Title

Optimal Use of Changeable Message Signs for Displaying Travel Times

Authors

Ban, Xuegang (Jeff)
Li, Yuwei
Margulici, Jean-Davi

Publication Date

2009-08-10

CALIFORNIA PATH PROGRAM
INSTITUTE OF TRANSPORTATION STUDIES
UNIVERSITY OF CALIFORNIA, BERKELEY

Optimal Use of Changeable Message Signs for Displaying Travel Times

Xuegang (Jeff) Ban
Yuwei Li
Jean-Davi Margulici

California PATH Research Report
UCB-ITS-PRR-2009-33

This work was performed as part of the California PATH Program of the University of California, in cooperation with the State of California Business, Transportation, and Housing Agency, Department of Transportation, and the United States Department of Transportation, Federal Highway Administration.

Final Report for Task Order 6303

August 2009

ISSN 1055-1425

Optimal Use of Changeable Message Signs for Displaying Travel Times

Xuegang (Jeff) Ban
Yuwei Li
Jean-Davi Margulici

Institute of Transportation Studies
University of California at Berkeley

Final Report for PATH Task Order 6303

August 10, 2009

ACKNOWLEDGEMENTS

This work was performed by the California PATH Program of the University of California at Berkeley, in cooperation with the State of California Business Transportation and Housing Agency, Department of Transportation (Caltrans), Division of Research and Innovation Division (DRI) (Task Order 6303, Interagency Agreement #65A0208). The contents of this report reflect the views of the authors who are responsible for the facts and the accuracy of the results presented herein. The contents do not necessarily reflect the official views or policies of the State of California. This report does not constitute a standard, specification or regulation.

Xuegang (Jeff) Ban led the research efforts in the evaluation of the travel time estimation methods and the development and application of the network model. Yuwei Li led the effort in the survey of the Bay Area commuters and the analysis of the results. They were assisted by JD Margulici of CCIT, and Nina Harvey, graduate student in transportation engineering. Alexander Skabardonis served as the Principal Investigator for the study.

The authors especially acknowledge Asfand Siddiqui of Caltrans DRI and David Lively of Division of Traffic Operations, Caltrans Headquarters for their cooperation and support throughout the project. We also thank Dr. Michael Mauch of DKS Associates for his assistance with the development of the San Francisco Bay Area network model.

Optimal Use of Changeable Message Signs for Displaying Travel Times

Xuegang (Jeff) Ban, Yuwei Li, Jean-Davi Margulici

February 2009

ABSTRACT

The report describes the research performed and the findings of the PATH Task Order 6303 "Optimal Use of Changeable Message Signs for Displaying Travel Times". First, travel time estimation methods from various sources are evaluated against probe vehicle data. Next, a survey of San Francisco Bay Area commuters was undertaken regarding the perception and preference for displaying travel times on CMS. Finally, a model that evaluates the CMS impacts on network traffic and determines optimal CMS configurations was developed and tested. Deployment guidelines for CMS are presented based on these results.

Key Words: *Travel Time, Variable Message Signs, Freeways, Travel Behavior*

EXECUTIVE SUMMARY

This goal of the research described in this report was to develop guidelines for displaying travel times on Changeable Message Signs (CMS) with regard to a few key issues: what method to use in estimating travel time, whether to display travel times if no alternative routes exist, how frequently the display should be updated, and where the CMS should be activated to display travel times.

In the first part of this research, we evaluated the performance of existing methods that estimate freeway route travel times using data from loop detectors. The accuracy of travel time estimates are evaluated with probe vehicle data (the "ground truth") obtained from the *FasTrak* system in the San Francisco Bay Area. Next, we conducted a survey of Bay Area commuters regarding their perception and preference for displaying travel times on CMS. In the third part, we studied the problem of optimal CMS configuration for displaying travel times. Our approach is based on a Stochastic Network – Stochastic User Equilibrium (SN-SUE) model to capture how commuters make route choice decisions which considers both travel time variability and travelers' perceptions errors. To solve the optimal CMS configuration problem, we developed a heuristic method based on simulated annealing (SA). The model and solution method are tested on a hypothetical network and a real world network in the San Francisco Bay Area.

The study revealed that: 1) compared to ground truth travel times from probe vehicles, travel times estimated from loop detectors for different lanes may be significantly different based on the time of day; 2) Perception of accuracy of estimated travel time greatly affects the perceived usefulness of travel time display. However, even when the perceived accuracy is within 10-15 minutes, the vast majority (over 80%) of the commuters who hold the accuracy perception still regards the travel time display on CMS as useful; 3) a large threshold for the difference between expected route travel times needs to be reached before commuters switch from their usual routes. The effectiveness of CMS alone in persuading drivers to divert is likely to be small; and 4) the network-wide performance of the system depends on the driver's perceptions and travel time variability. Based on results from the three parts of this research, we recommend the following guidelines for displaying travel times on CMS:

1. Travel time estimation methods:
   During off-peak, the differences using different estimation algorithms are not significant, and the instantaneous travel time can be adopted for its simplicity.
   During peak hours, if the route travel time is relatively short and the transition from free-flow to maximum congestion is slow, the instantaneous travel time still can be adopted for its simplicity.
   During peak hours, when the route travel time is long (e.g. from Emeryville to SFO), linear regression travel times (generated using both real-time data and historical data) should be considered.
   Using speed data from different lanes makes significant difference in the accuracy of the travel time estimates. If patterns similar to those shown in this study are found, lane-by-lane loop data can be used to improve travel time estimation.

When relying on loop detectors alone, loop spacing should not be too large. Use multiple data source when loop coverage is poor.

2. Display travel times in the presence of alternative routes:
   Benefits besides route choice, such as being able to plan ahead and having peace of mind, are as important as route choice to users when they consider the usefulness of travel time display on CMS. When determining whether travel times should be displayed on a CMS, all these benefits needs to be considered. Even if there is no route choice, it is beneficial to have travel time display on CMS if there is great variability and uncertainty in traffic condition on a freeway segment.

3. Frequency of travel time displays:
   Estimated travel time should be displayed as an exact number of minutes, not as a range. Information should be updated every 2 minutes. More frequent updates may confuse drivers and reduce their confidence in the accuracy travel time display.

4. CMS location for displaying travel times:
   Exactly which CMS should be activated to display travel times (or whether new CMS needs to be installed for this purpose) and which destinations/routes to display on a CMS need to be determined based on the following factors: roadway geometry, traffic conditions especially travel time variability, travelers' perceptions over the actual travel times, and CMS installation/activation cost.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1
INTRODUCTION

Changeable Message Signs (CMS), also called Dynamic Message Signs (DMS) or Variable Message Signs (VMS) or Electronic Message Signs (EMS), are commonly-used ways today for disseminating real-time traffic information to the driving public. Expected benefits of CMS are reduced delays and risks caused by incidents, construction, or other recurrent and on-recurring congestion. The assumption is that drivers will make informed decisions based on information provided by the signs (such as diverting to an alternative route in case incidents happened on the freeway). CMS can also help alleviate drivers' stress and better manage their time (e.g. one can call in advance if he/she is going to be late for work based on the information from the signs). In the US, CMS have been widely deployed as shown in Figure 1, which depicts an increasing rate of CMS deployment.

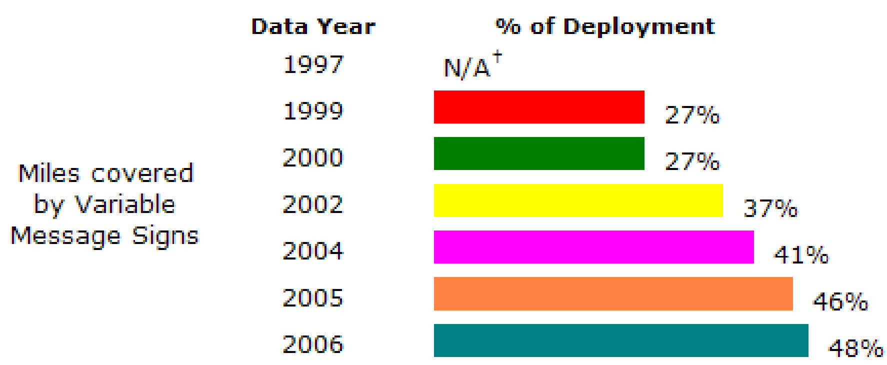

Figure 1: CMS Deployments by Miles Covered (Source: RITA (2006))

Travel time and delay messages are considered to be valuable information and an efficient use of CMS in the absence of adverse traffic incidents or events. In this manner, travel times (or delays) not only give the estimated time between a CMS and a point downstream, the presence of travel time information also gives the implicit message that there are no adverse conditions affecting traffic (see PBSJ(2004)). In 2004, a memo[1] released by the Federal Highway Administrations (FHWA) recommends local agencies display travel times on CMS. The memo states that "[travel times] have proven successful in regions or corridors that experience periods of recurring congestion - congestion generally resulting from traffic demand exceeding available capacity and not caused by any specific event such as a traffic incident, road construction or a lane closure." It further recommends that "no new Changeable Message Signs (CMS) should be installed in a major metropolitan area or along a heavily traveled route unless the operating agency and the jurisdiction have the capability to display travel time messages."

[1] http://ops.fhwa.dot.gov/travelinfo/resources/cms_rept/travtime.htm

In California, expanding the scope and coverage of roadway travel times is a top-priority of the *GoCalifornia* initiative. Supporting statements for accurate travel time estimates and traveler information have come all the way from the Governor's office.

The purpose of this research is to develop guidelines for displaying travel times on Changeable Message Signs. This rest of report is organized as follows. Chapter 2 evaluates the performance of travel time estimation methods for real-time traffic applications. Chapter 3 analyzes a survey of bay area commuters regarding the perception and preference for displaying travel times on CMS. Chapter 4 presents a model that evaluates the CMS impacts on network traffic and determines optimal CMS configurations. Chapter 5 offers deployment guidelines for displaying travel times on CMS and concluding remarks.

CHAPTER 2
PERFORMANCE EVALUATION OF TRAVEL TIME ESTIMATION METHODS

2.1 Background

Travel time for selected itineraries constitutes one of the most relevant roadway traffic metrics. This is because (1) travel time is a crucial measure of traffic conditions and system performance; (2) travel time represents information that is easy for the driving public to understand and process; and (3) travel time information can arguably enable travelers to make educated choices about their itinerary, departure time or even transportation mode, with the result of bringing about a form of "system self-management." Numerous studies reveal that commuters appreciate and value travel time information, which reduces their uncertainty and their stress (Peng et al., 2004; Lindveld et al., 2000; Khattak et al., 1994). Many researchers have contributed to developing algorithms for accurate and reliable travel time estimation (Oda, 1990; Smith and Demetsky, 1997; Huisken and Maarseveen, 2000; Rice and Zwet, 2001; Hartley, 2003; Hinsbergen et al, 2007, to name just a few).

Measuring the quality of travel time estimates is therefore important because it helps to understand the performance of travel time estimation and point to needed improvements in traffic data collection. In the past, despite the extensive research efforts on developing travel time estimation methods, studies devoted specifically on evaluating the performances of these methods were rare. Although many studies did provide certain discussions on performances of the proposed methods, the discussions were limited in the sense that (1) "ground truth" travel times (e.g. those from probe vehicles) were not widely available due to technical or resource limitations; (2) in case of loop detectors, most studies used speed data averaged across multiple lanes without looking at lane-by-lane travel time variations; and (3) evaluations were mostly conducted on "baseline" scenarios, i.e., detector locations were assumed to be given and all detector data were used to estimate travel times.

Lindveld et al. (2000) is one of the few studies focusing on evaluating performances of several travel time estimation methods using loop detector data. They found that up to moderate congestion levels, travel time estimators could produce reasonably accurate results (10-15%); however, for heavy congestion, the results may degrade significantly. The ground-truth travel times in Lindveld et al. (2000) were collected via license plate readers, floating car runs, and toll ticket collection. However, the number of observed data points using floating cars is not sufficient; travel times from toll ticket collection have problems as well (Lindveld et al., 2000, pp. 46). Also, there is no discussion about removing outliers within the observed data. In addition, lane-by-lane travel times were not computed and studied in Lindveld et al. (2000). Zhang et al. (1999) studied travel time estimation methods based on single loop detector data. Floating car runs were conducted to gather the ground truth travel times. As pointed out in Kwon et al. (2006), however, limited floating car runs may be biased. Kwon et al (2006) and Fujito et al. (2006) studied the relationship between detector spacing and travel time estimation quality. However, they used travel times computed from the "baseline" detector spacing as the ground truth travel times. As shown in this study, this may be very different from actually experienced travel times by individual drivers. Note also that evaluating travel time estimation quality is essentially different from studying the variability of travel times which has been widely studied in the literature (Noland and Polak, 2002; Van Lint et al, 2004; Chen et al., 2003). The former focuses on the differences of estimated and actual travel times, while the latter is supposed to study the features (such as distribution) of the actual travel times. For the same reason, our study also differs from the travel time reliability studies that have gained much attention recently in the transportation research communities (Al-Deek and Emam, 2006; Chen et al., 2003; Chen et al., 1999; Liu et al., 2007).

In this section, we evaluate a set of "benchmark" methods, i.e. algorithms along with specific speed data from dual loop detectors, that give point estimate of route travel times. The evaluated algorithms include three that are popularly used in real time traffic applications such as displaying travel times on Changeable Message Signs (CMS): the instantaneous, dynamic and linear-regression travel times (Rice and Zwet, 2001). The data source to be evaluated is double loop detectors, and both average and lane-by-lane speeds are used to compute the average and lane-specific travel times. Travel time estimates are compared with probe vehicle travel times, obtained from *FasTrak* in the San Francisco Bay Area. Due to the large amount of data samples, *FasTrak* travel times are expected to provide more accurate representation of the ground truth travel times than limited floating car runs, although the data also contain outliers that need to be filtered.

We first propose a local Median Absolute Deviation (MAD) method with variable window length to remove outliers in *FasTrak* data. The method captures traffic characteristics properly and is thus appropriate for processing probe-vehicle data. The processed "ground truth" travel time has a clear time-dependent trend. The dispersion of travel time is larger in off-peak than in peak periods, which coincides with the theoretical explanation by Daganzo (1997, pp. 142).

Two measures of the point estimate accuracy are defined. One is based on the absolute difference between the loop-based travel time estimate and the estimate of medium travel time via FasTrak data. The other is based on whether the loop-based travel time estimate falls into certain range of travel time estimated from the dispersion estimated of *FasTrak* data.

We next conduct performance evaluation for a particular route in the San Francisco Bay Area. The evaluation is done for different traffic conditions represented by time-of-day periods and various detector spacing scenarios. The results show that 1) the accuracy of travel time estimates based on loop detector data is better in off-peak than in peak periods ; 2) lane-by-lane loop detector speed data may be utilized to improve travel time estimation; and 3) larger detector spacing negatively affect accuracy of travel time estimation.

The extent to which findings in this section apply to other sites may vary. We thus provide discussions at the end on how these findings can be further verified and applied to achieve improved travel time estimation quality.

2.2 Ground Truth Travel Times from Probe Vehicles

The travel time $T_r(t)$ for an arbitrary vehicle to travel from a CMS location to certain destination via route *r* at time *t*, is a random variable. The population of concern is the set of different people driving different cars who potentially may take route *r* at time *t* (i.e. starting from the CMS location at time *t*). In Noland and Polak (2002), this randomness is termed as "vehicle-to-vehicle travel time variation." Note that sometime it may make sense to impose further restrictions, such as "not using carpool lane", "with *FasTrak* onboard", and "not stopping over at a point along the route". Ideally, there should be sufficiently many drivers who take route *r* at time *t*, which constitute the samples of $T_r(t)$. However, if steps of time and distance are arbitrarily fine, at any time *t* there is only at most one vehicle passing through the exact CMS location. Thus strictly speaking, we can have at most one sample for $T_r(t)$ in reality.

The term "ground-truth" in this study refers to travel time values obtained via probe vehicles, each generating one observation of the random variable $T_r(t)$. A "ground-truth' travel time value may be the best-available estimate of the medium of $T_r(t)$ if there is no other relevant information. On the other hand, if observations for travel time $T_r(t+dt)$ are available, the "ground-truth' travel time value may not be the best-available estimate of the medium of $T_r(t)$.

In order to obtain estimates of not only the medium but dispersion of the travel time variable, we may define $T_r(t)$ as the (random) travel time for a vehicle to travel route *r*, starting from the CMS location at some time $t' \in [t-\Delta/2, t+\Delta/2]$. This is feasible when we are only interested in (or limited to) discrete points on the time scale. Under this alternative definition, we may have multiple samples for $T_r(t)$. However, now the difference of the observed travel times may be due not only to driver/vehicle differences but largely to traffic condition change from time $t-\Delta/2$ to time $t+\Delta/2$ if Δ is too large.

The "ground truth" travel times in this section are obtained from probe vehicles, particularly the *FasTrak* data in the San Francisco Bay Area. *FasTrak* is used statewide in California to automatically collect road and bridge tolls. *FasTrak* readers are currently installed at each toll booth, as well as along the road side every 5 to 10 miles. It has a large market penetration in the Bay Area: nearly 50% of drivers use *FasTrak* to pay their bridge tolls in 2007 (http://goldengate.org/board/2007/Min-Sum/fa070426min.php). *Fastrak* data contain individual vehicle travel times between two consecutive readers.

The "ground-truth' travel time values obtained from *FasTrak* data may be interpreted as observations of $T_r(t)$ under the alternative definition (i.e. for the period of $[t-\Delta/2, t+\Delta/2]$), with Δ being sufficiently small. In this case, $T_r(t)$ may contain multiple observations but the differences between observed values are largely due to driver/vehicle differences. However, because of the population of concern, certain *Fastrak* data points should be excluded: some vehicles may have used the carpool lane, or the itinerary between the origin and the destination is not the intended one (i.e. route *r*). Thus the raw *FasTrak* data need to be processed to remove outliers.

In addition, if Δ is too small, there will be too few observations in $[t-\Delta/2, t+\Delta/2]$ to characterize the random variable $T_r(t)$. Thus choosing appropriate Δ is important in using the *FasTrak* data to estimate the medium and dispersion of travel times. This issue will be further discussed below.

2.2.1 A Local MAD Method for Probe-Vehicle Data Processing

Figure 2 depicts the raw *FasTrak* travel times for a route from the City of Albany to the Carquinez Bridge along I-80 EB. The time-dependent pattern of the travel time is obvious in Figure 2, but the raw data contain a significant amount of outliers. Outliers include those vehicles that took excessively long time to travel the route, possibly because they left and re-entered the freeway at some intermediate points. Outliers may also come from detection errors that result in those "negative" travel times in Figure 2. Vehicles that used the HOV lane during PM peak hours are also treated as outliers since their travel times are much lower than those using general-purpose lanes.

To remove outliers, we applied the Median Absolute Deviation (MAD) method (Hoaglin et al., 1983). MAD is a statistical measure to capture the variation of a given set of data points. Assume $\{x_i, i=1,\cdots,N\}$ is the set of data points. Then MAD can be defined as

$$MAD = median(|x_i - \tilde{x}|). \tag{2.1}$$

Here $\tilde{x}$ is the median value of $\{x_i, i=1,\cdots,N\}$. To detect whether x_i is an outlier, a *z*-score needs to be computed for each data point:

$$z_i = \frac{|x_i - \tilde{x}|}{MAD} \tag{2.2}$$

Then if $z_i \geq \bar{z}$ for a given threshold $\bar{z}$, x_i can be regarded as an outlier. Here we use $\bar{z} = 4.5$.

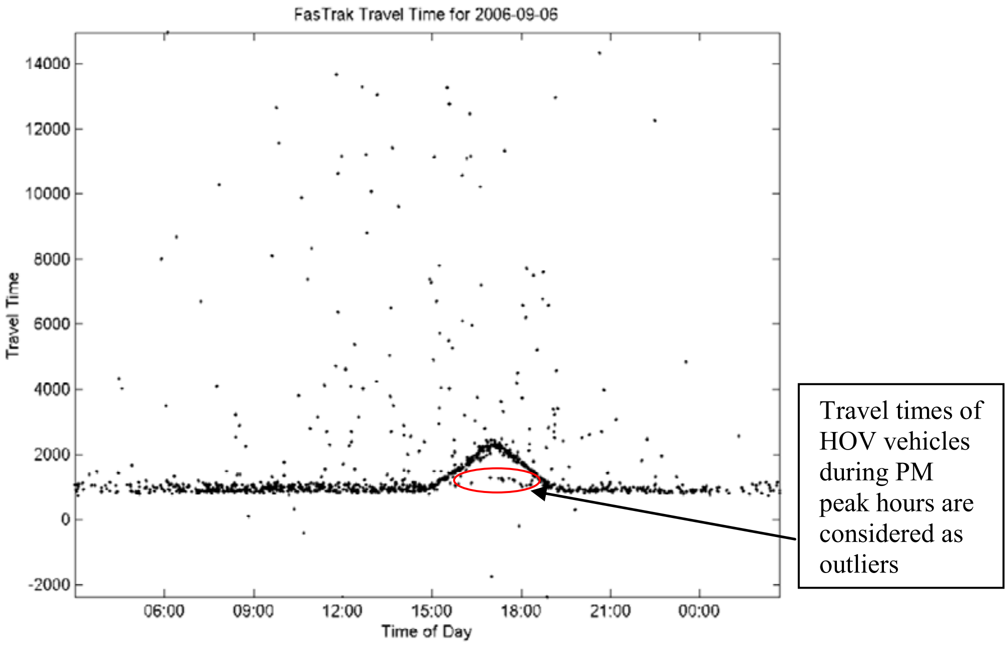

Figure 2: Raw FasTrak data

Since the raw *FasTrak* data in Figure 2 have a clear time-dependent pattern, we apply the MAD method locally to data points in a time window (band) with a proper "bandwidth." We call this method the *Local MAD Method (LMM)*. Choosing the bandwidth, however, is not trivial. First, to adequately capture the time-dependent trend, it is natural to use a small bandwidth; but this may result in only a few data points within a band for which the MAD method may not be properly applied. On the other hand, a large bandwidth can certainly make the MAD method statistically meaningful, but the time-dependent pattern may be smoothed out.

To account for the above two issues, a method of variable bandwidth is adopted in this study. We first set the default bandwidth as h_0 minutes, which is expansible so that each band contains at least *M* data points. This is to make the bandwidth as small as possible, while still keeping the processing statistically meaningful. Such a bandwidth will be constructed for each of the data points in the set $\{x_i, i = 1, \cdots, N\}$, as illustrated in the *LMM* algorithm below.

LMM Algorithm

Step 1. Initialization. Set *i*=1, h_0=10 minutes, *M*=25, Δ*h*=2 minute, and $\bar{z}$=4.5.

