A multi-resolution approach to investigate the impacts of pre-planned road capacity reduction based on smartphone GPS trajectory data

- A case study of MoPac Expressway in Austin, Texas

Xianbiao Hu, PhD
Metropia Inc.
1790 East River Road, Suite 140, Tucson, AZ 85718
Email: xb.hu@metropia.com

Xiaoyu Zhu*, PhD
Metropia Inc.
1790 East River Road, Suite 140, Tucson, AZ 85718
Email: xiaoyu.zhu@metropia.com

Yifei Yuan
Department of Systems and Industrial Engineering, University of Arizona
1127 E. James E. Rogers Way, Tucson, AZ USA
Email: yifei.f.yuan@gmail.com

Asif Rehan
Metropia Inc.
1790 East River Road, Suite 140, Tucson, AZ 85718
Email: asif.rehan@metropia.com

Yi-Chang Chiu, PhD
Department of Civil Engineering and Engineering Mechanics, University of Arizona,
1209 E. Second St, Tucson, AZ 85721
Email: chiu@email.arizona.edu

Mia Zmud
Metropia Inc.
1790 East River Road, Suite 140, Tucson, AZ 85718
Email: mia.zmud@metropia.com

Word Count: 5,745 words text + 7 tables/figures x 250 words (each) = 7,495 words
Submitted for Presentation and Publication at the 96th Annual Meeting of Transportation Research Board
Submission Date: November 15, 2016
ABSTRACT

Pre-planned events, such as construction or special events, lead to road capacity reductions and create bottlenecks in the traffic network. The traffic impact of such events goes beyond local areas, as informed drivers may detour to alternative corridors for faster travel speed and subsequently the traffic congestion may propagate to region-wide. Traditional traffic impact analyses are typically based on simulation models, fixed-location sensor data or survey questionnaires, which have various shortcomings respectively. In this research, we use the vehicle trajectory data collected via a smartphone GPS module, which is capable of keeping track of individual driver’s behavior change before and after road capacity reduction, combined with system-wide dynamic traffic condition and roadway geometry network to investigate the impacts of pre-planned events in a multi-resolution manner. First, the traffic impact of such events at the network level is analyzed, indicating how traffic may propagate to alternative corridors from the system perspective. Second, at the individual driver level, behavior changes and corresponding outcomes are examined by the comparison of before and after travel behavior. For this purpose, a functional data analysis (FDA) approach based clustering method is used to cluster trajectory data and identify detour patterns, and then two logistic and a LASSO regression models are used to explain drivers’ detour behavior choice for each pattern with spatial and temporal features of interest. A case study based on a lane closure event on MoPac Expressway in Austin, TX is used as an example in this research. The case study demonstrates that: 1) the freeway capacity reduction has a significant impact on other corridors nearby, 2) driver detour behavior exhibits three major patterns, and 3) each detour pattern highly depends on spatial features such as trip length, distance to freeway entrance and distance to other alternative freeways, in addition to the temporal features when the trip happens.

Keywords: Preplanned capacity reduction; traffic impact analysis; behavioral change; trajectory data; multi-resolution analysis; regression analysis;
1. INTRODUCTION

Pre-planned events, such as construction or special events, can lead to road capacity reductions and create bottlenecks in the traffic network. The traffic impact of such events goes beyond local areas, as informed drivers may detour to alternative corridors in search for faster travel speeds and consequently the traffic congestion may divert or propagate to other corridors. The impact of such types of events is usually assessed at the traffic level, meaning using traffic detectors to measure the change of traffic flows before and after the event. Few prior studies were able to investigate this topic in a holistic multi-resolution manner that combines both system-level and driver-level analyses. The data needed for the driver-level analysis is mostly non-existent.

