Relationships between Crash Involvement and Temporal-Spatial Driving Behavior Activity Patterns Using GPS Instrumented Vehicle Data

Jungwook Jun
Graduate Research Assistant
School of Civil and Environmental Engineering, The Georgia Institute of Technology
790 Atlantic Drive Atlanta, GA 30332-0355
TEL: (404) 385-2376
FAX: (404) 894-2278
jungwook.jun@ce.gatech.edu

Jennifer Ogle
Assistant Professor
Department of Civil Engineering, Clemson University
116 Lowry Hall, Clemson, SC 29634-0911
TEL: (864) 656-0883
FAX: (864) 656-2670
ogle@clemson.edu

Randall Guensler
Professor
School of Civil and Environmental Engineering, The Georgia Institute of Technology
790 Atlantic Drive Atlanta, GA 30332-0355
TEL: (404) 894-0405
FAX: (404) 894-2278
randall.guensler@ce.gatech.edu

Submitted to ANB20 Safety Data, Analysis, and Evaluation (Subcommittee: Surrogate Measures of Safety)
Date of submittal: November 15, 2006
Word count: 7,900 words (5,150 words in the manuscript + 2,750 for Figures and Tables)

ABSTRACT
Although knowledge about characteristics of crash-involved drivers and their driving behavior can enhance countermeasure programs designed to improve safety on our roadways, relatively few studies have been able to analyze detailed individual behavior activity data (travel mileage, speed patterns, and acceleration activities) due to the lack of reliable data and data collection technology. However, recent advancement in mobile technology and accuracy of global positioning systems allow researchers to monitor driving activity of large fleets of vehicles over long-time study periods.

This study compared driving behavior activity patterns of crash-involved drivers with those of drivers who were not involved in crashes for a 14-month study period based on GPS-observed travel mileage, speed, and acceleration patterns to assess which of driving behavior activities are significantly different. As a result, this study found that crash-involved drivers had usually traveled longer mileage, normally traveled at higher speeds than non-crash drivers, and frequently engaged in hard deceleration events.

In addition, although this study evaluated driving behavior activity patterns between crash-drivers and non-crash drivers to understand their different patterns, those behavior activity measures may be employed as roadway safety surrogate measures for selecting locations (roadways or intersections) where hard deceleration events or high speeding patterns frequently occur.

Finally, this study suggests that transportation safety engineers and policy makers continue to aim anti-speeding campaigns to drivers and that driving behavior activity metrics of individual drivers be incorporated into education campaigns and driver evaluation or monitoring programs.

INTRODUCTION

Motor vehicle traffic crash is one of the leading causes of death in the United States and in the world. Based on a recent NHTSA report [1], motor vehicle crashes were ranked 3rd, behind only cancer and heart diseases and also showed that about one out of every 50 deaths was caused by motor vehicle traffic crashes in 2003. The State of Georgia had found [2] in 2003 that on average, 30 people lost their life in crashes each week and about 2,555 people were injured by motor traffic crashes. Totally, 1,610 people died (4 per day) and 132,879 people were injured (364 per day) on the roadways [2].

Safety in the field of transportation is one of the biggest issues that raises attention and awareness [3]. To decrease a number of motor vehicle traffic crashes and fatality rate, various efforts have been made with regard to three major components (i.e., the roadway, the vehicle, and the driver) [4], which are key parts of the general transportation environment and transportation safety improvement programs. Thus, engineers have focused on improving roadway designs, adopting enforcement programs, and developing safety devices. Vehicle manufacturers have also tried to develop various safety techniques and devices such as anti-lock brake system (ABS) or intelligent cruise control (ICC) system [4].

Since identifying locations or roadways where crashes may frequently occur and understanding the causes of crashes are very important processes, various crash prediction models and surrogate safety measures have been examined over several decades. Traffic volume, driving speed, and speed variation are mainly used for predicting crash rate and severity of crashes, especially on freeways as explanatory variables in crash prediction models [4-6]. Conflict events, deceleration rates, braking power distributions, speed variations, number of vehicles caught in dilemma zone, and numbers of signal violations are employed for arterials and local roadways [3]. Although improvements of roadway designs and development of vehicle safety devices based on those existing surrogate safety measures can increase transportation safety, parallel changes in driver behavior should be met since increases in the frequency of unsafe driving behavior may lead to increase the number of crashes [4]. However, much of the safety research on developing surrogate safety measures has focused on the vehicle or the roadways, and relatively few studies have been performed on the driving behavior activity patterns due to the lack of reliable behavioral data and the absence of effective data collection techniques. In addition, since previous studies have used only simple exposure measures such as travel mileage or speeds observed at only fixed points, they did little explain about detailed relationships between crash risk and behavior activity exposure patterns representing when, where, how, and under what conditions drivers traveled [7-9]. When drivers travel at different time (day or night), on different roadways (freeway, arterial, or local roadway), at different regions (urban or rural), or in different traffic conditions (congested or uncongested), individual drivers may expose to different crash risk [6, 10].

To collect travel mileage data, previous studies generally utilized Nationwide Personal Transportation Survey (NPTS) [11], short-time travel diary survey [12], and telephone interview (or questionnaire) [10, 13]. However, those self-reported mileage data have been shown to significantly underestimate actual vehicle miles traveled. Ogle et al. [4, 14] recently conducted an analysis of the accuracy of household trips reported in a standard two-day travel diary survey by comparing simultaneous GPS-measured trip data in the Commute Atlanta program. A total of 2,292 trips were found from the GPS-measured trip files, but only 1,622 trips were reported by two days travel diary survey, which was the under-reporting rate of 29.2% based on total GPS trips. In addition, after comparing trip duration and mileage estimates from both methods, the researchers concluded that the trips reported by survey diary produced only 90% of total travel duration and 78% of total mileage. When comparing only reported survey trips with corresponding GPS-measure trips, respondents overestimated their travel duration and mileage by 15% and 2%, respectively. The total travel mileage obtained from the surveys is off by 20% from actual, which is a problem if used in defining a driver's risk level.

