead end and the routing fails. If such route exists, go to step 4. Please note that the generalized cost is not re-computed to avoid intense computation. Therefore, the overhead of the re-calculation of routes is just the computation of the shortest path (e.g. Dijkstra), which is small.

Mandatory re-routing can also be based on shortest-distance if we choose the shortest-distance path instead of the least generalized cost path each time when we recalculate an optimal route after a segment failure (each time when we encounter a failed segment in the calculated path, we need to recalculate). Section 6.3 gives a comparison of the performance. The mandatory re-routing algorithm applies to navigation under congestion in a similar way.

6.3 Experimental Results

In this section, disaster evacuation is selected for performance comparison, but the navigation under congestion is expected to have similar results.

6.3.1 Disaster Evacuation: Performance Comparison

In a set of experiments, we use the freeway network of Riverside County in Southern California, using the Census 2000 TIGER/Line shapefile format [53] as a case study.

Using ArcGIS 9.1 Geographical Information System tools[20], we extracted all the roadway segments whose Census Features Class Codes (CFCC) belongs to categories A1, A2, and A3 corresponding to the freeway categories. Abstract graphs were then built using the corresponding database files to describe the roadway connectivity. We assume that in a hypothetical emergency, people in the west area need to be relocated to the east. As shown in Figure 6.1, the map was divided into 8 sub areas manually according to the geographic clustering. People in area 1 through area 5 need to be relocated to area 6, area 7, or area 8. Two or three random nodes in each of the areas 1-5 and areas 6-8 are selected. There are 66 origin-destination pairs in total.
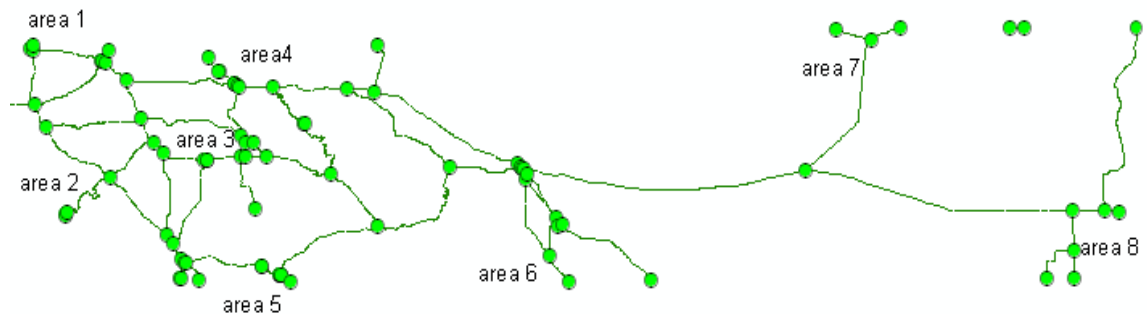

Figure 6.1: Map of Riverside County, southern California, with separate sub areas.

The following damage probabilities were used: 5%, 10%, 15%, 20%, 25%, 30%, 35%, and 40%. Each damage probability was applied to each direction of roadway segments (i.e. each direction of roadway segments is likely to be destroyed with a probability equal to the damage probability). Various seeds were used to generate the random destruction. The destruction of each direction of one roadway segment is

assumed to be independent of each another.

In this experiment, the generalized cost function is set as

$$Cost = \sqrt{length} * e^{\frac{min*max}{NMI}}$$

(6.9)

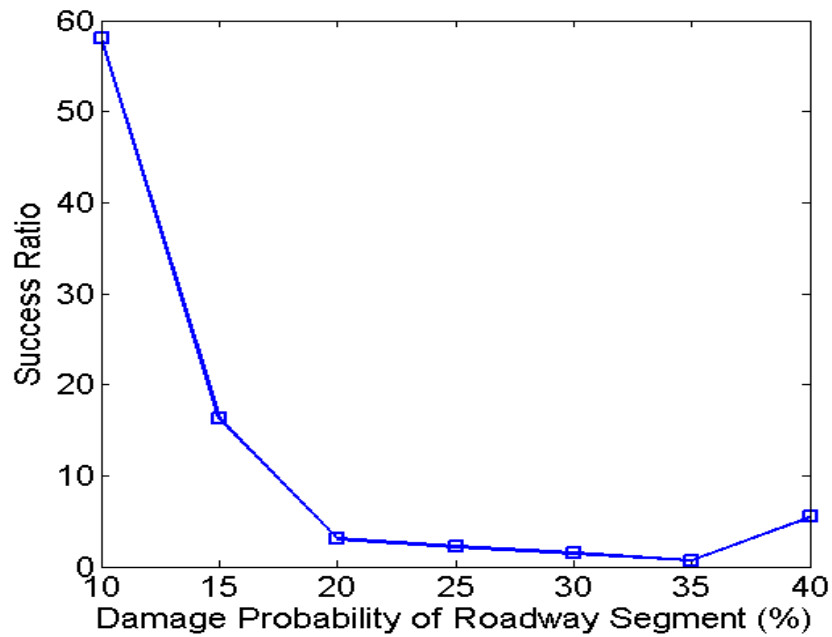putting a small weight on the length.

