University Transportation Centers Program

Project Title:
Exploring Innovational Solutions in Multimodal OD Data

Principal Investigator (PI) Name, Title and Contact Information:

Li Zhang, Ph.D., P.E., Associate Professor

CO-PI Dennis Truax, Ph.D., P.E., Professor

Co-PI Thomas D. White, Ph. D., P.E., Professor

Department of Civil and Environmental Engineering

P.O. Box 9546, Mississippi State University

220 Walker Engineering Building

Mississippi State, MS 39762

Phone: (662) 325-9838

Fax: (662) 325-7189

Email: lzhang@cee.msstate.edu

Submission Date:
June 15, 2016

Project Period
March 2016, June 2016

Acknowledgement

The authors wish to acknowledge that the Planning Division of the Mississippi Department of Transportation (MDOT) provided pivotal support for this research. This study was performed to augment the findings of an MDOT study, "Origin Destination (OD) and Travel Time (TT) Data." This latter effort provided a foundation for the work performed and essential matching funds for this study. Special thanks are extended to Mr. Jeff Elly, Director, MDOT Planning Division, and division staff members Sammy Holcomb, Trung Trinh, and Susannah Seal, for their boundless support of this project.

Furthermore, MDOT contracted with Michael Baker International to collect traffic volume data for this project in Desoto County, Mississippi. Hence, we would also wish to express our gratitude to the staff of Michael Baker International's Mississippi office; specifically, Ray Balentine, Miller Brooks and Luke McMahen for their flexibilities in planning and conducting traffic data under a tight schedule.

Finally, the authors wish to acknowledge William Case Fulcher, Lei Zhang, Amos Hou, Sara Randall, John Green, Paul Moore, John Swain, Kelly Truax, Chastain Hughes, and Christian Shelby for their participation in this research. All of these students, with the exception of Sara and Amos, graciously gave three days of their summer to collect real world data in Desoto County, Mississippi. All of this data was processed by Amos under the supervision of Case and Lei who developed the model. Lei also programed the model and performed all of the simulation runs. All of this information was then compiled into this paper by both Case and Lei with assistance from Sara who aided in the literature review.

Disclaimer

Contents

Introduction

Origin destination data (O/D) is critically important in transportation planning. Real time O/D data can contribute to improving traffic operations and performance measures as well; however, this data is difficult to obtain. In the transportation planning process, it is usually estimated through an expensive survey with only a small response percentage. While Bluetooth O/D equipment can provide some percentages of O/D data, it is not able to provide the comprehensive O/D data needed for transportation planning and operations. In addition, real time O/D data could be useful at diamond interchanges which consist of two signalized intersections. These traffic signals provide safety and mobility benefits; however, they are currently treated as an individual intersection on the arterial and the OD traffic flow patterns are not considered to establish the traffic signal timing plan. Real time O/D data estimation at interchanges such as diverging diamond interchanges (DDI) are important for traffic signal optimization which could alleviate congestion and remove the typical DDI bottleneck, but OD is difficult to collect and even estimate in the field. This report presents two models to improve O/D estimates for both of these scenarios. The first model provides a means to determine an accurate O/D estimate from available O/D data (less than 100%) and traditional loop detector data while the second model presents a method of determining real time O/D including congestion effects through a linear system at a diamond interchange. These two models are then verified through simulation, and with the assistance of the Mississippi Department of Transportation (MDOT) field data was collected and used to validate the model estimations and simulation results.

Literature Review

In this section, a comprehensive literature review is conducted.

Van der Zijpp conducted a study on the dynamic origin-destination matrix estimation from traffic counts and automated vehicle identification data. The purpose of his study was to prove that the Bayesian updating method is the preferred method for the estimation process. He conducted an experiment to using the DCLS method, Kalman filter, and the Bayesian updating method. In the experiment he allowed the Bayesian updating method to use only link volume counts and the DCLS and the Kalman method to use both link counts and the trajectory information. The Bayesian method is the preferred method for using nonnegativity constraints, but a more traditional method, such as the DCLS or the Kalman method, would need to be used to if nonnegativity constraints were not being used. It has been determined that the Bayesian method is the most preferred method because it eliminates the amount of errors and produces a more correct estimation with less information. However, if the Bayesian method were to have link volume counts along with trajectory counts the estimation would be even more correct. [1]

In another paper, Tsekeris and Stathopoulos conducted a study on the real-time dynamic origin-destination matrix adjustment with simulated and actual link flows in urban networks. In their paper they compared different algorithms such as the MPP, MART, DIMAP, and RMART. The algorithms used were based on a quasi-DTA model. The quasi-DTA model is a 'time-dependent estimation of link-use proportions.' For the MPP, MART, RMART and DIMAP algorithms there were different processes used. Each algorithm has a process with five to nine steps to be sure it is computed correctly. These algorithms were tested in Athens on an urban road network. It was found that all algorithms were affected by a number of different objects. Some of these include: the time interval duration, source of ground-truth flows, the DTA procedure, and the link count availability. It was proven that for real-time problems the MART and RMART were the best algorithms to use, while the DIMAP was the best to use when wanting a matrix for simulated link flows. [2]

Dan, Zhicai, and Hongfei studied the adaptive-filtering based dynamic origin destination matrix estimation process. They came up with a method based on the Sage-Husa adaptive filtering algorithm. Throughout the article the algorithm is referenced as the "dynamic OD matrix estimate adaptive filtering process" or the "improved adaptive filtering process." It was found that this new algorithm was not only more consistent than the previous algorithms, but it also had better accuracy than the Sage-Husa adaptive filtering algorithm. [3]

Also, White and Wells conducted a study on the extraction of origin-destination information from mobile phone data. For their study to work they had to research mobile phones and how to get their locations. First off, a location can only be received from a mobile phone that is turned on. Once the phone is on it has to "poll" with a transmitter. This "polling" process happens every five to ten minutes to make the locations more accurate. The best way to get the location of a vehicle through a cellular phone is for the phone to make a phone call. The location of the phone can be traced throughout the entire call. A big problem with using cellular phones to get locations is the legal aspects. To avoid legal situations, the phone number, name of the owner, or any other personal information is encrypted. This means that when a call is made only the accurate locations come in. White and Wells collected and studied every aspect of using mobile phones for extracting origin destination information before conducting a "pilot study." In their pilot study they compared the OD matrices of CONTRAM8 to billing data from the mobile phones. They found that a much larger sample was needed for the billing data than for the CONTRAM8. [4]

Lou and Yin performed a study on the real-time estimation of origin-destination flows for actuation-controlled intersections. In their study they use time sensitive data to determine origin destination flows that are not relative to time. Lou and Yin have invented a new procedure that will fix the problems with "complete entering counts but incomplete exiting counts." Their procedure has two steps. In the first step they estimate two columns of the origin-destination matrix by using equation (8). The second step gets rid of the non-negativity constraint and solves for each additional column in the origin-destination matrix by using equation (9). Lou and Yin conducted an example to demonstrate their new two-step procedure. In this example they use a typical four-way intersection that does not allow U-turns. This simulated five different experiments in their example. In each experiment it was proven that the new two-step method outperformed the GLS method that was previously used in this situation. Although their procedure worked, Lou and Yin know that there are a few problems and that there is always room for improvement. [5]

Hazelton conducted a study on the statistical inference for time varying origin-destination matrices. In his study, he based the origin-destination matrices on daily link count sequences. Because of this, he cannot effectively track a certain vehicle from departure to arrival. While developing his method, Hazelton decided that using a Bayesian approach to inference would be the best way to go. He tested his new method on a road network in Leicester. His test concluded that the traffic flow between weekdays and weekends was very methodical. [6]

Calabrese, Di Lorenzo, Lou, and Ratti also performed a study on the estimation of origin-destination flows using opportunistically collected mobile phone location data from one million users in the Boston Metropolitan area. They developed an algorithm specifically for determining the origins and destinations of individual trips by using mobile phone locations. The location of a mobile phone can be found when the user does any one of the following three: 1) when a call is placed or received, 2) when a short message is sent or received, and 3) when the user connects to the internet. These network connections allow the location to be determined and help to find the origins and destinations. To test the algorithm that they developed, they conducted a study in the Boston Metropolitan area. Throughout their study, Calabrese, Di Lorenzo, Lou, and Ratti found that there were several advantages and disadvantages to their algorithm. Both the advantages and disadvantages are listed below: [7]

Advantages: [7]

1. It can capture the weekday and weekend patterns as well as seasonal variations.
2. It can capture work and nonwork trips.
3. It can produce real time, continuous origin-destination matrices which can capture the very fine grain spatialtemporal patterns of urban mobility.

Disadvantages: [7]

1. The market share of the mobile phone operator from which the dataset is obtained.
2. The potential non-randomness of the mobile phone users.
3. Calling plans which can limit the number of samples acquired at each hour or day.
4. Number of devices that each person carries.

In another paper, Caceres, Wideberg, and Benitez conducted a study on the derivation of origin destination data from a mobile phone network. When beginning their study, they were looking for a more cost-efficient and reliable way to estimate origin-destination matrices. They performed a feasibility study on their new global system for mobile communication (GSM) network. The test area of the study was between Huelva and Seville, both being Spanish cities. In this area there were four location areas. In both the origin destination matrices and traffic counts it is noticed that "estimation errors tend to decrease as observation intervals increase." By using the anonymous phone location data it was determined that the method developed produced accurate estimations and will be a good method to use in the future. [8]

Tuydes-Yaman, Altintasi, and Sendil conducted a study to better estimate the origin-destination matrix by using automated intersection movement count data. In their paper they proposed both a mathematical formulation and a new model. These will work together to perform estimates of a static origin-destination flow. While developing the models, Tuydes-Yaman, Altintasi, and Sendil assumed that traffic counts and the total production were observed. They also assumed that there was no prior information on the area studied. The group performed two test cases to better determine how well the models worked. They used thirty different scenarios in each test case. The results of the two test cases were successful. The new models performed much better than the previously used link count based models. Although it was successful, the authors still realize that the models have about a ten percent measurement error and can be worked on. [9]

Perrakis, Karlis, Cools, Janssens, Vanhoof, and Wets researched a Bayesian approach for modeling origin-destination matrices. In their study, they use a statistical Bayesian approach to determine origin-destination matrices based on census data. This study is very cost-efficient and will work for even large dimensioned matrices. The group first researched the Poisson model and different forms of this model. They wanted to know every aspect to be sure that there were no mistakes in the study. Next, they studied the Metropolis-Hastings simulation. This was studied because it is involved in the formulation of the Bayesian approach. The overall result of their study was that the Bayesian method was satisfactory. [10]

Cao, Miwa, Yamamoto, and Morikawa studied the estimation of dynamic link flows and origin-destination matrices from lower polling frequency probe vehicle data. In their study they looked at many aspects pertaining to dynamic link flows and origin-destination matrices. They "analyzed the effects of polling frequency and method of decomposing travel time on the derived travel time and then explored methods of estimating dynamic link flows an origin-destination matrices from lower polling frequency probe vehicle data." The group discussed each of these issues and proposed a method to give more reliable estimates. They had a four-step process in developing the method. This process included: step 1-travel time allocation, step 2-link performance function fitting, step 3-dynamic link flows estimation, and step 4-dynamic origin-destination matrices estimation. After they went through the four-step process and had a method, the team conducted a numerical experiment. From the experiment, it was learned that the estimates are more accurate if the polling intervals are longer rather than shorter. Overall, this is a great method to use when estimating dynamic link flows and dynamic origin-destination matrices. [11]

Parry and Hazelton studied the estimation of origin-destination matrices from link counts and sporadic routing data. They propose three questions that they wish to solve in their research. These questions are: How should link and routing data be combined, particularly when collected contemporaneously? How much routing data is necessary to overcome the traditional identifiability problems in origin-destination matrix estimation? To what extent will it matter if we have imprecise information about the penetration rate and its possible variation across the network?" They were able to answer the questions throughout the article. They also determined that the statistical model that they developed worked very well. There was only a five percent error rate. Parry and Hazelton realize that their model may have a few errors and can always be updated. [12]

Hazelton studied the estimation of origin-destination matrices from link flows on uncongested networks. The purpose of this study was based off of two reasons. The first reason was because "the uncongested case is of practical interest itself" and the second was because "studying the problem without congestion will hopefully provide insight into the considerably more complicated question of inference for congested networks." He presents an origin-destination estimation method to combine variations in route choice probabilities. Hazelton knows that there could be a measurement error in his model, but this is the case in any method involving link flow data. In his study, he conducts an experiment based on traffic flows that are recorded by electronic vehicle detectors. The purpose for this is because conducting the experiment based on the electronic vehicle detectors is a much more "concrete" example. There were three tests experiments shown in the study. These tests were very consistent; however, it is strictly theoretical. [13]

Li and De Moor studied the dynamic identification of origin-destination matrices in the presence of incomplete observations. The purpose of their study was to investigate these matrices when there were no traffic counts. Two different types of incomplete observation situations are considered in the study. The first one is "insufficient installment of sensors of traffic flows" and the second one is " failures of some sensors in a traffic system." Li and De Moor conducted experiments with what they researched and developed. In these experiments it was learned that estimates of an origin-destination matrices could still be determined with missing information. The accuracy o the estimation varies depending on how much information is missing. One of the big 'selling points' of the algorithm that Li and De Moor developed is that the journey time from an entrance to an exit does not need to be assumed. [14]

Tamin, Hidayat, and Indriastuti developed a maximum-entropy and Bayesian–inference estimation method for calibrating transport demand models based on link volume information. In their method they used traffic counts to obtain the information needed for the origin-destination matrices. The group used traffic counts because they are inexpensive, collected relatively often, and are easier to organize and manage. While researching, Tamin, Hidayat, and Indriastuti studied many different theorems, methods, and models that would playa part in the method that they developed. They went through four different steps in order to determine the perfect or best location for the traffic counts. These four steps were: 1) proportion of trip interchanges on a particular link, 2) inter-link relationships, 3) optimum number of traffic counts, and 4) the determination of the optimum number of traffic counts. From their extensive research and after developing their method, the group learned that there were several factors that determined the accuracy of the origin-destination estimation. [15]

