# Modeling Long Distance Highway Passenger Travel: A National Data Framework Approach

Runze Yu

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2012

Reading Committee:

Yinhai Wang, Chair

Edward D. McCormack

Qing Shen

Program Authorized to Offer Degree:

Department of Civil and Environmental Engineering

University of Washington

**Abstract**

Modeling Long Distance Highway Passenger Travel: A National Data Framework Approach

Runze Yu

Chair of the Supervisory Committee:

Professor Yinhai Wang

Department of Civil and Environmental Engineering

Long distance travel is a critical component of American life, generating immense impacts on a variety of aspects of the society. In 1995, American households took about 656 million long distance domestic trips (100+ miles, one way), which totaled 1 billion person trips. Compared to the well-studied intra-urban travel, long distance travel needs further in-depth analysis and modeling efforts, for which the long distance travel data serve as a substantial basis. While MPOs periodically collect both supply and demand data to support the development of regional models, supply-side data reflecting transportation service and costs at corridor, interregional, and national level are scarce. To merge the gap, there is a need to establish a national data framework, with available datasets integrated to support implementing quantitative methods on a national highway network to resolve network loading issues for long distance travel.

This research intends to address this imperative. The research scope embraces the long distance passenger travel on national highways within the contiguous United States. The framework developed comprises not only a data warehouse where a plethora of datasets reside and interact, but also a complete modeling methodology to infer network loading conditions of long distance travel. Applying the national data framework approach, this research answers a pivotal question: how can we gain knowledge and insights of distributive patterns of long distance passenger travel at interregional level given existing data resources?

This research elaborates the methodologies for the framework design and addresses the technical and theoretical challenges particularly in issues of travel impedance estimation and long distance travel network loading. Findings from this research lay a solid foundation for building the fundamental theoretical framework urgently needed for long distance travel studies and contribute significantly to the understanding of the supply side of long distance travel, from available infrastructure capacity and networks loading conditions to the induced environment challenges, infrastructure congestion, and investment costs, as well as facilitating policy making and evaluation pertain to long distance travel at national level.

## TABLE OF CONTENTS

## LIST OF FIGURES

## LIST OF TABLES

# ACKNOWLEDGEMENTS

It is a long distance trip to complete this dissertation and the subsequent Ph.D. degree, probably the most challenging activity of my first 26 years of my life. The best and worst moments of my doctoral journey have been shared with many people. It has been a great privilege to spend five years in the Department of Civil and Environmental Engineering at the University of Washington, and its members will always remain dear to me.

My first debt of gratitude must go to my adviser, Prof. Yinhai Wang. He patiently provided me the vision, encouragement, and advice to fight through the graduate program and complete my dissertation. I want to thank him not only for being a strong and supportive adviser to me throughout my graduate school career, but also for being a role model as a great human being.

Special thanks to my supervisory committee, Dr. Qing Shen, Dr. Edward D. McCormack, and Dr. Timothy L. Nyerges for their support, guidance, and helpful suggestions. Their guidance has served me well and I owe them my heartfelt appreciation.

Members of the STAR Lab at the University of Washington also deserve my sincerest thanks, their friendship and assistance has meant more to me than I could ever express. I could not complete my work without invaluable friendly assistance of Ms. Sa Xiao, Ms. Xiaoyue Liu, and Mr. Yegor Malinovskiy. I should also mention Mr. Patrick Zhang from the FHWA and Mr. Roger Mingo from the R.D. Mingo and Associates for their valuable comments and crucial help with the research.

My friends in the U.S., China and other parts of the world were sources of laughter, joy, and support. Special thanks go to Li Ma, Li Ying Leung, Ti Wu, Yen-Yu Lin, and Christine Liu; my friends from the Chinese Students and Scholars Association and the ITE UW Chapter, my co-workers from the World Bank, and all the 2012 Eno Fellows. I am very happy that, in many cases, my friendships with you have extended well beyond our shared time. I also wish to thank my entire family for providing a loving environment for me.

Lastly, and most importantly, I wish to thank my parents, Qun Song and Weidong Yu. They bore me, raised me, supported me, taught me, and loved me. To them I dedicate this dissertation.

## DEDICATION

To my parents.

# Chapter 1 Introduction

## 1.1 Problem Statement

### *1.1.1 Data needs for long distance travel*

Long distance travel is an indispensable component of American life. In 1995, American households took about 656 million long distance domestic trips (100+ miles, one way), which totaled 1 billion person trips (BTS, 1997). Long distance passenger travel connects far flung families and friends, attracts people to natural sceneries and cultural attractions, supports the national and regional economy and boosts the local business through tourism. According to the 1995 American Travel Survey (ATS), primary trip purposes of long distance travel involve visit friends and relatives (33%), leisure (30%), business (22%), and personal business (15%). A strong reliance on automobile is reflected in long distance passenger travel: personal use vehicles account for about 75% of household trips and resulting in over 280 billion vehicle miles of travel on the nation's highways. The massive travel generates extensive impacts on broad aspects including infrastructure operation, travel reliability, urban congestion, and energy use.

To understand and address those impacts induced by long distance travel on our society, we need a comprehensive picture portraying long distance movements. On the one hand, this comprehensive picture depicts the operational reality of the existing infrastructure, by defining markets for major transportation facilities and corridors of national significance, and identifying congestion bottlenecks within urban areas and on rural portions of intercity links. On the other hand, this comprehensive picture helps allocate potential investments to the future needs for maximizing the network performance. Altogether it facilitates justification or evaluation of a variety of potential or existing policies and regulations, embracing highway design, environmental planning, and safety improvement.

The comprehensive picture is too enormous to be drawn from any single data source. Over the past few decades, massive amount of data has been accumulated through various forecast models, national and local travel surveys, administrative records, roadside counts, etc. However, from a travel demand/supply perspective we find that for long distance travel, there is a plethora of data sources on the ***demand*** side and limited data sources on the ***supply*** side. Specifically, we know relatively well about the origin-destination (O-D), mode share, and estimated trip generation for long distance travel; yet regarding the available infrastructure capacity and networks loading conditions, our knowledge is constrained mainly due to the incompleteness of a national data framework.

### *1.1.2 National data framework*

A national data framework integrates diverse datasets. Particularly in a long distance travel context, it includes an abstraction of the national highway network as common ground of basic geographic representation to which all the datasets refer. The framework can be stored in one geodatabase, or through access to several external sources with a collaborative mechanism. Within the framework there are also algorithms manipulating the datasets for trip assignment and network loading. As a ***data warehouse***, the framework exports popular data themes regarding long distance travel to different stakeholders including government, industry, and academia. As a ***computational infrastructure***, the framework facilitates portraying the supply side of the comprehensive picture, to resolve network loading and distribution issues of long distance passenger travel on the national highway network.

Currently there is a scarce of knowledge regarding national data framework for long distance travel in the United States, especially in the passenger travel field. In the freight world, the Freight Analysis Framework 3 ($FAF^3$) presents a fairly complete travel data framework. It integrates a variety of data sources to yield estimates for commodity tonnage and value; it also provides methodology for assigning freight trips to the highway network (Battelle, 2011). However, $FAF^3$ embraces the entire freight travel rather than focuses on the long distance ones. On the passenger travel side, Georgia Department of Transportation conducted a Southwest Georgia Interstate Study (Georgia Department of Transportation, 2009) by building a geodatabase covering a five-level network in multiple resolutions. However, there was no computational function developed in the study. By contrast Nielsen and Burgess (2008) led the efforts of establishing the European TRANS-TOOLS Transport Model covering all European Union member countries, with both a geodatabase implementation and travel demand and assignment models. Although the model is not long distance travel oriented, the effort indicates that a large-scale travel data framework can go beyond the basic needs for network performance monitoring and transportation planning; the framework itself could also leverage numerous potentials for energy preservation, environmental analysis, and system reliability studies as a computational base.

### *1.1.3 Long Distance Travel Network Loading*

Network loading, traffic assignment, and route choice have been referred interchangeably in previous literature; however, the author would like to differentiate their emphases. Route choice focuses on individual choices, which reflects the interaction between the individual (the traveler profile and traveling awareness) and the surrounding environment (the resources and characteristics of the route); traffic assignment is an aggregated outcome of individual behavior, usually summarized as link flows in an intra-urban context; network loading exhibits the distributive pattern in an interregional resolution, where detailed choice information is abstracted into macroscopic statistics. Overall all the three different terms reflect travelers' choices in different geographic and behavioral dimensions, and obviously this research concentrates on the network loading issues on a national highway network.

Resolving network loading issues in a long distance travel context presents unique challenges from the well-studied intra-urban traffic assignment problems. It is still unclear that whether intra-urban research findings can be well adapted to long distance travel scenarios. Several critical issues are enumerated and discussed as follows.

#### 1.1.3.1 Long distance travelers: a very diverse portfolio

People who travel long distances constitute a very diverse traveler portfolio, even if they only account for 40% of the U.S. population. While the intra-urban travel is mostly centered around people's daily life (work, school, home, shopping, etc.), long distance trips are pursued to embrace a wide spectrum of activities. The 1995 American Travel Survey (ATS) roughly categorized long distance trips by four trip purposes: business, visiting friends or relatives, leisure, and personal business, of which leisure trips refer to the ones for rest or relaxation, sightseeing, outdoor recreation, entertainment, or shopping, and personal business trips involved school-related activities as well as other personal or family matters like seeking for medical care.

Additionally, life styles and values of travelers significantly influence their travel behavior when planning for long distance travel. The idiosyncrasy in life styles and values can be captured by some socio-demographic characteristics, including age, race, education, working status, household type (family/nonfamily; with/without children; living alone/not living alone), household structure (house, apartment, etc.), household size, household income, etc. For example, Georggi and Pendyala (2001) found that both the elderly and the low income undertake significantly fewer long-distance trips than other socioeconomic groups. It was also found that these groups traveled less for leisure purposes, and buses were frequently used as the traveling mode.

#### 1.1.3.2 Pleasure travel account for a major proportion of long distance travel

Pleasure travel summarizes trips for leisure and those for the purposes of visiting friends and relatives, since both categories share considerable similarities in terms of travelers' sensitivity level to travel cost. Pleasure travel takes a substantial share of all trip purposes in the long distance passenger travel domain. According to the 1995 ATS, 62.9% of long distance trips are for pleasure across all the modes. Particularly, leisure travel increased by 122% between 1977 and 1995, as a major driver for the total growth of long-distance travel (BTS, 1998).

And it is generally believed that pleasure travel will continue to grow, despite the current economic downturn. A survey conducted by the Priceline.com has shown that 84% of the travelers were still planning for vacation trips even with increasing travel cost (Hotel News Resource, 2007). Traveling for exploration seems to be an integral part of American life and people are spending substantially on it (Focalyst, 2007). The U.S. Bureau of Labor Statistics (BLS, 2010) reported that in 2008, a U.S. household spent an average of $1,415 on vacation and pleasure trips. Notably, the retiring baby-boomer generation will serve as a driving force to the growing demand for pleasure travel in the coming decade (Mallet and McGuckin, 2000; Davies 2005).

Pleasure travel possesses several unique characteristics. First, higher utilization of surface transportation modes: two-thirds of all long-distance person trips by personal use vehicle in 1995 were for pleasure (35% to visit friends or relatives and 31% for leisure travel). Especially for leisure travel, personal use vehicles are the dominating mode with an 82% mode share. Second, pleasure travel is usually associated with bigger travel groups. The average size of the travel party on leisure trips in 1995 was 3.9 persons, highest among all the four trip purposes. Third, people on leisure trips take longer time and distance on the road. On average 3.3 nights away from home were spent on leisure travel according to the 1995 ATS. Fourth, pleasure travel supports local economy through tourism due to its highly consumption-oriented nature. Capturing those intrinsic characteristics of pleasure travel comprises another challenge for this research.

#### 1.1.3.3 Assignment procedures and travel impedance

Loading long distance trips to the national highway network presents a large-scale trip assignment problem. While concepts of traffic assignment are straightforward, details pertaining to the assignment on a national basis were barely endeavored in previous research and need further exploration. A methodological framework is in need to address several issues in this procedure: first, will intra-urban assignment techniques be readily transferred to a long distance travel context? Second, will Static Traffic Assignment (STA) meet the needs for large-scale assignment? Or does it require a more complicated Dynamic Traffic Assignment (DTA) approach at increased computational cost? Third, how to generalize travel cost for long distance travel, given the aforementioned diversity in the travel behavior?

Of all the issues generalized travel cost is considered as the major subject in this study. The very much diversified trip purposes result in a significant divergence in people's perception and sensitivity to travel time and travel distance, two major measurements of travel cost in intra-urban traffic assignment. Only taking those conventional measurements into account will not sufficiently represent long distance travel cost. For example, of four trip purposes the leisure travel is the least constrained by the time and financial budget, considering the very nature of its. Leisure travelers may prefer scenic highways over interstate freeways even though the latter is generally associated with higher reliability and shorter travel time. Quantifying the "pull" and "push" factors and integrating them into the methodological framework will be detailed in the following chapters.

#### 1.1.3.4 Seasonality in travel demand

Very strong seasonality effect is revealed from long distance passenger travel, due to the large share of pleasure travel. The 1995 ATS shows that long distance travel peaked during the third quarter, July through September, when 35% of vacations and 30% of weekend trips occurred. The largest share of leisure trips was also observed during the third quarter (38%), along with 27% of trips to visit friends or relatives. The fourth quarter, when holiday travel is prevalent, was the peak period to visit friends or relatives (30%). Business travel was evenly distributed during the first three quarters of the year at 26% in each quarter, but somewhat lower, about 22%, in the fourth quarter. Temporal resolution of loading long distance travel should take into account of the varying seasonality across different trip purposes.

## 1.2 Research Background

Long distance travel fulfills some unique needs for individual and commercial movements in this nation. People travel longer distance to make business happen, visit friends or relatives, and explore the nation's natural sceneries and cultural attractions. The long-haul trucking transports all kinds of commodities across the country to activate our economy. Those needs are beyond the scope of urban travel and generate immense impacts on the service of transportation infrastructure and facilities. Air pollution, travel delay, and safety exposure induced by the burgeoning demand are turning the joy and freedom of traveling into chores gradually.

Collecting information pertain to long distance travel to estimate those impacts has been a dedicated mission and practice of U.S. Department of Transportation (DOT). Since the 1960s a series of nationwide travel surveys have been conducted, of which three most recent ones are the 1995 American Travel Survey (ATS), 2001-2002 National Household Travel Survey (NHTS), and 2008-2009 NHTS.

The 1995 ATS is a legendary travel survey aiming at providing accurate and comprehensive long distance passenger flow information to assist policy making. Conducted by the Bureau of Transportation Statistics (BTS) in 1995, the ATS is over 15 years old and still the latest large scale long distance travel survey to date and serves as the estimation basis of many studies in the long distance travel analysis and forecast field. The 2001-2002 NHTS updates information from previous years' Nationwide Personal Transportation Surveys (NPTSs) and ATSs, with a long distance survey component (Travel Period trips) but smaller sample size. The 2008-2009 NHTS abandons this component and only focuses on Travel Day trips in the main part of the survey and only collects very limited samples for long distance travel in some of its add-on programs.

Long distance trips are defined differently in those surveys. A 1995 ATS trip refers to a trip from home to the farthest destination that is 100 miles or longer. The trip includes all the overnight stops made along the way if there is any and excludes commuting to work trips. The 2001 NHTS defines long distance trips as 50+ miles (one way) including the commuting trips. Long distance trips discussed in this proposal follow the 1995 ATS definition, which reflects the research focus on inter-regional and interstate travel, regardless of the daily commute in urban or suburban areas.

The burgeoning demand, the aging infrastructure with limited expansion potential, and the recent interest in high-speed rail investment accentuates the continuing imperative of long distance passenger flows data. To address the need, the Federal Highway Administration (FHWA) initiates a collective of research efforts under the theme of establishing a Traveler Analysis Framework, which is analogous with the Freight Analysis Framework (FAF) in the freight operations and management field. The initiative does not include a new survey but explores the possibility of extrapolation from existing data resources. One of the major efforts within the initiative is the establishment of Long Distance Multimodal Passenger Travel (LDMPT) OD data program for 2008 and 2040. This program will comprehensively summarize and compile long distance travel data from diverse sources and greatly expand the availability of such data.

An expert panel from the LDMPT project also identified the need to establish a highway network modeling framework that utilizes the most updated traffic data for researching network loading and distribution for both passenger and freight movements. Part of this research's motivations resonates to the panel discussion findings to develop a national data framework that takes traffic data, congestion data, and roadway capacity information into account for long distance travel assignment.

## 1.3 Research Objectives

The ultimate objective of this study is to explore the importance and feasibility of establishing a national data framework for long distance highway travel modeling, ***with a passenger travel***

***focus***; this study also aims to bring new knowledge about multi-scale network representation and long distance network loading.

Specific objectives of this study are as follows:

- Identify existing datasets regarding the long distance travel through a comprehensive literature review;
- Identify and resolve the challenges of describing and modeling large-scale network in multiple resolutions;
- Expand the existing traffic assignment theory to a long distance travel context, taking account of data and network reality;
- Demonstrate the feasibility of the proposed methodology in a case study.

Such a modeling framework will be very valuable for long distance network loading issues at national scale and hence leverage benefits in both short and long terms. In the short term, the framework and related theories developed in this research will enhance academic understanding and field practice in the following aspects:

- Catalog existing datasets regarding the long distance travel, particularly passenger travel;
- Design a complete procedure elaborating the conceptual design of a modeling framework that embraces multi-scale views of the physical infrastructure in a geodatabase, incorporates the most updated traffic data and other travel impedance information, and also resolves computational needs for network loading and trip assignment across the network;

- Develop a “working solution” for estimating the distributive pattern of long distance passenger travel flows, as well as system-wide indicators and performance measures to facilitate policy and investment decision making.

In the long term, research efforts committed by this study will hopefully radiate the benefits beyond the long distance travel context to other related areas including but not limited to large-scale system reliability evaluation, environmental analysis, and energy preservation. Because this study will lay a solid foundation for those related future research and applications as a national data framework, which not only interoperates various datasets in the geodatabase depicting the large-scale national highway network, but also enables the computational capabilities, to support project appraisal and decision making for highway related investment and policies at a national level.

### 1.4 Research Scope

This research focuses on long distance passenger travel in the mode of automobiles. The network study is national highways across 48 continental states and Washington D.C. in the U.S. The modeling framework focuses on a national (or at least interregional) level system. While considerable network analyses work continues to be done at local level, this study is more experimental than advanced, with primary concerns of system-wide indicators and performance measures to foster an understanding of long distance travel patterns and trends as a national scale. For example, this study endeavors to reveal the distributive patterns and total Vehicle Hours Traveled (VHT) at regional or national level, while an MPO may be interested in the travel time on a specific urban corridor.

### 1.5 Dissertation Organization

To fulfill the research objectives, a four-step work flow is proposed: literature review and data identification, network representation in geodatabase, methodology development for long distance travel network loading, and case study of long distance travel network loading. The remainder of this dissertation is organized in the order of this work flow. Chapter 2 identifies relevant datasets and summarizes relevant research on long distance travel data collection, methodologies for quantifying travel impedance, approaches for long distance traffic assignment, and frameworks for traffic data archival and analyses. Chapter 3 focuses on the multi-scale network representation in geodatabases and elaborates the network model design following a “conceptual – logical – physical” sequence, with a highlight of linking traffic data to the network. Chapter 4 starts with a tentative work flow clarifying each step in the process of long distance travel network loading, and then elaborates the theory and algorithms to be applied. After investigating different attributes within various long distance traveling populations, Chapter 4 develops impedance models for network loading in a long distance travel context. Chapter 5 details a case study where the whole modeling framework is practiced, and summarizes findings from the case study. Finally, Chapter 6 concludes the dissertation.

# Chapter 2 Literature Review and Data Identification

This chapter comprehensively summarized the state of art in establishing a national data framework. It is recognized that the efforts are based on review of the five following respects: ***infrastructure***, ***information***, ***externalities***, ***computation***, and ***implementation***. ***Infrastructure*** gives an overview of the national highway system, whose abstraction will serve as the base network in the framework. ***Information*** is a variety of datasets reflecting the volume to be loaded onto the network (demand) as well as the operational status of the network (supply). ***Externalities*** refer to the travel impedance, embracing recurrent and non-recurrent congestion that influences traffic assignment among alternative routes. ***Computation*** involves a series of methodologies to assign the demand to the roadway network, constrained by externalities. Lastly, ***implementation*** incorporates the aforementioned aspects through reviewing some established large-scale geodatabases and regional traffic data platform.

## 2.1 Overview of Our National Highway System

Accessible, robust, and efficient highway network is essential to a nation's economy, defense, and public welfare, especially for a highly motorized country like the United States, with 828 vehicles in operation per 1,000 people (Davis *et al.*, 2011) and 61.6% of total tonnage of freight moved by trucks (FHWA, 2010). There are over four million miles of public roads in the United States (FHWA, 2010), which are generally categorized into local roads, collectors, and arterials on a functional basis. While arterials and local roads connect urban centers and small communities, respectively, collectors provide access from local roads to arterials. Of all the public roads, more than 163,000 miles belong to the National Highway System (NHS), which serves as the backbone of the nation's transportation network.

The Dwight D. Eisenhower National System of Interstate and Defense Highways is the major component of the NHS, consisting of 30,600 miles of rural highways and 16,000 miles in urban areas. The rest of the 82,300 miles of rural and 34,100 miles of urban highways are in the following four subsystems that include the Strategic Highway Network (or STRAHNET, which includes other important highways serving for defense purposes), major strategic highway network connectors (highways that channels traffic from major military installations to STRAHNET highways), other principle arterials (rural and urban highways connecting arterials and major transportation facilities and hubs), and intermodal connectors (highways providing access between major intermodal facilities and the other four subsystems of the NHS).

On the freight front, the establishment of the National Network for large trucks was authorized by the Surface Transportation Assistance Act in 1982. The network has a length of more than 200,000 miles of Interstate and other highways (FHWA, 2010).

## 2.2 Available Datasets for a National Long Distance Travel Framework

### *2.2.1 Data requirements, concerns, and challenges*

A national long distance travel data framework integrates a diversity of data sources; meanwhile the advancing of survey science and detection technologies keeps generating new data exponentially. Therefore, before reviewing relevant data sources, it is beneficial to make scoping efforts to clarify what types of data are needed for developing such a data framework. Roughly the target data can be categorized in: ***geography data***, ***demand data***, ***externality data***, and ***validation data***. ***Geography data*** include both linear data providing basic geometry of the nation's highways, and zonal data indicating major origins and destinations. ***Demand data*** are the outputs from the first three steps in the classic four-step transportation planning model, represented by the OD table with the estimated numbers of trips for each OD pair. ***Externality data*** reflect the detected or estimated volume traveling on the highway network, which to large extent determines travel impedance and influence mode and route choice. ***Validation data*** are those data for local validation purposes, which rely on temporary data collection techniques.

Some concerns and challenges emerge given the current national data reality. First, a national multi-scale zoning system is in absence. The current Metropolitan Planning Organization (MPO) Traffic Analysis Zones (TAZs) and Regional Model TAZs are too small for Multistate Highway Corridor Models, Long-Distance/Tourist Trips, or even for Statewide Freight Models (Cambridge Systematics, Inc., 2008). A balance must be achieved whereby model geography is appropriate for a multi-scale data framework and the number of OD matrix cells at the most detailed level is still manageable. Second, there is still ambiguity in the existing modes data. For the highway sector, major modes include motorcycles, passenger cars, light trucks, buses, single-unit trucks, and combination-unit trucks. Currently there is no data collection effort to differentiate passenger travel from freight movement for the small-size trucks. Third, trip assignment on a national network is a multi-class problem; because both passenger and freight travels use the same networks (Gudzinas, 2011).

### 2.2.2 *Available datasets*

#### 2.2.2.1 Geography Data

- National Highway Planning Network

The National Highway Planning Network (NHPN) is a 1:100,000 scale network database that contains line features representing just over 450,000 miles of the current and planned highways in the U.S. NHPN consists of interstates, principal arterials, and rural minor arterials and serves as a geospatial basis for many other critical national transportation databases, including Highway Performance Monitoring System (HPMS) and Freight Analysis Framework 3 ($FAF^3$).

In NHPN, roadway functional classification is coded in seven categories as summarized in Table 2.1, following the convention set up in the Functional Classification Guidance Document (FHWA, 1989).

Table 2.1 Roadway Functional Classification Coded in NHPN

| Code | Description |
|---|---|
| 1 | Rural Principal Arterial - Interstate |
| 2 | Rural Principal Arterial - Other |
| 6 | Rural Minor Arterial |
| 7 | Rural Major Collector |
| 8 | Rural Minor Collector |
| 9 | Rural Local |
| 11 | Urban Principal Arterial - Interstate |
| 12 | Urban Principal Arterial - Other Freeways and Expressways |

| | |
|---|---|
| 14 | Urban Principal Arterial - Other |
| 16 | Urban Minor Arterial |
| 17 | Urban Collector |
| 19 | Urban Local |

- Core Based Statistical Areas

Metropolitan and micropolitan Statistical Areas (MSA, referred generically as "Core Based Statistical Areas" or CBSAs) are statistical geographic areas defined by the Federal Office of Management and Budget and maintained by the Bureau of the Census, 2003. Metropolitan statistical areas contain at least one U.S. Census Bureau-defined urbanized area of a population 50,000 or more; micropolitan statistical areas contain at least one Census Bureau-defined urban cluster of at least 10,000 and less than 50,000. An MSA may span over several counties and is a good geographical representative for populated urban areas.

#### 2.2.2.2 Demand Data

- The 1995 American Travel Survey

The 1995 ATS collected cross-sectional and longitudinal estimates on the *origin, destination, volume*, and *characteristics of long-distance travel* in the United States. A total of 80,000 households nationwide were randomly selected to participate in the survey. The survey consisted of four detailed interviews conducted approximately every three months from April 1995 to March 1996.

The 1995 ATS adopted the convention of *census region* by the U.S. Census Bureau and MSA

(1995 ATS Technical Documentation). The 1995 ATS collected more than half a million person trips from 163 MSAs (U.S. Department of Transportation, 1999). The database incorporate both the number of trips between each pair of MSAs and the mode split.

As the most conclusive long distance travel database so far, the 1995 ATS also suffers from several shortcomings: first, some large travel volumes between certain OD pairs exceed with the physical distance from origin to destination; second, some suspiciously higher volumes are reported from lower population density area; third, "*significant differences are observed in the modal splits between the two directional traffic flows for the same OD pair, inherently claiming that trip distance has insignificant influence on transportation mode choices for those OD pairs.*" (Lim, 2008)

- FHWA Traveler Analysis Framework

Due to the massive price tag of new national travel surveys and imperative needs for long distance passenger travel OD data, FHWA initiated the taskforce to develop a methodology to produce 2008 passenger OD data by mode (air, rail, and auto). The major contractor Wilbur Smith Associates team applies growth factor methods to the 1995 ATS data for developing the 2008 auto trip. "*Data from the 2001 National Household Travel Survey (NHTS) will also be used to determine if travel trends hold between 1995 and 2001 and to grow the table to 2008.*" (Wilbur Smith Associates, 2011)

- Existing Long Distance Travel Demand Forecast Model

There have been extensive literatures dedicated to extending the well-developed urban travel demand forecast to a statewide or nationwide context (Liu, *et al.*, 2006; Ashiabor, 2007; Henderson and Trani, 2008; Souleyrette, *et al.*, 1996; Giaimo and Schiffer 2005; Horowitz, 2006; Cohen, *et al.*, 2008, Horowitz, 2008). Some proposed methodologies could be applied to fill the gap or void in the current OD table (Monzon and Rodriguez-Dapena, 2006) or to overcome the shortcomings of the current travel demand forecast techniques. Also, notably many of those methods/algorithms are based on manipulation of some major national databases including the 1995 ATS or 2001 NHTS (LaModia and Bhat, 2007). Additionally, stated preference surveys are frequently applied in search of travel demand impact factors and mode choice (Peeta *et al.*, 2007 and 2008; Ashiabor *et al*., 2007; Srinivasan, *et al.*, 2006; Liu and Li, 2005).

#### 2.2.2.3 Externality data

- Highway Performance Monitoring System (HPMS)

Administered by the FHWA, HPMS is supported by collaborative efforts from state transportation agencies and MPOs and serves as a national data source that provides traffic volumes on the national highways with broad coverage. There are three major components in the HPMS database: 1) the *full extend data* are collected (or estimated) section by section along the entire NHS, some of the attributes include AADT of general traffic, AADT of single-unit trucks, AADT of combination trucks, section length, etc.; 2) the *sample panel data* provide statewide statistics of the highway network, based on randomly selected samples; and 3) the *summary data* reveal information of rural minor collector and local roads (both urban and local) in aggregated form.

- Freight Analysis Framework 3 ($FAF^3$)

$FAF^3$ integrates data "*from a variety of sources to create a comprehensive picture of freight movement among states and major metropolitan areas by all modes of transportation*" (Oak Ridge National Laboratory, 2012). The data set covers the 48 contiguous States plus the District of Columbia, Alaska, and Hawaii. $FAF^3$ processes a richness of traffic volume data, including ***Year 2007 Truck Volume*** estimated using a combination of HPMS 2008 database, State truck percentage, and functional class specific defaults, ***FAF 3.1 long distance truck volume*** estimated based on the FAF 3.1 Origin-Destination truck tonnage and includes empty trucks, ***Link specific peak capacity*** estimated using the procedures outlined in HCM 2000 and the arc geometry provided in 2008 HPMS database, and ***Estimated service flow*** using the procedures outlined in HCM 2000 and arc geometry, FAF truck, non-FAF truck and passenger volume.

- Traffic Flow Sensor Data

Counting traffic started from the 1920s in the United States. Initially people collected travel data for transportation system monitoring purposes. With intensive development of new technologies and computer networks, Intelligent Transportation Systems (ITS) were initiated and have been under development since 1980s. ITS significantly refreshed the concept and practice in travel data collection. Data collected by various types of ITS sensors has accumulated significantly over the past three decades. Traffic flow sensor data cover basic information including traffic volume, speed, and occupancy, serving as substantial piece of externality information. Widely deployed detection techniques are summarized in Table 2.2.

Table 2.2 Summary of Traffic Flow Sensor Characteristics (excerpted from Klein, 2001)

| Technology | Available Data | Advantages | Disadvantages |
|---|---|---|---|
| Inductive Loop Detector | • Count<br>• Presence<br>• Occupancy<br>• Average vehicle speed with data processing algorithm or two ILDs<br>• Queue length with multiple sensors | • Standardization of loop electronics units<br>• Excellent counting accuracy<br>• Mature, well understood technology<br>• Some models provide classification data | • Not suitable for bridges, over passes, viaducts, poor roadbeds<br>• Reliability and useful life are dependent on installation procedures<br>• Installation and maintenance require lane closure<br>• Decreases life of pavement<br>• Susceptible to damage by heavy vehicles, road repair, and utilities<br>• Multiple sensors usually required at a site |
| Magnetometer | • Count<br>• Presence<br>• Occupancy<br>• Speed with multiple sensors or knowledge of detection zone length and vehicle length | • Less susceptible than loops to stresses of traffic<br>• Detects stopped and moving vehicles<br>• Some models transmit data over wireless RF link | • Small detection zone<br>• Installation requires pavement cut<br>• Installation and maintenance require lane closure<br>• Decreases pavement life<br>• Multiple sensors usually required at a site |
| Magnetic | • Count<br>• Occupancy<br>• Speed with multiple sensors or knowledge of detection zone length and vehicle length | • Can be used where loops are not feasible (e.g., bridge decks)<br>• Some models installed under roadway without need for pavement cuts<br>• Less susceptible than loops to stresses of traffic | • Small detection zone<br>• Installation requires pavement cut or tunneling under roadway<br>• Cannot detect stopped vehicles (exception for 1 model using multiple sensors and application-specific software from vendor) |
| Video Image Processor (Machine Vision Processor) | • Count<br>• Presence<br>• Occupancy<br>• Speed<br>• Queue length<br>• Instantaneous vehicle density<br>• Incident evaluation<br>• Turning movements | • Single camera and processor can service multiple lanes and multiple zones/lane<br>• Rich array of traffic data provided<br>• Easy to add and modify detection zones | • Large vehicles can mask smaller vehicles, leading to undercounting<br>• Tall vehicles can project their image into adjacent lanes, leading to over counting<br>• Shadows, reflections from wet pavement, vehicle-to-road contrast, headlight projection into adjacent lanes on curved road sections, sun glint, day/night transitions, camera vibration, and debris on camera lens can affect performance<br>• Side mount requires high [50+ ft (15+ m)], stable camera mounting platform for most accurate data |

| | | | |
|---|---|---|---|
| | • Class by vehicle length | | • Over-roadway camera mounting requires lane closure for installation and maintenance<br>• Reliable nighttime signal actuation requires street lighting<br>• Weather, but effects ameliorated by recall modes |
| Microwave–Presence Detecting | • Count<br>• Presence<br>• Occupancy<br>• Speed<br>• Range<br>• Instantaneous vehicle density<br>• Class by vehicle length | • Good performance in inclement weather<br>• Detects stopped vehicles<br>• Can operate in side-looking mode to service multiple lanes | • Vehicle occlusion may occur in distant lanes in side-looking mode when congestion is heavy<br>• Vehicles undercounted more in heavy congestion<br>• Offset mounting distance must be accommodated |
| Microwave–Doppler | • Count<br>• Occupancy<br>• Speed | • Good performance in inclement weather<br>• Direct measurement of speed | • Cannot detect stopped or very slow-moving vehicles |
| Active Infrared (Laser Radar) | • Count<br>• Presence<br>• Occupancy<br>• Speed<br>• Range<br>• Classification | • Direct measurement of speed<br>• Provides vehicle classification data | • Performance degradation by heavy fog [visibility < ≈20 ft (6 m)] and blowing snow<br>• Installation and maintenance require lane closure |
| Passive Infrared | • Count<br>• Presence<br>• Occupancy<br>• Speed with multi-zone sensor | • Compact size, ease of installation | • Performance possibly degraded by heavy rain, fog, overcast skies, or snow |
| Acoustic | • Count<br>• Presence<br>• Occupancy<br>• Speed | • Insensitive to precipitation<br>• Services multiple lanes | • May under count in congested flow<br>• Cold temperature has been reported as affecting data accuracy |

#### 2.2.2.4 Validation Data

Some newly developed technologies can meet temporary data collection needs for validation. Representatives of these technologies are:

- **Bluetooth Traffic Monitoring**

  Bluetooth Traffic Monitoring is based on the Media Access Control (MAC) address matching. Every Bluetooth or WIFI enabled device is associated with a unique 48-bit MAC address for data communications. A vehicle carrying a Bluetooth device under the discoverable mode can be observed by Bluetooth readers in its detectable range. If the MAC address and time of detection is logged, a sample travel time for a roadway segment can be extracted by comparing the timestamps of each MAC read at both ends. Observations of multiple vehicles containing Bluetooth devices provide a highly accurate estimate of traffic conditions. Experiments have indicated that approximately one in ten vehicles contains a Bluetooth device that can be detected and roughly half of the detected vehicles can be matched to obtain travel time information (Malinovskiy *et al*., 2010; Wang *et al*., 2010; Young, 2008; Sawant *et al*., 2004).

- **Probe Vehicles with GPS Devices**

  Probe vehicles equipped with Global Positioning System (GPS) receivers can be used to collect travel trajectory and travel time information. The challenge lies in that GPS devices output latitude and longitude pairs. Therefore an Linear Referencing System (LRS) is needed as the base to organize and analyze GPS data. An ongoing trend is to combine the floating vehicle information with the GIS interface (Tong, *et al*., 2005; Schäfer *et al*., 2002; Quiroga and Bullock, 1998).

## 2.3 Travel Impedance in Long Distance Travel

### *2.3.1 Definition of travel impedance*

Travel impedance is also known as friction factor, applied in the famous and most commonly used *gravity model* (Hutchinson, 1974; Haynes and Fotheringham, 1984) for trip distribution. It is an inverse function of travel cost between zones *i* and *j*, for example, travel time, distance, fuel cost, or any combination of the above. This friction factor needs to be calibrated through comparing the observed and predicted trip length distributions (Meyer and Miller, 2001; Levinson, 1998).

In trip assignment, travel impedance quantifies the travel cost when using specific routes and significantly influence drivers' decision making. Previous research has identified a series of impact factors, such as the O/D distance, congestion level, etc. Notably, most of those factors are indeed associated with travel time. Ewing *et al*. (2004) used the inverse of travel time from the trip generation zone to the attraction zone as the travel impedance in a school zone travel mode choice study. Similar manipulation can also be found in Black (1972) and Levinson (1993). Alternatively, travel impedance might take the form as a function of travel time with a negative exponent (Park and Smith, 1997).

As mentioned in Section 1.3, for recreational long distance travel, travel time is not the preliminary travel impedance. Other factors will be explored in this research. However the following literature review mainly investigates travel time as the major impedance measure, and reviews the travel time estimation for nodes and links separately.