Figure 6.2: Success ratio vs. damage probability of roadway segments.

Figure 6.2 shows the results. The X axis is the damage probability, and y axis is the success ratio, which is defined as

$$SuccessRatio = \frac{N(True(NMI) \,\&\&\, False(distance))}{N(True(distance) \,\&\&\, False(NMI))}$$

where *N(True(NMI)&&False(distance))* means the number of occurrences when the NMI-based mandatory routing can lead the driver to the destination successfully without

being trapped in a dead end while shortest-distance -based mandatory routing failed (leading to a dead end). We can see that the NMI-based mandatory routing is universally better than the shortest-distance based mandatory routing (the ratio is always larger than *1*). The result of the 5% damage probability is not shown in Figure 6.2 because *N(True(distance)&&False(NMI))* is equal to zero. If either the roadway infrastructure is completely undamaged or a large portion (e.g. over 50%) of the roadway infrastructure is damaged by disaster, the two navigation algorithms will give pretty much the same results. However, when the infrastructure is destroyed mildly or moderately (i.e. 10-40%), the NMI-based mandatory routing will give better results.

Since we put a small weight on the length, one question might be asked: will the NMI-based mandatory routes be unreasonably longer than the shortest-distance-based mandatory routes? The answer is no. Figure 6.3 presents the statistics of trip length differences between all the successful trip pairs (for a successful trip pair, both the NMI-based and the shortest-distance-based mandatory routing are successful in leading the driver from the origin to the destination without being trapped in a dead end). The trip length difference is presented as a percentage, and is based on the distance of shortest-distance-based mandatory routing. We can see that for 48% of the trips, the two routing algorithms are the same distance-wise. For most of the trips, the trip length difference is within 3%.
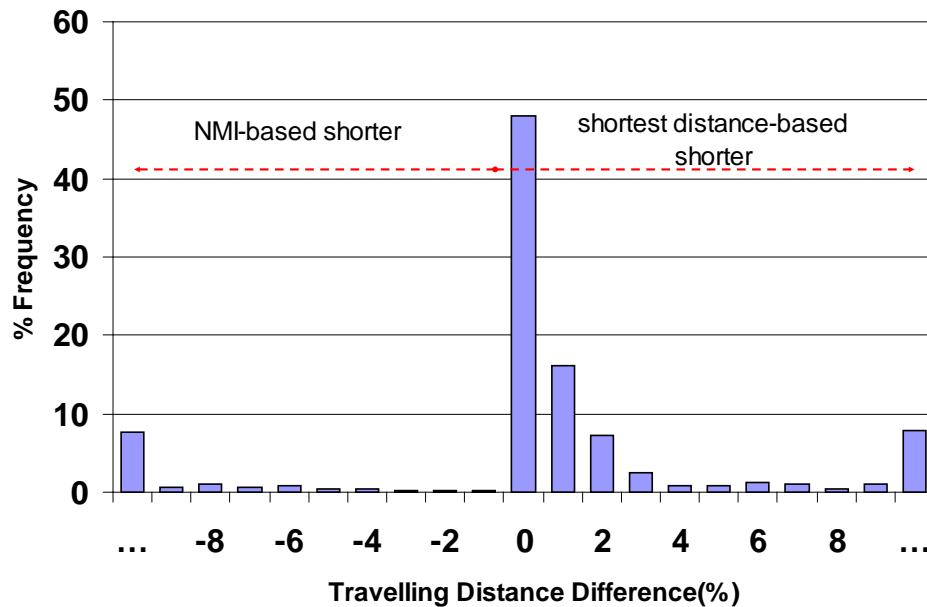
Figure 6.3: Histogram of trip length difference between successful trip pairs.