Bera and Roa studied the estimation of origin-destination matrix from traffic counts. They are trying to find the best or "state of the art" method to determine the origin-destination matrix (ODM) estimate. In their study they compared the static ODM to the dynamic ODM. It was found that the static origin destination matrices are a better and more realizable option. There still needs to be more tests ran to check the static ODM for long time transportation planning, but overall the static ODM is the best option for estimating origin destination matrices. In the process of trying to determine which option was the best, Bera and Roa went through every aspect of both the static and dynamic ODM. They considered models with and without congestion effects, travel demand models, information minimization (IM) and entropy maximization (EM) approaches, combined distribution and assignment based problems, bi-level programming approaches, fuzzy based approaches, and a few more. They went very in-depth in their research to help better determine the best way to estimate origin-destination matrices. [16]

Sherali and Park conducted a study on the estimation of dynamic origin-destination trip tables for a general network. In this study, they propose a bound-constrained least squares model. While developing this model they ran into a few problems. One problem, known as the "master problem" in the article, is a bound constrained quadratic problem. To solve this problem they developed a four-step process based off of a conjugate gradient algorithm that Fletcher and Reeves' developed. Sherali and Park wanted to demonstrate their model so they performed two test problems. The first test problem is strictly hypothetical and the second is based on the Massachusetts Turnpike. For the simulated problem they presented the 'known' data and used their four-step process. From the first test problem it was proved that the four step process that Sherali and Park developed would generate the exact solution. On the second experiment about the Massachusetts Turnpike they had five problems to solve. When looking at the results from problems one through five they found that problem three was the best. It generated a "closer replication of a true solution." [17]

Zhang, Qin, Dong, and Ran performed a different study focusing on the daily origin destination matrix estimation using cellular probe data. In their study they compared the cellular probe data method that they introduced to the simple random sampling (SRS) method that was more commonly used. In the introduction the team presented three major limitations that currently existed and propose that their cellular probe data method can cover the limitations. For the cellular probe data method to work properly they had to adopt the cellular trading method to find. The trip origins, even though they adopted this method for the trip origins, they still needed a method to identify the trip ends. After some research it was determined that the Transportation Analysis Zone (TAZ) was the best method to identify trip ends for their study. In order to complete this study the team had to make a few assumptions. The two most important assumptions are as follows: (1) “There might be multiple cellular carriers existing in the research areas. Each of the carriers is operated independently.” (2) “The cell-phone ownership pattern is identically distributed among different cellular carriers.” After these two assumptions, three more had to be made for the cellular probe trajectory. These assumptions are listed on page eight of the article and are important to the study, but are not as important as the two assumptions mentioned above. To test their method Zhang, Qin, Dong, and Ran conducted a simulation experiment to test their proposed method. In the experiment it was proven that the cellular probe data method that the team developed was more effective than the SRS method that was previously being used. This was proven when the results of the SRS method were twice as high as the cellular probe data methods results in the average percentage error. [18]

Rilett and Dixon conducted a study on real-time OD estimation using automatic vehicle identification and traffic counts. Their study primarily focuses on combining many different aspects of intelligent transportation systems into one process and using real-time information to do so. A few more exact focal points are: determining if automatic vehicle identification tags can be used in real-time, how much work would need to be done to use AVIs, and determining the most appropriate method to use. Rilett and Dixon decided to use Lagrange multipliers instead of Kalman filter to produce a non-negativity constraint in their functions. In the study, there were 26 different types of scenarios studied. They used a section of I-10 through Houston, Texas about 15 miles long. This made the test bed to be 9 links and 15 origin destination pairs. The 26 different scenarios were tested differently based on the information known. Rilett and Dixon state in their article that “although the test bed is a linear network, the proposed methodology could be applied easily to a general network because the dynamic link choice proportion information can be identified from the AVI data.” The algorithm and information learned from this study will make it much easier and more reliable to collect data based on real-time, but if something were to happen and the detectors failed this method would not work. [19]

Sherali, Narayanan, and Sivanandan researched the estimation of origin-destination trip-tables based on a partial set of traffic link volumes. During their research, they did three different studies or ‘test networks’. For test network one, they found that when the algorithms SA and THE are compared that THE is much slower. The linear programming approach used in test network one is ten times faster than the maximum entropy approach. For test network two, they used a network which has three links and two origin-destination changes. Sherali, Narayanan, and Sivanandan found that test network two can not only find user-equilibrium solution, but also reproducing link flows. In test network three, they compared test network one to test network two by comparing the linear programming approach from test network one and the Bielefeld programming method from test network two. From comparing these two, they found that the approach used in test network one could perfectly match the set that was given. Overall, it can be determined that the sequential linear programming approach from test network one is a faster, better approach. This was proven in test network two and three. [20]

Cascetta and Postorino conducted a study on fixed point approaches to the estimation of origin-destination matrices using traffic counts on congested networks. In their study they focused on fixed point problems of an implicit function. They developed a way to produce flows and costs in a matrix that align with the assignment. In their study they used small networks and the results verified that every algorithm comes to the same solution, just at a different time. Specifically, the functional iteration algorithm and the proposed MSA algorithm both outperform the traditional MSA algorithm. To prove that each algorithm came to the same conclusion, they used a three step process that does not always work the first time and may be repeated until a suitable answer is achieved. [21]

Hellinga and Van Aerde organized a study on estimating dynamic O-D demands for a freeway corridor using loop detector data. Their study was based off of a multi-lane freeway in Toronto, Canada. They briefly describe the freeway site and the traffic management system of the freeway to provide a background into the study. To be able to conduct the study, they had to make a few assumptions. The first assumption was that drivers would access the express lane as soon as possible and stay in that lane as long as they could. The second assumption was that drivers would stay in the collector lane the entire trip. The third assumption was that drivers do not switch lanes. The third assumption was necessary for assumptions one and two to work. There were a few problems with the study. These basically include onto having all of the information needed and having to assume information. Hellinga and Van Aerde thought this was typical of most sections studied even though it was not ideal for estimating origin-destination demands. They chose to study a 15 minute time interval to decrease the amount of errors, but said a shorter time interval would compute a more accurate estimate if there were few to no errors. In their study, they determined that both the LSE and LRE models can 'estimate a time-varying O-D demand that successfully reflected the observed link flows.' [22]

Yang, Meng, and Bell developed a general model and an efficient solution procedure for simultaneous estimation of the O-D matrices and travel-cost coefficient for congested networks in a stochastic user equilibrium. The model that was proposed is a logit-based model that allows consistency between trip estimation and trip assignment procedures. To be able to randomly determine the flow or traffic assignment on congested networks, the stochastic user equilibrium (SUE) must be followed. They demonstrate how they formulated their simultaneous estimation model and how it works. The model can resolve the discrepancy between trip estimation and trip assignment procedures and, using θ to determine more flexibility, it can estimate the route-choice proportions. Yang, Meng, and Bell determined that all of the functions used in their model are differentiable. Once that was determined, they could use the successive quadratic programming (SQP) method. The SQP method can be used to find an 'exact local solution'. [23]

Model Methodology and Proof of Concept

Diamond Interchange

Figure 1 shows the origin destination (OD) traffic flow map of a typical diamond interchange (US-78 and East Goodman Road). In Figure 1, each traffic flow source could be seen as one origin and each traffic flow target could be seen as one destination.

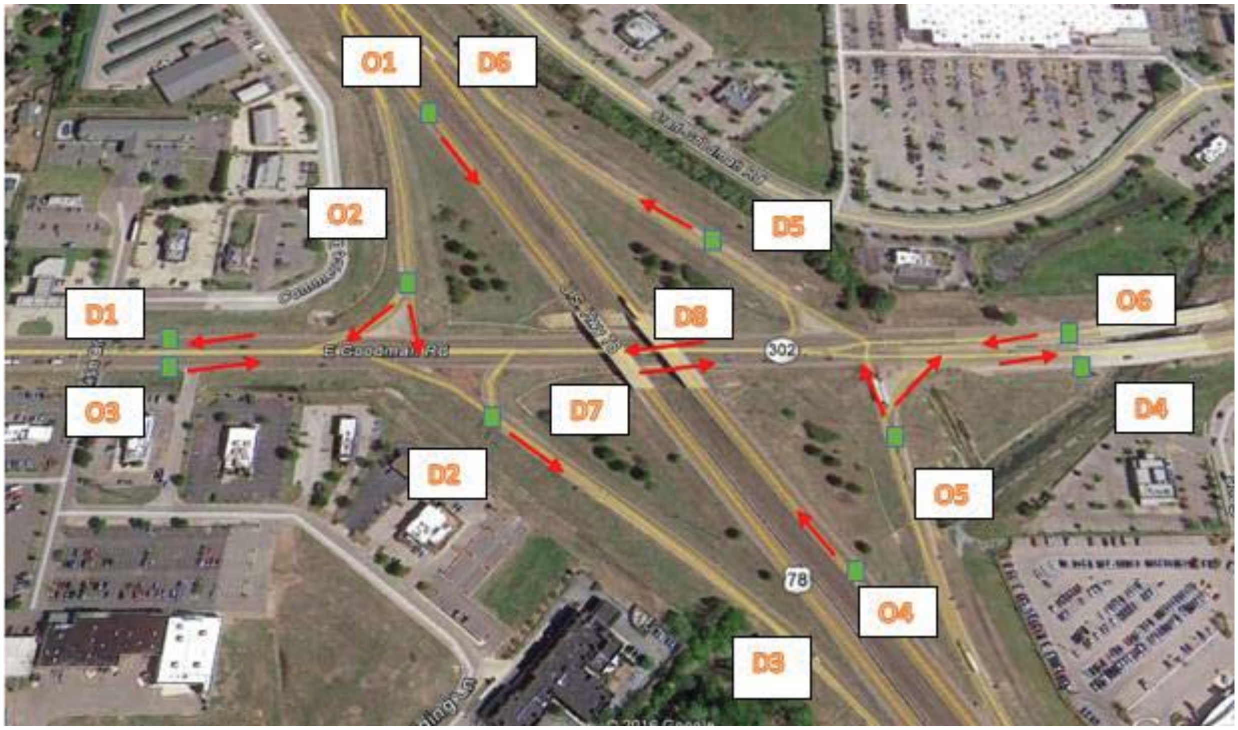


Figure 1 Origin Destination Map of diamond interchange

Based on the OD map, an OD estimation model could be represented as one OD table. Table 1 shows this OD estimation table.

Table 1 OD estimation table

Origin Sta / Dest Sta.	O1	O2	O3					On
D1	T11	T21	T31					Tn1
D2	T12	T22	T32					Tn2
	…	…	…					…
	…	…	…					…
	…	…	…					…
	…	…	…					…
Dm	T1m	T2m	T3M					Tnm

In the OD table, Tij is the vehicle trips from origin i (Oi) to destination j (Dj). The constraints of the OD estimation model are the sums of the traffic volume from each origin to all destinations should be equal to each origin's total traffic volume, if there is no queue at the studied area. The sums of the traffic volume from all origins to each destination should be equal to each destination's total traffic volume. The formulas for each constraint are shown below.

$$O1 = \sum_{j=1}^{m} T1j = T11 + T12 + T13 \ldots + T1m$$

$O2 = \sum_{j=1}^{m} T2j = T21 + T22 + T23 \ldots + T2m$

…

$Oi = \sum_{j=1}^{m} Tij = Ti1 + Ti2 + Ti3 \ldots + Tim$

$D1 = \sum_{i=0}^{n} Ti1 = T11 + T21 + T31 \ldots + Tn1$

$D2 = \sum_{i=0}^{n} T2j = T12 + T22 + T32 \ldots + Tn2$

…

$Dj = \sum_{i=1}^{n} Tij = T1j + T2j + T3j \ldots + Tnj$

Where i = 1 … n,

j = 1 … m

Vehicle trip detectors could be used to collect origin and destination data.

In addition, several constraints are related to special scenarios unique to diamond interchanges. First, the origin and destination traffic volumes are zero for regular traffic. For example, T18 that is from origin 1 and destination 8 should be zero. Second, traffic volume from an origin to a destination at the same location should be zero. T16 is a good example of this. Third, the vehicle trips of the on-ramp and off-ramp in the same direction should be zero. T22, which is from origin 2 to destination 2, should be zero.

Using a linear programming (LP) solver could generate the OD table. Excel solver is good enough to solve this LP problem.

The objective function is,

$$Minimize \sum_{i=0}^{n} \sum_{j=0}^{m} (\mathrm{Tij} - \mathrm{Tdij} + \epsilon)$$

Where i is the number id of the origin

j is the number id of the destination

Tij is the feasible solution of the OD table (Control variables)

Tdij is the detector data for vehicle trips

ϵ is an error variable for a congested network.

Diamond Interchange Model Proof of Concept

The model is implemented and independently evaluated in a microscopic simulation at the US-78 diamond interchange shown in Figure 1. The OD estimation model is implemented as an interface with the Enhanced Traffic Flow Open-source Microscopic Model (ETFOMM), an open source microscopic traffic simulation software sponsored by U.S. DOT. Detector locations are the same as the field data collection plan. The simulation output data contains detector count data which is used in the OD estimation model and real OD data which is used to evaluate model performance.

The results of the US-78 diamond interchange are shown below,

OD Table based on model results from detector data (vehs/15 mins)

	D1	D2	D3	D4	D5	D6
O1	0.0	0.0	875.0	0.0	0.0	0.0

O2	99.0	0.0	0.0	101.0	0.0	0.0
O3	0.0	78.3	0.0	265.0	156.8	0.0
O4	0.0	0.0	0.0	0.0	0.0	875.0
O5	107.8	0.0	0.0	109.8	0.0	0.0
O6	360.3	72.8	0.0	0.0	67.0	0.0

Real OD trips from the simulation trajectory (vehs/15 mins)

	D1	D2	D3	D4	D5	D6
O1	0	0	880	0	0	0
O2	99	0	0	96	0	0
O3	0	82	0	266	155	0
O4	0	0	0	0	0	874
O5	108	0	0	107	0	0
O6	364	75	0	0	63	0

Comparing the OD table from the model and the actual OD trips from the simulation trajectory (15 mins)

	D1	D2	D3	D4	D5	D6
O1	0.00%	0.00%	-0.51%	0.00%	0.00%	0.00%
O2	-0.25%	0.00%	0.00%	4.93%	0.00%	0.00%
O3	0.00%	-5.16%	0.00%	-0.47%	1.29%	0.00%
O4	0.00%	0.00%	0.00%	0.00%	0.00%	0.11%
O5	-0.23%	0.00%	0.00%	2.81%	0.00%	0.00%
O6	-0.96%	-2.68%	0.00%	0.00%	6.78%	0.00%

To build a congested network, the volume is increased, and the time interval is decreased to 5 min to better capture vehicles stored within the system.