### *2.3.2 Travel impedance in urban areas*

Travel impedance estimation at a junction requires an integrated measure over the whole urban area. Several congestion indices included in the annual *Urban Mobility Report* published by Texas Transportation Institute (TTI) provide critical reference on this matter. TTI develops methodologies to generate several key measures, e.g. travel time index, delay per auto commuter, cost per auto commuter, total delay, and total cost), which are categorized as “intensity” and “magnitude” measures. Collectively those indices can be used to interpret and quantify the area-wide congestion and mobility levels. They also allow for comparisons across similar urban areas. The ranking is population based, where urban areas are classified into very large (over 3 million), large (over 1 million but less than 3 million), medium (over 500,000 but less than 1 million), and small (less than 500, 000) categories. Some key congestion measures are listed below:

- **Travel Delay**: the total amount of extra travel time due to congestion. Most of the key measures presented in the Urban Mobility Report were developed based on calculating travel delay. For example,

$$\text{Daily Vehicle Hours of Delay} = \frac{Daily\ Vehicle\ Miles\ of\ Travel}{Speed} - \frac{DailyVehicle\ Miles\ of\ Travel}{FreeFlow\ Speed}$$

- **Annual Persons Delay**: the yearly amount of extra travel time for freeways and arterial streets in each study area.

$$\text{Annual Persons Hours of Delay} = \text{Daily Vehicle Hours of Delay on freeways and Arterial Streets} \times \text{Annual Conversion Factor(50 working weeks per year)} \times \text{Average Vehicle Occupancy(1.25 Persons per Vehicle)}$$

- **Travel Time Index**: the average amount of extra time spent on travel relative to free-flow travel.

$$\text{Travel Time Index} = \frac{Peak\ Travel\ Time}{FreeFlow\ Travel\ Time}$$

$$\text{Travel Time Index} = \frac{Delay\ Time + FreeFlow\ Travel\ Time}{FreeFlow\ Travel\ Time}$$

- **Annual Delay per Auto Commuter**: a yearly sum of the extra travel time for those commuters who travel during the peak period. All of the delay that occurs during the peak period is assigned to the auto commuters. Furthermore, since commuters also contribute to the delay occurs outside of the peak period, the off-peak delay should consequently be considered and is assigned to the entire population of the urban area. This measure illustrates the effect of the per-mile congestion as well as the length of each trip.

$$\text{Annual Delay per Auto Commuter} = \frac{Peak\ Period\ Delay}{the\ number\ of\ auto\ commuters} + \frac{Remaining\ Delay}{Population}$$

- **Cost of Congestion**: the value of extra time and fuel consumed during the congestion period.

### *2.3.3 Travel impedance along roadways*

Travel time estimation for roadway links is an area with extensive research because it is an essential piece of information for ITS applications and route travel time estimates. In freight movements, resolving the issue is straightforward due to the wide use of GPS devices in freight vehicles (Cambridge Systematics, Inc., 2008). For passenger traffic, however, travel time estimation either relies on the Bureau of Public Roads (BPR) functions and its variations (summarized in Table 3.3) due to its simplicity or on the travel speed estimation.

There is abundant literature related to speed estimation from single loops, the primary detection means by most of transportation agencies (Klein *et al.*, 2005). These research efforts mainly include: (1) advance detectors based approach (Cheung *et al.*, 2005; Haoui *et al.*,2008; Varaiya, 2004; López-Valcarce *et al*., 2004);  (2) volume and occupancy data-based methods (Athol, 1965; Pushkar *et al*., 1994; Dailey, 1999; Wang and Nihan, 2000 and 2003; Coifman *et al*., 2001 and 2003); and  (3)  Analog signal-based methods (Sun and Ritchie, 1999; Oh *et al*., 2002; Fang *et al*., 2007).

In addition to point speed estimation, other researchers have developed several approaches for link speed estimation (or space mean speed), with which link travel time can be directly computed. Some have proposed extrapolation methods/trajectory methods to estimate average speeds between two point measurements (Quiroga 2000; Dhulipala 2002; Cortes *et al*. 2002; Van Lint and van der Zijpp 2003; Lindveld *et al*. 2000; Wu *et al*., 2011). Others have built up statistical black box models with the link speed as the output and sensor measurements as inputs. Various statistical inference techniques were also developed for speed estimation and prediction based on Kalman filters (Wang and Papageorgiou, 2005), auto regressive moving average time series models (Van Arem *et al*., 1997), artificial neural networks (Blue *et al*., 1994), and fuzzy logic algorithms (Palacharla and Nelson 1999). Some research efforts (Bovy and Thijs, 2000;

Petty *et al*.,1998; Hoogendoorn, 2000; and Nam and Drew, 1998) were made to develop speed estimation methods based on traffic flow theory. However, very little research was found to address the methods to estimate network level speed data. Most researchers still break down the network into different links. For example, Lindveld et al. (2000) aggregated the link level speed estimations to obtain route travel data. Clark and Watling (2005) took a planning perspective and gave a distribution of network travel time without real-time detection involvement. Hence, it is desirable to develop techniques combining point or link estimations and generating network level speed data are needed to monitor traffic performance at the network level.

Table 2.3 Summary of BPR Function Variations

| Function | Inputs | Outputs | Notes | Reference | Limitations |
|---|---|---|---|---|---|
| $tt_j = \frac{2 \times (x_{i+1} - x_i)}{V_L(x_i, t_i) + V_L(x_{i+1}, t_i)}$<br>$v = \frac{\sum_{i=1}^{L}(q_i \times v_i)}{\sum_{j=1}^{L} q_j}$ | Time and milepost, number of cars. Observed by detectors. | Travel Time, time mean speed. | Assumes that the speed is constant | Liu *et al.*, 2009 | Actually the speed is not constant, and the time mean speed only contains local info. |
| $tt_j = \frac{2 \times (x_{i+1} - x_i)}{V_M(x_i, t_i) + V_M(x_{i+1}, t_i)}$<br>$V_L = \frac{\sigma_M^2}{V_M} + V_M$<br>$\sigma_M^2 \approx \frac{1}{2N}\sum_{i=1}^{n} \beta_i (v_{i+1} - v_i)^2$<br>$\beta_i = \frac{1}{v_i}(\frac{1}{N}\sum_{j=1}^{N}\frac{1}{v_j})^{-1}$ | Time, milepost, Observed by detectors. Number of subsequent observations, weighting factor. | Travel time, Calibrated mean speed. | Still assumes that the speed is constant, but the speed is calibrated.<br>N denotes the number of subsequent observations, and $\beta_i$ is a weighting factor for each observation. | | Actually the speed is not constant. |
| $tt_j = \frac{2 \times (x_{i+1} - x_i)}{V_M(x_i, t_i) + V_M(x_{i+1}, t_i)}$<br>$V_M = 2(\frac{1}{V_L(x_i, t_i)} + \frac{1}{V_L(x_{i+1}, t_i)})^{-1}$ | Time, milepost. Observed by detectors. | Travel time, Calibrated mean speed. | Just to simplify the equation above. | | Actually the speed is not constant. |
| $v_j(t) = v(x_i, t_i) + \frac{x_j(t) - x_i}{x_{i+1} - x_i}(v(x_{i+1}, t_i) - (v(x_i + t_i))$<br>$t(t) = \int_0^{L_i} \frac{d_x}{v_j(t)} = \int_0^{L_i} \frac{d_x}{v_i + \frac{x}{L}(v_{i+1} - v_i)} = \begin{cases} \frac{L_i}{v_i}, & v_{i+1} = v_i \\ \frac{L_i}{v_{i+1} - v_i} \ln \frac{v_{i+1}}{v_i}, & v_{i+1} \neq v_i \end{cases}$ | Time, milepost. Observed by detectors. | The mean speed of the section between two detectors, Travel time. | Assumes that the speed is not constant, but the speed of the section between of two detectors is constant. | | The speed within a section might not be constant. |
| $\overline{T}(t) = \frac{1}{\omega}\sum_{i}^{\omega} T(i, t)$ | Times, Number of cars. | Travel time. | The travel time can be get by probe or historical data | Wu et al., 2004 | There might not be enough historical data. |

| | | | | | |
|---|---|---|---|---|---|
| $f^c(x) = 2 + \sqrt{\alpha^2(1-x)^2 + \beta^2} - \alpha(1-x) - \beta$<br>t=$t_0$*$f^c$(x)<br>x=v/c , $\beta = \frac{2\alpha-1}{2\alpha-2}$ Conic function | Volume, capacity, α | Travel time | This function used v/c to estimate the travel time, and it might overcome some known disadvantages<br>Of BPR function.<br>It is more computational efficient than the BPR. | Spiess, 1990 | |
| $T = T_0\left(1 + \alpha * (\left(\frac{V}{C} - \sigma\right)^{\beta} + \sigma^{\beta}\right) + \varepsilon * \frac{V}{C} + (V > C) * (V - C) * \gamma$ | Volume, capacity, parameters | Travel time | Allow to assign bigger trip matrices without over-assignment | Jastrzebski, 2000 | Needs more iteration to reach equilibrium state in the network |
| $LS(V) = T_0 + \frac{(T_s - T_0)}{1 + e^{\tau(1-\frac{V}{C})}}$<br>The Logit S-curve function | Free flow delay, Delay for saturated flow, volume, capacity, parameter | Travel time | It assumes that there is a constant delay for saturated flow, so the time spent on a link in the network is never infinite. | Babin and Tremblay,1994 | The delay for saturated flow is not always constant.<br>We have to assume a minimum τ to ensure LS(0)≈$T_0$ |
| $PS(V) = T_0 + [\frac{(T_s - T_0)}{1 + \left(\frac{V}{C}\right)^{-\tau}}]$ | Free flow delay, Delay for saturated flow, volume, capacity, parameter | Travel time | The characteristics are similar with the Logit S-curve, and LS(0)=$T_0$. | | The τ parameter still should be greater than 4~5 in order to obtain a correct “S” shape. |

## 2.4 Traffic Assignment Algorithms

Traffic assignment procedure reflects the interactions between demand and supply, where demand represents trip volume, origin, destination, mode, and routes, with various underlying behavioral assumptions on route choice (e.g. user equilibrium/system optimization); and the supply side is reflected by the available capacity for serving the demand which determines how the transportation network operates.

### *2.4.1 Static traffic assignment*

#### 2.4.1.1 <u>All or Nothing Assignment</u>

In this approach, minimum travel time (ideal and uncongested) paths are computed for each OD pair, and all the flows of these pairs are loaded onto the routes. A given route receives "all or nothing" of the flow for a given OD pair. While simple and inexpensive to use, and easy to interpret, this approach is clearly unrealistic in situations where capacity constraints and congestion effects do exist.

#### 2.4.1.2 <u>Equilibrium Assignment</u>

This approach explicitly recognizes the transportation network link costs are generally influenced by the volume using that link. The two most widely accepted underlying behavioral assumptions are user equilibrium (UE) and system optimization (SO). In a UE network, no user can improve his/her travel time (cost) by unilaterally changing routes (Wardrop, 1952). In a SO network, the system users would be assigned to routes so as to minimize the system average travel costs. The bright side regarding the UE assignment procedures is they are capable of handling large, real-world network, and they are routinely available within most commercially available transportation modeling software packages.

#### 2.4.1.3 Stochastic Assignment

The equilibrium assignment methods briefly sketched in the preceding discussion are more properly referred to as deterministic user equilibrium (DUE) methods, since they assume that all users in the system have perfect information about the travel times on alternatives routes within the network and that they make perfectly correct route choices based on this information. Obviously this is inconsistent with the reality, where randomness and misinformation exist. Realizing that, a set of solutions have been proposed with travel costs being random variables that can vary among individuals based on their individual preferences, experiences, and perceptions, rather than deterministically as in the DUE framework. Suggested procedures include application of an incremental assignment in a stochastic simulation procedure (Burrell, 1968), employ of a multinomial logit model to predict route choice probabilities (Dial, 1971), and use of discreet choice models within a stochastic user-equilibrium (SUE) framework (Sheffi, 1985).

In practice, both DUE and SUE are categorized as Static Traffic Assignment (STA) methods, which are widely used to determine important infrastructure investment decisions, because of its well-defined user behavior principle, elegant mathematical formulation, and efficient algorithms.

### *2.4.2 Dynamic traffic assignment*

For many regional transportation planning applications, the static assignment assumption is acceptable and, with a properly validated network, can yield very useful results. For many other applications (ATIS, Route guidance, ramp metering, dynamic traffic signals, managed lanes, dynamic pricing etc.), however, the static representation of network performance is not sufficiently accurate. Because STA assumes that the flows of vehicles or people enter the network "simultaneously". This is obviously an unrealistic assumption because each vehicle can be on only one link at a time and each vehicle must travel through times as well as space as it moves from its origin to its destination.

In such cases, a dynamic representation of route choice behavior and resulting network performance (congestion, speeds, etc.) is required to explicitly describe movements of vehicles along their chosen paths in a timely manner. Dynamic Traffic Assignment (DTA) intends to address this problem by considering ***time varying flows***. Spatially, it models traffic flow phenomena such as queuing, spillovers, shockwaves, etc.; temporally, it includes temporal choice dimensions, such as departure time, schedule delay, etc. Unlike STA, DTA does not have a single, universally accepted formulation, or well-established solution properties, and it is difficult to solve mathematically, although tremendous efforts have been invested in developing efficient algorithms for DTA.

Methodology-wise, DTA research has progressed in two major directions: analytical models (Friesz *et al*., 1993; Ukkusuri and Waller, 2008) and simulation models, where a selective list includes DynaSMART (Mahmassani, 2001), DynaMIT (Ben-Akiva *et al*., 2002), VISTA (Ziliaskopoulos and Waller, 2000), and Dynameq (Mahut *et al*., 2005). While the former gives strong mathematical proofs, the latter reflects more of the traffic flow evolvement resulted from vehicle interactions. Both kinds of research are based on certain assumptions. Content-wise contemporary researchers keep pushing the limits to incorporate more travelling issues in the DTA framework, including departure time (Friesz, *et al*., 1993; Ran *et al*., 1996; Huang and Lam, 2002; Wie *et al*., 2002;  Szeto and Lo, 2004; Zhang and Zhang, 2007), activities/locations choices (Abdelghany, 2001 and 2003; Lam and Huang, 2003; Kim *et al*., 2006; Rieser *et al*., 2007), and activity duration (Ramadurai and Ukkusuri, 2008).  Table 2.4 illustrates a breakdown of DTA models.

Table 2.4 Breakdown of DTA Models

| | | **Computational Methodology** | | | |
|---|---|---|---|---|---|
| | | **Mathematical Programming** | **Optimal Control** | **Variational Inequality /Congestion Pricing** | **Algebraic/Graphical** |
| Traffic Flow Model | Physical Queue | Ukkusuri and Waller, 2008; Ziliaskopoulos. 2000 | | Lo and Szeto. 2002 | |
| | Point Queue | | | | Vickrey, 1969; Muñoz and Laval, 2005; |
| | Link Exit Function | Merchant and Nemhauser, 1978; Janson, 1991 | Friesz, *et al*., 1989; Ran and Shimazaki, 1989; Ran *et al*., 1993 | Ban et al., 2008; Friesz, *et al*., 1993; Wie, *et al*., 1995 | |

### *2.4.3 Applications of traffic assignment*

Although hot in research as discussed above, only few applications were found in applying DTA to a large-scale network with support of regional travel demand models. Dynameq has been applied to a large subarea of Calgary and to analyses of the Rue Notre-Dame in Montréal (Donnelly *et al*., 2010). Yet results from the work are currently unpublished and inaccessible.

The largest known DTA application to date is described by Hicks (2008). The network from the Atlanta Regional Commission's (ARC's) travel model served as the base for the DTA network. In this application trip matrices from the ARC model were divided into 15-min intervals for the specification of demand. The drawback of this practice is that it is very time-consuming.

A number of cities are currently testing DTA models, but the progress is still far away from even preliminary results. At least a dozen such cases are known to be in varying stages of planning or execution, suggesting that the use of DTA models in planning applications is about to expand dramatically. However, there is still a series of issues regarding applying DTA in an urban context, not to mention its application in a long-distance travel context. An attempt in applying DTA to long distance travel is I-285, a 64-mile circumferential freeway, where DTA serves as the mesoscopic layer of the analysis, connecting the demand matrices (divided into 15-min intervals) and the microscopic scale models (VISSIM simulations) (Simons, 2006). As for state models, in many aspects the travel models used at the state level are simple extensions of traditional sequential models (four-step) due to the cost of increased data requirements and computational burden of DTA.

## 2.5 Large-Scale Geodatabase and Regional Traffic Data Platform

Currently there is a void of knowledge regarding national data model for long distance travel, however, similar efforts have been committed from regional to continental scales. For example, Lin *et al*. (2008) mapped GPS recorded travel speed data from 27,471 trucks onto the entire U.S. highway network. Haider and Spurr (2006) devised the operational details of a large-scale (metropolitan-wide) traffic assignment model using TransCAD software, which comprises 240,000 unidirectional links for the Greater Montreal Area, Canada. Of these studies two are selected and summarized in the following table, one for the southwest Georgia region (Georgia Department of Transportation, 2009) and the other for the whole European Union (Nielsen and Burgess, 2008).

Table 2.5 Two Large-Scale Transportation Network Models

| | The European TRANS-TOOLS Transport Model | Southwest Georgia Interstate Study |
|---|---|---|
| Geographic Coverage | All European Union member countries | 32 counties in Southwest Georgia study area region, with the rest of Georgia State as buffer area and other states of continental US |
| Modeling Objective | Both passenger and freight | Both passenger and freight |
| Modes | For freight transport:<br>• *Trucks*<br>• *Rail*<br>• *Inland waterways*<br>• *Ships*<br>For passenger transport:<br>• *Cars*<br>• *Rail*<br>• *Air* | Automobiles |
| Primary Network Data Source | Zonal database:<br>• *NUTSII database*<br>• *Russia and Turkey as one-country zones*<br>Roadway database:<br>• *CISCO version II Ten_STAC* | National Highway Planning Network (NHPN)<br>Georgia DOT Road Characteristic (RC) file |
| Model Resolution | 1270 European zones and 16 port-zones<br>Roadway network:<br>• *35,079 nodes* | Five-level network in varying resolutions<br>Study area (32 counties) |

| | | |
|---|---|---|
| | • *47,373 links*<br>• *1,265 zonal connectors*<br>• *Network details vary from country to country*<br>Rail network:<br>• *18,851 nodes*<br>• *19,867 links*<br>• *1,269 zonal connectors*<br>Maritime network:<br>• *747 nodes*<br>• *812 links*<br>Air network<br>• *522 airports*<br>• *8,507 links*<br>• *7,962 zonal connectors* | • *All functional classified roadways plus some local roads*<br>Rest of Georgia State (127 counties)<br>• *All NHPN links validated by Georgia DOT RC file*<br>50-mile buffer area (including portions of f neighbor states)<br>• *All of the existing NHPN links*<br>Rest of the five neighbor states<br>• *Interstates and major state routes*<br>Outlying states (43 states and District of Columbia)<br>• *Interstates only*<br>The final network consists of 1,569 centroids and 82,360 miles of roadways |
| Computational Model | Travel demand model:<br>• *Spatial Computable Generalized regional Equilibrium model (SCGE)*<br>Assignment models<br>• *Roadway: Multi-class mixed probit stochastic user equilibrium procedure*<br>• *Rail: Mixed multi-class probit*<br>• *Inland waterway: shortest path procedure* | N/A |
| Implementation | ArcGIS for geoprocessing<br>Assignment models built in C# together with Rapidis | ArcGIS for network building<br>CUBE software for centroid and centroid connector generation |

The Southwest Georgia Interstate Study lacks computational modules for trip assignment, and both this study and the European TRANS-TOOLS Transport Model are not long-distance oriented. The bright side is that both studies well elaborated the efforts for constructing a geodatabase. Notably nationwide several state and local agencies pioneered the development of state- and city-wide travel data systems for system monitoring and network performance evaluation purposes. Although they may not be directly related to long-distance travel, they do provide invaluable insight in how a variety of datasets can be organized, retrieved, and analyzed on a GIS basis. Those systems are reviewed in the following sections.

- Visual Interactive System for Transportation Algorithms (VISTA)

A mesoscopic traffic simulator (RouteSim), together with algorithms in STA, DTA, control, and routing, have been implemented and embedded in the VISTA framework (Mouskos, *et al.*, 2003). The framework is based on the Common Object Request Brokerage Architecture (CORBA) specification that allows the modules to be written in separate programming languages, and run on different machines over a network. Its user interface was based on a Java-enabled GIS platform with zooming, panning, and query capabilities and can be accessed through the Internet.

- Freeways, Arterial Performance Measurement System (APeMS), California

APeMS (Petty *et al.*, 2005) is a software system that collects and archives transportation data ranging from conventional point sensors, special point sensors (Weigh-In-Motion stations), link-based values (toll tag data, Bluetooth-based), arterial data, events (lane closures, incidents), to transit information (schedules, load, location). APeMS computes and stores performance measures, and provides multiple ways to visualize transportation data online.

- Portland Oregon Regional Transportation Archive Listing (PORTAL), Oregon

PORTAL (Tufte *et al.*, 2010) archives 20-second speed, count, and occupancy data from the approximately 600 inductive loop detectors in the Portland, OR and Vancouver, WA metropolitan region. Additionally, PORTAL stores other transportation related data including weather, incidents, and variable message sign displays in addition to bus AVL and truck Weigh-In-Motion records. The web-based interface of PORTAL provides easy access to both raw data and a wide range of common summary data and standard performance measures.

- Regional Integrated Transportation Information System (RITIS), Maryland

The RITIS system (Pack *et al.*, 2008) archives transportation operations data, like traffic volume and speed, incident information, weather data, device operational status (traffic detectors, VMS, traffic signals, highway advisory radio, and CCTV cameras), managed lane status (high-occupancy vehicle, high-occupancy toll, and reversible lanes), surveillance video, transit alerts, automated vehicle locations, signal status, signal timing plans, computer-aided dispatch (CAD) information, statistic and descriptive information. RITIS information is available through a read-only web interface. RITIS has two primary capabilities: the exchange of real-time transportation related information and the archival of regional transportation-related data.

- Digital Roadway Interactive Visualization and Evaluation Network (DRIVE Net), Washington

The Digital Roadway Interactive Visualization and Evaluation Network (DRIVE Net), currently under development at the STAR Lab of the University of Washington, serves as a new online platform for transportation data sharing, visualization, analysis, and modeling (Ma, *et al.*, 2011). DRIVE Net currently stores various transportation data in its backend database, including freeway loop detector data, traffic signal control data, traffic incident data, travel time data, and truck GPS data. These data can be used to support online computing for other variables of interest, for example incident induced delay, dynamic routing for specific OD pair, and traffic emissions.

# Chapter 3 Network Representation in Geodatabase

## 3.1 Brief Introduction

As mentioned in Section 1.2, a critical component in the national data framework is an abstraction of the national highway network as the common ground of basic geographic representation to which all the datasets related. Portraying a network is conceptually simple through the notion of junctions (nodes) and edges (links), with junctions representing places where redistribution and consolidation activities occur, and edges representing the medium that connects junctions and depicts the trajectory of flow. The pattern reflected by how junctions and edges are organized is a network, where the geometric pattern designates the geospatial locations (geometric network, as shown on the right in Figure 3.1) while the topological pattern indicates the underlying connectivity information (logical network, as shown on the left in Figure 3.1).

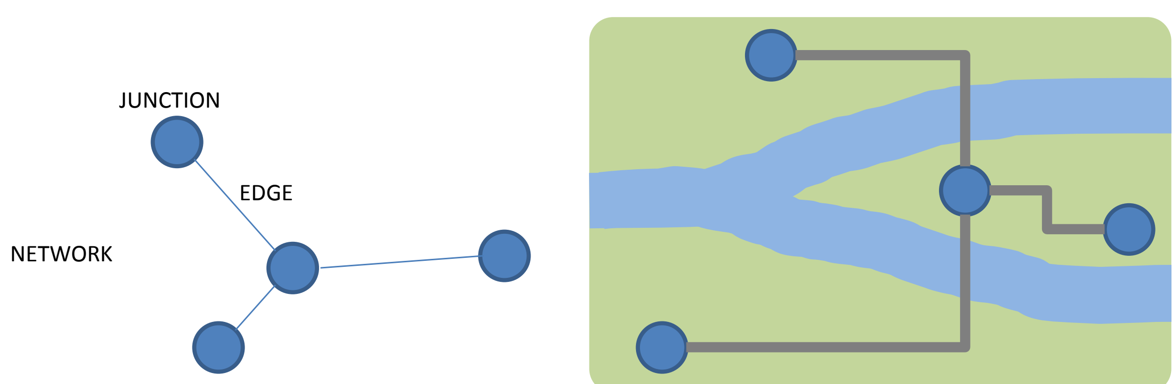



Figure 3.1 Edges and Junctions in a Geometric Network (Right) and Logical Network (Left)

Data regarding the national highway network can be stored and managed in a Geographic Information System (GIS) with geodatabase support. In ArcGIS, a prevailing implementation environment for geospatial analyses, junctions and edges in a transportation network are instantiated as feature classes and stored in a geodatabase. A feature class is a table with a special column called "*shape field*," indicating the geometry of the feature. Features in one feature class share the same geometry, including *point*, *polyline*, *polygon*, *multipoint*, and *multipatch*. Besides feature classes, another distinct feature of how geodatabase organizes information is the relationship class. A relationship class is a table that keeps track of complex and attributed relationships explicitly in the geodatabase. A relationship class implements association rules between feature classes and tables.

With these two distinct features, a geodatabase is indeed an object-relational database design, where an object class encapsulates some data together with functions storing, querying, and manipulating the data. Geodatabase extends from classical relational database design to embrace the needs to represent geographic objects, and therefore gains the capability for advanced geospatial analyses. Traditional data formats in files and tables are also incorporated in geodatabases. Application of geodatabases provides a working solution for network representation and multi-data source integration and thus will play a key role in this research.

## 3.2 Network Model Design

### *3.2.1 User requirements and project objectives*

Section 3.2 has identified the source data as well as the conceptual data model, which consist of entities and their relationships, and established the general structure (the skeleton) of the database. Designing more details for this data model requires a clarification on the user requirements and project objectives.

A national highway network data model should support a wide variety of existing and potential future applications. Concentrating on network loading and distribution issues for both passenger and freight travel needs, this study identifies specific user requirements for the data model:

- A national highway network model should represent the infrastructure reality as well as address the implementation feasibility;
- In the data model, "friction" factors should be developed and self-corrected through historical archived data and most updated network traffic data; and
- The network should be multi-scale with geographic coverage at county, metropolitan planning organizations, state, regional, corridor, and national levels.

Particularly, such a data model should support answering the following example questions: at individual level, if a person is traveling from Seattle to Chicago (junctions), which route (edge) should they take? At system level, how are different routes utilized by long distance passenger and freight travels between Seattle and Chicago?

Considering answering these questions, the data model design should be capable of:

- Not only depicting a network, but also reflecting traffic operational status on this network by incorporating most updated traffic data and other travel impedance measures;
- Not only embracing multiple scales to meet the needs for trip assignment at multiple network levels, but also maintaining network integrity at each level;

- Not only tailoring for the existing data and application reality, but also preparing for future utilization and expansion.

Constructing such a computational infrastructure necessitates the adoption of geodatabase architecture, mainly for the following reasons:

- **Geodatabase offers extra functionalities** over relational databases, such as geospatial analyses and queries, as well as advanced information visualization and presentation capabilities. Particularly, the ArcGIS geodatabase configuration maintains good compatibility with various geospatial data file types, including shapefiles, coverages, CAD files, and imagery. ArcGIS is also flexible to operate with other applications for complicated computations, like TransCAD and Rapidis for traffic assignment needs.
- **Geodatabase integrates transportation information alongside the geospatial distribution of roadway network.** Modeling transportation network is idiosyncratic from the classic theme-based database design, where multiple tables are related or a collection of coverages are overlaid. A generic network model contains facility centerlines attached with event tables using linear referencing system. Geodatabase supports such capabilities to reduce the database redundancy and facilitate efficient database revision and expansion.
- **Geodatabase supports framework standards.** Methodologies of managing information and knowledge evolve rapidly with the growing diversity of data collection techniques and emerging business needs in transportation. The object-relational database design concept, which is the core of geodatabases, not only groups the real-world objects in a

geospatial fashion, but also explicitly illustrates relationships between those objects, including association, type inheritance, instantiation, aggregation, and composition.

With user requirements and project objectives clarified, the next steps include:

- Conceptual data model design: The conceptual data model identifies the entities in a network and specifies relationships between entities. Specify the entities in the geodatabase at different levels: the network at each level of the framework corresponds to a distinct set of edges and junctions. Those features represent different real-world objects. On the one hand, the representation should be consistence with the data sources available; on the other hand, topological integrity (network connectivity) on each level should be enabled. A base network should be established to maintain sufficient accuracy when representing the national highway network, and also keep the network size manageable to ensure the computational efficiency.
- Logical data model design: The logical data model specifies attributes for the geodatabase based on the conceptual design and the base network specified.
- Physical data model design: The physical data model further specifies the geodatabase based on the logical data model, while considering the implementation environment. The choice of RDBMS (relational database management software), network structure, and organizational behaviors will influence how the logical design is translated into a physical implementation. Detailed geodatabase configurations will be specified in the physical data model.

### *3.2.2 Conceptual data model design*

The conceptual data model design developed in this study follows a triple-level system to represent a multi-scale national highway network desired by long distance travel analyses. The top level (highlighted by the top dash line box in Figure 3.2) represents the macroscopic representation of the national highway network, where the Interstate freeways connect states. Although the top level network is too coarse to reflect detailed network loading conditions, it can yield meaningful knowledge of major passenger and freight flows loaded onto the Interstate system between different states. This level of network could also be a good prototype system due to its network-wide simplicity for demonstration purposes. The middle and bottom levels (highlighted by the bottom dash line box in Figure 3.2), representing region and county level networks respectively, provide scaling steps for updated network information to be integrated for quantifying “friction” factors at the corresponding network level. Specifically, the middle level comprises MSAs and principal arterials, addressing the needs for depicting the travel between major urban clusters. The bottom level embraces counties and all arterials, in the finest resolution in this model design, addressing the needs for loading county level OD demand to the national highway network. In summary, all three levels attempt to capture inter-state travel, inter-regional travel, and inter-county travel respectively. As illustrated in Figure 3.2, across all three levels there are corresponding logical networks and geometric networks, where the former outlines the topical relationships between edge junction **elements**, while the latter specifies the edge and junction **features** that abstract the real-world places and roadways.

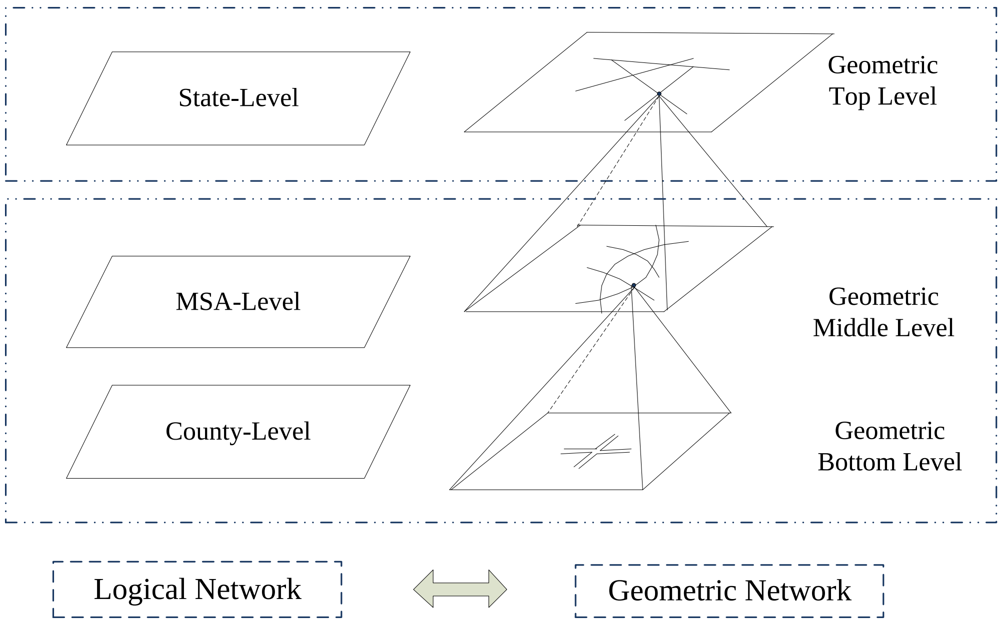



Figure 3.2 Triple-Level System for National Highway Network Model

Based on the triple-level system, the conceptual data model design for national highway network is devised as shown in Figure 3.3. The conceptual data model identifies the entities in a network and specifies relationships between entities. The Interstates, principal, and minor arterials for both urban and rural areas in the NHPN shapefile (a polyline feature class) constitute the base network, as shown in Table 3.1. Orphan junctions are automatically generated from the base network to ensure the network's topological integrity. Centroids of polygons from MSA and county shapefiles are generated to represent OD junctions on different levels and connect to the network through centroid connectors. Centroid connectors are virtual links that represent generalized local roads that have no physical correspondence in the real-world roadway network. This allocation attempts to maintain sufficient accuracy when representing the national highway network, and also keep the network size manageable to ensure the computational efficiency.

Table 3.1 NHPN Functional Classes included in the Base Network and the Mileage Status

| Functional Class | Mileage |
|---|---|
| Urban Principal Arterial – Interstate (11) | 16,134 |
| Rural Principal Arterial – Interstate (1) | 33,918 |
| Urban Principal Arterial - Other Freeways and Expressways (12) | 12,070 |
| Urban Principal Arterial – Other (14) | 58,955 |
| Rural Principal Arterial – Other (2) | 105,273 |
| Urban Minor Arterial (16) | 2,146 |
| Rural Minor Arterial (6) | 125,346 |
| Total | 353,842 |

Centroid and Network Node (orphans) are instances of JunctionFeatureSoure (a super class); while CountyCentroidConnector, MSACentroidConnector, and NHPN are instances of EdgeFeatureSoure (another super class). Centroid is connected to Network Node (orphans) by CountyCentroidConnector and MSACentroidConnector. Travel Impendence refers to the resistance drivers may experience to complete a long distance trip, such as travel time, toll, lane restriction, etc. The Travel Impendence information is linked to the base network from NHPN. As an example dataset for computing Travel Impendence, loop data is discussed in Chapter 5 to demonstrate how traffic data can be attached to the network.

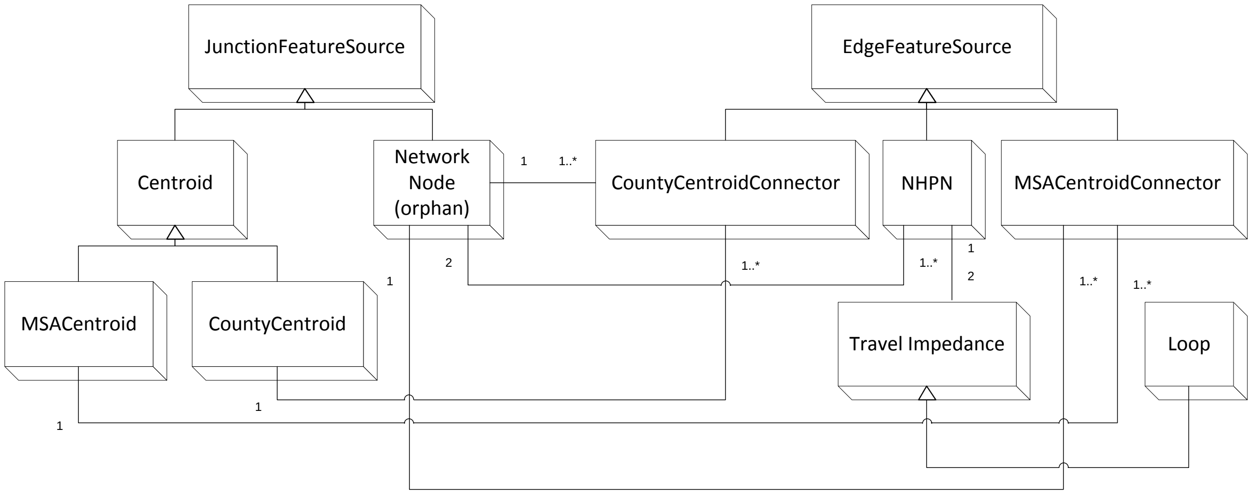



Figure 3.3 Conceptual Data Model Design for National Highway Network

Table 3.2 specifies the source datasets for edge and junctions features on each geometric level. Since NHPN dataset serves as a geospatial basis for both passenger travel and freight transportation ($FAF^3$) statistics, different subsets of it are referred to in Table 3.2 for different levels of network following the functional classification codes specified in Table 3.1.

