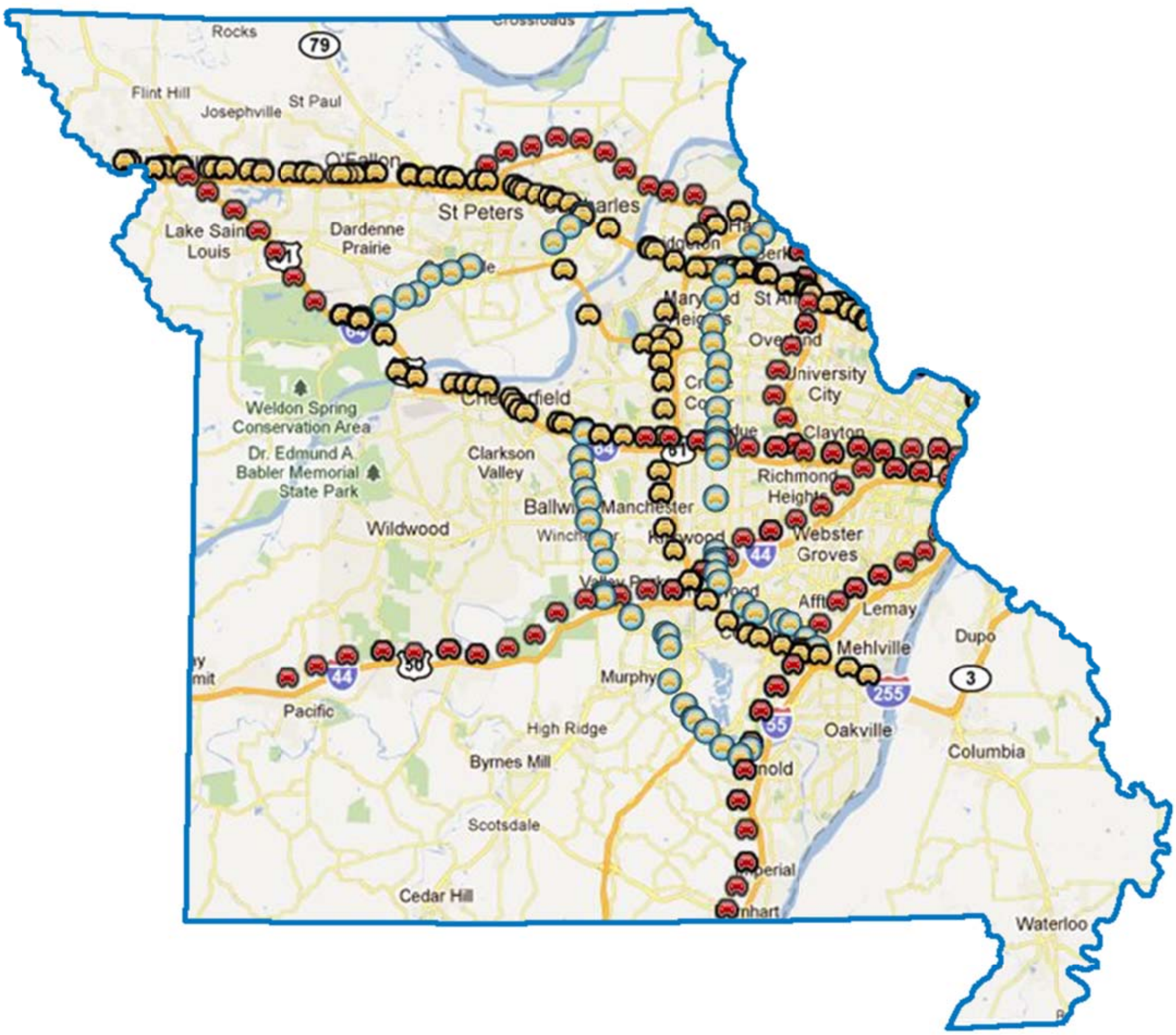
Freeway Travel Time Estimation Using Existing Fixed Traffic Sensors – Phase I



Prepared By

Yao-Jan Wu, Ph.D., P.E.
Shu Yang, Graduate Research Assistant
Saber Abdoli, Graduate Research Assistant
Saint Louis University Department of Civil Engineering

Final Report Prepared for Missouri Department of Transportation
2014 February   Project TR201311   Report cmr14-008

Final Report

Project Number TRyy1311

Freeway Travel Time Estimation using Existing Fixed Traffic Sensors – Phase 1

Prepared for
Missouri Department of Transportation
Traffic Management Center

By

Yao-Jan Wu
Assistant Professor, PhD, P.E.

Shu Yang
Graduate Research Assistant

Saber Abdoli
Graduate Research Assistant

Department of Civil Engineering
Parks College of Engineering, Aviation and Technology
Saint Louis University

August 2013

TECHNICAL REPORT STANDARD TITLE PAGE

1. REPORT NO.	2. GOVERNMENT ACCESSION NO.	3. RECIPIENT'S CATALOG NO.
cmr14-008		

4. TITLE AND SUBTITLE	5. REPORT DATE
Freeway Travel Time Estimation Using Existing Fixed Traffic Sensors – Phase 1	August 2013
	6. PERFORMING ORGANIZATION CODE
7. AUTHOR(S) Yao-Jan Wu, Shu Yang, and Saber Abdoli	8. PERFORMING ORGANIZATION REPORT NO.
9. PERFORMING ORGANIZATION NAME AND ADDRESS Smart Transportation Lab, Department of Civil Engineering Parks College of Engineering, Aviation and Technology Saint Louis University, 3450 Lindell Boulevard St. Louis, MO 603103	10. WORK UNIT NO. 11. CONTRACT OR GRANT NO. TR201311
12. SPONSORING AGENCY NAME AND ADDRESS Missouri Department of Transportation Research, Development and Technology P.O. Box 270 Jefferson City, MO 65102	13. TYPE OF REPORT AND PERIOD COVERED Final Report 14. SPONSORING AGENCY CODE

15. SUPPLEMENTARY NOTES

16. ABSTRACT

Freeway travel time is one of the most useful pieces of information for road users and an important measure of effectiveness (MOE) for traffic engineers and policy makers. In the Greater St. Louis area, Gateway Guide, the St. Louis Transportation Management Center (TMC) operated by the Missouri Department of Transportation (MoDOT), already collects traffic data from more than 700 fixed traffic sensors located along major freeways and arterials. Due to their significant investment in fixed freeway sensors, MoDOT has been striving to more fully utilize this data to extract useful information for stakeholders. One potential application is to convert the information collected by the fixed traffic sensors (point detectors) directly into travel times without the need to install additional sensors. This project developed a database specifically designed and optimized for the purpose of freeway performance measurement and implemented a data quality assurance procedure to examine the quality of the data collected. After a comprehensive literature review, the instantaneous travel time estimation model was selected and implemented in a stand-alone MATLAB-based system and a Microsoft Excel VBA-based tool. Several case studies were conducted to test the applicability of the system. The estimated travel time results were then verified using travel time data collected by a Bluetooth-based method and a video-based manual-vehicle-matching method. This Phase 1 project has proved the feasibility of the proposed approach. The travel time estimation method will be improved and the coverage of travel time estimation expanded to the city's entire freeway network in Phase 2.

17. KEY WORDS	18. DISTRIBUTION STATEMENT
Travel time estimation, traffic sensors, data collection, data quality control and database design.	No restrictions. This document is available to the public through the National Technical Information Service, Springfield, VA 22616

19. SECURITY CLASSIF. (of this report)	20. SECURITY CLASSIF. (of this page)	21. NO. OF PAGES	22. PRICE
None	None		

Table of Contents

Executive Summary 5

Section 1 Introduction 7

1.1 Project Background 7

1.2 Problem Statement 7

1.3 Research Objectives 8

Section 2 Literature Review 9

2.1 Travel Time Estimation 9

2.2 Ground Truth Data Collection 10

2.2.1 Conventional Techniques 10

2.2.2 Bluetooth-based Travel Time 11

Section 3. Data Collection 13

3.1 Traffic Data 13

3.1.1 Traffic Sensors 13

3.1.2 Archiving Data at Saint Louis University (SLU) 15

3.2 Bluetooth-based Travel Time Data 16

Section 4. Traffic Database Design and Implementation 18

4.1 Database Design 18

4.2 Database Optimization 18

4.3 Traffic Data Quality Assurance 19

4.3.1 Real-time Traffic Data Missing 20

4.3.2 Traffic Sensor Failure Rate 21

4.3.3 Abnormal Data 22

Section 5. Travel Time Estimation System........ 25

5.1 Model Selection........ 25

5.2 Implementation........ 26

5.2.1 Prototype system........ 26

5.2.2 Excel VBA-based Tool........ 27

5.3 Case Studies........ 29

Section 6. Model Verification and Calibration........ 35

6.1 Model Verification using Bluetooth-based Travel Time........ 35

6.2 Model Verification using Vehicle-matching-based Travel Time........ 37

Section7. Conclusions and Recommendations........ 41

7.1 Conclusions........ 41

7.2 Recommendations........ 42

Acknowledgements........ 42

References........ 43

Table of Figures

Figure 3-1: RTMS and their locations along freeways in the Greater St. Louis area.................. 14

Figure 3-2: Traffic data collection flowchart.. 16

Figure 3-3: Bluetooth-based travel time data collection sites. The background image is from Google Earth.. 17

Figure 4-1: Database Schema. 19

Figure 4-2: Daily data quality control.. 20

Figure 4-3: Abnormal Data Examples.. 24

Figure 5-1: The Interface of the Travel Time Estimation System... 27

Figure 5-2: The Excel VBA-based tool for travel time estimation... 29

Figure 5-3: Travel time estimation system 32

Figure 5-4: Case study of traffic bottleneck identification on I-64 33

Figure 5-5: Comparisons of before-and-during snowstorm travel times on entire Length of I-64 (westbound).. 34

Figure 6-1: Locations of Bluetooth and point-speed sensors along the data collection segment of I-70 for travel time estimation result verification 35

Figure 6-2: Comparison of Bluetooth-based and estimated travel times..................................... 37

Figure 6-3: Collecting travel time ground truth data through the Vehicle-matching-based technique. 38

Figure 6-4: Location of MoDOT surveillance cameras along I-64 39

(image obtained from Google Maps).. 39

Figure 6-5: Comparison of Vehicle-matching-based travel times and estimated travel times along the second case study corridor (Jun. 27th, 2013) 39

Figure 6-6: Linear model vs. non-linear model 40

Table of Tables

Table 4-1: Real-Time Data Missing Events and Corresponding Weather Conditions................. 21

Table 4-2: Traffic Sensor Failure Rate, August 2012.. 22

Table 5-1: Travel Time Comparisons Before, During and After the Snowstorm 34

Table 6-1: Quantitative Comparison Between Bluetooth-based and Estimated Travel Times (January 2013)... 37

Executive Summary

Freeway travel time is one of the most useful pieces of information for road users and an important measure of effectiveness (MOE) for traffic engineers and policy makers. In the Greater St. Louis area, Gateway Guide, the St. Louis Transportation Management Center (TMC) operated by the Missouri Department of Transportation (MoDOT), collects traffic data from more than 700 fixed traffic sensors along the city's major freeways and arterial roads. Due to their significant investment in fixed freeway sensors, MoDOT has been striving to more fully utilize the data collected to extract useful information for stakeholders. Making this data more available to other transportation agencies will also enable them to investigate traffic issues and take more effective action to address them.

This project was aimed at providing an automatic and systematic approach to estimating travel time on the section of Interstate 64 (I-64) located in the St. Louis area using existing fixed sensors. Travel time data used to be collected by designated sensors specifically designed to collect travel time data, but the coverage of these travel time sensors is usually limited. A successful outcome for this project would allow MoDOT to evaluate the performance of the entire freeway network at a very low or no additional cost by accessing the data already collected by the fixed traffic sensors. The database created for this project should also facilitate data exchange within and between the state's universities and transportation agencies.

The project began by conducting a comprehensive literature review regarding travel time estimation to explore two important aspects of travel time estimation, namely fixed-sensor-based travel time estimation and ground truth data collection methods. The literature review not only assisted in the selection process for the travel time estimation model to be implemented but also facilitated the verification process at the end of the project.

No travel time estimation system can be built without suitable data support. Once the project's dedicated data server had been set up at Saint Louis University (SLU), a custom computer program was developed to receive the data from MoDOT via the File

Transfer Protocol (FTP). The data is stored in the original Extensible Markup Language (XML) format as flat files in the SLU data server and also imported into the Database Management System (DBMS) developed for this project. To improve the query performance, the database was optimized based on the most frequently used queries. The custom computer program is also designed to monitor the traffic data quality, including the number of missing flat files, and automatically produces a daily data quality summary which is sent to the data server administrator.

Once the data support was firmly in place, the instantaneous model was implemented on two platforms: a custom MATLAB-based travel time estimation system and an Excel VBA-based travel time estimation tool. Only the former was used in the verification and case studies because of its high performance. Three case studies were conducted to demonstrate the applicability of the proposed system. These case studies consisted of: 1) a fundamental traffic analysis; 2) bottleneck identification; and 3) a snow storm impact study.

During the verification process, the performance of the estimated travel time obtained using the instantaneous model was compared with the times measured directly from the Bluetooth-based travel time data and video-based vehicle-matching-based travel time data. The results show that the travel time results tended to be slightly under-estimated due to errors associated with the linear interpolation. Overall, however, the performance results from the optimized database were satisfactory and facilitated an easy-to-understand visualization of the traffic data. Because of the high performance of this system and the proper data quality control procedures implemented, the system was deemed capable of providing fairly useful information for MoDOT and enabling quick traffic analyses.

Moving toward Phase 2, the travel time estimation method will be further improved to increase its accuracy and expanded to cover the entire freeway network in the St. Louis district. This expansion will allow traffic engineers to oversee the overall freeway network performance, rather than the performance of a single freeway segment.

Section 1

Introduction

This section provides an overview of the project background, problem statement and research objectives.

1.1 Project Background

Travel time is one of the key performance measures used to evaluate transportation systems, but travel time information is generally very difficult to collect. High-tech traffic sensors such as Global Positioning Systems (GPS) and manual/automatic vehicle identification using videos and Bluetooth are now being used to collect/ estimate travel times and recently these traffic sensors have begun to be widely deployed to monitor traffic flow in the U.S. (Klein, 2001) for many applications of Intelligent Transportation Systems (ITS). In many major cities in the U.S., traffic data are also being collected from fixed traffic sensors (e.g. inductance loop detectors embedded in road surfaces) and fed into a server located in the city's Transportation Management Center (TMC) in real time. As these fixed sensors are point detectors, the data collected consist only of vehicle spot speed and volume, but this type of "spot" data could not only be used to conduct a variety of traffic analyses but potentially also to estimate travel time without the need to install additional travel time data collectors. However, this type of advanced data application has rarely, if ever, been considered by practitioners.

1.2 Problem Statement

In the St. Louis area, traffic data are collected from more than 700 fixed traffic sensors located along major freeways and arterials by the Missouri Department of Transportation (MoDOT) TMC. The traffic flow data are fed into the TMC server in real time and used to support MoDOT's daily traffic operations. MoDOT is seeking ways to utilize the traffic data from these sensors more effectively, and this project is designed to extract useful information that can be applied by MoDOT traffic engineers and the city's policy makers to improve traffic flows and minimize congestion. It is also expected to help TMC staff deal more efficiently with the many data requests they receive from different agencies. At present, when a request is accepted staff must manually download the data, a very time-consuming process. It is therefore desirable to have an integrated data analysis platform that incorporates a sustainable data portal that other agencies can access to download the data they require without the need to send multiple data requests to the TMC.

1.3 Research Objectives

The major goal of this project was to develop an efficient traffic data platform to provide a research foundation for advanced research in transportation engineering. A travel time estimation method suitable for St. Louis transportation network was identified and a computer program was developed to automate the analysis process, thus facilitating freeway performance evaluation in the St. Louis area.

Specifically, this study had the following objectives:

- Design an efficient database schema based on the characteristics of the traffic data;
- Select and implement a travel time estimation model, and apply the model to traffic travelling along Interstate 64 (I-64);
- Implement the travel time estimation analysis tool using the selected travel time estimation model;
- Integrate the tool into the new database to improve query performance;
- Compare the estimated travel time results with the results of other travel time estimation methods

Section 2

Literature Review

This study was designed to develop a way to estimate travel times using existing traffic sensors on freeways. A comprehensive literature review was therefore conducted to investigate existing methods of travel time estimation. In order to verify the performance of the proposed travel time estimation method, ground truth travel time collection methods were also reviewed.

2.1 Travel Time Estimation

A number of different data sources and techniques, such as automated vehicle identification (Ma and Koutsopoulos, 2008), floating car data (Ehmke et al., 2012) and electronic toll collection systems (Ozbay and Yildirimoglu, 2011), can be applied to estimate travel times. Since this project sought to utilize existing traffic sensor data to estimate travel times on freeways, those travel time estimation models based on fixed traffic sensors were the primary focus of this literature review. For the purposes of this research, a fixed sensor is defined as a sensor that can only collect traffic data such as speed and volume at a particular spot.

The instantaneous model (Li et al., 2006) is widely used to estimate travel times. It uses the average of upstream and downstream speed data collected by traffic sensors to calculate the average link speed and the length of the link divided by the average link speed is then the estimated travel time. The total travel time on a segment is calculated by summing the link travel times. However, Li et al. (2006) pointed out that the error inherent in travel time estimation by the use of the instantaneous model can be substantial as it depends on the speed at which vehicles are travelling and this can vary considerably over relatively long link lengths. In spite of their acknowledgement of this problem, the group did not investigate the relationship between the link length and the error in travel time estimation.

Other sensor-based travel time estimation models are based on vehicle trajectory construction using point speed data. For example, the time slice model (Li et al, 2006) is based on the instantaneous model, but unlike the instantaneous model, it captures the speed at which a vehicle is moving when it enters each link and then uses that speed data to calculate the travel times on individual links. Assuming all the vehicles travel at the same speed on a link, traffic flow theory is then applied to the speed data to construct a pseudo vehicle trajectory on segments (Coifman, 2002). He concluded that the performance of the proposed model during non-rush hours, when traffic is flowing more smoothly, was better than during rush hours. A piecewise truncated quadratic speed trajectory has also been proposed to simulate vehicle trajectory on a link (Sun et al., 2008). Coifman (2002) concluded that as his proposed method required only speed data, it would be easy to implement for online real time travel time estimation application.

2.2 Ground Truth Data Collection

Many techniques have been suggested to assist both researchers and practitioners to collect ground truth travel time data. This sub-section provides a literature review of research in this area.

2.2.1 Conventional Techniques

The *Travel Time Data Collection Handbook* (Turner et al., 1998) provides an overview of most of the travel time collection techniques that have been used to date. The four basic approaches described in the *Handbook* are as follows:

- Test vehicles
- License plate matching
- Intelligent Transportation Systems probe vehicles
- Emerging and non-traditional techniques

Test vehicles, or “active test vehicles”, have been used for travel time data collection since the late 1920s. Although several methods can be used for this technique, its basic principle is for a test vehicle to travel along a segment and record the time stamps as it passes predefined checkpoints on the segment. This technique, however, can result in errors caused by both human and electronic devices.

License plate matching techniques require license plate information to be collected from two or more sites. The travel time can then be directly obtained from the times at which the vehicles with those license plates pass each site. However, the major disadvantage of this approach is that the sample size is usually very limited due to the high data collection costs.

ITS probe vehicle techniques, or "passive probe vehicle" techniques, are used to collect travel time in real time. Automatic Vehicle Location (AVL) and GPS are both classified as ITS probe vehicle techniques. Here, the data collection cost can be high if more data samples are required and the travel time data collection process can be tedious.

Emerging and non-traditional techniques include the use of weight-in-motion stations, video cameras, and electronic toll collection (ETC), among others. However, these techniques require more advanced algorithms and models if they are to be used to estimate travel time.

2.2.2 Bluetooth-based Travel Time

Recently, travel time measurements using Bluetooth (Wasson et al, 2008) have become popular due to the widespread use of Bluetooth devices in our daily lives. Bluetooth-based travel time collection is a new technique that utilizes enabled Bluetooth portable devices such as mobile phones, computers, personal digital assistants, and car radios to identify specific vehicles at downstream and upstream locations by tracking their unique 48-bit Machine Access Control (MAC) addresses. Figure 2-1 shows how the travel time can be "calculated" by matching Bluetooth MAC addresses at consecutive detection locations along the road according to the time stamps associated with those MAC addresses. Bluetooth-based travel time data was used in this project to provide the ground truth travel times.

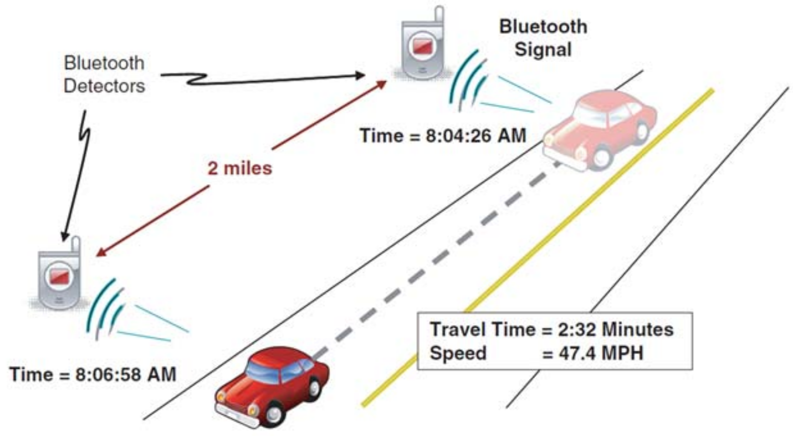


Figure 2-1: Bluetooth-based travel time measurement (Haghani et al., 2010)

Section 3.
Data Collection

This section summarizes the collection procedure and the results of the traffic data. The Bluetooth-based travel time data collected for the verification process is also introduced.

3.1 Traffic Data

3.1.1 Traffic Sensors

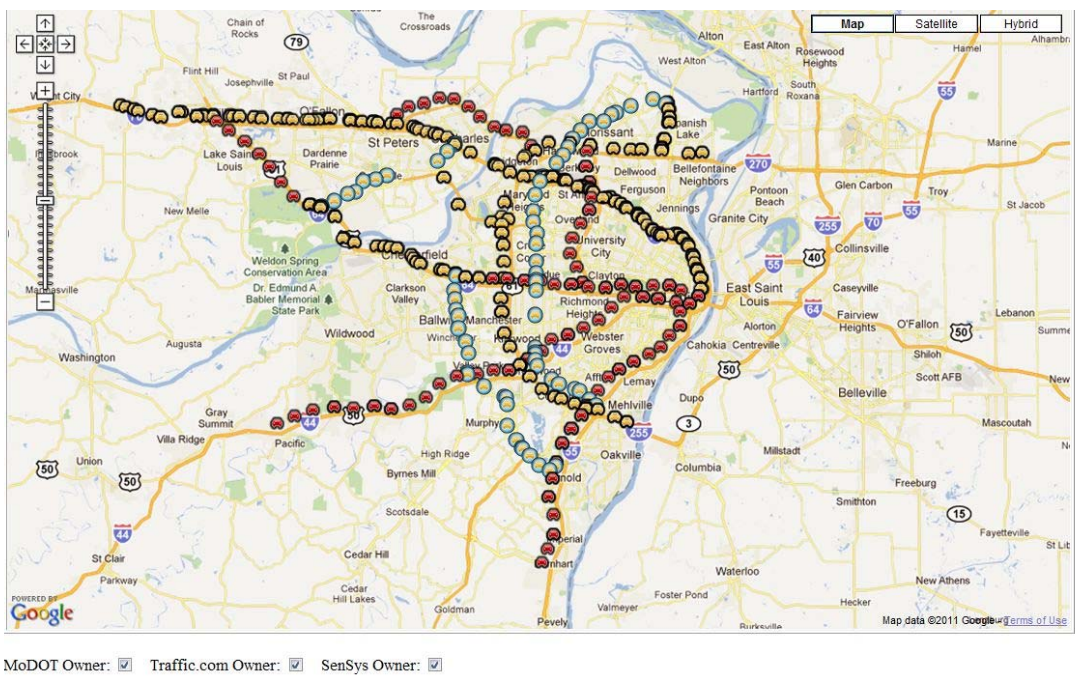
The data used to estimate travel time on the interstate was collected from Remote Traffic Microwave Sensors (RTMS®), a type of ITS traffic sensor, that have been installed along the major freeways in the Greater St. Louis area for traffic data collection. These sensors transmit a low-power microwave signal of constantly varying frequency in a fixed fan-shaped beam. The beam "paints" a long elliptical footprint on the road surface. Any non-background targets will reflect the signal back to the RTMS, where the targets are detected and their range measured.

RTMSs are used to collect traffic flow data, including traffic volume, speed, occupancy[1] and vehicle length, during a user defined time period. This time period is set at 30 seconds for the real-time feed sent to the data server located in MoDOT’s TMC. Two examples of the RTMSs deployed by the MoDOT are shown in Figures 3-1(a) and (b) and the locations of the RTMSs in St. Louis are depicted in Figure 3-1(c). Each icon represents a RTMS sensor monitoring all of the lanes of the freeway for both directions. For example, the RTMS shown in Figure 3-1 (a) monitors three lanes of westbound I-64 and three lanes of eastbound I-64. Three fundamental traffic parameters are collected from the RTMSs for each lane: aggregated volume, average speed and average occupancy every 30 seconds. Information on the vehicle classification by vehicle length is also currently collected by the RTMSs but is not stored in the data server.