The result of US-78 diamond interchange shows below,

OD Table basing on detector data (vehs/ 5 mins)

	D1	D2	D3	D4	D5	D6
O1	0.0	0.0	291.7	0.0	0.0	0.0
O2	33.0	0.0	0.0	33.7	0.0	0.0
O3	0.0	26.1	0.0	88.3	52.3	0.0
O4	0.0	0.0	0.0	0.0	0.0	291.7
O5	35.9	0.0	0.0	36.6	0.0	0.0
O6	120.1	24.3	0.0	0.0	22.3	0.0

Real OD trip from simulation trajectory (vehs/5 mins)

	D1	D2	D3	D4	D5	D6
O1	0	0	292	0	0	0
O2	32	0	0	35	0	0
O3	0	27	0	88	53	0
O4	0	0	0	0	0	293
O5	35	0	0	38	0	0
O6	121	25	0	0	22	0

Comparing with OD table and OD trip from simulation trajectory (5 mins)

	D1	D2	D3	D4	D5	D6
O1	0.00%	0.00%	0.06%	0.00%	0.00%	0.00%
O2	3.67%	0.00%	0.00%	-3.81%	0.00%	0.00%
O3	0.00%	-4.87%	0.00%	0.09%	-1.26%	0.00%
O4	0.00%	0.00%	0.00%	0.00%	0.00%	-0.40%
O5	2.87%	0.00%	0.00%	-3.72%	0.00%	0.00%
O6	-0.55%	-1.36%	0.00%	0.00%	2.29%	0.00%

Area Wide Estimate of OD from Available Data

The idea of an area-wide OD model is to use blue-tooth detector data and loop detector data to estimate the OD table of a large area-wide network. The area-wide model is shown below:

$$Tij = Vij/2 * \frac{Vli * \sum_i Vij + Vlj * \sum_j Vij}{\sum_j Vij + \sum_i Vij}$$

Tij is the vehicular trips from origin i to destination d

Vij is the vehciluar trips from Bluetooth detector, from origin i to destination d

Vli is the vehciluar trips from loop detector at origin i

Vlj is the vehciluar trips from loop detector at destination j

Because Bluetooth detectors only collect vehicle information with Bluetooth devices, the Bluetooth detectors' OD data is partial vehicle OD trip data, or OD percentage data. Traditional loop detector can count all vehicles at a specific location but will not be able to produce any OD information. The actual OD table could be generated by considering OD percentage data from the Bluetooth device and vehicle counts from loop detector.

Area Wide Model Proof of Concept

The area-wide OD estimation model is also implemented in the ETFOMM simulation. ETFOMM chooses different Bluetooth penetration percentages. The results are then used to compare actual traffic flow data from the simulation and area-wide OD model results. Figure 2 shows traffic network of simulation. In

Figure 2, three OD areas are origins and destinations in the network. The OD table would three origins and three destinations matrix.

Figure 2 Area-wide traffic network for simulation

The results of area-wide traffic network are show below

5% Bluetooth detected vehicles

Bluetooth detected vehicles	O1	O2	O3
D1	0	20	24
D2	21	0	21
D3	27	26	0
Bluetooth estimated vehicles	O1	O2	O3
D1	0	400	480
D2	420	0	420
D3	540	520	0
Real traffic vehicles	O1	O2	O3
D1	0	336	478
D2	375	0	468
D3	532	498	0
Comparing	O1	O2	O3
D1	0.0%	-19.0%	-0.4%
D2	-12.0%	0.0%	10.3%
D3	-1.5%	-4.4%	0.0%
Average differences	7.94%		

10% Bluetooth detected vehicles

Bluetooth detected vehicles	O1	O2	O3
D1	0	35	46

D2	39	0	45
D3	52	48	0
Bluetooth estimated vehicles	O1	O2	O3
D1	0	350	460
D2	390	0	450
D3	520	480	0
Real traffic vehicles	O1	O2	O3
D1	0	336	478
D2	375	0	468
D3	532	498	0
Comparing	O1	O2	O3
D1	0.0%	-4.2%	3.8%
D2	-4.0%	0.0%	3.8%
D3	2.3%	3.6%	0.0%
Average difference	3.61%		

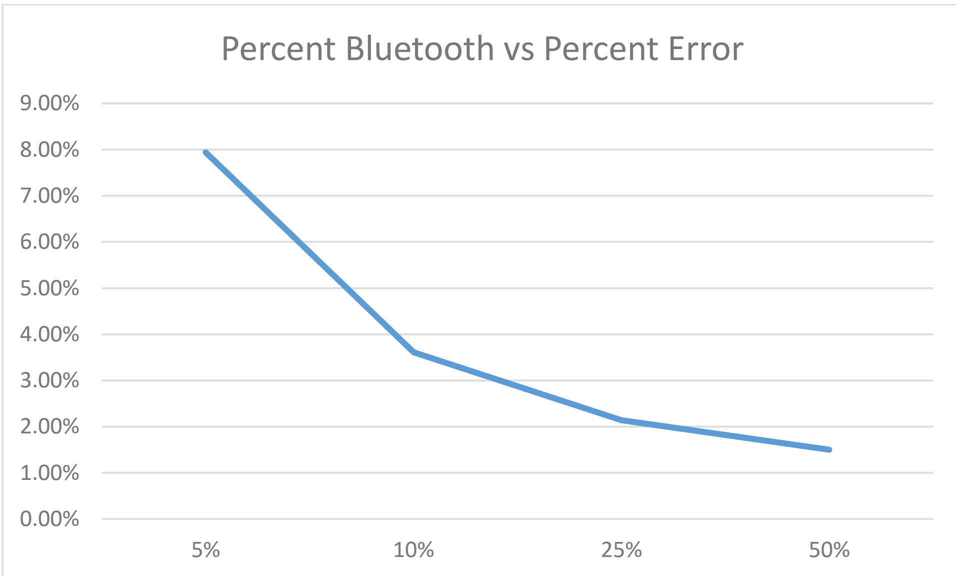
25% Bluetooth detected vehicles

Bluetooth detected vehicles	O1	O2	O3
D1	0	81	120
D2	98	0	116
D3	136	123	0
Bluetooth estimated vehicles	O1	O2	O3
D1	0	324	480
D2	392	0	464
D3	544	492	0
Real traffic vehicles	O1	O2	O3
D1	0	336	478
D2	375	0	468
D3	532	498	0
Comparing	O1	O2	O3
D1	0.0%	3.6%	-0.4%
D2	-4.5%	0.0%	0.9%
D3	-2.3%	1.2%	0.0%
Average difference	2.14%		

50% Bluetooth detected vehicles

Bluetooth detected vehicles	O1	O2	O3
D1	0	168	235
D2	186	0	225
D3	260	250	0

Bluetooth estimated vehicles	O1	O2	O3
D1	0	336	470
D2	372	0	450
D3	520	500	0
Real traffic vehicles	O1	O2	O3
D1	0	336	478
D2	375	0	468
D3	532	498	0
Comparing	O1	O2	O3
D1	0.0%	0.0%	1.7%
D2	0.8%	0.0%	3.8%
D3	2.3%	-0.4%	0.0%
Average difference	1.50%		



Field Data Collection

Field data was collected for both the area wide and intersection specific models. This data was collected on May 25th, area wide, and 26th, interchange specific, of 2016 on a stretch of HWY-302 between I-55 and US-78 in Northwestern Mississippi. The Mississippi Department of Transportation provided assistance in collecting the traditional loop detector data through Michael Baker Engineering. This data was collected using nine video cameras. The Bluetooth data was then collected using probe vehicle data since Bluetooth detectors were not available during our project time frame. The probe vehicle was used by following a vehicle from system entry to exit multiple times. Enough drivers were used to provide 10% of the total volume during the peak hour. This data was recorded on three different apps and in individual manual logs. The app data provided KML files which were processed to give the actual OD which was then used in the case study.

The field data collection plan contains two parts. The first is the probe vehicle data plan to collect OD trips and the travel time of each OD trip. The second is the video camera detector plan to collect the total volume of each origin and destination of both the area wide study and diamond interchange study.

Day 1: Area Wide

The defined study area is shown below:

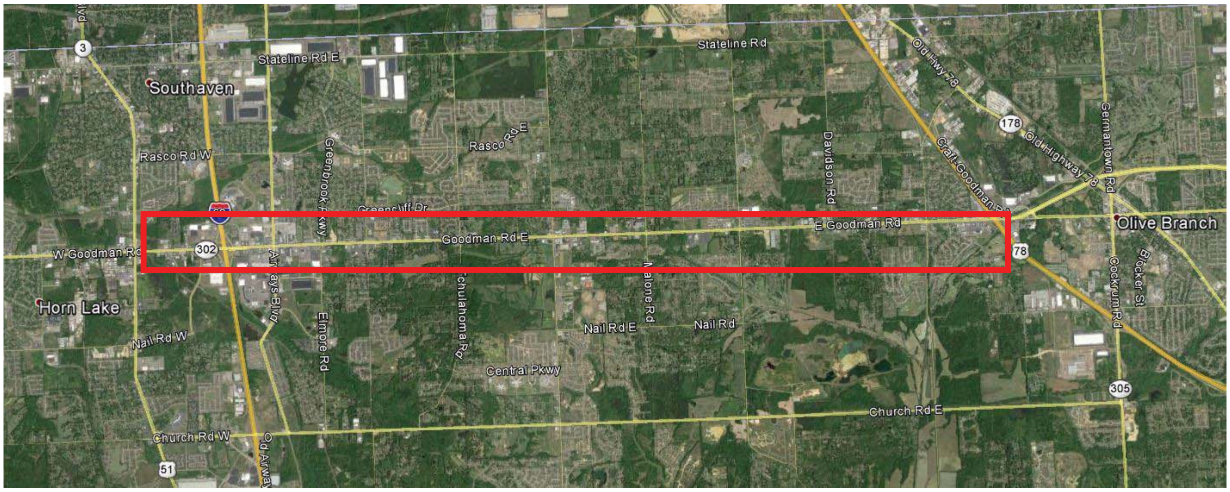
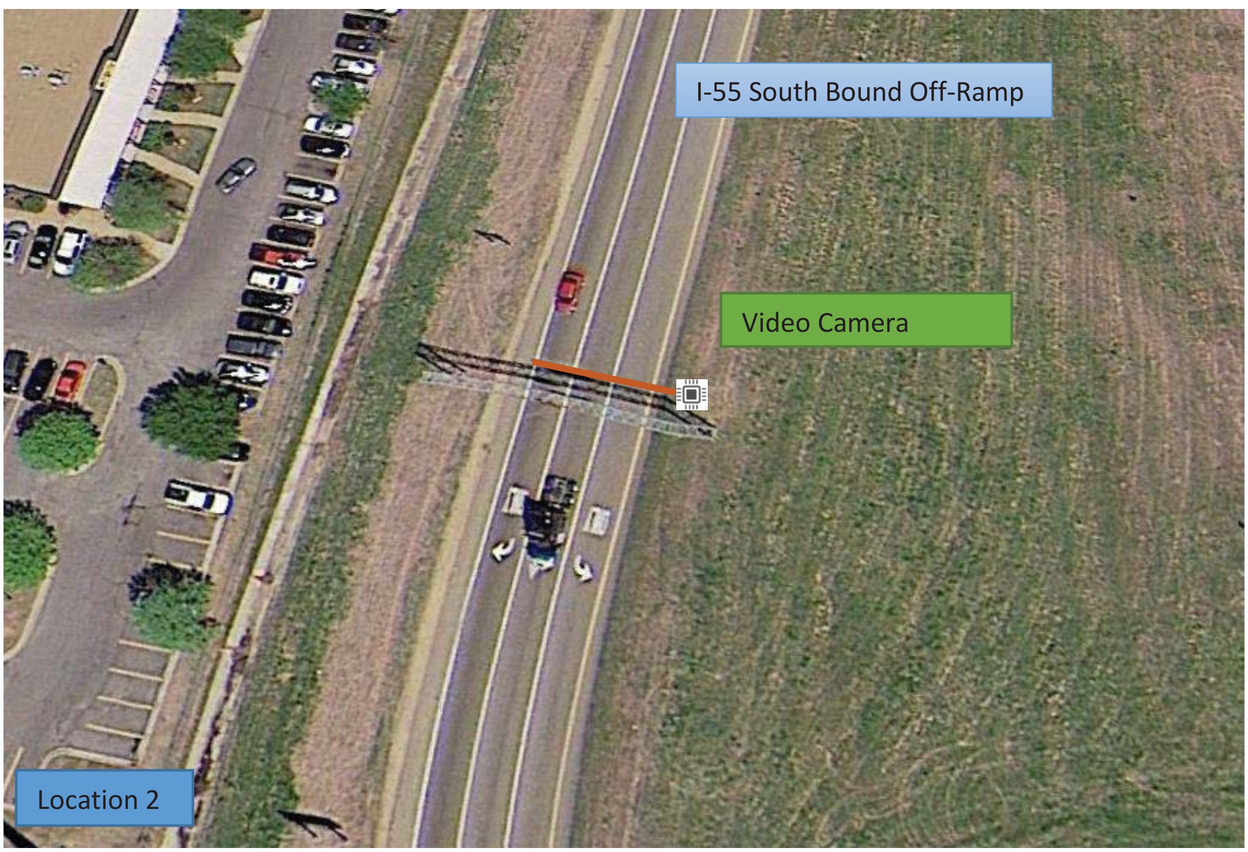