Table 3.2 Specifications of Geometric Networks

| Network Scale | Junction Feature | Edge Feature (both Urban and Rural) |
|---|---|---|
| State-Level | State centroids | NHPN Principal Arterial – Interstate (1 &11) |
| MSA-Level | MSA centroids | NHPN Principal Arterial – All (1, 2, 11, 12, & 14) |
| County-Level | County centroids | NHPN Principal and Minor Arterial (1, 2, 6, 11, 12, 14, & 16) |

Number of lanes for different type of facilities is summarized in Table 3.3, as they are critical reference to determine the capacity of roadway segments. The final base network is illustrated in Figure 3.3. For illustrative purposes, minor arterials are not shown to make the network more legible.

Table 3.3 Observed Network Functional Classes and Number of Lanes

| Functional Class | Number of Through Lanes in Both Directions | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
| Urban Principal Arterial – Interstate (11) | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ |
| Rural Principal Arterial – Interstate (1) | √ | √ | √ | √ | √ | √ | √ | √ | √ | | | | | |
| Urban Principal Arterial - Other Freeways and Expressways (12) | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | √ | | | |
| Urban Principal Arterial – Other (14) | √ | √ | √ | √ | √ | √ | √ | √ | √ | | √ | | | |
| Rural Principal Arterial – Other (2) | √ | √ | √ | √ | √ | √ | √ | | √ | | | | | |
| Urban Minor Arterial (16) | √ | √ | √ | √ | √ | √ | | | | | | | | |
| Rural Minor Arterial (6) | √ | √ | √ | √ | √ | | | | | | | | | |

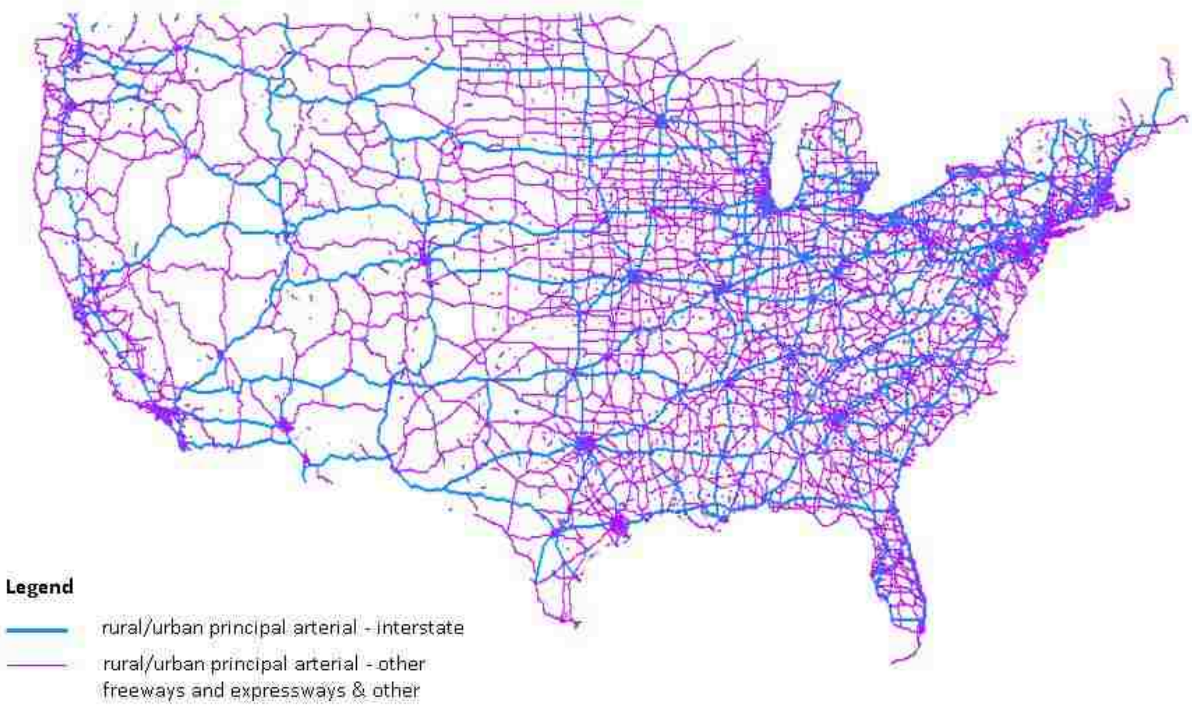



Figure 3.3 Base Network for National Highways

### *3.2.3 Develop logical data model*

The logical data model specifies attributes for the geodatabase based on the conceptual design and the base network specified in Table 3.2. The base network is the bottom level (county level) in the triple-level system, as the two upper levels are its subsets. The truck route source data (FAF$^3$) also overlaps significantly with the NHPN network, therefore the following can be derived from the base network: (1) different levels of highway networks depending on the road way functional classes; (2) the truck route network by relating to the FAF$^3$ network. Since other networks share great similarities with the base network in terms of data model design, further discussions in this report only focus on the base network.

Figure 3.4 shows the logical data model design. As in the conceptual design, the generic building blocks for the base network are **JunctionFeatureSource** and **EdgeFeatureSource**. The *Shape* attribute specifies a feature's geometry. The *IsEnabled* attribute can remove a network element from consideration in traffic assignment at different levels by setting its value to False, and thus materialize the multi-scale network design.

The superclass stereotype **JunctionFeatureSource** has two subclasses **Centroid** and **NetworkNode**. The *AncillaryRole* attribute in the **Centroid** class indicate whether the corresponding MSA or county is origin or destination (source or sink following the definition of *AncillaryRole*). The **Centroid** class further extends the type inheritance to the **MSACentroid** and **CountyCentroid** classes, which are simple junction feature classes generated from MSA and County polygons. Both the **MSACentroid** and **CountyCentroid** classes need an attribute to store a foreign key pointing to the corresponding polygon it represents in all cases. For the **MSACentroid** classes the foreign key is *CBSA* identifying code, which is a five-digit code assigned to each metropolitan and micropolitan statistical area. For the **CountyCentroid** classes the foreign key is the *CTFIPS* code, which represents each county.

The superclass stereotype **EdgeFeatureSource** has three subclasses: **NHPN**, **MSACentroidConnector**, and **CountyCentroidConnector**. The *FromJunctionID* and *ToJunctionID* attributes in the **NHPN** class indicate the two ends for an edge, and relate to the **NetworkNode** class through a one-to-two cardinality. The *FCLASS* attribute in the **NHPN** class is used to set preference in different levels of network. The LRSKEY, LRSSEQ, BEGMP, and ENDMP attributes in the NHPN class form a linear referencing system to attach traffic information (the **TravelImpedance** classes), which will be explained in detail in Chapter 5. The **MSACentroidConnector** and **CountyCentroidConnector** classes are "virtual links" connecting **MSACentroid** and **CountyCentroid** to the **NHPN** class. In real-world, they represent the local roads that are not physically specified and included in this data model due to their lower hierarchy in roadway functional classes.

The authors would like to include a special note regarding the network connectivity. Table 3.2 specifies the feature source at different levels and their corresponding codes in the NHPN dataset. Building a network dataset begins with specifying at least one edge feature classes (e.g. all NHPN principal arterials), which will form the network at a particular level (e.g. MSA level). This is done by assigning the classes to connectivity groups. In this study, there are three distinct levels of network; therefore, there are three connectivity groups. Each connectivity group corresponds to one edge feature source. Explicit junction feature sources (the **MSACentroid** and **CountyCentroid** classes in this design) are required to connect to edges in different connectivity groups directly or indirectly. Also, to differentiate passenger travel from freight transport, a dualmodal network design may involve the highway edge feature sources and truck route edge feature sources. To ensure that trucks traverse only on truck route segments, the two edge feature sources are placed in different connectivity groups. The junction feature sources are the same for both connectivity groups at all levels in this case, as both passenger and freight transport activities share the same OD configuration.

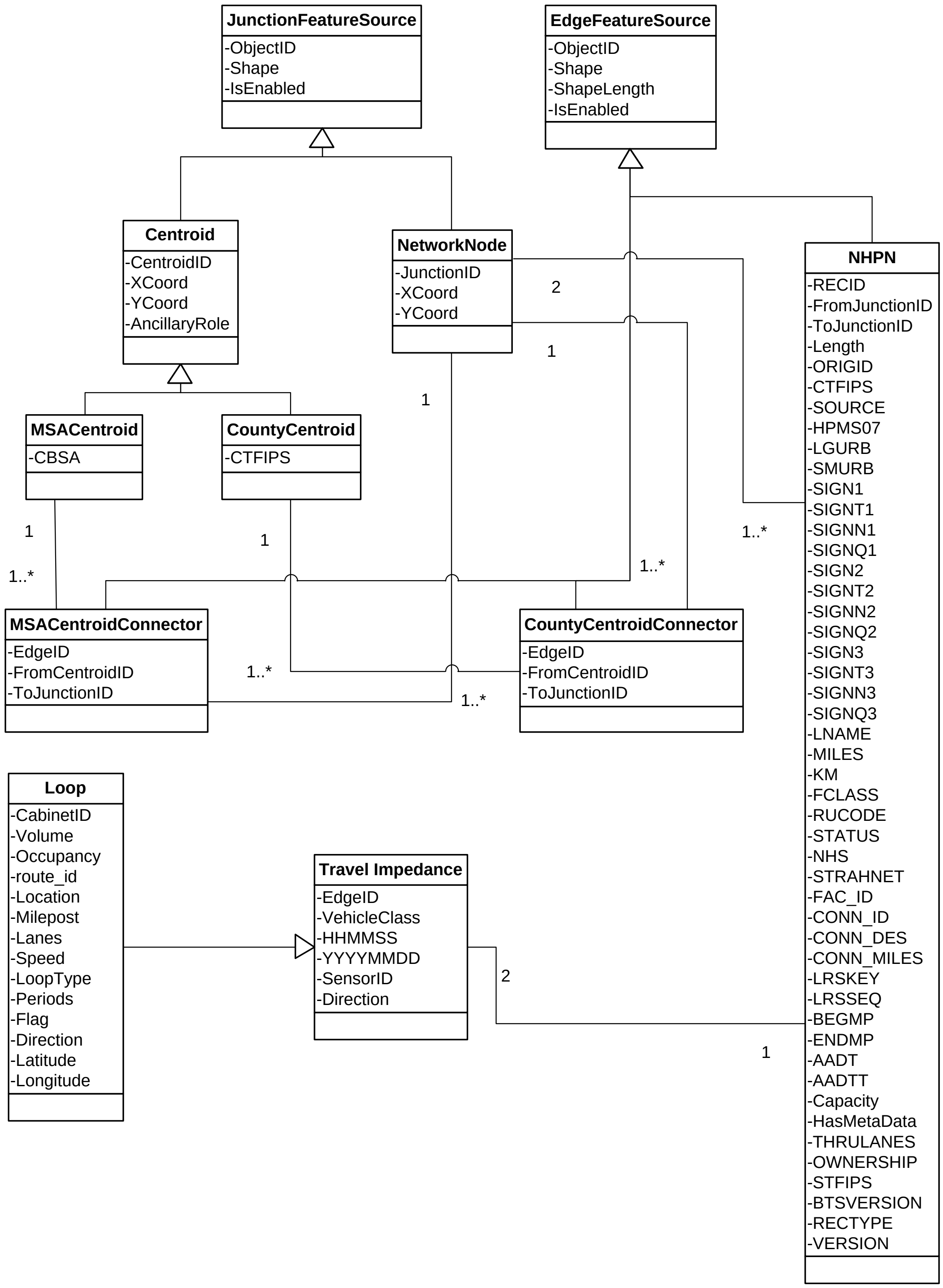



Figure 3.4 Logical Data Model

### 3.2.4 *Develop physical data model*

The physical data model further specifies the geodatabase based on the logical data model, while considering the implementation environment. The physical data model shown in Figure 3.5 is for an ArcSDE geodatabase that includes a geometric network called HwyNetwork, plus two polygon feature classes defining the **MSA** and **County** coverages. The geometric network does not contain those polygons, only their centroids, which are stored in a simple junction feature class, called **MSACentroid** or **CountyCentroid**. Each polygon feature class is related to the corresponding centroids explicitly by a relationship class **CentroidRepresentsMSA** or **CentroidRepresentsCounty**. The **NetworkNode** feature class is an implementation of the NetworkNode class in the logical data model, which defines all the orphan nodes automatically generated with creating a network dataset using the NHPN base network shapefile, which is stored in a simple edge feature classes **NHPN**. Two other simple edge feature classes store **MSACentroidConnector** and **CountyCentroidConnector**.

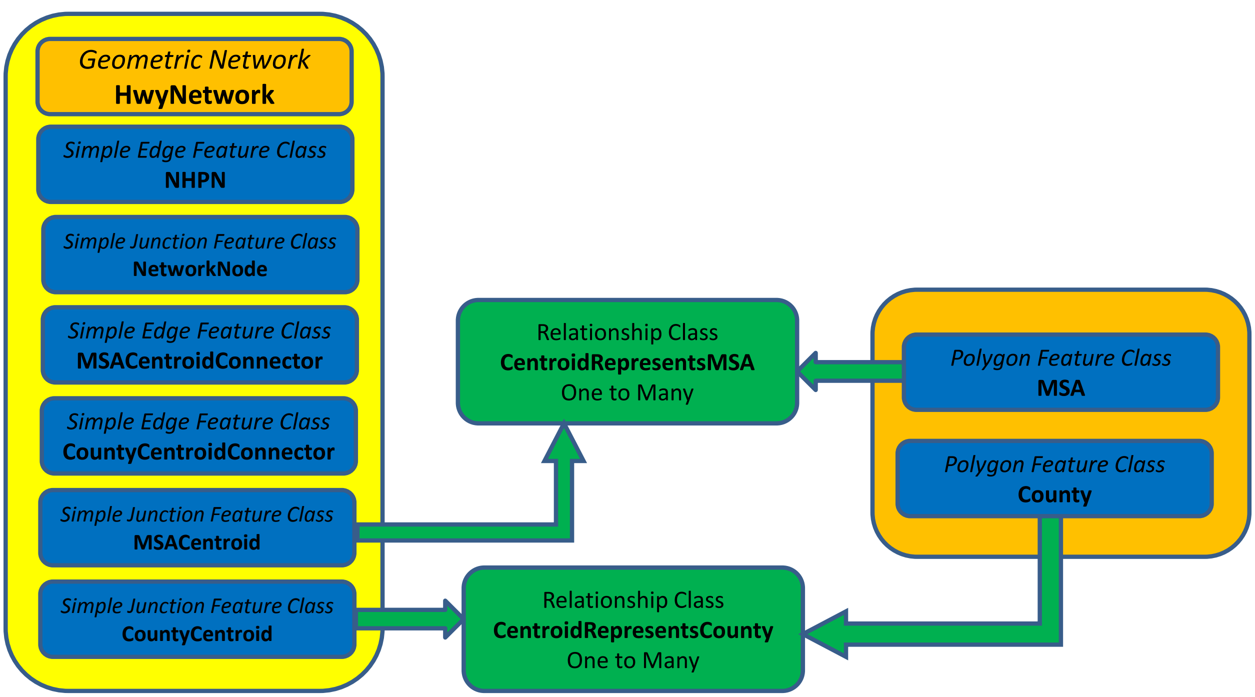



Figure 3.5 Physical Data Model in ArcGIS

The relationship class **CentroidRepresentsMSA** connecting **MSA** and **MSACentroid** is illustrated in Figure 3.6. The other relationship class **CentroidRepresentsCounty** follows the same fashion.

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Simple Feature Class | | | | Geometry | | *Polygon* | |
| **MSA** | | | | Contains M values | | *No* | |
| | | | | Contains Z values | | *No* | |
| Field name | Data Type | Allow nulls | Default value | Domain | Precision | Scale | Length |
| ObjectID | ObjectID | | | | | | |
| Shape | Geometry | Yes | | | | | |
| Shape_Length | Double | Yes | | | 0 | 0 | |
| Shape_Area | Double | Yes | | | 0 | 0 | |
| CBSA | String | No | | | | | 9 |
| CBSA_Name | String | No | | | | | 56 |
| TYPE | String | No | | | | | 29 |
| STATUS | String | No | | | | | 11 |
| ST_POSTAL1 | String | No | | | | | 2 |
| ST_POSTAL2 | String | No | | | | | 2 |
| ST_POSTAL3 | String | No | | | | | 2 |
| ST_POSTAL4 | String | No | | | | | 2 |
| STFIPS1 | String | No | | | | | 2 |
| STFIPS2 | String | No | | | | | 2 |
| STFIPS3 | String | No | | | | | 2 |
| STFIPS4 | String | No | | | | | 2 |
| VERSION | String | No | | | | | 2 |

| | | | |
|---|---|---|---|
| Relationship class | **MSAHasCentroid** | | |
| Type | *Simple* | Forward label: | *Is represented by* |
| Cardinality | *One to many* | Backward label: | *Represents* |
| Notification | *Forward* | | |
| Origin table | | Destination table | |
| name | MSA | name | MSACentroid |
| Primary key | CBSA | | |
| Foreighn key | CBSA | | |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Simple Feature Class | | Geometry | | *Point* | | | |
| **MSACentroid** | | Contains M values | | *No* | | | |
| | | Contains Z values | | *No* | | | |
| Field name | Data Type | Allow nulls | Default value | Domain | Precision | Scale | Length |
| ObjectID | ObjectID | | | | | | |
| Shape | Geometry | Yes | | | | | |
| MSACentroidID | Long Integer | No | | | 0 | | |
| CBSA | String | No | | | | | 20 |
| Xcoord | Short Integer | Yes | | | 7 | 2 | |
| Ycoord | Short Integer | Yes | | | 7 | 2 | |
| AncillaryRole | Short Integer | No | 0 | NetwrokRole | 0 | | |
| IsEnabled | Short Integer | Yes | 1 | EnabledDomain | 0 | | |

| | | | | |
|---|---|---|---|---|
| Subtypes of Centroid | | | | |
| subtype field: | *CentroidType* | Default subtype: | *0* | |
| Subtype Code | Subtype Description | Field name | Default value | Domain |
| 0 | Neither | | | |
| 1 | Source | | | |
| 2 | Sink | | | |

Figure 3.6 Relationship Class for CentroidRepresentsMSA

In Figure 3.6, feature classes MSA and MSACentroid are also specified, with field name, data type, allow nulls, default value, domain, precision, scale, and length for each attributes of the feature class. Attributes for other entities in Figure 3.6 are specified in Table A-1 through Table A-7 in Appendix A, in the format of how they are stored in an ArcSDE geodatabase. Loop data as an instance of the super class **TravelImpedance** is stored as a table rather than feature class, as shown in Table A-7. "*An attribute domain is a rule that limits users' data entries to a specific set of valid choices.*" (Bulter, 2008) Domains specified in Table A-1 to Table A-7 are listed in Table A-8 to A-20 in Appendix A.

## 3.3 Linking Traffic Data to the Network

The logical and physical data model outline the "skeleton" of the national highway network data model, however, to enable this "skeleton" with "mobility", "muscles" should be developed. In the long distance travel context, the "mobility" refers to the computational power for trip assignment and network loading, and the "muscles" refer to other travel impedance measures including traffic data. In this chapter, we focus on elaborating the procedure of linking traffic data to the network. Linking other travel impedance measures shares similarities with this procedure to be described.

The Linear Referencing System (LRS) is critical in this procedure, from concepts to practices. LRS facilitates (1) consolidating different network datasets. The transportation network is widely represented as 1-D cartography with the roadway centerlines. Therefore operations such as merging, relating, and data transfer on different datasets (e.g. HPMS, NHPN, and FAF) can only be completed by LRS rather than map overlay; (2) linking traffic information to the network. A majority of traffic data are universally collected and stored as "events" along the roadway network, associated by a diversity of Linear Referencing Methods (LRMs). In the following sections, concepts of LRS, LRM, and event data are first introduced. Then the authors discussed the discrepancy between the LRS in the NHPN and State DOT practices, as well as the challenges for linking traffic data induced by this discrepancy. Finally two-layer solutions are proposed to cope with the challenges.

### *3.3.1 Brief introduction of linear referencing system and event data*

The practice of locating objects along a network based on a series of anchor points and the distance between objects and anchors dates back to the early days before the advent of computer and modern GIS technology. The primal LRS was developed in the railway industry and the idea was adopted by more and more transportation agencies across diverse modes. The technique "*to identify a specific location with respect to a known point*" is referred as Linear Referencing Method (LRM) (Vonderohe, A.P., *et al*., 1993); and a Location Referencing System (LRS) is "*a set of office and field procedures that include a highway location reference method*" (Baker and Blessing, 1974).

In general LRM consists of two components: facility identifier and linear measure, which is the distance from the pre-defined origin to the point of interest (Butler, 2008). Data collected are stored in event tables, often categorized as point events with one linear measure and linear events with two linear measures (Fekpe, *et al*., 2003).

A diversity of LRMs are utilized by state and local DOTs, some examples enumerating route-milepost, route-offset distance, route-milepost-offset, and methods based on link-node models (Baker and Blessing, 1974). Adams *et al*. (1999) described the various LRMs used by states and compared their advantages and disadvantages. It was found that not only each state was developing their own data models, but the terminologies were also not consistent from one state to another. This situation hindered the interoperability among datasets based on different LRMs. In the 1990s, the need for a national LRS model grew, with the most influential research resonating to the needs being NCHRP20-27(2), entitled "*Development of System and Application Architectures for Geographic Information Systems in Transportation*".

The logical data model of NCHRP20-27(2) is shown in Figure 3.7, which includes three primary components: first, the LRS and its linear datum; second, business data, and third cartographic representation. A detailed explanation can be found in the NCHRP20-27(2) report but the major purpose of such an initiative is "*to integrate increasing amounts of linearly referenced data used by the transportation community*" (Vonderohe *et al.*, 1997) by proposing a generic and comprehensive data model with an agreed technical standard, and to facilitate sharing those data across modes and agencies.

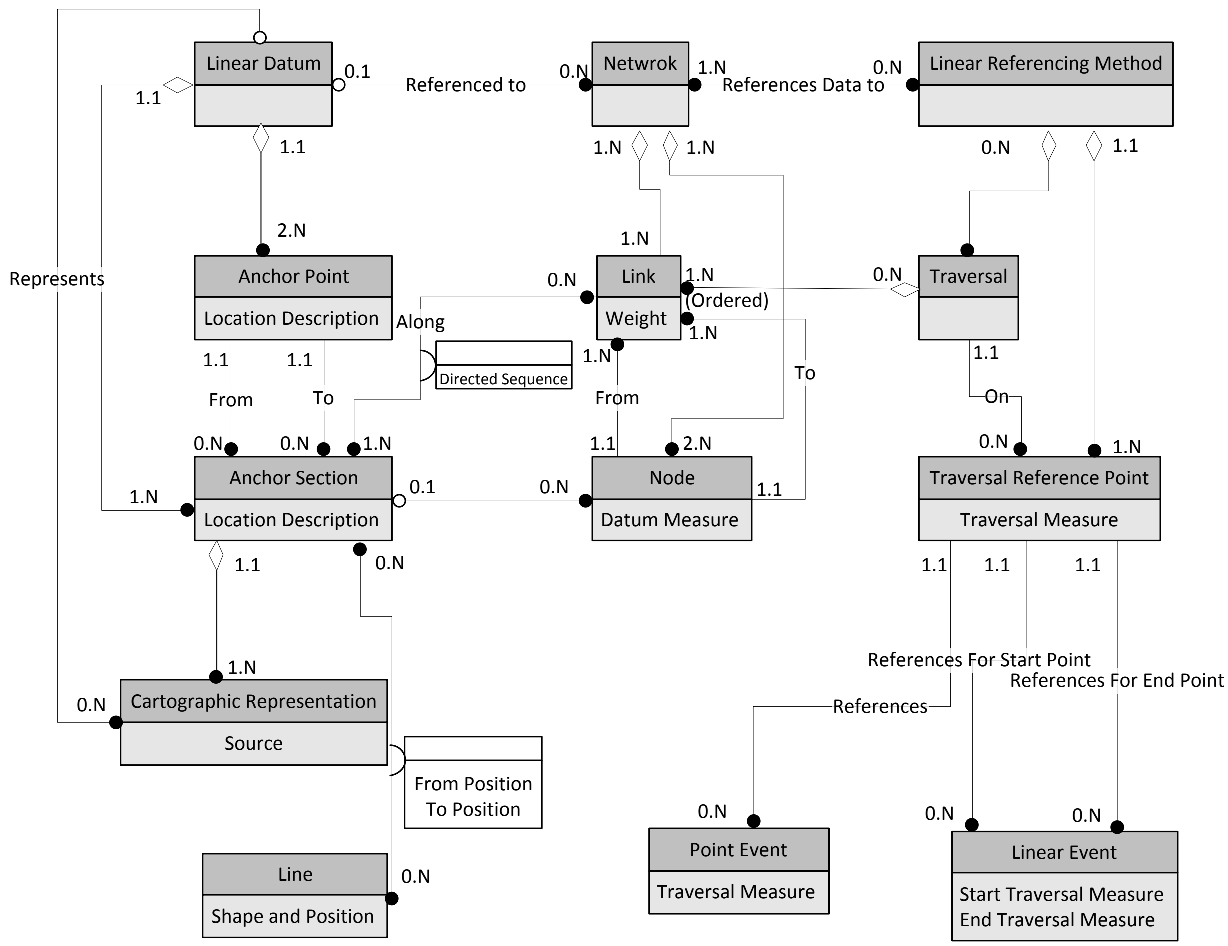


Figure 3.7 NCHRP 20-27(2) Proposed National Transportation Data Model

The data model proposed in NCHRP20-27(2) as a “universal” LRS prototype provides substantial guidance for standardizing state level LRS practices. However, while the design procedure and specification well conformed to the national model, actual LRS practices still varies from state to state (Fekpe, *et al.*, 2003). Even in the same state, it may vary on different issues (pavement management, traffic monitoring, performance reporting, traffic incidents, etc.) or by different Distance Measuring Instruments (DMIs).

To overcome the idiosyncrasies across different state level LRSs, the Federal Highway Administration (FHWA) created a national LRS on the basis of the NHPN network, which serves as the geometric base of HPMS. The FHWA requires each state to conform to this national LRS when submitting annual HPMS report. The national LRS also links NHPN to HPMS and $FAF^3$, with four major components (FHWA, 2010):

- County FIPS: 3-digit County Federal Information Processing Standards codes created by Bureau of the Census;
- Inventory Route Number: 10-digit code uniquely identifying a route for inventory purposes, not necessarily the same as that posted along the roadway;
- Inventory Subroute Number: 2-digit code uniquely identifying portions of an inventory route;
- Milepoint/Kilometerpoint (MPT/KMPT): actual milepost.

The first three data items (Inventory Route Number + Inventory Subroute Number + County FIPS) form a 15-digit unique code called **LRSKEY** in the NHPN dataset and together with BEGMP and ENDMP (see Table 3.1) identifies a particular portion of a route. The critical function of **LRSKEY** is discussed in the following section.

### *3.3.2 Handling traffic data*

This study divides traffic data into two categories: historical data and most updated data. The former one refers to the dataset routinely collected and reported, like the AADT in HPMS and Truck Volume in $FAF^3$. The advantage of historical data is their wide coverage due to FHWA's requirement on data submittal. The disadvantage of historical data is that the annual average cannot fully reflect the weekday/weekend and seasonal traffic variations, which are critical impact factors in long distance travel decision making. The most updated traffic data are available in higher temporal and spatial resolutions, collected by a variety of traffic detection techniques. However, since those data collections are conducted by different transportation agencies, it is challenging to link them to the NHPN network using the national LRS introduced in Section 5.1. The following sections propose two-layer solutions with both breadth and depth, by incorporating both historical data and most updated data.

#### 3.3.2.1 Basic Data Layer

The objective of constructing the basic data layer is to ensure that each edge on the base network is associated with AADT and capacity information. For freight network, truck volume should also be included as an attribute. The **base network** is the county level NHPN network as specified in Table 3.5, which includes principal and minor arterial (coded as 1, 2, 6, 11, 12, 14, & 16 in the **FCLASS** attribute), as high levels of network are only subsets of this source network and can be retrieved with the **FCLASS** codes as specified in Table 3.5. The historical traffic data are stored as attributes in the HPMS and $FAF^3$ datasets. Feeding the HPMS and $FAF^3$ data to the base network is a five-step procedure:

- **Match routes**: both HPMS and $FAF^3$ can be related to NHPN using the **LRSKEY** attributes. Note that **LRSKEY** specifies a route in any of the three datasets, each of which may consist of different number of edges. For example, the **LRSKEY** "*000000000500033*" specifies the Interstate-5 section in the metropolitan Seattle area, which corresponds to 99 edges (segments) in the NHPN dataset while 28 edges (segments) in the HPMS dataset. Both routes share the same beginning and ending milepost, only the NHPN with shorter segments;

- **Transfer historical data**: since there is no one-to-one cardinality between edges across different datasets, historical data are transferred on a mileage basis. Taking the previous example, the 99 NHPN edges will be aggregated into 28 groups to match the 28 HPMS edges based on the beginning and ending milepost, and edges in each group will be assigned the same traffic data as specified in HPMS;
- **Handle missing data**: after step 2 there are still some edges with missing traffic data. In this case average of neighboring edges will be assigned to those edges;
- **Handling non-HPMS/FAF edges**: after step 1 there are a number of edges in the source network without corresponding HPMS/FAF coverage. Estimated AADT from the edge location (CFIPS), number of lanes, and functional class will be assigned to those links;
- **Quality check**: in this step the source network equipped with historical data will be visualized in ArcGIS for visual check for data anomalies.

After the five-step procedure the source network should include the following attributes shown in Table 3.4 with full coverage and relative accuracy. Note that attributes shown in this table are extra attributes that are added to the NHPN Feature Class shown in Table A-1 after the five-step procedure. The **HasMetaData** attribute indicates whether the most updated traffic data are available.

Table 3.4 Extended Edge Feature Class to the Base Network

| | | |
|---|---|---|
| Simple Feature Class | Geometry | *Polyline* |
| **Edge** | Contains M values | *No* |

| | | Contains Z values | | *No* | | | |
|---|---|---|---|---|---|---|---|
| Field name | Data Type | Allow nulls | Default value | Domain | Precision | Scale | Length |
| ObjectID | ObjectID | | | | | | |
| Shape | Geometry | Yes | | | | | |
| ShapeLength | Double | Yes | | | 0 | 0 | |
| EdgeID | Double | No | | | 0 | 0 | |
| FromJunction | Long Integer | No | | | 0 | | |
| ToJunction | Long Integer | No | | | 0 | | |
| AADT | Long Integer | No | | | 0 | | |
| AADTT | Long Integer | No | | | 0 | | |
| Capacity | Double | No | | | 0 | 0 | |
| HasMetaData | Short Integer | No | | MetaDataCode | | | |

Table 3.5 MetaDataCode Domain

| | |
|---|---|
| Coded value domain | **MetaDataCode** |
| Description | *Whether has meta Data* |
| Field Type | *Short Integer* |
| Split policy | *Default value* |
| Merge policy | *Default value* |
| code | Description |
| 0 | Current Edge has meta data |
| 1 | Current Edge has no meta data |

### 3.3.2.2 Advanced Data Layer

Although the basic data layer gives full data coverage across the source network, annual average traffic counts cannot reflect seasonal and weekly fluctuations in traffic, which greatly affect long-distance travel behavior. Meanwhile, the diversity of traffic detection technologies results in a variety of high-resolution traffic data, reflecting the most updated traffic information.

However, two major issues emerge when linking most updated traffic data: first, the massive size of high-resolution data; and second, mapping data to the source network. To resolve the first issue, most updated traffic data are summarized in daily profiles with hourly variations, by day-of-week and month-of-year in the advanced data layer. Therefore, for an entire year, there are just 84 profiles for a roadway segment (in one direction). Regarding the second issue, considering the discrepancy between the state LRS for traffic monitoring systems and the NHPN LRS, it is cumbersome to convert multiple LRS to a uniform one. It is proposed to use the longitude and latitude coordinates of traffic monitoring station for mapping purposes, due the wide availability of this information. In the ArcGIS implementation, stations can be processed as a point feature class that overlays with the base network, a simple geospatial analysis will "snap" the closest edges to the target stations. Using the 2-D referencing method also incorporates the GPS-based mobile sensing technologies. A table for stations and their corresponding network edges is then created, as shown in Table 3.6.

Table 3.6 Traffic Monitoring Station Table

| Table | **Cabinet** | | | | | | |
|---|---|---|---|---|---|---|---|
| Field name | Data type | Allow nulls | Default value | Domain | Precision | Scale | Length |
| CabinetID | String | No | | | | | 10 |
| Xcoordinate | Double | No | | | 0 | 0 | |
| Ycoordinate | Double | No | | | 0 | 0 | |
| EdgeID | Long Integer | No | | | 0 | | |
| Direction | Short Integer | No | | DirectionCode | 0 | | |

For each station, traffic data are summarized into multiple profiles from the raw data collected, following the procedure illustrated in Figure 3.8. Here we use the Washington State Department of Transportation (WSDOT) case as an example to illustrate the approach. Since many State DOTs follow a similar process, the approach based on the WSDOT practice can be applied to other state DOTS without significant modifications. WSDOT collects traffic data from thousands of loops continuously and processes the raw data into 20-second archives at the control cabinets. Unique identifiers are assigned to each loop and a group of loops for one particular roadway segment is associated with a cabinet, with its configuration and location also stored in the database. Considering the scope of project as well as the data quality issues, it is necessary to conduct data aggregation and quality control before further processing 20-second loop data. The Daily Statistics Algorithm (DSA) used in the California PeMS database (Chen *et al.*, 2003) can be implemented to identify erroneous data and aggregate 20-second loop data into daily loop data profiles. Details of the output profile table are shown in Table 3.7.

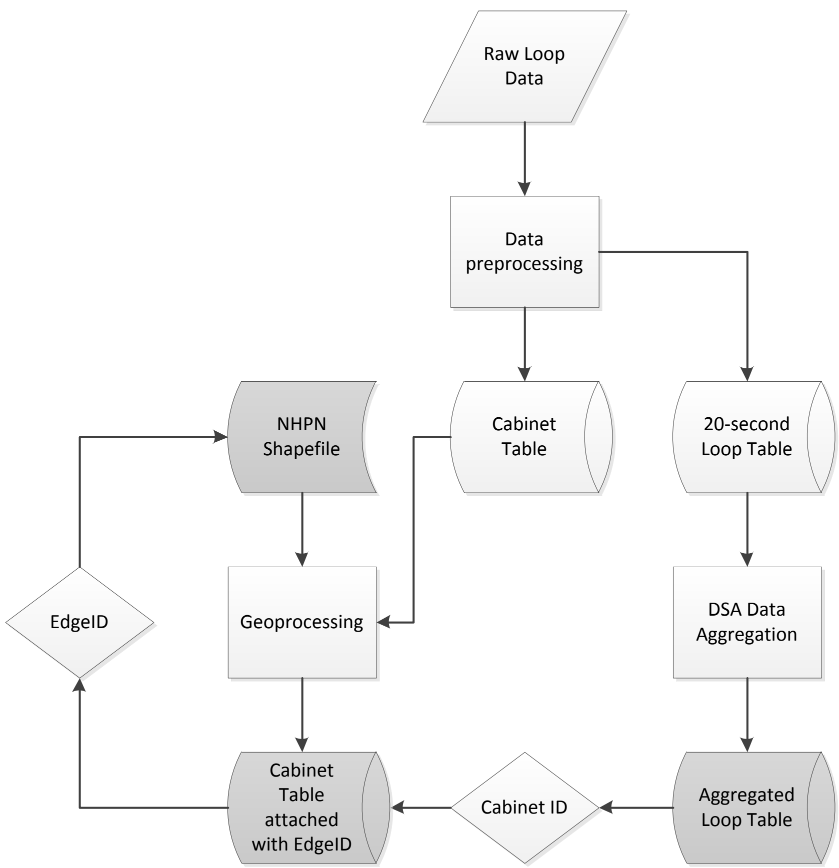



Figure 3.8 Feed Loop Data to the NHPN Shapefile

Table 3.7 AggregatedLoopData Table

| Table | **AggregatedLoopData** | | | | | | |
|---|---|---|---|---|---|---|---|
| Field name | Data type | Allow nulls | Default value | Domain | Precision | Scale | Length |
| CabinetID | String | No | | | | | 10 |
| Month | Short Integer | No | | MonthCode | 0 | | |
| Day | Short Integer | No | | DayCode | 0 | | |
| Hour | Short Integer | No | | HourCode | 0 | | |
| HourlyLoopData | Double | Yes | | | 0 | 0 | |

The domains for Month, Day, and Hour are specified in Tables A-21 through A-23. This two-layer traffic data linking mechanism utilizes both the historical data from the HPMS and FAF[3] datasets and the most updated traffic data when available. Meanwhile, for edges with both historical data and the most updated traffic data, these two datasets can be compared as self-correcting measures. Although this chapter as a whole mainly focused on traffic congestion data, it should be recognized that for traffic assignment there are a plethora of other impedance measures, including tolls, lane restrictions, etc. Those measures can also been linked to the source network in a similar fashion.

