USDOT Region V Regional University Transportation Center Final Report

NEXTRANS Project No. 126UMUY2.1

Using Naturalistic Driving Performance Data to Develop an Empirically Defined Model of Distracted Driving

By

C. Raymond Bingham, PhD

Research Professor

University of Michigan Transportation Research Institute

rbingham@umich.edu

and

Shan Bao, PhD

Assistant Research Scientist

University of Michigan Transportation Research Institute

shanbao@umich.edu

and

Carol Flannagan, PhD

Associate Research Scientist

University of Michigan Transportation Research Institute

cacf@umich.edu

and

Anuj K. Pradhan, PhD

Assistant Research Scientist

University of Michigan Transportation Research Institute

anujkp@umich.edu

DISCLAIMER

Funding for this research was provided by the NEXTRANS Center, Purdue University under Grant No. DTRT12-G-UTC05 of the U.S. Department of Transportation, Office of the Assistant Secretary for Research and Technology (OST-R), University Transportation Centers Program.

USDOT Region V Regional University Transportation Center Final Report

TECHNICAL SUMMARY

NEXTRANS Project No. 126UMUY2.1 Final Report, 08/05/2016

Title

Using Naturalistic Driving Performance Data to Develop an Empirically Defined Model of Distracted Driving

Introduction

Driver distraction is defined as a diversion of attention away from the primary driving activity toward non-driving related tasks (Lee et al., 2008). Multiple resource theory (MRT) describes this diversion as a process of competition for attentional resources (Wickens, 2002). When the non-driving related tasks compete for the same resource (e.g., visual or cognitive), performance of the primary task is very likely to degrade. In 2009, NHTSA reported highlights of analyses of crash databases for that year as related to distracted driving (Ascone, 2009). For example, in 2009, 5474 people were killed on U.S. roadways in motor vehicle crashes that were reported to have involved distracted driving. Of these, 18% (995) involved reports of cell phone as a distraction. Thus, cell phones were involved in approximately 3% of all fatalities. Of those injured in crashes in 2009, 20% involved reports of distraction. Of those, 5% involved cell phones. Thus, approximately 1% of injuries were reported as involving cell phones.

Cell phone use and other driver distractions have been the subject of many studies resulting in a range of findings (Bao, Flannagan, & Sayer, submitted; Liang & Lee, 2010; Nemme & White, 2010; Redelmeier & Tibshirani, 1997; Strayer & Drews, 2007; Strayer & Johnston, 2001). However, the most challenging element of the science of driver distraction is that while most simulator studies clearly show performance deficits with secondary tasks (Drews, Yazdani, Godfrey, Cooper, & Strayer, 2009; Liang & Lee, 2010; Owens, McLaughlin, & Sudweeks, 2011), the crash data show steady decreases in total crashes, fatalities, and crash rates (IIHS, 2010; Ascone, 2009).

One of the difficulties in understanding the effect of distraction, particularly cell phone use, on crashes has been that police reports have historically under-represented is traction or not coded various sources of distraction. As this issue has become more public, coding of distraction has increased in quantity and quality. The National Motor Vehicle Crash Causation Survey (NMVCCS) was conducted between 2005 and 2007 and involved in-depth investigation of the causation of a set of 6,949 crashes. At that time, 22% of drivers were distracted by one or more sources. Of these, 16% were conversing with a passenger and about 3.4% were either talking on or dialing a cell phone. Because in-depth investigations were done on-scene, these estimates are much less likely to be undercounting distraction. The objective of this study is to apply a stochastic modeling method, Hidden Markov Modeling method, to naturalistic driving data analysis, and to develop algorithms to identify distract driving by using vehicle kinematic variables only.

Method

Data

The data used for this study was collected from the Integrated Vehicle-Based Safety Systems Field Operational Test (IVBSS FOT) (Sayer et al., 2011). The IVBSS FOT involved 108 primary drivers (54 male, 54 female) who drove instrumented vehicles for 6 weeks each. The vehicles were 16 Honda Accord EX sedans (models years 2006 and 2007) equipped with a variety of prototype warning systems, including: forward radar measuring distance, closing speeds, and relative accelerations of leading vehicles; side and rear radars that track other vehicles and roadside objects alongside the vehicle, as well as closing speeds of trailing vehicles; a forward lane-tracking camera that gathered data on vehicle position within the lane as well as the type of lane markings (dashed, solid); five camera streams at rates of 10 Hz (forward and driver-face) and 2 Hz (other views); vehicle motion sensors, including accelerometers and a yaw rate sensor; GPS and onboard digital map information, including roadway attributes (e.g., posted speed limits, number of lanes); and, signals from the OEM automotive data bus, including speed, accelerator and brake pedal status, turn signals, and wipers. Audio data were not recorded. Participants were fully informed of the data recording and the camera locations