[1] “Occupancy” is defined as the percentage of time the sensor detection area is occupied by vehicles during a specific time period

(a)

(b)



(c)

Figure 3-1: RTMS and their locations along freeways in the Greater St. Louis area

(a) the RTMS located at milepost 29.8 on I-64 westbound; (b) the RTMS located at milepost 28.6 on I-64 westbound; (c) locations of the RTMSs in St. Louis. (a) and (b) are from Google Street View and the background image in (c) is from Google Maps.

3.1.2 Archiving Data at Saint Louis University (SLU)

The MoDOT server receives traffic data from all of the existing ITS sensors in St. Louis and generates two Extensible Markup Language (XML) files every 30 seconds from the system, consisting of: 1) Real-time traffic data that contains fundamental traffic parameters (e.g. volume, speed, and occupancy), and 2) the meta data storing basic information for each ITS sensor (e.g. location, number of lane being detected). Both the real time traffic data and the meta data are transferred to a File Transfer Protocol (FTP) server that has been physically located at the Smart Transportation Lab at Saint Louis University (SLU) since June, 2012. More than one gigabyte of real time traffic data is pushed from the MoDOT server to the SLU server via FTP every day. To increase data redundancy, the data is stored in the local SLU server, the shared drive and the MS SQL database. Figure 3-2 illustrates an overview of the traffic data collection flow.

Since the flat files are overwritten every 30 seconds on the SLU FTP server, in order to archive the two types of flat files, a custom C# computer program was developed to perform the tasks listed below:

- Automatically monitor changes in the flat files. When the flat files are overwritten, a signal will be sent to the computer program;
- Transfer the files from the SLU FTP server to both of the local and remote storage systems if the files change;
- Rename the files based on the time stamp in the flat files;
- Parse the transferred flat files to obtain the traffic information;
- Tabulate the parsed traffic information;
- Store the tabulated traffic information in the database (more details can be found in Section 4: Traffic Database);

- Generate a daily data quality summary and send it to the SLU team (more details of this process can also be found in Section 4).

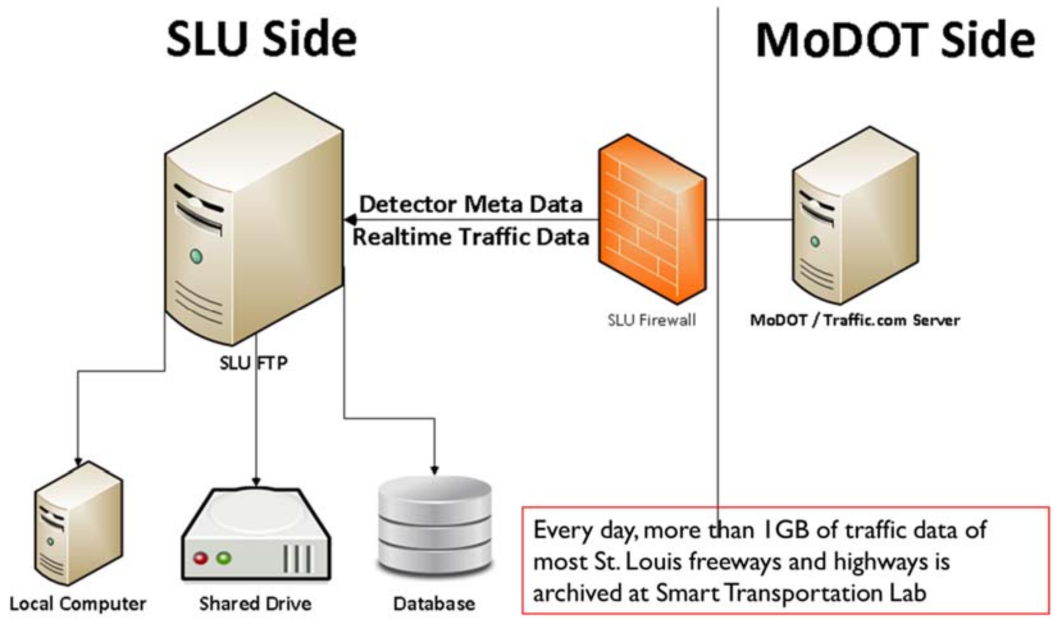


Figure 3-2: Traffic data collection flowchart

In addition to the real-time data, the historical traffic data was also successfully migrated from the MoDOT server to the SLU traffic data server to extend the data coverage back to Jan. 1st, 2008. This data provides additional information for future applications and analyses.

3.2 Bluetooth-based Travel Time Data

As discussed in Section 2, Bluetooth-based travel time data was used as the ground truth data for this project. The Bluetooth-based data was collected by Bluetooth Travel-time Origination and Destination (BlueTOAD™), operated by the MoDOT.

Four sets of Bluetooth-based travel time data were provided by the Transportation Management Center (TMC) in St. Louis. The data sets were collected during the period November 2012 to March 2013. The roadway segments are listed below:

- Segments on US 67 Highway from New Florissant to New Halls Ferry;

- Segments on I-70, between SR-94 and I-270, both eastbound and westbound; and
- Roadway segment on I-64 between I-55 and Ewing Ave.

Figure 3-3 shows the roadway segments where the Bluetooth travel time data was collected. According to the scope of this project, only the freeway data was to be used. However, the Bluetooth-based travel time collection segment on I-64 only partially overlaps the fixed sensors. Two countermeasures were taken: 1) an alternative Interstate 70 dataset was used to verify the estimated travel time and was thus considered an additional case study for this project, and 2) additional ground truth data was manually collected for I-64 since this was the primary focus of this project. Details of the ground truth data collection process will be provided in Section 6.

Figure 3-3: Bluetooth-based travel time data collection sites. The background image is from Google Earth.

Section 4.
Traffic Database Design and Implementation

This section describes the traffic database design and implementation, focusing particularly on the database schema design and optimization.

4.1 Database Design

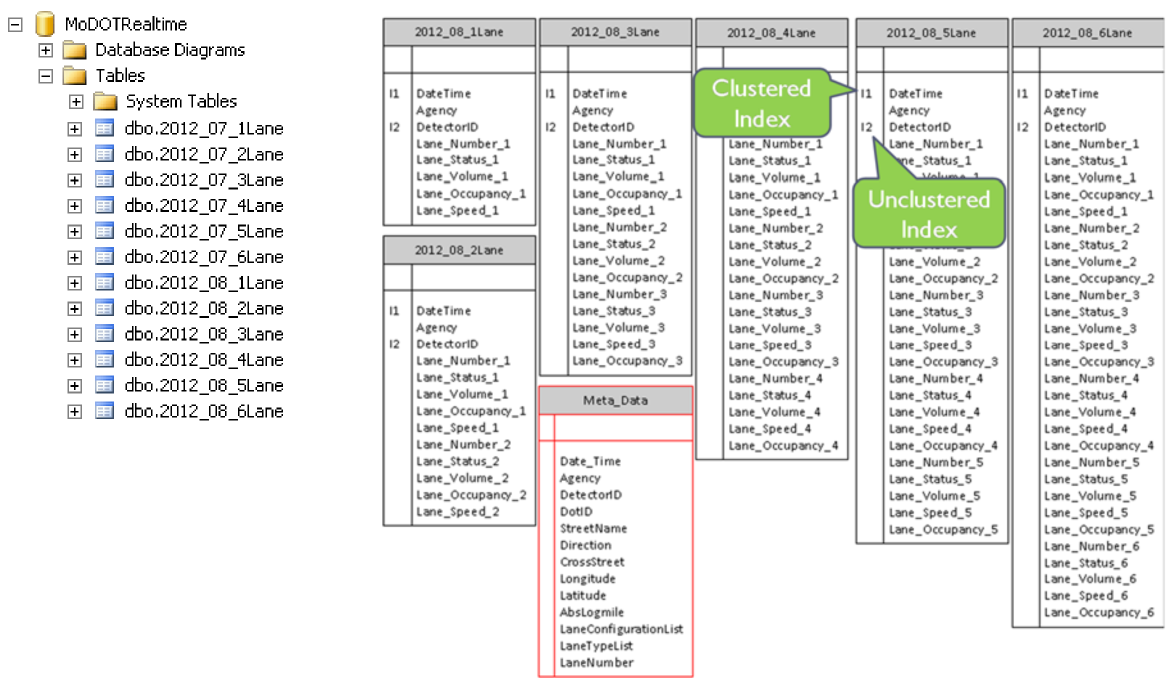
The efficiency of the travel time estimation calculations requires strong support from efficient database operations, so the SLU traffic database (Figure 4-1) needed to be re-designed in order to achieve a satisfactory performance. Microsoft SQL Server 2008 was used to build the new traffic database management system (DBMS).

Figure 4-1 shows a schematic diagram of the data tables and the new database designed for this project. All the data collected from a RTMS with the same lane configuration is grouped into one specific table. . For example, for all RTMSs monitoring three lanes on I-64 both westbound and eastbound on July, 2012, the traffic data collected from these RTMSs would be stored in the table “2012-07-3Lane”. , Since the highest number of lanes per direction is six, six data tables are accordingly created for each month. In addition to the real time traffic data collected, the metadata is exclusively stored in the table “Meta_Data”.

4.2 Database Optimization

The database optimization design is based on the most frequently used SQL queries. In order to calculate the travel time on either a specific segment of I-64 or for the entire length of I-64, the speed information is extracted for a given combination of a specific time period and the IDs of consecutive RTMSs from upstream through to downstream. In an SQL query, the performance of the clause “order by”, which is used to sort the records in a data table by key words, mainly determines the response time from a DBMS. In the most frequently used SQL queries, the RTMS must be spatially sorted from upstream to downstream. Microsoft SQL Server 2008 requires that only one clustered index can be created per data table, so a clustered index based on the attribute “DateTime” is created for each table. In addition to the clustered index, an unclustered index is created on the attribute "DetectorID" in order to improve the performance of the clause "where" in the SQL queries.

The SQL queries, which are most commonly used to extract the essential data for estimating travel time, were tested to examine the database performance enhancement. Before these two indexes were created, an SQL query took 50 seconds to return the corresponding results, but once the indexes had been created, the response time dropped to just 1 second.



(a) **(b)**

Figure 4-1: Database Schema.

(a) Examples of data tables, (b) Database schema

4.3 Traffic Data Quality Assurance

Traffic data quality assurance was conducted to assess the data quality, data accuracy, data completeness and data reliability of the database. A custom C# computer program was developed to fulfill this task.

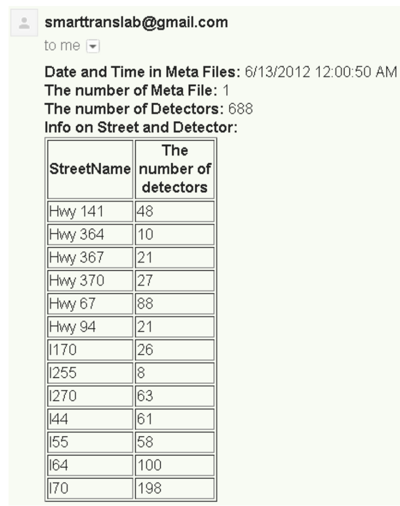
The new program monitors the real time traffic data and automatically sends a mail notification to the administrator(s) if an abnormal situation is detected. At the end of each day, the daily metadata quality summary and the daily real time traffic data quality summary are also sent to the system administrators. The basic daily summaries of the traffic data are intended to provide a preliminary evaluation of the traffic data quality. Figure 4-2 shows examples of the daily real time data summary and the daily metadata summary.


Realtime Data Summary

smarttranslab@gmail.com
to me

Date and Time: 7/11/2012 12:00:00 AM
The number of Realtime Files: 2821
The number of Missing Files: 3

(a)


smarttranslab@gmail.com
to me

Date and Time in Meta Files: 6/13/2012 12:00:50 AM
The number of Meta File: 1
The number of Detectors: 688
Info on Street and Detector:

StreetName	The number of detectors
Hwy 141	48
Hwy 364	10
Hwy 367	21
Hwy 370	27
Hwy 67	88
Hwy 94	21
I170	26
I255	8
I270	63
I44	61
I55	58
I64	100
I70	198

(b)

Figure 4-2: Daily data quality control
(a) Daily real-time data quality summary (b) Daily metadata quality summary

4.3.1 Real-time Traffic Data Missing

Due to network communication issues, the update frequency of the real time traffic data feed occasionally becomes 31 seconds rather than 30 seconds. Thus, the total number of daily real-time traffic data feeds theoretically ranges from 2,788 to 2,880. The number of daily existing files can be easily counted and the number of missing real time traffic data files calculated accordingly.

The research team also found an interesting correlation between the number of missing real time traffic data files and local weather conditions. It was found that inclement weather events were associated with days with a higher number of dropped data feeds. A daily summary of weather conditions is also recorded in a log file, based on weather information provided by wunderground.com. Table 4-1 lists all the days on which there were five or more missing files for the period from July 2012 to Oct. 2012. Therefore, a closer investigation was necessary to determine why inclement weather conditions have a negative impact on data transfer.

Table 4-1: Real-Time Data Missing Events and Corresponding Weather Conditions

Date	# of existing files	# of missing files	Weather conditions
08/09, 2012	2749	77	Wind Dir: WNW Max Wind Speed: 20 mph Max Gust Speed: 29 mph Rain, Thunderstorm
08/21, 2012	455	--*	Wind Direction: ESE Max Wind Speed: 9mph Max Gust Speed: - mph Rain
09/07, 2012	2819	7	Wind Direction: NW Max Wind Speed: 30mph Max Gust Speed: 38mph Rain, Thunderstorm
09/16, 2012	2799	31	Wind Direction: SE Max Wind Speed: 8mph Max Gust Speed: 12 mph ----
10/03, 2012	2653	183	Wind Direction: SSW Max Wind Speed: 14mph Max Gust Speed: 17mph Rain
10/13, 2012	2768	70	Wind Direction: S Max Wind Speed: 23mph Max Gust Speed: 29 mph ----
10/16, 2012	2430	415	Wind Direction: SSE Max Wind Speed: 21mph Max Gust Speed: 29mph ----
10/22, 2012	1762	1078	Wind Direction: SSE Max Wind Speed: 18mph Max Gust Speed: 24 mph Rain, Thunderstorm
10/23, 2012	1555	1301	Wind Direction: S Max Wind Speed: 17mph Max Gust Speed: 22mph Rain, Thunderstorm

--*: too many files were missing

4.3.2 Traffic Sensor Failure Rate

The traffic sensor working status is also indicated in one attribute, lane status, of the real-time data feed, with either “OK”, “Failed” or “Disabled” being displayed in the attribute. When the "lane status" of an RTMS is not shown as "OK", the RTMS is considered to be malfunctioning. The RTMS failure rate can be calculated as the number of sensors labeled "Failed" or "Disabled" divided by the number deemed "OK". Table 4-2 gives an example of the real time traffic data failure rate during August 2012. Note that the freeways in St. Louis have at least two lanes each direction; RTMSs installed on ramps monitor only one lane. Since the project focused on freeways, Table 4-2 excludes the failure rate of traffic sensor on ramps. Overall, the real time traffic data failure rate was 3.94% for the month. This low percentage indicates the overall good data quality achieved.

Table 4-2: Traffic Sensor Failure Rate, August 2012

# of Lane	# of Total Records	# of Failure Records	Percentage
2	9,451,854	457,004	4.84%
3	15,421,446	699,049	4.53%
4	10,861,341	318,879	2.94%
5	2,818,974	51,296	1.82%
6	165,822	37	0.02%
total:	*387,194,37*	*1,526,265*	*3.94%*

One possible reason for the low incidence of real-time traffic sensor failures may be MoDOT's good maintenance program for the traffic sensors. Network communication malfunctions may also adversely affect data completeness. Those data labeled "Failed" or "Disabled" are currently not used in estimating travel times.

4.3.3 Abnormal Data

In some situations, the data is considered "abnormal". This normally arises in one of two ways, categorized as follows:

1) Type 1: When the lane status is 'OK', the traffic flow data (volume, occupancy and speed) is unavailable; and

2) Type 2: When the values of volume, occupancy and speed are incompatible.

Figure 4-3 shows some examples of abnormal data. For example, it is not reasonable that vehicle speeds of greater than 90 mph or ‘occupancy’ levels of more than 90% would be achieved during rush hours. It is also not reasonable to see volume and speed values of zero during rush hours.

The possible causes of the abnormal data are as follows:

1. The network communication infrequently fails (Fig.4-3a);
2. Incidents may account for inconsistent travel speeds in each lane (Fig. 4-3b); and
3. When the traffic sensor detection zone is occupied by vehicles for longer than 30 seconds (e.g. a crashed or broken down vehicle), the resulting “occupancy” value is 100 and the volume is 0 (Fig. 4-3c).

The speed information is essential for the travel time estimation model. In order to remove abnormal speed data and “Failed” (or “Disabled”) data mentioned above, the median speed readings of all the lanes are used to represent traffic speed on the freeways.

	Date_Time	Agency	DetectorID	Lane_Number_1	Lane_Status_1	Lane_Volume_1	Lane_Occupancy_1	Lane_Speed
1	2012-08-01 00:00:25.000	STLATMS	MR067N013.0D	1	OK	-1	-1	-1
2	2012-08-01 00:00:25.000	STLATMS	MR067S013.0D	1	OK	-1	-1	-1
3	2012-08-01 00:00:25.000	STLATMS	MR067N015.8D	1	H Failed	-1	-1	-1
4	2012-08-01 00:00:25.000	STLATMS	MR067S015.8D	1	H Failed	-1	-1	-1
5	2012-08-01 00:00:56.000	STLATMS	MR067N013.0D	1	OK	-1	-1	-1
6	2012-08-01 00:00:56.000	STLATMS	MR067S013.0D	1	OK	-1	-1	-1
7	2012-08-01 00:00:56.000	STLATMS	MR067N015.8D	1	H Failed	-1	-1	-1
8	2012-08-01 00:00:56.000	STLATMS	MR067S015.8D	1	H Failed	-1	-1	-1
9	2012-08-01 00:01:27.000	STLATMS	MR067N013.0D	1	OK	-1	-1	-1
10	2012-08-01 00:01:27.000	STLATMS	MR067S013.0D	1	OK	-1	-1	-1
11	2012-08-01 00:01:27.000	STLATMS	MR067N015.8D	1	H Failed	-1	-1	-1
12	2012-08-01 00:01:27.000	STLATMS	MR067S015.8D	1	H Failed	-1	-1	-1
13	2012-08-01 00:01:58.000	STLATMS	MR067N013.0D	1	OK	-1	-1	-1
14	2012-08-01 00:01:58.000	STLATMS	MR067S013.0D	1	OK	-1	-1	-1

(a)

	Date_Time	DetectorID	Lane_Volume_1	Lane_Occupancy_1	Lane_Speed_1	Lane_Volume_2	Lane_Occupancy_2	Lane_Speed_2
2...	2012-08-12 17:12:31.000	MI064W010.0D	9	1	37	10	2	99
2...	2012-08-12 17:13:02.000	MI064W010.0D	4	1	37	5	5	99
2...	2012-08-12 17:13:32.000	MI064W010.0D	6	1	37	8	3	94
2...	2012-08-12 17:14:03.000	MI064W010.0D	9	1	37	9	4	98
2...	2012-08-12 17:14:33.000	MI064W010.0D	2	1	37	9	2	98
2...	2012-08-12 17:15:04.000	MI064W010.0D	5	1	37	7	2	93
2...	2012-08-12 17:15:34.000	MI064W010.0D	3	1	37	7	3	93
2...	2012-08-12 17:16:04.000	MI064W010.0D	5	1	37	10	3	91
2...	2012-08-12 17:16:35.000	MI064W010.0D	4	1	36	7	2	91
2...	2012-08-12 17:17:05.000	MI064W010.0D	6	1	36	8	5	91
2...	2012-08-12 17:17:35.000	MI064W010.0D	4	1	36	4	1	91
2...	2012-08-12 17:18:06.000	MI064W010.0D	11	2	36	9	2	94
2...	2012-08-12 17:18:36.000	MI064W010.0D	1	1	36	3	2	94
2...	2012-08-12 17:19:07.000	MI064W010.0D	4	1	36	4	1	94
2...	2012-08-12 17:19:37.000	MI064W010.0D	3	1	36	5	1	92
2...	2012-08-12 17:20:07.000	MI064W010.0D	4	1	36	5	3	92
2...	2012-08-12 17:20:37.000	MI064W010.0D	5	1	36	7	2	92
2...	2012-08-12 17:21:08.000	MI064W010.0D	3	1	39	4	1	92
2...	2012-08-12 17:23:09.000	MI064W010.0D	10	1	39	8	2	91
2...	2012-08-12 17:23:40.000	MI064W010.0D	3	1	39	4	1	91

(b)

	Date_Time	DetectorID	Lane_Volume_1	Lane_Occupancy_1	Lane_Speed_1	Lane_Volume_2	Lane_Occupancy_2	Lane_Speed_2	Lane_Volume_3	Lane_Occupar
22	2012-08-05 02:44:37.000	MI044W288.2U	5	7	29	7	93	13	1	1
23	2012-08-05 02:54:15.000	MI044W288.2U	7	16	33	9	100	12	0	0
24	2012-08-05 02:54:45.000	MI044W288.2U	7	16	33	9	100	12	0	0
25	2012-08-13 02:33:22.000	MI044W288.2U	6	6	42	10	90	14	3	2
26	2012-08-13 02:33:53.000	MI044W288.2U	6	6	42	10	90	14	3	2
27	2012-08-02 02:49:49.000	MI270N016.6D	0	0	0	0	100	0	0	100
28	2012-08-02 02:50:19.000	MI270N016.6D	0	0	0	0	100	0	0	100
29	2012-08-02 02:50:50.000	MI270N016.6D	0	0	0	0	100	0	0	100
30	2012-08-02 02:51:20.000	MI270N016.6D	0	0	0	0	100	0	0	100
31	2012-08-04 04:53:31.000	MI064E035.7U	6	9	26	11	99	10	1	2
32	2012-08-04 04:54:01.000	MI064E035.7U	6	9	26	11	99	10	1	2
33	2012-08-04 04:50:26.000	MI064E035.7U	6	10	29	7	100	14	0	0
34	2012-08-04 04:50:57.000	MI064E035.7U	6	10	29	7	100	14	0	0
35	2012-08-04 05:31:14.000	MI064E035.7U	5	19	12	10	94	7	1	1
36	2012-08-04 05:31:45.000	MI064E035.7U	5	19	12	10	94	7	1	1
37	2012-08-04 05:34:17.000	MI064E035.7U	6	14	9	10	100	6	4	4
38	2012-08-04 05:34:48.000	MI064E035.7U	6	14	9	10	100	6	4	4
39	2012-08-04 05:38:22.000	MI064E035.7U	11	24	6	8	94	9	3	3
40	2012-08-04 05:38:53.000	MI064E035.7U	11	24	6	8	94	9	3	3
41	2012-08-04 05:39:24.000	MI064E035.7U	11	46	10	13	100	10	2	2

(c)

Figure 4-3: Abnormal Data Examples

(a) Failure of data acquisition when the sensor status is OK, (b) Unreasonable speed data, and (c) Unreasonable volume and occupancy data

Section 5.

Travel Time Estimation System

This section presents the travel time estimation model selected for this project, the implemented system and three case studies that demonstrate the feasibility and applications of the proposed system.

5.1 Model Selection

The literature review described in Section 2 identified several models that have been specifically developed for travel time estimation using fixed point sensors. Based on the findings reported by Li et al. (2006), the instantaneous model and the time slice model were deemed the most suitable candidates for this implementation.

The instantaneous model uses real-time speed data from the upstream and downstream sensors of each link at time *k*. The link travel time can then be calculated through dividing the link length by the average of the collected speed data as formulated in Equation (1):

$$T(i,k) = \frac{2l_i}{v(i_u,k) + v(i_d,k)} \qquad (1)$$

where $v(i_u,k)$ and $v(i_d,k)$ are the measured speeds at the upstream and downstream end points of link *i* at a time *k*; l_i represent the length of the link *i*, and $T(i,k)$ is the link travel time. Accordingly, the total travel time $T(k)$ for a vehicle beginning its trip at time *k* is the summation of the estimated travel time of *n* links:

$$T(k) = \sum_{i=1}^{n} t(i,k) \qquad (2)$$

Unlike the instantaneous model, the time slice model attempts to account for variations in speed over time by constructing a vehicle "trajectory" using downstream speed values. Assuming the trip start time is *k*, the first link travel time is calculated similarly to the instantaneous model in Equation (1) and denoted as *t*(1, *k*), so the arriving time at the second link travel time can be expressed as *k* + *t*(1, *k*). Therefore, the travel time on the second link is $t(2,t_2) = \frac{2*l_2}{v[2_a,k+t(1,k)]+v[2_b,k+t(1,k)]}$. Generally, the travel time on link *n* can be written as Equation (3):

$$T(n,t_n) = \frac{2l_n}{v(n_u,t_n) + v(n_d,t_n)} \qquad (3)$$

where $t_n = k + t(1,k) + \sum_{i=2}^{n-1} t(i,\ t_i)$. As in the instantaneous model, the total travel time is calculated by summing all the link travel times.