# Chapter 4 Long Distance Travel Network Loading: Methodologies

## 4.1 Resolving Network Loading in Long Distance Travel: A Work Flow

Long distance passenger travel embraces fairly idiosyncratic traveler behavior, depending on trip purposes and travelers' socio-demographic characteristics. Perceptions to "pull" and "push" factors regarding alternative routes are quite contrastive across different traveler groups. Additionally, it should be recognized that interactions between different traveler groups are negligible. For example, loading leisure trips to the network will not substantially influence the route choice of business travelers. It is not only because long distance trips merely account for about one percent of total trips (as shown in Figure 4.1), but also because the seasonality characteristics vary for different trip purposes. Therefore it is reasonable to study each traveler group individually and assume that travelers within a certain group select routes in a similar way. Based on this assumption a tentative work flow for resolving the problem is proposed as follows:

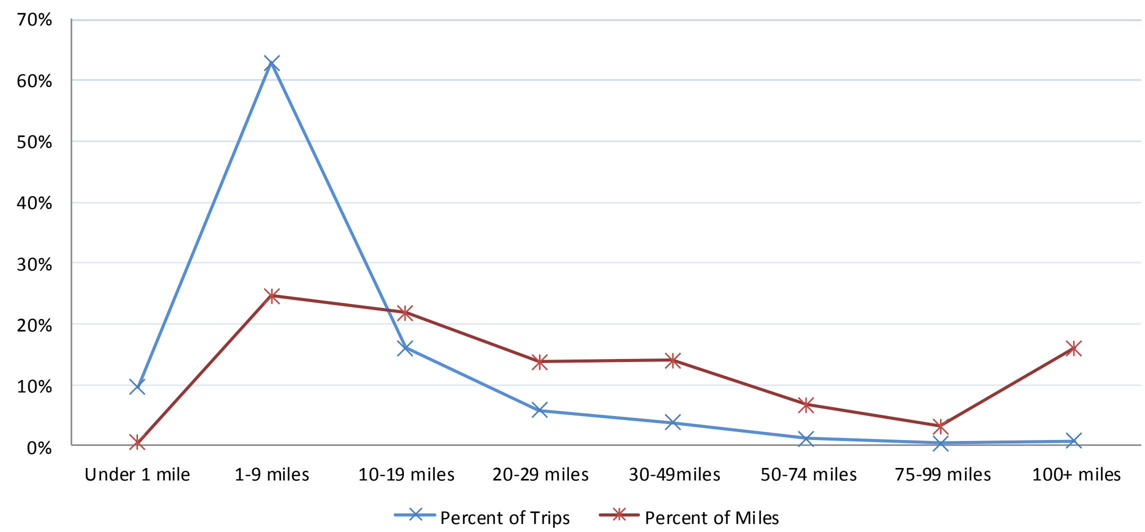



Figure 4.1Vehicle Trips and VMT by Trip Length, 2009 NHTS

***Market Segmentation***: The idiosyncrasy within the whole population of long distance travelers results in significant divergence in their route choice behavior. Market segmentation can be well applied to identifying internally homogeneous and externally heterogeneous traveler groups, by investigating the correlation between multiple characteristics of travelers (e.g. gender, interests, location, religion, income, size of household, age, education, occupation, social class, ethnicity, etc.). Trip purposes can be considered as a rough segmentation criteria. However, it should be noted that even for one trip purpose (e.g. leisure travel) traveler behavior still varies due to different underlying value structures. A finer segmentation seems difficult to achieve given the existing data reality. However, some clues can be found to reflect different groups' sensitivity to travel cost in the 1995 ATS. For example, number of stops was recorded, and it can be reasonably assumed that the more stops made by a traveling party, the less sensitive it was towards travel time.

***Identify attributes for each customer base***: through market segmentation several groups (customer bases) will be differentiated. For each customer base, impact factors to route choice can be roughly divided into concrete (or tangible) attributes and abstract (or intangible) benefits, needs, motivations, or personal values. The former category reflects some physical properties of the route. For example, travel time and travel distance are two major factors that influence all the customer bases. Personal characteristics, however, reflect people's perception towards a particular route, determined by the needs, motivations, or personal values carried within one customer base. Influence of personal characteristics can be captured by either deterministic values or random variables. In this step, different attributes for each customer base will be identified.

***Account for different attributes in one generalized cost model***: Link travel impedance measures the resistance or deterrence for a certain link (Chin and Hwang, 2001) and can be depicted as a generalized cost function. The route impedance is then obtained through summation of link impedances. This study will extend the exclusively time-based cost function by involving both concrete and abstract characteristics as aforementioned. After the second step where different sets of attributes are identified, those attributes can be integrated with in a discrete choice model framework, where the concept of utility function can be materialized as travel impedance for the use of traffic assignment in the next step.

***Incorporate the generalized cost model into traffic assignment algorithms***: A diversity of traffic assignment algorithms have been proposed, developed, and practiced since the 1950s. From the simplest All-Or-Nothing (AON) algorithms, to the equilibrium family – Deterministic User Equilibrium (DUE), Stochastic User Equilibrium (SUE), and System Optimum (SO) algorithms, to the more advanced yet usually computationally cumbersome algorithms – Boundedly Rational User Equilibrium (BRUE), Behavioral User Equilibrium (BUE), and Dynamic Traffic Assignment (DTA) algorithms, each algorithm is on the basis of certain behavioral assumptions trying to describe travelers' choices (Zhang, 2011). A working solution in a long distance travel network loading context should achieve a balance between accuracy and feasibility, given the network size, temporal resolution, and computational cost of the problem. This step shall cover detailed equations, algorithms, steps, and theories. Strength and weakness of various approaches shall be discussed. Proper algorithm should be selected to include the impedance function developed towards the next step.

***Comparison of assignment scenarios***: Conducting large scale travel survey for model calibration is beyond the scope of this research; therefore implementation of the proposed modeling framework is mainly for a hypothesis testing purpose. As an experimental effort to explore the long distance travel's distributive pattern under different assumptions, a series of scenarios describing different model specifications will be implemented and compared to the base scenario, where an exclusively time-based impedance function is applied as how most of long distance or large-scale (interstate travel) assignment models were done. Through the implementation several critical questions can be answered:

1) The validity of extension to the exclusively time-based impedance function;
2) The significance and sensitivity of newly proposed parameters;
3) The range and magnitude of those parameters;
4) The change in the distributive patterns caused by different assignment algorithms (Deterministic User Equilibrium, SUE, and System Optimum);
5) System-wise performance measurement (total VHT and VMT, total delay, trip length distribution, etc.)

## 4.2 Theories and Algorithms

Network loading, traffic assignment, and route choice have been referred interchangeably in previous literature; however, the authors would like to differentiate their emphases. Route choice focuses on individual choices, which reflects the interaction between the individual (the traveler profile and traveling awareness) and the surrounding environment (the resources and characteristics of the route); traffic assignment is an aggregated outcome of individual behavior, usually summarized as link flows in an intra-urban context; network loading exhibits the distributive pattern in an interregional resolution, where detailed choice information is abstracted into macroscopic statistics. Overall all the three different terms reflect travelers' choices in different geographic and behavioral dimensions, and obviously this research concentrates on the network loading issues on a national highway network.

This chapter first briefly introduces the discrete choice model, the theoretical basis for route choice; then formulates the Stochastic User Equilibrium-based (SUE-based) traffic assignment, and discusses other traffic assignment approaches based on Deterministic User Equilibrium (DUE) and System Optimum (SO); finally it reviews practices in large-scale network loading, with an emphasis on configuring travel impedance functions.

### *4.2.1 Discrete choice model*

Discrete choice model simulates individuals' choices between competing alternatives. People's valuation about alternative options is described as utilities. Utility is a relative measure representing the attractiveness associated with each alternative, which involves both attributes of alternatives and the decision maker's characteristics. The behavioral foundation for discrete choice model is to assume each decision maker as a utility maximizer. As utility is usually summarized in the format of a function, with both tangible and intangible attributes and characteristics, it is modeled as a combination of a systematic (deterministic) component and an additive random "error term" (Sheffi, 1985), which translates utility into a random variable on the basis of different assumptive distributions. Consequently, discrete choice models only give the probability with which alternatives are chosen, not the choice itself.

To summarize the discrete choice model mathematically, denote the vector $\vec{U}=\{U_1,U_2,\ldots,U_k\}$ as the utilities associated with a set of alternatives $\kappa=\{1,2,\ldots,k\}$. The utility function comprises a systematic component, $V_k(\vec{a})$, where $\vec{a}$ is the vector of alternatives' attributes and individuals' characteristics, as well a random error $\xi_k(\vec{a})$, and

$$U_k(\vec{a})=V_k(\vec{a})+\xi_k(\vec{a}),\forall k\in\kappa \tag{4.1}$$

Usually it is assumed that the probabilistic distribution for the random error satisfies that $E\left[\xi_k(\vec{a})\right]=0$, resulting in that $E\left[U_k(\vec{a})\right]=V_k(\vec{a})$. In this context the error term actually reflects the difference between perceived utility $U_k(\vec{a})$ and measurable utility $V_k(\vec{a})$.

Introducing a probabilistic distribution for the random error in the utility function enables calculation for the probability of a particular alternative $k$ being selected. Based on the utility maximizer assumption, alternative $k$ is superior to any other alternative $j$ only because it is associated with the highest utility, which is

$$\begin{aligned}&P_k(\vec{a})=\Pr\left[U_k(\vec{a})\geq U_j(\vec{a}),\forall j\in\kappa\right],\forall k\in\kappa\\&\Rightarrow P_k(\vec{a})=\Pr\left[V_k(\vec{a})+\xi_k(\vec{a})\geq V_j(\vec{a})+\xi_j(\vec{a}),\forall j\in\kappa\right],\forall k\in\kappa\\&\Rightarrow P_k(\vec{a})=\Pr\left[\xi_k(\vec{a})-\xi_j(\vec{a})\geq V_j(\vec{a})-V_k(\vec{a}),\forall j\in\kappa\right],\forall k\in\kappa\end{aligned} \tag{4.2}$$

Given the distribution of the random error, $\xi_k(\vec{a})$, the joint distribution of $\xi_k(\vec{a})-\xi_j(\vec{a})$ can be specified and consequently $P_k(\vec{a})$ can be calculated explicitly.

Logit and Probit models are widely used discrete choice models. Logit models assume that random errors are independently and identically distributed Gumbel variables. With this assumption the choice probability has a closed form given by

$$P_k = \frac{e^{V_k}}{\sum_{j=1}^{K} e^{V_j}}, \forall k \in \kappa \tag{4.3}$$

The probit models, by contrast, assume that random errors are jointly distributed following a multivariate normal (MVN) density function. The MVN distribution is characterized by a ($K$-length) vector of means, $\vec{\mu}$, and a ( $K \times K$ ) covariance matrix, $\Sigma$, in the notation of $\vec{\xi} \sim MVN(\vec{\mu}, \Sigma)$, where $\vec{\xi} = (\xi_1, \ldots, \xi_K)$. The covariance matrix includes the variances of random errors, and the covariance between them, which is $(\Sigma)_{kk} = \mathrm{var}(\xi_k), \forall k$ and $(\Sigma)_{kj} = \mathrm{cov}(\xi_k, \xi_j), k \neq j$.

Since the MVN distribution conserves its shape under linear transformation, given a covariance matrix (associated with the randomness in the perception of a set of alternatives) and a vector of known alternatives' attributes, the distribution of the utility vector, $U(\vec{a}) = [U_1(\vec{a}), U_2(\vec{a}), \ldots, U_k(\vec{a})]$, can be modeled as multivariate normal; in other words, $U(\vec{a}) \sim MVN(V(\vec{a}), \Sigma)$. With probit models choice probability $P_k(\vec{a})$ cannot be expressed analytically since the cumulative normal distribution function cannot be evaluated in closed form. Therefore the calculation is usually conducted using either an analytical approximation or a Monte Carlo simulation.

So far the discussion about discrete choice models is at individual level, and how can we aggregate choice predictions? One analytical solution involves capturing the distribution of $\vec{a}$ among the population using a density function $f(\vec{a})$ and calculating the share of the population who selects alternative k as $\bar{P}_k = \int_{\vec{a}} P_k(\vec{a}) f(\vec{a}) d\vec{a}$. While this multiple integral cannot be easily estimated, a practical solution is to divide the population into groups of similar values of $\vec{a}$, computing $P_k(\vec{a})$ for each group, and averaging the results. Alternatively, $P_k(\vec{a})$ can be computed by Monte Carlo simulation as well.

Either at individual level or aggregated level, a variety of discrete choice models has been applied, with variations and extensions to the basic logit and probit models, to a broad spectrum of areas. Some examples include: logit with attributes of the person but no attributes of the alternatives; logit with variables that vary over alternatives (also called conditional logit); nested logit and generalized extreme value (GEV) models; multinomial probit; mixed logit (Train, 2009; Sethi and Koppelman, 2001; Train, 1998; Bierlaire, *et al.*, 2006; Dafermos, 1972; Hensher and Rose, 2006; Huang and Gao, 2012; Taplin and Qiu, 1997). One of the important applications is traffic assignment, which will be detailed in the following section.

### *4.2.2 Formulate traffic assignment in a discrete choice framework*

As the aggregated outcome of individual route choice behavior, traffic assignment yields link flows over a network. Since traffic assignment is essentially a reflection of the interaction between individuals and the network, traffic assignment algorithms can be categorized as equilibrium and non-equilibrium approaches, depending on whether the approach considers capacity-constrained flow (the volume-delay dependency). For example, the All-Or-Nothing (AON) algorithm simply assigns all the traffic volumes to the shortest path between an OD pair, which is a non-equilibrium approach. Obviously equilibrium approaches are closer to the reality and are more widely applied in practice.

Prevailing equilibrium approaches include Deterministic User Equilibrium (DUE), Stochastic User Equilibrium (SUE), and System Optimum (SO), each on the basis of a certain behavioral assumption. The SO method portrays an ideal scenario by assuming there were a central controller for the network coordinating all the travelers to minimize the total travel time; the DUE method assumes every traveler on the network has a consistent understanding of the traffic situations and perfect information about the network; when equilibrium is reached, no one can improve his/her travel time by switching to alternative routes, and all the utilized routes for an O-D pair have the same minimum travel time. SUE method (Daganzo and Sheffi, 1977) discards this unrealistic assumption by recognizing travelers' perception error and modeling the error with randomness. At SUE, no travelers *believe* that they can save travel time by choosing a different route.

Discrete choice models can be well applied to formulate the SUE mechanism, which better simulate real-world scenarios than DUE and SO methods. The choice set under consideration is a number of alternative routes connecting a certain O-D pair; each alternative route is associated with some travel time as the utility. Due to the variations in perception and exogenous factors, the travel time for the same route is perceived differently by each traveler. While the actual travel time can be measured as a systematic component, the difference between people's perceived travel time and actual travel time can be modeled as a random error distributed across the traveling population. The SUE problem is to determine the probability of each route being utilized, based on the random error distribution and the volume-delay dependency. Then the link flows can be obtained from the O-D demand and estimated probability.

To formulate the SUE mechanism in a mathematical form, the utility function $U_k\left(\vec{a}\right)$ in the discrete choice model, is then translated into the travel impedance functions $C_k^{od}$ (generalized travel cost) on route *k* between origin *o* and destination *d*, where $\forall k \in \kappa$. $C_k^{od}$ is a random variable. There are two components constituting $C_k^{od}$: the measurable travel impedance $c_k^{od}$ and the random error term $\xi_k^{od}$, whose distribution reflects the difference of perceiving travel time across the traveler population.

$$C_k^{od} = c_k^{od} + \xi_k^{od}, \forall k, o, d \tag{4.4}$$

If the population of travelers between *o* and *d* is large, the share of travelers choosing the *k*th route, $P_k^{rs}$, is given by

$$P_k^{od} = \Pr(C_k^{od} \le C_l^{od}, \forall j \in \kappa_{od}), \forall k, o, d \tag{4.5}$$

In other words, the probability that a given route is chosen is the probability that its travel impedance is perceived to be the lowest of all the alternative routes. A number of algorithms have been proposed for solving SUE problems, and the method of successive averages (MSA) is applied in this research (Sheffi, 1985).

### 4.2.3 *A review on large-scale network loading*

So far very little literature can be found relevant to the traffic assignment at an interregional scale. One of well-established existing works is the Freight Analysis Framework 3 ($FAF^3$) completed by the Oak Ridge National Laboratory (ORNL). The assignment for freight traffic was accomplished by SUE traffic assignment procedure in TransCAD 5.0 with user defined volume-delay function (VDF) (Battelle, 2011). The assignment is constrained by the highway network's current capacity. The selected VDF for FAF assignment is the Bureau of Public Roads (BPR) function and it is an exclusively time-based impedance function. The highlight is ORNL conducted extensive data cleansing and imputation, which laid a quality foundation for a routable network. Due to the network size and inconsistency between the truck flow O-D data source and the baseline traffic data (HPMS truck volume), calibrating the $FAF^3$ model experience considerable difficulty. The calibration effort only involved adjusting the link travel time or capacity or the parameters in the BPR functions so that the assigned link flows were as close as possible to the baseline flows.

On the passenger travel side, Chin and Hwang (1999) assigned 1995 ATS highway trips to the Oak Ridge National Highway Network based on AON algorithm. ATS auto trips were routed between centroids of the origin and destination zip codes. Similarly to the FAF3, the impedance function was also exclusively time-based, where the travel time is determined by distance and speed limit adjusted by the physical and functional characteristics of the road. Some criticism regarding this work include: first, over-simplified highway network, the final assignment was represented by a match-stick flow map, not explicitly with the corresponding network; second, the impedance function did not consider the "attractions" of some roadway sections to the pleasure travel; third, the assignment algorithm is “all-or-nothing” (shortest path), which further lost some validity of the assignment results.

To summarize, the network loading issues for long distance travel can be resolved by existing traffic assignment algorithms, following similar procedures widely practiced in the intra-urban traffic assignment. SUE is selected as the algorithm for this study, mainly for the following reasons: first, this study focuses on annual travel patterns and does not require a finer temporal modeling resolution; second, SUE models are associated with well-established calculation procedures with converging solutions; third, SUE models are widely applied and successfully implemented in similar large-scale studies. In the next Chapter, travel impedance in long distance travel will be specially discussed, since it is the travel impedance that distinguishes the long distance travel network loading from the intra-urban traffic assignment, and also captures the behavioral difference across traveler markets.

## 4.3 Long Distance Travelers

### *4.3.1 Overview*

Route choice, as a salient feature in people’s travel behavior, reflects people’s underlying values and needs. Compared to intra-urban travel, long distance travel embraces a more significant behavioral divergence. This diversity can be well captured by the change in travelers’ sensitivity to travel impedance when trip purposes vary. Therefore it is logical to segment the entire long distance traveling population by trip purpose for separate analyses. The 1995 ATS collected long distance trips in 12 sub-categories, as shown in Table 4.1. In this study the research team generalized all the sub-categories into three main purposes: business, pleasure, and other. This categorization not only simplifies further analyses without obfuscating the differentiations and similarities between sub-categories, but also keeps the analyses consistent with partner research teams (Viswanathan and Vary, 2012).

Table 4.1Trip Purpose Definition and Distribution – Only Auto Trips (1995 ATS)

| Purpose | Analysis Purpose | Frequency | Percent |
|---|---|---|---|
| Business | Business | 125,664,045 | 15% |
| Combined Business/Pleasure | Business | 18,527,520 | 2% |
| Convention, Conference, Or Seminar | Business | 8,556,234 | 1% |
| Visit Relatives Or Friends | Pleasure | 285,404,623 | 35% |
| Rest Or Relaxation | Pleasure | 98,461,499 | 12% |
| Outdoor Recreation (Sports, Hunting, Fishing, Boating, Camping, Etc.) | Pleasure | 59,320,149 | 7% |
| Entertainment (Attend The Theater Or Sports Event, Etc.) | Pleasure | 45,744,170 | 6% |
| Sightseeing, Or To Visit A Historic Or Scenic Attraction | Pleasure | 34,313,300 | 4% |
| Shopping | Pleasure | 16,811,484 | 2% |
| Personal, Family, Or Medical (Wedding, Funeral, Health Treatment, Etc.) | Other | 106,614,688 | 13% |
| School-Related Activity | Other | 19,108,160 | 2% |
| Other Reason | Other | 30,798 | 0% |

While the business/pleasure/other categorization portrays a rough segmentation of the long distance travelers, it should be recognized that within each category there still exists a very diverse traveler portfolio. This is especially true for the pleasure travel, where numerous researchers endeavored to explain tourist behavior by developing tourist typologies. For example, Cohen (1972) suggested four general types of tourists: the organized mass tourist, the individual mass tourist, the explorer, and the drifter. Perreault *et al.* (1977) segmented the tourist population by their travel inclination into the budget travelers, adventurers, homebodies, vacationers, and moderates. Stewart (1993) focused on the holiday travel and summarized four phases of holiday taking: the bubble travelers, idealized-experienced seekers, wide-horizon travelers, and total immersers.

A finer segmentation of the long distance traveler group requires extra data collection (e.g. stated preference survey) and analysis efforts, which is beyond the scope of this study. Therefore a dichotomy of business travel vs. pleasure/other travel is adopted as shown in Table 2.1, which is adequate to meeting needs for demonstrating the varying distributive patterns of different trip purposes over the national highway network.

The remaining part of this Chapter will elaborate efforts of describing characteristics of two major customer bases: business travelers and pleasure/other travelers. Exploration for each customer base will be carried out in two dimensions: the concrete attributes travelers would consider when selecting routes, as well as abstract benefits, needs, motivations, or personal values within the traveling population. Understanding those characteristics builds a foundation for making assumptions when developing travel impedance functions.

### *4.3.2 Business travelers*

A business trip is defined as "*any trip where the purpose of the trip is given as business, combined business with pleasure, or convention, conference or seminar*" (BTS, 1997). It is attributable to 18% of all long distance trips in the 1995 ATS. Contrary to the stereotypical image, very few business travelers make cross-continent trips. According to the 2001 NHTS, 84% of business trips were within census region boundaries, which were less than 250 miles in length. Males in their 30s and 40s who work in a professional, managerial, or technical position are the dominating population in the business travel group (BTS, 2003). The personal vehicle is the dominant travel mode for business trips.

Travel distance and travel time are two controlling impact factors on route choice. In general business trips are made with constrained time windows (e.g. travelers meeting clients at a specific time, attending conferences, etc.), and by people with above-average income level Additionally although some firms continue to provide vehicles to full-time business travelers, a growing number of employees use their own vehicles for company business and receive reimbursement for travel cost. Therefore compare to other customer bases, business travelers are less sensitive to monetary cost of travel like tolls or driving cost but prefer to choose route with shorter travel distance/time and higher traffic reliability (usually with roadways with higher functional class). Moreover, business travelers valuate travel time more than pleasure travelers since the opportunity cost for travel is associated with productive working hours. Both travel distance and travel time can be quantified as tangible network attributes, which will be detailed in Chapter 4.

### *4.3.3 Pleasure travelers*

Pleasure trips refer to "*any trip where the purpose of the trip is given as visiting friends or relatives, rest or relaxation, sightseeing, outdoor recreation, entertainment, or shopping*" (BTS, 1997). As summarized in Chapter 1, pleasure travelers account for a substantial percentage of the entire long distance traveling population, and the travel demand shows resilience in a nose-diving economy, being expected to keep growing in the following years.

Although compared to business travelers, pleasure travelers are less sensitive to the monetary travel cost but rather value pleasure components along the route in general; the diversity of travel behavior within this customer base should be recognized. For example, after retrieving the leisure travel subset from the 1995 ATS samples, Nostrand *et al*. (2011) found that the elderly (65 or older) and lower-income households made fewer leisure trips. Also leisure travel presented a “variety-seeking” property, meaning that people tend to visit multiple destinations per year rather than the same destination for multiple times. Obviously lower income households would be more sensitive to additional travel distances than higher income travelers; the middle age group inclines to travel shorter distance due to professional or family obligations than younger (< 25 years) and older (> 64 years) age groups.

The diversity can be partially attributed to route characteristics : route proximity, route attraction, safety, and accessibility rate as summarized by Javaheri (2011). It is also caused by the travelers’ personal attributes. In a more particular case study, Kemperman *et al*. (2009) described and predicted tourist shopping route choice behavior in a downtown historic center. Some impact factors they identified include: shopping motivations, familiarity with the area, and planning of the route. Demographic heterogeneity is considered as a general impact factor. Anable (2002) concluded that “*increases in disposable income and demographic factors such as an aging population with decent incomes, abundant leisure time, and increasing confidence to travel are some of the more direct and obvious factors influencing the form and structure of leisure travel*”.

Newman (2001) argued that "*retirees tend to view road trips as an adventure. They are more relaxed, willing to go at a slower pace, and spend more time exploring. … Young travelers, with or without children, resent being in the car over long periods of time and just want to get there*". While Schneider and Vogt (2005) recognized household composition's impacts on recreational activities, Lanzendorf (2002) indicated that "*orientations, lifestyles, and mobility styles*" are the essential determinants and "*including orientations or styles is useful for explaining travel behavior*".

It is also believed that the form and structure of pleasure trips is essentially a collective reflection of varying value structures across the traveling population. Cho and Jang (2008) identified five value dimensions through an extensive review of the literature: ***utilitarian***, ***risk avoidance***, ***hedonic***, ***sensation seeking***, and ***social***. ***Utilitarian*** people constantly try to accomplish their goals and avoid undesirable outcomes. ***Risk avoidance*** people justify their choices by maximizing the risks to avoid the anxiety induced. ***Hedonic*** people are consumption-oriented and seek for the entertainment value of their choice. ***Sensation seeking*** people are willing to "*take physical, social, legal, and financial risks*" to achieve sensational arousal. Social interaction and communicative features determine ***social*** people's behavior. For example, involving families and friends has a significant impact on their decision making process.

In summary, long distance route choice reflects people's behavior under combined effects of route characteristics and personal characteristics of the travelers. Attributes enumerated in this chapter are certainly not exhaustive, and it should be realized that while route characteristics can usually be measured physically, quantifying personal characteristics needs to reply on behavioral interviews and surveys. Means-end analytic approach can be readily applied to identifying those intangible characteristics; and discrete choice model can be used to estimate the equivalent monetary value of personal characteristics. Both efforts require large-scale data collections to draw meaningful conclusions, which are beyond the scope of this study; however, they will be discussed as further steps in Chapter 5.

## 4.4 Impedance Models

As shown in the last Chapter, the SUE mechanism can well serve the needs for loading long distance trips to the national highway network. The pivot in resolving the issue lies in well-defined impedance functions that not only capture the behavioral attributes and route characteristics of different traveler markets, but also can be readily integrated the SUE mechanism as a working solution. This chapter first elaborates a generalized cost function, which extends the exclusively time-based impedance used by previous studies, and then introduces two new parameters accounting for the intangible attributes in pleasure travel; finally this chapter highlights the process of capturing and integrating the impedance on centroid connectors.

### *4.4.1 Generalized cost function for travel impedance*

Selecting a route between an O-D pair involves a process of evaluating the travel impedance of each possible route. Impedance quantifies the “cost” or penalty for using any link along a route. A critical question to be answered by configuring the impedance functions involves what impact factors the function should represent, and to what extent these impact factors influence the likelihood of selecting the route. Answering this question for business travel is fairly straightforward, based on the assumption that long distance travelers for business are seeking for the "shortest path," both temporally and spatially, to minimize travel time and travel distance, as their travel is usually time sensitive and constrained. Also, those travelers consider the fixed cost associated with certain routes or modes (e.g. road tolls; airfare; Amtrak ticket prices, etc.). Therefore, in this study, the exclusively time-based travel impedance is extended to ***composite impedance*** accounting for fixed cost, travel distance, and travel time. All the impact factors are translated into monetary values to formulate a unified impedance measure, as shown in Equation 4.6.

$$c_i = k_i + \delta L_i + \varphi t_i \left[ 1 + \alpha_i \left( \frac{x_i}{C_i} \right)^{\beta_i} \right] \tag{4.6}$$

where:

$c_i$ = Generalized cost on link *i*

$k_i$ = Fixed cost on link *i*

$\delta$ = Constant such as the operating cost per unit of length

$L_i$ = Length of link *i*

$\varphi$ = Constant representing the value of time

$t_i$ = Free-flow travel time on link *i*

$\alpha$ = Constant

$x_i$ = Flow on link *i*

$C_i$ = Capacity of link *i*

$\beta_i$ = Constant

The impedance function in Equation 4.1 summarizes the BPR delay function, a fixed cost component, and a distance-based operating cost as a generalized travel cost model in the unit of dollars. Each component is further explained as follows:

#### 4.4.1.1 Fixed cost

$k_i$ represents the fixed cost. In a highway travel context, it mainly involves road tolls. Congestion pricing has been more and more practiced in the U.S. as a countermeasure against urban congestion, as well as a critical means of revenue collection in some jurisdictions. Currently most of the states in the U.S. with toll facilities are collecting roadway tolls at flat rate or time-of-day dependent rate along the full-span of the tolling facility, therefore some aggregation of links in the base network is necessary to accurately reflect the monetary cost of using a specific facility. The International Bridge, Tunnel and Turnpike Association (IBTTA) summarizes and updates toll rates for major tolling facilities periodically. It is assumed that fixed cost will be counted as a component of pleasure travel impedance but not in business travel, considering the fact that most business travelers are getting reimbursed from their employers and therefore insensitive to the fixed cost.

#### 4.4.1.2 Operational cost

$\delta$ represents the average operational cost per mile for driving. The American Automobile Association (AAA) has been published *Your Driving Cost* since 1950 (that year driving a car

10,000 miles cost 9 cents a mile, and gasoline sold for 27 cents per gallon). The process used to estimate annual driving cost is proprietary to AAA therefore the calculation formula is not publicly accessible. However, it is know that this process does consider the cost of fuel, maintenance, tires, insurance, license, registration and taxes, depreciation, and finance. National passenger vehicle composition data is retrieved from the Bureau of Transportation Statistics (BTS), and in 2008 there are 196,762,927 sedans and 39,685,228 minivans and SUVs. Based on this data, the estimated operational cost is 74.4 cents per mile for urban travel and 48 cents per mile for rural travel in 2008.

#### 4.4.1.3 Monetary cost for travel time

The third component converts flow-dependent travel time generated by the BPR function to monetary cost using the value of time (VOT) represented by $\varphi$. In transport economics, the VOT is the opportunity cost of the time that a traveler spends on his/her journey. In essence, it is the amount of money that a traveler would be willing to pay in order to save time, or the amount they would accept as compensation for lost time. The VOT varies by trip purpose, as well as by urban/rural roadway type.

Although there is extensive literature about VOT estimation, fewer people investigated VOT in a long distance travel context. Maki *et al*. (2007) estimated the VOT using multinomial logit (MNL) and mixed logit (MXL) models based on both revealed preference and stated preference data in Japan. Their contribution was to show the practical validity of VOT estimated using a route and modal choice model for both intra-city and inter-city trips, also by different trip purpose. Although the VOT estimated cannot be directly translated into a U.S. context, their major findings are inspirational: 1) for intra-city trips business travel VOT is lower than pleasure travel VOT; for inter-city trips, the situation is reversed with higher business travel VOT; 2) inter-city trips VOT is generally higher than intra-city trips VOT, with VOT increasing with longer travel distance. Tsukai and Okumura (2003) also made a very interesting comparison between business travel and non-business travel by rail. "*According to t-value, most important LOS factor for business passenger is line-whole trip time, frequency is second, and fare is not so important. However, sight-seeing & private passengers consider fare the most important, then secondly line-whole trip time, and lastly frequency. Value of time calculated from time and fare parameters are 3,448 yen/hour (29.5 $/hour) for business trip, and 2,983 yen/hour (25.5 $/hour) for sightseeing & private trip.*"

In the U.S., the relativities of VOT between business travel vs. pleasure travel, and intra-city trips vs. inter-city trips are very similar to the research findings in Japan. Constraints on the total time available (e.g., school vacations or allowed leave) also increase the business travel VOT, therefore pleasure travel VOT is generally estimated lower than business travel VOT and associated with household income. Besides distinctions based on transportation mode and trip purpose (business or pleasure), a major source of variation in the VOT is the large differences between local and intercity trips. Because intercity travel is usually consumed jointly with expensive services such as hotel rooms, restaurant meals, and entertainment, travel time saved is freed for purposes that travelers value highly. Intercity travel time is, therefore, likely to be more valuable than time spent on local travel.

VOT is estimated based on traveler's wages (for business travel) or household income (for pleasure travel), given its nature of opportunity cost: when a trip is undertaken during work or when the traveler is free to vary his or her work hours, an important measure of the value of time is the wage paid for the productive work that is sacrificed to travel. The U.S.DOT recommended travel time factors based on Employer Cost for Employee Compensation (ECEC) are summarized in Table 4.2 (U.S.DOT, 1997).

Table 4.2 U.S.DOT Recommended Value of Time Factors

| Time Component | Reference | Value |
|---|---|---|
| In-Vehicle Personal (local) | Of wages | 50% |
| In-Vehicle Personal (intercity) | Of wages | 70% |
| In-Vehicle Business | Of total compensation | 100% |
| Excess (waiting, walking, or transfer time) | Personal Of wages | 100% |
| Excess (waiting, walking, or transfer time) Business | Of total compensation | 100% |

For business travel, the ECEC figures are supplied by the Bureau of Labor Statistics. For pleasure travel (personal travel by surface modes), the standard adopted is the median annual household income, as reported by the Bureau of the Census, divided by 2,000 hours. Based on those references, VOTs in 2008 are determined as shown in Table 4.3.

Table 4.3 Value of Time by Trip Purpose and Traveling Area in 2008

| Trip Purpose | Business | | Leisure/Other | |
|---|---|---|---|---|
| | Urban | Rural | Urban | Rural |
| VOT in $/hr | 29.18 | 29.18 | 12.58 | 17.61 |

BPR function represents the penalty (increase) in travel time when congestion builds up. Free flow travel time is first calculated by dividing the link length by the reference speed limits

(Hwang and Rollow, 2000). Then $\alpha_i$ and $\beta_i$ are introduced to reflect the reaction of different types of highway facilities to the increasing traffic volume in terms of travel delay. In general $\alpha_i$ and $\beta_i$ represent a linear and an exponential increase respectively. While $\alpha_i$ penalizes the entire travel time increase curve, $\beta_i$ penalizes the part of curve with higher v/c ratio (meaning a more steep increase when v/c ratio increases). Reference values of $\alpha_i$ and $\beta_i$ (Martin and McGuckin, 1998; Fitzpatrick *et al.*, 2003), together with speed limits by roadway functional class, are summarized in Table 4..

Table 4.4 BPR Coefficients

| FCLASS | Note | Speed Limit | Alpha | Beta |
|---|---|---|---|---|
| 1 | Rural Interstate | 70 | 0.88 | 9.8 |
| 2 | Rural Principal | 55 | 0.83 | 2.7 |
| 6 | Rural Minor Arterial | 45 | 0.71 | 2.1 |
| 11 | Urban Interstate | 60 | 0.83 | 5.5 |
| 12 | Urban Other Freeways | 55 | 0.56 | 3.6 |
| 14 | Urban Principal | 35 | 0.15 | 4 |
| 16 | Urban Minor Arterial | 25 | 0.15 | 4 |

Variations of $\alpha_i$ and $\beta_i$ by functional class reflect the fact that even though dominated by time, impedance also includes several other considerations and adjustments. For instance, it is commonly believed that travelers will tend to use freeways and interstates even when doing so will result in a route that is slightly longer in time than an alternative surface route. This may be due to considerations of safety, ease of travel, or lack of familiarity with local roads. Therefore, congestion on interstate freeways is punished more significantly than other types of roadways.

Another critical assumption in long distance network loading is that the "impedance" or "resistance" resulted from congestion (represented by $x_i$ in Equation 4.1) is not from the long-distance trips, because it is really only a small proportion (less than one percent as shown in Figure 1.1). The interaction between long distance trips and non-long distance trips is negligible. Therefore the link travel time is basically determined by the general traffic (preloaded volume). In this case, other factors (e.g. travel distance, tolls, roadside scenery) are playing a more significant role in the network loading.

### 4.4.2 *Travel impedance for pleasure trips*

Except for the tangible attributes captured by the impedance function developed in the last section, there are many other intangible attributes that play critical roles in the long distance network loading process, especially for the pleasure travel. Some intangible attributes include: traveler socioeconomic profiles, guided and unguided travelers, typologies of tourists, travel awareness, confidence in travel intermediaries, perceived risk and uncertainty of travel, domestic pressures, trip pressure, political, economic, and social value structure, etc. Travelers' evaluation on the "attractiveness" of an alternative route is essentially a comprehensively external representation of those internally intangible attributes. Since it is not realistic to enumerate all the impact factors and quantify their impacts, this study proposes a framework to accounting for travelers' perceptions and behavior using ***scenic byways*** and ***roadside attractions*** as two case studies.