Step 2 Major Iteration.

Step 2.1 Determine the Bandwidth. Assume the timestamp of the current data point x_i is t_i. Set *h*=h_0. Assume the number of data points in the time window $[t_i - h/2, t_i + h/2]$ is *m*. If $m \geq M$, go to Step 2.2; otherwise, set *h*=*h*+ Δ*h*, update the bandwidth, recalculate *m*, and check again.

Step 2.2 Local MAD Method. Compute the z statistics using equations (2.1) and (2.2) for the current data point x_i using all points in the band $[t_i - h/2, t_i + h/2]$. If $z \geq \bar{z}$, record x_i as an outlier.

Step 3. If i=N, go to Step 4; otherwise, set i=i+1 and go to Step 2.1.

Step 4. Remove all recorded outliers from the set $\{x_i, i = 1, \cdots, N\}$.

The above LMM algorithm shows that we initially set the bandwidth as 10 minutes, which is expanded if needed symmetrically to both sides of the band (using 1 minute as the increment for each side) until the band contains at least 25 data points. Figure 4 illustrates that if we apply this method to the *FasTrak* data of a particular day, how the bandwidths will change over time. It clearly shows that in order to have enough data samples (25 in this study) the bandwidth is much larger in off-peak hours than peak hours. For example, nearly 60% of bandwidths in PM peak hours are less than or equal to 16 minutes and the minimum bandwidth is 10-minute; whereas for other periods of the day, the minimum bandwidth are much larger. This is consistent with the trip characteristics of commuter traffic: there are usually more travelers, especially those using *FasTrak* toll tags, in peak hours (the PM peak in this case) than off-peak hours. As a result, there are more *FasTrak* data samples per unit time during peak hours than off-peak hours. As discussed later in Section 2.2.2, the bandwidth generated this way is suitable for processing *FasTrak* data.

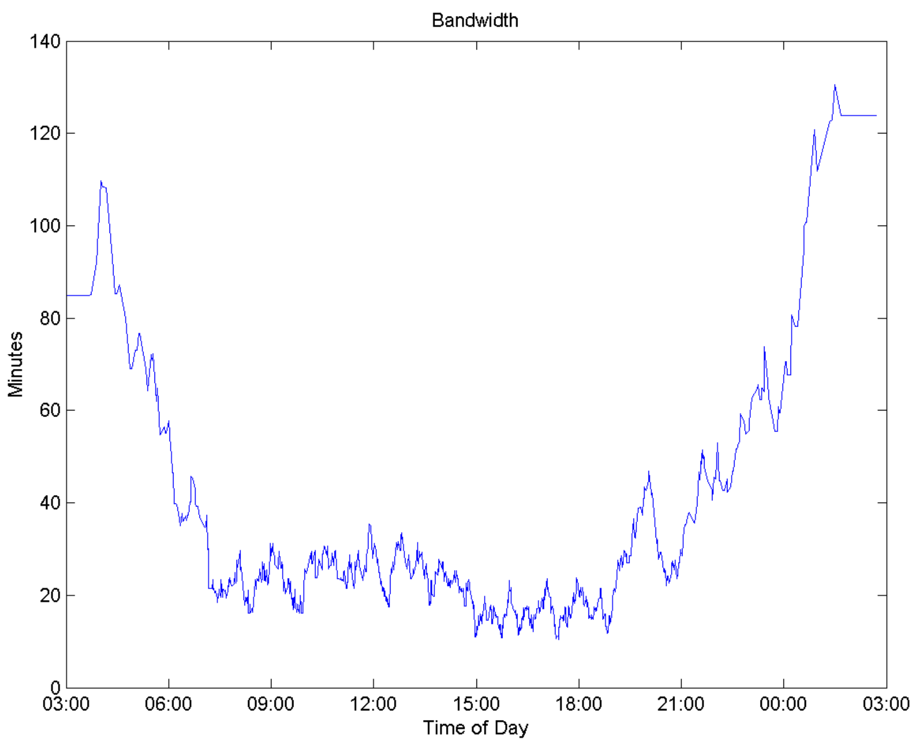

Figure 3: Variable Bandwidth

Figure 4 shows the processed *FasTrak* data, which we refer to as the "ground truth" travel times thereafter in this study. It illustrates that it is inadequate to characterize the travel time at an instant with a single value, as many previous studies did. It is also worthwhile to differentiate the travel time variations at any given time in Figure 4 with those studied in travel time reliability. In Figure 4, it comes from vehicle-to-vehicle variations (i.e. even for the same time instant, travel times from different vehicles are different due to varied driving behaviors, see Noland and Polak (2002)). In travel time reliability studies, however, travel time variations across different days (i.e., day-to-day) are considered and for a particular day travel times are treated as a single value. Therefore, *for a given time instant*, travel time variations in travel time reliability studies mainly reflect traffic condition changes across different days; in Figure 4 however, the variations are mainly due to varied driving behaviors (i.e., aggressiveness) of individual drivers, not changes of traffic conditions. We also note in Figure 4 that the dispersion of travel time is much smaller during PM peak than other periods of the day for the study route. We will show in Section 2.2.2 why this is the case.

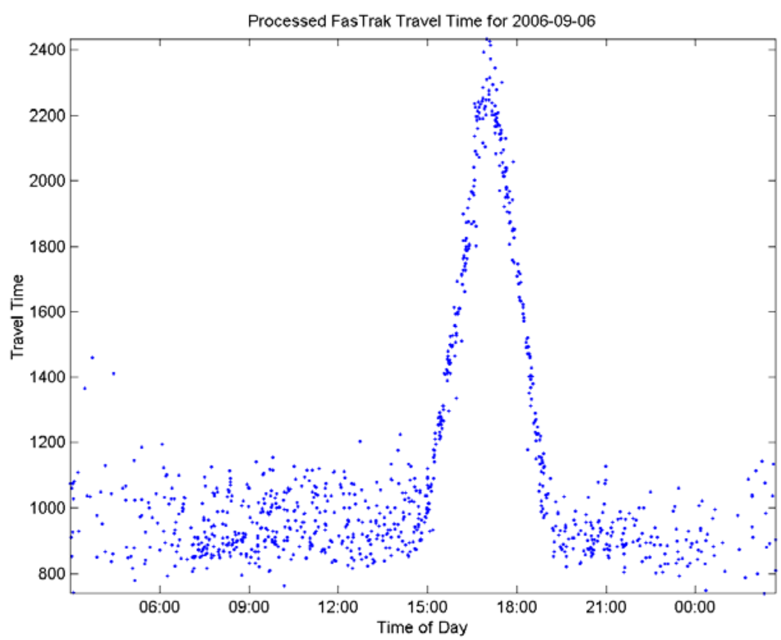

Figure 4: Processed Fastrak data

2.2.2 Characterization of Travel Time

The travel time dispersion in Figure 4 imposes difficulty on how to properly characterize travel time. One may use a single "representative" value such as the median without capturing the dispersion. To capture the dispersion, we estimate the interval of the 15th and 85th percentile travel times. Figure 5 illustrates the estimated 15th, median (50th), and 85th percentile travel times based on the processed *FasTrak* data in Figure 4. The use of the 15th and 85th percentiles is somewhat arbitrary. However, this gives us an interval that encloses the travel times of the middle 70% of drivers (in terms of driving aggressiveness).

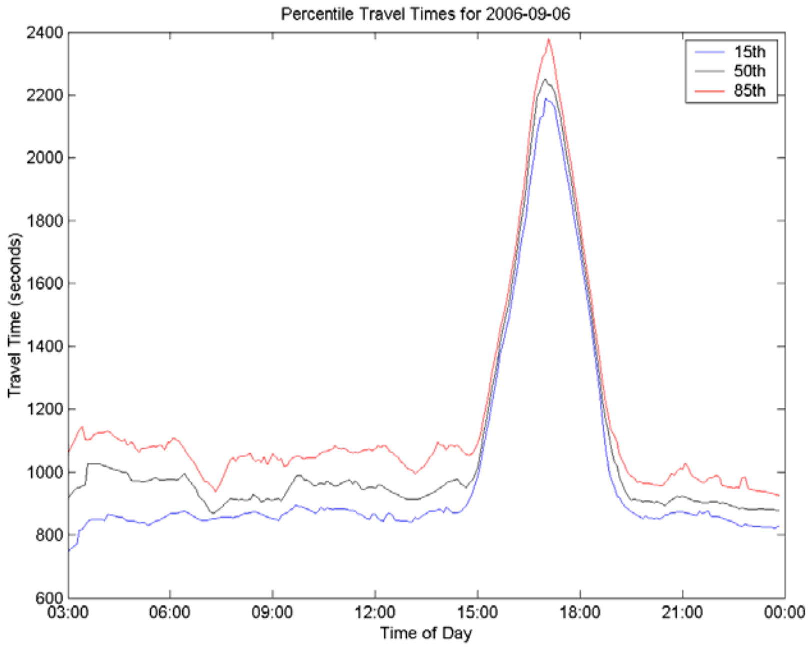

Figure 5: Percentile Travel Times

The percentile travel times in Figure 5 were obtained via the method of local linear fit, originally proposed by Koenker and Bassett (1978). This method was later applied by Small et al. (2005) and Liu et al. (2007) to process travel times computed from loop data. Details of the method are omitted here and one can refer to (Small, 2005) for more discussions.

Several observations follow Figure 4 and Figure 5. First, the most congested period for this route is the PM peak hours (15:00 – 19:00) during which the time-dependent trend of travel times is evident. For example, the median travel time increases from free-flow (around 15 minutes) to almost 35 minutes from 15:00 to 17:00 and then decreases to about 15 minutes after 19:00. For the other periods of the day, travel times are fairly stable, i.e., no obvious trend is observed. This pattern can also be observed on the other days. Second, for the congested period (i.e., PM peak hours), the travel time dispersion at a given instant is small, while it is much larger during non-congested periods. The second observation may be explained by different driving behaviors at congested and non-congested periods. That is, under non-congested periods, drivers have more freedom to stick to their preferred driving styles (aggressive or not). Therefore, the resulting travel times have more variations. During heavily congested periods (which is the case for the PM peak hours of the studied route), however, what drivers can do most times is to keep “flowing” with the congested traffic. Hence, their individual driving preferences may not be reflected at all, resulting in the nearly homogeneous travel times during the congested period. This second observation has been previously suggested by Daganzo (1997, pp. 142) who believes that in a queue, delays (and thus travel times) can be predicted “independent of drivers’ shenanigans.” Our results here provide empirical evidences for that in Daganzo (1997). Note that the travel time dispersion in off-peak hours is expected – the range is from about 800 seconds to 1000 seconds, corresponding to speeds from 55 to nearly 70 MPH.

Our observation is not inconsistent with previous findings in Chen et al. (2003) who reported that travel time variability is proportional to the mean travel time, because in Chen et al. (2003), the variability is obtained from multiple days, which as aforementioned, is mainly day-to-day variability. One should note that our observation is for a heavily congested route (like the one in this study); for lightly congested routes, large dispersions may still be observed.

The above discussions also imply that *LMM* is appropriate for processing *FasTrak* data. This is because although *LMM* generates large bandwidths during off-peak hours as shown in Figure 3, the trend of ground truth travel times does not change too much as well. Therefore, the large time-window in off-peak hours will likely not smooth out the trend of travel times. While the trend does change fairly rapidly during peak hours, *LMM* generates rather small bandwidth (most less than 16 minutes), which should be adequate to capture the time-dependent trend of travel times during peak hours.

2.3 Methodology for Performance Evaluation of Travel Time Estimation Methods

This section describes how the evaluation is conducted. The travel time estimation methods include three travel time estimation algorithms applied to speed data for individual lanes and the average of all lanes. Two quality measures are defined and applied. Following a description of the study site (route), the two scenarios used to examine the impact of different detector spacing are described.

2.3.1 Travel Time Estimation Algorithms

We test on three benchmark travel time estimation algorithms: the *instantaneous*, *dynamic*, and *linear-regression* (*LR*) algorithms. The *instantaneous* travel time assumes traffic conditions remain unchanged from the time a vehicle enters a route until it leaves the route. Therefore, route travel time can be computed by simply summing travel times of the constituent links at the time the vehicle enters the route. The *dynamic* route travel time is also a summation of travel times of its constituent links; however, the link travel time is computed using the latest traffic condition at the time a vehicle enters a particular link.

The *LR* algorithm, on the other hand, combines (linearly) the instantaneous and dynamic travel times so that the historical trend of travel times for a given route can be considered to certain extent (Rice and Zwet, 2001; Chen et al., 2004). The LR algorithm can be expressed using the following equation.

$$\tau_r^l(t) = \bar{\tau}_r^d(t) + (\tau_r^i(t) - \bar{\tau}_r^i(t)) \cdot \lambda(t) . \qquad (2.3)$$

Where we have

$\tau_r^l(t)$: the LR travel time for route *r* at departure time *t*,
$\bar{\tau}_r^d(t)$: the average dynamic travel time at time *t* computed from historical data,
$\tau_r^i(t)$: the instantaneous travel time computed at time *t*,
$\bar{\tau}_r^i(t)$: the average instantaneous travel time computed at time *t* from historical data,
$\lambda(t)$: parameter that needs to be estimated.

The parameters can be estimated via a linear regression model using historical data. For details, one can refer to Chen et al. (2004). In practice, *t* is discretized into five-minute intervals, i.e., we will have 288 parameters for a given route for an entire day.
Note that computing the instantaneous and LR travel times only requires real time speeds. Therefore, they are suitable for real time traffic applications, such as posting travel times on CMS. The dynamic travel time needs speeds in the future and thus may not be used in real time applications. However, we include the dynamic travel time algorithm in our evaluation since it provides benchmark travel times that the other two algorithms can compare with.

2.3.2 *Quality Measures*

The travel time estimates given by the algorithms are point estimates. In this section, we define two measures of the point estimate accuracy: the relative error and the accuracy index. The former is based on the absolute difference between the loop-based travel time estimate and the estimate of medium travel time based on *FasTrak* data. The latter is based on whether the loop-based travel time estimate falls into certain range of travel time estimated from the dispersion of the *FasTrak* data.

2.3.2.1 The Relative Error

In this study, we use the median of the processed *FasTrak* data as the estimate of the ground truth travel time. First, we denote $\overline{T}_r(t)$ the estimate of the median travel time for route *r* for vehicles entering the first link of *r* at time *t*. Here *t* is the discrete time instant (e.g., in every five minutes). Similarly, $\hat{T}_r(t)$ is the loop-based *estimated* travel time for the same route at time *t*. Then the relative error can be defined as follows:

$$E_r(t) = \left| \frac{\hat{T}_r(t) - \overline{T}_r(t)}{\overline{T}_r(t)} \right|. \qquad (2.4)$$

Equation (2.4) defines the accuracy measure for a particular time instant, which is referred as *disaggregated* measure. Sometimes aggregating quality measures over a certain time period may be of more interest, especially from practitioners' point of view. The typical periods of a day may include AM off peak, AM peak, mid-day, PM peak, and PM off peak (Fujito et al., 2006). For a given period *p*, the *aggregated* measure can be computed using the following equation:

$$E_r^p = \frac{\sum_{t=1}^{n} E_r(t)}{n}. \tag{2.5}$$

Here n is the total number of estimates within the period p.

2.3.2.2 The Accuracy Index

As aforementioned, the relative error measure does not capture the dispersion of the travel time random variable. In this study, we construct an interval by the estimated 15th and 85th percentile travel times based on *FasTrak* data, and define an *accuracy index* of the loop-based estimated travel time at time *t*, denoted as $A_r(t)$. The accuracy index is 1 if the loop-based estimated travel time lies in the interval; otherwise, it is zero. In other words,

$$A_r(t) = \begin{cases} 1, & \overline{T}_r^{15}(t) \le \hat{T}_r(t) \le \overline{T}_r^{85}(t) \\ 0, & \text{otherwise} \end{cases}. \tag{2.6}$$

Here $\overline{T}_r^{15}(t)$ and $\overline{T}_r^{85}(t)$ denote the estimated 15th and 85th percentile travel times at *t*, respectively, based on FasTrak data.[2] Hence, the accuracy index at a single time instant is a binary value (0 or 1). This definition can be extended to a time period, e.g., AM or PM peak hours, as follows.

$$A_r^p = \frac{\sum_{t=1}^{n} A_r(t)}{n}, \tag{2.7}$$

where n is the number of time instants in the time period p. The accuracy index over a time period, as defined in (2.7), may be more practical than what is defined in (2.6) for a single time instant.

2.3.3 The Study Site

As shown in Figure 6, a route along Interstate 80 EB from the City of Albany to the Carquinez Bridge was selected for the evaluation. In this figure, the dark arrow and “star” signs indicate, respectively, the origin and destination of the route. The length of the route is about 15 miles with the free flow travel time around 15 minutes (900 seconds) at 60 mph. We further selected four weekdays in Mid-September of 2006 for the evaluation. There are 33 double loop detectors deployed approximately evenly in this route and most of them worked properly during the four evaluation days. We used 5-minute loop speeds to compute the estimated travel times. The data can be downloaded from PeMS.
(http://pems.eecs.berkeley.edu/).

[2] For simplicity, we will subsequently refer the estimated 15th/50th/85th percentile travel time at *t* based on FasTrak data as the “15%/50%/85% ground truth”.

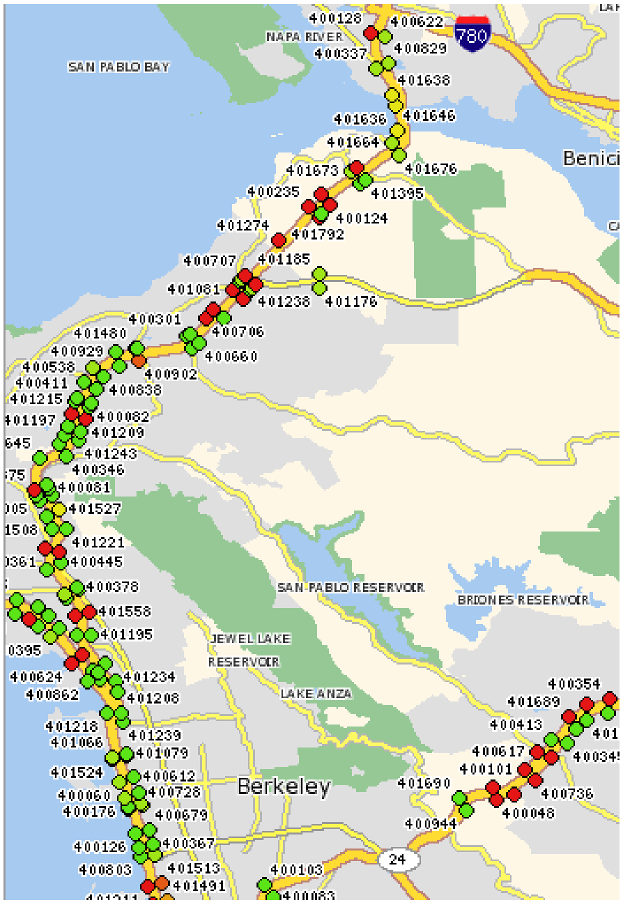

Figure 6: Evaluation Route and Loop Detectors

(source: http://pems.eecs.berkeley.edu)

2.3.4 Scenarios for Examining the Impact of Different Detector Spacings

Two scenarios are evaluated. First, we test the baseline configuration of detector spacing. The baseline configuration is what is currently deployed on the route. The spacing between adjacent loop detectors is approximately 0.5 mile. We are interested in assessing the performances of the three travel time estimation algorithms at different times of day, which we use as a proxy for recurrent congestion. In this study, different periods of a day are defined as AM off peak (00:00 – 07:00), AM peak (07:00 – 10:00), mid-day (10:00 – 15:00), PM peak (15:00 – 19:00), and PM off peak (19:00 – 0:00). Further, because there are loop detectors in each lane, we compute the lane-by-lane travel times and compare the resulting performances.

In the second scenario, we vary detector spacing and investigate how it will impact travel time estimations. For this purpose, we randomly take out detectors. This scheme was also used in Kwon et al. (2006) and Fujito et al. (2006). However, detectors were removed in Kwon et al. (2006) purely randomly so that the remaining detectors may be distributed highly unevenly. Thus for the same average spacing, the estimated travel times may be very different depending on whether sensors are evenly spaced or not. In Fujito et al. (2006), on the other hand, detectors were taken out in such a way that the remaining detectors are distributed almost evenly. Hence, the method in Fujito et al. (2006) resulted in only a few detector deployment settings for a given number of remaining detectors and the variation of performances may not be easily captured.

In this study, we generalize the random selection process in the above two studies. In particular, we randomly take out detectors in such a way that the remaining detector spacing satisfies the following condition:

$$\max_{\forall i \in N} s_i - \min_{\forall i \in N} s_i \le \rho \bar{s} \,. \tag{2.8}$$

Here s_i is the i-th spacing, $\bar{s}$ is the average spacing, and ρ is a constant. By using different values of ρ, one can control the variations of individual detector spacing for a given number of detectors. For example, if ρ=0, we require sensors to be deployed absolutely evenly (close to what was done in Fujito et al. (2006)); if $\rho = +\infty$, sensors can be selected completely randomly (close to what was done in Kwon et al. (2006)). In our study, $\rho = 2$ is used.

We implemented the above method as a random selection algorithm. For the route shown in Figure 6, we took out detectors in such a way that the resulting average detector spacing is 0.75 mile, 1 mile, 1.5miles, 2 miles, 2.5 miles, 3 miles, and 5 miles, respectively (i.e., the remaining number of detectors is correspondingly 20, 15, 10, 8, 6, 5, and 3). For each of the above seven scenarios, we ran the random selection algorithm for multiple times so that 100 *distinct* detector settings were obtained. These detector settings are used later to evaluate the performances of travel time algorithms under different detector spacing configurations.

2.4 Evaluation Results

In this section, we present the evaluation results that illustrate the performances of the travel time methods evaluated in this study.

2.4.1 The Baseline Scenario

Estimated travel times using the three algorithms for the evaluation route on September 6, 2006 are shown in three figures: Figure 7 for instantaneous travel times, Figure 8 for dynamic travel times, and Figure 9 for LR travel times. In each figure, estimated travel times using data on individual lanes are shown in different broken lines. Most part of the evaluation route has four lanes, but the last 1/3 of the route (about 6 miles) has only 3 lanes. Therefore, lane-specific travel times for lanes 1 – 3 are shown in the figures. Lane 1 is the left-most lane, and during peak hours only high-occupancy vehicles are allowed in this lane. Lane 2 is the second from the left, and Lane 3 is the third from the left. Estimated travel times for these three lanes were calculated. The line labeled as “All Lanes” resulted from the average speed across all lanes except for PM peak hours (from 3:00 PM – 7:00 PM). During the PM peak, the average speeds are calculated by excluding Lane 1 speeds. This is because, as mentioned in Section 2.2.1, the *LMM* method in this study effectively filtered travel times from Lane 1 during the PM peak. For a fair comparison, therefore, Lane 1 speeds are also excluded when calculating average speeds for “All Lanes” travel times. Further, the ground truth travel times are plotted using solid lines, each representing a different percentile.