Although travel mileage and duration could be obtained from the previous methods (with low accuracy), a driver's mileage and speed pattern on the certain types of roadways (specific travel routes such as freeways, arterials, or local roads) could not be obtained. Traffic speed depends not only on the driver, but also on the road condition, traffic density, time of day, etc. During peak-time conditions, a driver may be in congested traffic and therefore forced to travel at slower speeds, while a driver can choose his or her own speed in free flow conditions. Therefore, the type of roadway and time of day must be investigated in regards to the speeds in which drivers travel.

Unlike previous travel survey methods, the GPS-measured travel data provide abundant reliable information which help to identify the relationships between driving behavior and crash risk under varying conditions of facility type and time of day. Coupling the detailed travel information with known driver, household, and vehicle characteristics, operations can then be tied back to a wide variety of socio-demographic parameters. Furthermore, GPS-measured data can be used to identify how driving behavior changes during a trip in response to changes in roadway operating conditions. In this respect, to identify and substantiate driving behavior activity patterns that linked with crash-involved drivers is the challenge.

Recently, Klauer et al. [15] conducted an analysis of impact of inattention on near-crash and crash risk from naturalistic driving data from 100 GPS-instrumented vehicles. However, they focused on driver distractions such as fatigue, drowsiness, eye glance, and secondary task (eating or cell phone use), so they did not investigate driving behavior activity patterns such as travel mileage, speeding, and acceleration patterns. Wahlberg [16] examined the relationship between acceleration behavior and accident frequency for local buses. Accident data were obtained from the local bus company in Uppsala, Sweden and acceleration data were gathered by a researcher traveling on a bus with the accelerometer equipment that measures speed changes with 10 Hz. Based on the measure of celebration, which is the absolute mean of speed changes, this study did not find any strong relationships between them but suggested that celebration behavior has a higher predictive power than speed.

Thus, this current study will investigate the driving behavior activity patterns of drivers who have and who have not experienced a crash during a 14-months study period. This study will allow for an empirical investigation to determine if drivers with a crash experience have driven differently in terms of speed, time of day, and roadway types. In addition, this study provides useful techniques of implementing GPS data streams in safety research, especially in large-scale data collection processes.

DATA COLLECTION PROCESS

Behavioral Data Collection

Researchers from the DRIVE Atlanta Laboratory at the Georgia Institute of Technology developed a wireless data collection system known as the Georgia Tech Trip Data Collector (GT-TDC), which collects second-by-second vehicle activity data, including vehicle position and vehicle speed via the GPS technology. The GT-TDC collects ten engine-operating parameters from the onboard diagnostics (OBD II) system in post-1996 model year vehicles and also monitors redundant vehicle speed at 4Hz from the vehicle speed sensor (VSS). All activity data simultaneously collected are integrated into trip files, encrypted, and transmitted to the central server system at Georgia Tech using a wireless data transmit system via a cellular connection [4, 12]. The GT-TDCs have installed in over 460 light-duty vehicles in the metro Atlanta region covering 13 counties, through the commuter choice and value pricing insurance incentive program (Commute Atlanta program) since September 2003. Driver socioeconomic and demographic information were also collected at the initial phase of the program and tried to update if any changes are necessary.

Self-Reported Safety Data

As one of parts of the Commute Atlanta Program, a safety-related survey was conducted in November 2004 to obtain the information on crash involvements that had experienced during the first phase of the study (about 14 months between September 2003 and November 2004) [4]. From this self-reported safety survey, this study could categorize drivers into two groups, drivers who involved and not involved in crashes. Among 234 drivers who returned the survey form, drivers who shared a vehicle with another household member more than 10 % of the time were excluded because their individual driving data could not be identifying from that of other members. Some participants have purchased new vehicles and several GT-TDCs had malfunctioned during the study period, so those drivers could not be used for this analysis due to the gap of reinstallation time. After the data cleaning process, this study finally selected 167 drivers who had been continuously monitored through a whole 6-months period (January through June 2004) for which survey data were available. Among the 167 drivers, 26 drivers had experienced crashes.

In fact, the self-reported number of crashes could be underestimated, given that some of drivers may not accurately report their crashes or simply forget crash involvements [6]. However, the self-reported crash survey can include any minor crashes that cannot be obtained from the official crash report database since motor crashes resulting in minor property damage and occurring at non-public roadways are usually not reported [17]. After comparing the crash rates per licensed drivers in a 13-counties study area in 2002 (11.24 %) with the crash rate based on the self-reported crash survey (13.6 %), it was possible that the crash survey did not significantly underestimate the actual results[1].

After grouping drivers into two different sets (crash-involved and non-crash-involved drivers), this study estimates disaggregate mileage exposures, speed patterns,

[1] The researchers will request crash reports to DMVS in Georgia in order to verify the accuracy of the self-reported crash survey data under contract to participants.

acceleration patterns to assess differences across the two driver-groups. Table 1 shows the number of drivers used in this study by gender, age, and crash involvement status. Figure 1 illustrates residential locations and ethnic distribution in the Commute Atlanta program.