6.3.2 Disaster Evacuation: Case Study

The environment for the experiment in this section is as follows. The Riverside Country freeway network is again used, and people in the west need to be relocated to the east. The source node is Node 9 is area 1, and the destination node is Node 82 in area 7. The damage probability is 25%, i.e., each direction of each roadway segment is likely to be damaged with a probability of 25%. The destruction of each direction of one roadway segment is assumed to be independent from another.

From the two resulting routes shown in Figure 6.4, we can see that the NMI-based mandatory re-routing gives a reasonable route avoiding those links with low NMI. On the other hand, the shortest-distance-based routing provides a route that is trapped in area 4
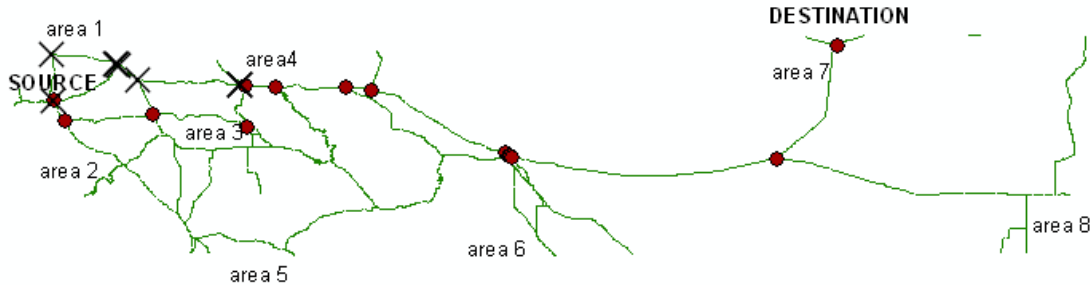
due to a roadway failure.

Figure 6.4: Example routing trajectories of NMI-based and distance-based mandatory re-route.

6.3.3 Navigation under Congested Conditions

The same algorithm can be applied to navigation under chaotic traffic conditions with a small modification. To avoid repetition, a pure simulation similar to section 6.3.1 is not included. However, for navigation under congestion conditions, it is possible to use real traffic data [67] for evaluation purposes, which is considered future work.

NMI-based mandatory re-routing is expected to be very useful since freeway incidents sometimes happen frequently. According to the incident data from PeMS [67], there were a total of 52,042 incidents that occurred on monitored Californian freeways from April 1, 2008 to April 30, 2008. Those incidents cause severe congestion, resulting in a lot of energy consumption and higher emissions. Figure 6.5 shows a plot of the number of incidents per day on a 59-mile section of State Route 91-E in Los Angeles area. The number of incidents each day is generally larger than 20.

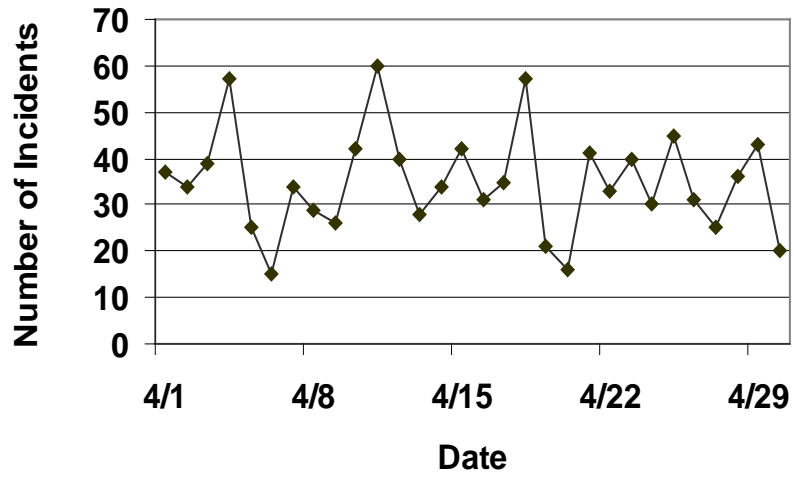**Freeway: SR91-E, 04/1/2008 to 04/30/2008**

Figure 6.5: Number of incidents for each day in April, 2008.

7. Conclusions and Future Work

In this dissertation, we have presented the development and application of an environmentally-friendly navigation system, described a wavelet-based roadway type classification, defined and justified the navigational mobility index concept, and proposed the NMI-based best degree of freedom routing algorithm. From this work, we can draw conclusion and identify some future work for each sub topic.