The study of roadway capacity reduction impact draws the considerable attention of researchers and practitioners. A popular area of research interest is the accurate estimation of freeway capacity as the result of vehicle crashes and driving behavior at the incident location, work zone activities, severe weather, and so on [1-7]. In the event that capacity reduction degree is known or could be assumed, simulation models are frequently used to analyze the traffic impacts of road capacity reduction at the network level. For example, Mahmassani and Jayakrishnan adopted a dynamic simulation model that has a macroparticle traffic simulator, and a user decision component to represent traffic system disruptions in the form of lane closures [8]. Mitra proposed a framework to evaluate the impact of variable speed limits on work zone traffic operation with Vissim [9], while Li et al. used PARAMICS to evaluate the control strategies of variable speed limits [10]. More studies applying a simulation approach can be found [11-13]. These studies, although capable of revealing system-wide traffic changes, are not based on real data and certain assumptions have to be made for the simulation model to approximate real-world traffic conditions. For example, it’s generally known that the calibration of a model, particularly O-D demand calibration, plays a significant role in the scenario analysis, and depending on the system setup, travelers may respond to the capacity reduction events differently.

Another network level analysis approach that is frequently seen in the literature is based on the traditional fixed-location sensor data. For example, Zhu et al. used loop detector data to analyze the traffic and behavioral effects of the I-35W Mississippi River bridge collapse [14]. Also by using the dataset collected by loop detectors, He and Liu modeled the day-to-day traffic evolution process after unexpected network disruption [15], Xie and Levinson modeled the economic loss incurred by increased travel delay in alternative scenarios [16], and Lee et al. analyzed the impacts of urban freeway rehabilitation on network traffic [17]. In addition, Victor et al. utilized high-quality videos of the traffic flow around two accidents recorded from a helicopter to calibrate the capacity reduction. Such analysis based upon fixed-location sensor data is able to measure the impact of the capacity reduction in a quantifiable way, however, due to the nature of the fixed-location sensor, the individual traveler’s behavior change cannot be captured with this approach, nor can the reasons for the detour behavior be revealed.
Besides the system-level analysis, understanding driver behavior, including how travel behavior changes due to road capacity reduction, and what the underlying features that determine behavior change at the individual level are will help agencies better cope with road incidents when they occur. To achieve this goal, the survey-based approach is frequently used. For example, Ye et al. [18] analyzed the impact on commuter’s behavior during the ‘Fix I-5’ project in California based on two waves of survey data. Nam et al. [19] analyzed the travel pattern changes during pavement of I-405 in Seattle. Fujii et al. [20] surveyed 335 drivers in Japan, which revealed that temporary freeway closure has a significant impact on the driver’s perception of commute time thereby influencing travel behavior decisions. Although it is a step further towards understanding traveler’s behavior changes, the survey-based approach is not based on observation and it also suffers from various problems such as miss reporting and nonresponse bias.

The objective of this research is to expand the understanding of the impacts of a preplanned road capacity reduction event by studying the traffic impact at both the network and individual driver level. The wide use of GPS-enabled smartphones and other GPS devices give rise to the new opportunities that allow us to access vehicle trajectory data containing valuable information. Such information includes both travel behavior of a specific person (e.g., origin-destination, departure time, route), and corresponding travel experience (e.g., driving speed, actual travel time, delays).

In this research, we use the vehicle trajectory data collected via the smartphone app Metropia that keeps track of individual driver’s behavior change before and after road capacity reduction, combined with system-wide dynamic traffic information and roadway geometry dataset, to investigate the effects of pre-planned events in a multi-resolution manner. The app provides travel time prediction for the best route given various departure times, and the various incentives are provided according to the traffic congestion associated with the departure time. At the system level, the traffic impact is analyzed by examining how traffic may propagate to adjacent corridors. At the individual driver level, route choice changes and the corresponding outcomes are examined by the comparing user’s before and after detour behavior. For this purpose, a functional data analysis (FDA) approach [21, 22] based clustering method is used to cluster trajectory data and identify detour patterns, and then two logistic and a Least Absolute Shrinkage and Selection Operator (LASSO) [23] regression models are utilized to explain drivers’ detour behavior for each pattern with spatial and temporal features of interest. A real-world lane reduction event on MoPac Expressway in Austin, TX is used as the case study of the proposed research methods. The case study shows that a specific freeway capacity reduction has a significant impact on other freeways in Austin and drivers’ detour behavior exhibits three major patterns that are highly dependent upon spatial features such as trip length, distance to freeway entrance and to other alternative freeways.
The remainder of this paper is organized as follows: Section 2 presents the analysis methodology used in this paper, including the data description, traffic impact analysis workflow and behavior change analysis workflow. The case study is presented in sections 3 and 4, focusing on the network level and individual driver level, respectively. Section 5 reviews and concludes this research.