It turns out that for the evaluation route, when the same speed data are used, different estimation algorithms do not make much difference, under both the free-flow and the recurring-congestion conditions. Theoretically, dynamic travel times should be superior to instantaneous travel times, when congestion forms or dissipates rapidly. However, because the route travel time is relatively short (15 minutes under free-flow condition, and 35 minutes when most congested), and the transition from free-flow to maximum congestion is slow (taking almost 2 hours), the results using different estimation algorithms are not significantly different. This suggests that under similar circumstances, using the instantaneous travel time algorithm is sufficient.

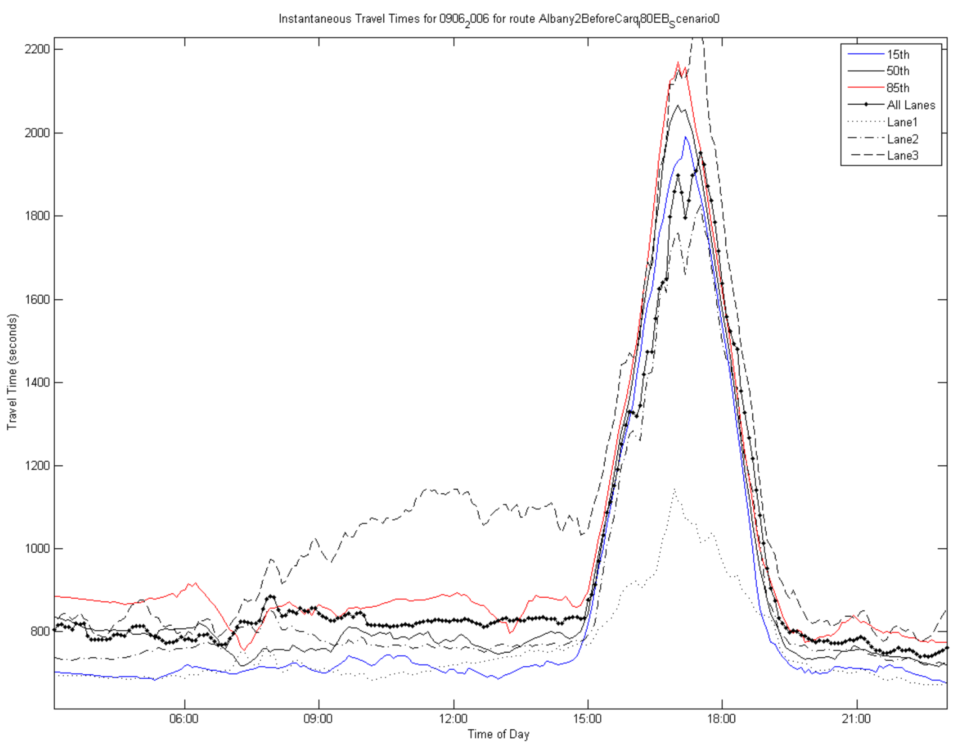

Figure 7: Instantaneous Travel Times Vs. Ground Truth Travel Times

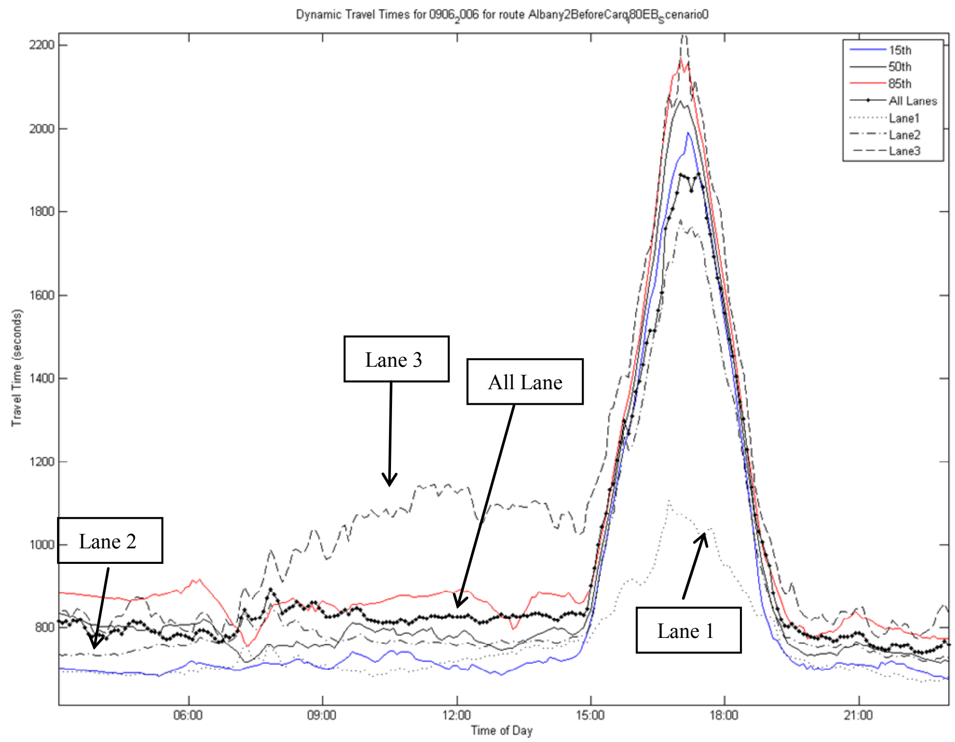

Figure 8 Dynamic Travel Times Vs. Ground Truth Travel Times

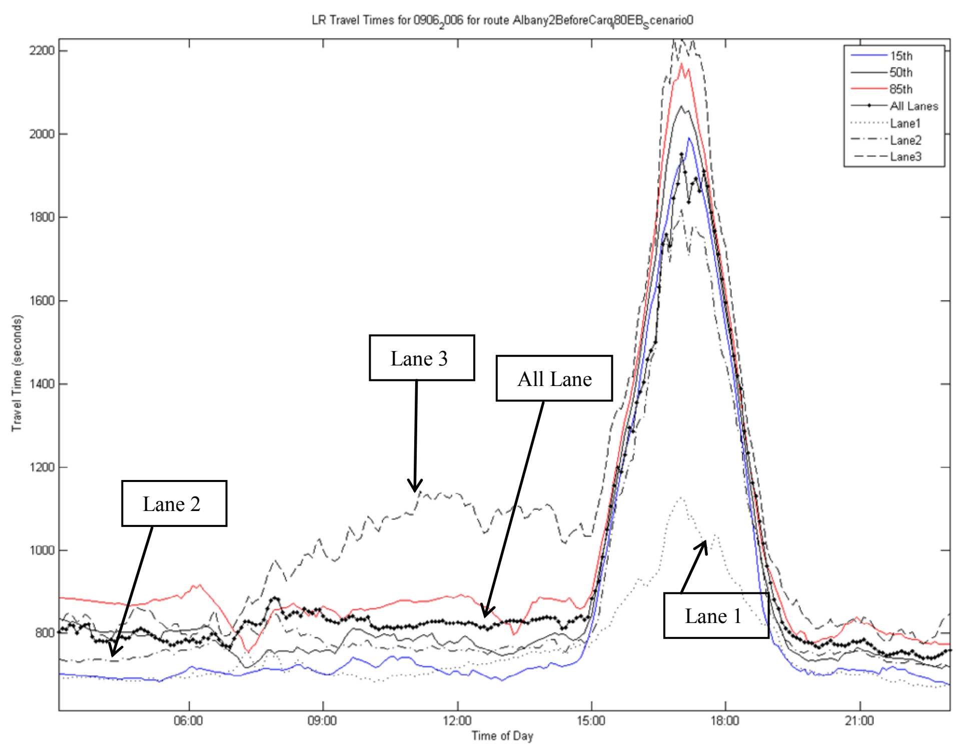

Figure 9: Linear-Regression Travel Times Vs. Ground Truth Travel Times

Our second observation is that data from different lanes affects estimated travel times greatly. Since all three estimation algorithms yield similar results, we focus on Figure 8 for simplicity. During the free-flow period, travel times calculated using Lane 1 data are the closest to the 15% ground truth line; travel times estimated with Lane 2 data are the closest to the 50% ground truth line; and travel times using Lane 3 data are often longer than the ground truth (except during 0-7am). This is plausible since the speed in the right lane is lower than others, and travelers going through the whole route (who are the most-relevant users of the travel-time display) tend to stay in left lanes, except during 0-7am, when a large percentage of through traffic are trucks, and they stay in right lanes more often than cars. Travel time estimated with data from all lanes (i.e., the average speed) is closest to the 85% ground truth in general. If we are mostly interested in displaying the travel time estimated for an "average" driver (or more precisely, a driver with median aggressiveness and speed), using data from Lane 2 is a better choice than using data for other lanes or for all lanes (for free-flow periods).

During the recurring-congestion period (i.e., PM peak), the percentile ground truth lines plotted should be interpreted a little differently. As explained in Section 2.2.1, we processed *FasTrak* data to eliminate outliners, which in this study effectively removed travel time from HOV vehicles during PM peak hours. When we turn to the loop-detector data, it is not surprising that estimated Lane 1 travel times are way below the plotted percentile ground truth lines, because the former is HOV travel times and the latter is for regular lanes. Meanwhile, travel times from "Lane 2" underestimate the ground truth travel times during this period, while "Lane 3" travel times generally overestimate the ground truth travel times. Travel times estimated with "All Lanes" data (note that Lane 1 speeds are excluded during this period), however, are the closest to the ground truth lines. The above findings can also be observed for the other evaluation days.

The second observation can also be easily seen from Table 1, which shows the aggregated relative errors and accuracy indexes of dynamic travel times for the average and lane by lane speeds computed using equations (2.5) and (2.7). In this table, we show in bold text the best performance for each time period. We can see, especially through the relative errors, that "All Lanes" travel times have the best performance during PM peak period, while "Lane 2" travel times are the best for the other periods of the day (except the AM off-peak period during which the "All Lanes" error is slightly smaller than that for "Lane 2"). Note that while one prefers small relative errors, larger values of the accuracy index represent higher possibilities that the loop-based estimated travel times lie in the interval of *FasTrak*-based estimated 15th and 85th percentile travel times and are thus more desirable.

Table 1: Relative Error and Accuracy Index of Dynamic Travel Times

Period of the Day	Relative Error				Accuracy Index			
	All Lanes	Lane 1	Lane 2	Lane 3	All Lanes	Lane 1	Lane 2	Lane 3
AM Off Peak	**0.02989**	0.12954	0.06225	0.044	**1**	0.3333	0.7917	0.9167
AM Peak	0.13502	0.06701	**0.06039**	0.26243	0.1667	**0.6944**	0.6667	0
Mid Day	0.10898	0.07101	**0.02768**	0.39492	0.5667	0.6667	**0.9667**	0
PM Peak	**0.05088**	0.34182	0.08531	0.09826	**0.3125**	0.0417	0.2917	0.1042
PM Off Peak	0.0771	0.05555	**0.01893**	0.11604	0.551	0.3469	**1**	0.3061

2.4.2 Impact of Different Detector Spacings

To evaluate the impact of different detector spacings, we only investigate the congested period (i.e., the PM peak hours). One may expect that detector spacing does not change much of the performance for non-congested periods when vehicles are in (nearly) free-flow. Our previous discussions showed that the average travel times estimated using "All Lanes" speeds are the closest to the ground truth travel times during PM peak hours (note that Lane 1 speeds were excluded in this period). Therefore, we focus on average travel times in this section. The performance measure we use for this purpose is the aggregated relative error defined in (2.5).

We first show, for each of the four weekdays, the aggregated relative error vs. detector spacing, as depicted in Figure 9. Since for each detector spacing scenario (except the baseline scenario), we randomly generated 100 detector deployment configurations, we show in each figure the minimum, maximum, median, and 25th and 75th percentile relative errors among these 100 configurations. Because there is only one baseline scenario (which is currently employed in the field), no variation exists for the baseline (i.e., the 0.5-mile spacing) in Figure 10. However, its relative error is still presented in the figure for comparison purposes.

Two observations thus follow. First, the median relative error increases, slowly and nearly monotonically, as detector spacing increases. This is intuitive since as the distance between detectors increases, less information can be collected regarding the traffic condition, which leads to less accurate travel time estimation.

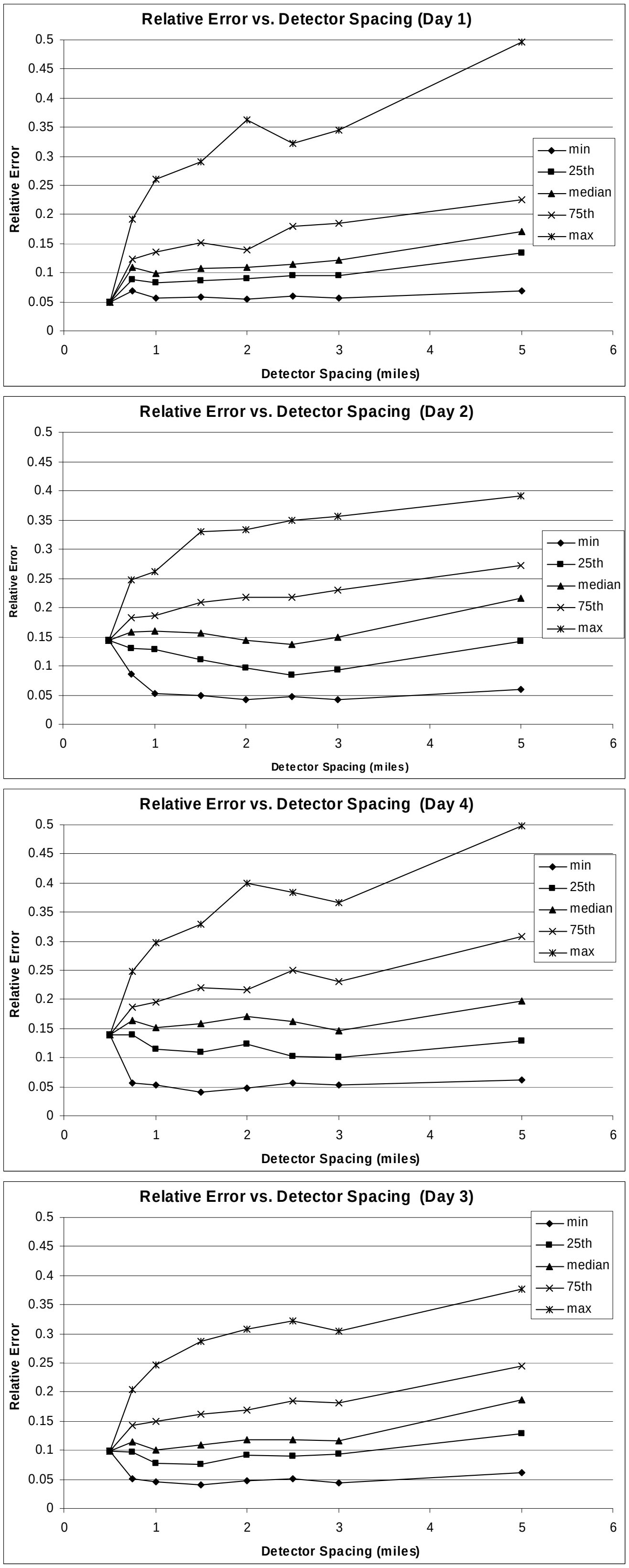

Figure 10: Performance of Detector Spacing – Relative Error

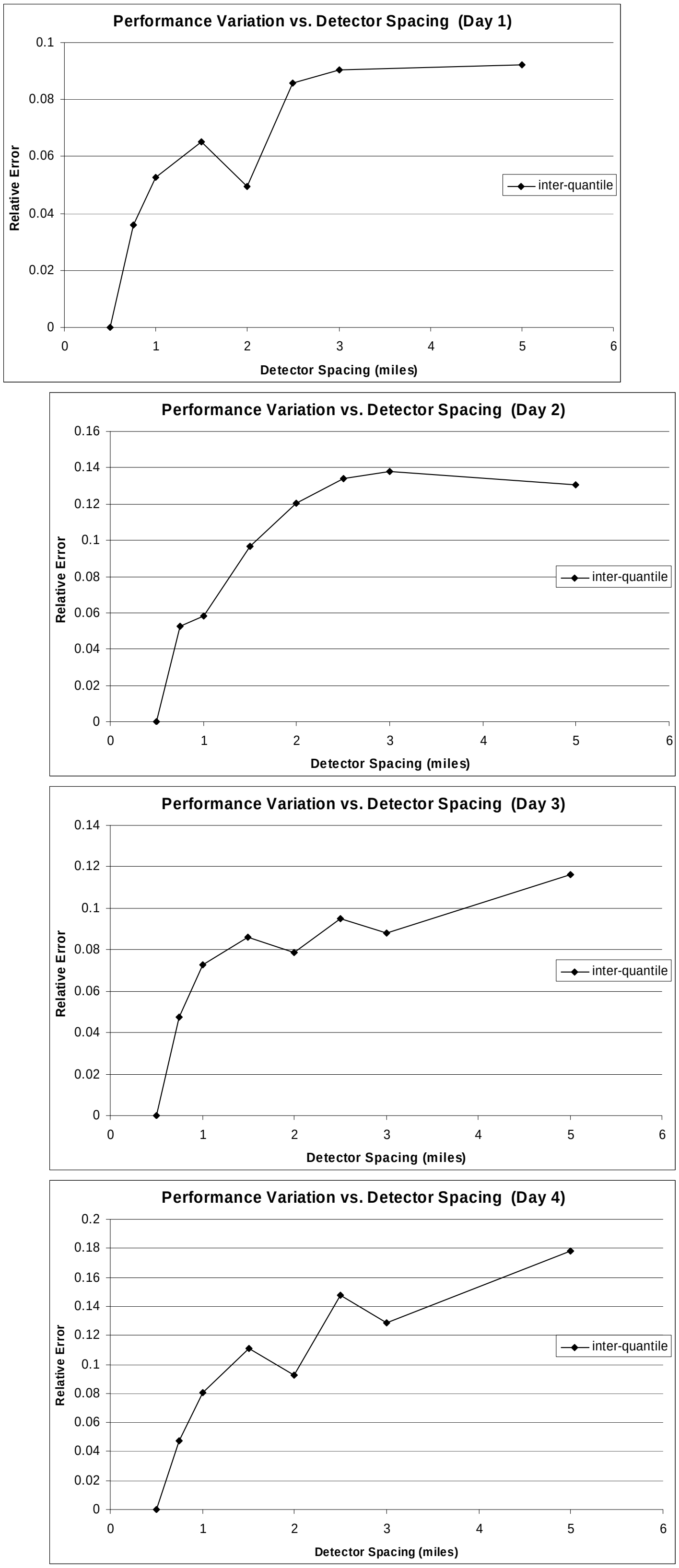

Figure 11: Performance of Detector Spacing – Variation

The second observation is that as detector spacing increases, so does the variation of the relative errors. This can be seen more clearly in Figure 11 which shows the difference of the 75th and 25th relative errors (i.e., the so-called inter-quartile of relative errors). From Section 2.3.4, the variation of relative errors for each detector spacing scenario for a single day is obtained from 100 detector configurations, which were randomly generated in such a way that they are nearly evenly distributed. Since the data used in this study cover only 4 days, the fact that a certain configuration of sparse detectors may generate small relative error for a day may merely be a coincidence, because such a configuration is picked out of 100 configurations based on the one-day performance. We cannot conclude from the data available whether there exists a sparser configuration that generates small relative error over longer periods.

2.5 Summary of Results

In this chapter, we evaluated the performance of a set of "benchmark" methods, i.e. algorithms along with specific speed data from dual loop detectors, which estimate route travel times for real time applications. We first proposed a local MAD method to process travel time data from probe vehicles. The method is effective since it captures the characteristics of commuter trips during both off-peak and peak hours, and allows us to estimate percentiles of travel times based on ground-truth data. We verified that during heavily congested peak hours, travel time dispersion in a small interval is small.

We then compared the performance of three travel time estimation algorithms that use speed data from loop detectors. We found that when the route travel time is relatively short and the transition from free-flow to maximum congestion is slow, the differences using different estimation algorithms are not significant, and the instantaneous travel time can be adopted for its simplicity. On the other hand, using speed data from different lanes makes significant difference. For example, for the evaluation route, we recommend using middle-lane (Lane 2) data during free-flow periods and "All Lanes" data during recurring-congestion periods. This finding may be site-specific. However for individual routes, one can always archive and study historical (average and lane-by-lane) travel times computed from loop detectors, and compare them to probe vehicle data. If patterns similar to those shown in this study are found, then lane specific loop data can be used to improve the travel time estimates. This procedure is useful for certain real time applications such as displaying travel times on CMS because only a set of routes need to be considered. We also evaluated the performance of travel time estimates with different detector spacings. We found that both the median relative error and the variation of relative errors increase as detector spacing increases.

In this study, we only investigated travel times computed using 5-minute loop detectors on a single route for four weekdays. The authors are currently evaluating performances of travel time methods using other types of data sources (such as 30-second loop data and speed radar sensors) on multiple routes and more evaluation days. Finer resolution data are expected to improve the performance of travel time estimation methods during peak periods, while data over multiple days can help determine if there exists a sparser detector configuration that generates similar relative error as the default configuration over longer periods. Some preliminary results in this regard can be found in Margulici and Ban (2008a & 2008b).

CHAPTER 3

USER PERCEPTION AND PREFERENCE FOR DISPLAYING TRAVEL TIMES ON CHANGEABLE MESSAGE SIGNS IN THE SAN FRANCISCO BAY AREA

3.1 Introduction

Changeable Messages Signs on freeways are usually deployed by State Departments of Transportation as a public service. For travel time display on CMS as a technology to be considered useful by travelers and worthy of tax dollars by the general public, it is important to understand user perception and preference for displaying travel time on CMS. For users to accept and use a technology, the technology must be perceived by the users as both useful and easy to use.

Do the travelers consider the travel time display on CMS useful? What contribute to the perceived usefulness of the display to travelers? What can be done to enhance the ease of use? These are questions that can be best answered through a user survey.

In addition, from a system management perspective, a user survey may also help study the impact of CMS on drivers' route diversion behaviors. Although route diversions do not necessarily lead to a reduction in the aggregate traffic delays experienced by all users, understanding in what situations a driver will decide to use an alternate route is an important building block for studying the impact of CMS on the overall network traffic conditions and for choosing the location and message content of a CMS.