7.1 Environmentally-Friendly Navigation

This research presents the development of an environmentally-friendly navigation system methodology, which can calculate not only distance- and time-minimizing paths, but also energy- and emissions-minimizing paths. In this navigation system, a standard route minimization algorithm is applied to a new set of cost functions based on energy and emissions. These link-based energy/emissions factors were derived from a state-of-the-art mobile-source energy/emissions model that has been calibrated with real-world vehicle activity patterns. It was found that in many cases, a time-minimization path also minimizes energy and emissions. However, when congestion occurs, there are cases where this is not true. Because energy and emissions are often higher at congested speeds, a heavily congested (but shorter) path may not be the most environmentally friendly. In contrast, moderate congestion often provides a better choice from an energy and environmental perspective. Moderate congestion generally reduces average traffic

116

speeds from higher free-flow conditions, where vehicles have increased fuel consumption and emissions due to higher loads placed on the vehicle engines. Using historical data of two comparable freeways in the Los Angeles area, it was shown that for almost half of the time, the developed energy/environmentally-friendly navigation system can help a driver choose the best travel route that consumes the least fuel and/or produces the least amount of emissions. This may be highly desirable during this time period when fuel prices continue to hit record high and global warming as well as climate change have become major concerns worldwide.

To date, our studies have taken place on roadways that have real-time traffic performance data available, i.e., mostly on freeways in southern California. However, the method is also applicable to other types of roadways. For those roadways that do not have real-time traffic data, an estimated average speed value can be used. As the coverage of real-time traffic data expands (to include arterials and other roadways), the energy- and emissions-minimization navigation will be even more beneficial.

Further, road grade can also have a significant impact on fuel consumption and emissions of vehicles, especially for heavy-duty trucks. Thus, the developed energy and environmentally-friendly navigation system can be expanded to take into account road grade information in its routing decision, which can contribute to significant fuel savings. In the future, we will develop a prototype on-board navigation system that not only provides the real-time shortest-duration and shortest-distance routing choices, but also the lowest energy and emissions choices.

7.2 Vehicle Velocity-Based Roadway Type Recognition by Wavelet Analysis

In this sub topic, a wavelet-based pattern recognition methodology has been proposed to classify real-time vehicle velocity trajectories from probe vehicles. This potentially allows the probe vehicle to estimate roadway facility type and possibly congestion conditions (e.g. level-of-service or LOS). This information is complementary to standard traffic monitoring systems (i.e., those that rely on embedded loop detectors in the roadway). The system can correctly estimate roadway conditions 90% of the time. This performance can be improved with additional data in the original training set.

In terms of future work, it is planned to expand the experimentation with additional training sets and evaluation runs. Congestion level classification is currently underway and will continue as part of the future work. Further, with the basic vehicle trajectory recognition system operating successfully, it is planned to develop functional relationships between the estimated congestion levels and traffic flow and speed data obtained from macroscale roadway sensors. In this way, it will be possible to estimate localized fuel consumption and pollutant emissions based simply on traffic flow and speed data.

7.3 Navigation Mobility Index

Due to dynamic traffic conditions, a certain pre-selected freeway route corresponding to standard measures of travel distance, time, or energy cost may not always be achievable. In this study, the NMI based on the availability of freeway route choices has been developed and proposed with demonstrated applications in roadway navigation. NMI can be based on the number of possible routes weighted by shared segments among routes from a source to a destination (node-to-node NMI), or a single source to a destination (node-NMI) or on an average of nodes in a specific area. The proposed NMI has several attractive properties, among which the most important one is adding additional infrastructures will not decrease the NMI.

The NMI can be used to: 1) measure the degree of freedom in which drivers can choose routes; 2) evaluate the potential effectiveness of navigation systems in an area; 3) assess the overall degree of freedom level of an area; 4) guide the movement of people during emergency evacuation; and 5) determine a route having the least generalized cost incorporating both the distance and the NMI in route selection. Experimental results were presented to show the effectiveness of the proposed concepts and algorithms. Specifically, the LA County network and the SF Bay Area network were compared, and it was found that the LA County network has better freeway connectivity in terms of re-routing options.

In terms of future work, we plan to: 1) consider additional cities/counties to be studied and determine their navigational mobility indices that can be compared. Additional

roadway types (e.g. arterials) will be included in the analysis; 2) develop a faster algorithm/implementation that enables real-time deployment in large networks; 3) investigate how to construct a more effective generalized cost based on NMI, and to investigate how to construct a more effective general function G_{l_a}.