2. ANALYSIS METHODOLOGY

In this section, the data used in this research will be reviewed, followed by the workflow to analyze the traffic impacts at the network level. Then, the individual detour behavior analysis methodology will be presented.

2.1. Data description

The primary data are segmented into three categories: vehicle trajectory data, roadway geometry, and time-varying traffic condition. More details on the data are explained in Zhu et al. [24].

- Vehicle trajectory data. When travelers use the Metropia smartphone app\(^1\) to travel from origin to destination, the internal GPS module is activated and starts to record second-by-second latitude/longitude data location, time stamp, and instantaneous moving speed, and sends these data to the cloud server until the trip is completed. The trip start and end are obtained when user start a trip and turn on and off the navigation mode.

- Time-varying traffic condition. The cloud server stores the traffic dataset including time-dependent traffic speed volume, with a time interval of five minutes at each traversed link, and continues updating every five minutes. There are multiple sources of traffic data. First, a national data provider covers most freeway and arterial links. Second, the GPS data collected is used to estimate the traffic conditions. Third, a data fusion method is used to generate the traffic condition for each link. For more details on the traffic dataset, please refer to Hu, et al. [25].

- Roadway geographic network. The geographic network is stored on the cloud server and includes a set of links and nodes, i.e. \(G = (N, A)\) where \(N\) is the set of nodes \(\{1, 2, \ldots, n\}\) and \(A\) is the set of links. The basic attributes include the latitude and longitude of the nodes; the link types (i.e. freeway, arterial, local streets), speed limits, number of lanes of a road segment, and so on.

2.2. Calculate system-level traffic impact

To analyze the impact of the road capacity reduction at the network level, we first identify the links of interest, including the links on and around the capacity reduction site, as well as the alternative corridors in the network. For example, if the capacity reduction happens on a freeway link, most likely we need to analyze the traffic condition changes over the entire freeway network, as well as the ramps and arterial links in that area. Then, for each link, we calculate how its travel speed and volume change by comparing the traffic conditions before and after a

\(^1\) http://www.metropia.com.
capacity reduction event. Here the capacity reduction event is recognized as the moment the
capacity is changed. Then before and after the event are referring to the status without capacity
reduction and with the capacity reduction (during the construction), respectively. The system
travel speed and volume for the studied link are calculated from the vehicle trajectory data
described in Section 2.1 from all the users using the app to complete a trip during the study
period. Next, the traffic changes over the entire network can be visualized and analyzed in GIS
platform, which allows us to observe the differences not only on the corridor where capacity
reduction happens, but also the alternative corridors and other road segments of interest in the
area.

2.3. Individual behavioral change analysis

To analyze the traveler’s behavior changes due to the capacity reduction events, we take three
major steps. First, the complete user set and trip set is divided into two categories, based on
whether they change routes or not. Then, their related travel time changes, i.e. the benefits of
switching routes, are observed. Finally, statistical models are built to understand the major
factors that influence traveler’s detour decision.