Both the instantaneous and time slice models assume that speeds are constant along each link when calculating travel times. As a consequence, a discontinuity in the speed occurs as a vehicle leaves one link and enters the downstream link. According to the preliminary results of these two models, the differences between the two models were fairly minor. To expedite data processing, the faster instantaneous model was selected for this project.

5.2 Implementation

Since the Phase 1 project focuses primarily on the feasibility of the proposed approach, one prototype system and one EXCEL VBA-based tool were implemented to not only demonstrate the feasibility but also to gather feedback from MoDOT TMC staff. Both the system and the tool serve the same purpose – travel time estimation - but they have different hardware and software requirements.

5.2.1 Prototype system

The travel time estimation prototype system was first developed using MATLAB®, a high-level technical computing language. Figure 5-1 shows the graphical user interface (GUI) of the developed prototype system. Users can freely select the start (origin) and end (destination) points on I-64, the dates of interest, and the time of day. The system can retrieve data from the database and then calculate the average/median travel time automatically.

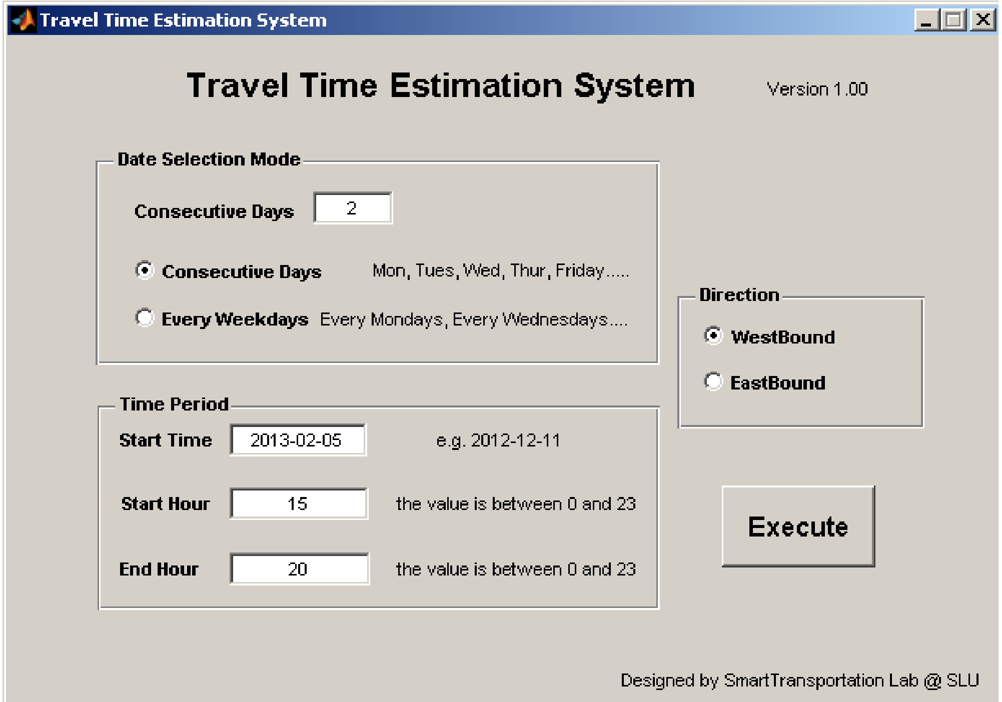
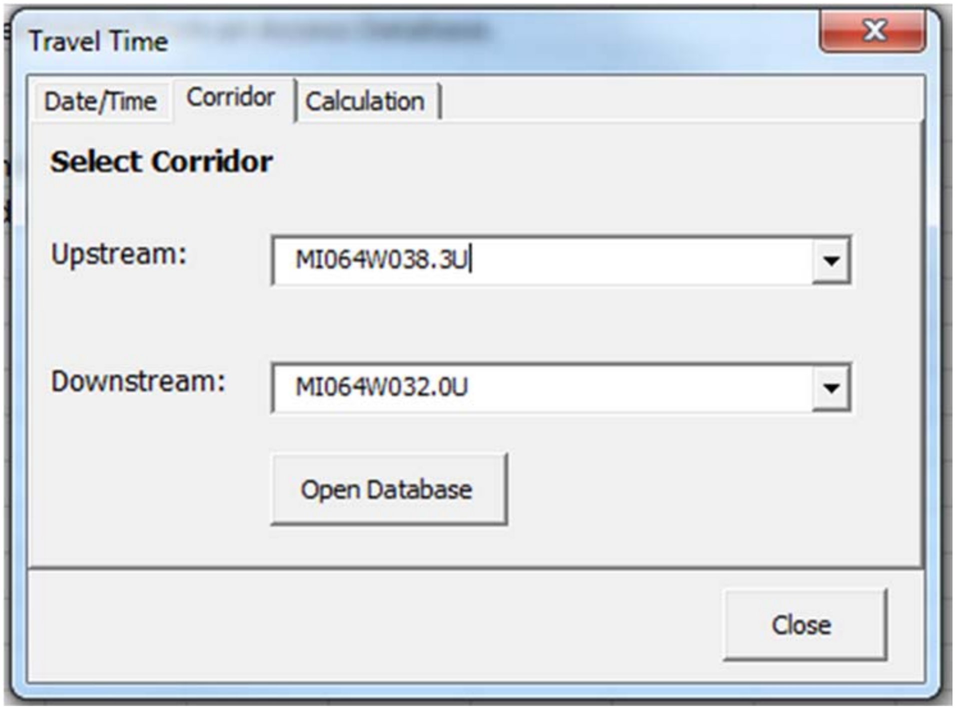


Figure 5-1: The Interface of the Travel Time Estimation System

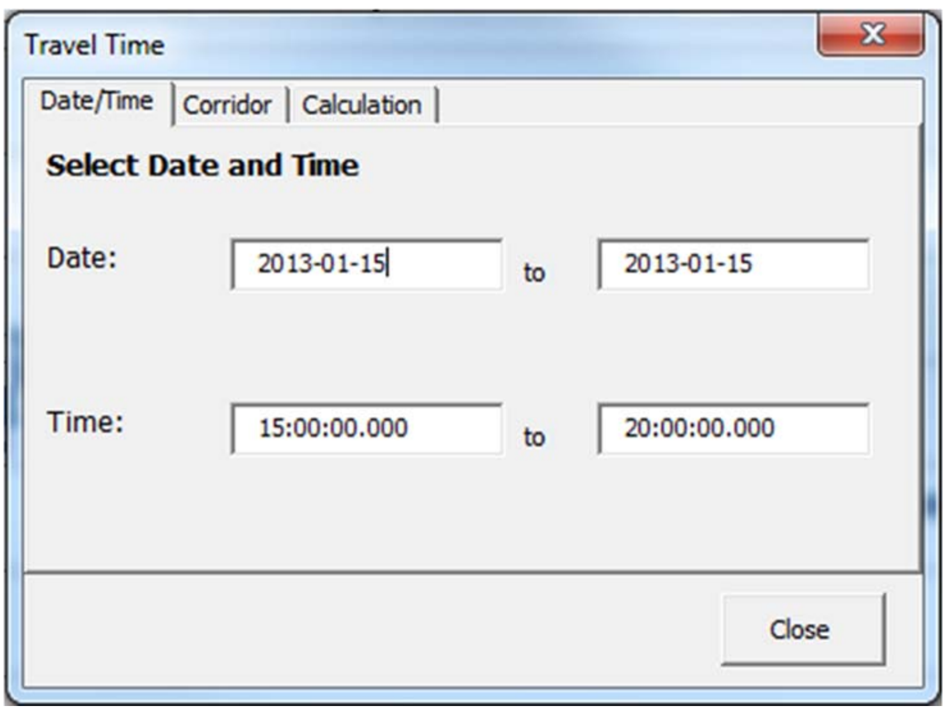
5.2.2 Excel VBA-based Tool

In addition to the MATLAB-based travel time estimation system, an EXCEL VBA-based tool was also developed. This tool is connected to a Microsoft ACCESS database rather than an SQL database. The advantage of this design is that the tool can run on any standalone computer without the need for any IT support. Its main disadvantage is that the MS ACCESS database requires frequent manual updates by traffic engineers because the tool is an offline system.

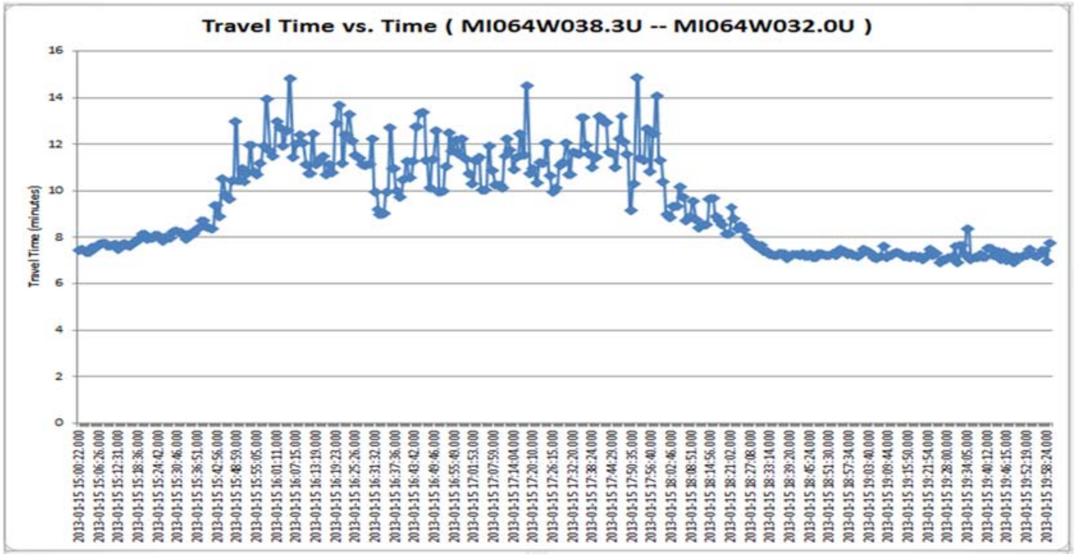
This tool performs similar functions to the MATLAB-based system in that it allows users to freely select any roadway segments between two sensors on I-64 and then select a date and time period to generate a travel time report. The GUI and the estimated travel time results are shown in Figure 5-2.


Travel Time
Date/Time
Corridor
Calculation
Select Corridor
Upstream:
MI064W038.3U
Downstream:
MI064W032.0U
Open Database
Close

(a)


Travel Time
Date/Time
Corridor
Calculation
Select Date and Time
Date:
2013-01-15
to
2013-01-15
Time:
15:00:00.000
to
20:00:00.000
Close

(b)



(c)

Figure 5-2: The Excel VBA-based tool for travel time estimation
(a) Upstream and downstream sensor selection, (b) Date and time selection, and (c) Results for the estimated travel time

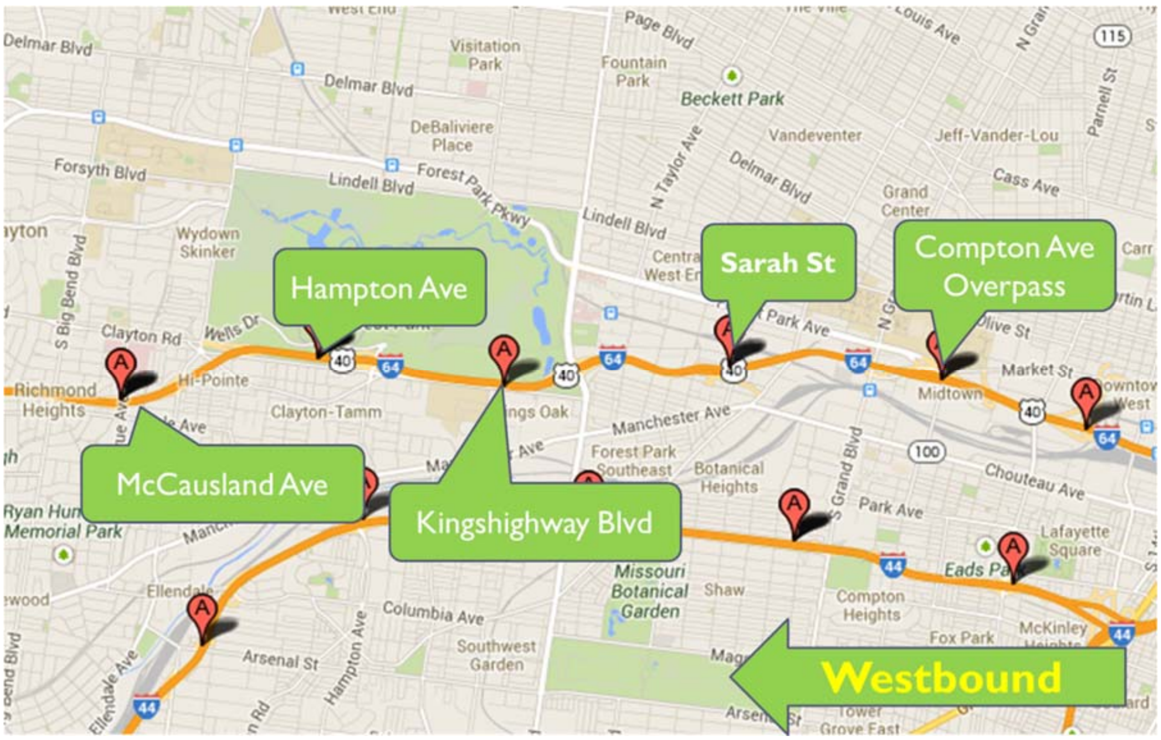
5.3 Case Studies

After the evaluation of the EXCEL VBA-based tool and the prototype system and a review of the feedback from the TMC staff, it was found that the prototype system provided more flexibility in function development. Therefore, several additional functions were developed for the prototype system to increase its usability. To demonstrate these functions, several case studies were conducted.

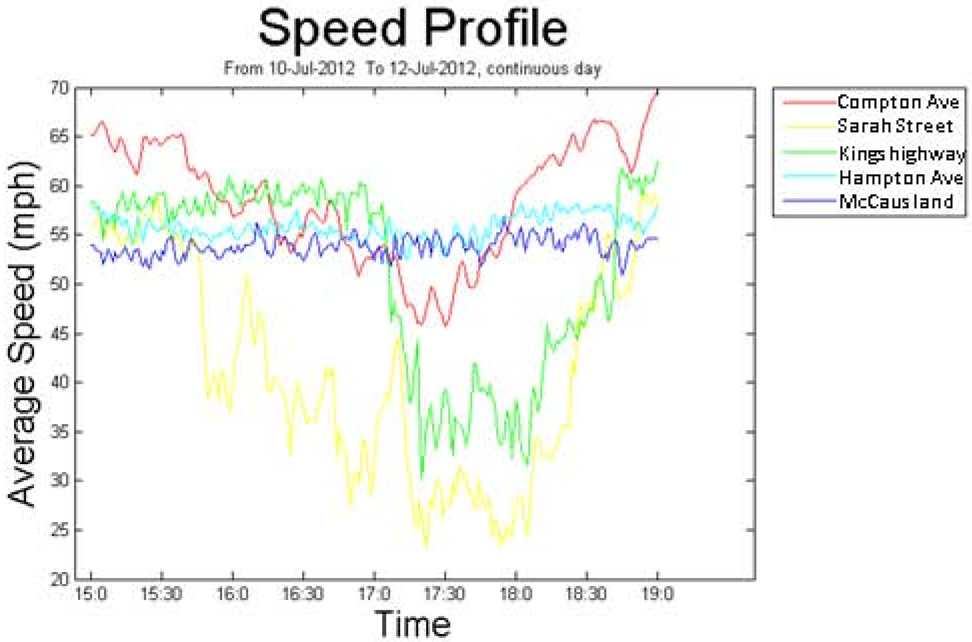
Case Study 1: Fundamental Traffic Analysis

The first case study investigated the segment of I-64 between Compton Ave and McCausland Ave. This segment was chosen because it is a well-known bottleneck that frequently suffers from recurrent congestion in the westbound direction during afternoon peak hours on weekdays. To investigate a traditional day of congestion, the data collected during the period 3pm~7pm for the three days from July 10, 2012 through to July 12, 2012, was extracted from the database and visualized by the prototype system. Figure 5-3 (a) shows the case study roadway segment on I-64, where five RTMS are located. Figure 5-3 (b) shows the speed profiles for all the RTMS on the study roadway segment from 3pm to 7pm for the three days of the case study. Figure 5-3 (c) shows the speed heat map, which is a fairly useful way for traffic engineers to identify the amount of delay caused by the congestion. Figure 5-3 (d) shows the travel time profile estimated by the system along this study roadway segment. Figure 5-3 (e) depicts a scatter plot showing the relationship between flow rate (volume) and speed data. This plot is also fairly useful and is commonly used by transportation researchers to investigate traffic flow problems.

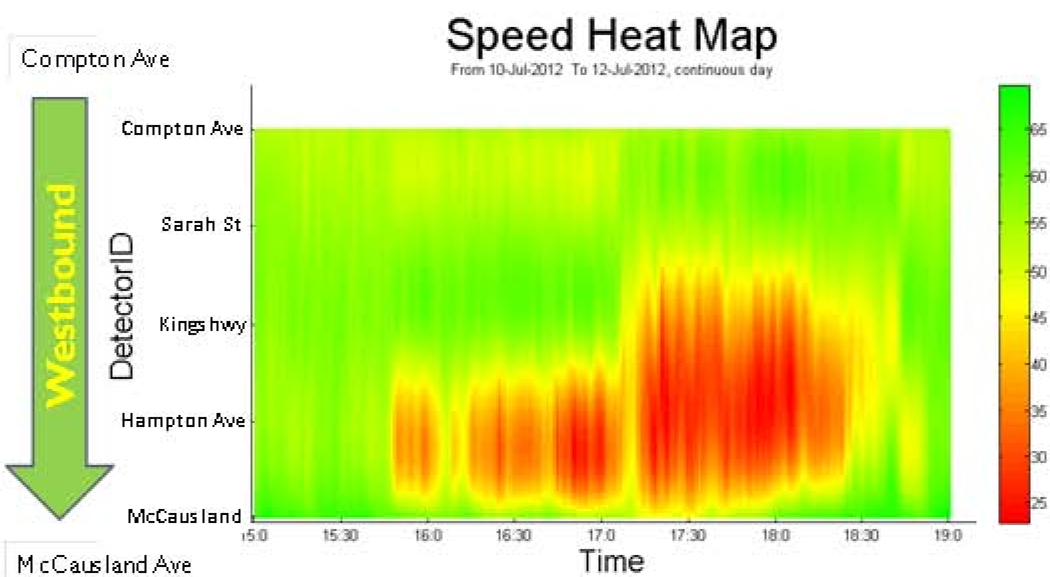
It should be noted that all the figures are generated automatically from the system based on the inputs from the users. This system is specifically designed to enable traffic engineers and researchers to conduct travel time analyses.



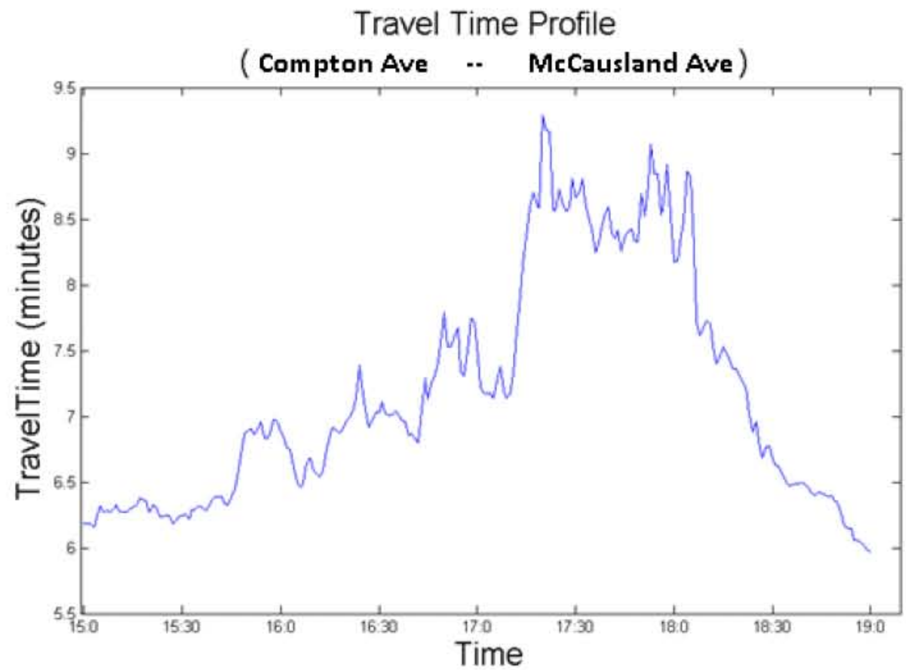
(a)


Speed Profile
From 10-Jul-2012 To 12-Jul-2012, continuous day
Average Speed (mph)
Time
Compton Ave
Sarah Street
Kings highway
Hampton Ave
McCausland
70
65
60
55
50
45
40
35
30
25
20
15:0
15:30
16:0
16:30
17:0
17:30
18:0
18:30
19:0

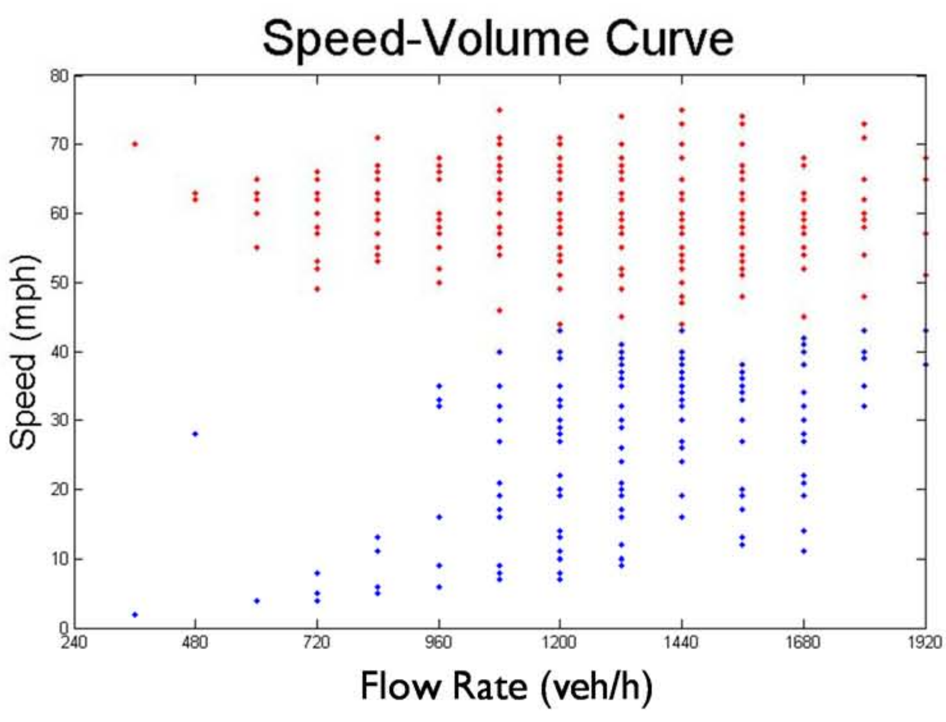
(b)


Speed Heat Map
From 10-Jul-2012 To 12-Jul-2012, continuous day
Compton Ave
Westbound
McCausland Ave
DetectorID
Compton Ave
Sarah St
Kings hwy
Hampton Ave
McCausland
Time
15:0
15:30
16:0
16:30
17:0
17:30
18:0
18:30
19:0
65
60
55
50
45
40
35
30
25

(c)



(d)



(e)

Figure 5-3: Travel time estimation system

(a) Case study site on I-64, (b) Speed profiles for five consecutive sensors, (c) Speed heat map (d) Travel time estimation results, and (e) Speed-volume relationship plot

Case Study 2: Bottleneck Identification

The first case study demonstrated a small scale traffic analysis, but the prototype system is expected to also be able to handle a larger query to visualize the results for the entire corridor. As illustrated in Figure 5-4, three bottlenecks can be visualized by querying three consecutive days of traffic speed data for the entire I-64 corridor. These three bottlenecks are at McCausland Ave (Case Study 1), the intersection with I-270 and the Daniel Boone Bridge on I-64.