7.4 NMI-Based Mandatory Re-Routing Algorithm

In Chapter 6 we discuss two possible scenarios of mandatory re-routing: traffic under severe congestion and evacuation in a disastrous event. Traffic conditions are often unpredictable in metropolitan areas as unexpected freeway incidents occur on a regular basis. During traffic congestion, drivers normally want to change routes because the pre-defined shortest-distance or shortest-duration route is no longer optimal. We propose an NMI-based mandatory re-routing method that gives drivers the highest degree of freedom in choosing routes if they would like to change routes while in travel. The NMI-based mandatory re-routing can also be used in evacuation applications. During disasters, roadway infrastructure may be damaged, and people in affected areas need to be relocated to safe area as soon as possible without an accurate map of the latest roadway conditions. Experimental results from a hypothetical case study show that the success ratio is universally higher than 1, which suggests a better performance of the NMI-based mandatory re-routing over the shortest-distance-based mandatory re-routing.

Future work includes the following aspects: 1) integrate a practical disaster (e.g. earthquake) roadway infrastructure destruction model. For example, in [68] the researchers developed methods for evaluating the performance of highway systems

subjected to a severe earthquake. They explicitly modeled the transportation network performance and introduced probabilistic earthquake scenarios. In this research, we assume that destruction of each road segment and destruction of each direction of one roadway segment are independent. However, this may not be the case; 2) integrate the real time traffic speed collected from loop sensors (e.g. PeMS system), and use those data to calculate the travel time for comparison purposes; 3) find new forms of a generalized cost function and associated parameters that gives better performance in terms of success ratio and more robust routes.

7.5 Publications Resulting From This Research

[1] W. Zhu, K. Boriboonsomsin, and M. Barth, "Defining a freeway mobility index for roadway navigation," *Journal of Intelligent Transportation Systems,* submitted for publication, third round review.

[2] W. Zhu, K. Boriboonsomsin, and M. Barth, "Development and application of an environmentally-friendly navigation system," *IEEE Trans. On Intelligent Transportation System,* submitted for publication, second round of review.

[3] W. Zhu, K. Boriboonsomsin, and M. Barth, "Mobility index-based navigation for mandatory re-routing scenarios," in *Proc. IEEE Int. Conf. Transp. Syst*., Beijing, China, Oct. 2008, pp.581-586.

[4] W. Zhu, K. Boriboonsomsin, and M. Barth, "Mobility index-based routing for dynamic navigation systems," *IEEE Trans. On Intelligent Transportation System,* in preparation, 2009.

[5] W. Zhu, K. Boriboonsomsin, and M. Barth, "A new methodology for processing time varying data in multiple states," in *Proc. IEEE Int. Conf. Transp. Syst.*, Beijing, China, Oct. 2008, pp.234-239.

[6] W. Zhu, K. Boriboonsomsin, and M. Barth, "Microscopic traffic flow quality of service experienced by drivers," in *Proc. IEEE Int. Conf. Transp. Syst.*, Seattle, Sep. 2007, pp. 47-52.

[7] W. Zhu and M. Barth, "Vehicle trajectory-based road type and congestion recognition using wavelet analysis," in *Proc. IEEE Int.. Conf. Transp. Syst.*, Toronto, Canada, Sep. 2006, pp 879-884.

[8] M. Barth, K. Boriboonsomsin, W. Zhu, A. Vu, A. Gerdes, C. Lee, and D. Rosario, "Environmental-friendly navigation: technology description and field operational testing plan," *15th World Congress on ITS*, New York City, Nov. 16-20, 2008.

Bibliography

[1] Department of Energy, Energy Information Administration (2001), *Transportation Energy Consumption by Vehicles*. Available: http://www.eia.doe.gov/emeu/rtecs/contents.html.

[2] D. Schrank and T. Lomax (2005), *The 2005 Urban Mobility Report*, Texas Transportation Institute. Available: http://www.apta.com/research/info/online/urban_mobility.cfm

[3] K. Anh, and H. Rakha, "The effects of route choice decisions on vehicle energy consumption and emissions," *Transportation Research Part D*, 13(3), 151-167, 2008.

[4] H. C. Frey, K. Zhang, and N. M. Rouphail, "Fuel use and emissions comparison for alternative routes, time of day, road grade, and vehicle based on in-use measurements," *Environmental Science & Technology*, 42, 2483-2489, 2008.

[5] E. Ericsson, H. Larsson, and K. Brundell-Freij, "Optimizing route choice for lowest fuel consumption – potential effects of a new driver support tool," *Transportation Research Part C*, 14, 369-383, 2006.