2.3.1 User set and trip set selection

We first divide the users and trips into two categories, representing those who change their route
choice after the closure and those did not, respectively. The process for this is outlined below.

a. Get a set of linkid (LS, i.e. link set) for the road segment where capacity reduction
happens
b. Find a set of trips (TSB, i.e. trip set before) that went through links in LS before the
capacity reduction event
c. Find a set of trips (TSA, i.e. trip set after) that went through links in LS after the event
d. Find a set of trips (TSN, i.e. trip set non-traversal) that did not go through links in LS
after event
e. Identify a set of users (USU) who continued to travel on the original route after capacity
reduction, and a set of their corresponding before-after trip pairs (TSS) using the
following criteria.
   i. User id in TSB equals to user id in TSA, and
   ii. Day of the week in TSB and TSA should be the same, and
   iii. Departure time of the trips in TSB is fairly close to that in TSA (e.g. within an hour),
and
   iv. Origin location in TSB is fairly close to origin location in TSA (e.g. a quarter mile),
and
   v. Destination location in TSB is fairly close to destination location in TSA (e.g. a
quarter mile)
f. Using the same criteria, identify a set of users (USD) who detoured from the original
route after capacity reduction as well as a set of corresponding trip pairs (TSD) can be
found.
2.3.2 Calculate travel time change

With user sets USS, USD and trip sets TSS, and TSD, we can look into the travel time changes for those who made behavior changes versus who did not, and see if there are any marginal benefits associated with their decision making. Travel time change can be calculated via the equations below:

\[ \text{dif}_{tt_k} = tt_{a,k} - tt_{b,k} \quad k \in \{ \text{TSS} \cup \text{TSD} \} \]

Where:

- \( k \) is the before-after trip pair index, representing a pair of similar trips before and after event selected by criterion in 2.3.1 (e,f);
- \( \text{dif}_{tt_k} \) is the differences in experienced travel time for trip \( k \);
- \( tt_{b,k} \) and \( tt_{a,k} \) represent experienced travel time of trip \( k \) before and after capacity reduction, respectively.

2.3.3 Behavior changes statistical analysis

Besides the observations of travel behavior changes and its corresponding outcomes, this section discusses the methodology to discover major factors that influence traveler’s detour decision.

The detour patterns are first identified with a functional data analysis (FDA) approach based on the clustering method proposed by Yuan [26]. The basic idea of FDA is to take a series of data points that could come from discrete or continuous observations, approximate them by a function or curve, and perform analysis of information on such function or curve. A simple example regarding the trajectory data in this study, is to use X-axis to denote latitude, Y-axis to denote longitude, and then the entire trajectory which may include thousands of GPS data points can be plotted on the figure and presented by a function through proper basis expansion techniques [21, 22]. Compared with the traditional methods that often rely on only the aggregated information such as average travel time or average trip length to represent a trip, and lack detailed information for each GPS data comprising the trip, under the FDA approaches every single GPS point, including its latitude and longitude information, is utilized in the modeling process and each trajectory is treated as one functional data, which makes it possible for us to take advantage of the detailed trajectory information in the clustering process and identify main detour patterns. Then, the clustering results are used as the known dependent variable in the subsequent statistical models.

Next, statistical models are built to identify and explain the main factors influencing detour decisions. Based on the trip records’ information, we collected system-estimated travel time, actual travel time, a distance of the route and true trajectory, departure time, day of the week and other such information. Also, spatial features of the trips are defined to capture the relative location of origin and destination, including trip length, the distance between origin/destination and capacity reduction site, and the distance between original/destination and alternative corridors. These spatial features are defined as the logarithm of distance and are further explained in Table 1 in section 4.2. Based on these trip-based collected features, logistic
regression models and LASSO regression are tested aiming to explain drivers’ detour behavior choices for each detour pattern with spatial and temporal features of interest. To select proper variables and build a robust model, LASSO multinomial regression is utilized. Section 4.2 further illustrates how the regression models are utilized and present the research findings.