TABLE 1 Sample Size of Drivers by Age, Gender, and Crash Involvement Status

	Drivers who were not involved in crashes		Drivers who were involved in crashes		Total
Age	Male	Female	Male	Female	
15 - 24	2	5	0	1	8
25 - 34	6	8	2	1	17
35 - 44	13	14	0	3	30
45 - 54	10	18	6	3	37
55 - 64	20	26	2	3	51
65 +	12	7	3	2	24
Subtotal	63	78	13	13	167
Total	141		26		167

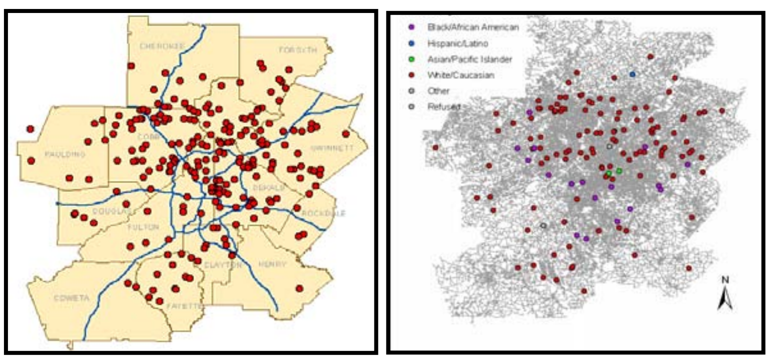

FIGURE 1 Participant residential locations and ethnic groups.

DATA PROCESSING

The Map-Matching Process

To evaluate facility-time-specific mileage, speeding, and acceleration behavior patterns using the GPS-observed data, the map-matching process requires with the correct roadway characteristics information. The research team in Commute Atlanta program developed a specialized automated map-matching algorithm to combine GPS-collected trip data with roadway characteristics (RC) information in the GIS system, using two map-matching methods, route method and buffer method [4, 18].

The University of Georgia (UGA) GIS Laboratory provided the most recent roadway maps of the 13 counties in the metro Atlanta region. Under contract to Georgia Department of Transportation (DOT), the UGA is continuously updating and managing the roadway network maps and roadway characteristics [4]. After the map-matching process, each GPS data point is associated with the corresponding roadway characteristic such as facility type, number of lanes, and lane width. Finally, those map-matched GPS data profiles are used to analyze disaggregated driving behavior activity patterns of drivers who involved in and not involved in crashes.

GPS Data Filtering Using the Kalman Filter

Similar to the map-matching process, to reduce data processing time and to estimate reliable vehicle mileage and speed estimates from the GPS-measured data, this study used an automated GPS data filtering algorithm by modifying the conventional Kalman filter [12]. Jun et al. [12] examined four statistical smoothing techniques (the least squares spline approximation method, the kernel-based smoothing technique, the conventional Kalman filter, and the modified Kalman filter) to the instrumented vehicle GPS speed data and evaluated the performance of the algorithms in minimizing the impact of GPS random error on the estimation of speed, acceleration, and distance estimates. While the conventional Kalman filter smooths all GPS data points with the constant rate (one measurement error), the modified Kalman filter selectively chose two measurement GPS errors based on the quality of GPS data points (number of satellites and Position Dilution of Precision (PDOP) value). This previous study [12] found that the modified Kalman filter provided the smallest differences from the vehicle speed sensor(VSS)-derived speed, acceleration, and travel mileage estimates and has a lesser impact on those accurate data points that reside near erroneous data points. Furthermore, the Kalman filter required less computational time than others, which indicates that this technique can be applied for the real time smoothing algorithm. This study used mileages and speeds filtered by the modified Kalman filter algorithm [12].

Test of Difference in Means Using the Wilks' Lambda Test and the Bootstrap Confidence Interval

Based on statistical literature reviews, the central limit theorem can be applied to not-normally distributed samples if the sample size is greater than 25 [19, 20]. Thus, this study used the Wilks' lambda test to verify differences in means of behavior activity metrics (mileage, speed, and acceleration) between the two driver-groups. The Wilks' lambda test is the popular asymptotic method in discriminant analysis to check the equality of means of groups and is the analog of the F-test for multivariate analysis of variance (ANOVA) [21]. Thus, the Wilks' lambda test examines if the means of behavioral metrics are equal across the two driver-groups (the null hypothesis). However, other statistical references also show that the number of samples needs to be greater than 100 to obtain a satisfactory result if the sample is not normally distributed [19, 22, 23]. The sample size of drivers who were not involved in crashes is larger than 100 (141 drivers), but that of drivers who were involved in crashes is small (26 drivers). Driving behavior activity metrics were not normally distributed (even after log transformation) after testing three nonparametric methods; Jarque-Bera test, Lilliefors test, and Kolmogorov-Smirnov test (KS-test) [24].

Due to the small sample size and non-normal distribution, this study also used an alternative method for the confidence interval estimation (for a means test). A nonparametric bootstrap resampling method was utilized to estimate the confidence intervals of sample means. The bootstrap method employs uniform random sampling with replacement method to create new data sets from the original sample data. Uniform resampling means that each data point has the same probability of being randomly selected. The bootstrap creates new sample data, and from each new bootstrap sample, the empirical distribution can be estimated. Martinez et al. [22] recommended using 1000 iterations of the bootstap reampling process to achieve stability.

Unlike to the bootstrap technique using resampling method, potential driver bias may be more pronounced in the Wilk's lambda method because it relies on the only original sample data collected without examining individual driver effects by pulling and entering their behavioral metrics. Due to the small sample size of drivers who were involved in crashes as well as the characteristics of parametric and non-parametric methods (the Wilks' lambda test and the bootstrap technique, respectively), this study selects behavior activity metrics showing significant differences between the two driver-groups based on results from either the bootstrap technique or the Wilks' Lambda test. This approach reduces the possibility of losing potential metrcis that may happen when relying on only one of methods. From the test result, any significant values less than the certain level such as 0.05 indicate that the means of behavior activity metrics estimated from the two driver-groups are statistically different.

RESULTS

For all of the following results, this study employed three roadway types, freeways, arterials, and locals, and six different time frames, AM peak (6 am ~ 9 am), morning time (9 am ~ 12 am), afternoon time (12 pm ~ 5 pm), PM peak (5 pm ~ 8 pm), night-time (8 pm ~ 12 pm), and early morning time (12 am ~ 6 am)[2].