[6] M. Barth, F. An, T. Younglove, C. Levine, G. Scora, M. Ross, and T. Wenzel., "The Development of a Comprehensive Modal Emissions Model," Final report submitted to the National Cooperative Highway Research Program, Nov. 1999.

[7] M. Barth, et al., "Modal Emissions Modeling: A Physical Approach," *Transport. Res. Rec.*, no. 1520, pp. 81-88, 1996.

[8] M. Barth, T. Younglove, T. Wenzel, G. Scora, F. An, M. Ross, and J. Norbeck, "Analysis of modal emissions from a diverse in-use vehicle fleet," *Transport. Res. Rec.*, no. 1587, pp. 73-84, 1997.

[9] M. Barth, T. Younglove, and G. Scora., "The development of a heavy-duty diesel vehicle model," *Transport. Res. Rec.*, no. 1880, pp. 10-20, 2004.

[10] M. Barth, C. Malcolm, T. Younglove, and N. Hill, "Recent validation efforts for a Comprehensive Modal Emissions Model," *Transport. Res. Rec.* no. 1750, pp. 13-23, 2001.

[11] R. Dowling, R. Ireson, A. Skarbardonis, D. Gillen, and P. Stopher., "Predicting Air Quality Effects of Traffic-Flow Improvements: Final Report and User's Guide," *NCHRP Report 535*, Transportation Research Board, Washington, D.C., 2005.

[12] P. C. Fernández, and J. R. Long, "Grades and other loads effects on on-road emissions: an on-board analyzer study," in *Proc. 5th CRC On-Road Vehicle Emission Workshop*, San Diego, April. 1995.

[13] S. Park and H. Rakha, "Energy and environmental impacts of roadway grades," in *Proc. 85th Annual Meeting of Transportation Research Board* , Washington, D.C., Jan. 2006.

[14] C. Chen, Z. Jia, K. Petty, J. Shu, A. Skabardonis, and P. Varaiya, "Freeway Performance Measurement System (PeMS) shows big picture," feature article, *California PATH Intellimotion*, vol. 9, no. 2, 2000.

[15] T. Choe, A. Skabardonis, and P. Varaiya., "Freeway performance measurement system (PeMS): an operational analysis tool," in *Proc. 81st Annual Meeting of Transportation Research Board*, Washington, D.C. Jan. 2002.

[16] M. Coelho, T. Farias, and N. Rouphail, "Impact of speed control traffic signals on pollutant emissions," *Transportation Research Part D*, no. 10, pp. 323–340, 2005.

[17] K. Boriboonsomin, and M. Barth, "Fuel and CO2 impacts from advanced navigation systems that account for road grade," in Proc. *88th Annual meeting of Transportation Research Board*, Washington, DC, Jan 2009, to appear.

[18] A. Levintin, *Introduction to the Design& Analysis of Algorithms*. Addison-Wesley, 2003.

[19] I. Chabini and S. Lan, "Adaptation of A* algorithm for the computation of fastest path in deterministic discrete-time dynamic networks," *IEEE Trans. Intell. Transp. Syst.,* vol. 3, no.1, pp. 60-74, Mar. 2002.

[20] Environmental Systems Research Institute, Inc., ArcGIS GIS Software Tools, Available: http://www.esri.com.

[21] M. Barth, and K. Boriboonsomsin, "Real-world CO2 impacts of traffic congestion," in *Proc. 87th Annual meeting of Transportation Research Board* , Washington, DC, Jan. 2008.

[22] L. A. Wischhof and H. Rohling, "Information dissemination in self-organizing intervehicle networks," *IEEE Trans. On Intelligent Transportation System*, Vol. 6, No.1, March 2005.

[23] H. Xu and M. Barth, "A transmission-interval and power-level modulation methodology for optimizing inter-vehicle Communications," in *Proc. of the first ACM International Workshop on Vehicular Ad Hic Networks*, Philadelphia, Pennsylvania, October 2004.

[24] Y. Ping, "Linear prediction of traffic volume," in *Proc. of 2005 Transportation Research Board Annual Meeting*, Washington D.C., January 2005, also at http://www.missouri.edu/~sunc/TRB/presentations.html.

[25] L. Yu, "Wavelet-based determination of aggregation levels for ITS data," in *Proc. of 2005 Transportation Research Board Annual Meeting*, Washington D.C., January 2005, also at http://www.missouri.edu/~sunc/TRB/presentations.html.