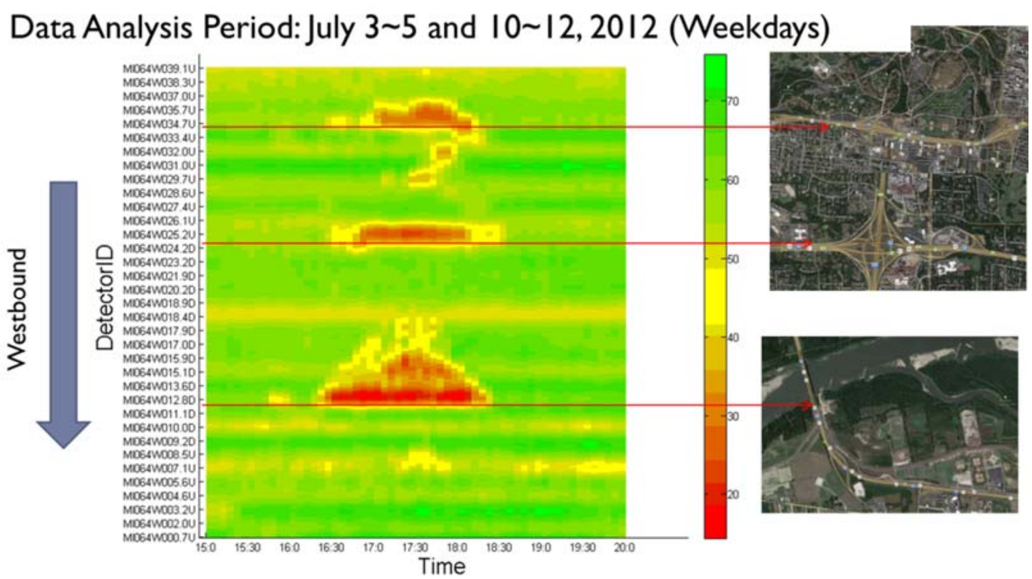


Figure 5-4: Case study of traffic bottleneck identification on I-64

Case Study 3: Snow Storm Impact Study

The third case study investigated the impact of a snowstorm on I-64 traffic. The freeway performance (travel time) before, during and after the snowstorm on Feb., 21st, 2013 is summarized in Table 5-1. The maximum travel time along the I-64 corridor increased to 111 minutes on the day of the snowstorm compared to 46 minutes on a regular day. In other words, many travelers were experiencing 2.4 times their normal travel time during the snowstorm. Interestingly, MoDOT was able to return the roadway conditions back to normal very soon after the snowstorm, because the minimum and maximum travel times the day after the snowstorm were effectively back to what they had been earlier in the week. Figure 5-5 compares the travel times on I-64 for the day before and the day of the snowstorm. Looking at the graph, it is clear to see that the snowstorm impacts started at 10am as the snowstorm hit the St. Louis area and were still delaying traffic that evening.

Table 5-1: Travel Time Comparisons Before, During and After the Snowstorm

	Westbound I-64			Eastbound I-64		
	Before-snowstorm traffic*	Snowstorm day traffic (Feb. 21st, 2013)	After-snowstorm traffic(Feb. 22nd, 2013)	Before-snowstorm traffic*	Snowstorm day traffic (Feb. 21st, 2013)	After-snowstorm traffic(Feb. 22nd, 2013)
Min travel time (minutes)	39	63	41	40	61	40
Max travel time (minutes)	49	105	47	46	111	46
Min speed (mph)	49	23	51	53	21	53

*Average travel times on Monday, Tuesday and Wednesday (Feb., 18th~20th, 2013)

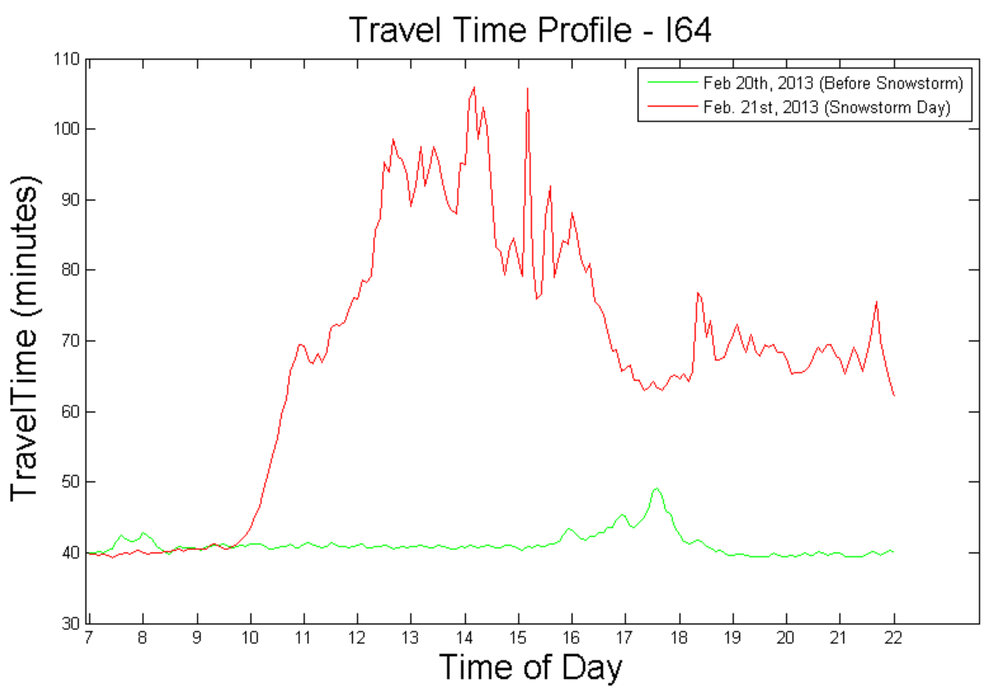


Figure 5-5: Comparisons of before-and-during snowstorm travel times on entire Length of I-64 (westbound)

Section 6.
Model Verification and Calibration

The estimated travel time results (using the instantaneous model) were compared with two sources of ground truth data**:** 1) Bluetooth-based travel time, and 2) vehicle-matching-based travel time.

6.1 Model Verification using Bluetooth-based Travel Time

Due to the unavailability of the Bluetooth-based travel time on I-64 (see Section 3 for more details), westbound and eastbound I-70 between SR-94 and I-270 were selected as a special case study to compare the results of the model with ground truth data. The study time period was the entire month of January 2013.

Figure 6-1 shows the locations of both the Bluetooth devices and RTMSs in the segment of I-70 between SR–94 and I-270 used in the study. Since the locations of Bluetooth devices and traffic sensors do not perfectly overlap, interpolation methods were used to estimate the travel times for the same segments using both the Bluetooth devices and the traffic sensors. The interpolation method is described by the small figure in the upper right corner of Figure 6-1.

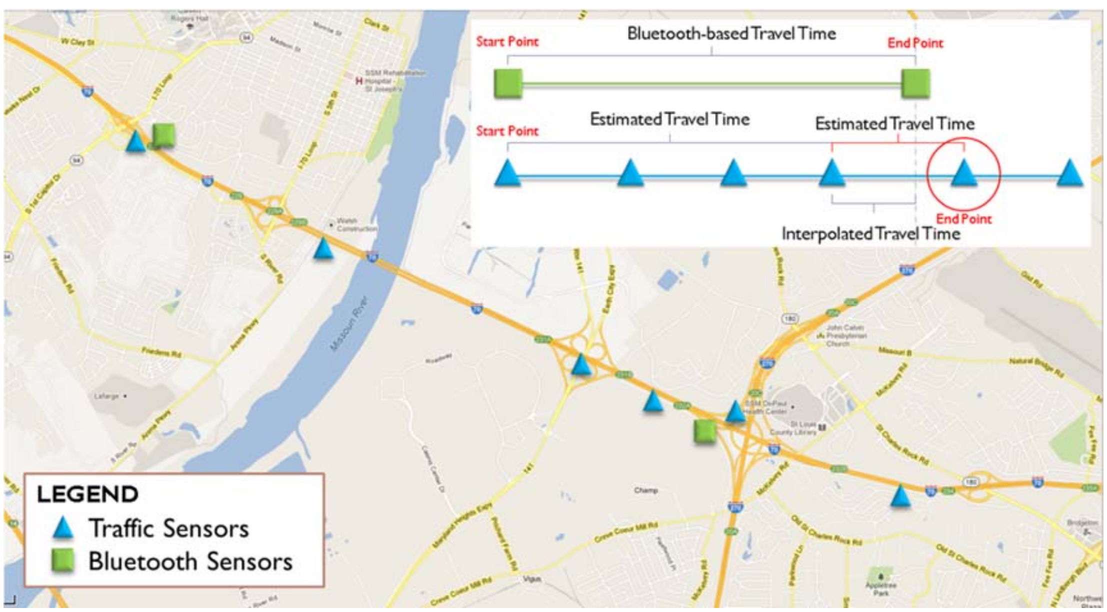


Figure 6-1: Locations of Bluetooth and point-speed sensors along the data collection segment of I-70 for travel time estimation result verification

Two measures of accuracy, Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE) are used in this study and defined as follows (Washington et al., 2010).

$$MAE = \frac{\sum_{t=1}^{n} \lvert F(t) - G(t) \rvert}{n}$$	(6-1)
$$MAPE = \frac{1}{n}\sum_{t=1}^{n} \left\lvert \frac{F(t) - G(t)}{G(t)} \right\rvert$$	(6-2)

where $G(t)$ is the ground truth data (in this case, the Bluetooth-based travel time) at a time interval t; $F(t)$ is the estimated travel time at time interval t; and n is the total number of samples.

MAE provides an overview of all the errors and shows the gaps between the estimated travel times and the collected travel times. MAPE, which shows the error as a percentage, is a scale independent measure of accuracy.

The Bluetooth-based travel time and the estimated travel time were quantitatively compared using data sets for each collected during January 2013. Figure 6.2 graphically depicts the comparison between the two types of travel times for the period 6:00 am to 9:00 pm on Jan 8th, 2013. The comparison results are summarized in Table 6-1. Both MAE and MAPE were calculated for both directions of I-70 for three types of time periods, namely weekdays, weekends, and the entire month of January. In the eastbound direction, MAE and MAPE were 0.22 minutes and 4.6%, respectively, on weekdays, while for the entire month they were 0.19 minutes and 4.1%, respectively. In the westbound direction, MAE and MAPE were 0.83 minutes and 12.3%, respectively, on weekdays and 0.73 minutes and 11.1%, respectively, for the entire month. These MAE and MAPE values indicate that the estimated travel time can be used effectively to represent the Bluetooth travel time, with fairly small gaps. It would therefore be reasonable for MoDOT to use the estimated travel time obtained using the data from the estimation model developed for this project to effectively evaluate freeway performance without the need to install additional Bluetooth-based travel time sensors. However, the estimated travel times from the model were generally lower than the Bluetooth-based travel times. This may have been because one Bluetooth detector was installed close to an intersection. Vehicle delays caused by the signalized intersection were also captured by the Bluetooth detector.

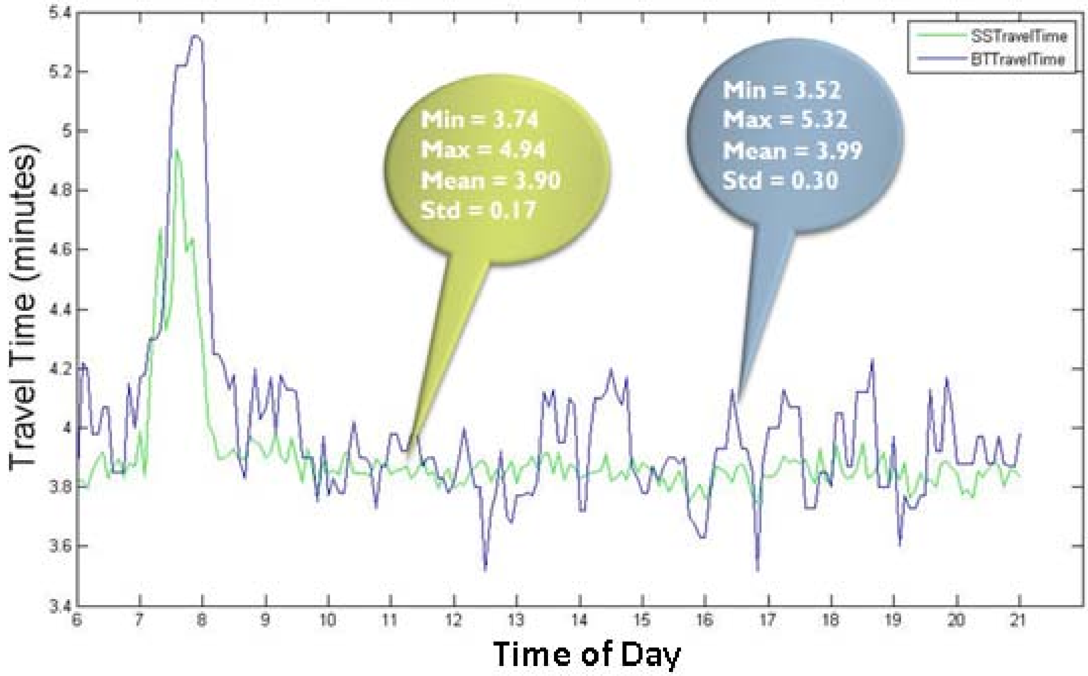


Figure 6-2: Comparison of Bluetooth-based and estimated travel times

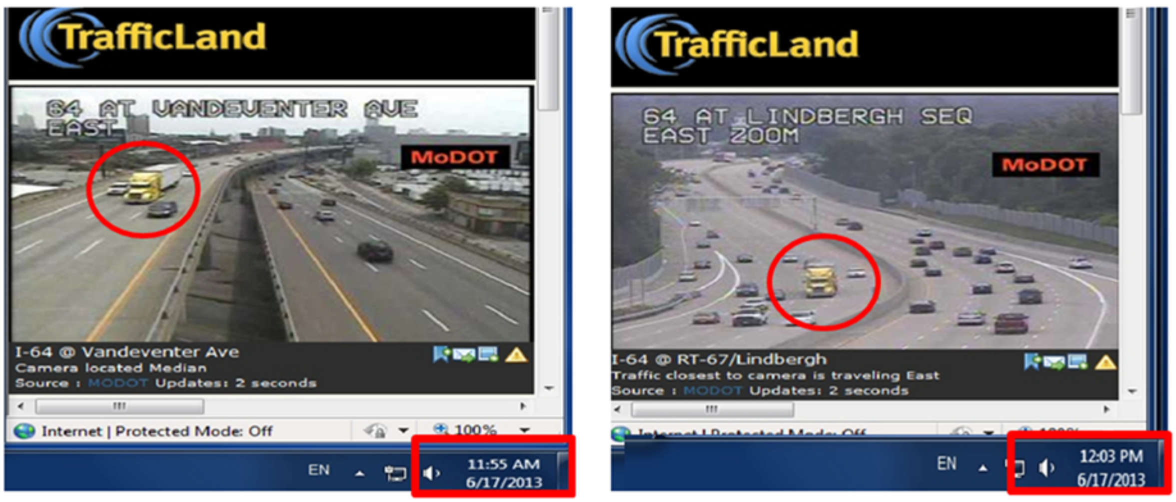
Table 6-1: Quantitative Comparison Between Bluetooth-based and Estimated Travel Times
(January 2013)

I - 70	**Eastbound**		**Westbound**	
Weekday	**MAE (min)**	**MAPE (%)**	**MAE (min)**	**MAPE (%)**
Weekdays	0.22	4.6	0.83	12.3
Weekends	0.11	2.8	0.45	7.6
Entire Month	0.19	4.1	0.73	11.1

6.2 Model Verification using Vehicle-matching-based Travel Time

A second verification exercise was deemed necessary for this project for two reasons: 1) The Bluetooth-based travel time may not be truly representative of the ground truth travel time, and 2) the I-70 corridor is outside the research scope of this project. The segment manual video-based travel time collection method described in Section 4 was therefore implemented to verify the results of the estimated travel time model using real-time traffic data and manual vehicle-matching.

The video stream was provided by TrafficLand.com. Figure 6-3 shows how the travel time was collected through vehicle matching of the vehicles shown by the surveillance cameras. For example, the yellow truck in the first image was observed at the Vandeventer Ave camera on I-64 at 11:55 am, June 17, 2013 (Figure 6-3a), and reappeared at the Lindbergh Blvd camera on I-64 eight minutes later, at 12:03 pm (Figure 6-3b). Since the distance between the two cameras locations is known, the travel time of the yellow truck for this segment can be estimated.



(a) **(b)**

Figure 6-3: Collecting travel time ground truth data through the Vehicle-matching-based technique.

(a) at Vandeventer Ave &I-64 and (b) at Lindbergh & I-64

An additional verification study on I-64 from Kingshighway to Bellevue for the period from 4 pm to 7 pm on 27th, June 2013 was conducted to examine the travel time discrepancy during rush hours. Due to the low resolution of the TrafficLand videos, identifying and capturing the same vehicle from two videos is a fairly time consuming task, so the resulting number of valid samples was relatively low, thus impacting the statistical analysis. Consequently, the SLU research team decided to record the necessary videos at the TMC on June, 27, 2013. Figure 6-4 shows the corridor used for this verification study.

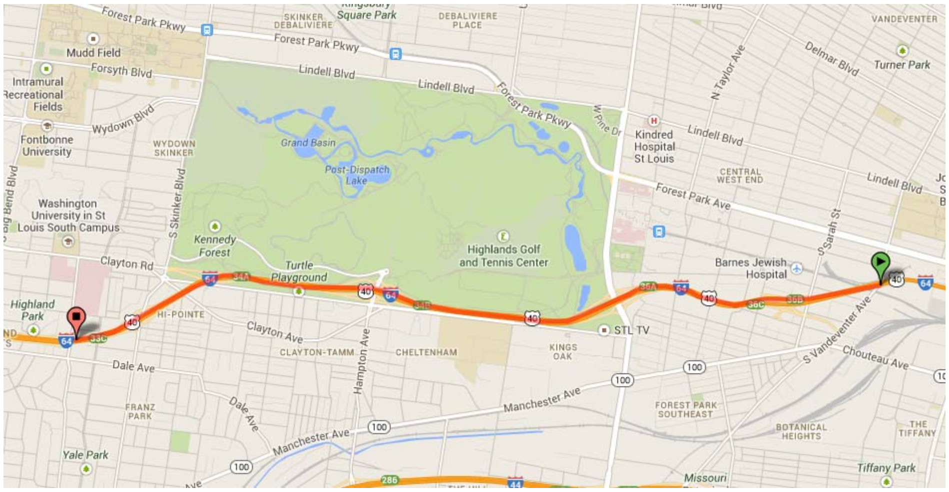


Figure 6-4: Location of MoDOT surveillance cameras along I-64
(Image obtained from Google Maps)

Figure 6-5 shows the two types of travel times, vehicle-matching-based travel time and estimated travel time. Consider that only trucks were tracked in the videos, and that truck speeds are usually lower than for passenger cars, the vehicle-matching-based travel time should be higher than the estimated travel time. However, the plots for the two travel times shown in the figure converged during the rush hours, indicating that the estimated travel times were overestimates.

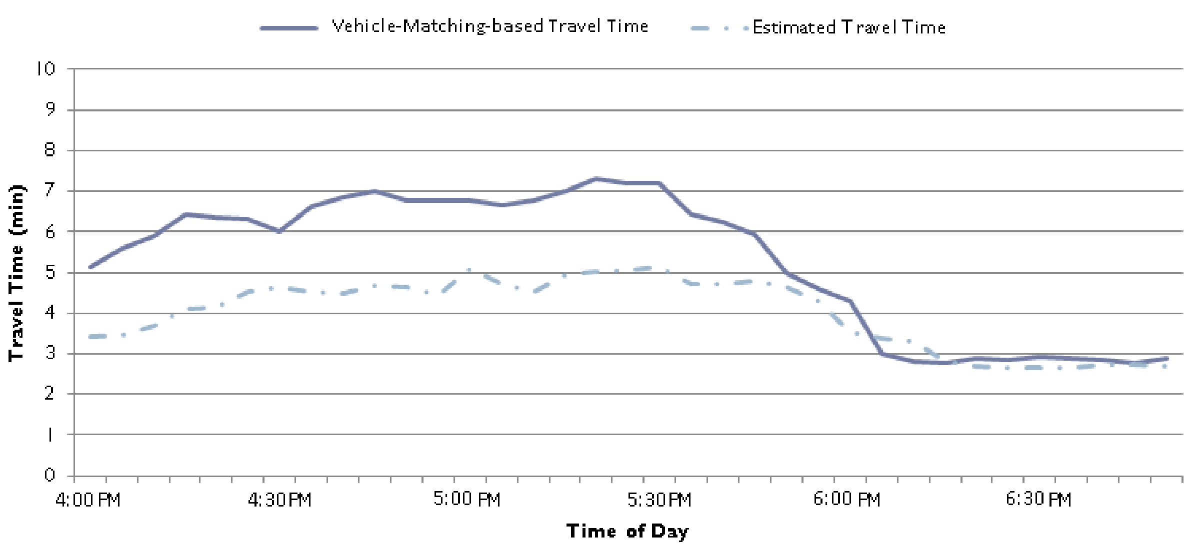


Figure 6-5: Comparison of Vehicle-matching-based travel times and estimated

travel times along the second case study corridor (Jun. 27th, 2013)

One possible reason for this overestimation was that any variations in the speed between the two traffic sensors is unknown. A linear travel time model was applied for the travel time estimation, with one of the underlying assumptions of the model being that vehicles travel at the same, uniform speed on segments. In fact, travel speed will vary based on both individual driver behavior and traffic conditions. Figure 6-6 shows an example explaining potential reasons why the speed may be underestimated. The dashed line shows the estimated speed curve (a linear plot). Since it is assumed to be linear, it is difficult to identify any bottlenecks that may occur between two sensors. In this case, a nonlinear model could be developed in future research to more accurately capture real traffic behavior such as bottlenecks.

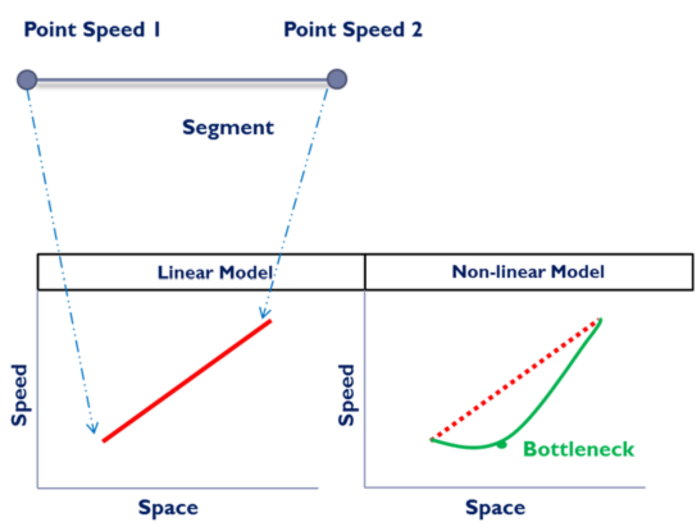


Figure 6-6: Linear model vs. non-linear model

Section7.

Conclusions and Recommendations

7.1 Conclusions

Travel time is widely believed to be a key performance measure for the evaluation of transportation systems. In the Greater St. Louis area, traffic data has been collected from fixed traffic sensors along the major freeways and arterial roads by the Transportation Management Center (TMC) operated by the Missouri Department of Transportation (MoDOT). This project sought to develop a new data handling model that will more fully utilize existing traffic sensor data to estimate travel times along the portion of I-64 that passes through the city. The following products were developed as a result of this project.

1. The traffic data has been archived in flat files in the XML format, and also parsed and stored in a DBMS physically located at SLU. This storage system was intentionally designed to allow easy and flexible sharing of the traffic data;

2. A new database schema was designed based on the traffic data characteristics and optimized according to the structure of the most commonly used SQL queries. The database designed was then used for estimating the travel time in real-time with much shorter response times for obtaining travel time information;

3. An appropriate travel time model was selected and implemented. In order to verify the travel times generated by the model, both Bluetooth-based travel time data and travel time data obtained through vehicle matching from video footage were used to test the accuracy of the traffic-sensor-based travel time estimation;

4. A custom Matlab-based prototype system and an Excel VBA-based tool were successfully developed. Based on feedback received from TMC staff and the superior flexibility of the MATLAB development environment, several additional functionalities were developed for the Matlab-based prototype system to allow it to handle both historical and real time traffic data, and publish the resulting travel time information. In addition to its basic functions, extra functions such as speed profiles were incorporated into the program. The Excel VBA-based tool was used only to process the historical data.

5. The travel time on I-64 was successfully estimated to an acceptable degree of accuracy utilizing traffic data collected from existing fixed traffic sensors along I-64.

7.2 Recommendations

The recommendations indicated by the above research results can be summarized as follows:

- The travel time estimation results show that the estimated travel time may be underestimated during rush hours. Other travel time estimation models might offer alternative ways to handle this issue. It is suggested the research team conduct a more comprehensive analysis to investigate this issue in the future.
- The data used for the model verification may not be sufficient. More data was required from both the Bluetooth devices and the video footage. Moreover, for convenience truck travel times were often used as ground truth information even though trucks generally travel more slowly than passenger cars, The potential solution to this conundrum could be to record high quality videos and hire additional student workers to “process” the videos.
- The user interfaces still have some room for improvement. More advanced techniques can be used in conjunction with other programming languages to improve the look and feel of the user interfaces.

Since the project has been extended to Phase 2 to cover travel time estimation for the entire freeway network in the Greater St. Louis region, the issues mentioned above will be addressed in Phase 2.