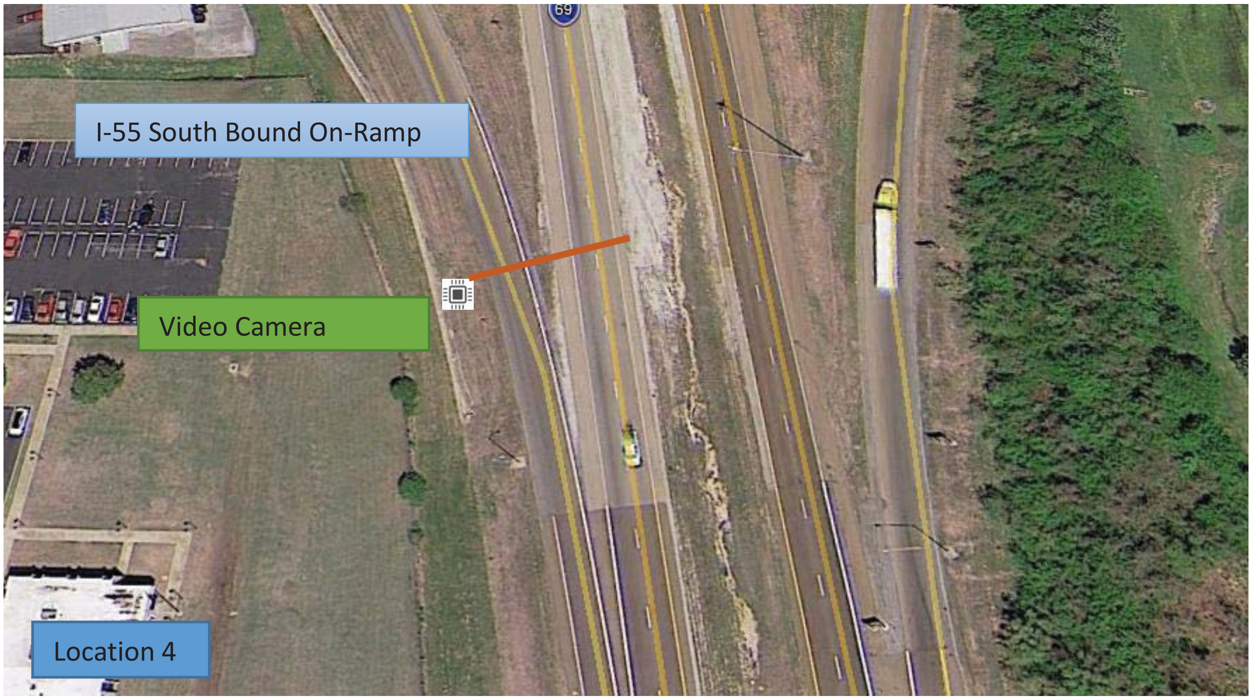
Figure 3 Area wide study location

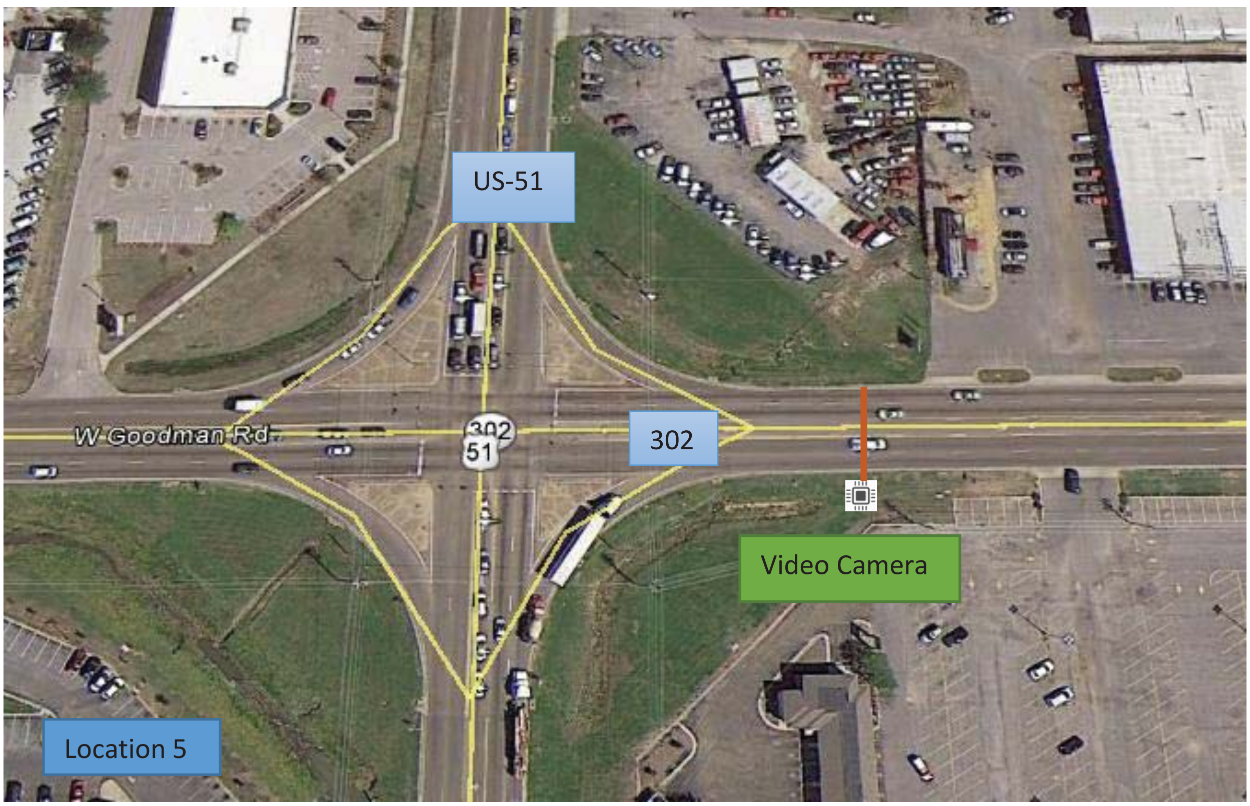
This is an area of HWY 302 between I-55 and US-78 in the South Haven/ Olive Branch area.

Drivers are to enter the system at any of the entry points of concern following another vehicle. Drivers need to follow this vehicle until it exists the system at any exit point. Once probe vehicles have exited the system, drivers would turn around and return to the closest entry point of concern and repeat the process. The entries of concern are shown in the following figures.

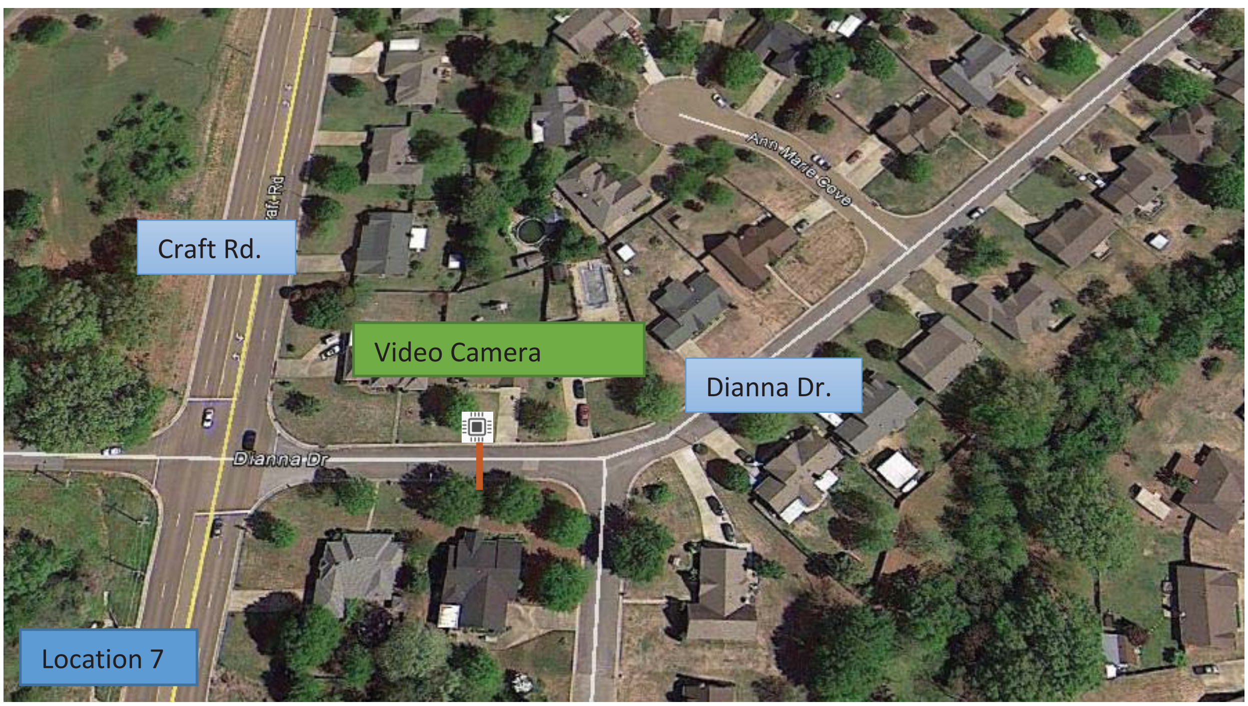

Interstate Blvd.
Video Camera
Location 1
302


I-55 South Bound Off-Ramp
Video Camera
Location 2


I-55 South Bound On-Ramp
Video Camera
Location 4
69


US-51
W Goodman Rd
302
51
302
Video Camera
Location 5

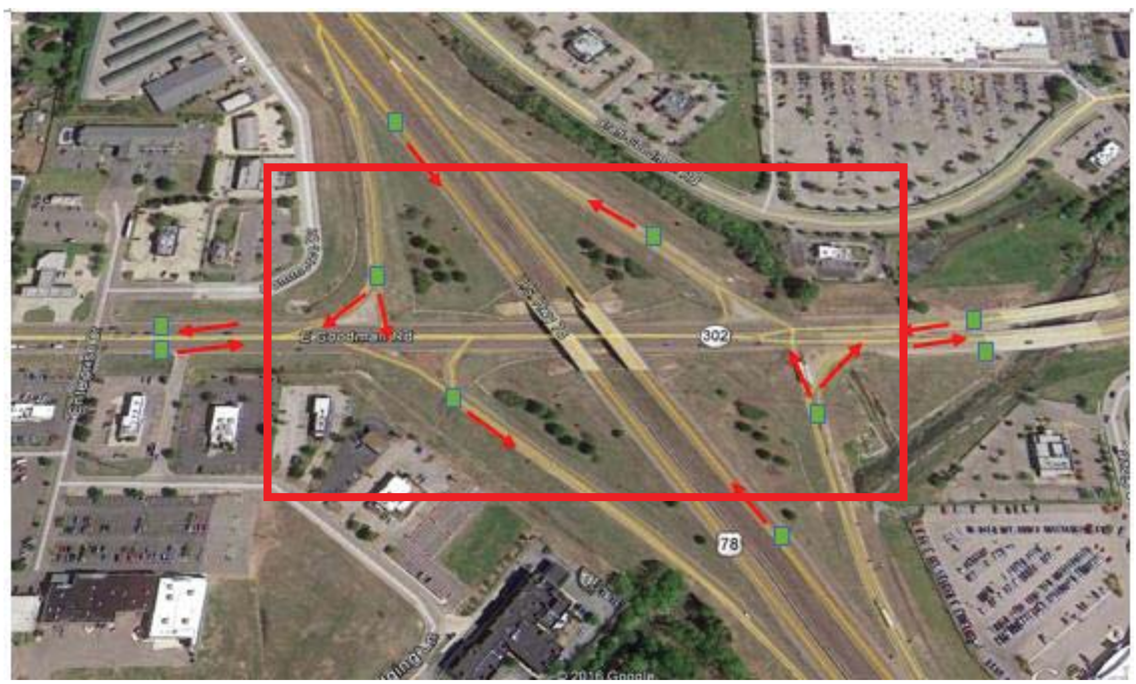
302
Video Camera
Oak Forest Dr.
Location 6
Oak Forest Dr.



Craft Rd.
Video Camera
Dianna Dr.
Location 7

Day 2: Individual Interchange

For this day, drivers are to do exactly the same thing however the location changes, the interchange of US-78 and HWY 302, and decreases in size with no non-important entries and exits. The figure below shows the area of interest.



Six video cameras are set within the interchange to collect traffic volume and turning movement data.

Model Case Study Results:

In the Day 1 area wide study, the OD table includes eight points where traffic volume and OD data were collected by the above method. Each of the eight point's descriptions are shown in the table below.

Table 2 Eight points of traffic volume capture

ID	Description
A	Interstate Dr Near Goodman Rd
B	Goodman Rd at I-55 SB Ramps
C	I-55 SB Loop and SB Lanes
D	Goodman Rd Near US 51
E	Oak Forest Dr Near Goodman Rd
F	Dianna Dr Near Craft Rd
G	Lauren Ln at Crumpler Blvd
H	Goodman Rd Near Crumpler Blvd

Figures 4-11 display the traffic volume data of the eight points.