[26] X. Jiang and H. Adeli, "Wavelet packet-autocorrelation function method for traffic flow pattern analysis," *Computer-Aided Civil and Infrastructure Engineering*, Vol. 19, 2004, pp. 324-337.

[27] Y. Freund and R. E. Schapire, "Experiments with a new boosting algorithm," in *Proc. Of the 13th Intl. Conf. on Machine Learning*, 1996, pp. 148-156.

[28] R. E. Schapire, "The strength of weak learnability," *Machine Learning*, vol. 5, no. 2, pp. 197-227, 1990.

[29] P. Viola, M. Jones, "Robust real-time object detection," in *Proc. of IEEE Workshop on Statistical and Computational Theories in Computer Vision*, 2001.

[30] Z. Qian, D. Metaxas, and L. Axel, "Boosting and Nonparametric Based Tracking of Tagged MRI Cardiac Boundaries," in *Proc. of MICCAI, LNCS 4190,* pp 636-644. 2006.

[31] R. Agrawal, C. Faloutsos, and A. Swami, " Efficient similarity search in sequence database," in *Proc. Fourth Int'l Conf. Foundations of Data Organization and Algorithm*, 1993.

[32] K. P. Chan and A. Fu, "Efficient time series matching by wavelets," in *Proc. Int'l Conf. Data Eng*., 1999.

[33] E. Hernandez and G. Weiss, *A first Course on Wavelets*, CRC Press, 1996.

[34] D. B. Percival and A.T. Walden, *Wavelet Methods for Time Series Analysis*, Cambridge University Press, 2000.

[35] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*, Wiley Press, New York, 2000.

[36] Y. Huhtala, J. Karkkainen, and H. Toivonen, "Mining for similarities in aligned time series using wavelets," *SPIE Conference on Data Mining and Knowledge Discovery*, Vol. 3695, pp. 150-160.

[37] R. E. Schapire, "Using output codes to boost multiclass learning problems," in *Proc. Of the 14th Intl. Conf. on Machine Learning*, 1997, pp. 313-321.

[38] T. J. Sejnowski and C. R. Rosenberg, "Parallel networks that learn to pronounce English text," *Complex Syst.,* vol. 1, pp. 145-168, 1987.

[39] T. G. Dietterich and G. Bakiri, "Solving multiclass learning problems via error-correcting output codes," *J. Artific. Intell. Res.,*,vol. 2, pp. 263-286, 1995.

[40] V. Guruswami and A. Sahai, "Multiclass learning, boosting, and error-correcting codes," in *Proceeding. Of the 12th Conference on Computational Learning Theory*, pp. 145-155, 1999.

[41] C. Chen, K. Petty, A. Skabardonis, P. Varaiya and Z. Jia, "Freeway performance measurement system: mining loop detector data," in *Proc. of the 80th Annual Meeting of the Transportation Research Board*, Washington D. C., January 2001.

[42] S. Handy, "Accessibility-vs-mobility-enhancing strategies for addressing automobile dependence in the US," *European Conference of Ministers of Transport*, Tokyo, Japan, 2002.

[43] M. Ahmed and D. M. Levinson, "Access to destination: development of accessibility measure," Technical Report, University of Minnesota, available at: http://www.lrrb.org/PDF/200626.pdf

[44] S. Baradran, and F. Ramjerdi, "Performance of accessibility measures in Europe," *Journal of Transportation and Statistics,* Sept-Dec., pp 31-48, 2001.

[45] W. Bruce, D. Liu, and S. Singer, "Accessibility measures of U.S. metropolitan areas," *Transportation Research B*, **27B**, No.6, pp. 439-449, 1993.

[46] F. Bruinsma, and P. Rietveld, "The accessibility of European cities: theoretical framework and comparison of approaches," *Environmental and Planning* A 30(3), pp.499-521, 1998.

[47] S. Hanson, and M. Schwab, "Accessibility and intra-urban travel," *Environment and Planning,* A, 19, pp.735-748, 1987.

[48] D. R. Ingram, "The concept of accessibility: a search for an operational form," *Regional Studies*, **5**, pp. 101-107, 1971.

[49] M. G. H. Bell, "A game theory approach to measuring the performance reliability of transport networks," *Transportation Research B*, **34**, pp. 533-545, 2000.

[50] A. Chen, H. Yang, H. Lo, and W. Tang, "Capacity reliability of a road network: an assessment methodology and numerical results," *Transportation Research B*, **36**, pp. 225-252, 2002.