During June to November 2007, we conducted a web-based user survey in collaboration with the Metropolitan Transportation Commission (MTC) of the San Francisco Bay Area. A promotional box was provided on the front-page of MTC's 511.org website, diverting visitors to take the survey. Over 1000 visitors took the survey. In this chapter, we analyze the results from survey, and give answers to the questions raised above.

The rest of this chapter is organized as follows. We first provide a literature review to summarize related research results. Then we proceed to describe the survey results and analysis.

3.2 Related Research

3.2. 1 User Perception and Preference for Displaying Travel Time on CMS

Usefulness and ease of use are two dominant factors that affect technology acceptance. Davis (1989) has shown that 1) both the perceived usefulness and the perceived ease-of-use were significantly and positively correlated with both self-reported current usage and self-predicted future usage; 2) usefulness has a significantly greater correlation with usage behavior than did ease of use; and 3) Perceived ease-of-use might actually be a causal antecedent to perceived usefulness, as opposed to a parallel, direct determinant of system usage.

Travel time display on CMS is useful because its potential impact on travel time reduction (when a driver chooses an alternate route) and on travel time predictability (regardless whether a driver chooses an alternate route). Small et. al .(2005) found that both travel time and its predictability are highly valued by drivers, and there is significant heterogeneity in these values. Jackson (1994) argued that driver's willingness to accept and use dynamic route guidance (DRG) in part would depend its usefulness, which is shaped not only by the obvious tangible benefits such as reductions in time and distance travelled, but also by a combination of the positive and negative psychological effects which DRG will have upon the individual. He urges greater consideration of the possible psychological benefits of DRG, such as a means of stress relief.

Chorus et. al. (2005) provided a theoretical framework for accessing the value (or usefulness) of travel information. The perceived benefits of ATIS are often loosely defined as “helping the traveler to make better choices”, “reduce uncertainty a traveler faces”, and “reduce traveler’s anxiety”. This paper provides representation of travelers’ perception of the value of acquiring information based on the concepts of initial and remaining uncertainty of choice and execution and integrates these with notions of Bayesian perception updating.

Using ordered-logit and regression analyses, Mannering et. al. (1995) showed that traveler’s socio-economics, habitual travel patterns, commute congestion levels, and attitudes toward in-vehicle technologies are significant determinants of travelers’ importance ratings of in-vehicle system information.

For travelers, the ease-of-use is a major advantage of CMS compared to other information sources such as radio, internet, phone or SMS (text messages). CMS does not require any equipment on the users’ side. Using other sources of information is often an interactive process and requires user input, which may distract the driver from driving safely. A European Commission study (EU 2006) reported that “Regarding the display of Travel Times on VMS, 90% of users are satisfied by the clarity of messages. 97% think Travel Times displayed on VMS are useful, 93% think the information is reliable and 71% would like a permanent display of Travel Times on VMS.” In contrast, “77% of users are satisfied by Travel Times given on radio (and another 12% does not listen to radio). 53% are satisfied by Travel Times given on Internet (and another 43% do not use internet). Only 9% of them are satisfied by Travel Times given on Foninfo (phone service) (with 88% not using this service)”.

3.2.2 Impacts of CMS on Drivers’ Route Diversion Behaviors

It is generally agreed that accurate and reliable information provided by ATIS (Advanced Traveler Information Systems, which include CMS as a special case) can benefit the user who received such information. For example, drivers may divert to alternative routes if severe incidents happened on freeways to avoid congestion if such information is provided on CMS. If the ATIS is properly deployed, the user response such as route diversion would also lead to improved system performance like reduction in the total system travel time (Khattak et al., 1994; Peeta and Ramos, 2006; Chavan et al., 2008).

In the literature, the diversion behavior of drivers under the influence of CMS information has been extensively investigated. In particular, many studies strived to find the most important contributing factors to the diversion behavior and how these factors impact diversion.

3.2.2.1 Factors that Impact Route Diversion

Existing studies have revealed that three major groups of factors do influence drivers' diversion, including trip characteristics, driver characteristics, and information characteristics (Khattak et al., 1994; Mahmassani et al., 1990). Trip characteristics represent whether the trip is a commuter trip or not, how flexible the work schedule is, the general traffic condition of the trip, etc. Driver characteristics include socioeconomics features of the drivers such as gender, age, income, etc., as well as drivers' attitude towards risks (i.e. risk-taking behaviors) and their general perception of the information displayed on CMS (whether they feel the information is accurate and/or useful). Information characteristics denote the type of messages displayed on the signs (e.g. incident alert, general traffic condition or explicit travel time information), the accurate and reliability of the information, etc.

3.2.2.2 Modeling Approaches

Two modeling techniques have been widely adopted to model the relation between the above factors and the diversion propensity. The first is linear regression models based on the assumption that the dependent variable (the propensity in particular) is a linear function of independent variables. These models were used in Mahmassani et al. (1990), Iida et al. (1994), Kraan et al. (2000), and Pan and Khattak (2008), to name just a few. In particular, Mahmassani et al. (1990) found that the most significant factors are the characteristics of the trip and information, while personal characteristics have less significant effects. Their findings do not completely agree with other studies, such as those by Khattak et al. (1994) and Vaughn et al. (1995) who discovered that driver characteristics are also important factors. Linear regression models were also used by Krann et al. (2000) to study how the route choice percentage is impacted by the queue length difference between two alternative routes.

The second modeling methodology is discrete choice models, such as logit or probit models. Uchida et al. (1994) used multinomial probit and logit analyses to study the strategic (long-term) and tactical (short-term) reactions of drivers under travel time information. Emmerink et al. (1996) applied ordered logit models to explore the impact of CMS information on route choices. The factors that are considered include demographic and those related to alternative routes. They found that women and commuters are less likely to be influenced by traffic information, which is not consistent with some of the previous studies. For example, Caplice and Mahmassani (1992) reported that women are more likely to divert; Mannering et al. (1994) found that commuters are more likely to divert. The ordered logit model was also used by Peng et al. (2004) to study the impact arterial CMS. They concluded that arterial CMS have more significant impacts than those on freeways. Recently, by comparing three types of binary probit models, Gan et al. (2008) found that travel time savings and driving ages are the most significant positive factors.

Other modeling approaches include hybrid tree models by Lee et al. (2008) and pure statistical analysis such as correlation analysis (Kawasaki et al., 2000) and ANOVA (ANalysis Of VAriation) analysis (Wang et al., 2006).

3.2.2.3 Data Collection Techniques

Travel survey has been the dominant means to obtain data in the aforementioned studies (Mahmassani et al., 1990; Khattak et al., 1994; Ullman et al., 1994; Iida et al., 1994; Emmerink et al., 1996; Peng et al., 2004; Peeta et al., 2000). Other researchers applied actual traffic measurements (Kawasaki et al., 2000) or traffic simulation (Vaughn et al., 1995; Diakaki and Papageorgiou, 1997). Uchida et al. (1994) and Kraan et al. (2000) utilized both survey data and observed data in their study. Wang et al. (2006) applied all three methods to study how the message contents of CMS may impact drivers' behaviors, although the simulation they applied is a lab-driving simulation instead of traffic simulation.

The survey method includes non-committal (stated intention and stated preference) surveys, revealed preference surveys, or a combination. The advantage of non-committal survey methods is that they can be used to study a wide spectrum of disaggregated factors such as personal characteristics which are not easily reflected in traffic data analysis or simulation. Also properly designed surveys can help surveyor only focus on polling data on the most significant factors. However, as pointed out in many previous studies (Peeta et al., 2000; Katsikopoulos et al., 2000), non-committal survey approaches also have drawbacks; in particular they may not be able to accurately reveal the actual choices of drivers. For example, Chatterjee et. al. (2002) reported that "A survey of drivers' actual responses to a message activation showed that only one third of drivers saw the information presented to them and few of these drivers diverted, although many found the information useful. Only one-fifth of the number of drivers diverted compared to that expected from the results of the stated intention questionnaire." However, "survey data for another UK city with a newly installed VMS system showed that the number of drivers diverting due to VMS information was very similar to that expected from the results of the stated intention questionnaire."

As non-committal surveys try to extrapolate results derived from the surveyed group of people to a larger population, they typically suffer from the following biases:

1. Self-selectivity bias: Only people with certain characteristics choose to take the survey or answer certain questions;
2. Policy response bias: People may answer strategically in order to affect future policy outcome;
3. Non-committal bias: This arises when the hypothetical choices do not reflect people's budget or other constraints on behaviors, or when cognitive burden may prevent the user from recognizing real-world situation that correspond to a hypothetical situation.

Using data collected from the field and conducting revealed preference study, on the other hand, are more objective, but are usually aggregated measures (such as volume, delay, etc.). Therefore, they cannot be used to link the disaggregated factors (such as age, gender etc.) to actual travel behaviors. As such, important determinants may not have been included. Also, in reality, there are always limitations on the coverage or resolution of the collected traffic data. Sufficient data are rarely available. In addition, actual choice of people may be habitual.

Simulation studies can provide detailed traffic measurements. For example, Chen and Mahmassani (2004) considered travel time perception and learning process, the triggering and terminating mechanisms that govern it, and the effect of the foregoing aspects on the day to day dynamic behavior of a traffic network, particularly convergence. This type of research is primarily exploratory in nature and hardly supported by evidence. One may question the practicality of the simulation results. In addition, usually drivers' personal characteristics are not reflected in many simulation models.

3.2.2.4 Specific Considerations on Travel Time Information

Only a few existing studies did explicitly consider travel time information on CMS (or via other ATIS channels). These studies revealed that the travel time difference between alternative routes is a major contribution factor to route diversion (Ullman et al., 1994; Kawasaki et al., 2000; Krann et al., 2000). For example, Ullman et al. (1994) reported that if the difference exceeds certain threshold (15 minutes was found in the study), drivers will be more likely to divert. The study by Kawasaki et al (2001) discovered that the travel time difference of alternative routes has a positive correlation with the diversion rate, indicating that the difference is indeed a significant factor for diversion. These findings are consistent with previous theoretical explorations and experimental analyses by Mahmassani and Chang (1985, 1986) and Chang (1985), which revealed that in order for drivers to divert, there is a threshold in terms of the travel time difference between the alternative routes. Notice that in Krann et al. (2000), a proxy of travel times (queue length) was considered instead since CMS in the study only displayed queue lengths.

The study by Katsikopoulos et al. (2000) explicitly analyzed the risk-taking behavior of drivers under travel time information. They found that drivers tend to reduce both average travel times and variations. In case the alternative route has a longer travel time than the regular commute route, the risk-taking behavior will be prone; if the travel time on alternative route is shorter, risk-averse behavior will be presented.

Most studies pointed out that information accuracy is critical to the diversion behavior. Iida et al. (1994) explicitly modeled the accuracy of travel time information in the route choice behavior. Warita et al. (2001) found that most drivers (80%) prefer to have travel time estimation errors no more than 5 minutes.

3.2.2.5 Summary

In summary, existing studies on the impacts of CMS information on route diversion identified the groups of factors that may impact the route diversion behavior. However, these findings are generally not conclusive. Especially, inconsistencies exist regarding what parameters are the most significant contributing factors for route diversion and how they actually impact the diversion. For example, some argued that female drivers would like to divert more than male drivers under information influence, but others reported the reverse; some studies found that familiar drivers are more likely to divert than un-familiar drivers, but other studies revealed the opposite.

Beside inherent biases of survey methods, there are multiple reasons for these inconsistencies. First, due to data collection limitations, a study can usually cover a subset of the above identified factors only. Those missing factors may have significant impacts but cannot be easily captured.

More importantly, the inconsistencies may indicate that the impacts of factors are content-sensitive (e.g. to the area of study). This implies that the results from one site may not be straightforwardly transferred to other sites. Therefore these inconsistencies impose challenges if one aims to generalize findings and apply the route choice models in a network-level. This is further exacerbated by the fact that some of the factors such as demographic factors are difficult to capture in most network-wide models even micro-simulation models.

As pointed out by Wardman et. al. (1997), one major contributor to the inconsistencies and wide ranges of estimates on the effectiveness of CMS in persuading drivers to divert is the varying and often unknown proportion of drivers whose destination makes the message relevant to them.

Another difficulty in association route choice with travel time display on CMS, or any ATIS technology, is that multiple technologies are available to a diverse group of users, and they jointly contribute to travelers' decisions and make isolating the effect of CMS difficult and futile.

3.3 Analysis of Survey Results

The web-based survey was designed to: 1) reveal actual user acceptance and usage patterns of travel time display on CMS in the San Francisco Bay Area; 2) uncover user perception and preference for various aspects of travel time display on CMS; 3) provide information on user's route diversion behavior.

The-based survey we conducted asks users to provide certain information about themselves, the characteristics of their commute, their awareness of travel time display on CMS, their perception of the usefulness and contributing factors for the usefulness, their route diversion behavior, and their use of other ATIS systems. In total, 1111 people completed the survey. Not all questions are answered by every person who completed the survey because some questions are follow-up questions to a previous question. The question that received least answers has answers from 585 persons. We consider this a sufficient sample size for statistical analysis that follows.

3.3.1 Demographic Characteristics

Since most people who took the survey were diverted by a promotional box from the front page of 511.org, it is not unreasonable to assume they are more likely to use 511.org than the general group of all drivers in the San Francisco Bay Area, although they may not be representative of people who visited 511.org.

Among the 1111 people who completed the survey, 90 do not commute. We call the other 1021 persons commuters. Among these commuters, 181 (18%) never notice travel time messages on CMS during their commute. Demographic characteristics for the commuters who have noticed travel time messages on CMS are as follows.

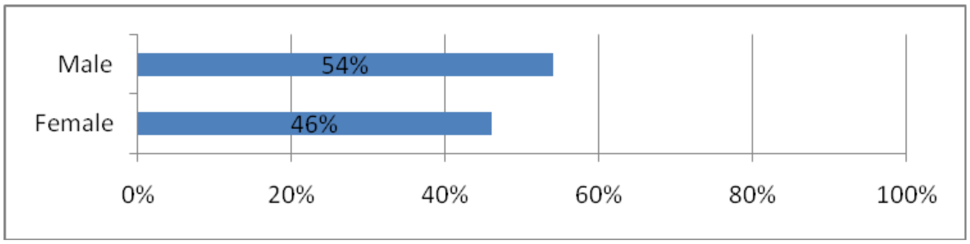

Figure 12: Gender Distribution (of Commuters Who Have Noticed Travel Time Messages on CMS)

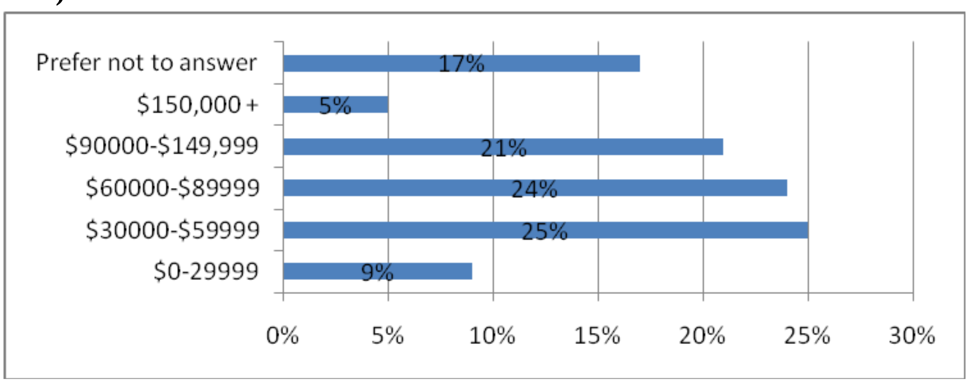

Figure 13: Income Distribution (of Commuters Who Have Noticed Travel Time Messages on CMS)

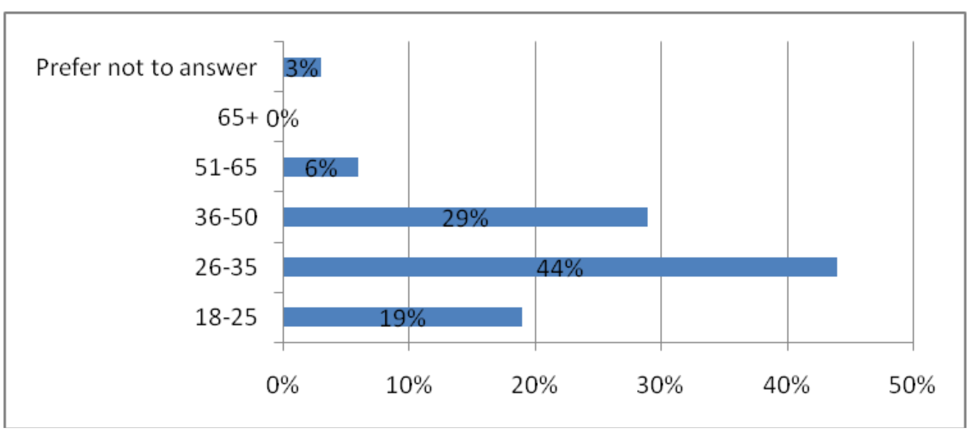

Figure 14: Age Distribution (of Commuters Who Have Noticed Travel Time Messages on CMS)

The apparent lack of respondents over 65 years old can be explained by that older people do not commute, and possibly those who do are less likely to visit 511.org. It is also possible that they are less likely to notice the CMS messages.

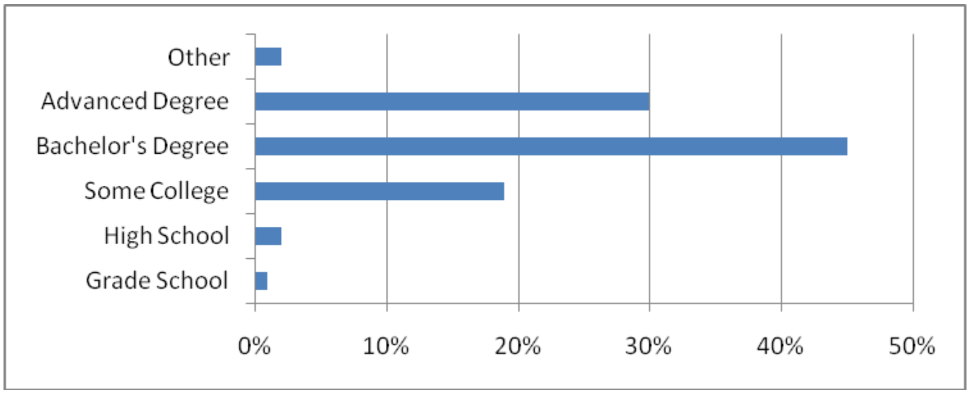

Figure 15: Education Level Distribution (of Commuters Who Have Noticed Travel Time Messages on CMS)

3.3.2 Commute Patterns

Among the 1111 people who completed the survey, their commute frequency is as follows.

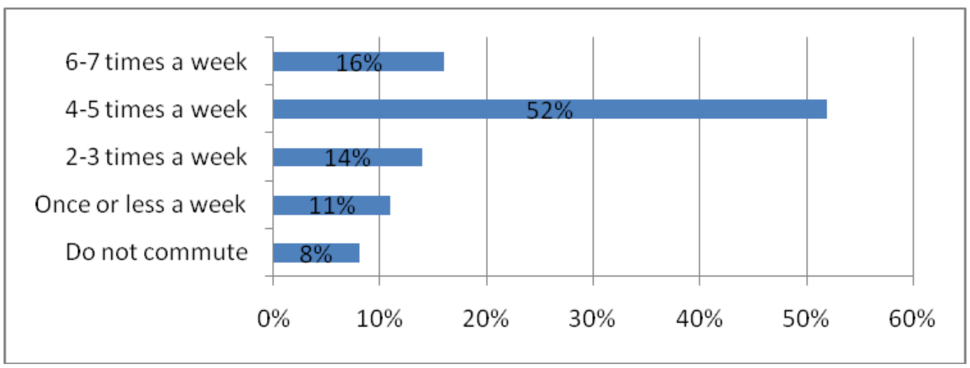

Figure 16: Commute Frequency of Survey Respondents

Among the 1021 survey respondents who commute, the distribution of the time they spend driving on an average day as follows:

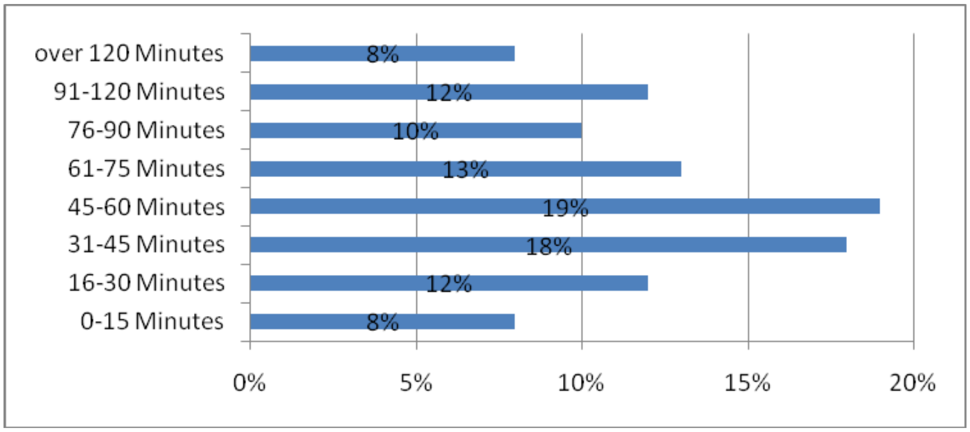

Figure 17: Time Spent Driving on an Average Day

The distribution of the length of commute routes is as follows. Note this distribution gives equal weight to each commuter, without considering how frequently a commuter commutes. So this is not a distribution of the length of all commute trips.