Differences in Total Travel Mileages between Groups

The differences of total travel mileage between the two driver-groups (with and without crash involvements) were significantly different (α = 0.05) (Figure 2), indicating that the total travel mileage may be used for clustering drivers who potentially have a high crash involvement rate from general driver population. The mean of travel mileage estimated from drivers who were involved in crashes was 7,718 miles and that of drivers who were not involved in crashes was 5,244 miles, showing a difference of 32%. This result still supports the conventional definition of the exposure, where higher exposure on roadways is linked to higher possibility of crash involvements.

[2] Readers are cautioned to keep these time-periods in mind when interpreting the results of the study (i.e., results related to night-time should not be misconstrued as the total dark/night period).

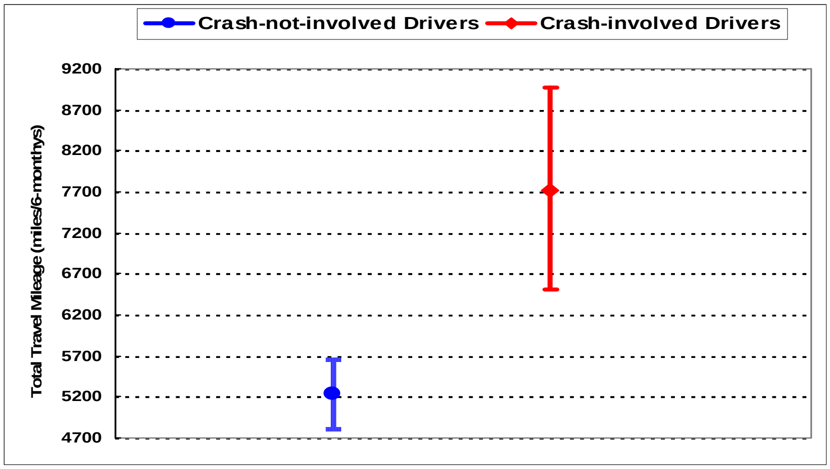

FIGURE 2 Differences of total travel mileages of two groups.

After assessing a positive relationship between total travel mileage and crash involvement, this study examined if the distributions of mileages by time-period were different to verify any preferred trip time-periods existed between two driver groups who had different crash involvements. This analysis also needed before estimating confidence intervals of means of time-specific travel mileages because although any differences in time-specific mileages existed between groups, those differences caused by the solely differences in total travel mileages if the distributions of time-specific mileages were not significantly different.

After utilizing the Kolmogorov-Smirnov (KS) test [22, 24] to verify the differences in distributions of time-specific mileage, this study found that the distributions regarding time-specific mileages between two driver groups were not significantly different (p-value: 0.07, $\alpha = 0.05$). Figure 3 shows the distributions of time-specific travel mileages between drivers who were involved and not involved in crashes, which may indicate that the choices on trip time by drivers between who were involved and not involved in crashes were not significantly different.

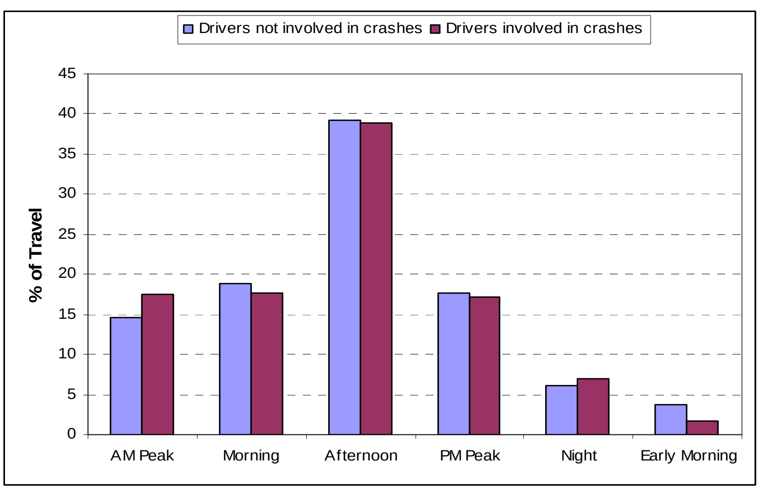

FIGURE 3 Distribution of time-specific mileage between groups.

Furthermore, this study examined the differences in mileage by facility type to evaluate if any preferences on roadway types existed between the two driver-groups. Utilizing the same method (KS test), this study found that the distributions of mileage from the two driver-groups were significantly different (p-value: 0.0326, α = 0.05) (Figure 4).

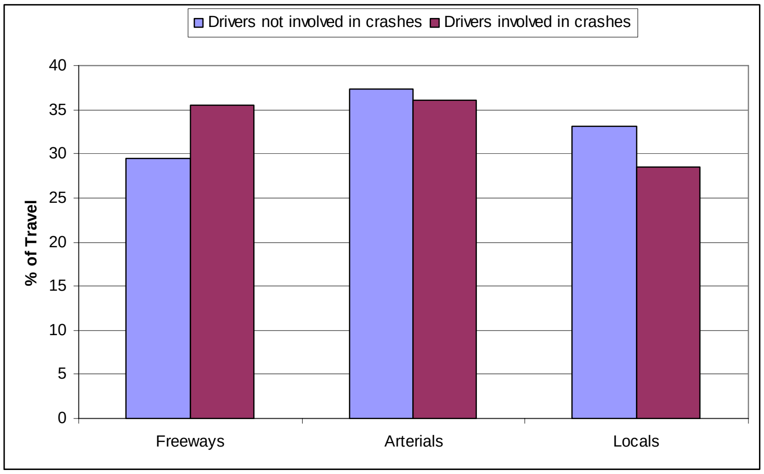

FIGURE 4 Distributions of facility mileages.