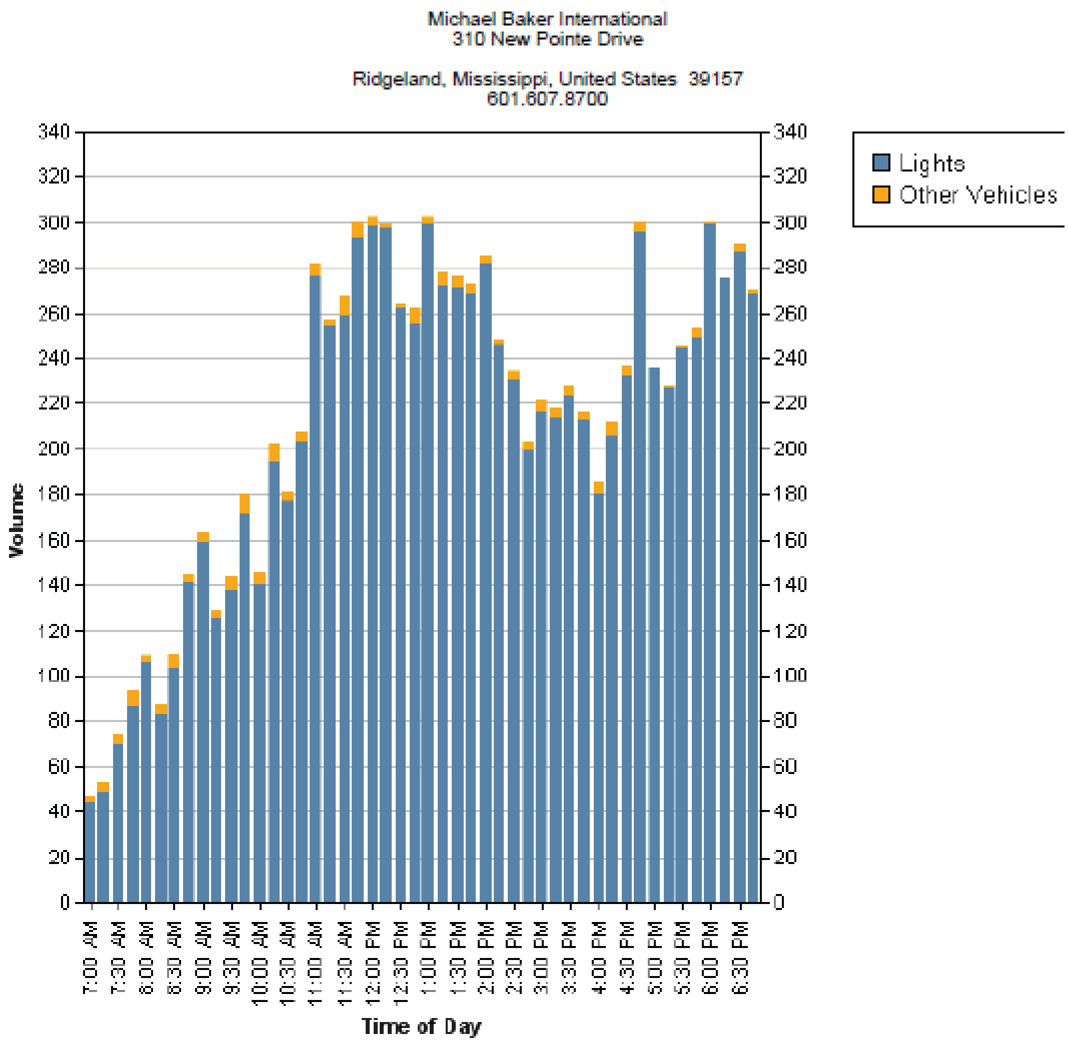


Figure 4 Point A traffic volume

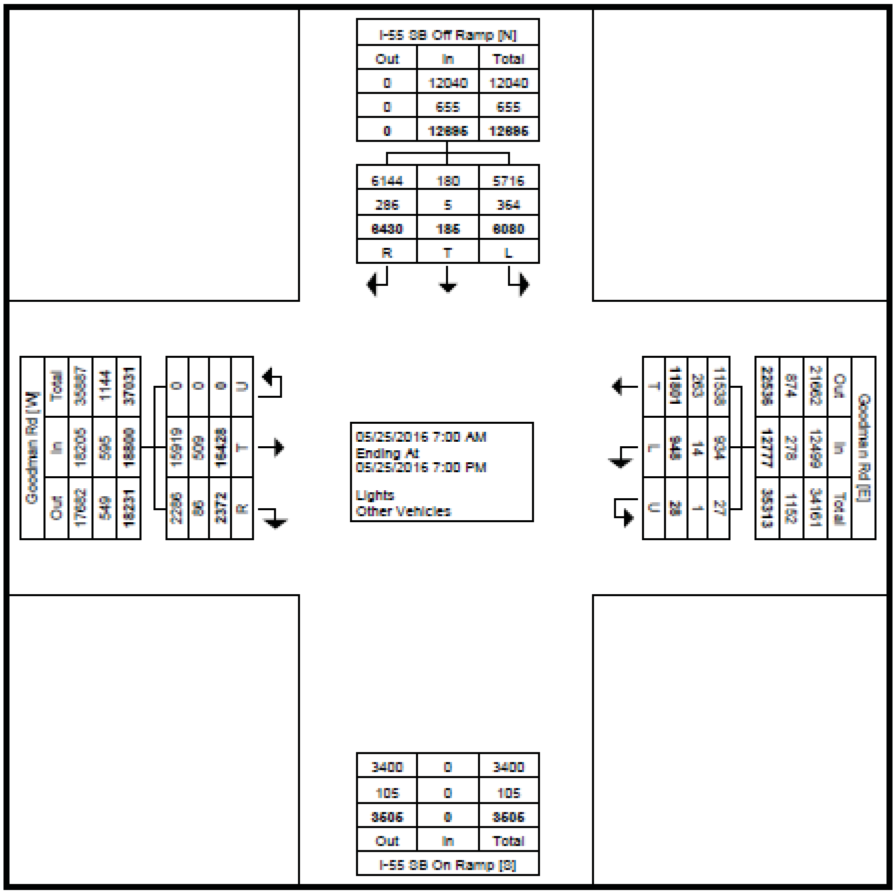


Figure 5 Point B traffic volume and turning percentage

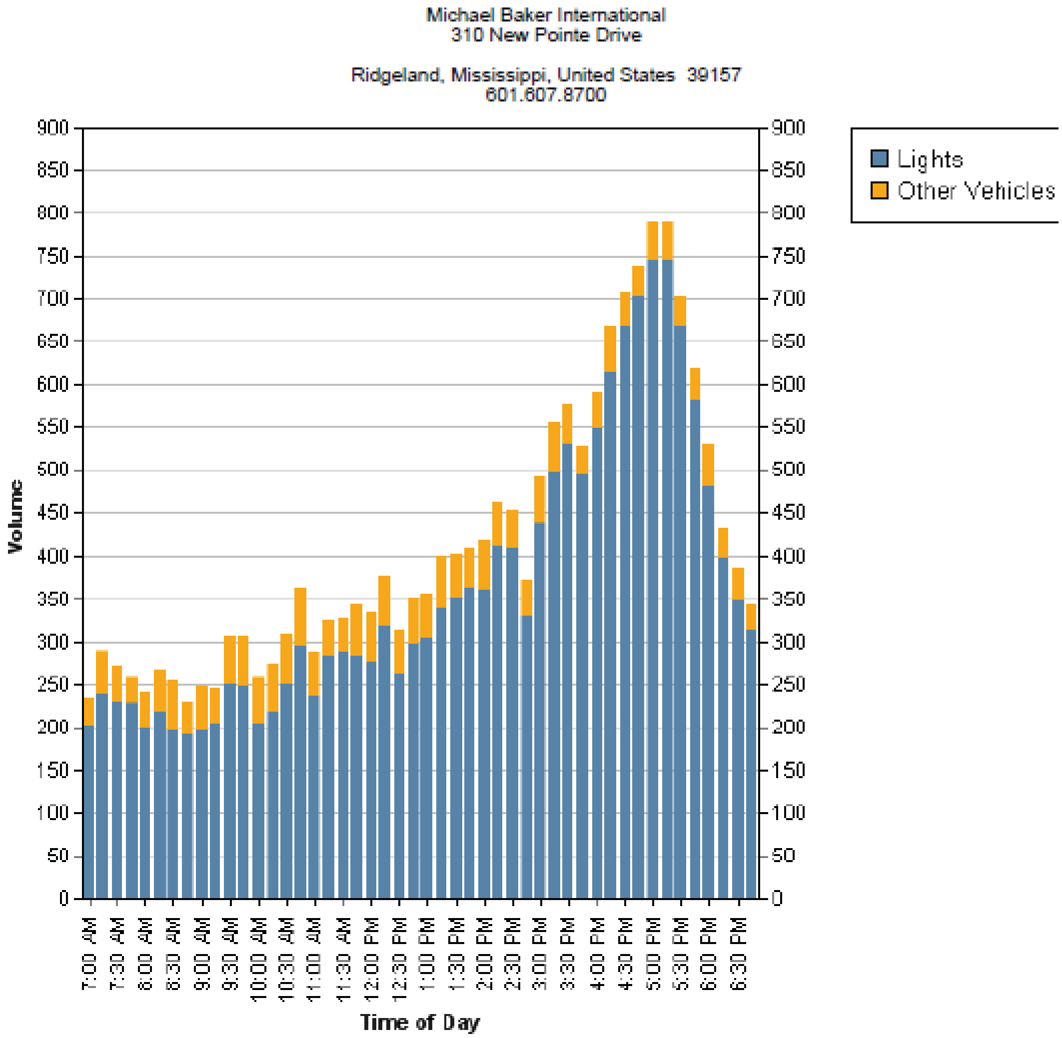


Figure 6 Point C traffic volume

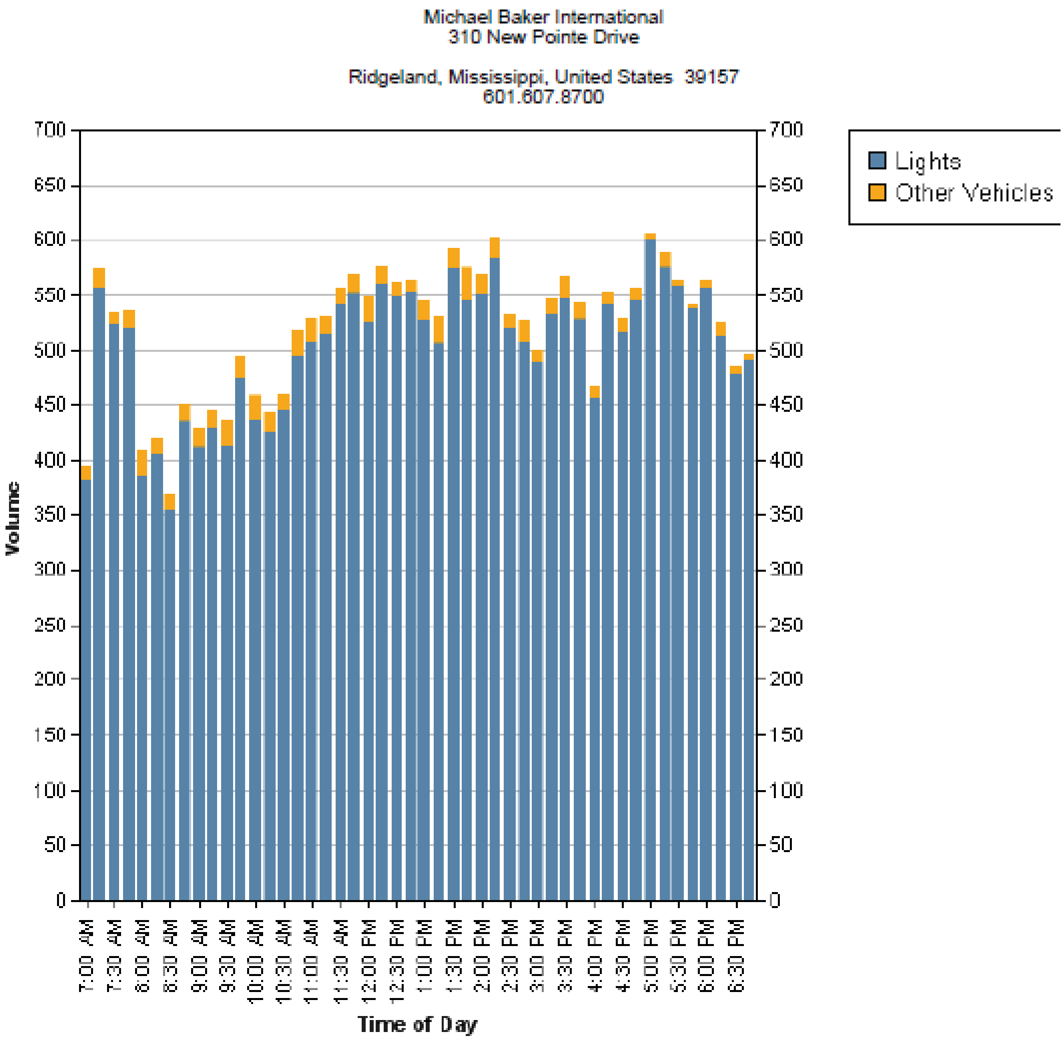


Figure 7 Point D traffic volume

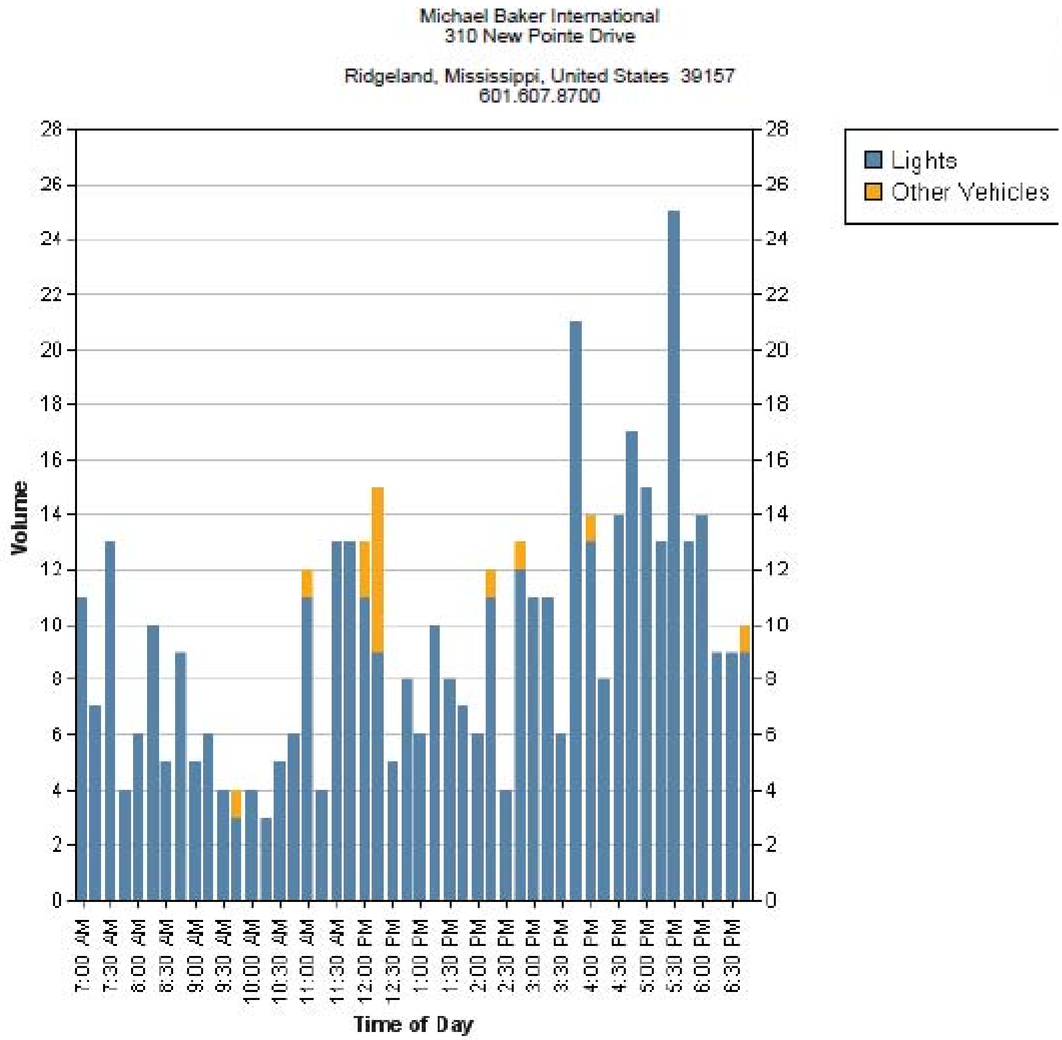


Figure 8 Point E traffic volume

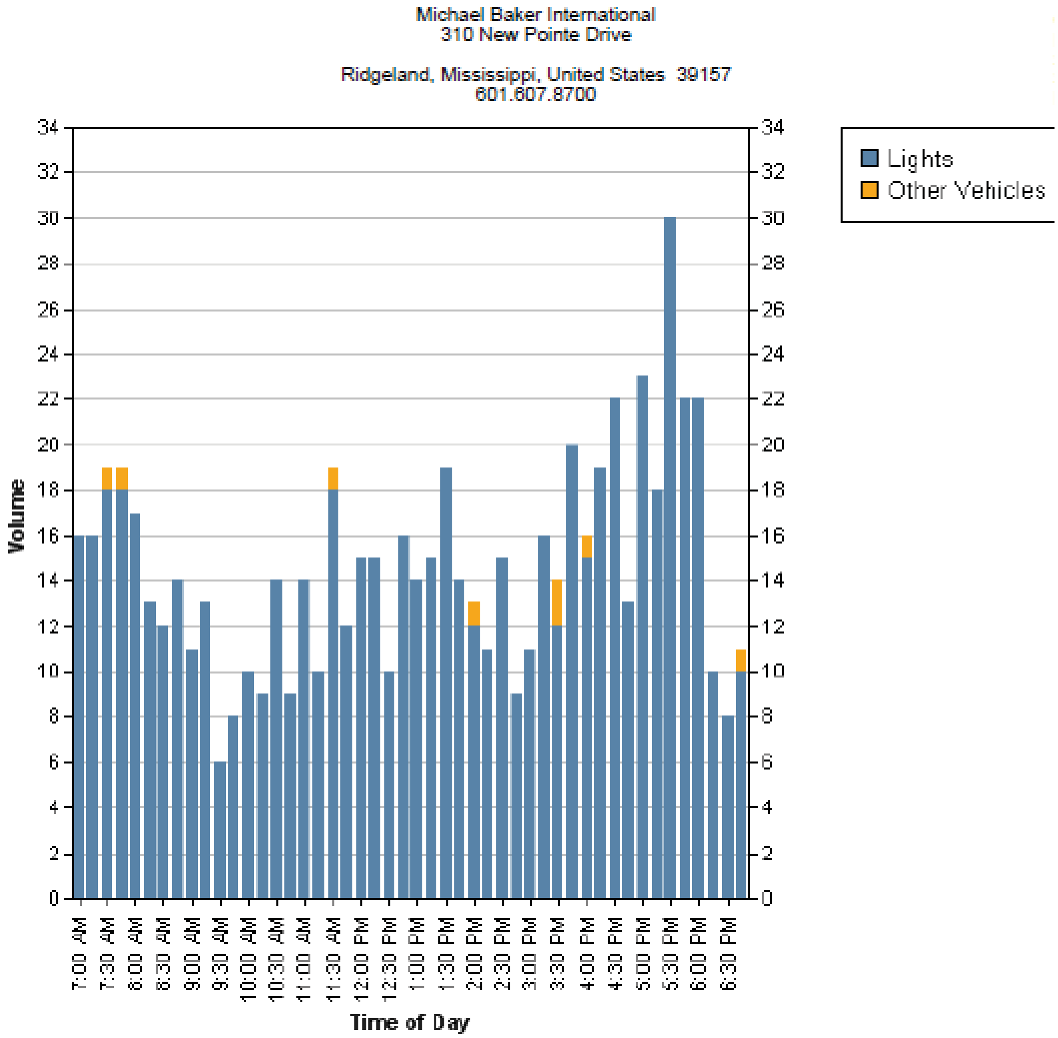


Figure 9 Point F traffic volume

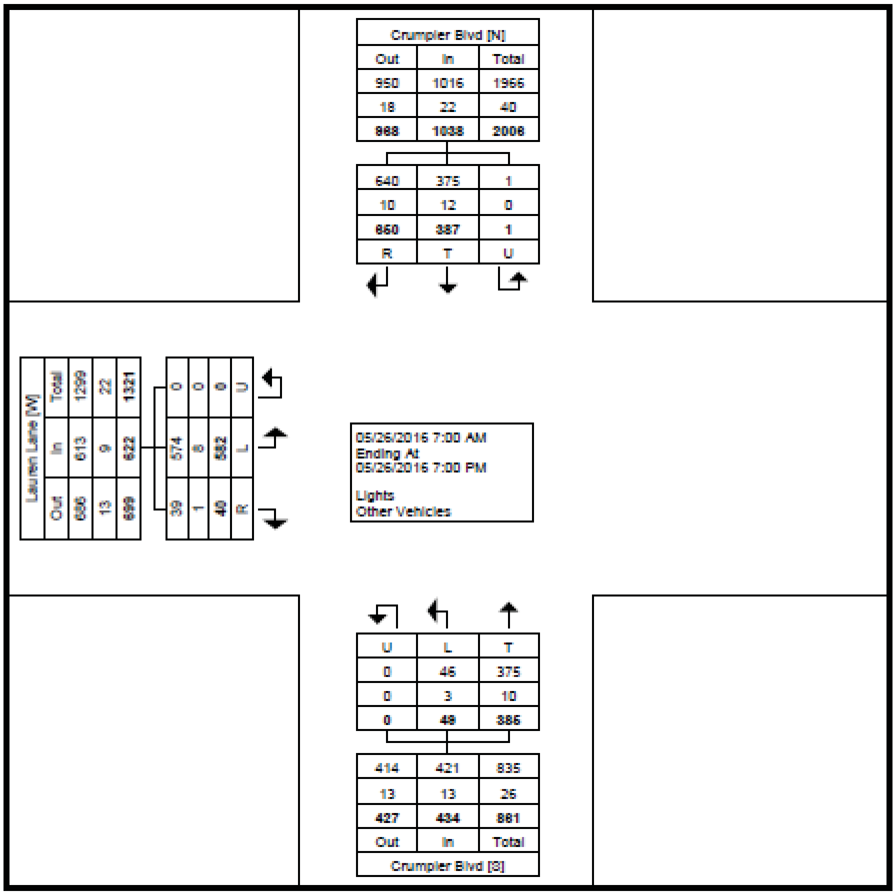


Figure 10 Point G traffic volume and turning percentage

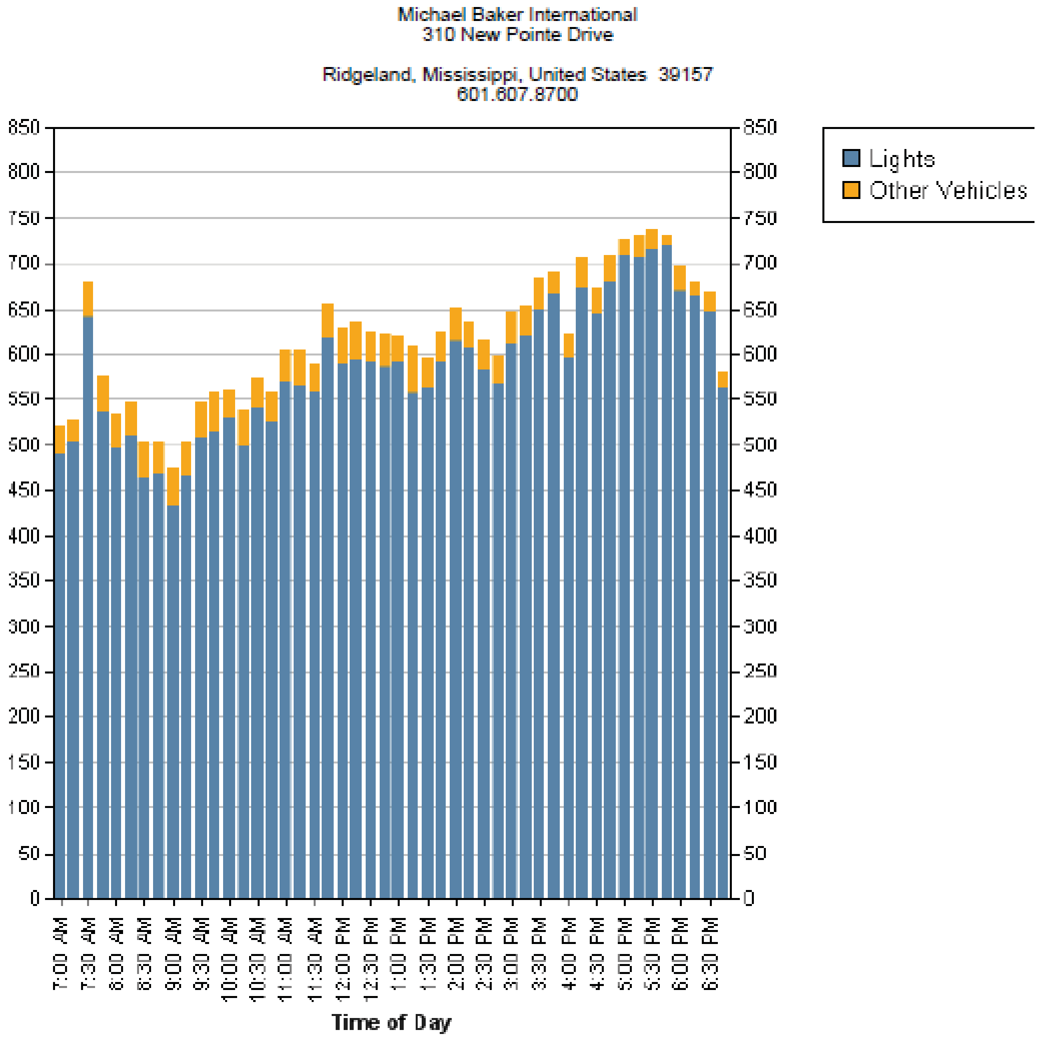


Figure 11 Point H traffic volume

Day 1 GPS data

Probe vehicles only follow the vehicles out of these eight points until they exit the system. This leaves many trips that are unusable since they do not go between two points of concern. The below table shows the OD trips between the eight points.

Table 3 OD trip table of 05/25/2016

	A	B	C	D	E	F	G	H
A			9	11		1		1
B	7			7	1			1
C				1				
D	3		4					3
E								8
F	1							8
G				1				9
H			1	1				

Although the area wide OD estimation model has a good performance in the simulation environment, the probe vehicle percentages are lower than one percentage of total traffic volume due to all of the unusable trips. This results in the area wide OD model performing poorly using this field data due to the low penetration rate and nature of probe vehicles not being able to evenly capture vehicles across every point. With Bluetooth detectors, this issue would be corrected and the model should perform as well as it proves capable of performing in the simulation.

Day 2 DDI OD table estimation study only set cameras to three locations that cover all origin and destination traffic flow volumes shown in Figure 1. The tables below show the origin and destination vehicle counts by the video detectors.

Table 4 Origins detected traffic volume

Time	O1	O2	O3	O4	O5	O6
7:00 AM	111	46	291	263	114	256
7:15 AM	127	57	280	280	130	283
7:30 AM	133	49	339	241	128	309
7:45 AM	140	62	333	241	117	312
8:00 AM	118	59	290	205	97	276