Acknowledgements

The authors would like to thank Tom Blair, Jeanne Olubogun, Greg Owens, Davar Divanbeigi and Cynthia Simmons at the TMC for technical support. We would also like to thank Rick Zygowicz, David Shackelford, Michael Bayer and Tiffany Rando at TransCore for providing real-time traffic data and technical support. We are grateful to William Stone, Jonathan Nelson and Andrew Hanks for the advice they have provided throughout. Special thanks to the Anis Dokic for video data processing.

References

Coifman, B. (2002). Estimating travel times and vehicle trajectories on freeways using dual loop detectors. *Transportation Research Part A: Policy and Practice*, *36*(4), 351–364.

Ehmke, J. F., Meisel, S., & Mattfeld, D. C. (2012). Floating car based travel times for city logistics. *Transportation Research Part C: Emerging Technologies*, *21*(1), 338–352.

Haghani, A., Hamedi, M., Sadabadi, K. F., Young, S., & Tarnoff, P. (2010). Data collection of freeway travel time ground truth with Bluetooth sensors. *Transportation Research Record: Journal of the Transportation Research Board*, *2160*, 60-68.

Klein, L. A. *Sensor technologies and data requirements for ITS*. Artech House, Boston, Mass., 2001.

Li, R., Rose, G., & Sarvi, M. (2006). Evaluation of speed-based travel time estimation models. *Journal of Transportation Engineering*, *132*(7), 540–547.

Ma, X., & Koutsopoulos, H. N. (2008). A new online travel time estimation approach using distorted automatic vehicle identification data., *ITSC 2008. 11th International IEEE Conference on Intelligent Transportation Systems*, Beijing, 12-15 October, pp. 204–209.

Ozbay, K., & Yildirimoglu, M. (2011). Comparison of real-time travel time estimation using two distinct approaches: Universal kriging and mathematical programming. *Intelligent Transportation Systems (ITSC), 2011 14th International IEEE Conference on* (pp. 1083–1088).

Sun, L., Yang, J., & Mahmassani, H. (2008). Travel time estimation based on piecewise truncated quadratic speed trajectory. *Transportation Research Part A: Policy and Practice*, *42*(1), 173–186. Turner, S. M., W. L. Eisele, W. L., Benz, R. J., & Holdener, D. J. (1998). *Travel time data collection handbook*. Report submitted to the Office of Highway Information Management, Federal Highway Administration, US Department of Transportation.

Washington, S. P., Karlaftis, M. G., & Mannering, F. L. (2010). *Statistical and econometric methods for transportation data analysis*, Chapman and Hall/CRC Press.

Wasson, J.S, Sturdevant, J.R, & Bullock, D.M (2008). Real-time travel time estimates using media access control address matching. *Institute of Transportation Engineers*, *78*(6), 20-23.

Final Report

Project Number TR201311

Freeway Travel Time Estimation using Existing Fixed Traffic Sensors – Phase 1

Prepared for
Missouri Department of Transportation
Traffic Management Center

By

Yao-Jan Wu
Assistant Professor, PhD, P.E.

Shu Yang
Graduate Research Assistant

Saber Abdoli
Graduate Research Assistant

Department of Civil Engineering
Parks College of Engineering, Aviation and Technology
Saint Louis University

August 2013

TECHNICAL REPORT STANDARD TITLE PAGE

1. REPORT NO.	2. GOVERNMENT ACCESSION NO.	3. RECIPIENT'S CATALOG NO.
cmr14-008		

4. TITLE AND SUBTITLE	5. REPORT DATE
Freeway Travel Time Estimation Using Existing Fixed Traffic Sensors – Phase 1	August 2013
	6. PERFORMING ORGANIZATION CODE
7. AUTHOR(S) Yao-Jan Wu, Shu Yang, and Saber Abdoli	8. PERFORMING ORGANIZATION REPORT NO.
9. PERFORMING ORGANIZATION NAME AND ADDRESS Smart Transportation Lab, Department of Civil Engineering Parks College of Engineering, Aviation and Technology Saint Louis University, 3450 Lindell Boulevard St. Louis, MO 603103	10. WORK UNIT NO. 11. CONTRACT OR GRANT NO. TR201311
12. SPONSORING AGENCY NAME AND ADDRESS Missouri Department of Transportation Research, Development and Technology P.O. Box 270 Jefferson City, MO 65102	13. TYPE OF REPORT AND PERIOD COVERED Final Report 14. SPONSORING AGENCY CODE

15. SUPPLEMENTARY NOTES

16. ABSTRACT

Freeway travel time is one of the most useful pieces of information for road users and an important measure of effectiveness (MOE) for traffic engineers and policy makers. In the Greater St. Louis area, Gateway Guide, the St. Louis Transportation Management Center (TMC) operated by the Missouri Department of Transportation (MoDOT), already collects traffic data from more than 700 fixed traffic sensors located along major freeways and arterials. Due to their significant investment in fixed freeway sensors, MoDOT has been striving to more fully utilize this data to extract useful information for stakeholders. One potential application is to convert the information collected by the fixed traffic sensors (point detectors) directly into travel times without the need to install additional sensors. This project developed a database specifically designed and optimized for the purpose of freeway performance measurement and implemented a data quality assurance procedure to examine the quality of the data collected. After a comprehensive literature review, the instantaneous travel time estimation model was selected and implemented in a stand-alone MATLAB-based system and a Microsoft Excel VBA-based tool. Several case studies were conducted to test the applicability of the system. The estimated travel time results were then verified using travel time data collected by a Bluetooth-based method and a video-based manual-vehicle-matching method. This Phase 1 project has proved the feasibility of the proposed approach. The travel time estimation method will be improved and the coverage of travel time estimation expanded to the city's entire freeway network in Phase 2.

17. KEY WORDS	18. DISTRIBUTION STATEMENT
Travel time estimation, traffic sensors, data collection, data quality control and database design.	No restrictions. This document is available to the public through the National Technical Information Service, Springfield, VA 22616

19. SECURITY CLASSIF. (of this report)	20. SECURITY CLASSIF. (of this page)	21. NO. OF PAGES	22. PRICE
None	None		

Table of Contents

Executive Summary 5

Section 1 Introduction 7

1.1 Project Background 7

1.2 Problem Statement 7

1.3 Research Objectives 8

Section 2 Literature Review 9

2.1 Travel Time Estimation 9

2.2 Ground Truth Data Collection 10

2.2.1 Conventional Techniques 10

2.2.2 Bluetooth-based Travel Time 11

Section 3. Data Collection 13

3.1 Traffic Data 13

3.1.1 Traffic Sensors 13

3.1.2 Archiving Data at Saint Louis University (SLU) 15

3.2 Bluetooth-based Travel Time Data 16

Section 4. Traffic Database Design and Implementation 18

4.1 Database Design 18

4.2 Database Optimization 18

4.3 Traffic Data Quality Assurance 19

4.3.1 Real-time Traffic Data Missing 20

4.3.2 Traffic Sensor Failure Rate 21

4.3.3 Abnormal Data 22

Section 5. Travel Time Estimation System.. 25
5.1 Model Selection 25
5.2 Implementation 26
5.2.1 Prototype system.. 26
5.2.2 Excel VBA-based Tool 27
5.3 Case Studies 29
Section 6. Model Verification and Calibration 35
6.1 Model Verification using Bluetooth-based Travel Time 35
6.2 Model Verification using Vehicle-matching-based Travel Time 37
Section7. Conclusions and Recommendations 41
7.1 Conclusions...41
7.2 Recommendations...42
Acknowledgements... 42
References... 43

Table of Figures

Figure 3-1: RTMS and their locations along freeways in the Greater St. Louis area 14

Figure 3-2: Traffic data collection flowchart 16

Figure 3-3: Bluetooth-based travel time data collection sites. The background image is from Google Earth. 17

Figure 4-1: Database Schema. 19

Figure 4-2: Daily data quality control 20

Figure 4-3: Abnormal Data Examples 24

Figure 5-1: The Interface of the Travel Time Estimation System 27

Figure 5-2: The Excel VBA-based tool for travel time estimation 29

Figure 5-3: Travel time estimation system 32

Figure 5-4: Case study of traffic bottleneck identification on I-64 33

Figure 5-5: Comparisons of before-and-during snowstorm travel times on entire Length of I-64 (westbound) 34

Figure 6-1: Locations of Bluetooth and point-speed sensors along the data collection segment of I-70 for travel time estimation result verification 35

Figure 6-2: Comparison of Bluetooth-based and estimated travel times 37

Figure 6-3: Collecting travel time ground truth data through the Vehicle-matching-based technique. 38

Figure 6-4: Location of MoDOT surveillance cameras along I-64 39

(image obtained from Google Maps) 39

Figure 6-5: Comparison of Vehicle-matching-based travel times and estimated travel times along the second case study corridor (Jun. 27th, 2013) 39

Figure 6-6: Linear model vs. non-linear model 40

Table of Tables

Table 4-1: Real-Time Data Missing Events and Corresponding Weather Conditions................. 21

Table 4-2: Traffic Sensor Failure Rate, August 2012.. 22

Table 5-1: Travel Time Comparisons Before, During and After the Snowstorm 34

Table 6-1: Quantitative Comparison Between Bluetooth-based and Estimated Travel Times (January 2013)... 37

Executive Summary

Freeway travel time is one of the most useful pieces of information for road users and an important measure of effectiveness (MOE) for traffic engineers and policy makers. In the Greater St. Louis area, Gateway Guide, the St. Louis Transportation Management Center (TMC) operated by the Missouri Department of Transportation (MoDOT), collects traffic data from more than 700 fixed traffic sensors along the city's major freeways and arterial roads. Due to their significant investment in fixed freeway sensors, MoDOT has been striving to more fully utilize the data collected to extract useful information for stakeholders. Making this data more available to other transportation agencies will also enable them to investigate traffic issues and take more effective action to address them.

This project was aimed at providing an automatic and systematic approach to estimating travel time on the section of Interstate 64 (I-64) located in the St. Louis area using existing fixed sensors. Travel time data used to be collected by designated sensors specifically designed to collect travel time data, but the coverage of these travel time sensors is usually limited. A successful outcome for this project would allow MoDOT to evaluate the performance of the entire freeway network at a very low or no additional cost by accessing the data already collected by the fixed traffic sensors. The database created for this project should also facilitate data exchange within and between the state's universities and transportation agencies.

The project began by conducting a comprehensive literature review regarding travel time estimation to explore two important aspects of travel time estimation, namely fixed-sensor-based travel time estimation and ground truth data collection methods. The literature review not only assisted in the selection process for the travel time estimation model to be implemented but also facilitated the verification process at the end of the project.

No travel time estimation system can be built without suitable data support. Once the project's dedicated data server had been set up at Saint Louis University (SLU), a custom computer program was developed to receive the data from MoDOT via the File

Transfer Protocol (FTP). The data is stored in the original Extensible Markup Language (XML) format as flat files in the SLU data server and also imported into the Database Management System (DBMS) developed for this project. To improve the query performance, the database was optimized based on the most frequently used queries. The custom computer program is also designed to monitor the traffic data quality, including the number of missing flat files, and automatically produces a daily data quality summary which is sent to the data server administrator.

Once the data support was firmly in place, the instantaneous model was implemented on two platforms: a custom MATLAB-based travel time estimation system and an Excel VBA-based travel time estimation tool. Only the former was used in the verification and case studies because of its high performance. Three case studies were conducted to demonstrate the applicability of the proposed system. These case studies consisted of: 1) a fundamental traffic analysis; 2) bottleneck identification; and 3) a snow storm impact study.

During the verification process, the performance of the estimated travel time obtained using the instantaneous model was compared with the times measured directly from the Bluetooth-based travel time data and video-based vehicle-matching-based travel time data. The results show that the travel time results tended to be slightly under-estimated due to errors associated with the linear interpolation. Overall, however, the performance results from the optimized database were satisfactory and facilitated an easy-to-understand visualization of the traffic data. Because of the high performance of this system and the proper data quality control procedures implemented, the system was deemed capable of providing fairly useful information for MoDOT and enabling quick traffic analyses.

Moving toward Phase 2, the travel time estimation method will be further improved to increase its accuracy and expanded to cover the entire freeway network in the St. Louis district. This expansion will allow traffic engineers to oversee the overall freeway network performance, rather than the performance of a single freeway segment.

Section 1

Introduction

This section provides an overview of the project background, problem statement and research objectives.

1.1 Project Background

Travel time is one of the key performance measures used to evaluate transportation systems, but travel time information is generally very difficult to collect. High-tech traffic sensors such as Global Positioning Systems (GPS) and manual/automatic vehicle identification using videos and Bluetooth are now being used to collect/ estimate travel times and recently these traffic sensors have begun to be widely deployed to monitor traffic flow in the U.S. (Klein, 2001) for many applications of Intelligent Transportation Systems (ITS). In many major cities in the U.S., traffic data are also being collected from fixed traffic sensors (e.g. inductance loop detectors embedded in road surfaces) and fed into a server located in the city's Transportation Management Center (TMC) in real time. As these fixed sensors are point detectors, the data collected consist only of vehicle spot speed and volume, but this type of "spot" data could not only be used to conduct a variety of traffic analyses but potentially also to estimate travel time without the need to install additional travel time data collectors. However, this type of advanced data application has rarely, if ever, been considered by practitioners.

1.2 Problem Statement

In the St. Louis area, traffic data are collected from more than 700 fixed traffic sensors located along major freeways and arterials by the Missouri Department of Transportation (MoDOT) TMC. The traffic flow data are fed into the TMC server in real time and used to support MoDOT's daily traffic operations. MoDOT is seeking ways to utilize the traffic data from these sensors more effectively, and this project is designed to extract useful information that can be applied by MoDOT traffic engineers and the city's policy makers to improve traffic flows and minimize congestion. It is also expected to help TMC staff deal more efficiently with the many data requests they receive from different agencies. At present, when a request is accepted staff must manually download the data, a very time-consuming process. It is therefore desirable to have an integrated data analysis platform that incorporates a sustainable data portal that other agencies can access to download the data they require without the need to send multiple data requests to the TMC.

1.3 Research Objectives

The major goal of this project was to develop an efficient traffic data platform to provide a research foundation for advanced research in transportation engineering. A travel time estimation method suitable for St. Louis transportation network was identified and a computer program was developed to automate the analysis process, thus facilitating freeway performance evaluation in the St. Louis area.

Specifically, this study had the following objectives:

- Design an efficient database schema based on the characteristics of the traffic data;
- Select and implement a travel time estimation model, and apply the model to traffic travelling along Interstate 64 (I-64);
- Implement the travel time estimation analysis tool using the selected travel time estimation model;
- Integrate the tool into the new database to improve query performance;
- Compare the estimated travel time results with the results of other travel time estimation methods

Section 2

Literature Review

This study was designed to develop a way to estimate travel times using existing traffic sensors on freeways. A comprehensive literature review was therefore conducted to investigate existing methods of travel time estimation. In order to verify the performance of the proposed travel time estimation method, ground truth travel time collection methods were also reviewed.

2.1 Travel Time Estimation

A number of different data sources and techniques, such as automated vehicle identification (Ma and Koutsopoulos, 2008), floating car data (Ehmke et al., 2012) and electronic toll collection systems (Ozbay and Yildirimoglu, 2011), can be applied to estimate travel times. Since this project sought to utilize existing traffic sensor data to estimate travel times on freeways, those travel time estimation models based on fixed traffic sensors were the primary focus of this literature review. For the purposes of this research, a fixed sensor is defined as a sensor that can only collect traffic data such as speed and volume at a particular spot.

The instantaneous model (Li et al., 2006) is widely used to estimate travel times. It uses the average of upstream and downstream speed data collected by traffic sensors to calculate the average link speed and the length of the link divided by the average link speed is then the estimated travel time. The total travel time on a segment is calculated by summing the link travel times. However, Li et al. (2006) pointed out that the error inherent in travel time estimation by the use of the instantaneous model can be substantial as it depends on the speed at which vehicles are travelling and this can vary considerably over relatively long link lengths. In spite of their acknowledgement of this problem, the group did not investigate the relationship between the link length and the error in travel time estimation.

Other sensor-based travel time estimation models are based on vehicle trajectory construction using point speed data. For example, the time slice model (Li et al, 2006) is based on the instantaneous model, but unlike the instantaneous model, it captures the speed at which a vehicle is moving when it enters each link and then uses that speed data to calculate the travel times on individual links. Assuming all the vehicles travel at the same speed on a link, traffic flow theory is then applied to the speed data to construct a pseudo vehicle trajectory on segments (Coifman, 2002). He concluded that the performance of the proposed model during non-rush hours, when traffic is flowing more smoothly, was better than during rush hours. A piecewise truncated quadratic speed trajectory has also been proposed to simulate vehicle trajectory on a link (Sun et al., 2008). Coifman (2002) concluded that as his proposed method required only speed data, it would be easy to implement for online real time travel time estimation application.

2.2 Ground Truth Data Collection

Many techniques have been suggested to assist both researchers and practitioners to collect ground truth travel time data. This sub-section provides a literature review of research in this area.

2.2.1 Conventional Techniques

The *Travel Time Data Collection Handbook* (Turner et al., 1998) provides an overview of most of the travel time collection techniques that have been used to date. The four basic approaches described in the *Handbook* are as follows:

- Test vehicles
- License plate matching
- Intelligent Transportation Systems probe vehicles
- Emerging and non-traditional techniques

Test vehicles, or "active test vehicles", have been used for travel time data collection since the late 1920s. Although several methods can be used for this technique, its basic principle is for a test vehicle to travel along a segment and record the time stamps as it passes predefined checkpoints on the segment. This technique, however, can result in errors caused by both human and electronic devices.

License plate matching techniques require license plate information to be collected from two or more sites. The travel time can then be directly obtained from the times at which the vehicles with those license plates pass each site. However, the major disadvantage of this approach is that the sample size is usually very limited due to the high data collection costs.

ITS probe vehicle techniques, or "passive probe vehicle" techniques, are used to collect travel time in real time. Automatic Vehicle Location (AVL) and GPS are both classified as ITS probe vehicle techniques. Here, the data collection cost can be high if more data samples are required and the travel time data collection process can be tedious.

Emerging and non-traditional techniques include the use of weight-in-motion stations, video cameras, and electronic toll collection (ETC), among others. However, these techniques require more advanced algorithms and models if they are to be used to estimate travel time.

2.2.2 Bluetooth-based Travel Time

Recently, travel time measurements using Bluetooth (Wasson et al, 2008) have become popular due to the widespread use of Bluetooth devices in our daily lives. Bluetooth-based travel time collection is a new technique that utilizes enabled Bluetooth portable devices such as mobile phones, computers, personal digital assistants, and car radios to identify specific vehicles at downstream and upstream locations by tracking their unique 48-bit Machine Access Control (MAC) addresses. Figure 2-1 shows how the travel time can be "calculated" by matching Bluetooth MAC addresses at consecutive detection locations along the road according to the time stamps associated with those MAC addresses. Bluetooth-based travel time data was used in this project to provide the ground truth travel times.

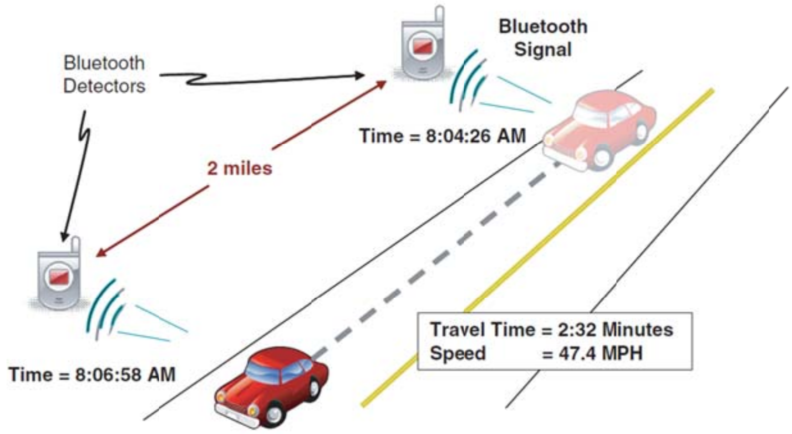


Figure 2-1: Bluetooth-based travel time measurement (Haghani et al., 2010)

Section 3.
Data Collection

This section summarizes the collection procedure and the results of the traffic data. The Bluetooth-based travel time data collected for the verification process is also introduced.

3.1 Traffic Data

3.1.1 Traffic Sensors

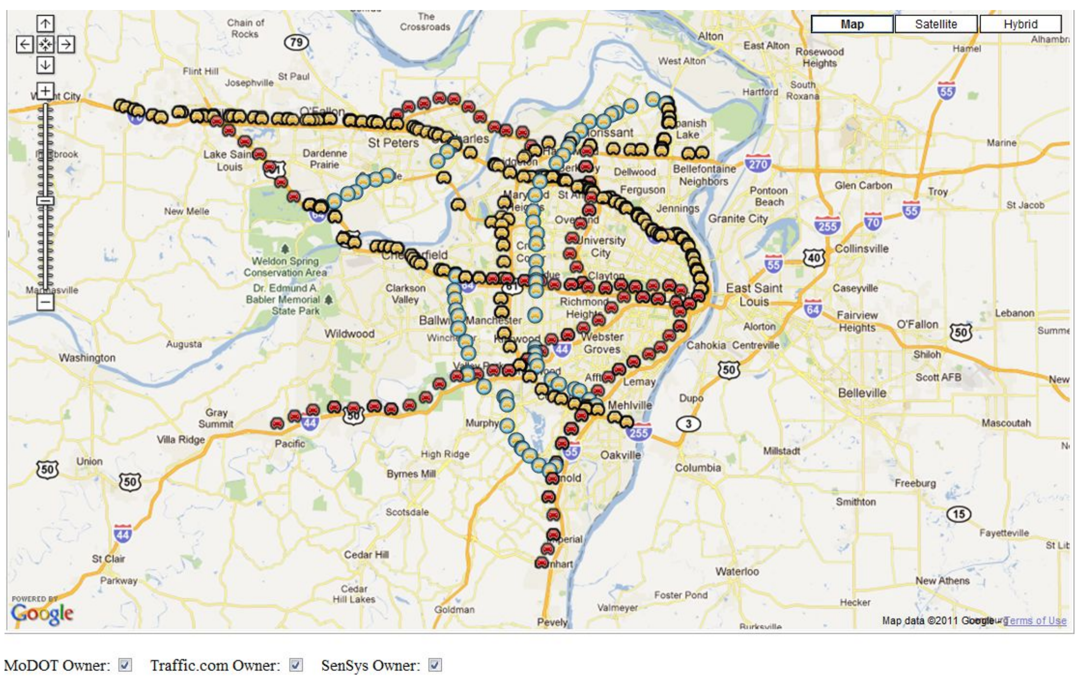
The data used to estimate travel time on the interstate was collected from Remote Traffic Microwave Sensors (RTMS®), a type of ITS traffic sensor, that have been installed along the major freeways in the Greater St. Louis area for traffic data collection. These sensors transmit a low-power microwave signal of constantly varying frequency in a fixed fan-shaped beam. The beam "paints" a long elliptical footprint on the road surface. Any non-background targets will reflect the signal back to the RTMS, where the targets are detected and their range measured.

RTMSs are used to collect traffic flow data, including traffic volume, speed, occupancy[1] and vehicle length, during a user defined time period. This time period is set at 30 seconds for the real-time feed sent to the data server located in MoDOT's TMC. Two examples of the RTMSs deployed by the MoDOT are shown in Figures 3-1(a) and (b) and the locations of the RTMSs in St. Louis are depicted in Figure 3-1(c). Each icon represents a RTMS sensor monitoring all of the lanes of the freeway for both directions. For example, the RTMS shown in Figure 3-1 (a) monitors three lanes of westbound I-64 and three lanes of eastbound I-64. Three fundamental traffic parameters are collected from the RTMSs for each lane: aggregated volume, average speed and average occupancy every 30 seconds. Information on the vehicle classification by vehicle length is also currently collected by the RTMSs but is not stored in the data server.

[1] "Occupancy" is defined as the percentage of time the sensor detection area is occupied by vehicles during a specific time period

(a)

(b)



(c)

Figure 3-1: RTMS and their locations along freeways in the Greater St. Louis area

(a) the RTMS located at milepost 29.8 on I-64 westbound; (b) the RTMS located at milepost 28.6 on I-64 westbound; (c) locations of the RTMSs in St. Louis. (a) and (b) are from Google Street View and the background image in (c) is from Google Maps.

3.1.2 Archiving Data at Saint Louis University (SLU)

The MoDOT server receives traffic data from all of the existing ITS sensors in St. Louis and generates two Extensible Markup Language (XML) files every 30 seconds from the system, consisting of: 1) Real-time traffic data that contains fundamental traffic parameters (e.g. volume, speed, and occupancy), and 2) the meta data storing basic information for each ITS sensor (e.g. location, number of lane being detected). Both the real time traffic data and the meta data are transferred to a File Transfer Protocol (FTP) server that has been physically located at the Smart Transportation Lab at Saint Louis University (SLU) since June, 2012. More than one gigabyte of real time traffic data is pushed from the MoDOT server to the SLU server via FTP every day. To increase data redundancy, the data is stored in the local SLU server, the shared drive and the MS SQL database. Figure 3-2 illustrates an overview of the traffic data collection flow.