The confidence interval estimation of means showed that the mean mileage on freeways of crash-involved drivers were significantly different from that of drivers without crash involvements, resulting a difference of 36% (Table 2). Thus, this study noted that crash-involved drivers might be more willing to travel on freeways comparing to freeway mileage of non-crash-drivers.

TABLE 2 Facility and Trip Time Mileage Differences in 13 Counties

	Mean of Travel Mileages (miles/6-months)		Mileage Difference (mile)	% Difference
	Drivers who were not involved in crashes	Drivers who were involved in crashes		
Freeways *	**1203.53 (29.47%)**	**1880.96 (35.46%)**	**677.42**	**36**
Arterials	1525.29 (37.35%)	1914.94 (36.10%)	389.65	20
Local Roads	1355.21 (33.18%)	1509.21 (28.45%0	154.01	10

* indicates a significant mean difference (α = 0.05).

In addition to the significant differences in facility-specific mileages, this study found that the distributions of time-specific mileages based on each facility type were also significantly different (p-value: 0.00002, α = 0.05) (Figure 5).

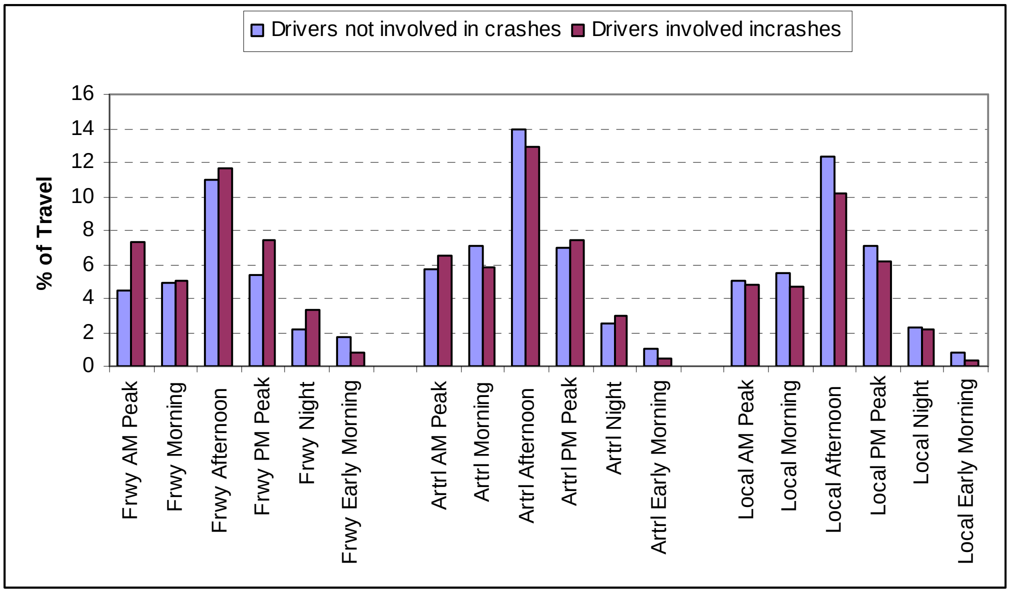

FIGURE 5 Distributions of facility-time-specific mileages.

Table 3 shows the results of the differences in travel mileage based on the pairs of facility and trip time. Based on the Wilks' lambda test and the bootstrap technique, freeways during AM peak provided the largest mileage difference (54%) between crash-involved drivers (388 miles/6-months) and drivers not involved in crashes (180 miles/6-months) and showed significant difference of the means between the two driver-groups. Travel mileages on freeways for PM peak and nighttime also showed the significant difference of 44% and 51%, respectively ($\alpha = 0.05$). This result explained that drivers who were involved in crashes traveled more on freeways, especially for peak time periods and nighttime.

TABLE 3 Facility-Time-Specific Mileage and Differences in 13 Counties

Facility Type	Trip Time	Mean of Travel Mileage (miles/6-months): Drivers who *were not involved* in crashes		Mean of Travel Mileage (miles/6-months): Drivers who *were involved* in crashes		Mileage Difference	% Difference
		Mile	%	Mile	%		
Freeways	**AM Peak ***	**180.04**	**4.41**	**388.18**	**7.32**	**208.14**	**54**
	Morning	198.66	4.86	264.05	4.98	65.39	25
	Afternoon	448.06	10.97	615.71	11.61	167.65	27
	PM Peak *	**218.29**	**5.34**	**392.26**	**7.39**	**173.97**	**44**
	Night *	**86.51**	**2.12**	**176.72**	**3.33**	**90.21**	**51**
	Early Morning[3]	71.97	1.76	44.04	0.83	-27.94	-63
Arterials	AM Peak	231.46	5.67	345.26	6.51	113.8	33
	Morning	288.02	7.05	308.57	5.82	20.55	7
	Afternoon	570.39	13.97	682.67	12.87	112.28	16

[3] Due to the very low activity during early morning, this study did not examine the difference between the two driver-groups during this period.

	PM Peak *	**286.5**	**7.02**	**395.94**	**7.46**	**109.44**	**28**
	Night	104.58	2.56	155.49	2.93	50.91	33
	Early Morning	44.34	1.09	27.01	0.51	-17.33	-64
Local Roads	AM Peak	206.73	5.06	256.52	4.84	49.79	19
	Morning	223.21	5.47	250.51	4.72	27.3	11
	Afternoon	506.34	12.40	537.87	10.14	31.53	6
	PM Peak	290.56	7.11	328.55	6.19	37.99	12
	Night	95.04	2.33	116.52	2.20	21.48	18
	Early Morning	33.33	0.82	19.24	0.36	-14.09	-73

* indicates a significant mean difference (α = 0.05).