8:15 AM	125	55	296	180	109	269
8:30 AM	134	36	305	163	114	255
8:45 AM	118	59	261	170	114	247
9:00 AM	113	50	273	163	109	229
9:15 AM	118	66	278	169	94	241
9:30 AM	138	51	330	173	104	292
9:45 AM	147	63	338	149	111	310
10:00 AM	126	65	334	146	122	289
10:15 AM	148	63	319	174	96	274
10:30 AM	145	60	366	147	114	311
10:45 AM	136	65	365	165	128	294
11:00 AM	170	60	352	153	128	304
11:15 AM	135	77	377	145	112	329
11:30 AM	147	85	387	153	120	312
11:45 AM	145	84	408	147	128	338
12:00 PM	150	76	417	156	137	342
12:15 PM	165	80	426	148	138	360
12:30 PM	140	83	441	150	117	372
12:45 PM	158	80	433	151	101	361
1:00 PM	172	77	426	167	141	347
1:15 PM	167	73	406	148	113	329
1:30 PM	163	66	422	160	121	359

1:45 PM	166	65	388	154	123	310
2:00 PM	147	70	442	162	89	373
2:15 PM	163	58	439	148	119	330
2:30 PM	176	68	392	182	112	346
2:45 PM	172	73	406	188	108	328
3:00 PM	198	90	425	184	91	361
3:15 PM	209	63	444	174	104	352
3:30 PM	213	104	388	147	106	340
3:45 PM	202	90	403	144	123	325
4:00 PM	205	99	433	168	95	373
4:15 PM	233	75	441	166	120	354
4:30 PM	254	81	474	188	112	384
4:45 PM	238	84	448	170	116	368
5:00 PM	250	87	531	154	121	390
5:15 PM	254	105	463	162	112	387
5:30 PM	257	108	487	166	102	415
5:45 PM	242	105	456	191	91	402
6:00 PM	253	92	394	163	126	349
6:15 PM	193	96	439	144	119	360
6:30 PM	162	73	412	133	90	365
6:45 PM	148	75	350	126	90	283

Table 5 Destinations detected traffic volume

Time	D1	D2	D3	D4	D5	D6	D7	D8
7:00 AM	326	66	111	264	68	263	260	296
7:15 AM	347	54	127	285	102	280	277	368
7:30 AM	435	72	133	317	73	241	306	422
7:45 AM	350	74	140	302	76	241	311	353
8:00 AM	367	66	118	274	72	205	275	352
8:15 AM	341	71	125	274	58	180	265	347
8:30 AM	343	77	134	257	61	163	268	342
8:45 AM	323	71	118	244	76	170	247	336
9:00 AM	313	96	113	234	46	163	226	314
9:15 AM	337	92	118	236	56	169	238	322
9:30 AM	331	71	138	303	46	173	293	301
9:45 AM	350	90	147	323	49	149	306	352
10:00 AM	360	98	126	304	55	146	293	342
10:15 AM	345	98	148	275	52	174	270	349
10:30 AM	344	87	145	329	34	147	332	346
10:45 AM	382	101	136	301	56	165	308	349
11:00 AM	390	99	170	308	65	153	296	395