#### 4.4.2.1 Scenic Highways and Roadside Attractions

Scenic highways and roadside attractions are categories of impact factors on pleasure travel route choice. Although people's needs, attitudes, and motivations are intangible and hard to quantify, people's reaction to those attractions (pull and push factors) can be captured and simulated to some extent. For example, a hedonic traveler may seek for more scenic routes by sacrificing travel time; and a utilitarian traveler tends to choose the route with less travel cost. Before introducing new parameters to quantify those reactions, it is necessary to study the roles of scenic highways and roadside attractions in influencing travelers' behavior.

Scenic highways are frequently used for pleasure travel. The 1991 Intermodal Surface Transportation Efficiency Act (ISTEA) established the National Scenic Byways Program within FHWA, who designate two types of byways: National Scenic Byways and All-American Road (Eby and Molnar, 2002). Currently there are 150 byways scattered across 46 states in the U.S. The designation entails multi-dimensional criteria: ***archaeological qualities*** (i.e., physical evidence of historic or prehistoric life), ***cultural qualities*** (i.e., evidence and expressions of the customs or traditions of a distinct group of people), ***historical qualities*** (i.e., legacies of the past that are distinctly associated with physical elements of the landscape), ***natural qualities*** (i.e., features in the visual environment that are relatively undisturbed), ***recreational qualities*** (i.e., outdoor activities directly associated with the natural and cultural elements of the byway corridor), or ***scenic qualities*** (i.e., heightened visual experience derived from the viewing of natural and manmade elements). Being designated reflects an overall evaluation of a roadway's attraction.

Travelers on scenic highways present diverse demographics distinct from users of non-scenic highways. Multiple researchers have identified the popularity of scenic highways in elderly populations compared to younger people (Briganti and Hoel, 1994; Dahlquist and Peterson, 1997; Sem *et al*., 1997). In an economic-impact study of Colorado scenic and historic byways, Sem *et al*. (1997) also found that scenic highways were more frequently used by higher income people than other groups. In a U.S. survey of the driving tourist's information needs and preference, Eby and Molnar (2003) also identified age and household income were significantly influencing people's rating of scenic highways, as well as that scenic highways are weighed with higher importance in vacation travel than other trip purposes. Briganti and Hoel (1994) also suggested that compared to non-scenic highways, scenic highways attracted a higher number of first-time drivers and a more diverse fleet. Considering the demographics of pleasure travelers as discussed in Chapter 2, all of the previous work pointed to the strong bond between scenic highway and pleasure travel.

Research also showed that people had a positive "willingness to pay" towards scenic byways. Tyrrell and Devitt (1999) conducted a stated preference survey to identify traveler types and their byway preferences. The survey investigated diverse roadway characteristics: distance view (forest, farm & fields), scenic designation, roadside view, speed limit, roadside rest areas, shoulders, toll cost per trip, by asking respondents to choose two specifically defined roadway scenarios. Despite the small and biased sample used in this study, results showed respondents were willing to pay approximately $0.77 per trip (1997 Dollar) to use a roadway that is designated as "scenic".

It should be recognized that compared to the tangible route characteristics such as travel distance and travel time, scenic factors are of secondary importance when travelers select routes (Eby and Molnar, 2003). Therefore it is reasonable to assume that the route choice in a long distance context is prone to compounding effects of the generalized travel cost and entertaining or pleasure factors and the impacts of the latter are generally more significant in pleasure travel.

While the scenic highway designation represents the attractiveness of the route itself, the proximity of the route to nearby attractions also hugely influences route choice, including closeness to tourist attractions (national/state parks, ski resorts, etc.) as well as closeness to the tourism facilities (hotels, etc.). This route proximity can be roughly captured by the number of attractions within certain distance of the roadway section. For example, a route with more accommodation facilities are more attractive to pleasure travelers, considering the fact the many pleasure trips involves multiple nights of stay along the route. Note that there exist correlation between route proximity and route attraction, scenic highway effects and roadside attraction effects are investigated separately in this study, and integrated to an adjusted travel impedance function.

#### 4.4.2.2 Integrating Scenic Byways and Roadside Attractions into Travel Impedance

Route proximity and route attraction effects are accounted for by introducing two new parameters to "discount" the generalized cost function, as shown in Equation 4.6. The behavioral assumption for this specification is that both route proximity and route attraction effects (when they exist) increase the attractiveness of the route and in turn decrease the perceived travel impedance for the route. For example, if a highway section is a scenic byway, the measured travel cost (derived from travel time and travel distance) will be multiplied with a percentage to make the section more “attractive” to pleasure travelers.

$$c_i = s_i \cdot e^{-a_i n_i} \left( k_i + \delta L_i + \varphi t_i \left[ 1 + \alpha_i \left( \frac{x_i}{C_i} \right)^{\beta_i} \right] \right) \tag{4.7}$$

where:

$s_i$ = scenic factor for route link $i$ that is part of scenic highways, $s_i \in (0,1]$;

$n_i$ = number of hotels or national/state parks within 5 miles;

$a_i$ = attraction factor for route link i;

All the other parameters and variables have the same meaning as in Equation 4.6.

Scenic factor $s_i$ and attraction factor $a_i$ are modeling parameters introduced as aggregated measurements across the traveling population for pleasure purposes. They are designed to possess several beneficial properties from a behavioral modeling point of view:

1), valuation of $s_i$ differentiates scenic highways and non-scenic highways: for non-scenic highways, $s_i = 1$; for scenic byways $s_i \in (0,1)$, which indicates that when using a scenic route, the generalized travel cost is discounted to make the route itself more attractive to pleasure travelers.

2), both parameters are within the interval $(0,1]$, which bounds possible scenarios to plausible ranges;

3), both parameters maintain the positivity of the adjusted travel impedance and are able to discount the measured impedance effectively to reveal attractiveness;

4), the multiplier $e^{-a_i n_i}$ decreases monotonically with an increasing number of roadside attractions;

5), for the attraction factor $a_i$, when $n_i = 0$, the multiplier $e^{-a_i n_i}$ goes to 1 and there is no discount effect, which is consistent with reality.

6), the multiplier $e^{-a_i n_i}$ is designed with the capability to entail different traveling behavior. As shown in Figure 4.2, when $a_i < 0.1$, the discounting effect has a ***quasi-linear*** decrease with an increasing $n_i$, meaning people have a steadily changing attitude towards roadside attractions; when $0.1 < a_i \le 1$, the discounting effect has an ***exponential*** decrease with an increasing $n_i$, reflecting the behavior that people have a strong reaction to a "with or without" roadside attractions scenario, yet a less sensitive reaction to when there are many attractions along the route (e.g. when $a_i = 0.6$, an increase of attractions from 1 to 2 will cause an extra discount in the travel impedance of 24%, while an increase of attractions from 11 to 12 will only cause an extra discount of 1.3%).

In summary, while Equation 4.6 extends the exclusively travel time-based impedance function by applying a generalized cost function, Equation 4.7 presents a novel way to count for the compounding effects of contextual factors and the generalized cost, by introducing two new parameters. In the next Chapter, distributive patterns under different magnitudes of scenic factors and attraction factors will be presented.

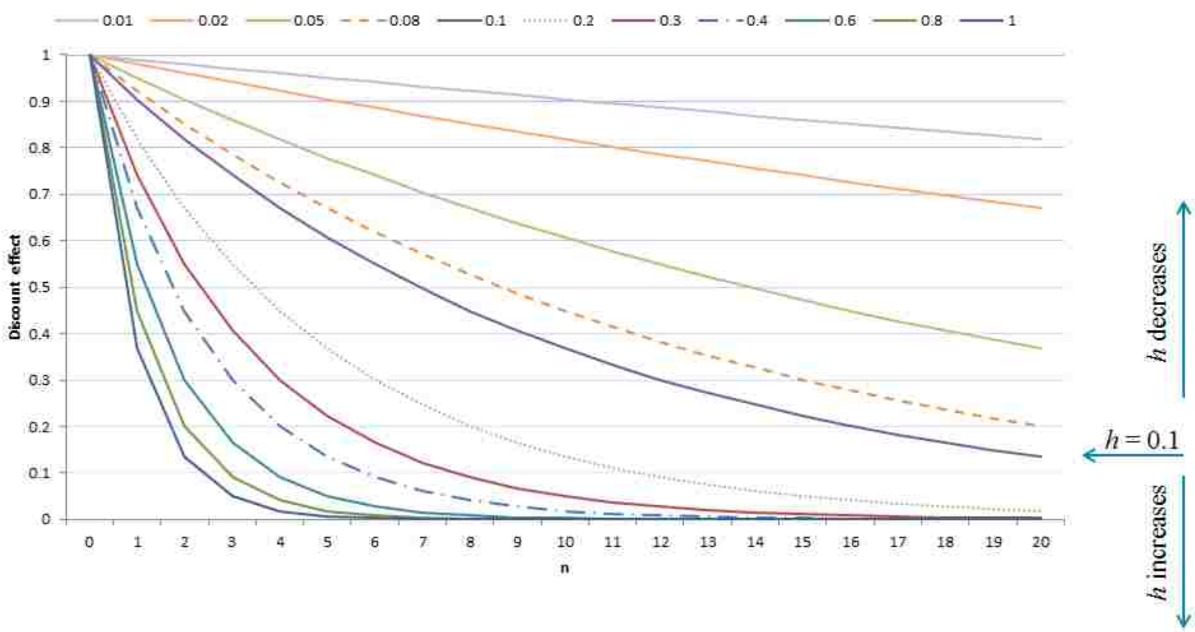



Figure 4.2 Attraction Factor Reflects Different Behavior

### *4.4.3 Impedance on centroid connectors*

So far we have integrated both tangible and intangible attributes into the travel impedance function, associated with each and every link on the network. However, there is another type of link in the network that is of critical importance to the modeling process. The centroid connector is the ***virtual link*** that connects a centroid to the rest of the network. The travel impedance on this virtual link should represent the generalized travel costs when traveling on local streets that are not explicitly included in the network. However there is very little literature theoretically documenting the process of generating centroid connectors or evaluating the impacts of placement and number of centroid connectors on the assignment results (Qian and Zhang, 2012). The common practice is to assign a constant travel time to the connector, the magnitude of which depends on the speed limit of the roadway link the connector it connects to.

To be consistent with the monetary cost developed in this study, the impedance on centroid connectors are determined by the *Annual Congestion Cost per Auto Commuter* calculated by TTI, as part of their congestion measures. The measure is determined by the value of lost time in passenger vehicles in congestion and the value of wasted fuel due to congestion. It is a fairly comprehensive reflection of how much travelers have to pay for living in the urban areas. The annual cost should be divided by 365, based on the assumption that this simulates the cost for a long distance traveler needs to pay on an average traveling day. For example, in 2008 the annual congestion cost per auto commuter in Seattle is $994, which indicates on an average traveling day a traveler would pay approximately $2.7. If we consider an origin or destination urban area as a node in the network, then both the node impedance and link impedance are unified under a monetary cost framework.

Although specifying node impedance by applying the congestion cost to centroid connectors facilitates in a unified monetary cost framework, some limitations from this specification should be recognized. First, currently the centroids are set up based on physical area of counties, rather than population based; second, due to the annual temporal resolution of modeling, seasonality of long distance travel cannot be reflected. Third, number and locations of centroid connectors are set up arbitrarily. Therefore it is worth exploring the impacts of allocating centroid connectors in future research.

In summary, this chapter first considers some tangible attributes in a generalized cost function, which extends the exclusively time-based travel impedance in previous studies. It then develops two factors reflecting people's behavior towards scenic highways and roadside attractions as a discounting effect to the generalized cost function, providing a solution for accounting for the pull and push factors in pleasure travel. Finally it translated one of the TTI's performance measurements into the monetary cost framework, which unifies the node impedance and link impedance. All the impedance components are placed on a monetary cost metric, which not only facilitates a smooth modeling process, but also brings insights from a cost analyses perspective. As shown in the next Chapter, the proposed framework will be implemented on a large-scale network, and different scenarios will be tested.

# Chapter 5 Long Distance Travel Network Loading: Case Study

This chapter demonstrates the implementation of the aforementioned models on a real-world large-scale network. The case study illustrates the network loading process of the long distance passenger travel within west-coast states of ***Washington***, ***Oregon***, and ***California*** at county level, using projected ***county-level*** long distance O-D demand data of year 2008. The temporal resolution of modeling in the case study is annual. Objectives of the case study are to validate the feasibility of proposed work flow, test the proposed hypotheses, investigate the sensitivity of newly introduced parameters, and reveal different distributive patterns under varying scenarios.

The case study is not for model calibration or validation purposes. Unlike the intra-urban traffic assignment, there is no well-established modeling procedure for the long distance travel network loading problem yet. As it is still in an initial modeling stage, there are strong needs for clarifying the model specification and testing parameter effectiveness, to explain the dynamics of the issue, discover new questions, and guide future data collection for calibration and validation. Some questions to be answered by this case study include: 1) is the proposed work flow feasible? 2) Are the newly introduced parameters significant and sensitive? 3) How do various assignment algorithms influence the distributive pattern? 4) How do variations of a model parameter produce different travel patterns? 5) Why and to what extent does the estimated distributive pattern matter in terms of policy and investment decision making?

In following sections, data acquisition is briefly introduced, followed by a specification of testing scenarios, each corresponding to the introduction or the variation of a model parameter; the implementation results are then presented before summarizing indications and further thoughts of the case study.

## 5.1 Data Acquisition

### *5.1.1 O-D demand data*

The 2008 long-distance auto trip O-D table at county level is provided by one of the major research partners, the Wilbur Smith Associates (WSA). This 2008 O-D matrix is populated from the 1995 ATS, the most thorough long distance survey as well as the mainstream dataset for long distance travel studies to date. To estimate the O-D matrix, WSA regressed population and employment against the 1995 ATS trip O-D demands. Subsequently, long distance trips were projected at county level based on 2008 employment and population data. State level totals were also estimated to help balance the county-level estimates. According to WSA's estimation, U.S. residents took 938,389,638 long-distance trips in 2008. Over 18% of those trips (173,945,591) were business-related travel while others (764,444,047) were for pleasure/other purpose, across all the travel modes. Considering the implementation scope of the case study, only trips between counties in states of Washington, Oregon, and California are selected. Additionally, average occupancy is estimated based on the 1995 ATS sample dataset: 2.30 persons per vehicle for business travel, and 3.35 persons per vehicle for pleasure/other travel, to convert person trips to vehicle trips.

### *5.1.2 Network data*

The Freight Analysis Framework (FAF$^3$) network is selected as the base network of national highways in this study. The FAF$^3$ network is a nationwide geospatial network developed by the FHWA to understand the performance of the national freight transportation system. Satisfactorily the network dataset includes such detailed and updated roadway features as length, functional class, AADT, link capacity, and link speed limit, which are salient for resolving the network loading issues of long distance passenger travel. Detailed attributes of FAF$^3$ can be found in Appendix B. Particularly the AADT07 is introduced as the preloaded volume to the network. In addition to the completeness of the dataset, another reason for choosing FAF$^3$ lies in its excellent network connectivity and topological integrity, as recommended by the FHWA expert team.

As specified in the *Network Design Report*, the county level FAF$^3$ network contains 170,994 links including both principal and minor arterials. A total of 22,014 links within states of Washington, Oregon, California, Idaho, and Nevada are selected as the base network for the case study. Idaho and Nevada are included to consider the possibility that some trips may traverse their jurisdictions. A network clean-up is conducted to remove redundant links and add some missing lines for solving the topological discrepancy and improving network connectivity.

Counties are treated as Traffic Analysis Zones (TAZs) in this study. There are 39 counties in Washington, 36 counties in Oregon, and 58 counties in California. Centroid is the network representation of the corresponding TAZ, where all long distance trips are generated from and attracted to. Each centroid is associated with a unique CTFIPS code. 680 centroid connectors are further generated arbitrarily to connect centroids to the base network. Each centroid connector is labeled with the CTFIPS code indicating the county with which it is associated. For example, link ID 176983 has value of 53033 stored in its *Connector* attribute, meaning the link is a centroid connector which connects to King county centroid.

Since there is no preloaded volume on centroid connectors, travel impedance on them cannot be determined by travel impedance functions. As proposed in the Chapter 4, the impedance on centroid connectors are determined by the *Annual Congestion Cost per Auto Commuter* calculated by TTI, as part of their congestion measures (see Appendix C). Since TTI's performance measures are at the MSA level, it is reasonable to assume that each county shares the same congestion cost as the MSA it belongs to. For the counties outside MSA, the annual cost is set arbitrarily as $180 for urban counties and $90 for rural counties. Additionally, a selected route will not go through the centroid connectors. Centroid connectors can only be traversed at the begging or in the end of a route; therefore it cannot represent the node impedance if the node is in the middle of a certain route, but can only reflect the impedance of origin and destination nodes. A list of rural counties is detailed in Appendix D.

### *5.1.3 Other datasets*

In addition to the O-D matrix and network data, several secondary data sources are utilized such as:

The ***1995 ATS sample data*** is retrieved as a comparison dataset for the trip length distribution information. The 1995 ATS randomly sampled approximately 80,000 households throughout

U.S. to participate in telephone interviews, which collected 556,026 person trips and diverse information, including the origin and destination of the trip, number of stops along the way, travel mode, size of travel party, trip purpose, etc. Considering the implementation scope of the case study, only trips between counties in states of Washington, Oregon, and California are selected. After excluding non-highway modes, there are 6,706 business person trips and 29,706 leisure/other person trips by auto within Washington, Oregon, and California. Additionally the Oak Ridge National Laboratory (ORNL) provided a rough estimate of the trip distances for all trips.

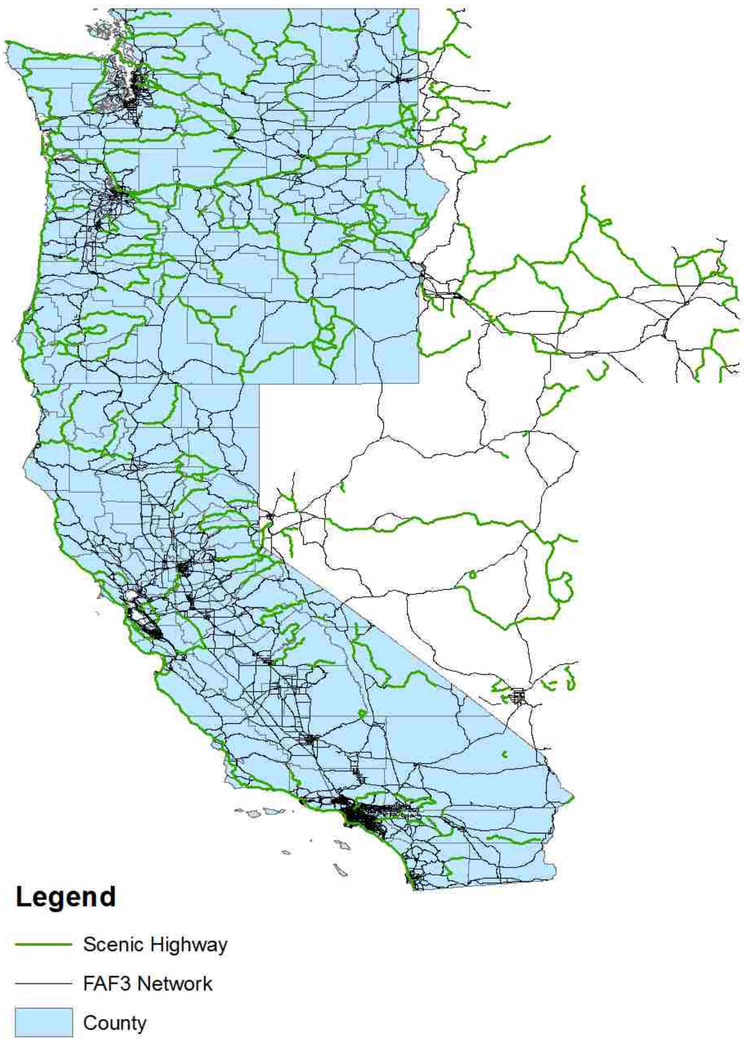



Figure 5.1 Scenic Highways in Five States

***Scenic highway dataset*** from National Scenic Byways Online (NSBO) is shown as green links in Figure 5.1. Geospatial analysis is performed in ArcGIS to identify scenic links in FAF$^3$ network by overlapping two datasets. Meanwhile scenic factor is set as 1 if the FAF$^3$ link is non-scenic. Additionally, this study acquires ***National and State Parks data*** from multiple state agencies and ***U.S. Hotels, Motels, and Accommodations Database*** obtained from online vendors to count number of roadside attractions within 5 miles of FAF$^3$ links (line buffers are applied). Relevant geo-processing tasks are also completed with ArcGIS.

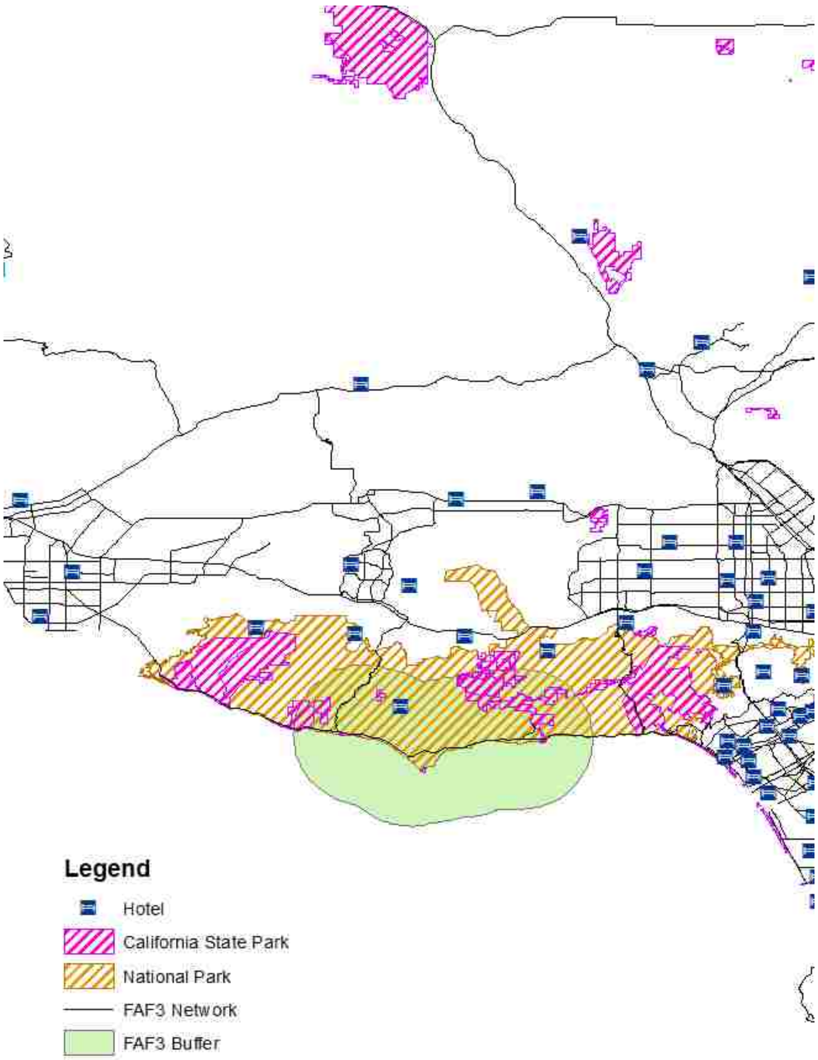



Figure 5.2 An Illustration Showing Roadside Attractions

## 5.2 Scenario Configurations

So far most of long distance or large-scale (e.g. statewide travel model) assignments were done using an exclusively time-based impedance function. This is considered as the base scenario and obviously ignores some important factors influencing long distance route choice. Besides testing the implementation feasibility of the proposed modeling framework, the following scenarios describing different configurations are analyzed, where the base scenario is successively refined:

i. The effect of taking distance-based operational cost into consideration;
ii. Distributive patterns of the scenarios where different assignment algorithms are applied;
iii. The difference of distributive patterns between business and pleasure/other travel;
iv. The effects of different distributions and magnitude of the error term;
v. The effects and sensitivity of newly introduced scenic factor and attraction factor.

Base Scenario:

Configuration 0: For both business and pleasure/other travel, only consider the impedance as the travel time-based cost, where $c_i = \varphi t_i \left[ 1 + \alpha_i \left( \frac{x_i}{C_i} \right)^{\beta_i} \right]$

Considering the travel distance and the operational cost induced:

Configuration 1: For both business and pleasure/other travel, add the distance-based operational cost component, where $c_i = \delta L_i + \varphi t_i \left[ 1 + \alpha_i \left( \frac{x_i}{C_i} \right)^{\beta_i} \right]$

Test different error term distributions:

Configuration 2.1 (normal distribution) and 2.2 (Gumbel distribution): based on configuration 1, introduce normal/Gumbel distributed error terms with varying magnitudes to the SUE algorithm;

Impedance function extension for pleasure/other travel:

Configuration 3.1: $c_i = s_i \left( k_i + \delta L_i + \varphi t_i \left[ 1 + \alpha_i \left( \frac{x_i}{C_i} \right)^{\beta_i} \right] \right)$ where $s_i = 0, 0.1, 0.2, \cdots, 0.9, 1.0$

Configuration 3.2: $c_i = e^{-a_i n_i} \left( k_i + \delta L_i + \varphi t_i \left[ 1 + \alpha_i \left( \frac{x_i}{C_i} \right)^{\beta_i} \right] \right)$ where $a_i = 1; 0.3; 0.1; 0.05; 0.01$

## 5.3 Implementation Results

TrasnCAD 5.0 is deployed as the implementation platform. This section summarizes implementation results. For each corresponding scenario total annual VMT, VHT, and delay are estimated as system-wide performance measures. The total annual VMT (VHT) is the total mileage (hours) traveled by all vehicles in one year. The total annual delay is calculated as the difference between the actual total annual VHT and the assumed total annual VHT if all the vehicles were traveling at free-flow speed. Centered on the SUE algorithm, the distributions and magnitudes of the error term are first investigated and its impact on distributive patterns is studied (Configuration 2.1 and 2.2). Configuration 0 and Configuration 1 are then compared across different assignment algorithms (SUE, DUE, and SO) to show the effect of including distance-based operational cost in travel impedance functions. Third, the impacts of newly introduced parameters (scenic factor and attraction factor) are studied, in terms of ***system-wide performance measures***, ***link usage***, and ***distributive patterns***. Finally the trip length distributions from this implementation are compared to the 1995 ATS data.

### *5.3.1 Error term of SUE*

TransCAD sets the standard deviation of the random error in the SUE mechanism (Equation 3.4) as $(e/100)\times link\_impedance$, where the $e$ is a modeling parameter referred as error term in the following discussion. As introduced in section 4.2, the magnitude of the random error roughly reflects how differently people perceive their travel time from the actual travel time. A larger error term indicates a significant difference exists between a traveler's perception and the actual travel impedance; as a result, route utilization is very diverse across the traveling population, including some that may be associated with significantly higher impedance than the true "shortest path". When the standard error $\xi \to \infty$, the ***share of flow*** on all routes will be equal, regardless of route travel impedance, meaning people are indifferent about travel impedance and just uniformly and randomly choosing routes. When the standard error $\xi \to 0$, the ***travel impedance*** on all routes will be equal, meaning people have perfect judgment on the travel impedance and the entire scenario becomes a DUE one. In this study, the magnitude of the error term is arbitrary set between 0 and 100.

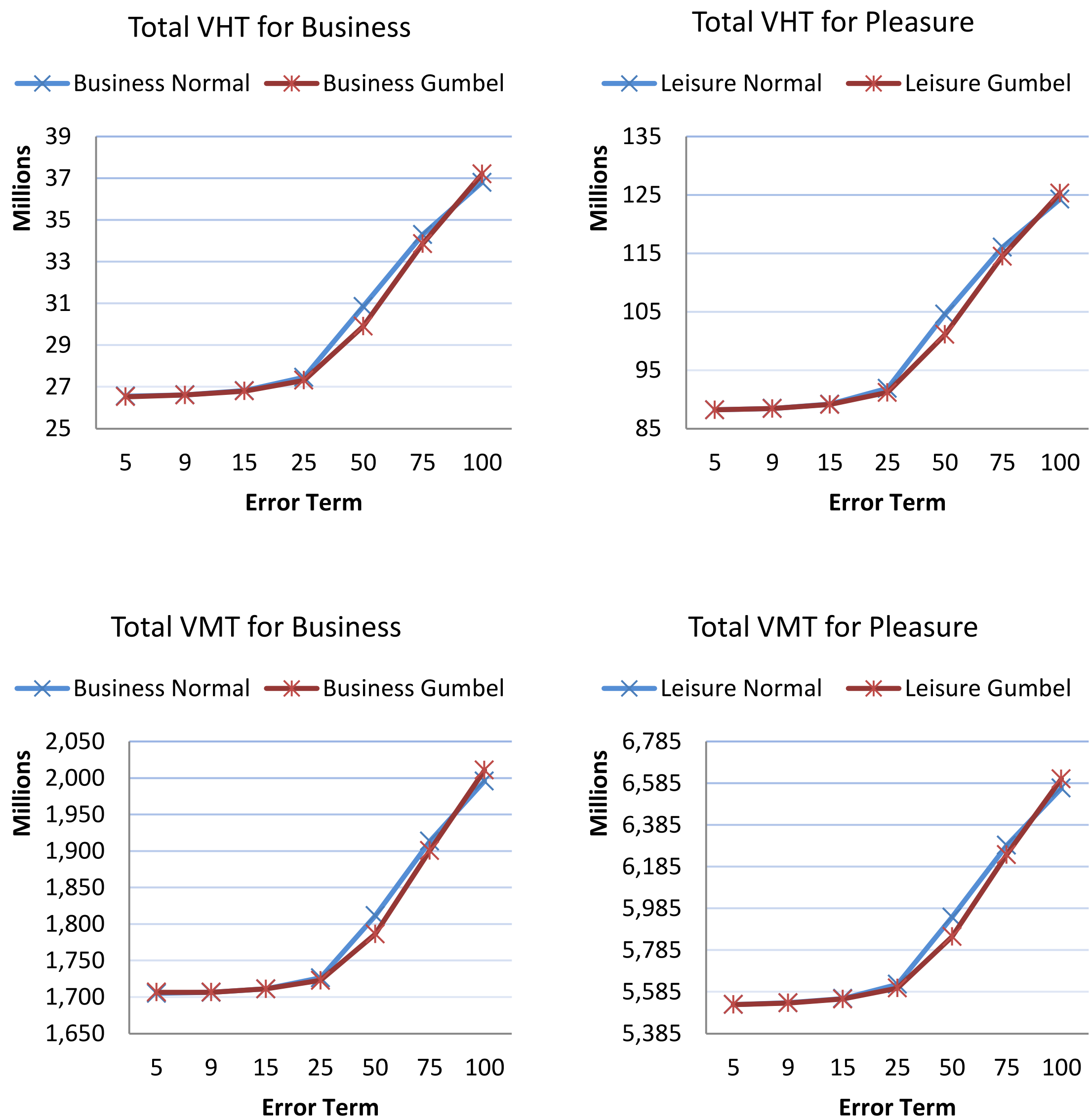



Figure 5.3 Total VMT and VHT Change with Varying Magnitudes of Error Terms

Figure 5.3 summarizes changes in total VMT and total VHT with varying magnitudes of error terms. First of all, regardless of the trip purpose or distribution of the error term, changes show a similar trend: with an increasing magnitude of the error term, total VMT and total VHT grow as well; because a larger error term will result in a more diverse route choice. Second of all, when the error term is less than 25 (meaning people's perceived travel impedance is within the range of

±25% of the actual impedance), changes in total VMT and total VHT are insignificant (increase of total VHT is less than 5%; increase of total VMT is less than 2%). However, when the error term is larger than 25, the total VMT and total VHT show an almost linear and dramatic increase. As the error term grows from 25 to 100, increase of total VHT is around 35% and increase of total VMT is around 16%. This indicates that system-wide performance measures present a range of “tolerance” in terms of long distance travelers’ perception errors, and this range is not influenced by the trip purpose or the assumptive distribution of the error term.

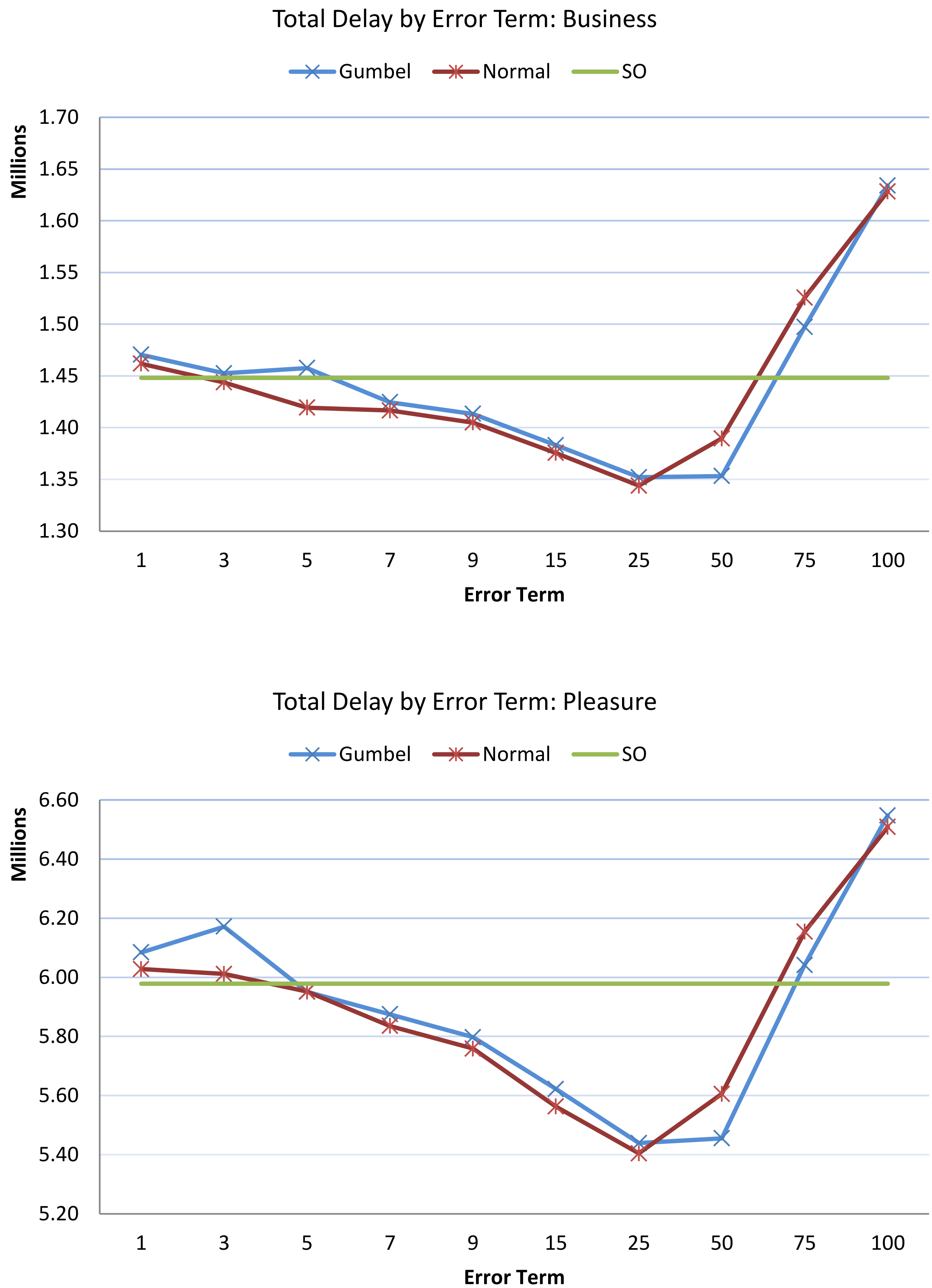


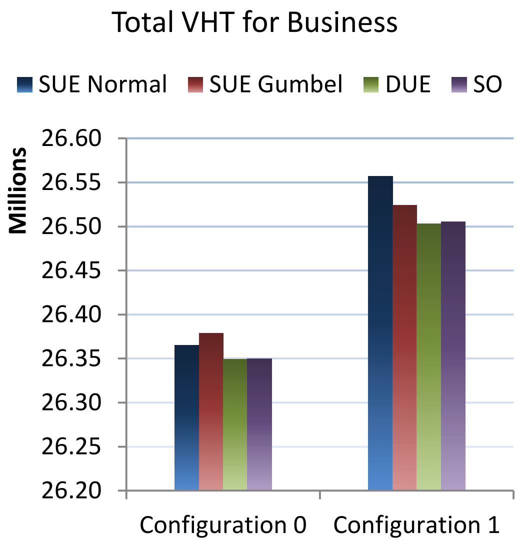

Figure 5.4 Total Delays by Error Term

Total delays are defined as the difference between travel time on the used route and the free-flow travel time. Total delays by error term are illustrated in Figure 5.4 for both business and pleasure/other travel. As a comparison, total delays in an SO scenario are also presented in the figure, where the minimum total VHT is achieved. Again Gumbel distributed error terms show a similar trend as the normally distributed ones. It is an very interesting observation that when error terms range from roughly 5 to 75, the SUE-based total delays are lower than SO-based total delays, as much as 9 for pleasure travel and 7 for business travel. The reason lies in that under the behavioral assumption of SUE, many travelers will not find the theoretically best routes, in which case certain individuals' non-optimal route choices ease the congestion on heavily traveled routes and enable others to travel much more efficiently, which creates a lower total delay than the SO scenario. When the error term is less than 5, the SUE scenarios become a DUE assignment, where most travelers hold similar perception toward travel impedance and result in higher congestions on more frequently traveled corridors and higher total VHT; when the error term is more than 75, travelers are very diversified in terms of perceptions toward travel impedance, therefore many longer routes are utilized, which also result in a higher total VHT than the SO scenario.