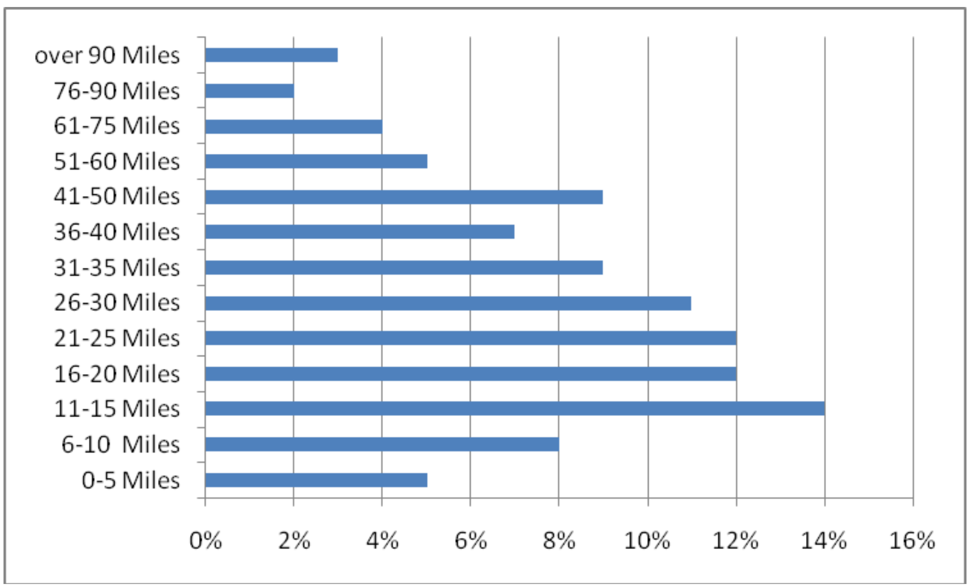

Figure 18: Distribution of Commute Route Length

A commuter may or may not be able to arrive late at his destination, with or without a severe consequence. The distribution of their schedule flexibility is as follows.

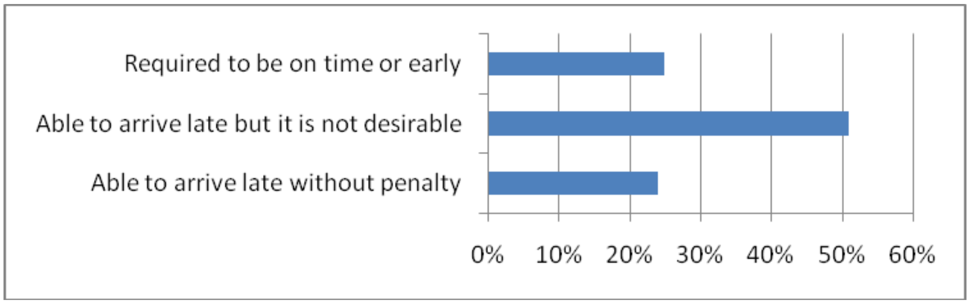

Figure 19: Commuter Schedule Flexibility

Commuters also answered questions regarding their commute origins and destinations. Major cities (reported over 50 times or more as origins and destinations) are: San Francisco (241 times as origins, 210 times as destinations), Oakland (90, 98), Richmond (30, 22), Berkeley (44, 48), Redwood City (36, 14), San Jose (43, 33). There is obvious origin-destination imbalance geographically. The following cities in the Peninsular and South Bay are far more often mentioned as destinations than as origins: South San Francisco, Foster City, Redwood City, Palo Alto, Redwood City, Mountain View, Menlo Park, Burlingame, Santa Clara, Sunnyvale, San Jose, and Cupertino. The following East Bay cities are far more often mentioned as origins than as destinations: Sacramento, Fairfield, Vacaville, Vallejo, Antioch, Concord, Pleasant Hill, Alameda, San Leandro, Castro Valley. This is consistent with our observation and experience of commute traffic patterns.

The major freeways that are used by 5% or more commuters are as follows.

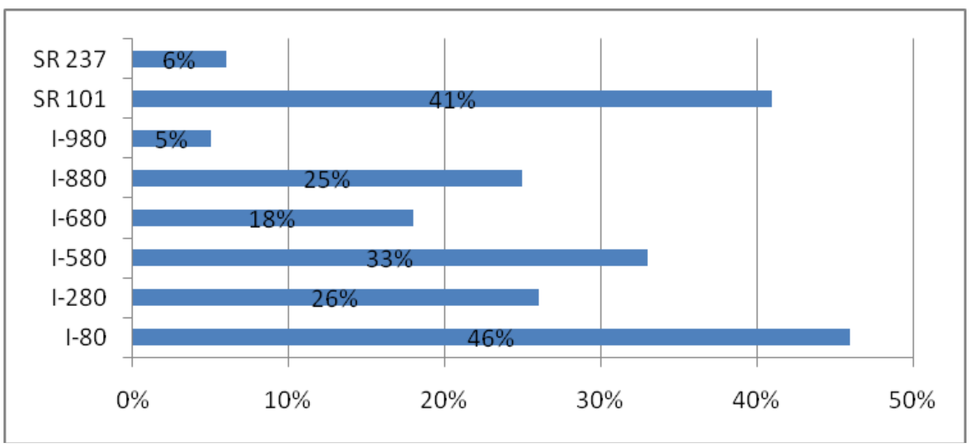

Figure 20: Percentages of Commuters Who Use Certain Freeways

Note the sum adds up to over 100% as a trip may use multiple freeways.

3.3.3 Perception of Availability and Relevance of Travel Time Display on CMS

Among the 1021 commuters, the frequency they notice travel time messages on CMS is as follows.

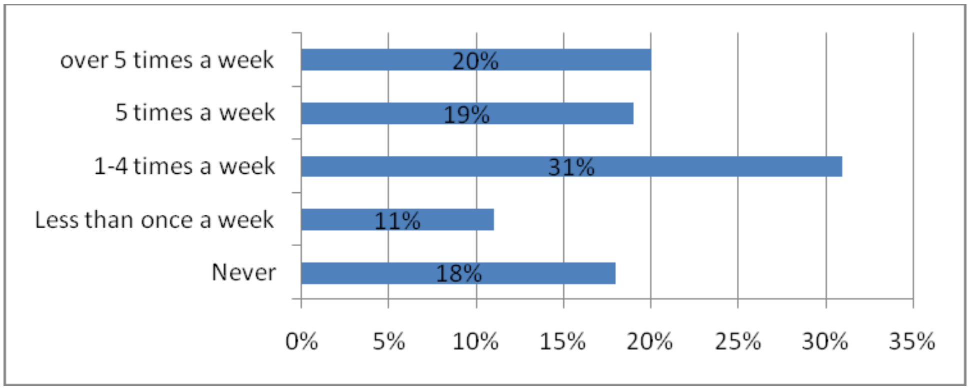

Figure 21: Frequency of Commuters Noticing Travel Time Messages on CMS

Figure 21 shows that travel time on CMS has reasonably good coverage in the San Francisco Bay Area. There is strong positive correlation between the frequency of noticing the messages and the commute route length. According to
Figure 16, 13% of commuters have a commute route length of less than 10 miles. Considering part of that 10 miles are on local streets, many of them would never see a CMS on their route. Among the 840 commuters who notice travel time messages on CMS, the perception of the relevance of shown messages to the commute route is as follows.

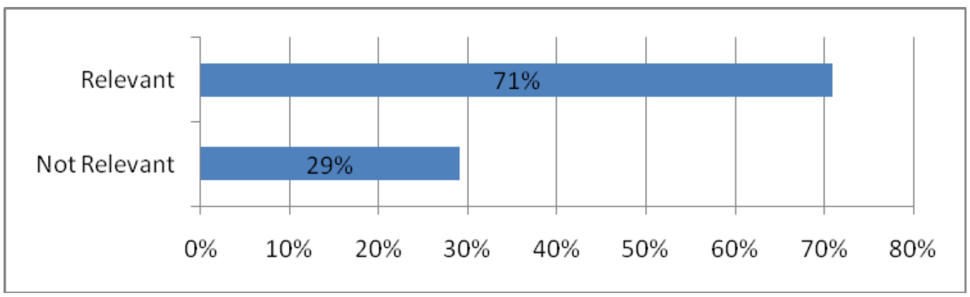

Figure 22: Relevance of Travel Time Messages to Commute Route

Although we did not ask for reasons for the travel time message not being relevant, the two figures below show, among the two groups who consider the messages relevant or not relevant, the percentages with different commute frequency and different commute route length. Among those who consider the messages not relevant to their commute routes, we see that a much larger percentage travel less frequently and a much larger percentage travel a short distance.

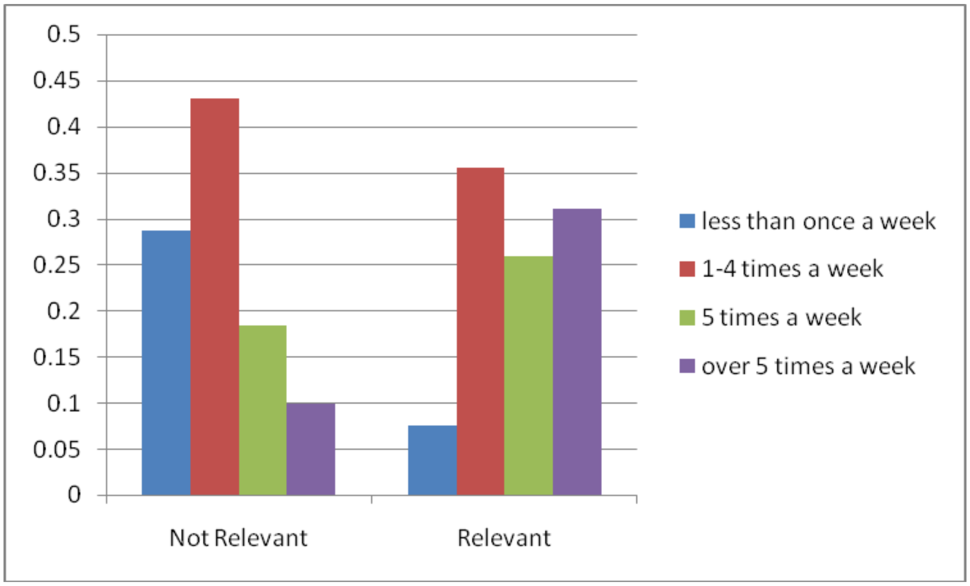

Figure 23: Frequency of Noticing TT on CMS vs Relevance of TT Message

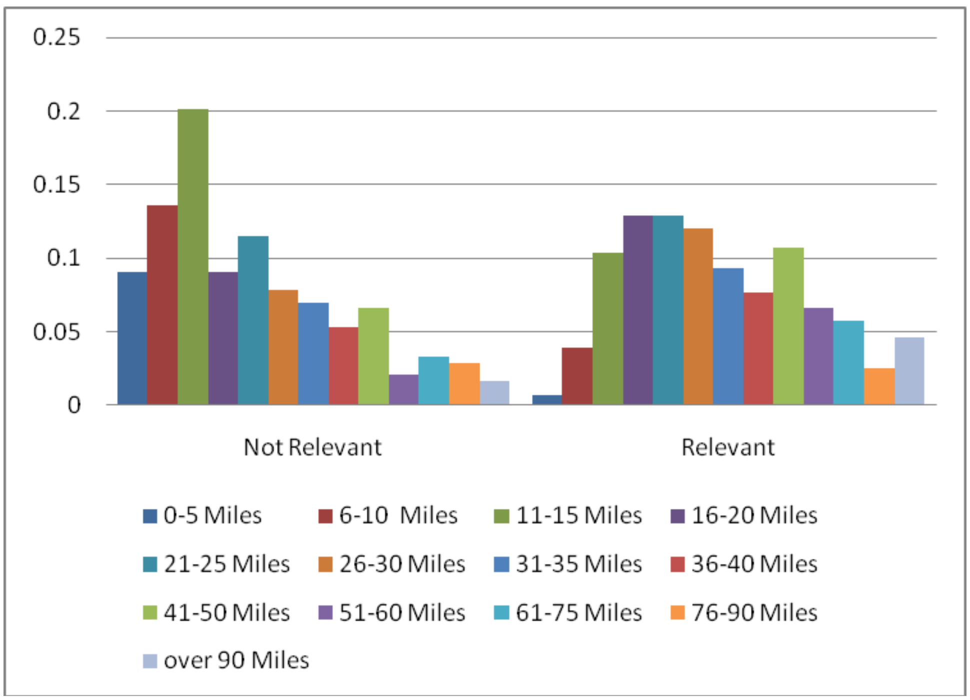

Figure 24: Commute Distance vs Relevance of TT Message

3.3.4 Perception of Accuracy and Usefulness of Travel Time Display on CMS

596 survey correspondents commute, have noticed travel time messages on CMS, and consider the messages relevant to their commute route. Their perception of the accuracy of the displayed travel time estimate is shown in the figure below. Notably, 63% of surveyed commuters believe the estimates are accurate within 5 minutes.

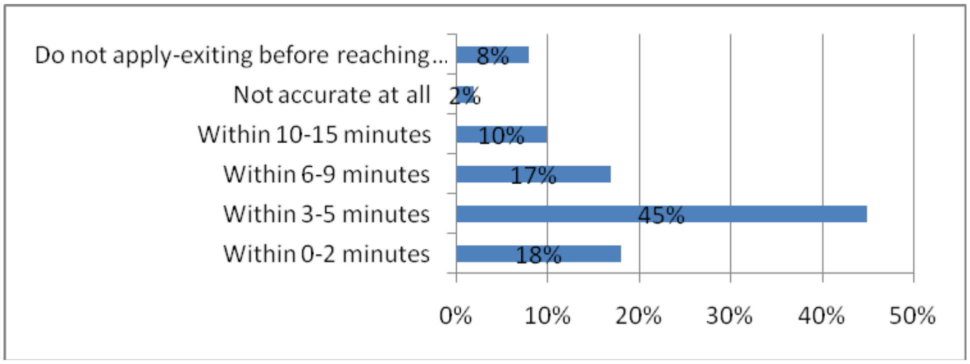

Figure 25: Perceived Accuracy of Travel Time Display on CMS

The 596 commuters were also asked to rate the usefulness of displayed messages on a scale between 1-5, with 1 being not useful at all and 5 being very useful. The result is shown in the figure below.

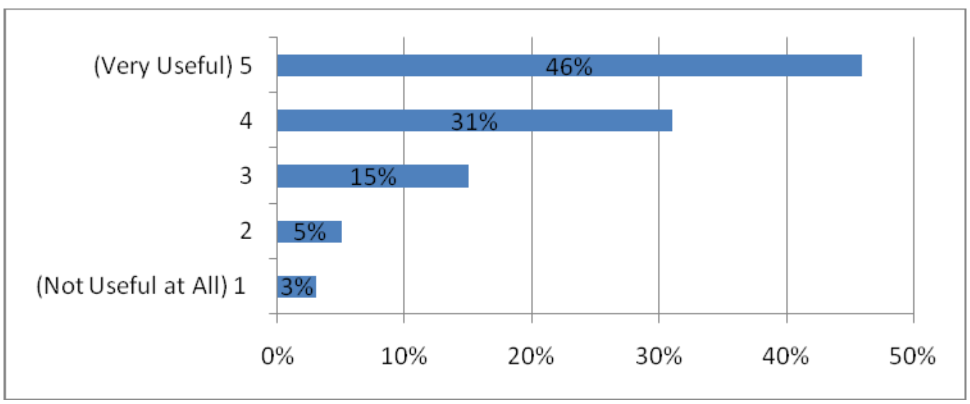

Figure 26: User Perception of the Usefulness of Travel Time Display on CMS

Compared to a survey conducted in Caltrans District 7 prior to 2007, this result is highly encouraging. In the District 7 survey, 60% of respondents are against displaying travel time on CMS. Major complaints were slow down due to the signs, hard to read due to sunshine, and accuracy. Our survey shows that an overwhelming majority of respondents consider the display useful.

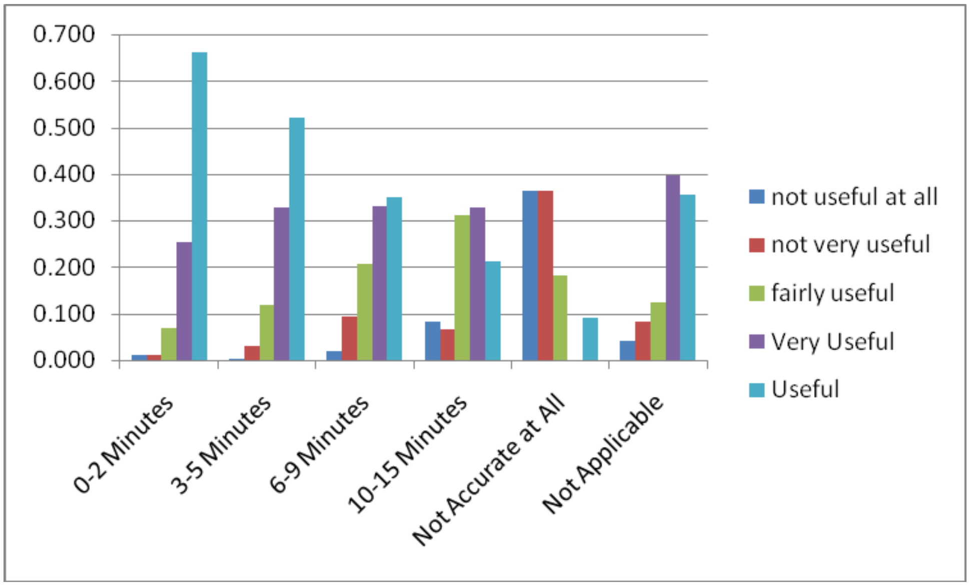

Figure 27: Impact of Estimate Accuracy on Perceived Usefulness

It is apparent from Figure 27 that the perceived accuracy of travel time estimate greatly affects the perceived usefulness of travel time display on CMS. However, even when the perceived accuracy is within 10-15 minutes, the vast majority (over 80%) of the commuters who hold the accuracy perception still regard the travel time display on CMS as useful.

In a previous meeting to discuss displaying travel time on CMS, representatives from Caltrans headquarters and districts hold different positions on whether alternative route travel time information should be displayed. Proponents of displaying travel times for alternative routes believe that the ultimate goal of providing travel information is to empower travelers to make better choices and decision, and not displaying alternative route travel time will not enable drivers to make any decision. The opponents believe displaying alternative route information is not a priority, partially because doing so may be too complicated as many alternative routes may be available.

In our survey, we asked users the principal reason for the usefulness of travel time. The results are shown in the figure below.

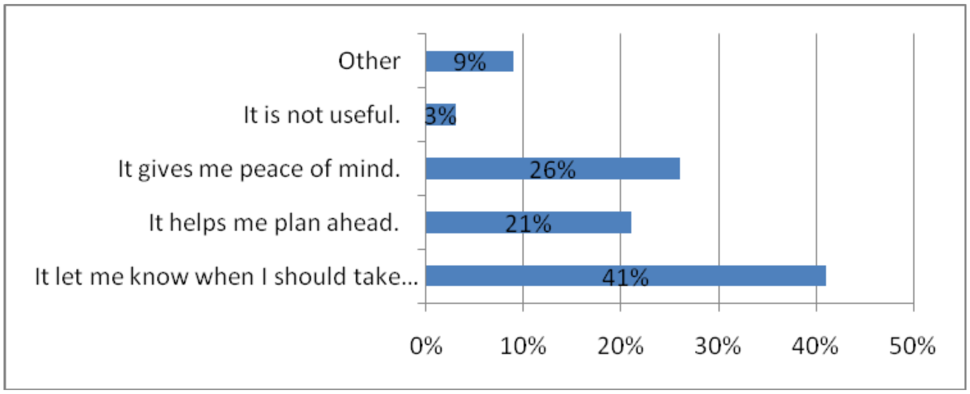

Figure 28: Principal Reasons for Usefulness of Travel Time Display on CMS

First of all, enabling route choice is the biggest reason for the usefulness of travel time display on CMS. However, the display also helps drivers to plan head (e.g. call in the expected arrival time), or gives the driver peace of mind even if no decision or choice is taken. Among the 9% who specify "Other" as the principal reason, the reasons they provided are mostly being able to call in late and knowing what to expect. Very few mentioned route choice. Only one mentioned ability to plan a mode change to BART.

The above results indicate that benefits besides route choice are as important as route choice to users when they consider the usefulness of travel time display on CMS. Even if there is no route choice, it is beneficial to have travel time display on CMS if there is great variability and uncertainty in traffic condition on a freeway segment.

3.3.5 Commuters' Usage of Other ATIS Systems

Notably, half of the surveyed commuters report that they are aware of that the travel time information displayed on CMS is generated by the Bay Area's 511 system. This suggests that they are most knowledgeable than the general public regarding traffic information sources and more active in seeking traffic information.

Although it is reasonable to believe that most surveyed commuters are directed to the survey were diverted by a promotional box from the front-page of 511.org, their reported usage of the 511 phone service or the 511.org website to obtain travel time estimates is not great, as shown in the figure below. This suggests that the usage of the phone or website for travel time information by the general public may be much lower. In light of this, information services such as CMS that requires no user input or interaction are of great value.

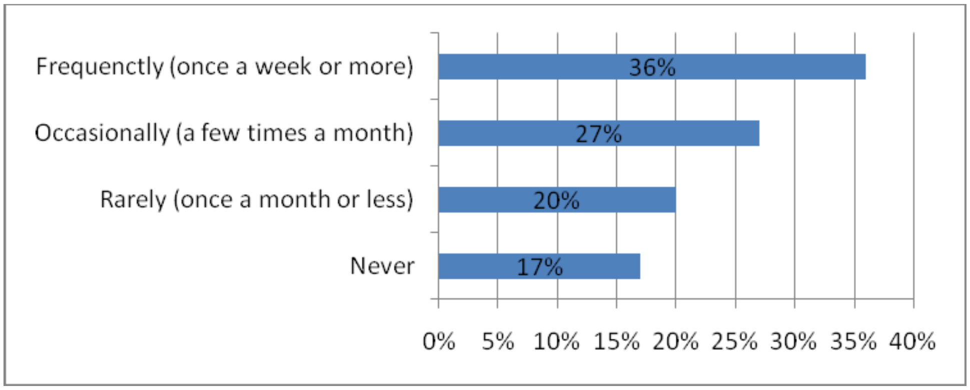

Figure 29: Usage of 511 Phone or Web Service for Travel Time Information

3.3.6 Commuter Preference for Travel Time Display Format and Update Frequency

A large majority (71%) of surveyed commuters prefer travel time estimates be displayed as an exact number in minutes instead of a range. The remaining 29% prefer a range. It is indeed difficult to interpret what the range mean exactly, and displaying a range will take more display space and less legible. Therefore, we suggest that the travel time estimate be displayed as an exact number of minutes.

When asked about what they feel about the current update frequency of travel time display on CMS, the surveyed commuters reported the following:

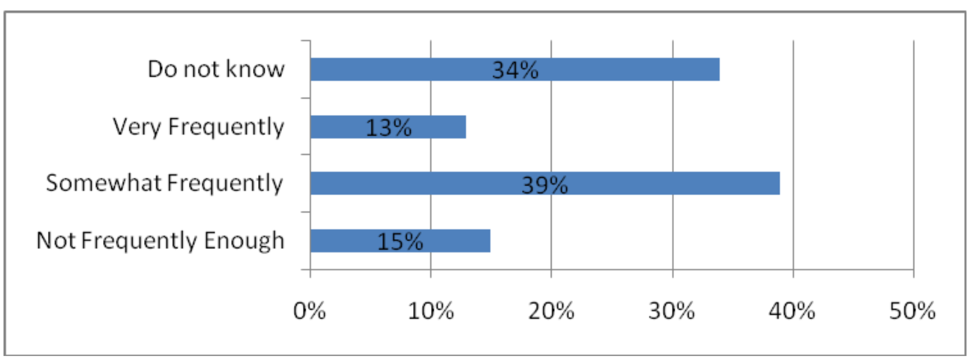

Figure 30: Commuter's perception about travel time update frequency

When asked to choose between an update frequency of every 2 minutes and a frequency of every minute, 63% of the surveyed commuters chose every 2 minutes. If the time is updated too frequently, it may confuse commuters. Thus an update frequency of every 2 minutes is recommended.