11:15 AM	383	115	135	326	56	145	333	384
11:30 AM	397	137	147	329	61	153	310	377
11:45 AM	439	110	145	337	63	147	338	417
12:00 PM	436	139	150	338	83	156	348	427
12:15 PM	430	134	165	351	79	148	356	405
12:30 PM	391	151	140	373	72	150	365	365
12:45 PM	401	152	158	355	66	151	350	412
1:00 PM	430	132	172	343	77	167	362	408
1:15 PM	390	152	167	328	66	148	329	397
1:30 PM	380	137	163	361	77	160	360	389
1:45 PM	410	136	166	319	59	154	314	415
2:00 PM	387	155	147	366	69	162	362	385
2:15 PM	389	169	163	339	51	148	334	390
2:30 PM	393	141	176	340	67	182	335	415
2:45 PM	402	145	172	334	53	188	330	390
3:00 PM	391	156	198	367	64	184	359	406
3:15 PM	423	170	209	351	68	174	342	417
3:30 PM	438	146	213	349	54	147	340	436
3:45 PM	500	146	202	328	65	144	333	473
4:00 PM	420	160	205	373	70	168	366	436
4:15 PM	471	169	233	376	54	166	351	472
4:30 PM	433	148	254	393	46	188	399	423
4:45 PM	452	157	238	380	58	170	369	449
5:00 PM	480	207	250	406	55	154	403	476
5:15 PM	488	186	254	398	51	162	380	507
5:30 PM	466	173	257	423	64	166	424	478
5:45 PM	506	175	242	414	54	191	395	502
6:00 PM	482	144	253	355	61	163	337	480
6:15 PM	445	169	193	380	61	144	357	433
6:30 PM	444	136	162	368	44	133	357	458
6:45 PM	383	129	148	293	41	126	286	370

In Table 4, D7 and D8 are not columns in the OD trip table, they are only helpful to solve the OD model to estimate OD tables. After processing the origin and destination volumes shown in the previous two tables, the results of the OD estimation model are shown in tables below:

Table 6 05/26/2016 7:00AM-7:00PM OD table

7:00AM	D1	D2	D3	D4	D5	D6
O1	0	0	111	0	0	0
O2	28	0	0	18	0	0
O3	0	50	0	202	39	0

O4	0	0	0	0	0	263
O5	72	0	0	42	0	0
O6	221	13	0	0	22	0
7:15AM	D1	D2	D3	D4	D5	D6
O1	0	0	127	0	0	0
O2	32	0	0	25	0	0
O3	0	34	0	204	41	0
O4	0	0	0	0	0	280
O5	78	0	0	52	0	0
O6	235	16	0	0	33	0
7:30AM	D1	D2	D3	D4	D5	D6
O1	0	0	133	0	0	0
O2	33	0	0	16	0	0
O3	0	40	0	243	56	0
O4	0	0	0	0	0	241
O5	82	0	0	46	0	0
O6	297	0	0	0	12	0
7:45AM	D1	D2	D3	D4	D5	D6
O1	0	0	140	0	0	0
O2	33	0	0	29	0	0
O3	0	55	0	227	51	0
O4	0	0	0	0	0	241
O5	70	0	0	47	0	0
O6	258	22	0	0	31	0
8:00AM	D1	D2	D3	D4	D5	D6
O1	0	0	118	0	0	0
O2	33	0	0	26	0	0
O3	0	42	0	211	37	0
O4	0	0	0	0	0	205
O5	62	0	0	35	0	0
O6	246	10	0	0	19	0
8:15AM	D1	D2	D3	D4	D5	D6
O1	0	0	125	0	0	0
O2	32	0	0	23	0	0
O3	0	55	0	206	36	0
O4	0	0	0	0	0	180
O5	69	0	0	40	0	0
O6	235	14	0	0	20	0
8:30AM	D1	D2	D3	D4	D5	D6
O1	0	0	134	0	0	0
O2	26	0	0	10	0	0

O3	0	62	0	205	39	0
O4	0	0	0	0	0	163
O5	76	0	0	38	0	0
O6	230	9	0	0	15	0
8:45AM	D1	D2	D3	D4	D5	D6
O1	0	0	118	0	0	0
O2	37	0	0	22	0	0
O3	0	54	0	182	25	0
O4	0	0	0	0	0	170
O5	75	0	0	39	0	0
O6	209	16	0	0	23	0
9:00AM	D1	D2	D3	D4	D5	D6
O1	0	0	113	0	0	0
O2	35	0	0	15	0	0
O3	0	60	0	183	29	0
O4	0	0	0	0	0	163
O5	75	0	0	34	0	0
O6	203	10	0	0	17	0
9:15AM	D1	D2	D3	D4	D5	D6
O1	0	0	118	0	0	0
O2	43	0	0	23	0	0
O3	0	62	0	185	31	0
O4	0	0	0	0	0	169
O5	67	0	0	27	0	0
O6	207	14	0	0	21	0
9:30AM	D1	D2	D3	D4	D5	D6
O1	0	0	138	0	0	0
O2	22	0	0	29	0	0
O3	0	54	0	229	47	0
O4	0	0	0	0	0	173
O5	59	0	0	45	0	0
O6	250	18	0	0	24	0
9:45AM	D1	D2	D3	D4	D5	D6
O1	0	0	147	0	0	0
O2	25	0	0	38	0	0
O3	0	67	0	235	36	0
O4	0	0	0	0	0	149
O5	59	0	0	52	0	0
O6	271	24	0	0	16	0
10:00AM	D1	D2	D3	D4	D5	D6
O1	0	0	126	0	0	0

O2	33	0	0	32	0	0
O3	0	77	0	222	36	0
O4	0	0	0	0	0	146
O5	72	0	0	50	0	0
O6	255	21	0	0	14	0
10:15AM	D1	D2	D3	D4	D5	D6
O1	0	0	148	0	0	0
O2	36	0	0	27	0	0
O3	0	79	0	209	31	0
O4	0	0	0	0	0	174
O5	64	0	0	32	0	0
O6	245	18	0	0	11	0
10:30AM	D1	D2	D3	D4	D5	D6
O1	0	0	145	0	0	0
O2	24	0	0	36	0	0
O3	0	75	0	244	46	0
O4	0	0	0	0	0	147
O5	63	0	0	51	0	0
O6	281	19	0	0	12	0
10:45AM	D1	D2	D3	D4	D5	D6
O1	0	0	136	0	0	0
O2	39	0	0	26	0	0
O3	0	87	0	230	48	0
O4	0	0	0	0	0	165
O5	82	0	0	46	0	0
O6	270	15	0	0	10	0
11:00AM	D1	D2	D3	D4	D5	D6
O1	0	0	170	0	0	0
O2	35	0	0	25	0	0
O3	0	79	0	228	45	0
O4	0	0	0	0	0	153
O5	80	0	0	48	0	0
O6	271	18	0	0	15	0
11:15AM	D1	D2	D3	D4	D5	D6
O1	0	0	135	0	0	0
O2	37	0	0	40	0	0
O3	0	88	0	241	48	0
O4	0	0	0	0	0	145
O5	66	0	0	46	0	0
O6	283	27	0	0	18	0
11:30AM	D1	D2	D3	D4	D5	D6

O1	0	0	147	0	0	0
O2	44	0	0	41	0	0
O3	0	99	0	239	48	0
O4	0	0	0	0	0	153
O5	73	0	0	47	0	0
O6	279	22	0	0	10	0
11:45AM	D1	D2	D3	D4	D5	D6
O1	0	0	145	0	0	0
O2	47	0	0	37	0	0
O3	0	97	0	254	57	0
O4	0	0	0	0	0	147
O5	81	0	0	47	0	0
O6	313	18	0	0	8	0
12:00PM	D1	D2	D3	D4	D5	D6
O1	0	0	150	0	0	0
O2	43	0	0	33	0	0
O3	0	107	0	249	60	0
O4	0	0	0	0	0	156
O5	85	0	0	52	0	0
O6	299	26	0	0	16	0
12:15PM	D1	D2	D3	D4	D5	D6
O1	0	0	165	0	0	0
O2	42	0	0	38	0	0
O3	0	104	0	258	63	0
O4	0	0	0	0	0	148
O5	81	0	0	57	0	0
O6	307	31	0	0	21	0
12:30PM	D1	D2	D3	D4	D5	D6
O1	0	0	140	0	0	0
O2	32	0	0	51	0	0
O3	0	120	0	264	57	0
O4	0	0	0	0	0	150
O5	59	0	0	58	0	0
O6	300	46	0	0	25	0
12:45PM	D1	D2	D3	D4	D5	D6
O1	0	0	158	0	0	0
O2	34	0	0	46	0	0
O3	0	116	0	264	53	0
O4	0	0	0	0	0	151
O5	55	0	0	46	0	0
O6	312	36	0	0	13	0

1:00PM	D1	D2	D3	D4	D5	D6
O1	0	0	172	0	0	0
O2	43	0	0	34	0	0
O3	0	111	0	254	61	0
O4	0	0	0	0	0	167
O5	85	0	0	56	0	0
O6	309	21	0	0	16	0
1:15PM	D1	D2	D3	D4	D5	D6
O1	0	0	167	0	0	0
O2	36	0	0	37	0	0
O3	0	115	0	242	49	0
O4	0	0	0	0	0	148
O5	66	0	0	47	0	0
O6	288	26	0	0	16	0
1:30PM	D1	D2	D3	D4	D5	D6
O1	0	0	163	0	0	0
O2	22	0	0	44	0	0
O3	0	110	0	257	54	0
O4	0	0	0	0	0	160
O5	61	0	0	60	0	0
O6	298	35	0	0	25	0
1:45PM	D1	D2	D3	D4	D5	D6
O1	0	0	166	0	0	0
O2	38	0	0	27	0	0
O3	0	115	0	230	43	0
O4	0	0	0	0	0	154
O5	77	0	0	46	0	0
O6	282	19	0	0	9	0
2:00PM	D1	D2	D3	D4	D5	D6
O1	0	0	147	0	0	0
O2	22	0	0	48	0	0
O3	0	119	0	271	52	0
O4	0	0	0	0	0	162
O5	42	0	0	47	0	0
O6	322	35	0	0	16	0
2:15PM	D1	D2	D3	D4	D5	D6
O1	0	0	163	0	0	0
O2	26	0	0	32	0	0
O3	0	141	0	254	44	0
O4	0	0	0	0	0	148
O5	67	0	0	52	0	0

O6	295	28	0	0	7	0
2:30PM	D1	D2	D3	D4	D5	D6
O1	0	0	176	0	0	0
O2	27	0	0	41	0	0
O3	0	109	0	245	38	0
O4	0	0	0	0	0	182
O5	59	0	0	53	0	0
O6	297	30	0	0	19	0
2:45PM	D1	D2	D3	D4	D5	D6
O1	0	0	172	0	0	0
O2	33	0	0	40	0	0
O3	0	120	0	247	40	0
O4	0	0	0	0	0	188
O5	61	0	0	47	0	0
O6	295	24	0	0	9	0
3:00PM	D1	D2	D3	D4	D5	D6
O1	0	0	198	0	0	0
O2	33	0	0	57	0	0
O3	0	120	0	261	44	0
O4	0	0	0	0	0	184
O5	44	0	0	47	0	0
O6	308	35	0	0	18	0
3:15PM	D1	D2	D3	D4	D5	D6
O1	0	0	209	0	0	0
O2	26	0	0	37	0	0
O3	0	133	0	261	51	0
O4	0	0	0	0	0	174
O5	57	0	0	47	0	0
O6	305	32	0	0	15	0
3:30PM	D1	D2	D3	D4	D5	D6
O1	0	0	213	0	0	0
O2	47	0	0	57	0	0
O3	0	114	0	242	32	0
O4	0	0	0	0	0	147
O5	58	0	0	48	0	0
O6	301	28	0	0	11	0
3:45PM	D1	D2	D3	D4	D5	D6
O1	0	0	202	0	0	0
O2	49	0	0	41	0	0
O3	0	125	0	240	39	0
O4	0	0	0	0	0	144

O5	76	0	0	47	0	0
O6	298	21	0	0	6	0
4:00PM	D1	D2	D3	D4	D5	D6
O1	0	0	205	0	0	0
O2	41	0	0	58	0	0
O3	0	123	0	262	48	0
O4	0	0	0	0	0	168
O5	49	0	0	46	0	0
O6	315	37	0	0	21	0
4:15PM	D1	D2	D3	D4	D5	D6
O1	0	0	233	0	0	0
O2	29	0	0	46	0	0
O3	0	132	0	265	44	0
O4	0	0	0	0	0	166
O5	62	0	0	58	0	0
O6	314	31	0	0	9	0
4:30PM	D1	D2	D3	D4	D5	D6
O1	0	0	254	0	0	0
O2	32	0	0	49	0	0
O3	0	127	0	291	55	0
O4	0	0	0	0	0	188
O5	58	0	0	54	0	0
O6	355	26	0	0	4	0
4:45PM	D1	D2	D3	D4	D5	D6
O1	0	0	238	0	0	0
O2	38	0	0	46	0	0
O3	0	124	0	277	47	0
O4	0	0	0	0	0	170
O5	65	0	0	51	0	0
O6	349	19	0	0	0	0
5:00PM	D1	D2	D3	D4	D5	D6
O1	0	0	250	0	0	0
O2	40	0	0	47	0	0
O3	0	174	0	303	54	0
O4	0	0	0	0	0	154
O5	67	0	0	54	0	0
O6	361	29	0	0	0	0
5:15PM	D1	D2	D3	D4	D5	D6
O1	0	0	254	0	0	0
O2	49	0	0	56	0	0
O3	0	151	0	280	32	0

O4	0	0	0	0	0	162
O5	63	0	0	49	0	0
O6	359	28	0	0	0	0
5:30PM	D1	D2	D3	D4	D5	D6
O1	0	0	257	0	0	0
O2	41	0	0	67	0	0
O3	0	140	0	301	47	0
O4	0	0	0	0	0	166
O5	49	0	0	53	0	0
O6	370	33	0	0	12	0
5:45PM	D1	D2	D3	D4	D5	D6
O1	0	0	242	0	0	0
O2	40	0	0	65	0	0
O3	0	129	0	290	36	0
O4	0	0	0	0	0	191
O5	43	0	0	48	0	0
O6	365	29	0	0	8	0
6:00PM	D1	D2	D3	D4	D5	D6
O1	0	0	253	0	0	0
O2	52	0	0	40	0	0
O3	0	127	0	252	16	0
O4	0	0	0	0	0	163
O5	79	0	0	47	0	0
O6	343	6	0	0	0	0
6:15PM	D1	D2	D3	D4	D5	D6
O1	0	0	193	0	0	0
O2	42	0	0	54	0	0
O3	0	142	0	267	31	0
O4	0	0	0	0	0	144
O5	64	0	0	55	0	0
O6	323	22	0	0	15	0
6:30PM	D1	D2	D3	D4	D5	D6
O1	0	0	162	0	0	0
O2	30	0	0	43	0	0
O3	0	120	0	265	27	0
O4	0	0	0	0	0	133
O5	49	0	0	41	0	0
O6	334	16	0	0	15	0
6:45PM	D1	D2	D3	D4	D5	D6
O1	0	0	148	0	0	0
O2	42	0	0	33	0	0

O3	0	113	0	230	7	0
O4	0	0	0	0	0	126
O5	60	0	0	30	0	0
O6	276	4	0	0	4	0

The accumulated OD trips are shown below.