Observed Differences in Speed Exposure by Crash Involvement Status

This study also examined differences of facility-time-specific speed behavior between the two driver-groups based on not only average driving speed but also average running speed. When studying central values of speeding behavior, it is important to consider the effects of stop delays and slow moving traffic on overall speeding pattern. Thus, average running speeds (excluding all speeds less than 5 mph) were used to examine the speeding behavior differences between the two driver-groups [4]. As shown in Table 4, two speed-related metrics (average driving speed and average running speed) resulted in the same result. Crash-involved drivers tended to travel at higher speeds on each facility type and time of day (a surrogate for traffic flow and speeds) than drivers who were not involved in crashes. Average driving speeds of crash-involved drivers during morning and nighttime on freeways and during AM peak on arterials showed that their speeding behaviors were significantly different from drivers who were not involved in crashes (α = 0.05). However, speeds on local roadways did not provide any significant differences.

TABLE 4 Facility-Time-Specific Average Running Speeds

Facility	Trip Time	Drivers who *were not involved* in Crashes		Drivers who *were involved* in Crashes		Speed Difference		% Difference	
		Average Speed	Average Running Speed	Average Speed	Average Running Speed	Average Speed	Average Running Speed	Average Speed	Average Running Speed
Freeways	AM Peak	51.64	52.88	53.08	54.07	1.45	1.19	3	2
	Morning*	**56.84**	**57.62**	**61.46**	**61.94**	**4.63**	**4.31**	**8**	**7**
	Afternoon	55.33	56.36	57.57	58.98	2.24	2.62	4	4
	PM Peak	51.01	52.35	53.95	54.98	2.94	2.63	5	5
	Night*	**57.69**	**58.17**	**62.91**	**63.08**	**5.22**	**4.91**	**8**	**8**
	Early Morning	62.55	62.73	64.73	65.33	2.18	2.6	3	4
Arterials	**AM Peak ***	**27.06**	**34.14**	**31.42**	**37.84**	**4.36**	**3.7**	**14**	**10**
	Morning	27.16	33.87	28.1	35.08	0.94	1.21	3	3
	Afternoon	25.65	32.76	26.64	33.87	0.99	1.11	4	3
	PM Peak	24.96	31.97	26.77	33.85	1.81	1.88	7	6
	Night	30.25	35.43	32.32	37.05	2.07	1.62	6	4
	Early Morning	35.08	39.38	39.28	42.78	4.21	3.4	11	8

Local Roads	AM Peak	19.94	31.14	19.17	31.84	-0.78	0.71	-4	2
	Morning	20.16	30.18	20.47	30.32	0.31	0.14	2	0
	Afternoon	20.28	29.62	19.12	29.48	-1.16	-0.14	-6	0
	PM Peak	18.71	29	18.61	29.42	-0.1	0.42	-1	1
	Night	18.76	30.3	18.99	30.71	0.22	0.41	1	1
	Early Morning	16	32.18	17.6	32.08	1.6	-0.1	9	0

* indicates a significant mean difference (α = 0.05).

Observed Differences in Acceleration Activities by Crash Involvement Status

In addition to the speed patterns analyses, this study also examined the differences in acceleration activity exposures. This study estimated accelerations of individual drivers by the central difference method [25] and compared differences between the two driver-groups through various measures such as mean, standard deviation (acceleration noise), and frequency of hard acceleration/deceleration events per mile traveled. This study employed several thresholds for defining the hard acceleration/deceleration events such as 4 mph/s, 6mph/s, and 8 mph/s, and the better result was obtained from hard deceleration activity using the threshold of 6 mph/s. Finally, this study found that hard deceleration events between the two driver-groups (with and without crash involvements) were significantly different, especially trips during morning on all facilities and nighttime on local roadways (Table 5).

TABLE 5 Frequencies of Hard Decelerations (6 mph/s) per Mile Traveled based on Facility and Time Period

Facility Type	Trip Time	Frequency of Hard Decelerations (6 mph/s) per Mile		Frequency Difference	% Difference
		Drivers who *were not involved* in Crashes	Drivers who *were involved* in Crashes		
Freeways	AM Peak	0.005	0.006	0.001	20
	Morning *	**0.009**	**0.029**	**0.021**	**71**
	Afternoon	0.008	0.018	0.010	55
	PM Peak	0.017	0.017	-0.001	-4
	Night	0.004	0.003	-0.001	-36
	Early Morning	0.070	0.010	-0.060	-618
Arterials	AM Peak	0.087	0.076	-0.011	-14
	Morning *	**0.131**	**0.221**	**0.090**	**41**
	Afternoon	0.132	0.172	0.040	23
	PM Peak	0.166	0.214	0.048	23
	Night	0.096	0.136	0.040	29
	Early Morning	0.611	0.138	-0.473	-341
Local Roads	AM Peak	0.122	0.100	-0.022	-22
	Morning *	**0.152**	**0.303**	**0.151**	**50**
	Afternoon	0.162	0.203	0.042	21
	PM Peak	0.179	0.272	0.092	34
	Night *	**0.091**	**0.161**	**0.070**	**44**
	Early Morning	1.006	0.255	-0.752	-295

* indicates a significant mean difference (α = 0.05).

Although the amount of accelerations and the frequency of hard acceleration and deceleration events on arterials or local roadways were larger than those on freeways, large behavioral differences between the two driver-groups were found from the activities on freeways instead of on arterials and local roadways. This result may imply that drivers with crash involvements may be undertaking tailgating behavior or other factors such as cellular phone use since the periods of morning and nighttime did not generally indicate the congested traffic conditions. Brookhuis et al. [26] investigated the relationship between cellular phone use and driver performance of 12 drivers in an instrumented passenger car on the road measured every work day for 3 weeks and found that statistically significant increase in brake reaction time to adapt to a slowing lead vehicle. They also found that drivers who were using a cellular phone while driving did not decrease their driving speeds. Further study may need to be evaluated the impacts of tailgating behavior and cellular phone use on hard acceleration and deceleration activities using the more sophisticated instruments.