3.3.7 Impact of Travel Time Display on Route Diversion

The 840 commuters who have noticed travel time messages on CMS during their commute were asked whether they are aware of alternative routes for their typical commute. 79% answered yes and 21% said no. Thus for most people, route diversion is a real possibility.

However, for those commuters who are aware of alternative routes, the distribution of the extent to which they rely on the CMS travel times to know when to take another route is shown in the figure below.

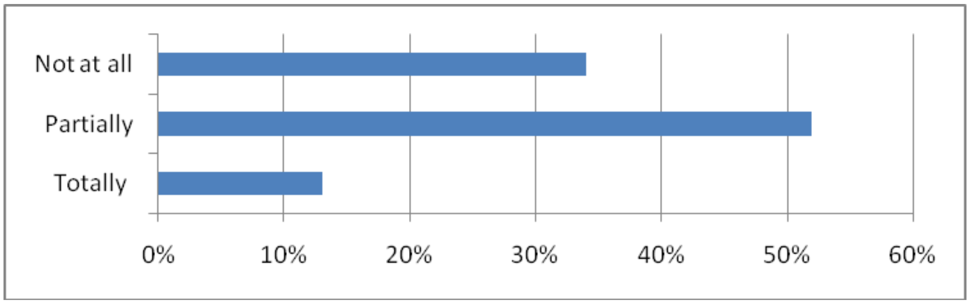

Figure 31: Extent to Which Commuters Rely on CMS Travel Times for Route Diversion

Only a very small percentage (13%) totally depends on CMS travel times.
For those who report not relying on CMS travel time at all, their diversion behaviors are given as following:
Divert only when there is big accident or total blockage;
Make divert decision by using 511, personal traffic information device, etc.
Divert based on experience (Friday, big games);
Divert when alternative route has bridge toll;
Divert when need to run errand on alternative route;
Work at home if there is traffic;
Cannot divert because CMS is too late to affect diversion;
Never divert;
By far the biggest behavior pattern is that many people will divert only in extreme traffic conditions, such as road closure or very heavy traffic. Even for people who totally depend on CMS travel times for diversion decisions, many reports they will divert only in accidents or very heavy traffic. This suggests a large threshold for the difference between expected route travel times needs to be reached before commuters switch from their usual routes.
The above results on diversion decision are for survey respondents who have noticed travel time messages on CMS during their commute, and who are aware of alternative routes and in general more savvy and active in seeking traffic information. The effectiveness of CMS alone in persuading typical drivers to divert is likely to be small.

3.4 Summary of Results

The major results we obtain from the survey are the following:
The San Francisco Bay Area achieved reasonably good coverage of travel time display on CMS;
Displayed travel times are relevant to most commuters' commute route, and among those who consider the messages not relevant to their commute routes, we see that a much larger percentage travel less frequently and a much larger percentage travel a short distance;

1. The majority of surveyed commuters believe the estimates are accurate within 5 minutes, and an overwhelming majority regard the travel time display as useful;

2. Perception of accuracy of estimated travel time greatly affects the perceived usefulness of travel time display. However, even when the perceived accuracy is within 10-15 minutes, the vast majority (over 80%) of the commuters who hold the accuracy perception still regards the travel time display on CMS as useful;

Benefits besides route choice, such as being able to plan ahead and having peace of mind, are as important as route choice to users when they consider the usefulness of travel time display on CMS.

3. Users need not to make a choice or decision to benefit from CMS travel time display. Even if there is no route choice, it is beneficial to have travel time display on CMS if there is great variability and uncertainty in traffic condition on a freeway segment.

Commuters use many sources for travel information. Travel time display on CMS requires no user input or interaction but offers great value compared to other services that require input and thus are used infrequently, by few, and unsafely.

Estimated travel time should be displayed as an exact number of minutes. An update frequency of every 2 minutes is recommended.

For most people, route diversion is a real possibility. However, a large threshold for the difference between expected route travel times needs to be reached before commuters switch from their usual routes.

The effectiveness of CMS alone in persuading drivers to divert is likely to be small.

CHAPTER 4

CMS IMPACT ON NETWORK-WIDE PERFORMANCE AND OPTIMAL CMS CONFIGURATIONS FOR DISPLAYING TRAVEL TIMES

4.1 Background

For a given CMS, there are usually possibilities to display travel times for multiple destinations and for each destination, there are multiple routes for which travel times can be posted. It is very likely that different combinations of locations, destinations and routes may result in different system performances. Therefore, *the optimal CMS configuration problem* is to determine, under normal traffic conditions, the best CMS locations and the travel time destinations and routes on each CMS. The objective is to optimize certain network-wide performance (such as minimizing the total system travel time). We also assume that all drivers are *commuters*, which is roughly the case for AM and PM peak hours.

Our proposed modeling framework for the problem contains two major components. The first one is to evaluate the network-wide performance for a given CMS configuration (i.e. locations and travel time destinations and routes). Based on this network model, the second component exploits the performances of different configurations and determines the "optimal" configuration. Before proceeding to describe our work, we summarize the literature below.

4.1.1 CMS Impacts on Network-Wide Performance

4.1.1.1 Traffic Data Analysis Method

There are many studies in the literature about how CMS information may impact network-wide performances. The most straightforward evaluation method is to analyze traffic data collected before/after the installation of CMS. For the newly installed 14 CMS in Amsterdam in the Netherlands, Kraan et al. (2000) evaluated their impacts on network-wide performances during peak hours and recurrent congestion. Total vehicle-miles-traveled (VMT), queue length, and travel time were used as performance measures for the evaluation. It was shown that after installing CMS, the total congestion decreased slightly, traffic performance increased slightly and travel time reliability also increased. The authors thus concluded that CMS has a positive impact on network-wide performances. The CMS however did not display travel time information during the evaluation.

4.1.1.2 Network Optimization Models

Network optimization models have also been developed to evaluate the network-wide performances of CMS (e.g. Peeta and Gedela, 2001). Lam and Chan (1994) modeled the CMS travel time information via both static and time-varying stochastic traffic assignment. Their key assumption is that drivers generally have perception errors regarding travel times. However, if travel times are posted on CMS for a given route, the perception over the route will be reduced to zero. It was found in the study that CMS affect more significantly the congested periods; if travel time information is provided on CMS for some OD pairs, the average trip times for these ODs can be reduced significantly. The study however did not consider stochasticity of the network such as travel time variability. In addition, the assumption that CMS information can completely eliminate drivers' perception error may need to be relaxed.

Network models have also been proposed to study ATIS (such as in-vehicle route guidance systems), a broader concept than CMS. Since ATIS usually has certain market penetration, most studies applied the so-called multi-class equilibrium models (Daganzo, 1983). That is, drivers are divided into (at least) those equipped with ATIS services and those who are unequipped. Usually equipped drivers are assumed to have perfect information and thus follow deterministic user equilibrium. Unequipped drivers, on the other hand, do not have perfect information and follow stochastic user equilibrium. Most of these studies adopted static traffic models and did not consider network stochasticity (Kousopoulos and Lotan, 1990; Maher and Hughes, 1995; Van Vuren and Watling, 1991; Yang, 1998; Lo et al., 1999). Yin et al. (2002), on the other hand, explicitly considered dynamic user equilibrium and travel time uncertainty. To capture drivers' attitude regarding travel time variability, the risk-taking behavior (in terms of how they deal with travel time uncertainty) was explicitly modeled. It was found that by assuming different risk taking behaviors (risk-averse, risk-prone or risk-neutral), the impact of ATIS will vary significantly; in certain cases, ATIS may worsen the system performance. This shows that for networks with large travel time variations, considering risk-taking behavior is necessary.

CMS information however is significantly different from other ATIS in at least one aspect. That is, one can reasonably assume that all drivers can access the information as long as they pass by the signs. Therefore, there is no need to divide drivers into groups, which implies that a single user class may suffice (e.g. the study by Lam and Chan, 1994). Also, since previous studies on ATIS have clearly indicated that in order to properly model the impact of CMS on networks with non-negligible travel time variability (which is true for most urban areas), one needs to capture both the network stochasticity and the perception errors of drivers. This results in the so-called Stochastic Network – Stochastic User Equilibrium (SN-SUE), which was first introduced by Mirchandani and Soroush (1987) and subsequently studied by Tatinei (1996), Tatinei et al. (1997), and Chen et al. (2002). In our study, SN-SUE will be adopted to evaluate CMS impacts in a network level.

4.1.2 Optimal CMS Locations

Although instructions on operational issues for CMS have been developed by MUTCD and other federal- and state-level guidelines, they usually cannot provide detailed guidance on optimal CMS configurations to effectively convey information (Chiu and Huynh, 2007). Specific studies for optimal CMS configurations are also sparse in the literature: only a few researchers addressed the optimal CMS location problem under incident scenarios (Abbas and McCoy, 1999; Chiu and Huynh, 2007). In particular, Abbas and McCoy (1999) proposed an analytical modeling framework to determine the optimal CMS locations for incident management. The potential benefits of CMS for displaying incident information were characterized by reduction of delay and accidents in freeway upstream and downstream of the incidents and that in the alternative route. The potential benefits were further adjusted by the proportion of traffic passing through the CMS that can actually divert. The model was solved using the Genetic Algorithm and the optimal locations were ordered so that a phased implementation of CMS installation was possible. The final selection criteria were based on a benefit/coast analysis: the (monetary) benefits that CMS can bring (by delay and accidents reduction) and the costs of installing these CMS. The study, however, did not consider the impact of the information on the route diversion behavior and the resulting network traffic condition changes.

Chiu and Huynh (2007) proposed methods to determine strategic CMS locations in a network level by considering stochastic incident occurrence. The authors argued that 1) CMS information impacts motorists' en-route route choices instead of pre-trip choices, and 2) dynamic traffic assignment concept is more appropriate to model such en-route choices since incident information may only be valid and displayed for a short period of time. Due to the complex interactions of user-to-user and user-to-system, CMS impacts on route choices were modeled using the DynaSmart-P simulation tool (Mahmassani, 2001). The problem was then formulated as a non-close-form minimization model, which was solved by the Tabu search. The authors claimed that the optimal results obtained by the model make sense. One major finding of the study is that ATIS and CMS nullify the marginal benefits of one another. This implies that one needs to consider the inter-dependency among different information strategies when evaluating the benefit of ITS technologies. The authors also recommended that users' choice under imperfect information be considered for future research, which is similar to the findings in Yin et al. (2002).

In summary, existing studies mainly focus on the determination of optimal CMS locations under incident scenarios. There is no modeling approach for finding optimal CMS configurations (including not only locations but also destinations and routes) for displaying travel times under recurrent congestion. This is particularly true for heavily congested urban areas (such as the San Francisco Bay Area) where travel time uncertainty is significant and needs to be considered. It is our understanding that under our specific consideration, previous assumptions made for incident scenarios may not hold or can be relaxed in certain sense.

4.2 The Network Model to Evaluate CMS Impacts

For the *network model*, several features distinguish our study from previous research. First, since the CMS travel time system is specifically modeled, we need to consider not only CMS locations but also travel time destinations and routes, which were not captured in previous research. Second, we focus on normal traffic conditions such as recurrent congestion, whereas most previous research on optimal CMS locations was concerned with incident scenarios (Abbas and McCoy, 1999; Chiu and Huynh, 2007). The reasons we study normal conditions are that 1) usually travel times are displayed only under normal traffic conditions since when incidents happen, incident alert messages will be posted instead, and 2) from our travel surveys, we found that travel times on CMS do impact positively drivers' route choices and we thus postulate that they should also help improve system performances.

The fact that we only focus on normal traffic conditions may relax previous assumptions made particularly for incident scenarios. Especially, the major argument for applying dynamic traffic assignment is that incident information is only valid for a (possibly short) duration of time (Chiu and Huynh, 2007). Depending on when a driver arrives at a CMS node upstream of the incident, he/she may or may not see the message. Static traffic assignment could not capture this difference. However, in our study, we assume travel time information is displayed all day (for example it is activated from 5:00 am to 9:00 pm on weekdays in the San Francisco Bay Area). Therefore, all drivers will see the information as long as they pass the signs. Under the assumption that the displayed travel times match well with (expected) actual travel times, commuters will learn and adjust their choice behavior. The decision can then be made pre-trip instead of en-route. This indicates that a static model may also be acceptable for our study. We note however that a dynamic model may do better and we will investigate this issue in future studies. Lastly, we consider heavily congested urban areas where travel time variability cannot be neglected. Therefore, we will explicitly consider divers' risk taking behavior in the modeling framework.

As a result, we formulate the network model as a SN-SUE problem for which both network travel time variability and drivers' perception errors can be captured. One critical issue here is to model the impact of CMS on drivers' route choices. In this study, we assume that drivers have regular perception errors and CMS will help reduce the error to certain level. Both the regular and CMS-impacted perception errors will be estimated from traffic data (such as travel times) collected from the field. Details of how to formulate and solve SN-SUE problems can be found in Mirchandani and Soroush (1987), Tatinei (1996), Tatinei et al. (1997), and Chen et al. (2002). In the following, we only illustrate how CMS influences drivers' travel time perceptions.

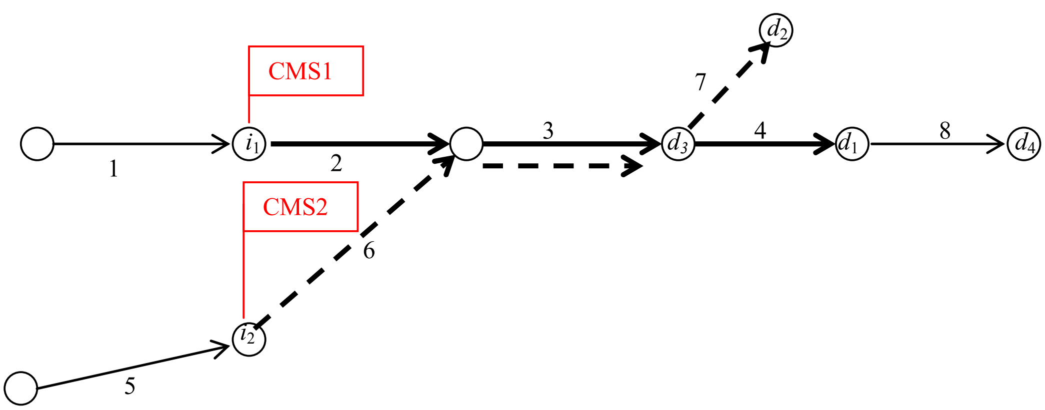

Figure 32: CMS Impact on Travel Time Perception

Figure 32 depicts two CMS installed at node i_1 and i_2 respectively. CMS1 displays travel times to destination node d_1 via a path including links 2, 3, and 4 (shown as the bold solid arrows), whereas CMS2 shows travel times to destination node d_2 via the path 6, 3, and 7 (shown as the bold dashed arrows). We assume drivers' regular (i.e. with no information provided) travel time perception errors over links are $\beta_a, \forall a \in \{1, \cdots, 8\}$ (modeled as a percentage). When CMS travel time information is provided, perception errors become

$\gamma_b = \bar{\gamma}, \forall b \in \{2,3,4,6,7\}$. Here we assume drivers have a uniform perception; in other words, when CMS provide travel time information, all drivers reduce their perception errors to the same level. This assumption is reasonable since for a given region, travel times on CMS are usually produced by a single calculation system which has a specific accuracy level depending on the utilized data sources and algorithms. As commuters interact with the travel time information, they will learn and build their perception regarding the accuracy level of the system. As a result, $\bar{\gamma}$ may be approximated as the drivers' perception regarding the travel time accuracy posted on the signs, which can obtained from travel survey data. Note that we assume $\bar{\gamma} \le \beta_a, \forall a$ in general so that travel time information on CMS always reduces drivers' perception errors.

In Figure 32, after the two signs are installed, drivers' perception errors for links 1, 5, and 8 still remain as β_1, β_5, and β_8 respectively. However, for other links, the perception error is reduced to $\bar{\gamma}$. Notice that whether a traveler is impacted by the travel time information on a CMS depends on whether the traveler passes by the CMS location. Therefore the actual path of each individual traveler has to be traced. In other words, the SN-SUE model will be implemented in path-based.

Note that the above scheme guarantees consistency for drivers traveling between different OD pairs or routes. For example, drivers from i_1 to d_1 and i_2 to d_2 will all traverse link 3. Although link 3 is covered by both CMS1 and CMS2, its unique perception error $\bar{\gamma}$ guarantees that drivers have the same perception error over the link no matter where they are from or heading to. Another observation is that the scheme can capture CMS' impact on destinations that are implicitly or partially covered by CMS information. For example, d_3 is implicitly covered by both signs in Figure 32 for commuters who know that d_3 is actually on the routes displayed on the signs. This fact can be captured by our modeling scheme since for drivers from i_1 to d_3 (or i_2 to d_3), their perception errors over links 2, 3, and 6 will be reduced to $\bar{\gamma}$. This is also true for destinations that are partially covered by the signs, such as d_4 in Figure 32. In this case, drivers from i_1 to d_4 will have reduced perception errors for links 2, 3, and 4, while maintain the regular perception for link 8 which is not covered by the signs.

4.3 Bi-Level Model to Determine Optimal CMS Configuration

Assume a given traffic network can be represented as $G(N, A)$, where N is the set of nodes and A is the set of links. We use index $i \in N$ to denote a node and $a \in A$ to denote a link. Further, $N_c \subseteq N$ a subset of N and for each node $i \in N_c$, a CMS may be installed or an existing CMS may be activated to display travel times. In other words, N_c is the set of potential CMS locations (nodes) and $i \in N_c$ is called a *CMS node*[3]. For each CMS node $i \in N_c$, there is a set of destination nodes D_i for which travel time may be displayed; for each destination node $d \in D_i$, we further assume that there are multiple (could be one)

[3] If an existing or newly installed CMS is not exactly located at a network node, an immediate node can always created at the location of the CMS. This way, a CMS is always on a network node.

possible routes (denoted as the set $P_{i,d}$) that can be displayed. In this project, we assume $N_c, D_i, \forall i \in N_c$, and $P_{i,d}, \forall i \in N_c, d \in D_i$ are given. In reality, this is usually the case since traffic operations already have some potential CMS locations in mind, and for each potential location, they often have candidate destination nodes and routes to display travel times.

We define a vector of binary variables ρ indexed by *i* to indicate whether a CMS is installed or activated at node $i \in N_c$ (ρ_i=1 for yes, and 0 otherwise). Assume there are at most *K* CMS that can be installed or activated for travel time display, we then have:

$$\sum_i \rho_i \leq K. \tag{4.1}$$

Denote another vector of binary variables λ which can be indexed by *i*, *d* and *p*. In particular, $\lambda_{i,d}^p$=1 represents that a CMS is installed (or activated) on node $i \in N_c$ that displays travel time information for destination $d \in D_i$ and route $p \in P_{i,d}$, and $\lambda_{i,d}^p$=0 otherwise. In practice, there are a maximum number of lines that can be displayed on CMS, denoted as *M*. Then we have:

$$\sum_{d,p} \lambda_{i,d}^p \leq M\rho_i, \forall i \in N_c. \tag{4.2}$$

Equation (4.2) requires that if CMS is not installed at location *i*, ρ_i=0, implying that $\lambda_{i,d}^p = 0, \forall d \in D_i, p \in P_{i,d}^p$. Otherwise, we have $\sum_{d,p} \lambda_{i,d}^p \leq M$ which satisfies the maximum number of lines constraints on each CMS. Notice that (2) allows the case when ρ_i=1 and $\lambda_{i,d}^p = 0, \forall d \in D_i, p \in P_{i,d}^p$. However, as will be shown in model (3) below, the objective function guarantees that this is not possible.

Hence, a given (ρ, λ) pair represents one possible CMS configuration if it also satisfies equations (4.1) and (4.2) above. The impact on the network can be obtained by solving the SN-SUE problem under the configuration. Clearly, for different pairs, the impact may be different and one may want to find the one with the most desirable performance. In this project, we consider both expected total system travel time and deployment cost of CMS. Denote $S(\rho, \lambda)$ the link flow solution of SN-SUE for a given (ρ, λ) pair, we may formulate the optimal CMS configuration problem as the following bilevel model:

$$\min_{\rho,\lambda} f(\rho,\lambda) = \theta \sum_{a \in A} [\bar{t}_a(x,\rho,\lambda) x_a] + \sum_{i \in N_c} \rho_i c_i \qquad (4.3\text{-a})$$

$$s.t.$$

$$\sum_{d,p} \lambda_{i,d}^p \le M\rho_i, \forall i \in N_c \qquad (4.3\text{-b})$$

$$\sum_i \rho_i \le K \qquad (4.3\text{-c})$$

$$x \in S(\rho,\lambda) \qquad (4.3\text{-d})$$

Here (4.3-a) is the objective function which is a weighted summation of expected total system travel time (the first term) and total deployment cost (the second term). The second term guarantees that if $\lambda_{i,d}^p = 0, \forall d \in D_i, p \in P_{i,d}^p$, we must have ρ_i=0. This implies that if ρ_i=0, then at least one of the $\lambda_{i,d}^p, \forall d \in D_i, p \in P_{i,d}^p$ must be 1. The weight θ denotes the "value-of-time" and c_i is the deployment cost at CMS node *i*. It may be the construction cost if a new CMS is to be installed or activation cost if a CMS already exists. (4.3-d) explicitly requires that the resulting flow *x* must satisfy SN-SUE, which is itself an optimization problem. Constraint (4.3-d) makes model (3) a bi-level problem.

4.4 Case Study I: A Hypothetical Network

We present an illustrative example in this section. The example shows how the model can be used to 1) evaluate the impacts of CMS travel times on drivers' route choice behaviors, and 2) how to determine the optimal CMS locations and displaying destinations/routes.

Figure 33 depicts the example network with 5 nodes, 5 links, and 1 OD pair from node 1 to 5. There are two routes between the origin and destination: 1->2->4->5 (i.e. links 1, 3, 5) and 1->2->3->4->5 (i.e. links 1, 2, 4, 5).