The data analyzed in the current study were from the first two weeks of data collection, during which all warning systems were disabled for the collection of baseline driving data. For this analysis, the distracted driving episodes of interest were cellular phone use by the driver. Cellular phone use was considered any use of a hand-held or hands-free device (dialing, browsing, texting, talking, listening, glancing/viewing, etc.). During video coding completed as part of prior analysis of these data, clips containing cellular phone use by drivers had already been identified. A sample of 5-second clips in which the drivers were using a cellular phone were randomly selected and were considered cellular phone cases or case clips.

Matched control clips, in which drivers did not engage in any secondary tasks, were identified for each case clip via a two-part matching and coding process. To optimize statistical power, the investigators attempted to identify as many matched control clips as possible for each case clip up to a maximum of five control clips per case. While it had already been determined from prior coding that there was no cellular phone use in these clips, new coding was conducting to confirm that drivers did not engage in other secondary tasks and that the clips were control cases.

Three matching strategies were ultimately attempted to obtain an evaluation sample with enough matched clips to provide adequate analysis power. In the first matching strategy, 519 case clips were identified from 65 drivers (mean = 8 cases per driver; maximum cases per driver = 10). Potential control clips were matched to case clips by driver, same trip, within 30 minutes of the case clip, same roadway, and same traffic density, resulting in 2,724 matched clips - potential control clips. Examination of the matched clips during coding revealed that, within drivers, multiple case clips had occurred within a close enough period of time that the clips matched to them were duplicates of the clips matched to other cases for that driver. Overall, 80% of the clips matched using the first strategy was duplicates. The investigators determined that the first strategy was too conservative and attempted a second round of matching with modifications.

In the second matching strategy, 323 case clips were identified from 35 drivers (mean = 9 cases per driver; maximum cases per driver = 10). The matching condition of same trip was dropped and potential control clips were matched to case clips by driver, within 30 minutes of the case clip, same roadway, and same traffic density, resulting in 2,555 matched clips. This strategy performed better but 71% of the clips matched were duplicates still leaving too few control clips to for adequate power.

In the third matching strategy, 349 cases were identified from 35 drivers (mean = 10 cases per driver; maximum cases per driver = 10). The matching condition of within 30 minutes of the case clip was broadened to same time of day (day or night). Potential control clips were matched to case clips by driver, same day/night, same roadway, and same traffic density, resulting in 10,054 matched clips. Overall, this strategy still resulted in 77% duplicate clips, but the starting volume of matched clips was so large that the coding process revealed an adequate number of control clips to complete the data analyses. The dataset generated using this strategy was used for all data analysis.

Driving performance measures include acceleration pedal use, driving distance, driving speed and lane offset in both time and frequency domain were used in this study. The measurement of lane offset was associated with a “confidence” level in the data collection process. The confidence level was calculated based on how well the forward-looking video camera could identity lane markings on the road. The Fast Fourier Transform (FFT) method was used to transform the four variables into frequency domain (See Bao et al., 2015). FFT is an algorithm to break down time-domain data, $\mathrm{h(t)}$, into constituent sinusoids of different frequencies, $\mathrm{X(f)}$.

Hidden Markov Model

The Hidden Markov model method (HMM) was used in this study. This method has been used by other studies to recognize driving patterns (Mitrovic, 2005). An HMM is a chain-style probabilistic graph model with variable length. Its structure is shown in Figure 1.