Since the flat files are overwritten every 30 seconds on the SLU FTP server, in order to archive the two types of flat files, a custom C# computer program was developed to perform the tasks listed below:

- Automatically monitor changes in the flat files. When the flat files are overwritten, a signal will be sent to the computer program;
- Transfer the files from the SLU FTP server to both of the local and remote storage systems if the files change;
- Rename the files based on the time stamp in the flat files;
- Parse the transferred flat files to obtain the traffic information;
- Tabulate the parsed traffic information;
- Store the tabulated traffic information in the database (more details can be found in Section 4: Traffic Database);

- Generate a daily data quality summary and send it to the SLU team (more details of this process can also be found in Section 4).

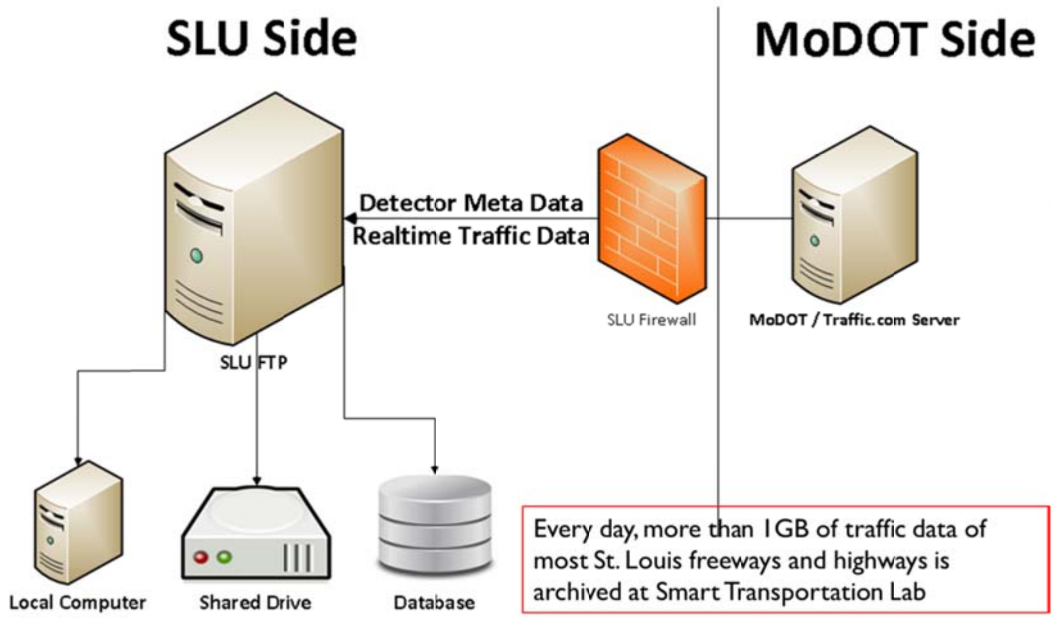


Figure 3-2: Traffic data collection flowchart

In addition to the real-time data, the historical traffic data was also successfully migrated from the MoDOT server to the SLU traffic data server to extend the data coverage back to Jan. 1st, 2008. This data provides additional information for future applications and analyses.

3.2 Bluetooth-based Travel Time Data

As discussed in Section 2, Bluetooth-based travel time data was used as the ground truth data for this project. The Bluetooth-based data was collected by Bluetooth Travel-time Origination and Destination (BlueTOAD™), operated by the MoDOT.

Four sets of Bluetooth-based travel time data were provided by the Transportation Management Center (TMC) in St. Louis. The data sets were collected during the period November 2012 to March 2013. The roadway segments are listed below:

- Segments on US 67 Highway from New Florissant to New Halls Ferry;

- Segments on I-70, between SR-94 and I-270, both eastbound and westbound; and
- Roadway segment on I-64 between I-55 and Ewing Ave.

Figure 3-3 shows the roadway segments where the Bluetooth travel time data was collected. According to the scope of this project, only the freeway data was to be used. However, the Bluetooth-based travel time collection segment on I-64 only partially overlaps the fixed sensors. Two countermeasures were taken: 1) an alternative Interstate 70 dataset was used to verify the estimated travel time and was thus considered an additional case study for this project, and 2) additional ground truth data was manually collected for I-64 since this was the primary focus of this project. Details of the ground truth data collection process will be provided in Section 6.

Figure 3-3: Bluetooth-based travel time data collection sites. The background image is from Google Earth.

Section 4.
Traffic Database Design and Implementation

This section describes the traffic database design and implementation, focusing particularly on the database schema design and optimization.

4.1 Database Design

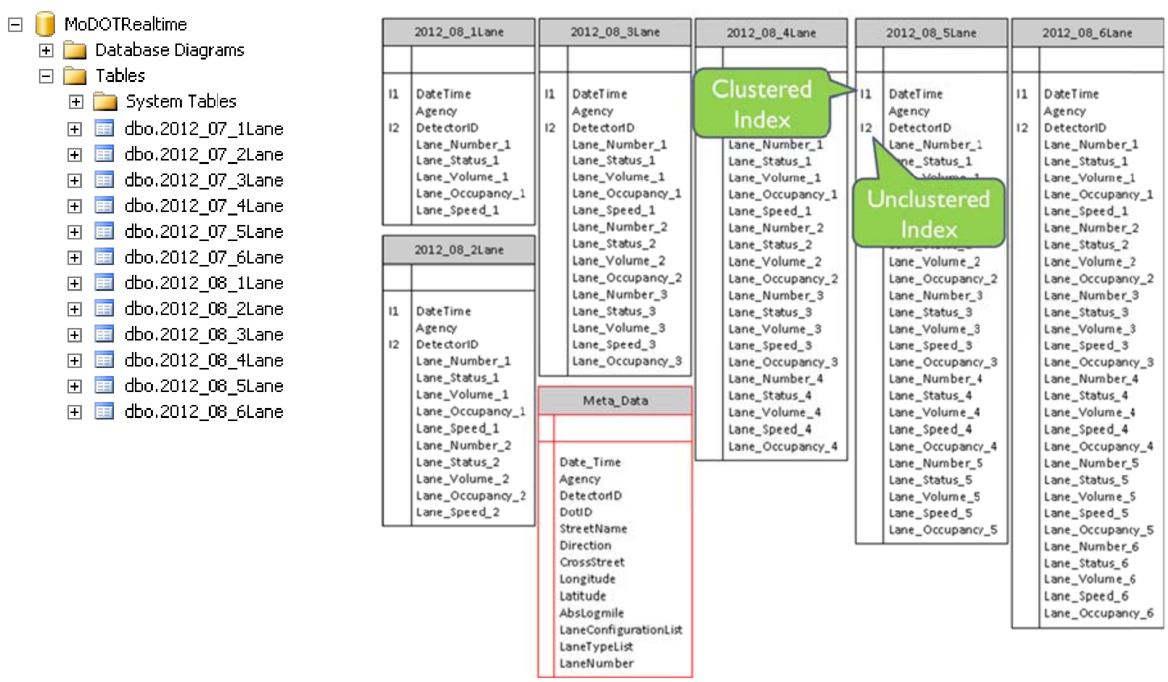
The efficiency of the travel time estimation calculations requires strong support from efficient database operations, so the SLU traffic database (Figure 4-1) needed to be re-designed in order to achieve a satisfactory performance. Microsoft SQL Server 2008 was used to build the new traffic database management system (DBMS).

Figure 4-1 shows a schematic diagram of the data tables and the new database designed for this project. All the data collected from a RTMS with the same lane configuration is grouped into one specific table. . For example, for all RTMSs monitoring three lanes on I-64 both westbound and eastbound on July, 2012, the traffic data collected from these RTMSs would be stored in the table “2012-07-3Lane”. , Since the highest number of lanes per direction is six, six data tables are accordingly created for each month. In addition to the real time traffic data collected, the metadata is exclusively stored in the table “Meta_Data”.

4.2 Database Optimization

The database optimization design is based on the most frequently used SQL queries. In order to calculate the travel time on either a specific segment of I-64 or for the entire length of I-64, the speed information is extracted for a given combination of a specific time period and the IDs of consecutive RTMSs from upstream through to downstream. In an SQL query, the performance of the clause “order by”, which is used to sort the records in a data table by key words, mainly determines the response time from a DBMS. In the most frequently used SQL queries, the RTMS must be spatially sorted from upstream to downstream. Microsoft SQL Server 2008 requires that only one clustered index can be created per data table, so a clustered index based on the attribute “DateTime” is created for each table. In addition to the clustered index, an unclustered index is created on the attribute “DetectorID” in order to improve the performance of the clause “where” in the SQL queries.

The SQL queries, which are most commonly used to extract the essential data for estimating travel time, were tested to examine the database performance enhancement. Before these two indexes were created, an SQL query took 50 seconds to return the corresponding results, but once the indexes had been created, the response time dropped to just 1 second.



(a) **(b)**

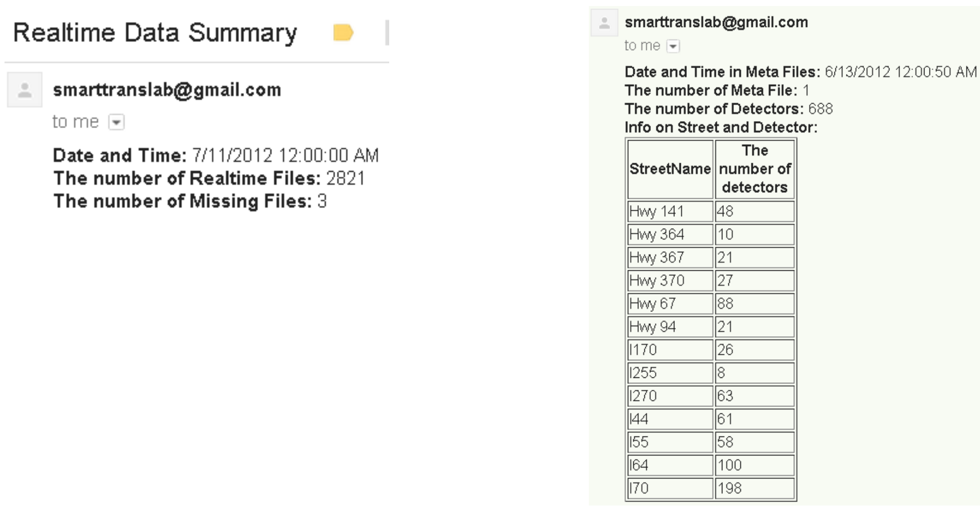
Figure 4-1: Database Schema.

(a) Examples of data tables, (b) Database schema

4.3 Traffic Data Quality Assurance

Traffic data quality assurance was conducted to assess the data quality, data accuracy, data completeness and data reliability of the database. A custom C# computer program was developed to fulfill this task.

The new program monitors the real time traffic data and automatically sends a mail notification to the administrator(s) if an abnormal situation is detected. At the end of each day, the daily metadata quality summary and the daily real time traffic data quality summary are also sent to the system administrators. The basic daily summaries of the traffic data are intended to provide a preliminary evaluation of the traffic data quality. Figure 4-2 shows examples of the daily real time data summary and the daily metadata summary.



Realtime Data Summary

smarttranslab@gmail.com
to me

Date and Time: 7/11/2012 12:00:00 AM
The number of Realtime Files: 2821
The number of Missing Files: 3

smarttranslab@gmail.com
to me

Date and Time in Meta Files: 6/13/2012 12:00:50 AM
The number of Meta File: 1
The number of Detectors: 688
Info on Street and Detector:

StreetName	The number of detectors
Hwy 141	48
Hwy 364	10
Hwy 367	21
Hwy 370	27
Hwy 67	88
Hwy 94	21
I170	26
I255	8
I270	63
I44	61
I55	58
I64	100
I70	198

(a) (b)

Figure 4-2: Daily data quality control

(a) Daily real-time data quality summary (b) Daily metadata quality summary

4.3.1 Real-time Traffic Data Missing

Due to network communication issues, the update frequency of the real time traffic data feed occasionally becomes 31 seconds rather than 30 seconds. Thus, the total number of daily real-time traffic data feeds theoretically ranges from 2,788 to 2,880. The number of daily existing files can be easily counted and the number of missing real time traffic data files calculated accordingly.

The research team also found an interesting correlation between the number of missing real time traffic data files and local weather conditions. It was found that inclement weather events were associated with days with a higher number of dropped data feeds. A daily summary of weather conditions is also recorded in a log file, based on weather information provided by wunderground.com. Table 4-1 lists all the days on which there were five or more missing files for the period from July 2012 to Oct. 2012. Therefore, a closer investigation was necessary to determine why inclement weather conditions have a negative impact on data transfer.

Table 4-1: Real-Time Data Missing Events and Corresponding Weather Conditions

Date	# of existing files	# of missing files	Weather conditions
08/09, 2012	2749	77	Wind Dir: WNW Max Wind Speed: 20 mph Max Gust Speed: 29 mph Rain, Thunderstorm
08/21, 2012	455	--*	Wind Direction: ESE Max Wind Speed: 9mph Max Gust Speed: - mph Rain
09/07, 2012	2819	7	Wind Direction: NW Max Wind Speed: 30mph Max Gust Speed: 38mph Rain, Thunderstorm
09/16, 2012	2799	31	Wind Direction: SE Max Wind Speed: 8mph Max Gust Speed: 12 mph ----
10/03, 2012	2653	183	Wind Direction: SSW Max Wind Speed: 14mph Max Gust Speed: 17mph Rain
10/13, 2012	2768	70	Wind Direction: S Max Wind Speed: 23mph Max Gust Speed: 29 mph ----
10/16, 2012	2430	415	Wind Direction: SSE Max Wind Speed: 21mph Max Gust Speed: 29mph ----
10/22, 2012	1762	1078	Wind Direction: SSE Max Wind Speed: 18mph Max Gust Speed: 24 mph Rain, Thunderstorm
10/23, 2012	1555	1301	Wind Direction: S Max Wind Speed: 17mph Max Gust Speed: 22mph Rain, Thunderstorm

--*: too many files were missing

4.3.2 Traffic Sensor Failure Rate

The traffic sensor working status is also indicated in one attribute, lane status, of the real-time data feed, with either "OK", "Failed" or "Disabled" being displayed in the attribute. When the “lane status” of an RTMS is not shown as “OK”, the RTMS is considered to be malfunctioning. The RTMS failure rate can be calculated as the number of sensors labeled “Failed” or “Disabled” divided by the number deemed “OK”. Table 4-2 gives an example of the real time traffic data failure rate during August 2012. Note that the freeways in St. Louis have at least two lanes each direction; RTMSs installed on ramps monitor only one lane. Since the project focused on freeways, Table 4-2 excludes the failure rate of traffic sensor on ramps. Overall, the real time traffic data failure rate was 3.94% for the month. This low percentage indicates the overall good data quality achieved.

Table 4-2: Traffic Sensor Failure Rate, August 2012

# of Lane	# of Total Records	# of Failure Records	Percentage
2	9,451,854	457,004	4.84%
3	15,421,446	699,049	4.53%
4	10,861,341	318,879	2.94%
5	2,818,974	51,296	1.82%
6	165,822	37	0.02%
total:	*387,194,37*	*1,526,265*	*3.94%*

One possible reason for the low incidence of real-time traffic sensor failures may be MoDOT’s good maintenance program for the traffic sensors. Network communication malfunctions may also adversely affect data completeness. Those data labeled “Failed” or “Disabled” are currently not used in estimating travel times.

4.3.3 Abnormal Data

In some situations, the data is considered “abnormal”. This normally arises in one of two ways, categorized as follows:

1) Type 1: When the lane status is ‘OK’, the traffic flow data (volume, occupancy and speed) is unavailable; and

2) Type 2: When the values of volume, occupancy and speed are incompatible.

Figure 4-3 shows some examples of abnormal data. For example, it is not reasonable that vehicle speeds of greater than 90 mph or 'occupancy' levels of more than 90% would be achieved during rush hours. It is also not reasonable to see volume and speed values of zero during rush hours.

The possible causes of the abnormal data are as follows:

1. The network communication infrequently fails (Fig.4-3a);
2. Incidents may account for inconsistent travel speeds in each lane (Fig. 4-3b); and
3. When the traffic sensor detection zone is occupied by vehicles for longer than 30 seconds (e.g. a crashed or broken down vehicle), the resulting "occupancy" value is 100 and the volume is 0 (Fig. 4-3c).

The speed information is essential for the travel time estimation model. In order to remove abnormal speed data and "Failed" (or "Disabled") data mentioned above, the median speed readings of all the lanes are used to represent traffic speed on the freeways.

	Date_Time	Agency	DetectorID	Lane_Number_1	Lane_Status_1	Lane_Volume_1	Lane_Occupancy_1	Lane_Speed
1	2012-08-01 00:00:25.000	STLATMS	MR067N013.0D	1	OK	-1	-1	-1
2	2012-08-01 00:00:25.000	STLATMS	MR067S013.0D	1	OK	-1	-1	-1
3	2012-08-01 00:00:25.000	STLATMS	MR067N015.8D	1	H Failed	-1	-1	-1
4	2012-08-01 00:00:25.000	STLATMS	MR067S015.8D	1	H Failed	-1	-1	-1
5	2012-08-01 00:00:56.000	STLATMS	MR067N013.0D	1	OK	-1	-1	-1
6	2012-08-01 00:00:56.000	STLATMS	MR067S013.0D	1	OK	-1	-1	-1
7	2012-08-01 00:00:56.000	STLATMS	MR067N015.8D	1	H Failed	-1	-1	-1
8	2012-08-01 00:00:56.000	STLATMS	MR067S015.8D	1	H Failed	-1	-1	-1
9	2012-08-01 00:01:27.000	STLATMS	MR067N013.0D	1	OK	-1	-1	-1
10	2012-08-01 00:01:27.000	STLATMS	MR067S013.0D	1	OK	-1	-1	-1
11	2012-08-01 00:01:27.000	STLATMS	MR067N015.8D	1	H Failed	-1	-1	-1
12	2012-08-01 00:01:27.000	STLATMS	MR067S015.8D	1	H Failed	-1	-1	-1
13	2012-08-01 00:01:58.000	STLATMS	MR067N013.0D	1	OK	-1	-1	-1
14	2012-08-01 00:01:58.000	STLATMS	MR067S013.0D	1	OK	-1	-1	-1

(a)

	Date_Time	DetectorID	Lane_Volume_1	Lane_Occupancy_1	Lane_Speed_1	Lane_Volume_2	Lane_Occupancy_2	Lane_Speed_2
2...	2012-08-12 17:12:31.000	MI064W010.0D	9	1	37	10	2	99
2...	2012-08-12 17:13:02.000	MI064W010.0D	4	1	37	5	5	99
2...	2012-08-12 17:13:32.000	MI064W010.0D	6	1	37	8	3	94
2...	2012-08-12 17:14:03.000	MI064W010.0D	9	1	37	9	4	98
2...	2012-08-12 17:14:33.000	MI064W010.0D	2	1	37	9	2	98
2...	2012-08-12 17:15:04.000	MI064W010.0D	5	1	37	7	2	93
2...	2012-08-12 17:15:34.000	MI064W010.0D	3	1	37	7	3	93
2...	2012-08-12 17:16:04.000	MI064W010.0D	5	1	37	10	3	91
2...	2012-08-12 17:16:35.000	MI064W010.0D	4	1	36	7	2	91
2...	2012-08-12 17:17:05.000	MI064W010.0D	6	1	36	8	5	91
2...	2012-08-12 17:17:35.000	MI064W010.0D	4	1	36	4	1	91
2...	2012-08-12 17:18:06.000	MI064W010.0D	11	2	36	9	2	94
2...	2012-08-12 17:18:36.000	MI064W010.0D	1	1	36	3	2	94
2...	2012-08-12 17:19:07.000	MI064W010.0D	4	1	36	4	1	94
2...	2012-08-12 17:19:37.000	MI064W010.0D	3	1	36	5	1	92
2...	2012-08-12 17:20:07.000	MI064W010.0D	4	1	36	5	3	92
2...	2012-08-12 17:20:37.000	MI064W010.0D	5	1	36	7	2	92
2...	2012-08-12 17:21:08.000	MI064W010.0D	3	1	39	4	1	92
2...	2012-08-12 17:23:09.000	MI064W010.0D	10	1	39	8	2	91
2...	2012-08-12 17:23:40.000	MI064W010.0D	3	1	39	4	1	91

(b)

	Date_Time	DetectorID	Lane_Volume_1	Lane_Occupancy_1	Lane_Speed_1	Lane_Volume_2	Lane_Occupancy_2	Lane_Speed_2	Lane_Volume_3	Lane_Occupar
22	2012-08-05 02:44:37.000	MI044W288.2U	5	7	29	7	93	13	1	1
23	2012-08-05 02:54:15.000	MI044W288.2U	7	16	33	9	100	12	0	0
24	2012-08-05 02:54:45.000	MI044W288.2U	7	16	33	9	100	12	0	0
25	2012-08-13 02:33:22.000	MI044W288.2U	6	6	42	10	90	14	3	2
26	2012-08-13 02:33:53.000	MI044W288.2U	6	6	42	10	90	14	3	2
27	2012-08-02 02:49:49.000	MI270N016.6D	0	0	0	0	100	0	0	100
28	2012-08-02 02:50:19.000	MI270N016.6D	0	0	0	0	100	0	0	100
29	2012-08-02 02:50:50.000	MI270N016.6D	0	0	0	0	100	0	0	100
30	2012-08-02 02:51:20.000	MI270N016.6D	0	0	0	0	100	0	0	100
31	2012-08-04 04:53:31.000	MI064E035.7U	6	9	26	11	99	10	1	2
32	2012-08-04 04:54:01.000	MI064E035.7U	6	9	26	11	99	10	1	2
33	2012-08-04 04:50:26.000	MI064E035.7U	6	10	29	7	100	14	0	0
34	2012-08-04 04:50:57.000	MI064E035.7U	6	10	29	7	100	14	0	0
35	2012-08-04 05:31:14.000	MI064E035.7U	5	19	12	10	94	7	1	1
36	2012-08-04 05:31:45.000	MI064E035.7U	5	19	12	10	94	7	1	1
37	2012-08-04 05:34:17.000	MI064E035.7U	6	14	9	10	100	6	4	4
38	2012-08-04 05:34:48.000	MI064E035.7U	6	14	9	10	100	6	4	4
39	2012-08-04 05:38:22.000	MI064E035.7U	11	24	6	8	94	9	3	3
40	2012-08-04 05:38:53.000	MI064E035.7U	11	24	6	8	94	9	3	3
41	2012-08-04 05:39:24.000	MI064E035.7U	11	46	10	13	100	10	2	2

(c)

Figure 4-3: Abnormal Data Examples

(a) Failure of data acquisition when the sensor status is OK, (b) Unreasonable speed data, and (c) Unreasonable volume and occupancy data

Section 5.

Travel Time Estimation System

This section presents the travel time estimation model selected for this project, the implemented system and three case studies that demonstrate the feasibility and applications of the proposed system.

5.1 Model Selection

The literature review described in Section 2 identified several models that have been specifically developed for travel time estimation using fixed point sensors. Based on the findings reported by Li et al. (2006), the instantaneous model and the time slice model were deemed the most suitable candidates for this implementation.

The instantaneous model uses real-time speed data from the upstream and downstream sensors of each link at time *k*. The link travel time can then be calculated through dividing the link length by the average of the collected speed data as formulated in Equation (1):

$$T(i,k) = \frac{2l_i}{v(i_u,k) + v(i_d,k)} \quad (1)$$

where $v(i_u,k)$ and $v(i_d,k)$ are the measured speeds at the upstream and downstream end points of link *i* at a time *k*; l_i represent the length of the link *i*, and $T(i,k)$ is the link travel time. Accordingly, the total travel time $T(k)$ for a vehicle beginning its trip at time *k* is the summation of the estimated travel time of *n* links:

$$T(k) = \sum_{i=1}^{n} t(i,k) \quad (2)$$

Unlike the instantaneous model, the time slice model attempts to account for variations in speed over time by constructing a vehicle "trajectory" using downstream speed values. Assuming the trip start time is *k*, the first link travel time is calculated similarly to the instantaneous model in Equation (1) and denoted as *t*(1, *k*), so the arriving time at the second link travel time can be expressed as *k* + *t*(1, *k*). Therefore, the travel time on the second link is $t(2,t_2) = \frac{2*l_2}{v[2_a,k+t(1,k)]+v[2_b,k+t(1,k)]}$. Generally, the travel time on link *n* can be written as Equation (3):

$$T(n,t_n) = \frac{2l_n}{v(n_u,t_n) + v(n_d,t_n)} \quad (3)$$

where $t_n = k + t(1,k) + \sum_{i=2}^{n-1} t(i,\ t_i)$. As in the instantaneous model, the total travel time is calculated by summing all the link travel times.

Both the instantaneous and time slice models assume that speeds are constant along each link when calculating travel times. As a consequence, a discontinuity in the speed occurs as a vehicle leaves one link and enters the downstream link. According to the preliminary results of these two models, the differences between the two models were fairly minor. To expedite data processing, the faster instantaneous model was selected for this project.