2.4. Case study setup

A lane closure event on MoPac, Austin, Texas, which started from Feb. 21 2016, and temporarily reduced MoPac northbound from three lanes to two lanes between Lady Bird Lake and Enfield Road (about 1 mile in distance), is used as a case study to demonstrate the multi-resolution analysis approach. The construction will last for months till Nov, 2016. MoPac Expressway, also called Texas State Highway Loop 1, is one of the key arteries across the downtown area of Austin. Figure 1 shows an overview of the network and the bottleneck location (black arrow) where the capacity reduction happens.

Figure 1 Overview of case study setup

Traffic dataset for Monday, February 22, the first weekday after lane closure is used to compare with the traffic condition on Monday, February 8. Data was not used from the immediate Monday before the lane closure event (February 15) because it was a national and state holiday and therefore had atypical traffic conditions. In other words, $v_{a,i}, q_{a,i}$ and $t_{a,k}$ refer to the speed,
volume and travel time information on Feb. 22, respectively, and $v_{b,i}$, $q_{b,i}$ and $tt_{b,k}$ refer to the speed, volume and travel time information on Feb. 8, respectively, based on which $dif_{vl}$, $dif_{ql}$ and $dif_{tt,k}$ can be computed via equations (1). Regarding the behavior change analysis, data of the completed trips from all the users are pulled from database. The period from Aug 1, 2015 till Feb 20, 2016 is defined as the before event to obtain TSB. And the period from Feb 21 2016 till Mar 13, 2016 is defined as the after event to obtain TSA. In total we found 261 users in TSB, 98 users in TSA, and 514 users in TSN. In terms of trips, 1,543 trips were identified in TSB, 306 trips in TSA, and 15,365 trips in TSN.

3. TRAFFIC IMPACT ANALYSIS

From the network-wide traffic speed data on the server, we compare the before and after traffic conditions during morning peak hours, mid-day, and evening peak hours, as shown in Figure 2 a, b, and c, respectively. The red color indicates the level of speed drop, with the darker color meaning a more severe drop in speed, or increase in the traffic flow.
A few observations and analysis from Figure 2:

1. In the morning peak hours, a dramatic speed drop is shown on MoPac NB approaching the construction location (circled in yellow), with an approximately 1-mile long queue observed. It is worth noting the speed drop mostly happened before the river, i.e. upstream of capacity reduction location, and after that the traffic becomes stable. This observation is consistent with the traffic flow theory that it is at the bottleneck upstream location that vehicles start to
merge and speed goes down, but as its throughput is constrained by the capacity reduction, traffic conditions for downstream links will not be too bad.

2. Speed drops were observed on other corridors during morning peak hour. A severe speed drop was observed on I-35 NB, a parallel freeway on the east side of Austin downtown. Slight speed drops were also observed east of the construction site on S Lamar, and on S 1st Street, both of which offer northbound entrances into the downtown across the river, and on TX-183, an eastern bypass and alternative to MoPac. This observation indicates that as the preplanned construction made MoPac more congested, some informed drivers detoured to other alternative corridors or local streets, where they encountered worse or equally bad traffic conditions.

3. In the mid-day, the nearby network does not show many locations with significant speed drop. This is probably because during the day, the baseline traffic is not bad, so the capacity reduction on MoPac is not significant enough to create noticeable changes in traffic conditions.

4. In the evening peak hours, sub-figure c shows the speed drop on MoPac and I-35 NB. In addition, Highway 290 to 360, which can be used as a detour route from central Austin, also exhibited a certain level of increase in congestion. This is consistent with the morning hour observation.

Similar patterns are also observed from the volume-based analysis, which reveals that the northbound MoPac corridor appears to have a sharp increase in demand upstream of the bottleneck location, whereas the downstream section enjoys a sudden release of traffic, exhibiting less demand than before. In addition, by observing the traffic flow increase and decrease information, it can be found that when MoPac gets congested, vehicles primarily detour to S Lamar Blvd and S 1st Street which both lead to downtown, to the 360 located on the left, and I-35 on the right, which go further north.