As this study showed that the frequency of hard deceleration events is strongly related with the crash involvement rate of individual drivers, this study tried to find locations where hard deceleration behaviors frequently occurred from individual GPS-instrumented vehicle data. Figure 6 shows the potential hazardous locations based on hard deceleration events and crash locations in previous occurred between 2000 and 2002 in the State of Georgia.

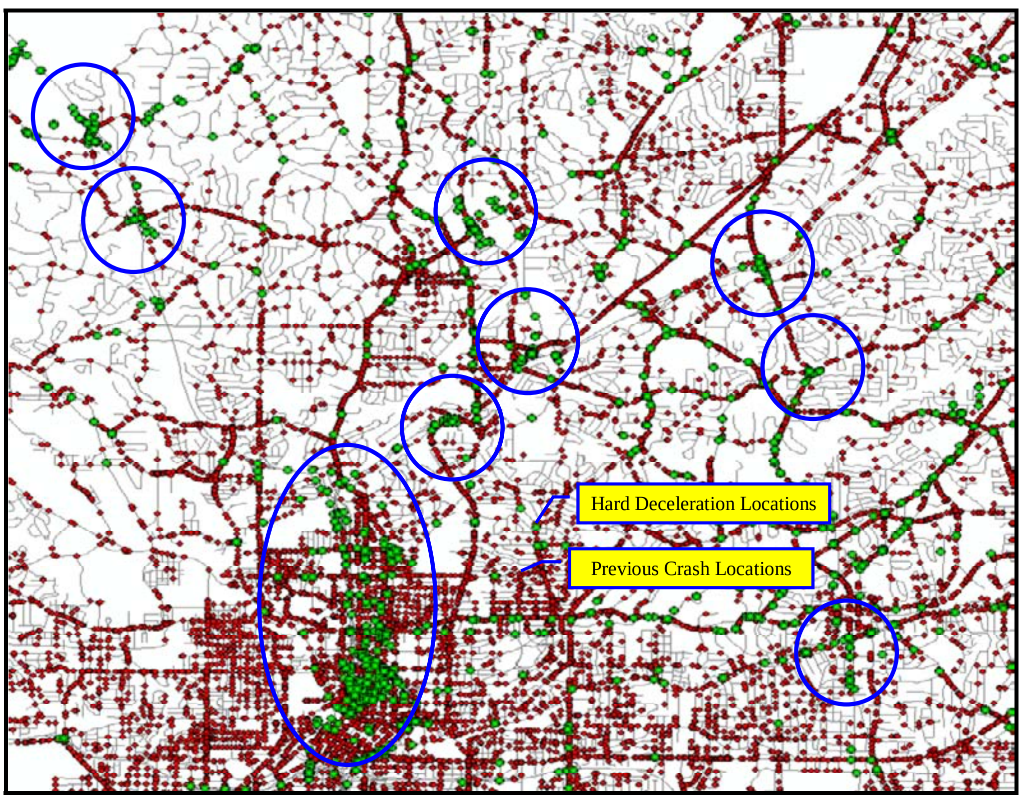

FIGURE 6 Potential hazardous locations based on hard deceleration events.

CONCLUSIONS

This study evaluated the individual driving behavior activity patterns based on crash involvement status. Each participant naturally drove his or her own vehicle in an urban setting for the 6-months period and were categorized into the "crash" group or the "non-crash" group by a self-reported crash survey. The primary variables of investigation include mileage traveled by roadways and time or day, speed patterns, and acceleration activities.

Overall, this study found that crash-involved drivers usually had traveled longer mileage, had normally traveled at higher speeds than non-crash drivers, and had frequently engaged in hard deceleration events. Driving behavior activities between the two driver-groups showed three significant differences, indicating that (1) crash-involved drivers had traveled significantly larger mileage during peak periods and nighttime on freeways than drivers who were not involved in crashes, (2) crash-involved drivers drove at faster speeds than drivers who were not involved in crashes during morning and nighttime on freeways and during AM peak on arterials, (3) crash-involved drivers frequently produced hard deceleration events than drivers who were not involved in crashes during morning on all roadway types and during nighttime on local roadways, and (4) the choices on trip time by drivers between who were involved and not involved in crashes were not significantly different.

In addition, although this study evaluated driving behavior activity patterns between crash-drivers and non-crash drivers to understand their different patterns, those behavior activity measures may be employed as roadway safety surrogate measures for selecting locations (roadways or intersections) where hard deceleration events or high speeding patterns frequently occur. Finally, this study suggests that transportation safety engineers and policy makers continue to aim anti-speeding campaigns to drivers and that driving behavior activity metrics of individual drivers be incorporated into education campaigns and driver evaluation or monitoring programs.

LIMITATIONS AND FUTURE STUDY

This study is an observational research evaluating naturalistic driving patterns by crash involvement status. In the experimental design research, researchers can recruit large number of drivers, crash-involved and non-crash-involved drivers, and evaluate their driving behavior patterns. However, those driving behavior patterns may not represent a normal driving behavior since drivers may modify their driving behavior patterns after being involved in a crash. The observational research like this study tries to randomly recruit participants and observe their normal driving behavior activity patterns and crash involvements. Thus, this observational study has relatively a small sample size regarding crash-involved drivers. This study collected crash data from the participants during the 14-months period. However, due to the rare event characteristics of crash involvements, this study employed the small sample size (only 26 drivers) for drivers who were involved in crashes. Thus, this study suggests that researchers in future who have larger sample data and longer period of data collection need to re-examine driving behavior activity patterns between crash-involved and non-crash-involved drivers.