5.2 Implementation

Since the Phase 1 project focuses primarily on the feasibility of the proposed approach, one prototype system and one EXCEL VBA-based tool were implemented to not only demonstrate the feasibility but also to gather feedback from MoDOT TMC staff. Both the system and the tool serve the same purpose – travel time estimation - but they have different hardware and software requirements.

5.2.1 Prototype system

The travel time estimation prototype system was first developed using MATLAB®, a high-level technical computing language. Figure 5-1 shows the graphical user interface (GUI) of the developed prototype system. Users can freely select the start (origin) and end (destination) points on I-64, the dates of interest, and the time of day. The system can retrieve data from the database and then calculate the average/median travel time automatically.

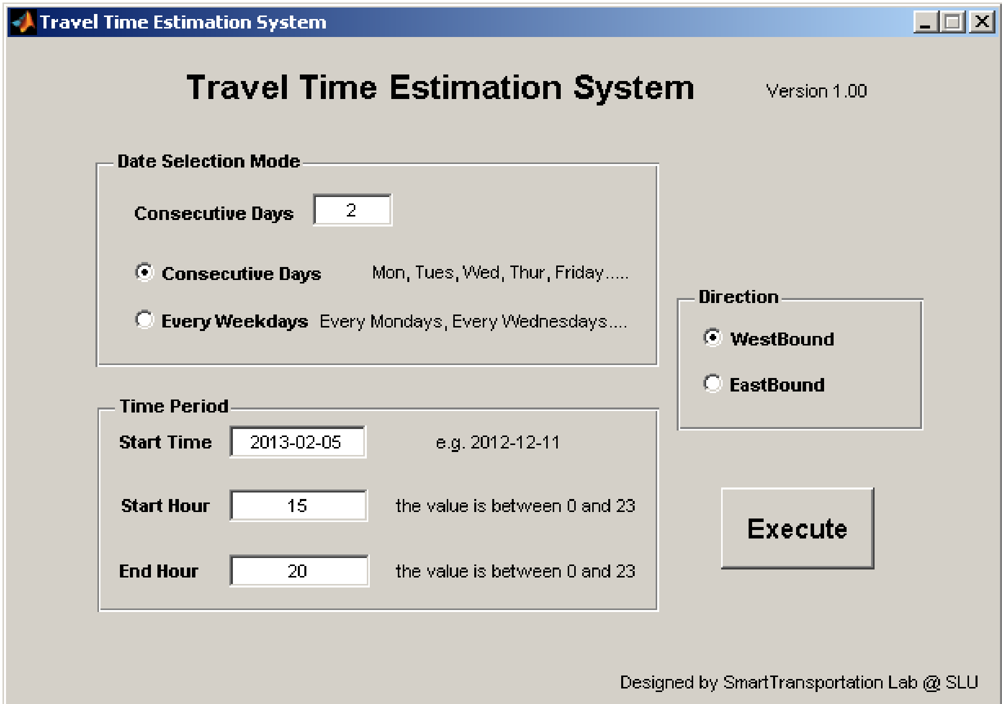
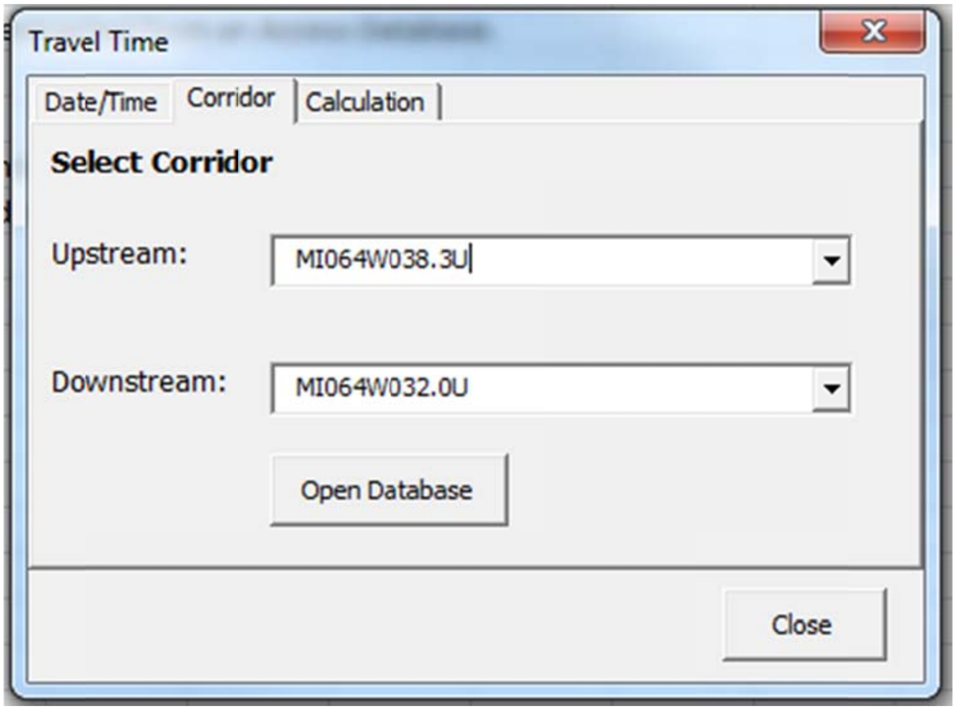


Figure 5-1: The Interface of the Travel Time Estimation System

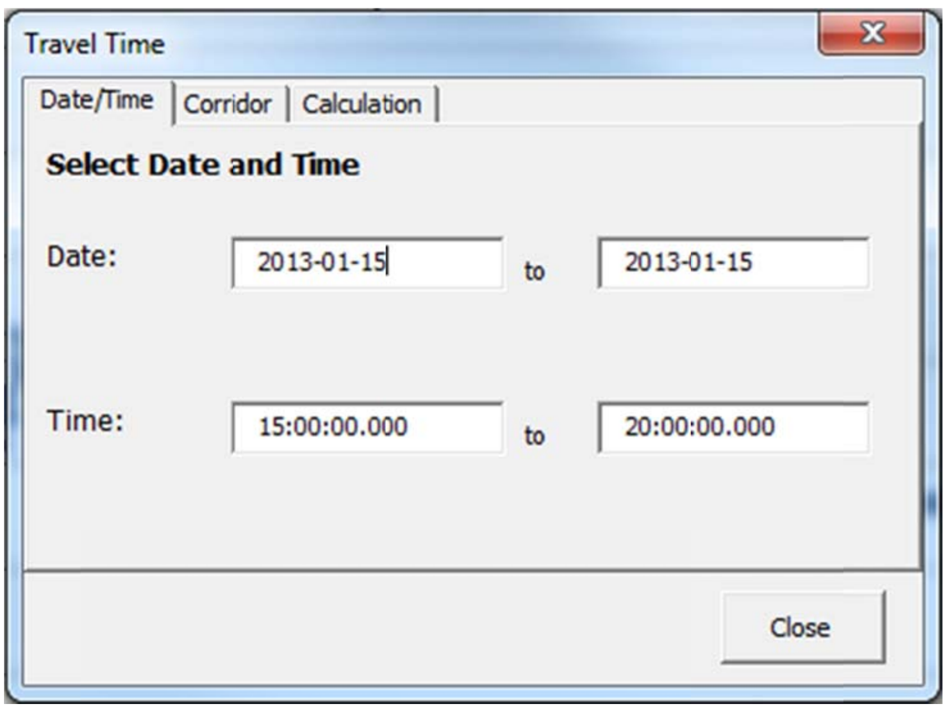
5.2.2 Excel VBA-based Tool

In addition to the MATLAB-based travel time estimation system, an EXCEL VBA-based tool was also developed. This tool is connected to a Microsoft ACCESS database rather than an SQL database. The advantage of this design is that the tool can run on any standalone computer without the need for any IT support. Its main disadvantage is that the MS ACCESS database requires frequent manual updates by traffic engineers because the tool is an offline system.

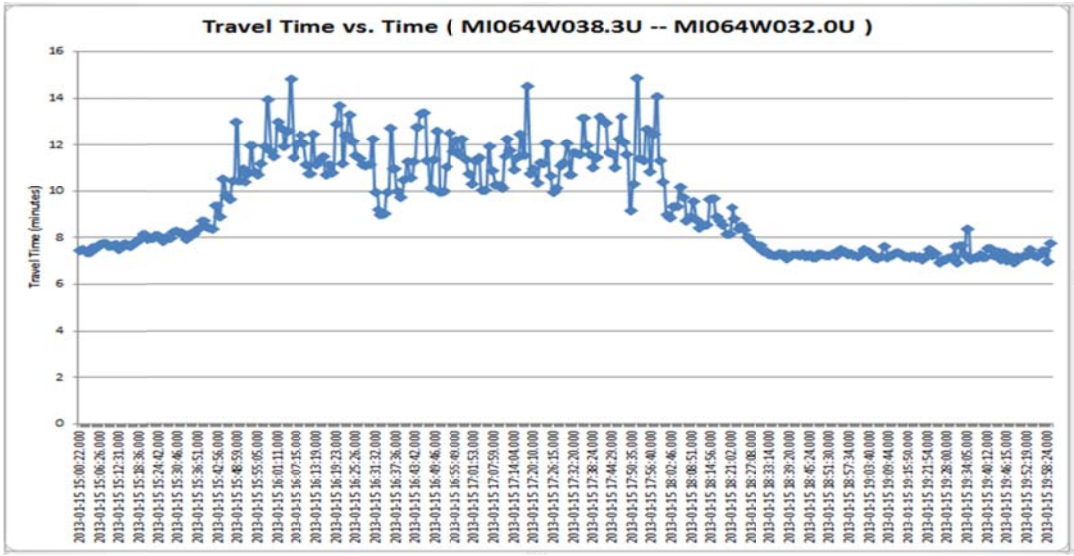
This tool performs similar functions to the MATLAB-based system in that it allows users to freely select any roadway segments between two sensors on I-64 and then select a date and time period to generate a travel time report. The GUI and the estimated travel time results are shown in Figure 5-2.


Travel Time
Date/Time
Corridor
Calculation
Select Corridor
Upstream:
MI064W038.3U
Downstream:
MI064W032.0U
Open Database
Close

(a)


Travel Time
Date/Time
Corridor
Calculation
Select Date and Time
Date:
2013-01-15
to
2013-01-15
Time:
15:00:00.000
to
20:00:00.000
Close

(b)



(c)

Figure 5-2: The Excel VBA-based tool for travel time estimation
(a) Upstream and downstream sensor selection, (b) Date and time selection, and (c) Results for the estimated travel time

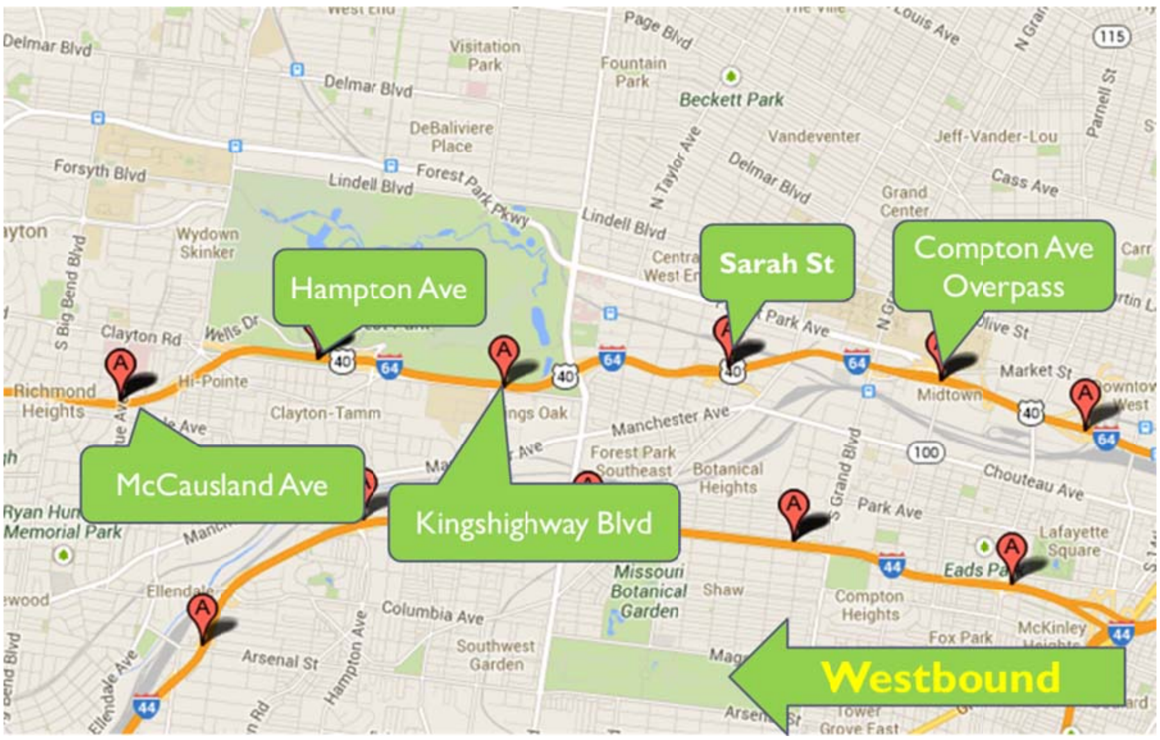
5.3 Case Studies

After the evaluation of the EXCEL VBA-based tool and the prototype system and a review of the feedback from the TMC staff, it was found that the prototype system provided more flexibility in function development. Therefore, several additional functions were developed for the prototype system to increase its usability. To demonstrate these functions, several case studies were conducted.

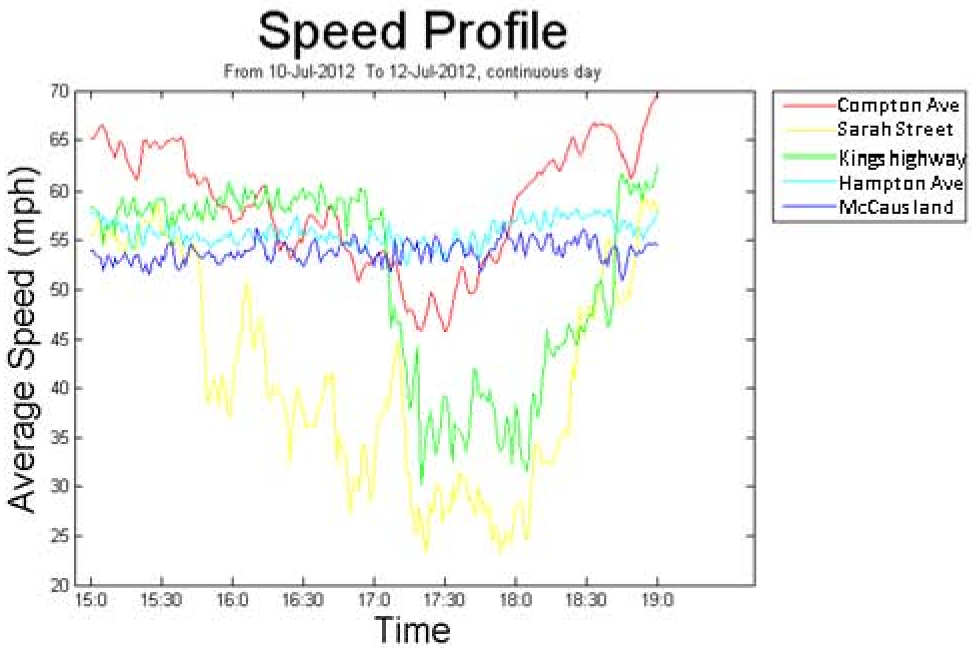
Case Study 1: Fundamental Traffic Analysis

The first case study investigated the segment of I-64 between Compton Ave and McCausland Ave. This segment was chosen because it is a well-known bottleneck that frequently suffers from recurrent congestion in the westbound direction during afternoon peak hours on weekdays. To investigate a traditional day of congestion, the data collected during the period 3pm~7pm for the three days from July 10, 2012 through to July 12, 2012, was extracted from the database and visualized by the prototype system. Figure 5-3 (a) shows the case study roadway segment on I-64, where five RTMS are located. Figure 5-3 (b) shows the speed profiles for all the RTMS on the study roadway segment from 3pm to 7pm for the three days of the case study. Figure 5-3 (c) shows the speed heat map, which is a fairly useful way for traffic engineers to identify the amount of delay caused by the congestion. Figure 5-3 (d) shows the travel time profile estimated by the system along this study roadway segment. Figure 5-3 (e) depicts a scatter plot showing the relationship between flow rate (volume) and speed data. This plot is also fairly useful and is commonly used by transportation researchers to investigate traffic flow problems.

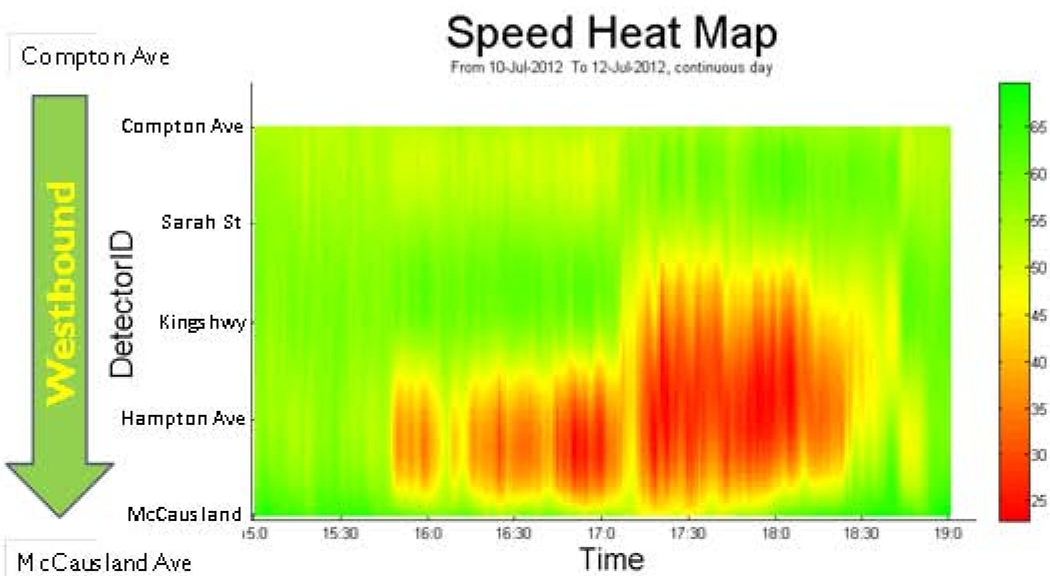
It should be noted that all the figures are generated automatically from the system based on the inputs from the users. This system is specifically designed to enable traffic engineers and researchers to conduct travel time analyses.



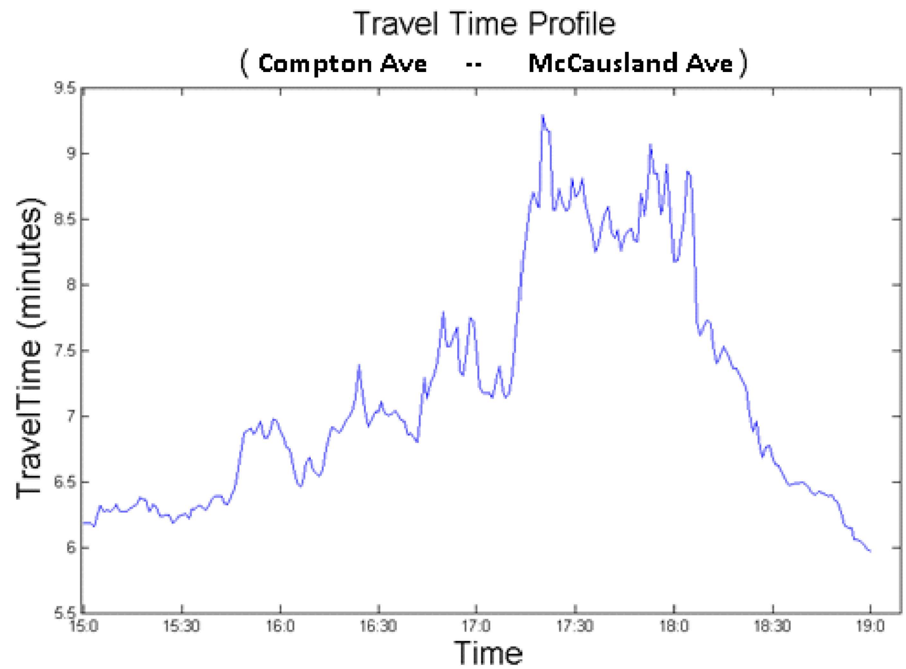
(a)


Speed Profile
From 10-Jul-2012 To 12-Jul-2012, continuous day
Average Speed (mph)
Time
Compton Ave
Sarah Street
Kings highway
Hampton Ave
McCausland

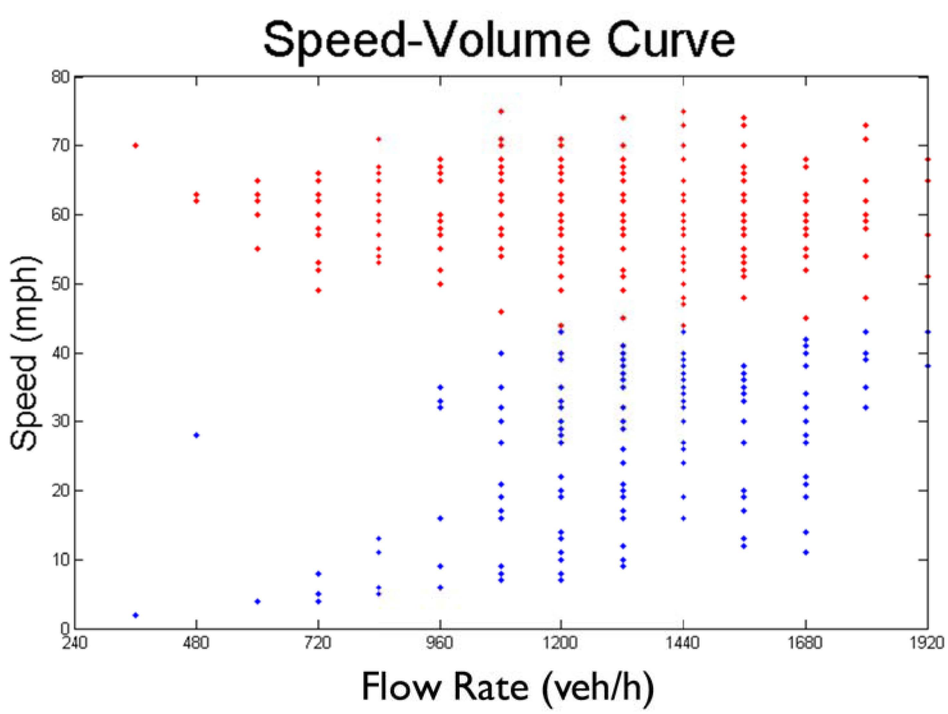
(b)


Speed Heat Map
From 10-Jul-2012 To 12-Jul-2012, continuous day
Compton Ave
Westbound
McCausland Ave
DetectorID
Compton Ave
Sarah St
Kingshwy
Hampton Ave
McCausland
Time

(c)



(d)



(e)

Figure 5-3: Travel time estimation system

(a) Case study site on I-64, (b) Speed profiles for five consecutive sensors, (c) Speed heat map (d) Travel time estimation results, and (e) Speed-volume relationship plot

Case Study 2: Bottleneck Identification

The first case study demonstrated a small scale traffic analysis, but the prototype system is expected to also be able to handle a larger query to visualize the results for the entire corridor. As illustrated in Figure 5-4, three bottlenecks can be visualized by querying three consecutive days of traffic speed data for the entire I-64 corridor. These three bottlenecks are at McCausland Ave (Case Study 1), the intersection with I-270 and the Daniel Boone Bridge on I-64.

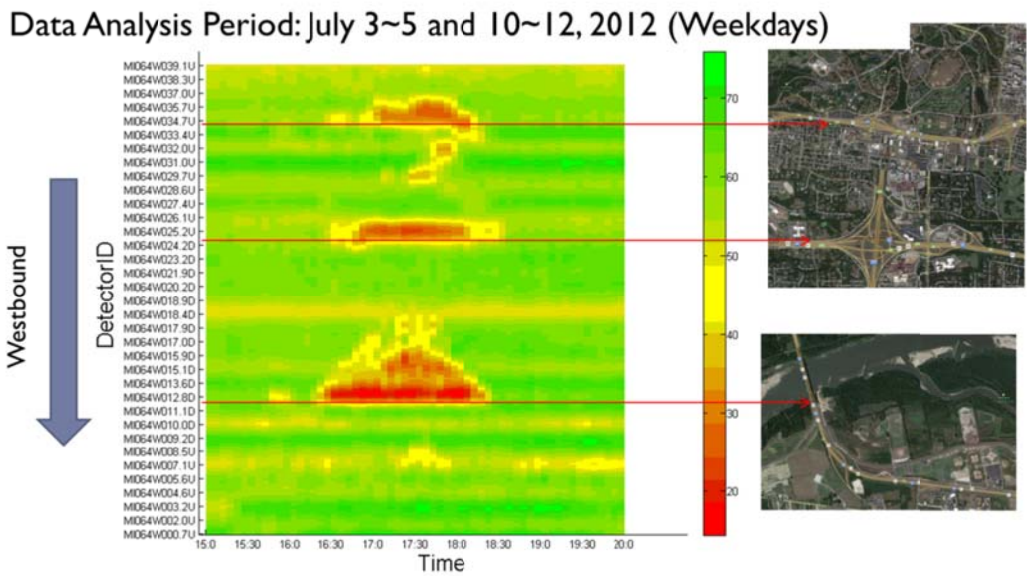


Figure 5-4: Case study of traffic bottleneck identification on I-64

Case Study 3: Snow Storm Impact Study

The third case study investigated the impact of a snowstorm on I-64 traffic. The freeway performance (travel time) before, during and after the snowstorm on Feb., 21st, 2013 is summarized in Table 5-1. The maximum travel time along the I-64 corridor increased to 111 minutes on the day of the snowstorm compared to 46 minutes on a regular day. In other words, many travelers were experiencing 2.4 times their normal travel time during the snowstorm. Interestingly, MoDOT was able to return the roadway conditions back to normal very soon after the snowstorm, because the minimum and maximum travel times the day after the snowstorm were effectively back to what they had been earlier in the week. Figure 5-5 compares the travel times on I-64 for the day before and the day of the snowstorm. Looking at the graph, it is clear to see that the snowstorm impacts started at 10am as the snowstorm hit the St. Louis area and were still delaying traffic that evening.