4. BEHAVIOR CHANGE ANALYSIS

With the data described in section 2.4, and following the criteria to match trip pairs in section 2.3.1, we were able to find 539 trips pairs in TSS (labeled as “Stay trips”), and 175 trip pairs in TSD (labeled as “Detour trips”).

4.1. Detour outcome analysis

Analysis results for 539 trip pairs that continue to use MoPac after lane closures indicates that, on average, travel time increased by 5.7 minutes per trip pair, which translates into a 27% increases overall. For 428 out of 539 trip pairs, or 79% of the trips, actual travel time after lane closure is longer than before. For the trips that avoided the MoPac construction site, travel times averaged 7.0 min longer than before, which translates into a travel time increase of 31%. These numbers reveal a somehow surprising fact that although road construction is ongoing, MoPac remained to be the best corridor to travel for the northbound traffic in Austin.
In terms of travel time prediction accuracy, by comparing the system estimated travel time with the driver’s experienced travel time, we can calculate that, on average the predicted travel time underestimates the actual travel time by 12.8% for the Stay trips in TSS, and underestimates by 8% for the Detour trips in TSD. These results indicate that MoPac traffic conditions during the construction related lane closure were less stable and more difficult to predict when compared with other corridors. The differences in prediction accuracy in a sense may be able to explain the travel behaviors of the detoured drivers. That is, although detouring to alternate routes may take more time than staying on MoPac, the travel time may be more reliable, and therefore makes the Detour trip attractive.

It is interesting to observe that drivers sometimes prefer to detour from MoPac to other alternative routes, even though they may experience a longer travel time. In section 4.2, we further investigate the main reasons affecting drivers’ detour behavior.

4.2. Detour Behavior Analysis

First, we applied the FDA approach based clustering method presented in section 2.3.3 to cluster trajectories, and three types of detour patterns are identified in Figure 3.

1. The first detour pattern is the group of drivers who have options to use a local arterial, such as Lamar Blvd, to avoid freeway traffic and the MoPac construction site. These drivers may originate from somewhere south of the river to go to downtown Austin or further north to destinations located between MoPac and I-35.

2. The second detour pattern is the group of drivers who shifted to other alternative freeways or highways. They could use TX-360, I-35, TX-183, or TX-130 to avoid the construction zone on MoPac.

3. The third detour pattern is the group of drivers who either have origin or destination addresses located downtown Austin. For these trips, most drivers choose to avoid the construction zone by leaving MoPac earlier or entering MoPac later, while some other drivers choose to use I-35 or other corridors.
Next, we build statistical models to explain what the attributes influencing the detour decision are. From the records of trip information from the smartphone app, we collected the estimated travel time, actual travel time, distance of the route and true trajectory, departure time, day of the week, and also whether the app suggests using the MoPac construction site or not. Also, from the three types of detours, spatial features of the trips can be generated to capture the location of origin and destination, as well as the reliability on the MoPac. These spatial features are defined as the logarithm of distance. By using the logarithm, the linear combinations of spatial features become the proportional relationship between distances. As distance proportion is of interest in this study, the logarithm is necessary here, which makes linear additive models such as logit regression applicable. The statistical description of the features is displayed in Table 1.

Table 1 Features and responses for modeling

<table>
<thead>
<tr>
<th>Notation</th>
<th>Definition</th>
<th>Variable name</th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Spatial Features</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$S_1$</td>
<td>Logarithm of trips’ OD distance</td>
<td>OD Distance</td>
<td>2.3853784</td>
<td>0.5265705</td>
</tr>
<tr>
<td>$S_2$</td>
<td>Logarithm of distance from MoPac to OD</td>
<td>MoPac to OD</td>
<td>1.2045100</td>
<td>0.7804345</td>
</tr>
<tr>
<td>$S_3$</td>
<td>Logarithm of distance from other alternative freeway to OD</td>
<td>Alternative to OD</td>
<td>1.3686915</td>
<td>0.4819097</td>
</tr>
<tr>
<td>$S_4$</td>
<td>Logarithm of minimum distance from construction area to Origin or Destination</td>
<td>Construction OD</td>
<td>1.0536233</td>
<td>0.7677018</td>
</tr>
<tr>
<td><strong>Temporal Features</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$T_1$</td>
<td>Trip’s duration (mins)</td>
<td>Trip Duration</td>
<td>26.1532979</td>
<td>11.5411078</td>
</tr>
<tr>
<td>$T_2$</td>
<td>Trip’s Estimated Travel Time (mins)</td>
<td>Estimated Travel Time</td>
<td>28.8566919</td>
<td>15.5391519</td>
</tr>
</tbody>
</table>