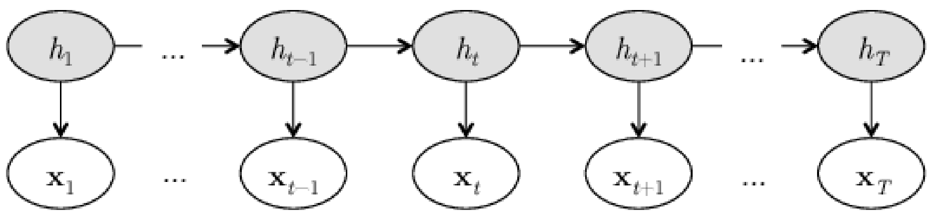

Figure 1: HMM structure

$\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_T$ are visible nodes, whose states (visible states) are the data series that we observed, and $h_1, h_2, \dots, h_T \in \mathcal{H}$ are the hidden nodes, which corresponds to the unobserved states (hidden states) that determines the underlying probabilistic distribution for generating the observations. $\mathcal{H}$ is usually a finite set with a reasonable small number of elements. The directed edge indicates two types of conditional dependency:

Transition dependency The hidden state at time $t+1$ depends on that at time t. The first hidden state has no dependency.

Emission dependency The visual state at time t depends on the hidden state at the same time.

Accordingly, three types of probability distributions need to be defined to specify an HMM:

Initial probability $p(h_1)$ can be specified by a vector $\mathbf{b} = \left(b_1, b_2, \dots, b_{|\mathcal{H}|}\right)^T$, where $|\mathcal{H}|$ denotes the number of hidden states, $b_i \geq 0$ is the probability of chosing the ith hidden state for h_1, and $\sum_{t=1}^{|\mathcal{H}|} b_t = 1$.

Transition probability $p(h_{t+1}|h_t)$ can be similarly specified by a square matrix $\mathbf{A} = \left[a_{ij}\right]_{i,j=1,2,\dots,|\mathcal{H}|}$, where $a_{ij} > 0$ is the probability of trasiting from the jth hidden state (at time t) to the ith hidden state (at time $t+1$), and $\sum_{i=1}^{|\mathcal{H}|} a_{ij} = 1$.

Emission probability $\mathbf{x}_t \in \mathbb{R}^D$ is a continous variable, so we cannot specify $p(\mathbf{x}_t|h_t)$ by a matrix, which is the common practical for discrete visible states. Instead, the Gaussion distribution is adopt, i.e. $\mathbf{x}_t|\hbar_t \sim \mathcal{N}(\mu(\hbar_t), \mathbf{\Sigma}(\hbar_t))$, where $\mu(\hbar_t), \mathbf{\Sigma}(\hbar_t)$ are respectively the mean vector and covariance matrix for the state of $\hbar_t$.

In summary, an HMM can be specified by a tuple of parameters $\Phi = (\mathbf{b}, \mathbf{A}, \{\mu(\hbar), \mathbf{\Sigma}(\hbar)\}_{\hbar \in \mathcal{H}})$.

Given the hidden states, the probability of observing a series $\mathbf{X}$ is

$$p(\mathbf{X}|\{\hbar_t\}_{t=1}^{T}, \Phi) = (p(\hbar_1)p(\mathbf{x}_1|\hbar_1)) \prod_{t=2}^{T} (p(\hbar_{t-1}|\hbar_t)p(\mathbf{x}_t|\hbar_t)). \quad (5)$$

The probability of observing an $\mathbf{X}$ with unknown hidden states can be calculated by marginalizing the hidden varibles:

$$p(\mathbf{X}|\Phi) = \sum_{\hbar_1, \hbar_2, \ldots, \hbar_T \in \mathcal{H}} p(\mathbf{X}|\{\hbar_t\}_{t=1}^{T}, \Phi). \qquad (6)$$

Given a set of training series[1] $\mathbf{X}^1, \mathbf{X}^2, \ldots, \mathbf{X}^n$, the standard way to learn Φ is the maximum likelihood (ML) criterion. More specifically, the optimal parameter tuple is

$$\widehat{\Phi} = \underset{\Phi}{\operatorname{argmax}} \sum_{k=1}^{n} \log p(\mathbf{X}^k|\Phi). \qquad (7)$$

This optimization problem can be solved by a standard expectation-maximization algorithm initialized with k-mean clustering algorithm. We used the implementation in Murphy's HMM toolbox for MATLAB [1].

Sliding window: classifier as a convolutional filter

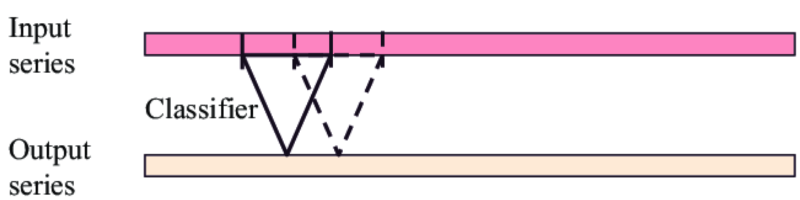

Figure 2: Sliding window scheme for generating a series of predictions.