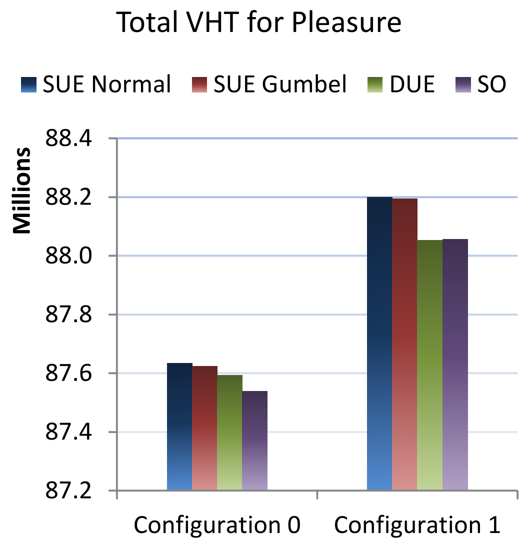
In a long distance travel context, given the long duration and distance spent on traveling, it can be assumed that the "relative error" in people's perception is lower than intra-urban travel. For example, a 10 minutes perceived error induces a 33% error in a 30-minute commute, yet only a 5.5% error in a 3-hour travel. Additionally given the findings from this case study, the system-wide performance measures are not varying significantly under the threshold of 25%. Therefore in the following discussion, a 5% error term is adopted. However, this investigation also indicates a more in-depth study on impact factors on long distance travelers' perception error. For example, people on business trips should have a smaller error than people on pleasure trips, considering the former group prefer to use higher functional class roadways with higher reliability and better roadway conditions. Business travelers are also more likely to be familiar with the planned route than pleasure travelers, because pleasure travel does not occur frequently and people tend to seek for new options when traveling for pleasure.

### *5.3.2 Comparison between Configuration 0 and Configuration 1*

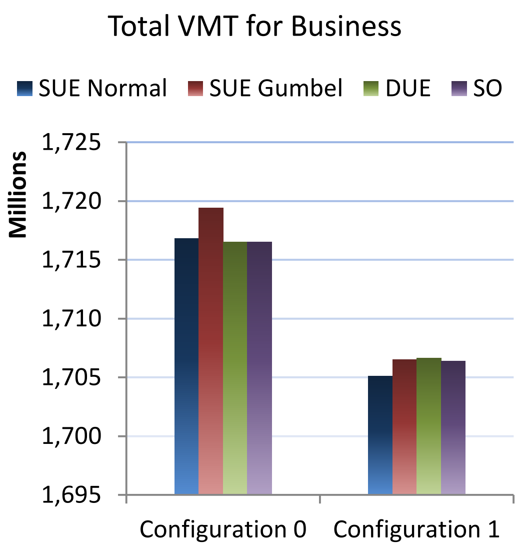
Different assignment algorithms (SUE, DUE, and SO) correspond to different behavioral assumptions. Speaking of the long distance network loading issue, a critical question to be answered is how those different behavioral assumptions alter system-wide performance measure and its indication to the interpretation of the model. Figure 5.5 summarizes total VHT, VMT, and delay for business and pleasure/other travel across four different assignment mechanisms, and also explores the effect of including a distance-based impedance component.



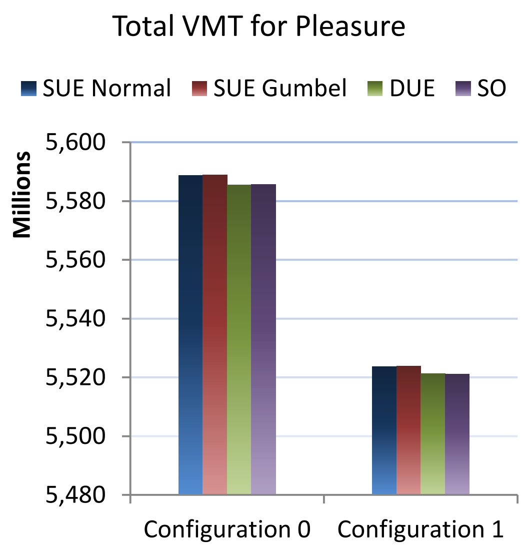
Total VHT for Business
SUE Normal
SUE Gumbel
DUE
SO
Millions
26.60
26.55
26.50
26.45
26.40
26.35
26.30
26.25
26.20
Configuration 0
Configuration 1

Total VHT for Pleasure
SUE Normal
SUE Gumbel
DUE
SO
Millions
88.4
88.2
88.0
87.8
87.6
87.4
87.2
Configuration 0
Configuration 1

Total VMT for Business
SUE Normal
SUE Gumbel
DUE
SO
Millions
1,725
1,720
1,715
1,710
1,705
1,700
1,695
Configuration 0
Configuration 1

Total VMT for Pleasure
SUE Normal
SUE Gumbel
DUE
SO
Millions
5,600
5,580
5,560
5,540
5,520
5,500
5,480
Configuration 0
Configuration 1

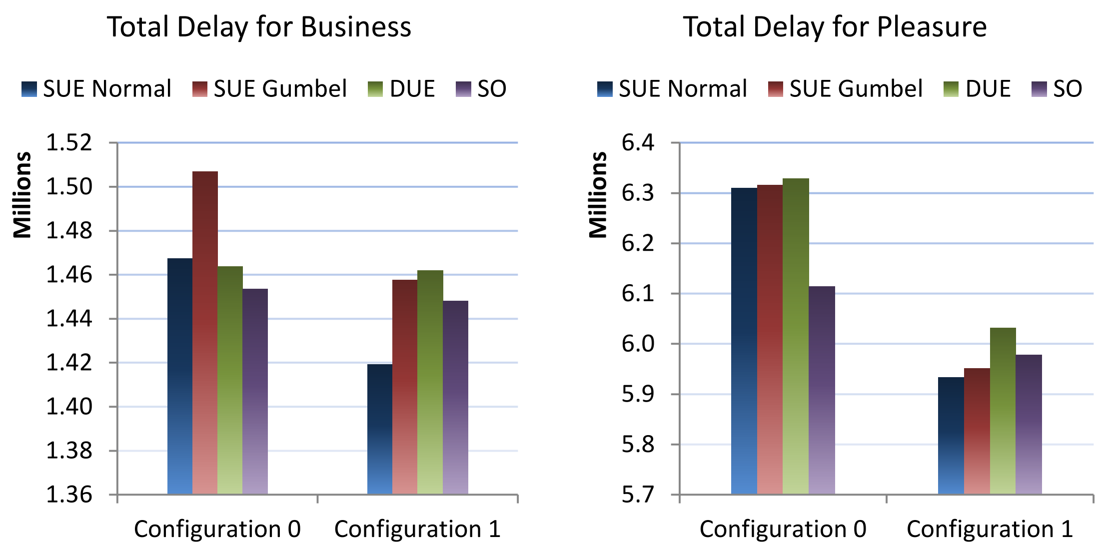



Figure 5.5 Total VHT, VMT, and Delay for Business and Pleasure Travel

At the system level, the average (over both configurations and four assignment mechanisms) total VHTs are 26.4 million vehicle-hours for business travel and 87.9 million vehicle-hours for pleasure. Total VMTs are 1.71 billion vehicle-miles for business travel and 5.55 billion vehicle-miles for pleasure. Total delays are 1.46 million vehicle-hours (5.5% of the total VHT) for business travel and 6.12 million vehicle-hours (7.0% of the total VHT) for business for pleasure. Apparently the long distance travel delay is lower than intra-urban travel delays.

Inclusion of the distance-based operational cost generally results in an increased total VHT, a decreased total VMT, and decreases of total delays. However, the change is not that significant. The average (over four assignment mechanisms) change in total VHT is 0.61% for business and 0.60% for pleasure; the average change in total VMT is -0.65% for business and -1.16% for pleasure; the average change in total delay is -1.77% for business and -4.69% for pleasure. Adding the distance-based operational cost to the travel impedance function essentially brings another dimension to people's route choice behavior, indicating in Configuration 1, travelers are explicitly "cautious" about the cost induced by the trip length. As a result, route choices are diversified compared to Configuration 0 and both total VMT and total delays decrease yet total VHT increases (due to a longer route selected). Although the relative change is small, considering the large size of the network and travel volume, the absolute changes are still considerable.

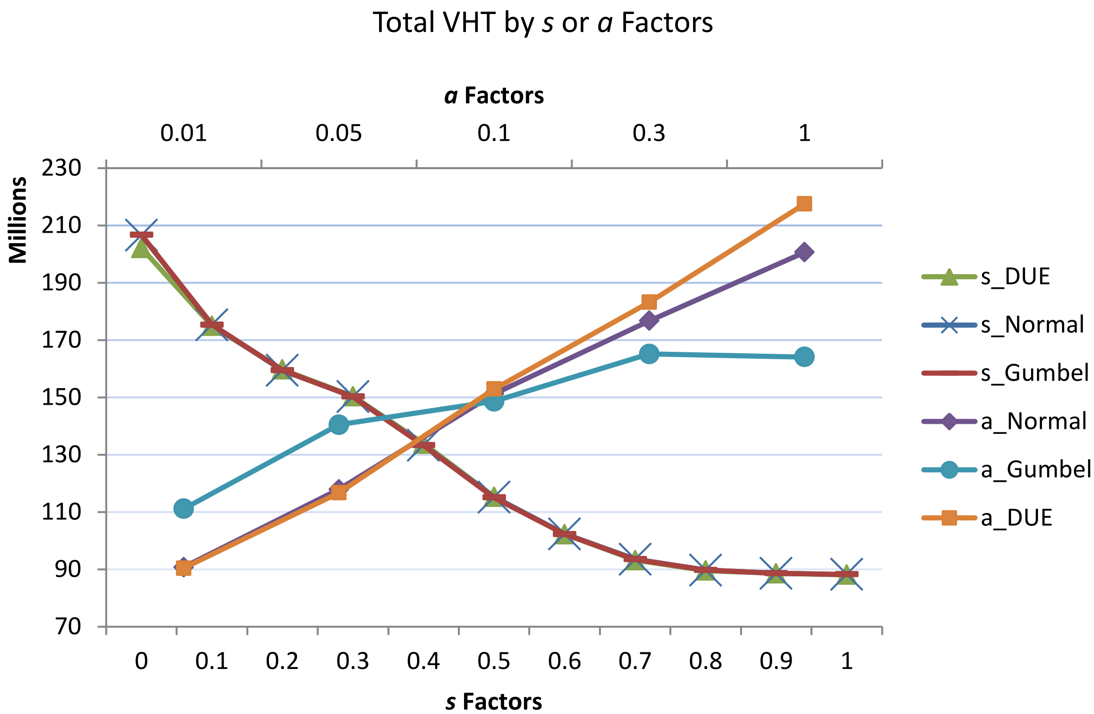
Results from different assignment mechanisms are very close to each other. Theoretically ① SO should yield minimum total VHT; ② SUE mechanisms should yield higher total VHT and VMT than the DUE method due to diversified route choices by perception error; and ③ compared to normally distributed error terms, Gumbel distributed ones will likely result in overestimated flow on overlapping routes (Sheffi, 1985), thus lessen the total VMT yet raise the total VHT, because overlapping routes will become more congested (when only travel time is considered as the travel impedance). Those trends can be slightly traced in Configuration 0 yet not significantly. Because when applied to large-scale and complex networks, closed-form solutions to any of the assignment mechanisms do not really exist therefore solving the network loading problem relies on an iterative process, which brings deviations from the theoretical assumption.

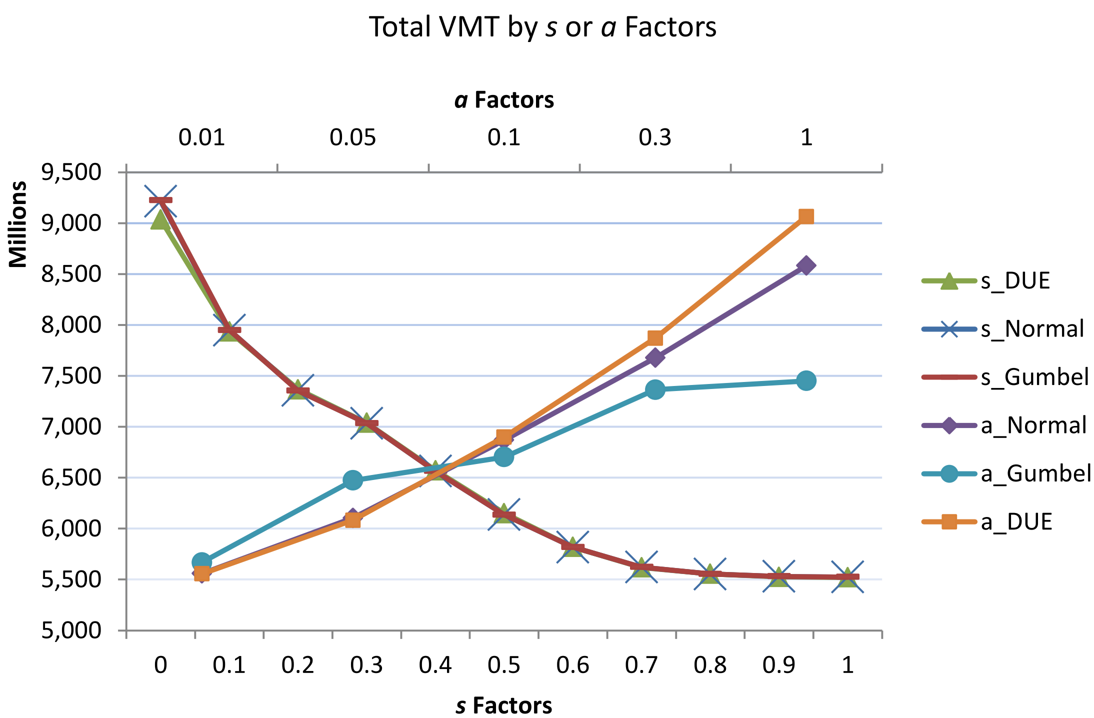
When taking distance-based operational cost into consideration of travel impedance, the second and third theoretical arguments do not hold, because the system-wide performance measures are now reflecting another dimension of travel cost induced by travel distance. As shown in Configuration 1, for business travel, the normally distributed error terms result in higher total VHT and lower total VMT than a Gumbel distribution. This brings insights to interpreting the network loading results when using composite travel impedance. Additionally the closeness of SO method results to ones by other assignment mechanisms suggests that the benefits of route guidance and traveler information systems are relatively low in the long distance travel market.

It is beyond the scope of this study to determine which assignment mechanism produces the best network loading results. However, without a doubt the discussion above helps improve our understanding on how different behavioral assumptions influence the loading result and how to interpret the result.

### *5.3.3 s and a factors in pleasure travel*

As introduced in section 4.4, the *s* factor and *a* factor represent the attractiveness of scenic highways and roadside attractions respectively. In Figure 5.6, the top horizontal axis is scale of the *a* factor, and the bottom one for *s* factors. Total VHT, VMT, and delay for pleasure/other travel are illustrated with varying magnitudes of those factors.


Total VHT by s or a Factors
a Factors
0.01
0.05
0.1
0.3
1
Millions
230
210
190
170
150
130
110
90
70
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
s Factors
s_DUE
s_Normal
s_Gumbel
a_Normal
a_Gumbel
a_DUE


Total VMT by s or a Factors
a Factors
0.01
0.05
0.1
0.3
1
Millions
9,500
9,000
8,500
8,000
7,500
7,000
6,500
6,000
5,500
5,000
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
s Factors
s_DUE
s_Normal
s_Gumbel
a_Normal
a_Gumbel
a_DUE

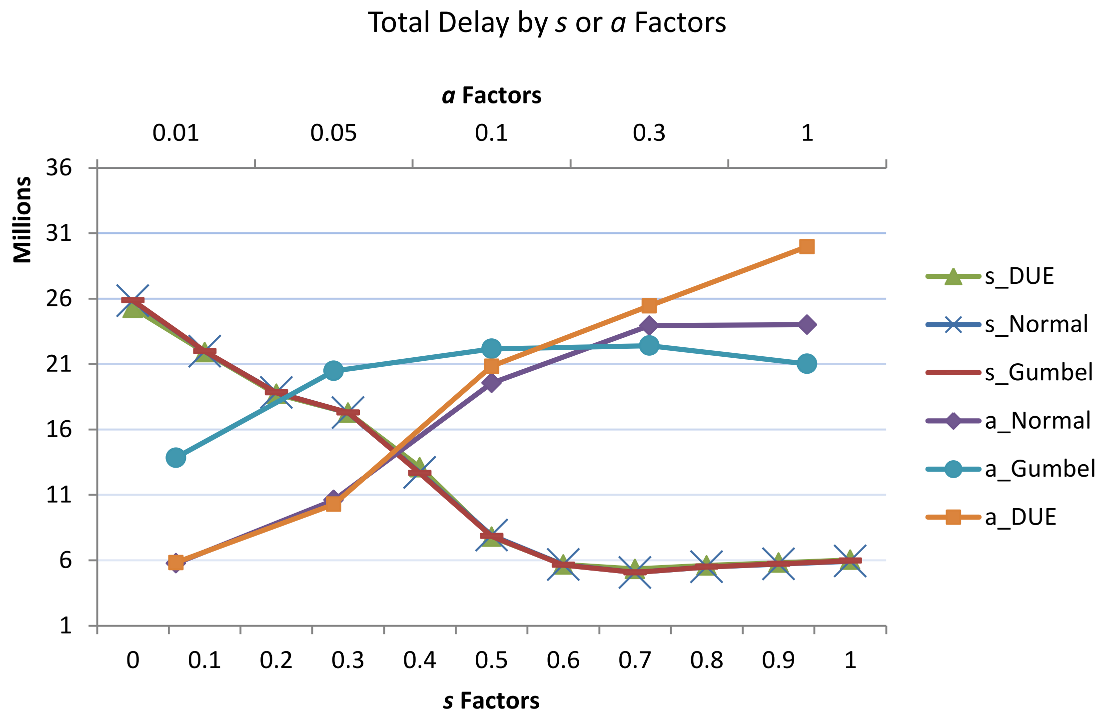



Figure 5.6 Impacts of *s* and *a* Factors

Both factors function as how they are designed. When *s* and *a* factors are applied, people are becoming less sensitive to the travel impedance. A higher *s* factor or a lower *a* factor reflects a lower valuation from travelers towards scenic highways and roadside attractions respectively, therefore people are less likely to select a "beautifully expensive" route, which is reflects by lower system-wide performance measures. Additionally, *s* factor presents a very consistent and stable characteristic across different assignment algorithms (DUE, SUE_Gumbel, and SUE_Normal). The *a* factor, however, presents a sensitive characteristics. As specified by Equation 4.2, when the *a* factor is less than 0.1, people's attitude change toward the increasing number of attractiveness is quasi-linear. Reflected to the system-wide performance measures, the left part of the *a* factor curve shows a steady decrease (and more consistent performance across different assignment algorithms) with in the decrease of *a* factor. When the *a* factor is larger than

0.1, people's attitude change toward the increasing number of attractiveness is exponential, which results in a bigger diverge on the left part of the *a* factor curve across different assignment algorithms. An important practical implication is that when applying the *a* factor, switching assignment algorithms will result significance difference in terms of system-wide performance measures. For example, when $a = 1$, compared to using the SUE_Gumbel, using DUE will increase the total VHT by 32%, the total VMT by 21%, and total delays by 42%!

### *5.3.4 Link usage*

Theoretically the random utility assumptions in SUE should produce more evenly distributed flows on alternative routes than DUE would. The implementation results validate this, as shown in Figure 5.7. Across different magnitudes of *s* and *a* factors, the difference in link usage between DUE and SUE methods ranges from 6% to 18.6%. In other words, DUE directs more long distance trips to higher-level roads than the SUE methods and thus overestimates the level of congestion on those roads. This systematic bias of flow and congestion on higher level roads with rational behavior assumptions could have important implications for transportation planning and policy analysis applications. Again as expected, a decreasing *a* factor or an increasing *s* factor triggers a growth in link usage, meaning the newly introduced parameters are sensitive and effective.

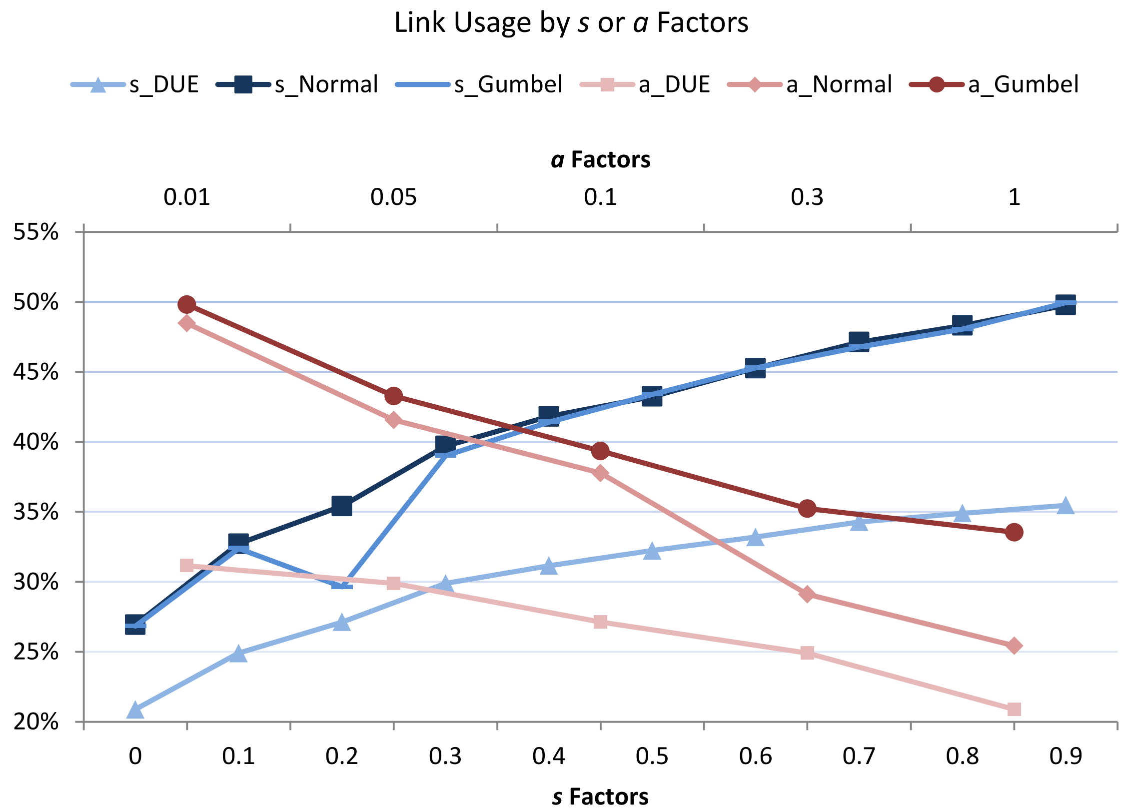



Figure 5.7 Link Usage

### 5.3.5 *Distributive patterns*

Varying magnitudes of *s* and *a* factors produce significant differences in the estimated long distance distributive patterns on a large-scale real-world network and potentially affect planning and policy decisions. Out of 17,689 O-D pairs in the demand matrix, pleasure trips from Seattle to Los Angeles are chosen to demonstrate the impact, as shown in Figure 5.8 and 5.9 (the original demonstration is in a GIF format; due to the limitation of MS Word they cannot be properly played in this document). When $s$ = 0.9, pleasure travelers only slightly value the attractiveness of scenic highways, therefore most of travelers still choose to travel Interstate 5, as they “perceive” the route minimizes their travel impedance. When $s$ = 0.4, most of travelers are using the U.S. Route 395 and the U.S. Route 101, both scenic routes. When $s$ = 0.1, another extreme representing that travelers show a high preference in scenic highways, the U.S. Route 101 is most traveled for the Oregon and California sections, while in the state of Washington, a number of travelers follow the Interstate-84, a beautiful drive along the Columbia River.

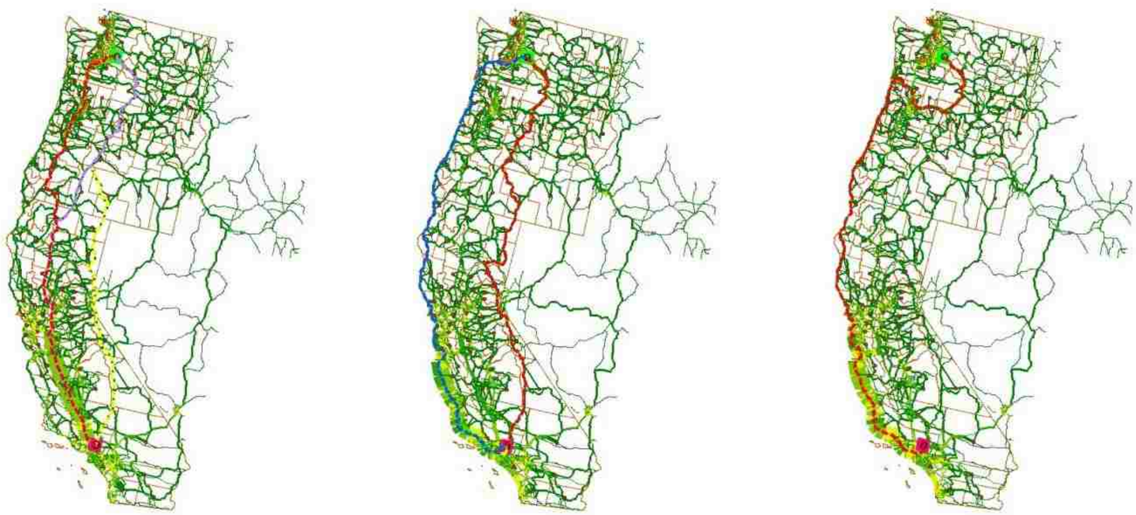


Figure 5.8 Distributive Patterns when $s$ = 0.9 0.4 0.1 (left to right); Trips from Seattle to LA

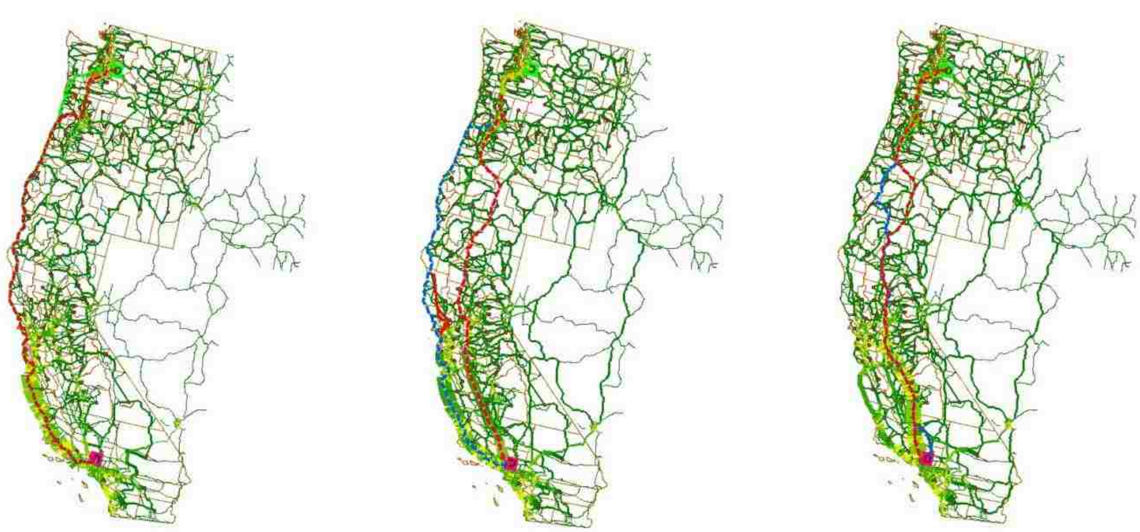


Figure 5.9 Distributive Patterns when $a$ = 1, 0.05, 0.01(left to right); Trips from Seattle to LA

Similar patterns are observed using the *a* factor. When $a = 1$, the U.S. Route 101 is most loaded route; while $a = 0.01$, most travelers selected Interstate-5; in between when $a = 0.05$, some travelers select U.S. Route 101 along the coast, while others still travel along the Interstate 5. Without further data collection and travel survey, it is unknown for now how much the *s* factor (or the *a* factor) exactly is, however, through this case study, it is shown both factors are effective in terms of reflecting different distributive patterns within a plausible range. For animated demonstrations, please visit http://www.uwstarlab.org/long_distance_travel.html

### *5.3.6 Trip length distribution*

The 1995 ATS sample data with trip length information is acquired to be compared with results from the case study. In general the 1995 ATS trip length estimation process involves an AON assignment method (allocate all the trips to the shortest path) and a travel time-based impedance with adjustments by highway functional class. A detailed description of the process can be found in Hwang and Rollow (2000). Contrastively this study applies UE assignment mechanisms with an enriched composite travel impedance function. And trip length distributions for business travel and pleasure/other travel are shown in Figure 5.10 and Figure 5.11 respectively.

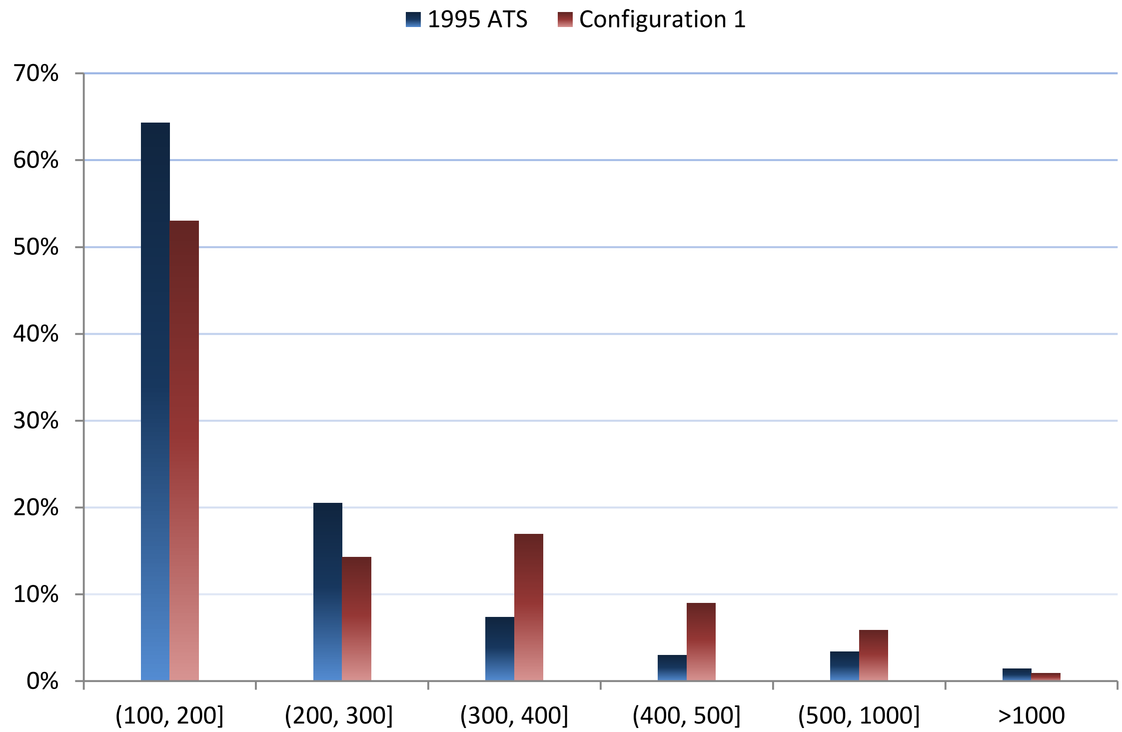



Figure 5.10 Trip Length Distribution – Business: Comparison with the 1995 ATS

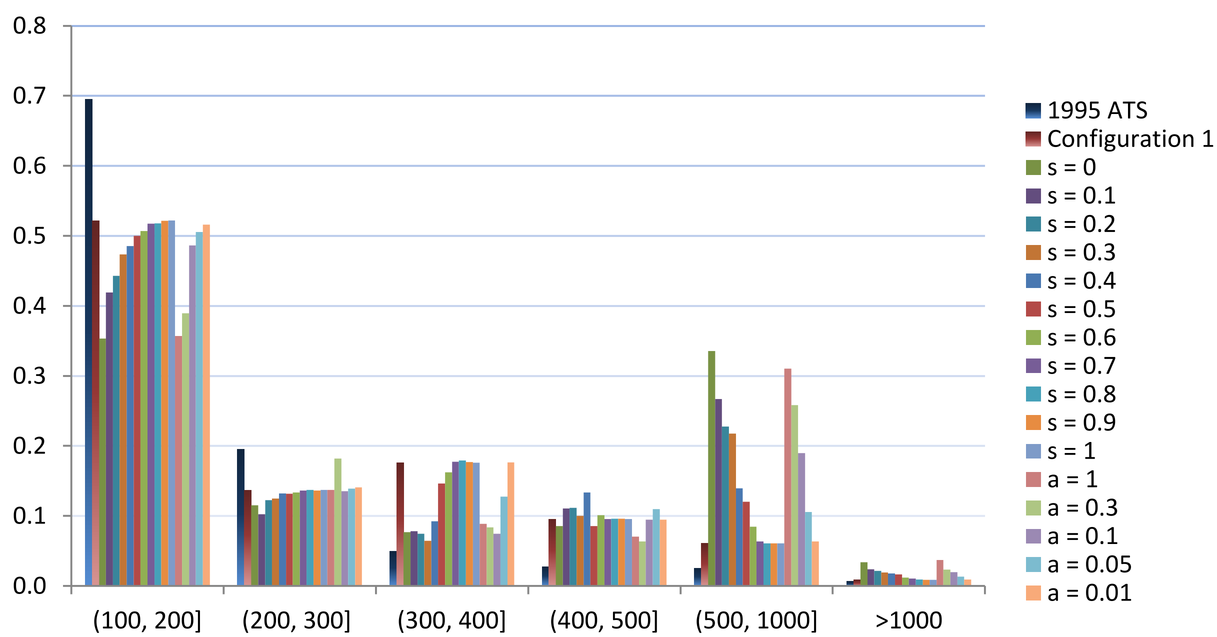



Figure 5.11 Trip Length Distribution – Pleasure: Comparison with the 1995 ATS

In general regardless of trip purpose, the case study yield less trips in “shorter than 300 miles” categories, and more trips in “longer than 300 miles” categories. The difference between the two distributions is more significant in the pleasure/other travel than the business travel. There are three potential reasons for the discrepancy: First, the 1995 ATS sample data is on a MSA O-D demand basis, with a bias towards long distance trips between urban areas and therefore “ignores” long distance trips between a rural O-D pair. The O-D matrix for this study, although derived the 1995 ATS data, has been averaged out at a county level. Some longer trips between remotely distributed counties are artificially “generated” through this averaging process, which add more trips to “longer than 300 miles” categories. Second, the 1995 ATS trip distance estimation applied a shortest path algorithm and biased over higher class roadways, which underestimate the trip lengths of many trips traveling the lower class roadways, and thus overestimated the number in “shorter than 300 miles” categories. Third, the 1995 ATS trip distance estimation did not involve a market segmentation process, as this study. In other words, it did not differentiate business travel from pleasure/other travel, nor consider the “discount effect” within the pleasure traveling population in the network loading process, therefore many lengthy pleasure trips were underestimated as shorter trips.

This discrepancy reveals important policy making indications, regarding the recently rising national interest in high-speed rail investment. Considering prevailing long distance traveling modes, two competitors of high-speed rail are autos and airplanes; 300 to 700-mile is the critical market for the high-speed rail. Correctly estimated travel volume in this market (especially for pleasure travel) is a salient quantitative reference for justification of massive investment.

## 5.4 Discussion

Through the case study, questions raised in the beginning of this chapter are all answered. The work flow proposed in Chapter 4 is a feasible solution to the long distance travel network loading. Variations in model parameters, including the error term, the *s* factor, and the *a* factor, generate significant impacts on the distributive patterns. Particularly, both newly introduced factors function as how they are designed, able to well generalize pleasure travelers' behavior and reflect it onto the distributive patterns. Different assignment algorithms based on varying behavioral assumptions are also influential. The case study not only brings insight to the planning practice, but also to the policy making process. For example, results show that the 1995 ATS trip length distribution underestimated the number of trips in "longer than 300 miles" categories, where the 300 to 700-mile is the critical market for the high-speed rail.