Categorical Percentage
Regression models are employed to further explore the relationship between these features and detour choice. First of all, two logit models are applied to model these features. The first model is a binomial logit model, where the response variable is $Y_1$ in Table 1. The second model is a multinomial logit model; in this model the response is $Y_2$. The results and significant level of these 2 models are demonstrated in Table 2:

Table 2 Coefficients and significant level of logit models

<table>
<thead>
<tr>
<th>Features</th>
<th>Model 1</th>
<th>Model 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Detour Coefficient</td>
<td>Detour Pattern 1 Coefficient</td>
</tr>
<tr>
<td>Intercept</td>
<td>0.8202</td>
<td>7.0439 ***</td>
</tr>
<tr>
<td>$S_1$</td>
<td>-1.6125 ***</td>
<td>-4.9128 ***</td>
</tr>
<tr>
<td>$S_2$</td>
<td>1.3437 ***</td>
<td>0.0921</td>
</tr>
<tr>
<td>$S_3$</td>
<td>-0.6421 *</td>
<td>-1.8927 **</td>
</tr>
<tr>
<td>$S_4$</td>
<td>-0.3531 *</td>
<td>-0.4997</td>
</tr>
<tr>
<td>$T_1$</td>
<td>0.0731 ***</td>
<td>0.0833 **</td>
</tr>
<tr>
<td>$T_2$</td>
<td>-0.0757 ***</td>
<td>-0.0706</td>
</tr>
<tr>
<td>$T_3$</td>
<td>0.1756</td>
<td>-0.3332</td>
</tr>
<tr>
<td>$T_4$</td>
<td>-0.6201</td>
<td>-0.9478</td>
</tr>
</tbody>
</table>

Model Evaluation

| AIC       | 600.97 | 805.4908 |
| Pseudo-R² | 0.2496 | 0.4027993 |

Significance level: * 0.1; ** 0.05; *** 0.01

These two models tell us that all four spatial features are significant for driver detour decision-making. As for the temporal features, the estimated travel time and actual travel time appear to be significant. Day of the week and time of day are less important when compared with travel time. Also, from AIC and Pseudo-$R^2$, Model 2 is a better fit than Model 1. The significance level of the variables shown in Model 2 is not sufficient for variable selection here. First, some of the features such as travel duration and estimated duration might be highly correlated, which will
result in a collinearity problem. Second, logistic regression includes all features simultaneously, while variable selection is usually a forward or backward procedure, and variable removal or addition will change the whole model every step of the way. These limitations of logit models may incorrectly omit some important features and make the model not robust enough.

To overcome the above-mentioned issues and select proper variables and build a robust model, LASSO multinomial regression is applied. Based on the previous analysis, a multinominal model is more adequate, and LASSO, is one of the most popular methods for variable selection. A 10-fold cross validation of the LASSO model is shown in Figure 4-a, where the upper bound of testing error indicates that 7 is the minimum number of variables. The significance of features can be determined by the LASSO path. Further, Figure 4-b shows an example of the LASSO path for detour pattern 3. The features are ordered by their significance as: $S_4 > T_3 > S_3 = S_1 > S_2 > T_4 > T_2 > T_1$. Based on cross validation and LASSO path, the final model is obtained as shown in Table 3.
Table 3 LASSO model results