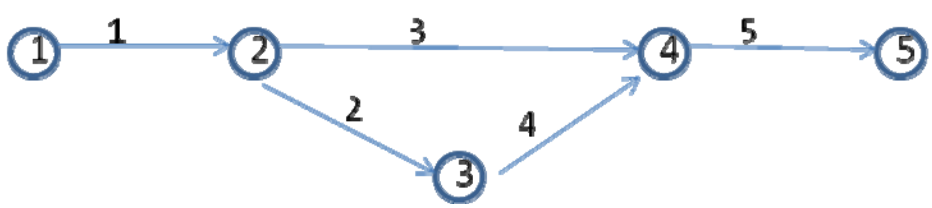

Figure 33: Layout of the Test Network

The travel time function of links is assumed to be the BPR function as follows:

$$\tau = a\left[1 + p\left(\frac{f}{C}\right)^b\right], \qquad (4.4)$$

where τ is the link travel time, a is the free flow travel time of the link, f is the link flow, C is the link capacity, and p and b is a coefficient. The values of these parameters (except flow f) for all links of the network are given in the table below.

Table 2: Parameters of the Example Network

Link(from.to)	*a*	*p*	*b*	Capacity
1.2	15	0.15	4	100
2.3	20	0.15	4	50
2.4	30	0.15	4	50
3.4	10	0.15	4	50
4.5	15	0.15	4	100

4.4.1 Impacts of CMS Travel Time on Drivers' Route Choice

It turns out that the impacts of CMS travel times on drivers' route choices depend on characteristics of the network (i.e. geometry of the network such as number of routes and lengths of the routes), traffic conditions (mean travel times, travel time variations), and characteristics of drivers (perceptions of travel times and risk-taking behaviors regarding travel time variations). The impacts of CMS travel times are a complex interaction of these factors. For the given example network, traffic conditions and driver characteristics will play the central role.

4.4.1.1 Deterministic Network

This is the most basic case, in which no stochasticity was assumed on link travel times or drivers' perception errors. That is, the network is deterministic. In this case, since the two routes have exactly the same length, the traffic should be evenly distributed over the two paths. The proposed model assigns 40 trips to each route, indicating that it works properly for the base case. Note that this is true for all three risk-taking behaviors.

4.4.1.2 Stochastic Network without CMS

This scenario tests whether the stochastic assignment component of the model works properly. First, we assume the perception error follows a normal distribution; the mean is sampled from a zero-mean normal distribution $N(0,\delta_1)$ and the variance is sampled from a Gamma distribution $G(\delta_2,\delta_3)$, where δ_1, δ_2, δ_3 are parameters. In particular, δ_2 and δ_3 are the shape and scale parameters of the Gamma distribution respectively. In this scenario, we use the same set of parameters for all links: $\delta_1 = 0.04$, $\delta_2 = 1$, $\delta_3 = 0.05$.

To illustrate how different risk-taking behaviors react to travel time distributions (i.e. risks), we assume the (actual) link travel time variation follows Gamma distribution $G(\mu_1,\mu_2)$, where μ_1,μ_2 are the shape and scale of the Gamma distribution respectively. All links except link 3 (2->4) follow the same Gamma distribution with $\mu_1 = 3.02, \mu_2 = 1$. For link 3, the parameter values are: $\mu_1 = 3.02, \mu_2 = 10$. Since the mean of a Gamma distribution is the product of its shape and scale, i.e. $\mu_1 \cdot \mu_2$ and the standard deviation is $\sqrt{\mu_1} \cdot \mu_2$, our setting implies that the distribution of travel time variation for link 3 has much larger mean and standard deviation. As a result, the sampled travel time variation should be generally larger for link 3 compared with other links. In other words, traveling on link 3 represents "more risks."

Column 2 of Table 3 depicts the flow on link 3 (2->4) of running the model using the above setting without activating CMS. The numeric values in the parentheses show the percentage of the total demand that was assigned to link 3. The three rows are for the three risk-taking behaviors (i.e. assuming drivers have same risk-taking behavior in each case). As we discussed previously, under no travel time variation and perfect perception (deterministic case), flow should be evenly distributed between the two routes, resulting in 40 on link 3. Under the stochastic assignment, the result slightly favors link 3 as indicated by the fact that for risk-neutral case, the flow assigned to link 3 is 43.88 which is slighter larger than 40. As risk-averse drivers generally dislike taking risks, fewer drivers will select link 3. Consequently, only 39.04 were assigned to link 3 for risk-averse drivers. For risk-prone drivers, nearly 60% of drivers chose link 3 as they prefer risks during their driving.

Table 3: Comparison of Risk-Taking and CMS Impacts

	Flow on Link 3 (W/O CMS)	Flow on Link 3 (With CMS)
Risk-Averse	39.04 (48.8%)	35.0 (43.8%)
Risk-Neutral	43.88 (54.9%)	40.7 (51.9%)
Risk-Prone	46.45 (58.1%)	45.0 (56%)

The above table shows that the model can properly capture the risk-taking behaviors of drivers during the assignment process.

4.4.1.3 Stochastic Network with CMS

To show the impacts of CMS, we assume a CMS is installed at node 2. Further assume the CMS displays travel times from node 2 to destination 5 for the route that contains links 3 and 5. As what we assumed in this study, due to the installation of the CMS and its information display, the perception error of drivers to links 3 and 5 will be reduced. In this scenario, we assume the reduction is 80%, i.e. the perception error of link 3 or 5 is 1/5 of its normal perception errors.

The results of the assignment are given in the third column of Table 3. We can see that in this case, only 43.8% of drivers chose link 3 if they are risk-averse. This intuitively makes sense as risk-averse drivers prefer less risk. Risk-prone drivers, on the other hand, prefer link 3 as they favor risks. As a result, 56% of risk-prone drivers chose link 3. For risk-neutral drivers, they are indifferent to risks and therefore they were almost evenly distributed between links 2 and 3 (about 51.9% to link 3). Notice that in this scenario, the results look more symmetric, i.e. with risk-neutral result being closer to 50%. This is because drivers' perception errors to the travel time variation of link 3 and 5 are reduced, or in other words, drivers are more certain that link 3 has more risks. This leads to more informed decisions to a driver based on his/her risk-taking behavior. Therefore, the distinction among the three risk-taking behaviors is more distinct.

4.4. 1. 4 Major Findings

As a summary, the impacts of CMS on route choice behaviors depend on 1) geometry of the roadway network, traffic conditions, and drivers' perceptions and risk-taking behaviors. For a given network and traffic condition, drivers' risk taking behavior will play a central role in the route choice process.

4.4.2 Optimal CMS Configuration

The optimal CMS configuration is a function of characteristics of the network, traffic conditions, characteristics of drivers, and CMS activation or construction cost. For the example network, we only focus on the optimal route to display on CMS.

In this case, we only assume a CMS will be installed at node 2, but will not specify to which route the travel time information will be displayed. Rather we will run an optimization routine to determine the "optimal" route. The optimization routine is based on simulated annealing, which is a heuristic-based optimization algorithm. For more details, one can refer to Kirkpatrick et al. (1983) and Friesz et al. (1992).

To determine the optimal CMS configuration, we need to consider the cost of installing a CMS, and the value of time. In this study, we assume the value of time factor $\theta = 2$, which may be interpreted as "the value of one minute is 2 dollars." We then set the CMS installation cost is 3. Notice that this value may be OK if we activate an existing CMS for displaying travel times, but may not be realistic for installing a new CMS. However, for illustrative purposes, we first set this value to be small. We will then change it and see how it may impact the optimal CMS configuration.

Table 4: Optimal CMS Configuration

	Optimal Route (link list)
Risk-Averse	2, 4, 5
Risk-Neutral	3, 5
Risk-Prone	None

Table 4 shows that for different risk-taking behaviors, the optimal routes to display travel times are different: for risk-averse drivers, showing travel times for the route comprising links 2,4,5 is most beneficial; for risk-neutral drivers, showing travel times for route (3,5) is most preferable. For risk-prone drivers, however, no travel time information should be displayed to obtain the best system performance. The risk-taking behaviors here mainly determine which route should be chosen to display travel times.

4.4.2.1 Impacts of CMS Construction Cost

Besides the drivers' risk-taking behaviors, the cost for constructing or activating CMS plays a critical role. To see this, we increase the cost of CMS from 3 to 30. In this case, the model suggests no route should be selected to display travel times, i.e. the CMS should not be installed or activated. This clearly shows that the CMS construction cost determines whether a CMS should be installed or not.

4.4.2.2 Major Findings

In summary, the optimal CMS configurations are determined by roadway geometry, traffic conditions, driver characteristics, and CMS installation/activation cost. In particularly, CMS construction cost determines whether a CMS should be installed/activated and drivers' risk-taking behaviors determine which optimal route should be selected to display travel times.

4.5 Case Study II: SF Bay Area Freeway Sub-Network

We further test the SN-SUE model and SA solution algorithm on a freeway sub-network in the San Francisco Bay Area. A sketch of the network is shown in Figure 34. The sub-network was produced via a sub-area analysis from a regional demand model for the entire Bay Area. The sub-area analysis resulted in a total of 3213 origin-destination (OD) pairs for the AM peak, of which about 20% have more than 100 trips. We therefore combine OD zones that are close to each other, resulting in a total of 81 OD pairs all of which have significant amount of trips. Figure 34 depicts how the original OD zones are grouped. The total number of trips for this sub-network after the adjustment is 261,194.

A CMS is currently installed at the location indicated in Figure 34. Here we focus on trips going SB from East Bay (*Zone 1*) to the other zones (*Zone 2- Zone 9*). The purpose is to test the impact of CMS on traffic distribution within the network and how the optimal CMS configuration can be determined. It can be seen that there are two routes from Zone 1 to the Zone 5, denoted as *R1* and *R2* respectively. The travel time of R1 is roughly 40 minutes, longer than that for R2 (about 30 minutes).

Similar to the analysis for Case Study I, we test on four scenarios. Scenario A assumes no travel time information is displayed on CMS, which provides a base case to test the SN-SUE model; Scenario B assumes the CMS displays travel time information for R1. Both the first two scenarios assume the same travel time variation parameters and perception parameters for every link. In the third scenario, we change the parameters for links along R1 to increase travel time variability. In the fourth scenario, we test how to determine the optimal CMS configuration, i.e. travel time information for which route should be displayed.

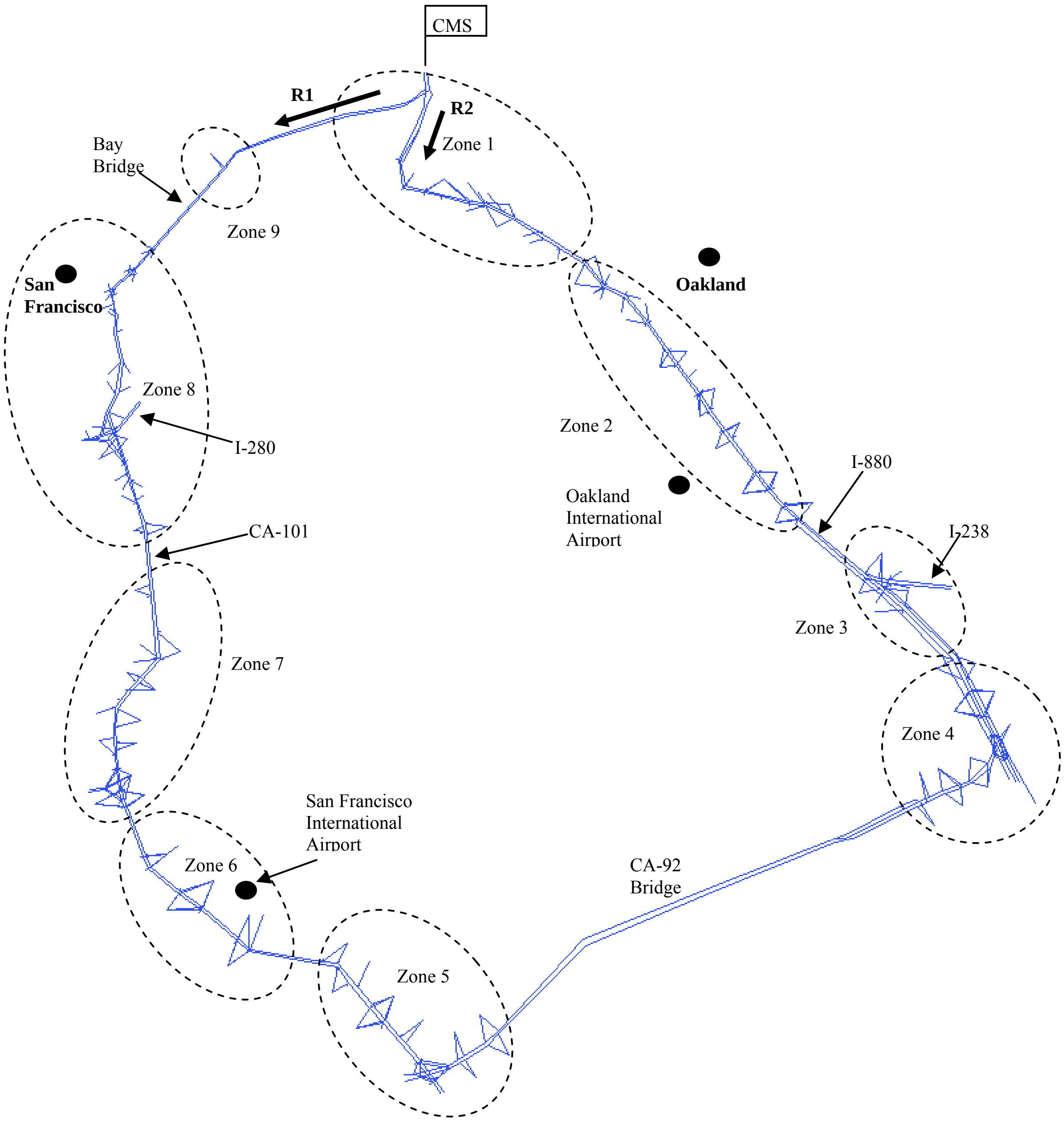

Figure 34: The San Francisco Freeway Sub-Network

4.5.1 Scenario A: No Travel Time Information Displayed on CMS

This scenario provides a baseline of the network traffic flow distribution according to the SN-SUE assumption. That is, we assume the perception error follows a normal distribution; the mean is sampled from a zero-mean normal distribution $N(0,\delta_1)$ and the variance is sampled from a Gamma distribution $G(\delta_2,\delta_3)$, where δ_1, δ_2, δ_3 are parameters. In particular, δ_2 and δ_3 are the shape and scale parameters of the Gamma distribution respectively. We also assume the (actual) link travel time variation follows Gamma distribution $G(\mu_1, \mu_2)$, where μ_1, μ_2 are the shape and scale of the Gamma distribution respectively. Similar to Case Study I, we use the same set of parameters for all links: $\delta_1 = 0.04$, $\delta_2 = 1$, $\delta_3 = 0.05$, and $\mu_1 = 3.02, \mu_2 = 1$.

Table 5 depicts the results of the SN-SUE assignment. It shows that about 45.5% of travelers going from East Bay (Zone 1) SB to all other potential locations (Zone 2 – Zone 9) will take R1. We can also see that there is almost no difference among different risk-taking behaviors (risk-averse, risk-neutral, or risk-prone) because the randomness for both perception errors and link travel time variations are the same for all links.

Table 5: Distribution of Traffic with/without CMS Travel Time Information

		Proportion of Traffic on R1 (%)	*Proportion of Traffic on R2 (%)*
Scenario A	*Averse*	45.46%	54.54%
	Neutral	45.47%	54.53%
	Prone	45.49%	54.51%
Scenario B	*Averse*	45.35%	54.65%
	Neutral	45.36%	54.64%
	Prone	45.27%	54.73%
Scenario C	*Averse*	39.54%	60.46%
	Neutral	45.73%	54.27%
	Prone	50.93%	49.07%

To see the convergence of the SN-SUE algorithm, we define a Gap function as follows:

$$G = | x^{k+1} - x^k |_2 . \tag{4.5}$$

Here x^k and x^{k+1} denote the link flow vector at the *k*-th and (*k*+1)-th iterations respectively. The gap function *G* is thus defined as the 2-norm of the difference of the link flow vectors for two consecutive iterations during the solution process. Figure 35 depicts the change of *G* vs. the number of iterations. We can see that the algorithm converges quickly after about 30 iterations.

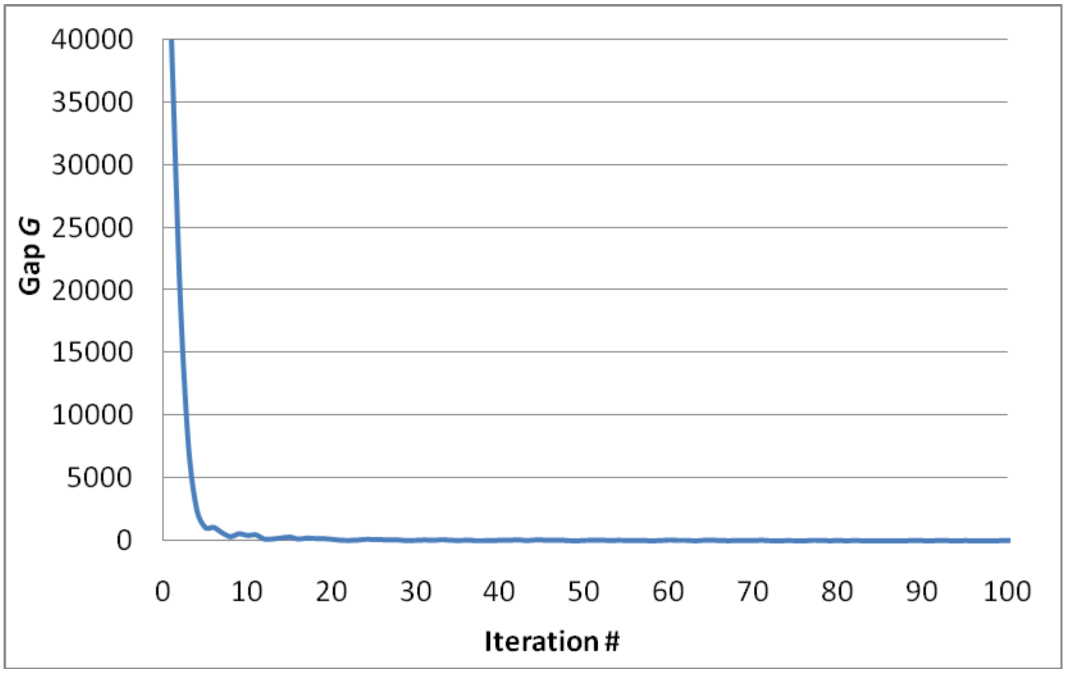

Figure 35: Convergence of the SN-SUE Algorithm

4.5.2 Scenario B: Travel Time Displayed on CMS for R1

In this scenario, we display travel times on the CMS for R1, but keep all the parameters unchanged for travel time variations and perception errors. The results in Table 5 show that the proportion of traffic via R1 goes down slightly, possibly because travelers have a better impression now that R1 is longer (especially for those going from Zone 1 to Zone 5) and therefore they switch to R2 instead even if they took R1 before. However, the difference is very minor, indicating that the perception errors do not play a major role in this particilar problem. We can also notice that the difference among different risk behaviors is negelible because all likes follow the same travel time distribution.

4.5.3 Scenario C: Increase of Travel Time Variation on R1

In Scenario C, we increase the parameters for link travel time variability of all links in R1 to $\mu_1 = 3.02, \mu_2 = 10$. Since the mean of a Gamma distribution is the product of its shape and scale, i.e. $\mu_1 \cdot \mu_2$ and the standard deviation is $\sqrt{\mu_1} \cdot \mu_2$, our setting implies that the distribution of travel time variation for R1 links has much larger mean and standard deviation. As a result, the sampled travel time variation should be generally larger for links on R1 compared with other links. Notice that the mean travel times of these links will remain the same as before. In other words, traveling on R1 now represents "more risks" to travelers.

As risk-averse travelers prefer less risk or smaller travel time variations, we expect that fewer will select R1. This is reflected in Table 5 in the "Scenario C" column, where only less than 40% travelers selected R1, about 5.5% reduction. The risk-prone travelers on the other hand prefers more risks, they tend to favor R1 more as they view travel time variation as an opportunity to get smaller (experienced) travel times. As a result, nearly 51% travelers selected R1 representing a 6.5% increase compared with the base case (Scenario A). The risk-neutral travelers however are insensitive to this as they mainly care about the mean travel time which is not changed. Therefore, the change of the proportion for R1 is negligible if all travelers are risk-neutral.

4.5.4 Scenario D: Optimal CMS Route to Display

In this scenario, we run the SA algorithm to find the optimal solution. The search space for this particular problem is small as it has only two possibilities: displaying travel times for R1 or R2. If we set the initial value as R2, Figure 36 shows the performance of the SA algorithm for 10 iterations. As indicated in the figure, the algorithm picks R1 at the third iteration and the remaining iterations will be just evaluating the objective values with travel time displayed for R1. The objective values fluncutate slightly due to the stochasticity involved in the travel times, which however stays fairly stable and significantly lower than the objective value if R2 is selected. Other parameters that are used to run the SA algorithm include: 1) travelers are assumed to be risk-averse, 2) costruction cost is zero (e.g. merely to enable the display of travel time messages on an exisiting CMS instead of installing a new CMS), and 3) the value of time is "$1 per minute."

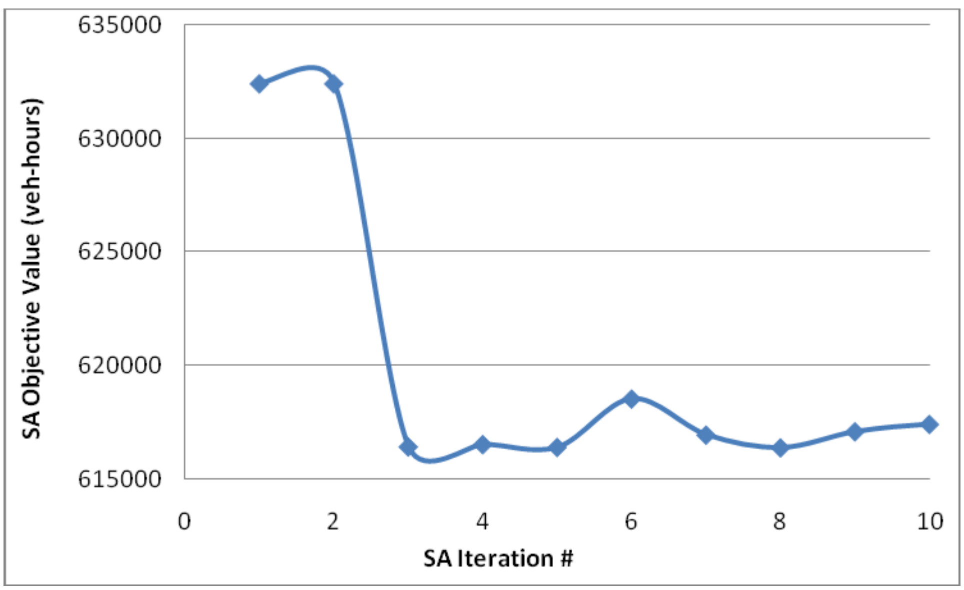

Figure 36: Performance of SA Solution Algorithm

4.6 Discussion and Recommendations

As shown via the two case studies in Section 4.4 and 4.5, the following factors determine how CMS travel times influence travelers' route choice behaviors:

- geometry of the roadway network,
- traffic conditions especially travel time variability,
- travelers' risk-taking behaviors, and
- travelers' perceptions over the actual travel times.