Note that without extra efforts and resources for new travel survey, it is very hard to calibrate the "real" values of the *s* factor and the *a* factor. However, it should be recognized that there exist well-developed methods for calibrating both factors, like stated preference survey and means-end analytical approaches. Even though their values are unclear at the current stage, due to the fact that both new factors are designed to be bounded within a certain range, decision makers are able to "tune" across varying magnitudes of both factors to speculate possible distributive patters and inter useful information to support policy making. Also this case study is not materialized on a multi-scale basis, and it would be interesting to apply similar methodological framework to MSA-level and state-level and compare the difference in system-wide performance measure resulted from varying geographic representations.

A byproduct from the modeling framework is all the travel impedance is in the unit of dollars, which provides a reference for pricing new facilities. For example, the median cost for business travel from Seattle to Los Angeles is around $1100, while the cost from Seattle to Portland is about $191. Since it is a composite cost by accounting for operational cost and value of time, a competing mode to highway travel may find these values useful to identify their market.

Smart and accountable policy making requires quantitative information detailing the operation of national transportation system. Traditional macroscopic statistics may not meet the diverse needs for answering questions regarding policy, public finance, and environmental issues. For example, while regional VMT and VHT will likely decrease in a nose-diving economy, congestion in areas of increased density will likely increase. Performance measures in multiple dimensions and at different scales are in great need. The modeling framework proposed in this research not only delineates the full procedure of depicting distributive flow patterns, but also is able to generate useful and comprehensive measurements at different geographic levels for accountability evaluation. Granted further funding and labor resources, the framework we developed can be readily extended to a broader spectrum of domains including intra-urban travels, to better support federal level decision making.

# Chapter 6 Conclusions

America's highways need a new vision, a vision that goes beyond facilitating individual and commercial movements and gives voice to new challenges brought by the aging infrastructure, mounting congestion, declining system reliability, changing demographics, and increasing planning and environmental awareness and concerns. While improving intra-urban travel has attracted great attention and massive investment, another integral part of American life, people and freight on the long haul, deserves more analyses and modeling, not only because it embraces all the challenges mentioned above, but also because it accounts for 25% of all person miles in the nation.

Long distance passenger travel is not a simple extension of travel distance to the well-studied intra-urban travel. Understanding long distance travel requires new data input, raises new questions, and presents new challenges. While extensive regional travel data collections are conducted at the MPO level periodically, transportation service and cost data at corridor, interregional, and national level are scarce. While intra-urban travel centers on people's daily life (home-based or non-home based trips), long distance travel embraces a much more diverse traveling needs, especially in recreational and tourism travel. Sequential transportation planning procedures have been well established and new theories keep advancing in an intra-urban context, but modeling framework for long distance passenger travel is absent. Addressing these needs, questions, and challenges will help support decisions about transportation policies, investments, and operations, as well as motivates the research efforts elaborated in this dissertation.

This research applies a national data framework approach to modeling long distance passenger travel over national highways. A pivotal question answered by the research is: how can we gain knowledge and insights of long distance passenger travel patterns at interregional level given existing resources? The framework developed comprises not only a data warehouse where a plethora of datasets reside and interact, but also a modeling framework to infer long distance travel network loading conditions. Particularly, the modeling framework delineates the procedure of deriving distributive patterns from individual behavior, with enriched impedance models to approximate the behavioral realism of long distance route choice. The model development is reasonably rigorous since most of model parameters are based on empirical justification of available datasets without arbitrary conjectures. Two new parameters are introduced to the impedance model as linkages between contextual data and planning data. Additionally the case study demonstrates the feasibility and practicality of the framework using a case study. The implementation on the county-level network within states of Washington, Oregon, and California generates system-wide performance measures, distributive patterns, and trip length distributions, providing quantitative references for policy making and investment justification in large-scale transportation systems. Future research should examine the likelihood of incorporating more advanced behavioral assumptions and the feasibility of deploying emerging data sources in the proposed modeling framework.

The following sections first summarize research efforts and findings, followed by highlighting research contributions in resolving data issues, modeling methods, and policy evaluation needs. A comprehensive picture of national travel is then presented as the Traveler Analysis Framework, and roles of this dissertation research in the big picture are stressed. Finally, future research is discussed, including an in-depth study on long distance travelers' behavior, potential data collections as well as utilization of emerging data sources, and further development of the modeling framework.

## 6.1 Summary of Research

### *6.1.1 Identify available data sources and review current practices*

Chapter 2 reviews available datasets and related literature regarding travel impedance and large-scale network loading from theories to practices, in five dimensions: ***Infrastructure*** gives an overview of the national highway system, whose abstraction will serve as the base network in the modeling framework. Recognizing the immense scale and complexity of the network, the review shows roadway functional classes can naturally serve as criteria for a multi-scale configuration, to appropriately involve different geographical layers into computations at varying resolutions. ***Information*** identifies available data sources on both the demand and the supply side of the large-scale system. Those data sources are categorized into geography data, demand data, traffic data, and validation data, with their roles, availability, and limitation clarified. This inventory efforts reveals the inconsistency between current practices of region-based data collection and needs for large-scale system modeling. A bright highlight to the data reality is that both HPMS and FAF$^3$ already make solid efforts in terms of data integrity, quality, and formatting issues, which may be utilized as the base network for this study. ***Externalities*** refer to the external factors creating resistance or attractiveness to long distance travel. Since very little literature addressed the route attractiveness, travel impedance related research is especially reviewed. Review shows travel time is overwhelmingly treated as travel impedance along roadways, and inclusion of other impact factors is rarely found. Various configurations regarding the volume-delay dependency are summarized. Additionally, research finds TTI's performance measures well serve the need for generalizing and quantifying congestion levels within urban areas. However, to date there is no algorithms or modeling procedures integrating link impedance and node impedance. After a thorough review on infrastructure/network, information/data, and externalities/impedance, ***computational methodologies*** are reviewed in the traffic assignment field. It is clear that traffic assignment research has been extensive and fruitful in an intra-urban context since its inception in the 1960s; however, it has not found its way into the long distance travel domain. It is unclear ① to what extent varying assignment algorithms would influence the distributive pattern of flows at interregional level; ② how well the traffic assignment procedures can be transferred to resolving the network loading issue in long distance travel; and ③ how to optimize the trade-off between the theoretical avant-garde and present implementation feasibility. Finally Chapter 2 summarizes the ***implementation*** of regional traffic data platforms, which highlight the value of GIS applications and multi-data-source merging. The thorough review in this chapter reveals the need of constructing a modeling framework, drawing analogies to the human body - with the geographic roadway network as the skeleton, diverse data sources as muscles, and more importantly a computational modeling approach as intelligence.

### *6.1.2 Represent the national highway network in a geodatabase*

Chapter 3 elaborates on research efforts for representing the national highway network to meet modeling needs: ① to depict a network's operational status as well as its linkage to contextual data – nearby land use, activity densities, and availability of facilities that support or attract long distance travel; ② to embrace multiple network scales and maintain topological integrity at each level; ③ to be tailored for the existing data availability and application feasibility and prepared for future utilization and expansion. It concludes that adoption of geodatabase architecture is necessary, because it offers extra functionality, integrates transportation information alongside the roadway network geometry, and supports framework standards. Based on the research findings of Chapter 2, this chapter evaluates available network datasets from a geodatabase construction perspective. With this detailed investigation, it is determined that $FAF^3$ provides desirable network geometry and details for geodatabase design. State, MSA, and County polygon shapefiles represent travel origins and destinations at different levels. A Triple-Level System is proposed, allocating varying datasets to different levels. For the state-level, node features are state centroids (geometric centers of state boundaries), while link features are Interstates only. For MSA-level, node features are MSA centroids, while link features are all of principal arterials. For county-level, node features are county centroids, while link features are both principal and minor arterials. The relationships between all the entities are outlined in the conceptual data model design. The logical data model specifies attributes for the geodatabase, which is independent of planned implementation platforms, while the physical data model takes ArcGIS as implementation platform into consideration. This chapter presents a thorough and detailed geodatabase design process, with all the feature classes, tables, and domains in the logical data model specified. Finally it discusses how traffic data, which include both the historical archive and the most updated traffic conditions, can be aligned alongside the network. A review on the Linear Referencing Systems (LRS) practices at federal and state levels reveals the discrepancy hindering a ready data transfer. The historical archive in HPMS and $FAF^3$ datasets also shows some deficiency in terms of reflecting weekly and seasonal traffic fluctuations. Two-layer solutions are proposed to tackle these challenges. On the basic data layer, a five-step procedure is devised to ensure a smooth transfer of archived data from HPMS and $FAF^3$ to the NHPN base network; on the advanced data layer, high-resolution traffic data are compiled into daily profiles and linked to the $FAF^3$ base network. Due to the inconsistency between state LRS and the $FAF^3$ LRS, the linkage is completed by the longitude and latitude of traffic monitoring stations using the geospatial analysis tool in ArcGIS. Self-correction can be conducted by comparing information from the two layers.

### *6.1.3 Long distance travel network loading: from model to implementation*

Distributive patterns of long distance travel are aggregations of individual route choice behavior. Recognizing the value of understanding long distance travelers' behavior in network loading, the modeling framework follows a workflow designed as: ① segment the entire traveling population, ② identify attributes for each customer base, ③ account for different attributes in one generalized cost model, ④ incorporate the generalized cost model into traffic assignment algorithms, and ⑤ comparison of assignment scenarios. A rough yet effective segmentation captures two major groups: business travelers and pleasure/other travelers. Travel distance and travel time are identified as the two controlling impact factors on route choice, as business trips are mostly made with constrained time windows by people with above-average income level. Pleasure travelers are more diverse, and there exist both tangible route characteristics (route proximity, route attraction, safety, and accessibility rate) and intangible personal attributes (values, needs, and motivations) that influence route choice. For both groups, a generalized cost function is configured as the representation of travel impedance by incorporating those characteristics and attributes within either group. The generalized cost function comprises fixed cost, distance-based operational cost, and time-based monetary cost. The parameter values are determined by empirical justification of available datasets, and also reflect the difference by traveling group. For example, unit operational cost is higher in urban areas than rural; business travel is associated with a higher value of time; BPR functions give stricter penalty on inter-city roadway congestion than within urban areas. Particularly, for pleasure travel, two new parameters are introduced to link contextual data (e.g. scenic byways, roadside attractions, etc.) to the route choice behavior. Both parameters reflect pleasure travelers' evaluation on the "attractiveness" of an alternative route, which is essentially a comprehensively external representation of those internally intangible attributes. Additionally, both parameters are configured to incorporate and simulate different behavior, as explained in Section 4.4.2, using scenic byways and roadside attractions as two examples. Impedance models developed are then integrated in the UE assignment mechanism. Various UE methods are discussed and SUE is further elaborated on how it leverages a random term to reveal people's perception errors in route choice. Integration of link and node impedance, an earlier question raised in Chapter 2, is resolved by applying TTI's urban congestion cost index. In summary, the modeling framework considers two major markets in long distance travel, develops a unified, concise, and behavior-oriented composite impedance model, and reveals the whole network loading procedure. From a theoretical perspective, it extends the intra-urban modeling procedure with an enriched impedance model; from a practical perspective, it demonstrates a working solution to yield service and cost measures over a large-scale network across a diversity of datasets.

The usefulness and feasibility of the modeling framework is demonstrated in a case study, where long distance passenger travel at county level is loaded to highway network within west-coast states of Washington, Oregon, and California, using projected long distance O-D demand data of year 2008. Extensive data acquisition is conducted to ensure the model portraying the reality as much as possible. Different scenarios are designed to ① validate the feasibility of modeling framework, ② test the proposed behavioral hypotheses, ③ investigate the sensitivity of newly introduced parameters, and ④ reveal different distributive patterns under varying scenarios. System-wide performance measures including total VMT, total VHT, total delays, and link usage are reported for each scenario. Finally the trip length distributions from this implementation are compared to the 1995 ATS data. Some key findings from the case study include: ① for SUE methods, 25% perception error is a threshold that triggers a rapid and linear increase in total VMT and VHT, regardless if the distribution of the error term; ② when error terms range from roughly 5% to 75%, the SUE-based total delays are lower than SO-based total delays, because certain individuals' non-optimal route choices ease the congestion on heavily traveled routes and enable others to travel much more efficiently; ③ inclusion of the distance-based operational cost generally results in an increased total VHT, a decreased total VMT, and decreases of total delays, all with mild changes, but the inclusion will cause counter-intuitive VMT – VHT relationship by different assignment algorithms and thus should be interpreted carefully; ④ long distance travel delay is lower than intra-urban travel delay; ⑤ both newly introduced parameters are capable of representing travelers' valuation towards scenic highways and roadside attractions; however, the *s* factor presents a very consistent and stable characteristic across assignment algorithms and the

$a$ factor can result in very different performance measures when assignment algorithms vary. For example, when $a$ = 1, compared to using the SUE_Gumbel, using DUE will increase the total VHT by 32%, the total VMT by 21%, and total delays by 42%. Both parameters are also effective of changing the distributive patterns of pleasure travel; ⑥ SUE diversifies route choice and yield higher lank usage than DUE method, and the differences range from 6% to 18.6%; ⑦ comparing trip length distributions from the 1995 ATS and this case study finds that the 1995 ATS trip length distribution underestimated the number of trips in "longer than 300 miles" categories, whereas the 300 to 700-mile travel market is critical to the high-speed rail.

## 6.2 Research Contribution

Modeling long distance travel presents a large-scale problem with great complexity. Through the aforementioned efforts, this research makes contributions to understanding the distributive pattern of long distance passenger flows in three dimensions: data issues, modeling framework, and policy evaluations.

***Data issues***: Supply-side information, namely, the service quality of the transportation infrastructure and costs for competing traveling options, is critical to supporting decisions about transportation policies, investments, and operations. While MPOs normally collect both supply and demand data to support the development of regional models, there is no detailed measure of highway performance and its costs for long distance travel. Utilizing available resources, this research fills this data gap. On the one hand, results from the modeling framework developed (at least as shown in the case study) portray various distributive patterns of passenger flows over the highway network. It explicitly highlights the variations of corridors selected by long distance travelers under different behavioral assumptions. System-wide performance measures, total VMT, VHT, and delays are estimated simultaneously as well. On the other hand, costs for different travel markets, covering a diversity of cost components, are explicitly estimated, which provides a critical pricing reference for justifying a competing travel mode (e.g. high-speed rail).

***Modeling framework***: this research presents a complete procedure for long distance network loading. Specifically, it develops the travel impedance models in a way that contextual data are incorporated into the framework, influencing and simulating pleasure travelers' behavior, a market with over 60% share of total number of long distance trips. As illustrated in the case study, contextual data like scenic highway designation and roadside attractions fundamentally influence many key travel choices including the route choice. Those contextual data, however, are seldom gathered in large-scale surveys or rarely applied to analyzing passenger travel and providing the information needed to probe these influences. Additionally, very little previous research gave attention to understanding what behavioral foundation motivates and causes route choice diversity in long distance travel. Although still in an exploratory stage, this research analyzes different travel markets and shows how different magnitude of the "behavioral" factor can change the distributive patterns. This understanding is important not only for designing and evaluating policies that involve changing travel behavior, but also for more basic purposes, such as designing travel surveys and other data collection activities.

***Policy evaluations***: Accounting for roughly 16% of the national total VMT, long distance travel generates immense impacts on the highway infrastructure in a broad spectrum of aspects:

tourism economy, travel patterns of an again population, high-speed intercity rail, energy efficiency, and greenhouse gas emissions. By providing quantitative system-wide indicators and performance measures, the modeling framework facilitates related policy evaluation. Particularly, by representing the network highway network in a geodatabase, the modeling framework adds another dimension of policy accountability: ***geographic specificity***. This dissertation study demonstrates how we can better understand the relationships between travel and contextual factors and to construct models for policy evaluation through geocoding and linking separate datasets. The geodatabase setup also supports map-based analysis and display of data, an important way to visualize and understand travel patterns.

## 6.3 An outlook on the Traveler Analysis Framework and roles of modeling supply-side data of long distance travel

Policy makers are looking for more knowledge and information rather than data about availability and performance of competing travel options, the political and economic context of traveling, and the complete and accurate description of travel behavior. Conventional travel surveys for transportation planning and data collection for traffic operations cannot meet all these needs. Meanwhile, fruitful development of regional travel demand models cannot well answer interregional questions, creating hurdles for federal level in understanding the big picture at national level. In a nose-diving economy, the idea of constructing a Traveler Analysis Framework incepts, exploring the potentials of utilize existing data resources to avoid the expense of "one-off" surveys. Hopefully a well-constructed Traveler Analysis Framework will address the growing importance of multi-regional, corridor and multinational analyses, as well as the need for policy analyses: transportation equity, pricing, financing, and planning.

This research serves as an organic component within the Traveler Analysis Framework. First of all, it answers some key questions on the supply side of long distance travel: what is the difference between business travelers and pleasure travelers? Which corridor are they choosing? What impacts are they generating? Second of all, it takes travelers' behavior into consideration: how do pleasure travelers respond to contextual factors? How to capture those factors in a quantitative way? Last but not the least, although the study does not yield formative answers to the questions above (given the reality of existing data sources); it enumerates a limited number of possible scenarios and extrapolates interesting observations.

## 6.4 Future Research

Long distance travel is a broad area accommodating multiple research efforts. Due to the complexity presented by the long distance travel network loading problem, this study is limited in scope, and is meant to provide a prologue for future adventure. Although a general modeling framework is developed, factors included in the travel impedance model and their quantifications are incomplete. The research can evolve in several directions:

- A finer segmentation of the traveling population: both means-end theory and Analytic Hierarchy Process (AHP) are potential approaches revealing the diversity across long distance travelers. New attributes influencing the individual route choice can also be identified through new surveys.
- Introduce randomness to the $s$ factor and the $a$ factor. Currently, they are considered as

deterministic values; however, consider the diverse traveling population, an assumed distribution would better capture the variations in people's choices.

- Incorporate more realistic behavioral assumptions in the network loading modeling. Boundedly Rational User Equilibrium (BRUE) and Behavioral User Equilibrium (BUE) proposed in last decade, as well as agent-based travel modeling could be promising next steps.

The author would like to conclude the dissertation by quoting President Dwight D. Eisenhower in one of his letters to the Congress in 1955. "*Our unity as a nation is sustained by free communication of thought and by easy transportation of people and goods. The ceaseless flow of information throughout the Republic is matched by individual and commercial movement over a vast system of interconnected highways crisscrossing the country and joining at our national borders with friendly neighbors to the north and south. Together, the united forces of our communication and transportation systems are dynamic elements in the very name we bear—the United States. Without them, we would be a mere alliance of many separate parts.*"

# Bibliography

Abdelghany AF, Mahmassani HS, Chiu YC (2001) *Spatial microassignment of travel demand with activity trip chains*. Transportation Research Record 1777:36-46

Abdelghany AF, Mahmassani HS, Chiu YC (2003) *Spatial microassignment of travel demand with activity trip chains.* Transportation Research Record 1831:89-97

Adams, T.M., A.P. Vonderohe, J.A. Butler, and N. Koncz. (1999). *Model Considerations for a Multimodal, Multidimensional Location Referencing System*. Proceedings GIS-T Symposium, San Diego, CA.

Anable, J. (2002). Picnics, Pets, and Pleasant Places: The Distinguishing Characteristics of Leisure Travel Demand. *Social Change and Sustainable Transport*. Indiana University Press, Bloomington, IN, pp. 181-190.

Ashiabor, S., A. A. Trani, H. Baik and N. K. Hinze. (2007). *Development of Intercity Model Choice Models for New Aviation Technologies*. Aviation: A World of Growth, Proceedings of the 29th International Air Transport Conference, pp. 61-77.

Ashiabor, S., H. Baik and A. A. Trani. (2007). *Logit Model for Forecasting Nationwide Intercity Travel Demand in the United States*. Transportation Research Record #2007, pp. 1-12.

Athol, P. (1965). *Interdependence of Certain Operational Characteristics within A Moving Traffic Stream*. Highway Research Record 72, 58-87.

Babin, A. and Pierre Tremblay. (1994). *Experimentation of "S-curve" type Volume/delay Functions* For the Montreal Region 9th International Emme/2 Users' Conference

Baker, W.T. and W.E. Blessing. (1974). *Highway Location Reference Methods*. NCHRP Synthesis of Highway Practice 21. Transportation Research Board, National Research Council, Washington, D.C.

Ban, X., Liu, H., Ferris, M.C., and Ran, B., (2008). *A link-node complementarity model and solution algorithm for dynamic user equilibria with exact flow propagations*. Transportation Research, part B, 42(9), 823-842.

Battelle. (2011). *FAF³ Freight Traffic Analysis. Final Report Submitted to Oak Ridge National Laboratory*. Accessed online from: faf.ornl.gov/fafweb/Data/Freight_Traffic_Analysis/faf_fta.pdf

Ben-Akiva, M., Bierlaire, M., Koutsopoulos, H. N., and Mishalani, R. (2002). *Gendreau, M. and Marcotte, P., editors, Transportation and network analysis: current trends, chapter Real-time simulation of traffic demand-supply interactions within DynaMIT*, pages 19–36.

Bierlaire, M., E. Frejinger, and J. Stojanovic. (2006). A latent route choice model in Switzerland. *Proceedings of European Transport Conference 2006: Advances in Discrete Choice Models*.

Black, W. R. (1972). *Interregional Commodity Flows: Some Experiments With the Gravity Model*. Journal of Regional Science. Volume 12, Issue 1, pages 107–118.

Blue, V., List, G. F., and Embrechts, M. J., (1994). *Neural net freeway travel time estimation*. Proceedings of Intelligent Engineering Systems through Artificial Neural Networks, vol. 4. American Society of Mechanical Engineers, New York, pp. 11135–11140.

Bovy, P. H. L., and Thijs, R., (2000). *Estimators of travel time for road networks: New developments, evaluation results, and applications*. Delft University Press, The Netherlands.

Briganti, C.M., and L.A. Hoel. (1994). *Transportation System Design and Information Requirements for Travel and Tourism Needs on Scenic Byways* (Report No. VTRC95-R1). Virginia Transportation Research Council, Charlottesville, VA.

Bureau of Transportation Statistics. (1997). 1995 *American Travel Survey Profile*. Accessed online from http://www.bts.gov/publications/1995_american_travel_survey/us_profile/entire.pdf

Bureau of Transportation Statistics. (1998). *Transportation Statistics Annual Report 1998 Long-Distance Travel and Freight*. Washington, D.C.

Bureau of Transportation Statistics. (2003). *America on the Go: Findings from the National Household Travel Survey- U.S. Business Travel*. U.S. Department of Transportation. Accessed online from http://www.bts.gov/publications/america_on_the_go/us_business_travel/pdf/entire.pdf

Butler, J.A. (2008). *Designing Geodatabases for Transportation*. ESRI Press. Redlands, California.

Cambridge Systematics, Inc. (2008). *National Travel Demand Forecasting Model Phase I Final Scope*. NCHRP Project 836-B, Task 70

Cambridge Systematics, Inc., (2008). *Estimated Cost of Freight Involved in Highway Bottlenecks*. Prepared for FHWA Office of Transportation Policy Studies.

Chen, C., J. Kwon, J. Rice, A. Skabardonis, and P. Varaiya. (2003). *Detecting errors and imputing missing data for single loop surveillance systems*. Transportation Research Record: Journal of the Transportation Research Board. Vol. 1855, pp. 160–167.

Cheung, S. Y., Coleri, S., Dundar, B., Ganesh, S., Tan, C-W., and Varaiya, P., (2005). *Traffic Measurement and Vehicle Classification with a Single Magnetic Sensor*. Technical Report, 84th Annual Meeting, Transportation Research Board.

Chin, S.M., and H.L. Hwang. (2001). Identifying Passenger Corridors on the U.S. Highway System Using ATS Data. *TRB Transportation Research Circular E-C026*. Transportation Research Board of the National Academies, Washington, D.C.

Cho, M.H. and S.C. Jang. (2008). Information Value Structure for Vacation Travel. *Journal of*

*Travel Research*.

Clark, S. and Watling, D. (2005) *Modelling network travel time reliability under stochastic demand*. Transportation Research Part B, Vol. 39, No. 2, 119-140.

Cohen, E. (1972). Toward a Sociology if International Tourism, Social Research, Vol. 39(1), pp. 164-182.

Cohen, H., Horowitz, A., and Pendyala, R. (2008). *Forecasting Statewide Freight Toolkit*. Transportation Research Board, Washington D.C.

Coifman, B., (2001). *Improved Velocity Estimation Using Single Loop Detectors*. Transportation Research: Part A, Vol. 35, No. 10, pp. 863–880.

Coifman, B., S. Dhoorjaty, and Z. Lee. (2003). *Estimating Median Velocity Instead of Mean Velocity at Single Loop Detectors*. Transportation Research: Part C, Vol. 11, No. 3-4, 211-222

Cortes, C. E., Lavanya, R., Oh, J. S., and Jayakrishnan, R., (2002). *A general purpose methodology for link travel time estimation using multiple point detection of traffic*. Presented at the 81st Annual Meeting (CD-ROM), Transportation Research Board, Washington D.C.

Dafermos, S.C. (1972). The Traffic Assignment Problem for Multiclass-User Transportation Networks. *Transportation Science*. Vol. 6, No. 1, pp. 73-87.

Daganzo, D.F. and Y.Sheffi. (1977). On Stochastic Models of Traffic Assignment. *Transportation Science*. Vol. 11, No. 3, pp. 253-274.

Dahlquist, D.L. and K. Peterson. (1997). Assessing visitor responses to Iowa scenic byways pilot program. *Transportation Research Record: Journal of Transportation Research Board*. National Research Council, Washington, D.C., Vol. 1599, pp. 81–95.

Dailey, D. J., (1999). *A statistical algorithm for estimating speed from single loop volume and* *occupancy measurements*. Transportation Research, Part B, Vol. 33, No. 5, 313-322.

Davies, C. (2005). *2005 Travel and adventure report: A snapshot of boomers' travel and adventure experiences*. Washington, D.C.: American Association of Retired Persons.

Davis, S. C., S. W. Diegel, and R. G. Boundy. (2011). *Transportation Energy Data Book: Edition 30*. Office of Energy Efficiency and Renewable Energy, U.S. Department of Energy.

Dhulipala, S., (2002). *A system for travel time estimation on urban freeways*. Master's thesis, Department of Civil and Environmental Engineering, Virginia Polytechnic Institute and State University, Blacksburg, Virginia.

Donnelly, Rick, Erhardt, Greg D, Moeckel, Rolf, and Davidson, William A. (2010). *Advanced Practices in Travel Forecasting*. NCHRP Project 20-5, Topic 40-06. Transportation Research Board.

Eby and Molnar. (2002). Importance of scenic byways in route choice: a survey of driving tourists in the United States. *Transportation Research Part A: Policy and Practice*. Vol. 36, pp. 95–106.

Ewing, R., William Schroeer, William Greene. (2004). *School Location and Student Travel Analysis of Factors Affecting Mode Choice*. Transportation Research Record #1895, pp. 55-63.

Fang, J., Meng, H., Zhang, H., and Wang, X., (2007). *A Low-cost Vehicle Detection and Classification System based on Unmodulated Continuous-wave Radar*. Intelligent Transportation Systems Conference, IEEE , pp.715-720.

Fekpe, E. S., T. Windholz, K. Beard, and K. Novak. (2003). *Quality and Accuracy of Positional Data in Transportation*. National Cooperative Highway Research Program Report 506, Transportation Research Board, Washington, D.C.

FHWA. (1989). *FHWA Functional Classification Guidelines*. Retrieved online at http://www.fhwa.dot.gov/planning/fcsec1_1.htm

FHWA. (2010). *Our Nation's Highways 2010*. Publication Number: FHWA-PL-10-023

Fitzpatrick, K., P. Carlson, M. A. Brewer, M.D. Wooldridge, and S.P. Miaou. (2003). *Design Speed, Operating Speed, and Posted Speed Practices*. Transportation Research Board, Washington, D.C.

Focalyst. (2007). *The sky's the limit: travel trends among the baby boom generation and beyond leisure travel trends*. American Association of Retired Persons.

Friesz, T. L., David Bernstein, Tony E. Smith, Roger L. Tobin and B. W. Wie. (1993). *A Variational Inequality Formulation of the Dynamic Network User Equilibrium Problem*. Operations Research, 41(1), 179-191

Friesz, T.L., Bernstein, D., Smith, T.E., Tobin, R.L. and Wie, B.W. (1993) *A Variational inequality formulation of the dynamic network user equilibrium problem*. Operations Research 41, 179-191.

Friesz, T.L., David Bernstein, Tony E. Smith, Roger L. Tobin and B. W. Wie. (1993). *A Variational Inequality Formulation of the Dynamic Network User Equilibrium Problem*. Operations Research, 41(1), 179-191

Georggi, N.L. and R.M. Pendyala. (2001). Analysis of Long-Distance Travel Behavior of the Elderly and Low Income. *TRB Transportation Research Circular E-C026*. Transportation Research Board of the National Academies, Washington, D.C..

Georgia Department of Transportation. (2009). *Southwest Georgia Interstate Study Technical Memorandum Highway Network Development*. Accessed online on Jan 5th, 2012 from http://www.dot.state.ga.us/informationcenter/programs/studies/SWGAInterstate/Documents/

ModelDevelopment/Development%20of%20Network.pdf

Giaimo, G.T. and Schiffer, R. (editors). (2005). *Statewide Travel Demand Modeling: A Peer Exchange*. Transportation Research Circular #E-C075, Transportation Research Board.

Gudzinas, B. (2011). *National Transportation Model: Needs of Long-Distance OD Data*. FHWA Office of Highway Policy Information, NHTS Workshop.

Haoui, A., Kavaler, R., and Varaiya, P., (2008). *Wireless magnetic sensors for traffic surveillance*, Transportation Research Part C: Emerging Technologies, Volume 16, Issue 3, pp.294-306.

Haynes, K. E., Fotheringham A. S. (1984). *Gravity and spatial interaction models*. Sage Publications, Beverly Hills, CA

Henderson, J. and A. A. Trani. (2008). *Temporal Transferability of Long-Distance Trip Rates*. Proceedings of the 87th Annual Meeting of the Transportation Research Board.

Hensher, D.A. and J.M. Rose. (2006). Respondent Behavior in Discrete Choice Modeling with a Focus on the Valuation of Travel Time Savings. *Journal of Transportation and Statistics*, Vol. 8 No. 2.

Hicks, J.E., (2008). *A Dynamic Traffic Assignment Model Breakdown*, In Conference Proceedings 42: Innovations in Travel Demand Modeling, Volume 2: Papers, Transportation Research Board of the National Academies, Washington, D.C., pp. 101–108.

Hoogendoorn, S. P., (2000). *Model-based multiclass travel time estimation*. Presented at the 79th Annual Meeting (CD-ROM), Transportation Research Board, Washington D.C.

Horowitz, A.J. (2006). *Statewide Travel Forecasting Models*. NCHRP Synthesis #358, Transportation Research Board.

Horowitz, A.J. (2008). *White paper: Statewide travel demand forecasting*. Requested by

AASHTO and presented at the Conference on Meeting Federal Surface Transportation Requirements in Statewide and Metropolitan Transportation Planning.

Hotel News Resource. (2007). *Consumers Aren't Letting Record Gasoline Prices Get In The Way of Their Next Vacation*. Accessed online from http://www.hotelnewsresource.com/article27802.html

Huang, H. and S. Gao. (2012). Optimal paths in dynamic networks with dependent random link travel times. *Transportation Research Part B*, Vol. 46, pp. 579–598.

Huang, H., Lam, H., (2002). *Modelling and solving the dynamic user equilibrium route and departure time choice problem in network with queues*. Transportation Research B 36, 257–273.

Hutchinson, B. G. (1974). *Principles of urban transportation systems planning*. McGraw-Hill, New York

Hwang, H.L. and J.Rollow. (2000). *Data Processing Procedures and Methodology for Estimating Trip Distances for the 1995 American Travel Survey (ATS)*. Report prepared by the Oak Ridge National Laboratory for the U.S. Department of Energy under Contract No. DE-AC05-96OR22464

Janson, B.N. (1991) *Dynamic traffic assignment for urban road networks*. Transportation Research 25B,143-161

Jastrzebski, W. P. (2000). *Volume delay functions*, 15th Inter EMME/2 Users' group conference

Jenkins, D. and D. Vary. (2010). *FHWA Traveler Analysis Framework Part II Status Report*. Presentation at the AMPO Travel Modeling Meeting. Washington, D.C.

Wilbur Smith Associates. (2011). Internal Memorandum: Traveler Analysis FrameworkTask 3 – 2008 Modal Passenger Data Methodology. A working draft acquired by the author.

Kemperman, A.D.A.M., A.W.J. Borgers, and H.J.P. Timmermans. (2009). Tourist shopping behavior in a historic downtown area. *Tourism Management*. Vol. 30, pp. 208–218.

Kim H, Oh JS, Jayakrishnan R (2006) *Activity chaining model incorporating time use problem and its application to network demand analysis*. Proceedings of the 85th Transportation Research Board Meeting, Washington DC

Klein, L.A. (2001). *Sensor Technologies and Data Requirements for ITS*. Artech House, Norwood, MA.

Lam WH, Huang HJ (2003) *Combined activity/travel choice models: Time-dependent and dynamic versions*. Network and Spatial Economics 3:323-347

LaModia, J. J. and C. Bhat. (2007). *Modeling Side Stop Purpose During Long-Distance Travel Using 1995 American Travel Survey*. Proceedings of 86th Annual Meeting of the Transportation Research Board.

Lanzendorf, M. (2002). Mobility Styles and Travel Behavior: Application of a Lifestyle Approach to Leisure Travel. *Transportation Research Record: Journal of Transportation Research Board*. National Research Council, Washington, D.C., Vol. 1807, pp. 163-173.

Levinson, D. (1993). *Integrating Feedback Into Transportation Planning Model: Structure and Application*. Transportation research record (0361-1981), p. 70.

Levinson, D. (1998). *Accessibility and the journey to work*. Journal of Transport Geography. Volume 6, Issue 1, Pages 11-21

Lim, C. G. (2008). *An Integrative Assessment of the Commercial Air Transportation System via Adaptive Agents*. PhD Dissertation. School of Aerospace Engineering, Georgia Institute of Technology

Lindveld, C. D. R., Thijis, R., Bovy, P. H., and Van der Zijpp, N. J., (2000). *Evaluation of online* *travel time estimators and predictors*. Transportation Research Record 1719, Transportation Research Board, Washington, D.C., 45-53.

Liu, F., R. G. Kaiser, M. Zekkos and C. Allison. (2006). *Growth Forecasting of Vehicle Miles of Travel at County and Statewide Levels*, Transportation Research Record #1957, pp. 56-65.

Liu, R. R. and G. Li. (2005). *The Safety Perceptions of Road, Rail and Air and Their Implication on Travel Choices*. Advances in Transportation Studies, Vol. 5, pp. 5-19.

Lo, H.K. and Szeto, W.Y. (2002a) *A cell-based variational inequality formulation of the dynamic user optimal assignment problem*. Transportation Research Part B, 36, 421-443.

Lo, H.K. and Szeto, W.Y. (2002b) *A cell-based dynamic traffic assignment model: formulation and properties*. Mathematical and Computer Modeling, 35, 849-865.

López-Valcarce, R., Mosquera, C., and Pérez-González, F., (2004). *Estimation of road vehicle speed using two omnidirectional microphones: a maximum likelihood approach*, EURASIP Journal on Applied Signal Processing, pp.1059-1077.

Mahmassani, H. S. (2001). *Dynamic network traffic assignment and simulation methodology for advanced system management applications*. Networks and Spacial Economics, 1(3):267–292.

Mahmassani, H.S. (2001). *Dynamic Network Traffic Assignment and Simulation Methodology for Advanced System Management Applications*. Networks and Spatial Economics. Volume 1, Numbers 3-4, 267-292, DOI: 10.1023/A:1012831808926

Mahut, M., M. Florian, D. Florian, S. Velan and N. Tremblay. (2005). *Equilibrium Dynamic Traffic Assignment for Large, Congested Network*. INRO White Paper

Maki, K., T. Kazuyuki, Y. Shohei, and S.Kazuya. (2007). Evaluation of Value of Travel Time Saving for Intra-/Inter-City Trip Using RP/SP Data. *Proceedings of the 11th World Conference on Transport Research*, Berkeley, CA.

Malinovskiy, Y., Yao-Jan Wu, Yinhai Wang, and Un-Kun Lee. (2010). *Field Experiments on Bluetooth-based Travel Time Data Collection*. Preprint CD-ROM, the 89th Annual Meeting of Transportation Research Board, Paper 10-3134.