<table>
<thead>
<tr>
<th>Features</th>
<th>LASSO Model</th>
<th>Detour Pattern 1 Coefficient</th>
<th>Detour Pattern 2 Coefficient</th>
<th>Detour Pattern 3 Coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>5.0318</td>
<td>-3.2654</td>
<td>-2.1215</td>
<td></td>
</tr>
<tr>
<td>$S_1$</td>
<td>-4.0005</td>
<td>0.0858</td>
<td>-0.0858</td>
<td></td>
</tr>
<tr>
<td>$S_2$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>$S_3$</td>
<td>-0.0396</td>
<td>0.0396</td>
<td>-0.0397</td>
<td></td>
</tr>
<tr>
<td>$S_4$</td>
<td>0</td>
<td>0</td>
<td>0.5608</td>
<td></td>
</tr>
<tr>
<td>$T_1$</td>
<td>0</td>
<td>0.0288</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>$T_2$</td>
<td>0</td>
<td>-0.0320</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>$T_3$</td>
<td>0</td>
<td>0</td>
<td>0.1319</td>
<td></td>
</tr>
<tr>
<td>$T_4$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Cross Validation Deviance</td>
<td>1.208</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

From the final LASSO model result,

1. The first detour pattern (avoiding freeways) is largely dependent on OD distance. The negative sign of $S_1$ coefficient suggests that a longer OD distance come with a lower possibility of detour for drivers who exhibit first detour pattern. On the other hand, whether a user can avoid freeway is highly correlated with trip distance.

2. The second detour pattern (using other freeways) appears more complicated with 4 features selected as the order: $T_2$, $T_1$, $S_1$, $S_3$. The coefficients $T_2$ and $T_1$ indicate that drivers who exhibit this pattern usually have longer travel times and shorter estimated travel times. In other words, those trips with significant travel time increase tend to detour to other freeways.

3. The third detour pattern (O or D in downtown and frequently leave MoPac earlier or enter MoPac later) is most related with $S_4$ (OD’s distance to construction area), i.e. if the user origin or destination location is close to the capacity reduction location, the possibility of them using a different ramp to enter or leave MoPac freeway and avoiding construction site will be high. The other significant features are $T_3$, $S_1$ and $S_2$. Another interesting finding is that drivers tend to detour with this pattern more on weekends (indicator $T_3$), which is likely due to the light traffic conditions on weekends, where using urban roads to bypass the construction area is a good choice.

5. CONCLUSIONS

To investigate the impacts of pre-planned road capacity reduction events, this paper documents the use of a multi-resolution analysis approach utilizing trajectory data combined with traffic and roadway geometry networks. The analysis methodology, including the data used in this research and the workflow to analyze the traffic impacts at a network level, behavior change at individual...
levels, and detour decision-making processes is discussed. Based on a real lane closure event on MoPac in Austin TX, the network-level analysis shows the traffic conditions over the whole traffic network are impacted, including the road segments nearby and the alternative corridors. The individual-level analysis indicates that drivers sometimes prefer to detour from MoPac to other alternative routes, even though they may experience a longer travel time. We then further categorized three different detour patterns with a FDA based trajectory-clustering model, and for each user group, LASSO is further applied to reveal their detour decision making process, and the statistical analysis result are intuitive and explainable. The case study demonstrated the feasibility of applying the proposed multi-resolution analysis approach in analyzing and presenting a holistic view of the impacts of preplanned roadway capacity reduction, which the traditional methods based on simulation model, fixed-location sensors, or survey data fell short. We hope this study can help researchers and practitioners gain a solid understanding of the traffic impact of such events at the system level, as well as the behavior changes at the individual driver level, which is critical for agencies to foresee outcomes and take corresponding strategies so as to better manage traffic congestions.

6. REFERENCES


[26]. Yuan, Y., *A Functional Approach of Driver’s Behavior Analysis Based on GPS Trajectories Data* 2016. (manuscript in progress)