For a given network and traffic condition, drivers' risk taking behavior will play an important role in the route choice process.

The optimal CMS configurations are determined by

- roadway geometry,
- traffic conditions especially travel time variability,
- travelers' risk-taking behaviors,
- travelers' perceptions over the actual travel times, and
- CMS installation/activation cost.

In particular, CMS construction cost determines whether a CMS should be installed/activated and drivers' risk-taking behaviors determine which optimal route should be selected to display travel times.

Based on these findings, we can make the following recommendations:

1. From the two case studies, displaying travel times is generally beneficial from a system perspective if the objective is to reduce the weighted summation of the expected total system travel time and the CMS installation/activation costs, unless the CMS installation/activation costs are too high (in this case the SA algorithm returns a solution which prefers no display).
2. Exactly which CMS should be activated to display travel times (or whether new CMS needs to be installed for this purpose) and which destinations/routes to display on a CMS need to be determined from a system point of view based on the factors identified above.
3. To fully explore the advantages of displaying travel times on CMS, thorough investigations on travel time variability of the study area and the perception of travelers are highly recommended. This will result in the needed parameters μ_1 and μ_2 to quantify the variability, and δ_1, δ_2, δ_3 to quantify travelers' perceptions.
4. Displaying travel times may bring other intangible benefits which may not be easily quantified. Some of those factors are captured by the associated survey study of this project such as public's acceptance, plan in advance, and reduction of stress, etc. When determining whether travel times should be displayed on a CMS, these factors need to be considered as well.
5. We studied the optimal CMS configuration problem for displaying travel times on freeway CMS. The study is based on a SN-SUE model to capture how commuters make route choice decisions which considers both travel time variability and

travelers' perceptions errors. To solve the optimal CMS configuration problem, we developed a heuristic method based on simulated annealing (SA). The model and solution method are tested on a hypothetical network and a real world network in the San Francisco Bay Area.

We recommend the following future studies to continue the research in this line:

- Expand the study area to a larger and more realistic network.
- Expand the model and algorithm to capture traffic dynamics.
- Investigate the sensitivity of the results to travel time variability and perception parameters, and to the "value of time" parameter.

CHAPTER 5

GUIDELINES FOR DISPLAYING TRAVEL TIMES ON CMS AND CONCLUDING REMARKS

In the literature, numerous studies have been conducted on CMS by both practitioners and the research community. These studies cover a broad range of issues related to CMS deployment, including CMS contents and design, operational and policy issues, impacts of CMS on drivers' behaviors, and impacts of CMS on network-wide performance. Issues such as CMS contents and design have been well documented in existing guidelines (NCDOT, 1999; FHWA, 2003; FHWA, 2004; Benz 1998).

In this project, we focused on a few key issues: the evaluation of travel time estimation methods, user perception and preference, the impacts of CMS on drivers' route choice behaviors, and the impact of CMS configurations (such as locations and destinations/routes to display) on the network-wide performances. Based on the findings from this research, the following guidelines are developed for deploying and displaying travel times on CMS:

Travel time estimation methods:

- During off-peak, the differences using different estimation algorithms are not significant, and the instantaneous travel time can be adopted for its simplicity.
- During peak hours, if the route travel time is relatively short and the transition from free-flow to maximum congestion is slow, the instantaneous travel time still can be adopted for its simplicity.
- During peak hours, when the route travel time is long (e.g. from Emeryville to SFO), linear regression travel times (generated using both real-time data and historical data) should be considered.
- Using speed data from different lanes makes significant difference in the accuracy of the travel time estimates. If patterns similar to those shown in this study are found, lane-by-lane loop data can be used to improve travel time estimation.
- When relying on loop detectors alone, loop spacing should not be too large. Use multiple data source when loop coverage is poor.

Display travel times in the presence of alternative routes:

Benefits besides route choice, such as being able to plan ahead and having peace of mind, are as important as route choice to users when they consider the usefulness of travel time display on CMS. When determining whether travel times should be displayed on a CMS, all these benefits needs to be considered. Even if there is no route choice, it is beneficial to have travel time display on CMS if there is great variability and uncertainty in traffic condition on a freeway segment.

Frequency of travel time displays:

- Estimated travel time should be displayed as an exact number of minutes, not as a range.
- Information should be updated every 2 minutes. More frequent updates may confuse drivers and reduce their confidence in the accuracy travel time display.

CMS location for displaying travel times:

- Exactly which CMS should be activated to display travel times (or whether new CMS needs to be installed for this purpose) and which destinations/routes to display on a CMS need to be determined based on the following factors: roadway geometry, traffic conditions especially travel time variability, travelers' perceptions over the actual travel times, and CMS installation/activation cost.

In addition, we have concluded that:

The San Francisco Bay Area achieved reasonably good coverage of travel time display on CMS; Displayed travel times are relevant to most commuters' commute route, and among those who consider the messages not relevant to their commute routes, we see that a much larger percentage travel less frequently and a much larger percentage travel a short distance.

The majority of surveyed commuters believe the estimates are accurate within 5 minutes, and an overwhelming majority regards the travel time display as useful. Perception of accuracy of estimated travel time greatly affects the perceived usefulness of travel time display. However, even when the perceived accuracy is within 10-15 minutes, the vast majority (over 80%) of the commuters who hold the accuracy perception still regards the travel time display on CMS as useful;

Benefits besides route choice, such as being able to plan ahead and having peace of mind, are as important as route choice to users when they consider the usefulness of travel time display on CMS. Users need not to make a choice or decision to benefit from CMS travel time display. Even if there is no route choice, it is beneficial to have travel time display on CMS if there is great variability and uncertainty in traffic condition on a freeway segment.

For most people, route diversion is a real possibility. However, a large threshold for the difference between expected route travel times needs to be reached before commuters switch from their usual routes. The effectiveness of CMS alone in persuading drivers to divert is likely to be small.

Commuters use many sources for travel information. Travel time display on CMS requires no user input or interaction but offers great value compared to other services that require input and thus are used infrequently, by few, and unsafely.

REFERENCES

1. Abbas, M.M., and McCoy, P. T. (1999) Optimizing variable message sign locations on freeways using genetic algorithms. In *Proceedings of the 78th Transportation Research Board Annual Meeting (CD-ROM)*, Washington, DC.
2. Al-Deek, H., and Emam, E. (2006). A new methodology for estimating transportation network reliability. *Journal of Intelligent Transportation System*, 10(3), 117-129.
3. Benz, R.J., W.L. Eisele, D.J. Holdener and S.M. Turner. (1998), Travel Time Data Collection Handbook, FHWA-PL-98-035, March 1998.
4. Caplice, C., and Mahmassani, H.S. (1992) Aspects of commuting behavior: preferred arrival time, use of information and switching propensity. *Transportation Research Part A* 26, 409-418.
5. Chang, G.L. (1985) *Departure Time Dynamics in Transportation Networks*, Ph.D. dissertation, University of Texas at Austin.
6. Chatterjee, K., Hounsell, N.B., Firmin, P.E., and Bonsall, P.W., (2002), Driver Response to Variable Message Sign Information in London, *Transportation Research C* 10 (2002) 149-169
7. Chavan, P., Tufte, K., Monsere, C.M., Stephens, M., and Bertini, R. (2008) Extending the use of archived ITS data as a potential management tool to evaluate traveler information on Dynamic Message Signs. *Proceedings of the 87th Transportation Research Board Annual Meeting (CD-ROM)*, Washington, DC.
8. Chen, A., Ji, Z., and Recker, W. (2002). Travel time reliability with risk-sensitive travelers. *In Proceedings of the 81st Transportation Research Board Annual Meeting (CD-ROM)*, Washington, DC.
9. Chen, A., and Recker, W. (2001). Considering risk taking behavior in travel time reliability. In *Proceedings of the 80th Transportation Research Board Annual Meeting (CD-ROM)*, Washington, DC.
10. Chen, A., Yang, H., Lo, H.K., and Tang, W. (1999) A capacity related reliability for transportation networks. *Journal of Advanced Transportation*, 33, 183-200.
11. Chen, C., Skabardonis, A., and Varaiya, P. (2003). Travel time reliability as a measure of service. *Journal of Transportation Research Board*, 1855, 74 – 79.
12. Chen, C., Skabardonis, A., and Varaiya, P. (2004) A system for displaying travel times on changeable message signs. In *Proceedings the 83rd Transportation Research Board Annual Meeting (CD-ROM),* Washington, DC.
13. Chen, R. and Mahmassani, H.S., (2004), Travel time perception and learning mechanisms in traffic networks, *Transportation Research Record*, No. 1894, pp. 209-221
14. Chiu, Y.C., and huynh, N. (2007) Location configuration design for dynamic message signs under stochastic incident and ATIS scenarios. *Transportation Research Part C* 15, 33 – 50.
15. Chorus, C., et al. (2005), Value of Travel Information, *Transportation Research Record*, No. 1926, pp. 142-151
16. Daganzo, C.F. (1983). Stochastic network equilibrium with multiple vehicle types and asymmetric, indefinite link cost Jacobians. *Transportation Science* 17 (3), 282 – 300.

17. Daganzo, C.F. (1997). *Fundamentals of Transportation and Traffic Operations*, Pergamon-Elsevier, Oxford, U.K.
18. Davis, D., 1989, Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology, *MIS Quaterly*, September 1989
19. Diakaki, C., and Papageorgiou, M. (1997) Simulation studies of integrated corridor control in Glasgow. *Transportation Research Part C* 5(3/4), 211-224.
20. Dudek, C.L. (1979) Changeable Message Signs—Synthesis of Highway Practice 61. National Cooperative Highway Research Program.
21. Emmerink, R.H.M., Nijkamp, P., Rietveld, P., and Van Ommeren, J.N. (1996) Variable message signs and radio traffic information: an integrated empirical analysis of drivers' route choice behavior. *Transportation Research Part A* 30(2), 135-153.
22. EU (2006), Travel Time Perception on Variable Message Signs, available at http://www.easyway-its.eu/1/index.php?option=com_docman&task=doc_download&gid=488&Itemid=103
23. FHWA (2003) Manual on Uniform Traffic Control Devices. U.S. Department of Transportation, Washington D.C.
24. FHWA (2004) Changeable Message Signs Operation and Messaging Handbook, Available at
25. http://tmcpfs.ops.fhwa.dot.gov/cfprojects/uploaded_files/CMS%20Operation%20and%20Messaging%20Handbook-Final%20Draft.pdf.
26. Friesz, T.L., Cho, H.J., Mehta, N.J., Tobin, R.L. and Anandalingham, G., (1992). Simulated annealing methods for network design problems with variational inequality constraints, *Transportation Science*, 26(1), 18-26.
27. Fujito, I., Margiotta, R., Huang, W., and Perez, W.A. (2006). The Effect of Sensor Spacing on Performance Measure Calculations. *In Proceedings of the 85th Annual Meeting of Transportation Research Board (CD-ROM)*, Washington, DC.
28. Gan, H.C., Ye, X., and Gao, W.S. (2008) Drivers' en-route diversion decisions under the influence of Variable Message Sign information: A empirical analysis. In *Proceedings of the 87th Transportation Research Board Annual Meeting (CD-ROM)*, Washington, DC.
29. Hartley, J.K. (2003) Prediction of link travel times in the context of Nottingham's urban road network. In Proceedings of the 17th European Simulation Multi-Conference, 423-428.
30. Hinsbergen, C.P.IJ, Lint, J.W.C., and Sanders, F.M. (2007). Short-term prediction models. In *Proceedings of the 14th world congress on ITS (CD-ROM)*.
31. Hoaglin, D.C., Mosteller, F., and Tukey, J.W. (1983). *Understanding Robust and Exploratory Data Analysis*. John Wiley & Sons.
32. Huisken, G., and M.v. Maarseveen (2000). Congestion prediction on motorways: a comparative analysis. In *Proceedings of the 7th world congress on ITS (CD-ROM)*.
33. Khattak, A., Kanafani, A., and Colletter, E.L. (1994) Stated and reported route diversion behavior: implications of benefits of advanced traveler information system. *Journal of Transportation Research Board* 1464, 28-35.
34. Iida, Y., Uno, N., and Yamada, T. (1994) Experimental analysis approach to analyze dynamic route choice behavior of driver with travel time information. In *Proceedings of Vehicle Navigation and Information Systems*, 377-382, Yokohama, Japan.

35. Jackson, P.G. (1994), Behavioural Responses to Dynamic Route Guidance (DRG) Systems, In: *Paper presented at the PICT International Doctoral Conference*, 28th-30th March 1994. Available at http://web.tongji.edu.cn/~yangdy/its6.htm

36. Katsikopoulos, K.V., Duse-Anthony, Y., Fisher, D.L., and Duffy, S.A. (2000) The framing of drivers' route choices when travel time information is provided under varying degrees of cognitive load. *Human Factors* 42(3), 470-481.

37. Khattak, A., Kanafani, A., and Colletter, E.L. (1994) Stated and reported route diversion behavior: implications of benefits of advanced traveler information system. *Transportation Research Record* 1464, 28-35.

38. Kirkpatrick, S., Gelatt, C.D., and Vecchi, M.P. (1983) Optimization by simulated annealing. Science 220, 671-680.

39. Koenker, R. and Bassett, G. (1978) Regression quantiles. Econometrica 46(1), 33-50.

40. Koutsopoulos, H.N., and Lotan, T. (1990) Motorist information systems and recurrent traffic congestion: Sensitivity analysis of expected results. *Transportation Research Record* 1281, 145-158.

41. Kraan, M., Zijpp, N.V.D., Tutert, B., and et al. (2000) Evaluating networkwide effects of variable message signs in the Netherlands. *Transportation Research Record* 1689, 60-67.

42. Kwon, J., McCullough, B., Petty, K. and Varaiya, P. (2006) Evaluation of PeMS to Improve the Congestion Monitoring Program. Final Report for PATH TO 5319.

43. Lam, W.H.K., and Chan, K.S. (1994) A stochastic traffic assignment model for road network with travel time information via variable message signs. In *Proceedings of Vehicle Navigation and Information Systems*, 99-104, Yokohama, Japan.

44. Lee, C., Ran, B., Yang, F., and Shaw, J. (2008) Analysis of Alternate Route Choice Behavior with Variable Message Signs Using Hybrid Tree Method. Presented at the *87th Annual Meeting of Transportation Research Board*.

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

46. Liu, X.H., He, X.Z., Recker, W. (2007) Estimation of the time-dependency of values of travel time and its reliability from loop detector data. *Transportation Research, Part B*, 41(4), 448-461.

47. Lo, H.K., Chen, A., and Yang, H. (1999) System time minimization in route guidance with elastic market penetration. In *Proceedings of the 78tg Transportation Research Board Annual Meeting (CD-ROM)*, Washington, DC.

48. Maher, M.J., and Hughes, P.C. (1995) Estimation of the potential benefits from an ATT system using a multiple user class stochastic user equilibrium assignment model. In *Applications of Advanced Technologies in Transportation Engineering*, Proceedings of the 4th International Conference, eds Y.J. Stephanedes and F. Filippi, Capri, Italy, 700-704.

49. Mahmassani, HS, Caplice, CG, and Walton, CM (1990) Characteristics of urban commuter behavior: switching propensity and use of information. *Transportation Research Record* 1285, 57-69.

50. Mahmassani, H.S., and Chang, G-L (1987) On boundedly-rational user equilibrium in transportation systems," *Transportation Science* 21(2), 89-99.

51. Mahmassani, H.S., and Chang, G-L (1985) Dynamic aspects of departure time choice behavior in a commuting system: Theoretical framework and experimental analysis. *Transportation Research Record* 1037, 88-101.

52. Mahmassani, H.S. (2001) Dynamic network traffic assignment and simulation methodology for advanced system management applications. *Networks and Spatial Economics*, 267-292.

53. Mannering, F., Kim, S.G., Barfield, W., and Ng, L. (1994) Statistical analysis of commuters' route, mode, and departure time flexibility. *Transportation Research Part C* 2, 35-47.

54. Mannering, F., Kim, S., NG, L., and Barfield, W., (1995), Travelers' Preferences for in-Vehicle Information Systems: an Exploratory Analysis, *Transportation Research C*, Vol 3 (6), 339-351.

55. Margulici, J.D., and Ban, X. (2008a) Benchmarking travel time estimates. *IET Journal of Intelligent Transportation Systems* 2, 228-237.

56. Margulici, J.D., and Ban, X. (2008b) Developing quality measures for evaluating travel time estimation methods. In *proceedings of the 15th ITS World Congress (CD-ROM)*.

57. Mirchandani, P., and Soroush, H. (1987). Generalized traffic equilibrium with probabilistic travel times and perceptions. *Transportation Science 21* (3), 133-152.

58. North Carolina Department of Transportation (1999).Operational Guidelines for the Use of Changeable Message Signs. *http://www.ncdot.org/doh/preconstruct/traffic/congestion/docs/cmsopera.pdf.*

59. Noland, R.B., and Polak, J.W. (2002) Travel time variability: a review of theoretical and empirical issues. *Transportation Review* 122 (1), 39-54.

60. Oda T. (1990). An algorithm for prediction of travel time using vehicle sensor data. *In Proceedings of Third International Conference on Road Traffic Control*, 40 - 44.

61. Pan, X.H., and Khattak, A. (2008) Evaluating traveler information impacts on commercial and non-commercial users. In *Proceedings of the 87th Transportation Research Board Annual Meeting (CD-ROM)*, Washington, DC.

62. PBSJ (2004), AMBER, Emergency, and Travel Time Messaging Guidance for Transportation Agencies, Final Report Prepared for FHWA Office of Operations, available at http://www.ops.fhwa.dot.gov/travelinfo/resources/cms_rept/cmspractices.htm#_ftn3

63. Peeta, S. and Ramos, J.L. (2006). Driver response to variable message signs-based traffic information, In *IEEE Proceedings of Intelligent Transportation Systems* 153 (1), 2-10.

64. Peeta, S., Ramos, J.L., and Pasupathy, R. (2000) Content of variable message signs and on-line driver behavior. *Transportation Research Record* 1725, 102-108.

65. Peng, Z.R., Guequierre, N., and Blakeman, J.C. (2004) Motorist response to arterial variable message signs. *Journal of Transportation Research Board*, 1899, 55-63.

66. Rice, J., and Zwet, E.V. (2001) A simple and effective method for predicting travel times on freeways. In *Proceedings of IEEE Intelligent Transportation Systems*, 227-232, Oakland, CA.

67. RITA (2006), Intelligent Transportation Systems Deployment Statistics, available at http://www.itsdeployment.its.dot.gov/Trends2006.asp?comp=FM

68. Small, K.A., Winston, C., and Yan, J. (2005) Uncovering the distribution of motorists references for travel time and reliability. *Econometrica* 73(4), 1367-1382.

69. Smith, B.L., and M.J. Demetsky (1997). Traffic flow forecasting: comparison of modeling approaches. *Journal of Transportation Engineering*, 123(4), 261-266.

70. Tatineni, M. (1996). *Solution Properties of Stochastic Route Choice Models*. Ph.D. dissertation, University of Illinois at Chicago.

71. Tatineni, M., Boyce, D.E., and Mirchandani, P. (1996) Comparisons of deterministic and stochastic traffic loading models. *Transportation Research Record* 1067, 16-23.

72. Uchida, T., Iida, Y., and Nakahara, M. (1994) Panel survey on drivers' route choice behavior under travel time information. In *Proceedings of Vehicle Navigation and Information Systems*, 383-388, Yokohama, Japan.

73. Ullman, G.L., Dudek, C.L., and Blake K.N. (1994) Effect of freeway corridor attributes on motorist diversion responses to travel time information. *Transportation Research Record* 1464, 19-27.

74. Van Lint, J.W.C, Tu, H., and Van Zuylen, H.J. (2004). Travel time reliability on freeways, In *Proceedings of 10th World Conference on Transport Research (CD-ROM)*, Istanbul, Turkey.

75. Van Vuren, T., and Watling, D. (1991) Multiple user class assignment model for route guidance. *Transportation Research Record* 1306, 22-31.

76. Vaughn, K.M., Reddy, P., Abdel-Aty, M.A., Kitamura, R., and Jovanis, P.P. (1995). Route choice and information use: initial results from simulation experiments. *Transportation Research Record* 1516, 61-69.

77. Wang, J.H., Yang, C.M., Collyer, C.E., and Waters, D. (2006) Effects of message display on motorist response to dynamic message signs. In *Proceedings of 85th Annual Meetings of Transportation Research Board (CD-ROM)*, Washington, DC.

78. Warita, H., Okada, T., and Tanaka, A. (2001) Evaluation of operation for travel time information on the metropolitan expressway. In *Proceedings of 8th ITS World Congess*.

79. Wardman, M., Bonsall, P.W., and Shires, J.D., (1997), Driver Response to Variable Message Signs: a Stated Preference Investigation, *Transportation Research C*, Vol. 5, No 6, 389-405

80. Yang, H. (1998) Multiple equilibrium behaviors and advanced traveler information systems with endogenous market penetration. *Transportation Research Part B* 32(3), 205-218.

81. Yin, Y., Lam, W.H.K., and Ieda, H. (2002) Modeling risk-taking behavior in queuing networks with advanced traveler information systems, *Proceedings of 15th International Symposium on Transportation and Traffic Theory*, (M.A.P. Taylor eds.), Pergamon, 309-328.

82. Zhang, X.Y., Rice, J., and Bickel, P. (1999) *Empirical Comparison of Travel Time Estimation Methods*. California Path Research Report, UCB-ITS-PRR-99-43

83. Zuurbier, F.S., Zuylen, H.J., Hoogendoorn, S.P., and Chen, Y. (2006) A generic approach to generating optimal controlled prescriptive route guidance in realistic traffic networks. *Proceedings of 85th Annual Meetings of Transportation Research Board (CD-ROM)*.