Table 5-1: Travel Time Comparisons Before, During and After the Snowstorm

	Westbound I-64			Eastbound I-64		
	Before-snowstorm traffic*	**Snowstorm day traffic (Feb. 21st, 2013)**	**After-snowstorm traffic(Feb. 22nd, 2013)**	**Before-snowstorm traffic***	**Snowstorm day traffic (Feb. 21st, 2013)**	**After-snowstorm traffic(Feb. 22nd, 2013)**
Min travel time (minutes)	39	63	41	40	61	40
Max travel time (minutes)	49	105	47	46	111	46
Min speed (mph)	49	23	51	53	21	53

*Average travel times on Monday, Tuesday and Wednesday (Feb., 18^{th}~20^{th}, 2013)

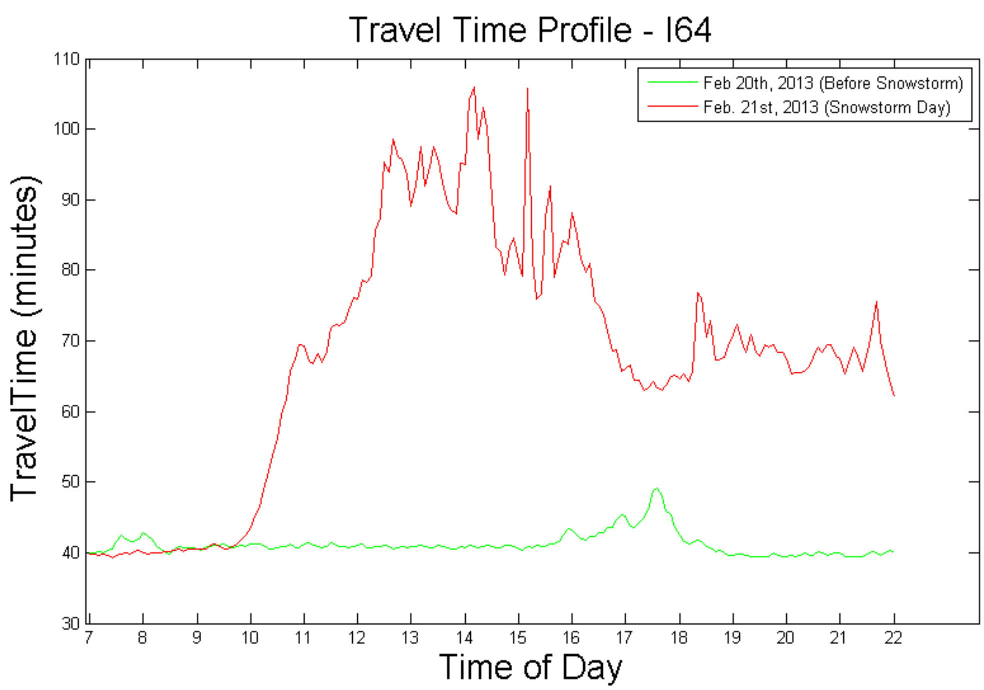


Figure 5-5: Comparisons of before-and-during snowstorm travel times on entire Length of I-64 (westbound)

Section 6.
Model Verification and Calibration

The estimated travel time results (using the instantaneous model) were compared with two sources of ground truth data**:** 1) Bluetooth-based travel time, and 2) vehicle-matching-based travel time.

6.1 Model Verification using Bluetooth-based Travel Time

Due to the unavailability of the Bluetooth-based travel time on I-64 (see Section 3 for more details), westbound and eastbound I-70 between SR-94 and I-270 were selected as a special case study to compare the results of the model with ground truth data. The study time period was the entire month of January 2013.

Figure 6-1 shows the locations of both the Bluetooth devices and RTMSs in the segment of I-70 between SR–94 and I-270 used in the study. Since the locations of Bluetooth devices and traffic sensors do not perfectly overlap, interpolation methods were used to estimate the travel times for the same segments using both the Bluetooth devices and the traffic sensors. The interpolation method is described by the small figure in the upper right corner of Figure 6-1.

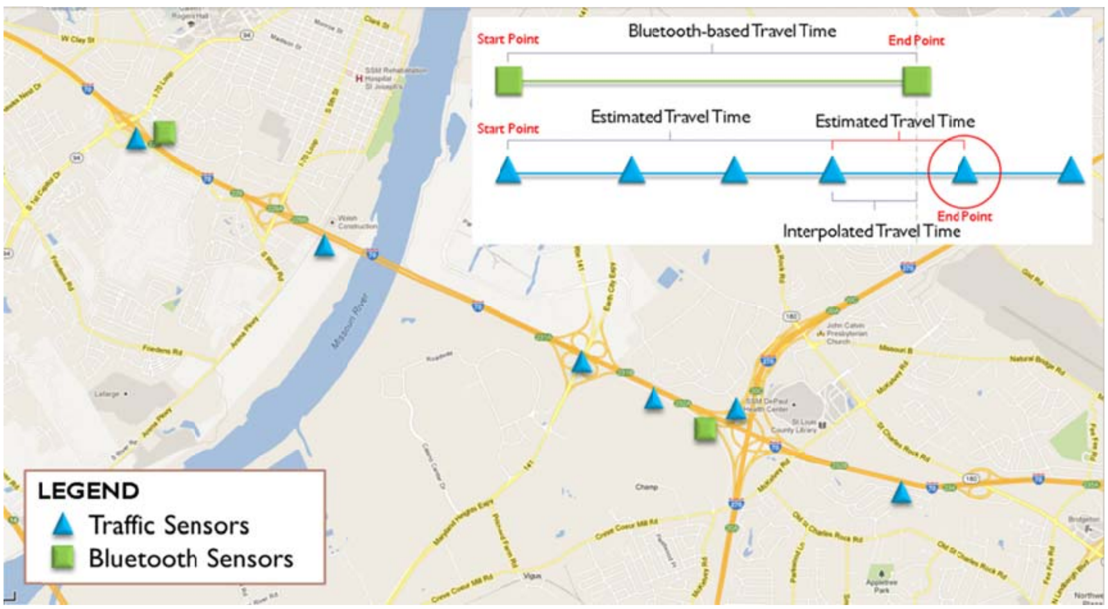


Figure 6-1: Locations of Bluetooth and point-speed sensors along the data collection segment of I-70 for travel time estimation result verification

Two measures of accuracy, Mean Absolute Error (MAE) and Mean Absolute Percentage Error (MAPE) are used in this study and defined as follows (Washington et al., 2010).

$$MAE = \frac{\sum_{t=1}^{n} \lvert F(t) - G(t) \rvert}{n}$$	(6-1)
$$MAPE = \frac{1}{n} \sum_{t=1}^{n} \left\lvert \frac{F(t) - G(t)}{G(t)} \right\rvert$$	(6-2)

where $G(t)$ is the ground truth data (in this case, the Bluetooth-based travel time) at a time interval *t*; $F(t)$ is the estimated travel time at time interval *t*; and *n* is the total number of samples.

MAE provides an overview of all the errors and shows the gaps between the estimated travel times and the collected travel times. MAPE, which shows the error as a percentage, is a scale independent measure of accuracy.

The Bluetooth-based travel time and the estimated travel time were quantitatively compared using data sets for each collected during January 2013. Figure 6.2 graphically depicts the comparison between the two types of travel times for the period 6:00 am to 9:00 pm on Jan 8th, 2013. The comparison results are summarized in Table 6-1. Both MAE and MAPE were calculated for both directions of I-70 for three types of time periods, namely weekdays, weekends, and the entire month of January. In the eastbound direction, MAE and MAPE were 0.22 minutes and 4.6%, respectively, on weekdays, while for the entire month they were 0.19 minutes and 4.1%, respectively. In the westbound direction, MAE and MAPE were 0.83 minutes and 12.3%, respectively, on weekdays and 0.73 minutes and 11.1%, respectively, for the entire month. These MAE and MAPE values indicate that the estimated travel time can be used effectively to represent the Bluetooth travel time, with fairly small gaps. It would therefore be reasonable for MoDOT to use the estimated travel time obtained using the data from the estimation model developed for this project to effectively evaluate freeway performance without the need to install additional Bluetooth-based travel time sensors. However, the estimated travel times from the model were generally lower than the Bluetooth-based travel times. This may have been because one Bluetooth detector was installed close to an intersection. Vehicle delays caused by the signalized intersection were also captured by the Bluetooth detector.

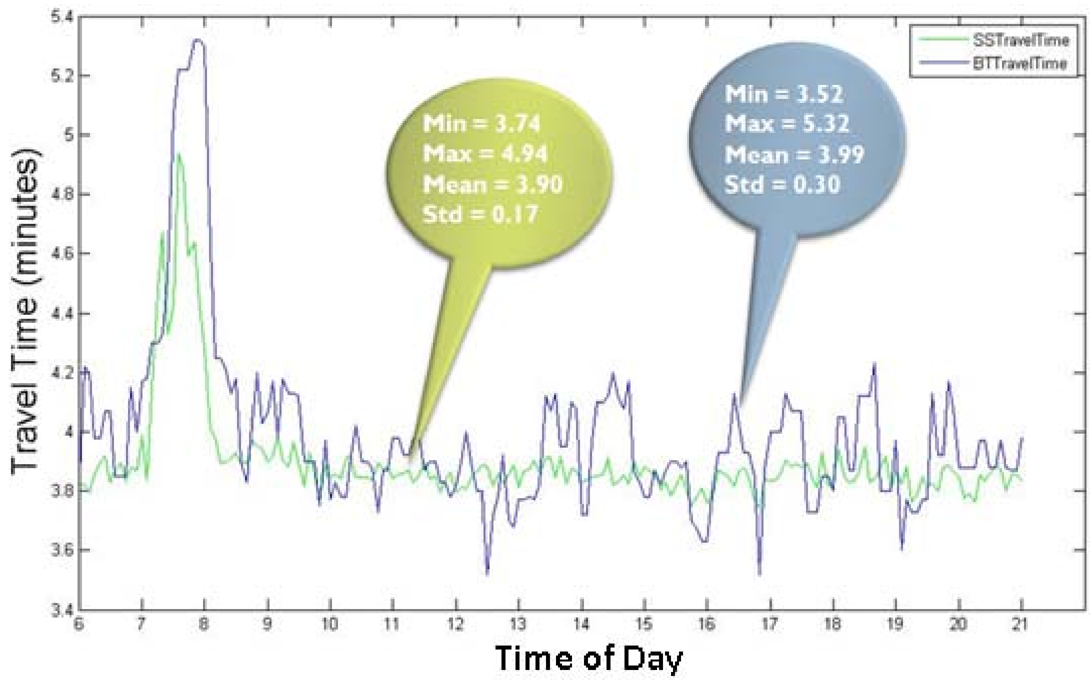


Figure 6-2: Comparison of Bluetooth-based and estimated travel times

Table 6-1: Quantitative Comparison Between Bluetooth-based and Estimated Travel Times
(January 2013)

I - 70	**Eastbound**		**Westbound**	
Weekday	**MAE (min)**	**MAPE (%)**	**MAE (min)**	**MAPE (%)**
Weekdays	0.22	4.6	0.83	12.3
Weekends	0.11	2.8	0.45	7.6
Entire Month	0.19	4.1	0.73	11.1

6.2 Model Verification using Vehicle-matching-based Travel Time

A second verification exercise was deemed necessary for this project for two reasons: 1) The Bluetooth-based travel time may not be truly representative of the ground truth travel time, and 2) the I-70 corridor is outside the research scope of this project. The segment manual video-based travel time collection method described in Section 4 was therefore implemented to verify the results of the estimated travel time model using real-time traffic data and manual vehicle-matching.

The video stream was provided by TrafficLand.com. Figure 6-3 shows how the travel time was collected through vehicle matching of the vehicles shown by the surveillance cameras. For example, the yellow truck in the first image was observed at the Vandeventer Ave camera on I-64 at 11:55 am, June 17, 2013 (Figure 6-3a), and reappeared at the Lindbergh Blvd camera on I-64 eight minutes later, at 12:03 pm (Figure 6-3b). Since the distance between the two cameras locations is known, the travel time of the yellow truck for this segment can be estimated.

(a) **(b)**

Figure 6-3: Collecting travel time ground truth data through the Vehicle-matching-based technique.

(a) at Vandeventer Ave &I-64 and (b) at Lindbergh & I-64

An additional verification study on I-64 from Kingshighway to Bellevue for the period from 4 pm to 7 pm on 27th, June 2013 was conducted to examine the travel time discrepancy during rush hours. Due to the low resolution of the TrafficLand videos, identifying and capturing the same vehicle from two videos is a fairly time consuming task, so the resulting number of valid samples was relatively low, thus impacting the statistical analysis. Consequently, the SLU research team decided to record the necessary videos at the TMC on June, 27, 2013. Figure 6-4 shows the corridor used for this verification study.

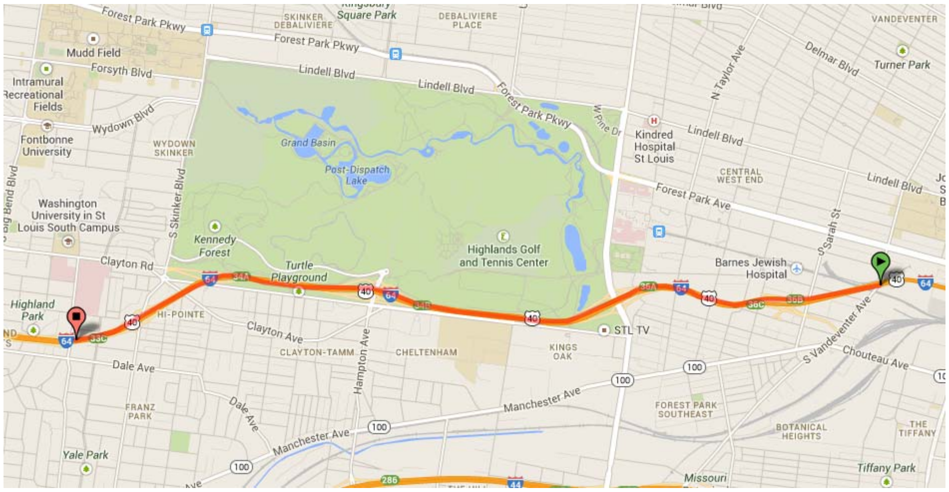


Figure 6-4: Location of MoDOT surveillance cameras along I-64
(Image obtained from Google Maps)

Figure 6-5 shows the two types of travel times, vehicle-matching-based travel time and estimated travel time. Consider that only trucks were tracked in the videos, and that truck speeds are usually lower than for passenger cars, the vehicle-matching-based travel time should be higher than the estimated travel time. However, the plots for the two travel times shown in the figure converged during the rush hours, indicating that the estimated travel times were overestimates.

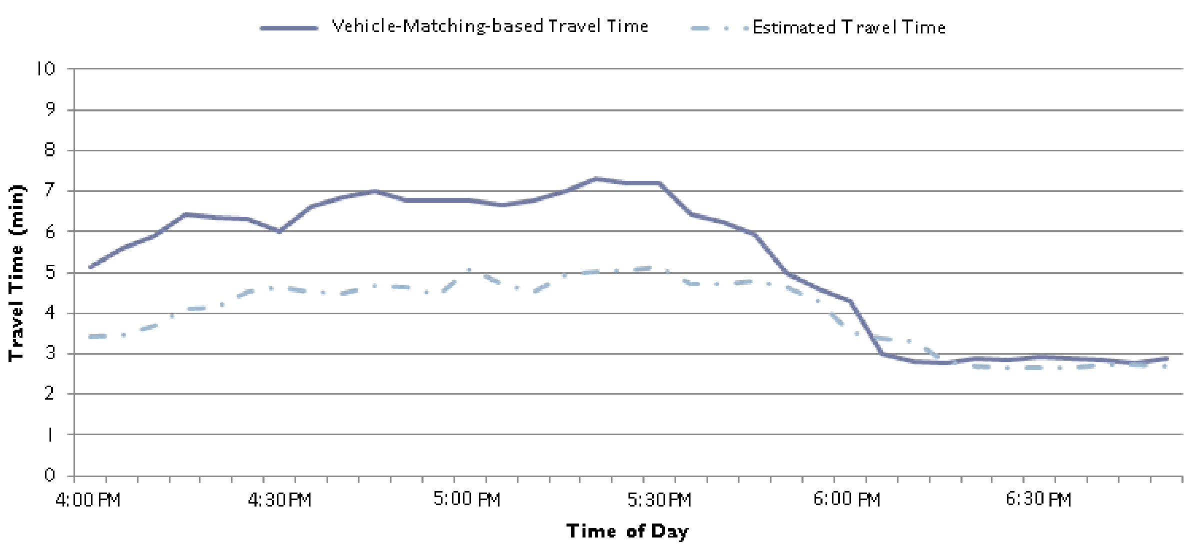


Figure 6-5: Comparison of Vehicle-matching-based travel times and estimated

travel times along the second case study corridor (Jun. 27th, 2013)

One possible reason for this overestimation was that any variations in the speed between the two traffic sensors is unknown. A linear travel time model was applied for the travel time estimation, with one of the underlying assumptions of the model being that vehicles travel at the same, uniform speed on segments. In fact, travel speed will vary based on both individual driver behavior and traffic conditions. Figure 6-6 shows an example explaining potential reasons why the speed may be underestimated. The dashed line shows the estimated speed curve (a linear plot). Since it is assumed to be linear, it is difficult to identify any bottlenecks that may occur between two sensors. In this case, a nonlinear model could be developed in future research to more accurately capture real traffic behavior such as bottlenecks.

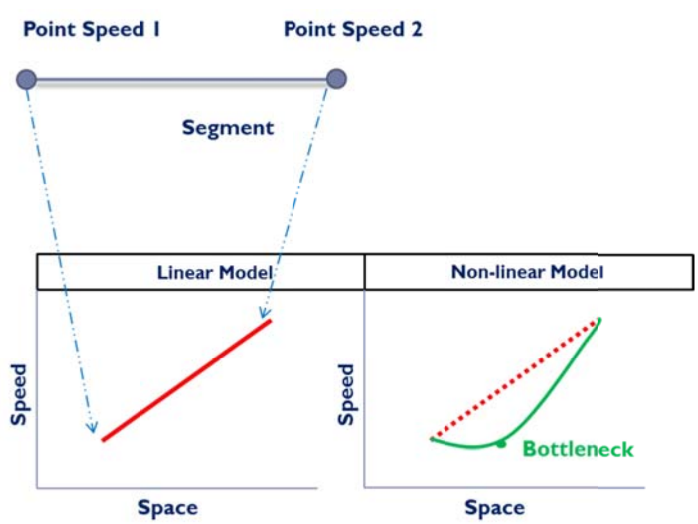


Figure 6-6: Linear model vs. non-linear model

Section7.

Conclusions and Recommendations

7.1 Conclusions

Travel time is widely believed to be a key performance measure for the evaluation of transportation systems. In the Greater St. Louis area, traffic data has been collected from fixed traffic sensors along the major freeways and arterial roads by the Transportation Management Center (TMC) operated by the Missouri Department of Transportation (MoDOT). This project sought to develop a new data handling model that will more fully utilize existing traffic sensor data to estimate travel times along the portion of I-64 that passes through the city. The following products were developed as a result of this project.

1. The traffic data has been archived in flat files in the XML format, and also parsed and stored in a DBMS physically located at SLU. This storage system was intentionally designed to allow easy and flexible sharing of the traffic data;
2. A new database schema was designed based on the traffic data characteristics and optimized according to the structure of the most commonly used SQL queries. The database designed was then used for estimating the travel time in real-time with much shorter response times for obtaining travel time information;
3. An appropriate travel time model was selected and implemented. In order to verify the travel times generated by the model, both Bluetooth-based travel time data and travel time data obtained through vehicle matching from video footage were used to test the accuracy of the traffic-sensor-based travel time estimation;
4. A custom Matlab-based prototype system and an Excel VBA-based tool were successfully developed. Based on feedback received from TMC staff and the superior flexibility of the MATLAB development environment, several additional functionalities were developed for the Matlab-based prototype system to allow it to handle both historical and real time traffic data, and publish the resulting travel time information. In addition to its basic functions, extra functions such as speed profiles were incorporated into the program. The Excel VBA-based tool was used only to process the historical data.

5. The travel time on I-64 was successfully estimated to an acceptable degree of accuracy utilizing traffic data collected from existing fixed traffic sensors along I-64.

7.2 Recommendations

The recommendations indicated by the above research results can be summarized as follows:

- The travel time estimation results show that the estimated travel time may be underestimated during rush hours. Other travel time estimation models might offer alternative ways to handle this issue. It is suggested the research team conduct a more comprehensive analysis to investigate this issue in the future.
- The data used for the model verification may not be sufficient. More data was required from both the Bluetooth devices and the video footage. Moreover, for convenience truck travel times were often used as ground truth information even though trucks generally travel more slowly than passenger cars, The potential solution to this conundrum could be to record high quality videos and hire additional student workers to “process” the videos.
- The user interfaces still have some room for improvement. More advanced techniques can be used in conjunction with other programming languages to improve the look and feel of the user interfaces.

Since the project has been extended to Phase 2 to cover travel time estimation for the entire freeway network in the Greater St. Louis region, the issues mentioned above will be addressed in Phase 2.

Acknowledgements

The authors would like to thank Tom Blair, Jeanne Olubogun, Greg Owens, Davar Divanbeigi and Cynthia Simmons at the TMC for technical support. We would also like to thank Rick Zygowicz, David Shackelford, Michael Bayer and Tiffany Rando at TransCore for providing real-time traffic data and technical support. We are grateful to William Stone, Jonathan Nelson and Andrew Hanks for the advice they have provided throughout. Special thanks to the Anis Dokic for video data processing.

References

Coifman, B. (2002). Estimating travel times and vehicle trajectories on freeways using dual loop detectors. *Transportation Research Part A: Policy and Practice*, *36*(4), 351–364.

Ehmke, J. F., Meisel, S., & Mattfeld, D. C. (2012). Floating car based travel times for city logistics. *Transportation Research Part C: Emerging Technologies*, *21*(1), 338–352.

Haghani, A., Hamedi, M., Sadabadi, K. F., Young, S., & Tarnoff, P. (2010). Data collection of freeway travel time ground truth with Bluetooth sensors. *Transportation Research Record: Journal of the Transportation Research Board*, *2160*, 60-68.

Klein, L. A. *Sensor technologies and data requirements for ITS*. Artech House, Boston, Mass., 2001.

Li, R., Rose, G., & Sarvi, M. (2006). Evaluation of speed-based travel time estimation models. *Journal of Transportation Engineering*, *132*(7), 540–547.

Ma, X., & Koutsopoulos, H. N. (2008). A new online travel time estimation approach using distorted automatic vehicle identification data., *ITSC 2008. 11th International IEEE Conference on Intelligent Transportation Systems*, Beijing, 12-15 October, pp. 204–209.

Ozbay, K., & Yildirimoglu, M. (2011). Comparison of real-time travel time estimation using two distinct approaches: Universal kriging and mathematical programming. *Intelligent Transportation Systems (ITSC), 2011 14th International IEEE Conference on* (pp. 1083–1088).

Sun, L., Yang, J., & Mahmassani, H. (2008). Travel time estimation based on piecewise truncated quadratic speed trajectory. *Transportation Research Part A: Policy and Practice*, *42*(1), 173–186. Turner, S. M., W. L. Eisele, W. L., Benz, R. J., & Holdener, D. J. (1998). *Travel time data collection handbook*. Report submitted to the Office of Highway Information Management, Federal Highway Administration, US Department of Transportation.

Washington, S. P., Karlaftis, M. G., & Mannering, F. L. (2010). *Statistical and econometric methods for transportation data analysis*, Chapman and Hall/CRC Press.

Wasson, J.S, Sturdevant, J.R, & Bullock, D.M (2008). Real-time travel time estimates using media access control address matching. *Institute of Transportation Engineers*, *78*(6), 20-23.