[51] S. Bekhor, M. Ben-Akiva, and M. Ramming, "Adaptation of logit kernel to route choice situation," *Transportation Research Record* **1805**, pp.78-85, 2002.

[52] S. Bekhor, M. E. Ben-Akiva, and M. S. Ramming, "Evaluation of choice set generation algorithms for route choice models," *Annals of Operations Research*. **144.** pp. 235-247, 2006.

[53] Tiger Map Shape files, available at http://arcdata.esri.com/data/tiger2000/tiger_statelayer.cfm?sfips=06

[54] J. Y. Yen, "Finding the k shortest loopless paths in a network," *Management Science*, pp. 712-716, 1971.

[55] E. Q. V. Martins and M. M. B. Pascoal, "A new implementation of Yen's ranking loopless paths algorithm," *Q. J. Belgian, French Ital. Oper. Res. Soc.* **1(2)**, pp. 122-133, 2003.

[56] R. B. Dial, "Probability assignment: a probabilistic multipath traffic assignment model which obviates path enumeration," *Transportation Research*, **5**, pp. 83-111, 1971.

[57] Google Map, available at http://maps.google.com.

[58] Mapquest, available at http://www.mapquest.com.

[59] Google Earth, available at: http://earth.google.com.

[60] M. Barth, K. Boriboonsomsin, and A. Vu, "Environmentally-friendly navigation," in *Proc. IEEE Int. Conf. Transp. Syst.*, Washington, Sept. 2007, pp.684-689.

[61] W. Zhu, K. Boriboonsomsin, and M. Barth, "Mobility index-based navigation for mandatory re-routing scenarios," in *Proc. IEEE Int. Conf. Transp. Syst.*, Beijing, China, Oct. 2008, pp.581-586.

[62] S. Russell, and P. Norvig, Artificial Intelligence: A Modern Approach, Prentice Hall, 2003.

[63] P. Makris, A. Makri, and C. Provatidis, "Energy-saving methodology for material handling applications," *Journal of Applied Energy*, vol. 83, pp. 1116-1124, 2006.

[64] K. Anh, and H. Rakha, "Energy and environmental impacts of route choice decisions", in *Proc. 86th Annual. Meeting of Transportation Research Board*, Washington D.C., Jan. 2007.

[65] W. Zhu, K. Boriboonsomsin, and M. Barth, "Defining a freeway mobility index for roadway navigation," *Journal of Intelligent Transportation Systems,* submitted for publication.

[66] J. Kwon, B. Coifman, and P. Bickel, "Day to day travel time trends and travel time prediction from loop detector data," *Transport. Res. Rec.* , no. 1717, pp. 120-129, 2000.

[67] Online PeMS Database. Available: http://pems.eecs.berkeley.edu/?dnode=State

[68] N. Shiraki, M. Shinozuka, J. Moore, S. E. Chang, H. Kameda, and S. Tanaka, "System risk curves: probabilistic performance scenarios for highway networks subject to earthquake damage," *Journal of Infrastructure Systems*, vol. 13, no. 1, pp. 43-54, 2007.

[69] M. Barth, E. Johnston and R. R. Tadi, "Using GPS technology to relate macroscopic and microscopic traffic parameters," *Transportation Research Record*, No. 1520, Washington D.C. 1996, pp. 89-96.

[70] Freeway Management and operations handbook, available at: http://ops.fhwa.dot.gov/freewaymgmt/publications/frwy_mgmt_handbook/toc.htm.

[71] N. Cohn, "Real time traffic information and navigation: an operational system," in *Proc. 88th Annual meeting of Transportation Research Board* , Washington, DC, Jan. 2009.

[72] Tomtom, available at: http://www.tomtom.com/hdtraffic.

[73] A D. May, *Traffic flow fundamentals*. Prentice Hall, 1990.

[74] Garmin Nuvi 270 Portable GPS Navigation, available at: http://www.amazon.com/dp/B000OH26OM/?tag=googhydr-20&hvadid=2689019717&ref=pd_sl_10wm9xrl2o_b

[75] W. Zhu, K. Boriboonsomsin, and M. Barth, "Development and application of an environmentally-friendly navigation system," *IEEE Trans. On Intelligent Transportation System,* submitted for publication, second round of review.

[76] W. Zhu and M. Barth, "Vehicle trajectory-based road type and congestion recognition using wavelet analysis," in *Proc. IEEE Int.. Conf. Transp. Syst.*, Toronto, Canada, Sep. 2006, pp 879-884.