Mallett, W. and N. McGuckin. (2000). Driving to distractions: recreational trips in private vehicles. *Transportation Research Record: Journal of the Transportation Research Board*, Vol. 1719, pp. 267-272. Washington, D.C.

Martin, W.A. and N.A. McGuckin. (1998). *NCHRP Report 365: Travel Estimation Techniques for Urban Planning*. Transportation Research Board, Washington D.C.

Merchant, D.K. and G.L. Nemhauser, (1978), *Optimality conditions for a dynamic traffic assignment model*. Transportation Science, 12 pp. 200–207.

Merchant, D.K. and G.L. Nemhauser, (1978), *A model and an algorithm for the dynamic traffic assignment problem*. Transportation Science, 12 pp. 183–199.

Meyer M. and Eric Miller. (2001). *Urban Transportation Planning*. McGraw-Hill.

Monzon, A. and A. Rodriguez-Dapena. (2006). *Choice of Mode of Transport for Long*- *Distance Trips: Solving the Problem of Sparse Data*. Transportation Research A, Vol. 40, No. 7, pp. 587-601.

Mouskos, K. C., Bahman Izadmehr, and Athanasios K. Ziliaskopoulos. (2003). *Implementation of the Visual Interactive System for Transportation Algorithms (VISTA) in the United States.* Iranian Academic Association (IAA), Application of Technology in Urban Development, December 21 – 28, 2003, IRAN

Muñoz, J. and J.A. Laval, (2005), *System optimum dynamic traffic assignment graphical solution method for a congested freeway and one destination*. Transportation Research Part B, 40 1 pp. 1–15.

Nam, D. H., and Drew, D. R., (1998). *Analyzing freeway traffic under congestion: Traffic dynamics approach*. Journal of Transportation Engineering, ASCE, 124(3), pp. 208-212.

Newman, C. (2001). Travelers Seek Byway Experiences. *Public Roads*. Vol. 05-64 (6), pp. 33-39.

Nielsen, O. A. and A. Burgess. (2008). *The European TRANS-TOOLS transport model*. Proceedings of the 87th Annual Meeting of the Transportation Research Board.

Nostrand, C.V., A. Pinjari, and V. Sivaraman. (2011). Analysis of annual, long-distance, vacation travel demand in the United States: a multiple discrete-continuous choice framework. *Proceedings of International Choice Modeling Conference 2011*.

Oak Ridge National Laboratory. (2012). *Introduction of Freight Analysis Framework 3*. Accessed online on Nov 25th, 2011 from http://faf.ornl.gov/fafweb/

Oh, S., Ritchiel, S.G., and Oh, C., (2002). *Real-Time Traffic Measurement from Single Loop Inductive Signatures*. Transportation Research Record 1804, TRB, National Research Council, Washington, D.C., 98-106.

Pack, M.L., Bryan, J., Steffes, A.M., (2008). An Overview and Status of the Regional Integrated Transportation Information System (RITIS) in the National Capital Region, Preprint CD-ROM for the 87th Annual Meeting of Transportation Research Board, Washington, D.C.

Palacharla, P. V., and Nelson, P. C., (1999). *Application of fuzzy logic and neural networks for dynamic travel time estimation*. International Transactions in Operational Research, 6, pp. 145-160.

Park, M.B. and Robert L. Smith. (1997). *Development of a Statewide Truck-Travel Demand Model with Limited Origin-Destination Survey Data*. Transportation research record (Volume 1602 /1997), p. 70

Peeta, S., A. Paz and D. DeLaurentis. (2007). *Stated Preference Analysis of New Microjet On-Demand Air Service*. Proceedings of the 86th Annual Meeting of the Transportation Research Board.

Peeta, S., A. Paz and D. DeLaurentis. (2008). *Stated Preference Analysis of a New Very Light Jet Based On-Demand Air Service*. Transportation Research A, Vol. 42, No. 4, pp. 629-645.

Perreault, W.D., D.K. Darden, and W.R. Darden. (1977). A Psychological Classification of Vacation Life Styles. *Journal of Leisure Research*, Vol. 9, pp. 208-224.

Petty, K., J. Kwon, and A. Skabardonis. (2005). *Arterial Performance Measurement System—APeMS*. In Workshop on Performance Measures, 85th TRB Annual Meeting. CD-ROM. Washington D.C., 2005.

Pushkar, A., F.L. Hall, and J.A. Acha-Daza., (1994). *Estimation of Speeds From Single-Loop Freeway Flow And Occupancy Data Using Cusp Catastrophe Theory Model*. Transportation Research Record 1457, TRB, National Research Council, Washington, D.C., 149-157.

Qian, Z., and M. Zhang. (2012). On centroid connectors in static traffic assignment: their effects on flow patterns and how to optimize their selections. *Transportation Research Part B*, in press.

Quiroga, C. A. and Bullock, D. (1998). *A Data Model for Managing GPS Travel Time Data*. Proceedings of the Fifth International Conference on Applications of Advanced Technologies in Transportation Engineering.

Quiroga, C., (2000). *Assessment of dynamic message travel time information accuracy*. Proceeding of the North American Travel Monitoring Conference and Exposition, Middleton, Wisconsin, 1- 13.

Ramadurai, G. and Satish Ukkusuri. (2008). *Dynamic User Equilibrium Model for Combined*

*Activity-Travel Choices Using Activity-Travel Supernetwork Representation*. Region 2 University Transportation Research Center Report. Retrieved online at http://www.utrc2.org/research/assets/132/DUE-Satish1.pdf

Ran B., Hall R.W. and Boyce D.E. (1996). *A link-based variational inequality model for dynamic departure time/route choice*, Transportation Research , 30B, 31-46.

Ran, B. and Shimazaki T. (1989). *A General Model and Algorithm for the Dynamic Traffic Assignment Problems*. Transport Policy, Management and Technology Towards 2001, Fifth WCTR, Yokohama, Japan, 463-477.

Ran, B., Boyce, D.E. and LeBlanc, L.J. (1993) *A New class of instantaneous dynamic user-optimal traffic assignment models*. Operations Research 41, 192-202.

Rieser M, Nagel K, Beuck U, Balmer M, Rmenapp J (2007) *Agent-oriented coupling of activity-based demand generation with multi-agent traffic simulation*. Transportation Research Record 2021:10-17

Sawant, H.; Jindong Tan; Qingyan Yang; Qizhi Wang. (2004). *Using Bluetooth and sensor networks for intelligent transportation systems*. Proceedings. The 7th International IEEE Conference on Intelligent Transportation Systems.

Schäfer, R. P., Kai-Uwe Thiessenhusen, Peter Wagner. (2002). *A Traffic Information System by Means of Real-Time Floating-Car Data*. The 9th World Congress on Intelligent Transport Systems Chicago.

Schneider, P.P., and C.A. Vogt. (2005). Leisure Activity Patterns and Household Composition: The Changing Household. *Proceedings of the 2005 Northeastern Recreation Research Symposium*, pp. 308-315.

Sem, J., P. Goff, and S. Pearce. (1997). *Interim report on Colorado scenic and historic byways* *economic-impact study*. *Transportation Research Record: Journal of Transportation Research Board*. National Research Council, Washington, D.C., Vol. 1599, pp. 86–95.

Sethi, V. and F.S. Koppelman. (2001). Incorporating Complex Substitution Patterns and Variance Scaling In Long Distance Travel Choice Models. *Travel Behavior Research: The Leading Edge*. pp. 375-395.

Sheffi, Y. (1985). *Urban Transportation Networks: Equilibrium Analysis with Mathematical Programming Methods*. Prentice-Hall, Inc.

Simons, C., (2006). *I-285 Matrix Variegator: Practical Method for Developing Trip Tables for Simulation Modeling from Travel Demand Modeling Inputs*, Transportation Research Record: Journal of the Transportation Research Board, No. 1981, Transportation Research Board of the National Academies, Washington, D.C., pp. 18–23.

Souleyrette, R. R., Z. N. Hans, S., Pathak. (1996). *Statewide transportation planning model and methodology development program*. Final Report to Iowa Department of Transportation and Midwest Transportation Center

Spiess, H. (1990). Conical Volume-Delay Functions. Transportation Science, Vol 24., No. 2.

Srinivasan, S., C. R. Bhat and J. Holguin-Veras. (2006). *An Empirical Analysis of the Impact of Security Perception on Intercity Mode Choice: A Panel Rank-Ordered Mixed Logit Model*. Transportation Research Record #1942, pp. 9-15.

Stewart, F. (1993). UK Leisure Trends and the Prospects after Skiing. Paper presented at the 1993 National Ski Conference, Olympia, London.

Sun, C., and Ritchie, S.G., (1999). *Individual Vehicle Speed Estimation Using Single Loop Inductive Waveforms*. UCB-ITS-PWP-99-14. California Partners for Advanced Transit and Highways (PATH), Institute of Transportation Studies, UC Berkeley

Szeto W.Y. and Lo H.K. (2004) *A cell-based simultaneous route and departure time choice model with elastic demand.* Transportation Research B, 38, 593-612.

Taplin J.H.E. and M. Qiu. (1997). Car trip attraction and route choice in Australia. *Annals of Tourism Research*. Vol.24, No. 3, pp. 624-637.

Tong, D., Carolyn J. Merry, and Benjamin Coifman. (2005). *Traffic Information Deriving Using GPS Probe Vehicle Data Integrated with GIS*. 2006 AASHTO GIS for transportation Symposium Proceedings. Retrieved online at: http://www.gis-t.org/files/13auY.pdf

Train, K.E. (1998). Recreation Demand Models with Taste Differences over People. *Land Economics*, Vol. 74, No. 2.

Train, K.E. (2009). *Discrete Choice Methods with Simulation*. Cambridge University Press; 2 edition

Tsukai, M. and M. Okumura. (2003). Analysis of Inter-City Passenger's Route Choice Behavior On Non-Shortest-Time Routes. *Journal of the Eastern Asia Society for Transportation Studies*, Vol.5.

Tufte, K. A., Bertini, R.L., Chee, J., Fernández-Moctezuma, R. J., Periasamy, S., Sarkar, S., Singh, P., Whiteneck, J., (2010). *Portal 2.0: Towards a Next-Generation Archived Data User Service*, Preprint CD-ROM for the 89th Annual Meeting of Transportation Research Board, Washington, D.C.

Tyrrell, T. J. and M. F. Devitt. (1999). Valuing Changes to Scenic Byways. *Consumer Behavior in Travel and Tourism*. The Haworth Hospitality Press, New York, 1999.

U.S. Department of Transportation. (1999). *The 1995 American Travel Survey*. Bureau of Transportation Statistics. Microdata CD-ROM.

U.S.DOT. (1997). *Departmental Guidance on the Evaluation of Travel Time in Economic*

*Analysis*. Memo used in STEAM software (www.ota.fhwa.dot.gov/steam).

Ukkusuri, S. V. S. K. & Waller, S. T. (2008), *Linear programming models for the user and system optimal dynamic network design problem: formulations, comparisons and extensions*, Networks and Spatial Economics, 8(4), 383–406.

Van Arem, B., Van der Vlist, M. J. M., Muste, M. R., and Smulders, S. A., (1997). *Travel time estimation in the GERDIEN Project*. International Journal of Forecasting, 13, pp. 73 - 85.

Van Lint, J. W. C., and van der Zijpp, N. J., (2003). *An improved travel time estimation algorithm using dual-loop detectors*. Presented at the 82nd TRB Annual Meeting (CDROM), Transportation Research Board, Washington D. C.

Varaiya, P., (2004). *Assessment of MeMS Sensors in an Urban Traffic Environment*. UCB-ITS-PRR-2004-8. California Partners for Advanced Transit and Highways (PATH), Institute of Transportation Studies, UC Berkeley

Vickrey, W.S. (1969). *Congestion theory and transport investment*. American Economic Review, 59 2, pp. 251–260.

Viswanathan, K. and D. Vary. (2012). Development of a Nationwide Long Distance Auto Trip Table. *Proceeding of the 4th Transportation Research Board Conference on Innovations in Travel Modeling (ITM)*. Accessed online from http://onlinepubs.trb.org/onlinepubs/conferences/2012/4thITM/Papers-R/0117-000030.pdf

Vonderohe, A.P., C.L. Chou, F. Sun, and T.M. Adams. (1997). *A Generic Data Model for Linear Referencing Systems*. NCHRP Research Results Digest 218. Transportation Research Board, National Research Council, Washington, D.C.

Vonderohe, A.P., L. Travis, R. L. Smith, V. Tsai. (1993). *Adaptation Of Geographic Information Systems For Transportation*. National Cooperative Highway Research Program Report 359, Transportation Research Board, Washington, DC

Wang, Y., and N.L. Nihan., (2003). *Can Single-Loop Detectors Do the Work of Dual-Loop Detectors?* ASCE Journal of Transportation Engineering, Vol. 129, No. 2, 169 - 176.

Wang, Y., Papageorgiou, M., (2005). *Real-time freeway traffic state estimation based on extended Kalman filter: a general approach*, Transportation Research Part B: Methodological, Volume 39, Issue 2, pp. 141-167.

Wang, Y., Yegor Malinovskiy, Yao-Jan Wu, and Un Kun Lee. (2010). *Error Modeling and Analysis for Travel Time Data Obtained from Bluetooth MAC Address Matching*. Transportation Northwest (TransNow) Research Report TNW2010-05.

Wang, Yinhai, and Nancy L. Nihan. *Freeway Speed Estimation Using Single Loop Outputs*. Transportation Research Record, No. 1727, 120-126, 2000.

Wardrop, J.G. (1952). *Some theoretical aspects of road traffic research*. Proceedings of the institute of Civil Engineers. Vol. 1.

Wie, B.W., Tobin, R., Bernstein, D. and Friesz, T. (1995) *A Comparison of system optimal and user optimal equilibrium dynamic traffic assignments with scheduled delays*. Transportation Research 3C, 389-411

Wie, B.W., Tobin, R.L., Carey, M., (2002). *The existence, uniqueness and computation of an arc-based dynamic network user equilibrium formulation*. Transportation Research B 36, 897–918.

Wu, Chun-Hsin, Jan-Ming Ho, and D. T. Lee. (2004). *Travel-Time Prediction With Support Vector Regression*. IEEE Transactions On Intelligent Transportation Systems, Vol. 5, No. 4.

Wu, Yao-Jan, Guohui Zhang, and Yinhai Wang. *Volume Data Correction for Single-channel Advance Loop Detectors at Signalized Intersections*. Transportation Research Record, No.

2160, 128 – 139, 2010.

Young, S. (2008). *Bluetooth Traffic Monitoring Technology: Concept Of Operation & Deployment Guidelines*. Research Project Brochure, retrieved online at http://www.catt.umd.edu/documents/UMD-BT-Brochure_REV3.pdf

Zhang X, and H.M. Zhang. (2007) *Simultaneous departure time/route choices in queuing networks and a novel paradox*. Networks and Spatial Economics Published Online: 25 July 2007

Zhang, L. (2011). Behavioral Foundation of Route Choice and Traffic Assignment: Comparison of Principles of User Equilibrium Traffic Assignment under Different Behavioral Assumptions. *Travel Forecasting*. Vol. 1, pp. 1-10.

Ziliaskopoulos, A.K. (2000). *A Linear Programming Model for the Single Destination System Optimum Dynamic Traffic Assignment Problem*. Transportation Science 34(1): 37-49

Ziliaskopoulos, A.K. and S.T. Waller. (2000). *An Internet Based Geographic Information System that Integrates Data, Models, and Users for Transportation Applications*. Transportation Research Vol. 8C pp. 427-444, 2000.

# Appendix A: Geodatabase Design

Table A-1 The NHPN Feature Class

| Simple Feature Class | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | Geometry | | *Polyline* | | | |
| **NHPN** | | Contains M values | | *No* | | | |
| | | Contains Z values | | *No* | | | |
| Field name | Data Type | Allow nulls | Default value | Domain | Precision | Scale | Length |
| ObjectID | ObjectID | | | | | | |
| Shape | Geometry | Yes | | | | | |
| ShapeLength | Double | Yes | | | 0 | 0 | |
| IsEnabled | Short Integer | Yes | 1 | EnabledDomain | 0 | | |
| RECID | Double | No | | | 0 | 0 | |
| RECTYPE | String | No | | | | | 1 |
| VERSION | String | No | | | | | 7 |
| ORIGID | Double | No | | | 0 | 0 | |
| CTFIPS | Short Integer | No | | | 0 | | |
| SOURCE | String | No | | SourceCode | | | 1 |
| HPMS07 | String | No | | IsHPMS07 | | | 1 |
| LGURB | Short Integer | No | | | 0 | | |
| SMURB | Long Integer | No | | | 0 | | |
| SIGN1 | String | Yes | | SignRouteCode | | | 6 |
| SIGNT1 | String | Yes | | | | | 1 |
| SIGNN1 | String | Yes | | | | | 5 |
| SIGNQ1 | String | Yes | | QualCode | | | 1 |
| SIGN2 | String | Yes | | SignRouteCode | | | 6 |
| SIGNT2 | String | Yes | | | | | 1 |
| SIGNN2 | String | Yes | | | | | 5 |
| SIGNQ2 | String | Yes | | QualCode | | | 1 |
| SIGN3 | String | Yes | | SignRouteCode | | | 6 |
| SIGNT3 | String | Yes | | | | | 1 |
| SIGNN3 | String | Yes | | | | | 5 |
| SIGNQ3 | String | Yes | | QualCode | | | 1 |
| LNAME | String | No | | | | | 30 |
| MILES | Double | No | | | 0 | 0 | |
| KM | Double | No | | | 0 | 0 | |
| FCLASS | Short Integer | No | | ClassCode | 0 | | |
| RUCODE | Short Integer | No | | URCode | 0 | | |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| STATUS | Short Integer | No | | StatusCode | 0 | | |
| NHS | Short Integer | No | | NHSCode | 0 | | |
| STRAHNET | Short Integer | No | | STRAHNETCode | 0 | | |
| FAC_ID | String | Yes | | | | | 10 |
| CONN_ID | String | Yes | | | | | 10 |
| CONN_DES | String | Yes | | | | | 200 |
| CONN_MILES | Double | No | | | 0 | 0 | |
| LRSKEY | String | Yes | | | | | 15 |
| LRSSEQ | Short Integer | No | | | 0 | | |
| BEGMP | Double | No | | | 0 | 0 | |
| ENDMP | Double | No | | | 0 | 0 | |
| AADT | Long Integer | No | | | 0 | | |
| THRULANES | Short Integer | No | | | 0 | | |
| OWNERSHIP | Short Integer | No | | OwnershipCode | 0 | | |
| STFIPS | String | No | | | | | 2 |
| BTSVERSION | String | No | | | | | 2 |

Table A-2 CountyCentroidConnector Feature Class

| Simple Feature Class | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | Geometry | | *Polyline* | | | |
| **CountyCentoridConnector** | | Contains M values | | *No* | | | |
| | | Contains Z values | | *No* | | | |
| Field name | Data Type | Allow nulls | Default value | Domain | Precision | Scale | Length |
| ObjectID | ObjectID | | | | | | |
| Shape | Geometry | Yes | | | | | |
| ShapeLength | Double | Yes | | | 0 | 0 | |
| IsEnabled | Short Integer | Yes | 1 | EnabledDomain | 0 | | |
| EdgeID | Long Integer | No | | | 0 | | |
| FromCentroidID | Long Integer | No | | | | | |
| ToJunction | Long Integer | No | | | 0 | | |

Table A-3 MSACentroidConnector Feature Class

| Simple Feature Class | | |
|---|---|---|
| | Geometry | *Polyline* |
| **MSACentroidConnector** | Contains M values | *No* |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| | | Contains Z values | | *No* | | | |
| Field name | Data Type | Allow nulls | Default value | Domain | Precision | Scale | Length |
| ObjectID | ObjectID | | | | | | |
| Shape | Geometry | Yes | | | | | |
| ShapeLength | Double | Yes | | | 0 | 0 | |
| IsEnabled | Short Integer | Yes | 1 | EnabledDomain | 0 | | |
| EdgeID | Long Integer | No | | | 0 | | |

Table A-4 NetworkNode Feature Class

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Simple Feature Class | | | | | | | |
| | | Geometry | | *Point* | | | |
| **NetworkNode** | | Contains M values | | *No* | | | |
| | | Contains Z values | | *No* | | | |
| Field name | Data Type | Allow nulls | Default value | Domain | Precision | Scale | Length |
| ObjectID | ObjectID | | | | | | |
| Shape | Geometry | Yes | | | | | |
| JunctionID | Long Integer | No | | | 0 | | |
| XCoord | Short Integer | Yes | | | 7 | 2 | |
| YCoord | Short Integer | Yes | | | 7 | 2 | |
| IsEnabled | Short Integer | Yes | 1 | EnabledDomain | 0 | | |

Table A-5 County Feature Class

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Simple Feature Class | | | | | | | |
| | | Geometry | | *Polygon* | | | |
| **County** | | Contains M values | | *No* | | | |
| | | Contains Z values | | *No* | | | |
| Field name | Data Type | Allow nulls | Default value | Domain | Precision | Scale | Length |
| ObjectID | ObjectID | | | | | | |
| Shape | Geometry | Yes | | | | | |
| Shape_Length | Double | Yes | | | 0 | 0 | |
| Shape_Area | Double | Yes | | | 0 | 0 | |
| STFIPS | String | No | | | | | 2 |
| CTFIPS | String | No | | | | | 5 |
| STATE | String | No | | | | | 66 |
| COUNTY | String | No | | | | | 66 |
| VERSION | String | No | | | | | 2 |

Table A-6 CountyCentroidConnector Feature Class

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Simple Feature Class | | | | | | | |
| | | Geometry | | *Point* | | | |
| **CountyCentroid** | | Contains M values | | *No* | | | |
| | | Contains Z values | | *No* | | | |
| Field name | Data Type | Allow nulls | Default value | Domain | Precision | Scale | Length |
| ObjectID | ObjectID | | | | | | |
| Shape | Geometry | Yes | | | | | |
| CountyCentroidID | Long Integer | No | | | 0 | | |
| CTFIPS | String | No | | | | | 20 |
| Xcoord | Short Integer | Yes | | | 7 | 2 | |
| Ycoord | Short Integer | Yes | | | 7 | 2 | |
| AncillaryRole | Short Integer | No | 0 | NetwrokRole | 0 | | |
| IsEnabled | Short Integer | Yes | 1 | EnabledDomain | 0 | | |

Table A-7 Loop Table

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Table | **Loop** | | | | | | |
| Field name | Data type | Allow nulls | Default value | Domain | Precision | Scale | Length |
| SensorID | Short Integer | No | | | 0 | | |
| RECID | Double | No | | | 0 | 0 | |
| VehicleClass | String | Yes | | | | | |
| Direction | Short Integer | No | | DirectionCode | | | |
| CabinetID | String | Yes | | | | | 10 |
| HHMMSS | Long Integer | Yes | | | 0 | | |
| YYYYMMDD | Long Integer | Yes | | | 0 | | |
| Volume | Short Integer | Yes | | | 0 | | |
| Occupancy | Short Integer | Yes | | | 0 | | |
| route_id | String | Yes | | | | | 254 |
| Location | String | Yes | | | | | 20 |
| Milepost | Double | Yes | | | 8 | 2 | |
| Lanes | Short Integer | Yes | | | 0 | | |
| Speed | Short Integer | Yes | | | 0 | | |
| LoopType | String | Yes | | | | | 7 |
| Periods | Short Integer | Yes | | | 0 | | |
| Flag | Short Integer | Yes | | | 0 | | |
| Latitude | Double | Yes | | | 8 | 4 | |
| Longitude | Double | Yes | | | 8 | 4 | |

Table A-8 NetworkRole Domain

| | |
|---|---|
| Coded value domain | **NetwrokRole** |
| Description | *Centroid type* |
| Field Type | *Short Integer* |
| Split policy | *Default value* |
| Merge policy | *Default value* |
| Code | Description |
| 0 | Neither source nor sink |
| 1 | Source |
| 2 | Sink |

Table A-9 EnabledDomain

| | |
|---|---|
| Coded value domain | **EnabledDomain** |
| Description | *Boolean logic value* |
| Field Type | *Short Integer* |
| Split policy | *Default value* |
| Merge policy | *Default value* |
| code | Description |
| 0 | Disabled |
| 1 | Enabled |

Table A-10 DirectionCode

| | |
|---|---|
| Coded value domain | **DirectionCode** |
| Description | *Driving Direction* |
| Field Type | *Short Integer* |
| Split policy | *Default value* |
| Merge policy | *Default value* |
| code | Description |
| 0 | FromJunction to ToJuction |
| 1 | ToJunction to FromJuction |

Table A-11SignRouteCode Domain

| | |
|---|---|
| Coded value domain | **SignRouteCode** |
| Description | *The type of the sign route* |
| Field Type | *String* |
| Split policy | *Default value* |
| Merge policy | *Default value* |
| code | Description |
| C | County Route |
| I | Interstate |

| | |
|---|---|
| M | Municipal |
| O | Off-Interstate Business Marker |
| P | Parkway or Forest Route Marker |
| S | State Route |
| T | Township |
| U | US route |
| N | none of above |
| | blank means not signed or not applicable |

Table A-12 QualCode Domain

| | |
|---|---|
| Coded value domain | **QualCode** |
| Description | *Route Qualifier Type* |
| Field Type | *String* |
| Split policy | *Default value* |
| Merge policy | *Default value* |
| code | Description |
| A | Alternative Route |
| B | Business Route |
| D | Temporary (Detour) |
| F | Proposed (Future) |
| L | Loop |
| P | Bypass |
| S | Spur |
| T | Truck Route |
| N | none of above |
| | blank means not signed or not applicable |

Table A-13 SourceCode Domain

| | |
|---|---|
| Coded value domain | **SourceCode** |
| Description | *Source of Arc* |
| Field Type | *String* |
| Split policy | *Default value* |
| Merge policy | *Default value* |
| code | Description |
| T | Tiger File |
| D | Digitized |
| S | State |

Table A-14 URCCode Domain

| | |
|---|---|
| Coded value domain | **URCode** |
| Description | *Urban type* |
| Field Type | *Short Integer* |
| Split policy | *Default value* |
| Merge policy | *Default value* |
| code | Description |
| 1 | Rural Area |
| 2 | Small Urban (1990 pop 5,000 -49,999) |
| 3 | Large Urban (1990 >= 50,000) |

Table A-15 StatusCode Domain

| | |
|---|---|
| Coded value domain | **StatusCode** |
| Description | *Describes availability of the arc* |
| Field Type | *Short Integer* |
| Split policy | *Default value* |
| Merge policy | *Default value* |
| code | Description |
| 0 | Proposed/Under construction |
| 1 | Open to traffic |

Table A-16 STRAHNETCode Domain

| | |
|---|---|
| Coded value domain | **STRAHNETCode** |
| Description | *Special subnetwork* |
| Field Type | *Short Integer* |
| Split policy | *Default value* |
| Merge policy | *Default value* |
| code | Description |
| 0 | Not on STRAHNET |
| 1 | Non-Interstate STRAHNET |
| 2 | STRAHNET Priority 1 Connector |
| 3 | STRAHNET Priority 2 Connector |
| 4 | STRAHNET Priority 3 Connector |
| 5 | Temporary STRAHNET Route |

Table A-17 OwnershipCode Domain

| | |
|---|---|
| Coded value domain | **OwnershipCode** |
| Description | *highway owner type* |
| Field Type | *Short Integer* |
| Split policy | *Default value* |

| | |
|---|---|
| Merge policy | *Default value* |
| code | Description |
| 1 | State Highway Agency |
| 2 | County Highway Agency |
| 3 | Town or Township Highway Agency |
| 4 | Municipal Highway Agency |
| 5 | Other State Agency |
| 6 | Other Local Agency |
| 7 | Federal Agency |
| 8 | Other |

Table A-18 NHSCode Domain

| | |
|---|---|
| Coded value domain | **NHSCode** |
| Description | *NHS type* |
| Field Type | *Short Integer* |
| Split policy | *Default value* |
| Merge policy | *Default value* |
| code | Description |
| 0 | Not on NHS |
| 1 | Interstate |
| 3 | Non-Interstate STRAHNET (Strategic Highway Network) |
| 4 | STRAHNET Connector |
| 7 | Other NHS |
| 8 | Approved Intermodal Connector |

Table A-19 IsHPMS07 Domain

| | |
|---|---|
| Coded value domain | **IsHPMS07** |
| Description | *updated with 2007 HPMS Data* |
| Field Type | *String* |
| Split policy | *Default value* |
| Merge policy | *Default value* |
| code | Description |
| Y | Record was updated with 2007 HPMS Data |
| N | Record was not updated with 2007 HPMS Data |

Table A-20 ClassTable Domain

| | |
|---|---|
| Coded value domain | **ClassCode** |
| Description | *Functional class* |
| Field Type | *Short Integer* |
| Split policy | *Default value* |
| Merge policy | *Default value* |
| code | Description |
| 1 | Rural Principal Arterial - Interstate |
| 2 | Rural Principal Arterial - Other |
| 6 | Rural Minor Arterial |
| 7 | Rural Major Collector |
| 8 | Rural Minor Collector |
| 9 | Rural Local |
| 11 | Urban Principal Arterial - Interstate |
| 12 | Urban Principal Arterial - Other Freeways and Expressways |
| 14 | Urban Principal Arterial - Other |
| 16 | Urban Minor Arterial |
| 17 | Urban Collector |
| 19 | Urban Local |

Table A-21 MonthCode Domain

| | |
|---|---|
| Coded value domain | **MonthCode** |
| Description | *month* |
| Field Type | *Short Integer* |
| Split policy | *Default value* |
| Merge policy | *Default value* |
| code | Description |
| 1 | January |
| 2 | February |
| 3 | March |
| 4 | April |
| 5 | May |
| 6 | June |
| 7 | July |
| 8 | August |
| 9 | September |
| 10 | October |
| 11 | November |
| 12 | December |

Table A-22 DayCode Domain

| | |
|---|---|
| Coded value domain | **DayCode** |
| Description | *Seven week days* |
| Field Type | *Short Integer* |
| Split policy | *Default value* |
| Merge policy | *Default value* |
| code | Description |
| 1 | Monday |
| 2 | Tuesday |
| 3 | Wednesday |
| 4 | Thursday |
| 5 | Friday |
| 6 | Saturday |
| 7 | Sunday |

Table A-23 HourCode Domain

| | |
|---|---|
| Coded value domain | **HourCode** |
| Description | *24 hours* |
| Field Type | *Short Integer* |
| Split policy | *Default value* |
| Merge policy | *Default value* |
| code | Description |
| 1 | 0:00 - 1:00 |
| 2 | 1:00 - 2:00 |
| 3 | 2:00 - 3:00 |
| 4 | 3:00 - 4:00 |
| 5 | 4:00 - 5:00 |
| 6 | 5:00 - 6:00 |
| 7 | 6:00 - 7:00 |
| 8 | 7:00 - 8:00 |
| 9 | 8:00 - 9:00 |
| 10 | 9:00 - 10:00 |
| 11 | 10:00 - 11:00 |
| 12 | 11:00 - 12:00 |
| 13 | 12:00 - 13:00 |
| 14 | 13:00 - 14:00 |
| 15 | 14:00 - 15:00 |
| 16 | 15:00 - 16:00 |
| 17 | 16:00 - 17:00 |
| 18 | 17:00 - 18:00 |

| | |
|---|---|
| 19 | 18:00 - 19:00 |
| 20 | 19:00 - 20:00 |
| 21 | 20:00 - 21:00 |
| 22 | 21:00 - 22:00 |
| 23 | 22:00 - 23:00 |
| 24 | 23:00 - 24:00 |

# Appendix B: Data Dictionary of FAF[3]

| Attribute | Domain Type | Description |
|---|---|---|
| ID | Integer | Unique identifier to link with FAF network arc |
| Version | Character | Used for maintaining consistency across data files containing alternate releases of the FAF. |
| AADT07 | Integer | HPMS annual average daily traffic for year 2007, derived from HPMS 2008 database. Volume/day/route |
| FAF07 | Integer | FAF 3.1 long distance truck volume estimated based on the FAF 3.1 Origin-Destination truck tonnage and includes empty trucks. Volume/day/route |
| NONFAF07 | Integer | Local truck traffic that is not part of FAF 3.11 O-D database. Volume/day/route |
| AADT40 | Integer | Year 2040 forecast Annual Average Traffic Volume estimated using the HPMS 20 years growth factors and projected to future using linear growth. Volume/day/route |
| AADTT40 | Integer | Forecast Annual Average Truck Volume estimated using the HPMS 20 years growth factors and projected to future using linear growth. Volume/day/route |
| FAF40 | Integer | Year 2040 FAF 3.1 long distance truck volume estimated based on the forecasted FAF 3.1 Origin-Destination truck tonnage and includes empty trucks. Volume/day/route |
| NONFAF40 | Integer | Year 2040 Local truck traffic that is not part of FAF 3.11 O-D database. Volume/day/route |
| CAP07 | Integer | Link specific peak capacity estimated using the procedures outlined in HCM 2000 and the arc geometry provided in 2008 HPMS database. Volume/hour/route |
| SF07 | Integer | Estimated service flow using the procedures outlined in HCM 2000 and arc geometry, FAF truck, non-FAF truck and passenger volume. Volume/hour/route |
| VCR07 | Real | 2007 estimated volume to capacity ratio, estimated by dividing SF07 with CAP07. Unit less |
| SPEED07 | Real | 2007 estimated peak period link speed, estimated using the procedures outlined in HCM 2000 and the arc geometry provided in 2008 HPMS database. miles/hour |
| DELAY07 | Real | 2007 estimated peak period link delay, estimated using the procedures outlined in HCM 2000 and the arc geometry provided in 2008 HPMS database. In hours |
| CAP40 | Integer | Link specific peak capacity estimated using the procedures outlined in HCM 2000. Volume/hour/route |
| VCR40 | Real | 2040 estimated volume to capacity ratio, estimated by dividing SF40 with CAP40. Unit less |
| SPEED40 | Real | 2040 estimated peak period link speed, estimated using the procedures outlined in HCM 2000. Miles/hour |
| DELAY40 | Real | 2040 estimated peak period link delay, estimated using the procedures outlined in HCM 2000. In hours |

## Appendix C: Annual Congestion Cost per Auto Commuter in Dollar (2008 Data)

| Urban Area | MSA | County | | | | |
|---|---|---|---|---|---|---|
| Seattle WA | 994 | King County, WA | Pierce County, WA | Snohomish County, WA | | |
| Spokane WA | 361 | Spokane County, WA | | | | |
| Portland OR-WA | 716 | Multnomah County, OR | Washington County, OR | Clackamas County, OR | Yamhill County, OR | Columbia County, OR |
| Salem OR | 447 | Marion County, OR | Polk County, OR | | | |
| Eugene OR | 212 | Lane County, OR | | | | |
| Sacramento CA | 491 | Sacramento County, CA | Placer County, CA | Yolo County, CA | El Dorado County, CA | |
| Stockton CA | 185 | San Joaquin County, CA | | | | |
| San Francisco-Oakland CA | 1020 | Alameda County, CA | Contra Costa County, CA | San Francisco County, CA | San Mateo County, CA | Marin County, CA |
| San Jose CA | 740 | Santa Clara County, CA | San Benito County, CA | | | |
| Fresno CA | 235 | Fresno County, CA | | | | |
| Bakersfield CA | 204 | Kern County, CA | | | | |
| Lancaster-Palmdale CA | 293 | | | | | |
| Oxnard-Ventura CA | 356 | Ventura County, CA | | | | |
| Los Angeles-Long Beach-Santa Ana CA | 1258 | Los Angeles County, CA | Orange County, CA | | | |
| Riverside-San Bernardino CA | 657 | Riverside County, CA | San Bernardino County, CA | | | |
| Indio-Cathedral City-Palm Springs CA | 277 | | | | | |
| San Diego CA | 856 | San Diego County, CA | | | | |
| Others/Urban | 180 | | | | | |
| Others/Rural | 90 | | | | | |

## Appendix D: Rural Counties

| |
|---|
| Stevens County, Washington |
| Okanogan County, Washington |
| Jefferson County, Washington |
| Pacific County, Washington |
| Klickitat County, Washington |
| Adams County, Washington |
| San Juan County, Washington |
| Pend Oreille County, Washington |
| Lincoln County, Washington |
| Ferry County, Washington |
| Columbia County, Washington |
| Wahkiakum County, Washington |
| Garfield County, Washington |
| Lincoln County, Oregon |
| Tillamook County, Oregon |
| Jefferson County, Oregon |
| Baker County, Oregon |
| Lake County, Oregon |
| Grant County, Oregon |
| Harney County, Oregon |
| Wallowa County, Oregon |
| Gilliam County, Oregon |
| Sherman County, Oregon |
| Wheeler County, Oregon |
| Calaveras County, California |
| Siskiyou County, California |
| Amador County, California |
| Glenn County, California |
| Colusa County, California |
| Plumas County, California |
| Mariposa County, California |
| Mono County, California |
| Trinity County, California |
| Modoc County, California |
| Sierra County, California |
| Alpine County, California |