The self-reported crash data used in this study did not contain other important information such as where, when, what conditions crashes occurred, crash type, and severity of a crash. Due to this limitation, this study evaluated driving behavior activity patterns based on only crash involvement status. If researcher can evaluate relationships between driving behavior activity patterns and fatality or severity of crashes, researchers may find other important behavior metrics. Such information may provide the relationships between behavioral metrics and crash situations in detail.

Although this study evaluated disaggregated behavioral exposures based on time of day and facility type, further investigations regarding exposures to roadways having different geometric designs (grade and curvature) and operational designs (speed limit and traffic volume) need to be performed. In addition, speed difference between individual driving speed and surrounding traffic speed may be one of potential behavioral crash-related exposure measures.

This study also suggests that the relationships between crash involvement rate and activities based on trip purpose (commuting or shopping) and vehicle types may be potential behavior exposure measure. Finally, this study recommends that difference in driving behavior activity patterns before and after crash involvements also need to be evaluated in future research

REFERENCES

1. Subramanian, R., "Motor Vehicle Traffic Crashes as a Leading Cause of Death in the U.S. 2002", NHTSA Technical Report, 2005.
2. Georgia Department of Motor Vehicle Safety, "Crash Analysis, Statistics, and Information Notebook 1996-2003", CASI Report, Georgia, 2004.
3. Gettman, D. and L. Head, "Surrogate Safety Measures From Traffic Simulation Models", Final Report. Federal Highway Administration, 2003.
4. Ogle, J., "Quantitative Assessment of Driver Speeding Behavior Using Instrumented Vehicles", Doctoral Dissertation, Civil Engineering, The Georgia Institute of Technology, Atlanta, 2005.
5. GAO, "Highway Safety - Research Continues on a Variety of Factors That Contribute to Motor Vehicle Crashes", United States General Accounting Office, GAO-03-436, 2003.
6. Crettenden, A. and A. Drummond, "The Young Driver Problem versus the Young Problem Driver - A Review and Crash Data Analysis.", Federal Office of Road Safety, 1994.
7. Janke, M., "Accidents, Mileage, and the Exaggeration of Risk. Accident Analysis and Prevention", Vol. 23, No. 3, pp. 183-188, 1991.
8. Litman, T. and S. Fitzroy, "Safe Travels - Evaluating Mobility Management Traffic Safety Impacts", Victoria Transport Policy Institute, Canada, 2005.
9. Greenshields, B.D. and F.N. Platt, "Development of a Method of Predicting High-Accident and High-Violation Drivers", Journal of Applied Psychology, Vol. 51, No. 3, pp. 205-210, 1967.
10. Jonah, B.A., "Age Differences in Risky Driving. Health Education Research - Theory and Practice", Vol. 5, No. 2, pp. 139-149, 1990.
11. Massie, D., K. Campbell, and A. Williams, "Traffic Accident Involvement Rates by Driver Age and Gender", Accident Analysis and Prevention, Vol. 27, No. 1, pp. 73-87, 1995.

12. Jun, J., R. Guensler, and J. Ogle, "Smoothing Methods Designed to Minimize the Impacts of GPS Random Errors on Travel Distance, Speed, and Acceleration Profile Estimates", Transportation Research Record, 2006. (In press).
13. Maxwell, J.P., S. Grant, and S. Lipkin, "Further Validation of the Propensity for Angry Driving Scale in British Drivers", Personality and Individual Differences, Vol. 38, pp. 213-224, 2005.
14. Ogle, J., J. Ko, and R. Guensler, "Travel Diary Underreporting in the Commute Atlanta Instrumented Vehicle Study", Journal of the American Planning Association, 2005. (Manuscript).
15. Klauer, S.G., et al., "The impact of driver inattention on near-crash/crash risk: An analysis using the 100-car naturalistic driving study data", NHTSA, April 2006.
16. Wahlberg, A.A., "The Stability of driver acceleration behavior, and a replication of its relation to bus accident", Accident Analysis and Prevention, Vol. 36, pp. 83-92, 2000.
17. Hauer, E. and A.S. Hakkert, "Extent and Some Implications of Incomplete Accident Reporting", Transportation Research Record, Vol. 1185, pp. 1-10, 1988.
18. Li, H., "Investigating Morning Commute Route Choice Behavior Using Global Positioning Systems and Multi-day Travel Data", Doctoral Dissertation, Civil Engineering, The Georgia Institute of Technology, Atlanta, 2004.
19. Washington, S., M. Karlaftis, and F. Mannering, "Statistical and Econometric Methods For Transportation Data Analysis", Chapman & Hall/CRC., 2003.
20. Miller, I., J.E. Freund, and R.A. Johnson, "Probability and Statistics for Engineers", Prentice Hall, 1990.
21. Hair, J., R. Anderson, R. Tatham, W. Black, "Multivariate Data Analysis", Prentice Hall, 1998.
22. Martinez, W.L. and A.R. Martinez, "Computational Statistics Handbook with Matlab", Chapman and Hall/CRC., 2002.
23. Hines, W.W., D.C. Montgomery, D.M. Goldsman, C.M. Borror, "Probability and Statistics in Engineering", John Wiley & Sons, Inc., 2002.
24. Matlab, Matlab Online Manual. 2005.
25. Jun, J., J. Ko, S. Yoon, and R. Guensler, "Impacts of Acceleration Calculation Methods on Onroad Vehicle Engine Power Estimation", CD-ROM, Air & Waste Management Association Annual Meeting (AWMA), New Orleans, 2006.
26. Brookhuis, K.A., G. De Vries, and D. De Waard, "The effect of mobile telephoning on driving performance", Accident Analysis and Prevention, Vol. 23, No. 4, pp. 309-316, 1991.