	D1	D2	D3	D4	D5	D6
O1	0	0	8124	0	0	0
O2	2963	0	0	545	0	0
O3	0	4718	0	11412	2419	0
O4	0	0	0	0	0	8121
O5	1520	0	0	3906	0	0
O6	14484	1211	0	1172	0	0

Probe Vehicle Data Processing

Introduction

For each probe vehicle runs, drivers are required to manually record the Oi and Dj and time stamps. The Apps installed on small phone records the GPS coordinates.

First, for each GPS record, the time step, as well as the latitude and longitude of the vehicle at that time step, was extracted. Next, the manual trip entries from each driver's log were compared with the GPS data. For each trip made, the start area, start time, end area, and end time were recorded. When comparing the trip information with the GPS data, for each trip, a "range" was built where the lower limit equals to the manual record's start time minus 1 minute, and the upper limit equals to the manual record's end time plus 1 minute. The GPS data whose time steps were in this range were traversed, and the distance between the vehicle location and the start area was calculated for each time step using the following formula:

DIST = acos(cos(90-Lat1) *cos(90-Lat1) +sin(90-Lat2) *sin(90-Lat1) *cos(Lon2-Lon1)) *6371*3280.84

Where:

DIST is the distance (ft) from the vehicle location to start area.

Lat1 is latitude of the vehicle location

Lat2 is latitude of the start area

Lon1 is longitude of the vehicle location

Lon2 is longitude of the start area

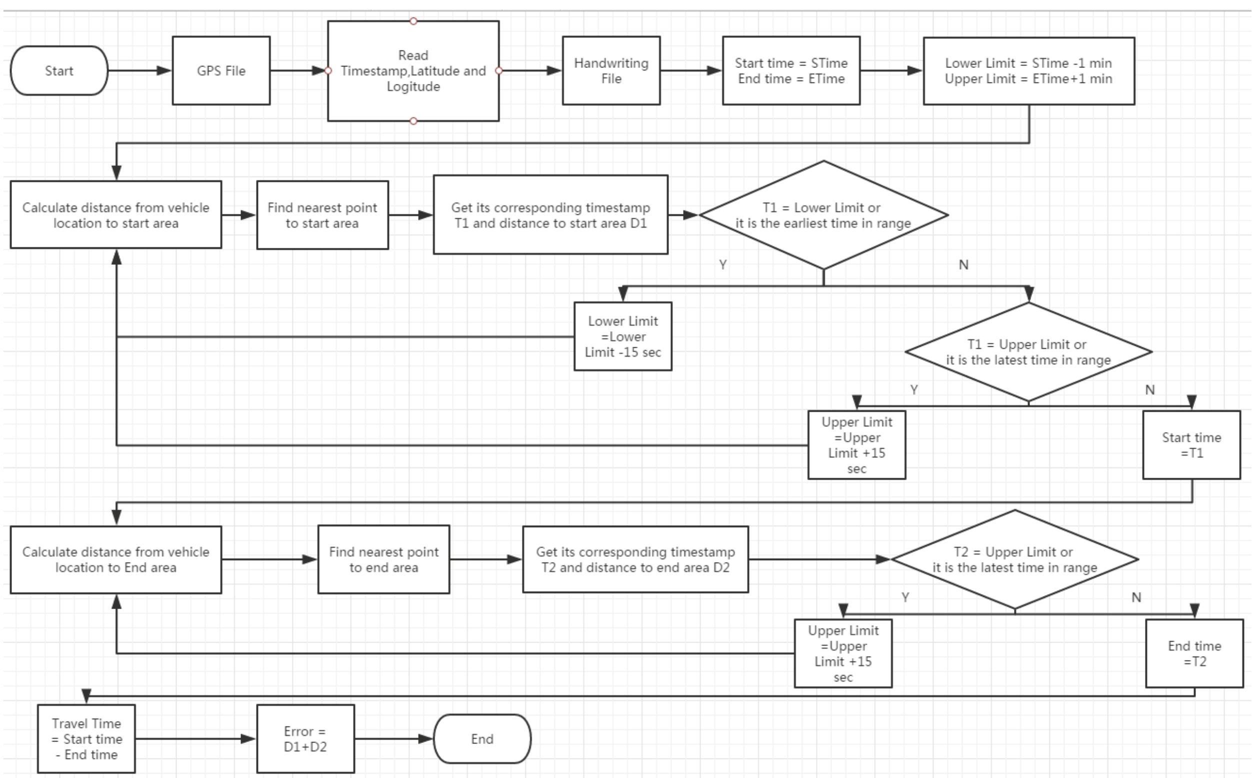
For these vehicle locations, the entry with the minimum distance from start area was considered the start point. The GPS time step of this entry was taken as the approximate time when the vehicle passed the start area. However, if this time step was equal to the lower limit of the range, or the closest value to the lower limit, an interval of 15 seconds would be further subtracted from the lower limit. Therefore, the lower limit would come to the start time minus 1 minute and 15 seconds. Another traversal is performed and if the one of the two situations above happens again, the lower limit would be reduced further by 15 seconds. This process continues until any of the two situations do not occur. Similarly, if the time step was equal to the upper limit of the range, or the closest value to the upper limit, the upper limit time would be put increased by an interval of 15 seconds and the traversal is executed another time until neither of these two cases occur.

After the starting time step is determined, another check from its next time step to the upper time limit, is performed to find the end time. Similarly, the distances from each of the vehicle locations to the end area are calculated and the closest vehicle location to the end area is chosen. The corresponding time step is considered to be the end time. The same process as above is taken when the time step is equal to the upper time limit or the last valid entry in the range.

After both the start time and end time are decided, the travel time is calculated. Further, the sum of the distance from the start area to its "closest point" and the distance between the end area and its "closest point", is recorded to for error checking purposes.

Flow Chart

The flowchart of the methodology used for data processing is shown below:



Results

The summary of the probe vehicle OD trip table is shown below.

	A	B	C	D
A		269	46	52
B	252		23	60
C	32	47		83
D	72	23	97	

The summary of the probe vehicle travel time is shown in the table below.

	A	B	C	D
A	\	00:34	01:09	00:27
B	00:32	\	00:27	01:00
C	00:42	01:47	\	00:25
D	01:56	00:22	00:23	\

A is the west side of Goodman Road. B is the east side of Goodman Road. C is the north side of US-78. D is the south side of US-78. To compare the probe vehicle and model OD table a percentage table of both is created and shown below.

Probe Vehicles	**A**	**B**	**C**	**D**
A		73%	13%	14%
B	75%		7%	18%
C	20%	29%		51%
D	38%	12%	51%	
OD Model				
	A	B	C	D
A		87%	5%	8%
B	63%		11%	26%
C	15%	16%		70%
D	23%	17%	60%	

Conclusions:

This project contains an OD estimation model based on real time detector data. The model is static with a congestion correction that uses vehicle count data of a diamond interchnage to create a real-time OD trip table. In addition, the model is adjusted to an area-wide OD estimation model that uses an acceptable number of traffic volume detectors and Bluetooth detectors to generate an area wide OD table. After a simulation evaluation, the OD estimation model proves to be feasible and shows very small errors between the model's predicted OD table and the real traffic OD given by the simulation.

Another important part of this project is field data collection. This collection process set video detectors and used probe vehicles instead of Bluetooth detectors. Although, the probe vehicle trips are lower than expected numbers, the results of probe vehicles and the OD model have similar percentage OD trip tables. With the usage of Bluetooth detectors, the results are expected to improve.

Future research using the results of this project has many potential benefits. Obtaining a real-time estimate of vehicle OD information could be used to establish delay/fuel consumption models. Signal timings for signalized interchanges could be optimized simultaneously, and this strategy appears to be able to adapt well to traffic variations, increased throughput, and other uncontrollable events. It could be used to assist in minimizing delay, number of stops, emission, and fuel consumptions at signalized interchanges.

References:

[1] Van Der Zijpp, N. (1997). Dynamic origin-destination matrix estimation from traffic counts and automated vehicle identification data. *Transportation Research Record: Journal of the Transportation Research Board*, (1607), 87-94.

[2] Tsekeris, T., & Stathopoulos, A. (2003). Real-time dynamic origin-destination matrix adjustment with simulated and actual link flows in urban networks.*Transportation Research Record: Journal of the Transportation Research Board*, (1857), 117-127.

[3] Dan, M., Zhicai, J., & Hongfei, J. (2011, May). Adaptive-filtering based dynamic OD matrix estimation. In *2011 Chinese Control and Decision Conference (CCDC)*.

[4] White, J., & Wells, I. (2002). Extracting origin destination information from mobile phone data. In *Road Transport Information and Control, 2002. Eleventh International Conference on (Conf. Publ. No. 486)* (pp. 30-34). IET.

[5] Lou, Y., & Yin, Y. (2006, September). Real-time estimation of origin-destination flows for actuation-controlled intersections. In *Intelligent Transportation Systems Conference, 2006. ITSC'06. IEEE* (pp. 827-832). IEEE.

[6] Hazelton, M. L. (2008). Statistical inference for time varying origin–destination matrices. *Transportation Research Part B: Methodological*, *42*(6), 542-552.

[7] Calabrese, F., Di Lorenzo, G., Liu, L., & Ratti, C. (2011). Estimating origin-destination flows using mobile phone location data. *IEEE Pervasive Computing*, *10*(4), 0036-44.

[8] Caceres, N., Wideberg, J. P., & Benitez, F. G. (2007). Deriving origin destination data from a mobile phone network. *Intelligent Transport Systems*, *IET*, *1*(1), 15-26.

[9] Tuydes-Yaman, H., Altintasi, O., & Sendil, N. (2015). Better estimation of origin–destination matrix using automated intersection movement count data.*Canadian Journal of Civil Engineering*, *42*(7), 490-502.

[10] Perrakis, K., Karlis, D., Cools, M., Janssens, D., Vanhoof, K., & Wets, G. (2012). A Bayesian approach for modeling origin–destination matrices.*Transportation Research Part A: Policy and Practice*, *46*(1), 200-212.

[11] CAO, P., MIWA, T., YAMAMOTO, T., & MORIKAWA, T. (2013). Estimation of Dynamic Link Flows and Origin-Destination Matrices from Lower Polling Frequency Probe Vehicle Data. *Journal of the Eastern Asia Society for Transportation Studies*, *10*(0), 762-775.

[12] Parry, K., & Hazelton, M. L. (2012). Estimation of origin–destination matrices from link counts and sporadic routing data. *Transportation Research Part B: Methodological*, *46*(1), 175-188.

[13] Hazelton, M. L. (2000). Estimation of origin–destination matrices from link flows on uncongested networks. *Transportation Research Part B: Methodological*, *34*(7), 549-566.

[14] Li, B., & De Moor, B. (2002). Dynamic identification of origin–destination matrices in the presence of incomplete observations. *Transportation Research Part B: Methodological*, *36*(1), 37-57.

[15] Tamin, O. Z., Hidayat, H., & Indriastuti, A. K. (2003, October). The development of maximum-entropy (ME) and bayesian-inference (BI) estimation methods for calibrating transport demand

models based on link volume information. In *Proceedings of the Eastern Asia Society for Transportation Studies* (Vol. 4, pp. 630-647).

[16]Bera, S., & Rao, K. V. (2011). Estimation of origin-destination matrix from traffic counts: the state of the art.

[17]Sherali, H. D., & Park, T. (2001). Estimation of dynamic origin–destination trip tables for a general network. *Transportation Research Part B: Methodological*, *35*(3), 217-235.

[18]Zhang, Y., Qin, X., Dong, S., & Ran, B. (2010). Daily OD matrix estimation using cellular probe data. In *89th Annual Meeting Transportation Research Board*.

[19]Dixon, M. P., & Rilett, L. R. (2002). Real-Time OD Estimation Using Automatic Vehicle Identification and Traffic Count Data. *Computer-Aided Civil and Infrastructure Engineering*, *17*(1), 7-21.

[20]Sherali, H. D., Narayanan, A., & Sivanandan, R. (2003). Estimation of origin–destination trip-tables based on a partial set of traffic link volumes.*Transportation Research Part B: Methodological*, *37*(9), 815-836.

[21]Cascetta, E., & Postorino, M. N. (2001). Fixed point approaches to the estimation of O/D matrices using traffic counts on congested networks.*Transportation science*, *35*(2), 134-147.

[22]Hellinga, B. R., & Van Aerde, M. (1998). Estimating dynamic OD demands for a freeway corridor using loop detector data. *Halifax, NS, Canada*, 185-198.

[23]Yang, H., Meng, Q., & Bell, M. G. (2001). Simultaneous estimation of the origin-destination matrices and travel-cost coefficient for congested networks in a stochastic user equilibrium. *Transportation Science*, *35*(2), 107-123.