The proposed classifier predicts a single classification label for an input series. However, in practice, the data under driver distraction and normal driving are usually mixed in a long series, where assigning a single label to the whole series makes no sense. To deal with this situation, we extend our method with a sliding window scheme to output a series of predictions, and enables real-time distraction alerts using the proposed classifier.

Specifically, we can get a series of classification score $(s_1, s_2, \dots, s_T)$ of the same length as a long input series $[\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_T]$ by

$$s_t = s([\mathbf{x}_{t-\delta}, \dots, \mathbf{x}_t, \dots, \mathbf{x}_{t+\delta}]), \qquad (8)$$

where t is the center of the window, and $\delta \geq 0$ determines the size of the window. Considering the two ends of the series, we set $\mathbf{x}_t$ to an empty vector ($\mathbf{x}_t \in \mathbb{R}^{D\times 0}$) if $t \notin \{1,2,\dots,T\}$. Figure 2 illustrates the sliding window scheme. It basically takes the HMM-based classifier as a convolutional filter, and feeds it only with small segments of the whole input series.

Modeling analysis and results

Parameters: 5 hidden states are used. The size of the sliding window is 4.

Evaluation protocol

The proposed method is evaluated on our dataset with three protocols.

Individual model: For each subject appearing in the dataset, we train an HMM-based classifier so that the driving patterns of each subject is specifically modeled with HMMs. As only one series exists for one subject, we use the same series for training and testing.

Generic model: With the aim to characterizing the generic driving patterns under distraction, an HMM-based classifier is trained with all the series in our dataset. For testing, we still use the whole dataset.

Leave-one-out model: To evaluate how challenging is it for a generic model to process driving patterns of unseen subjects, we adopt the leave-one-out strategy. In particular, we only excluding one subject in the dataset from training, and train an HMM-based classifier with the series of all the rest subjects. The excluded subject is then used for testing. All the subjects are enumeratively excluded from training once so that all of them can appear once for testing.

Model performance: Individual vs. Generic vs. Leave-one-out

The whole dataset is used for model performance evaluation. Performance was assed based on accuracy, equal error rate, Precision-recall (which is generated by tuning the threshold *C), and* AUC (i.e., area under P-R curve). In general, individual based model has a much higher accuracy (0.88) and lower error rate (0.27) than generic models (0.59 for accuracy, and 0.38 for error rate) (Figure 3).

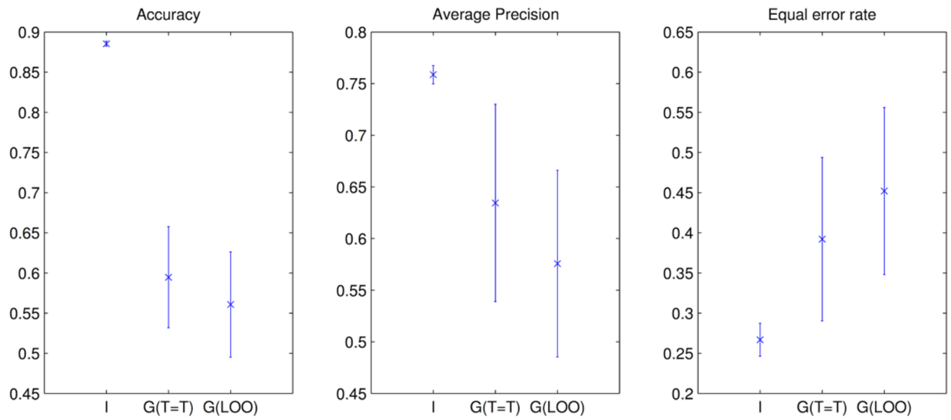

Figure 3: Performance measures output (I: individual T=T: generic Train=Test, LOO: generic Leave-one-out).

Figure 4 shows the ROC curves which is in the upper triangular region away from the diagonal line, meaning the metric being tested has useful information predicting texting behavior better than random guessing. The area under the ROC curve (AUC) was also calculated which indicates the probability that the classifier will rank a randomly chosen positive instance (texting while driving in our case) higher than a randomly chosen negative instance (baseline driving).

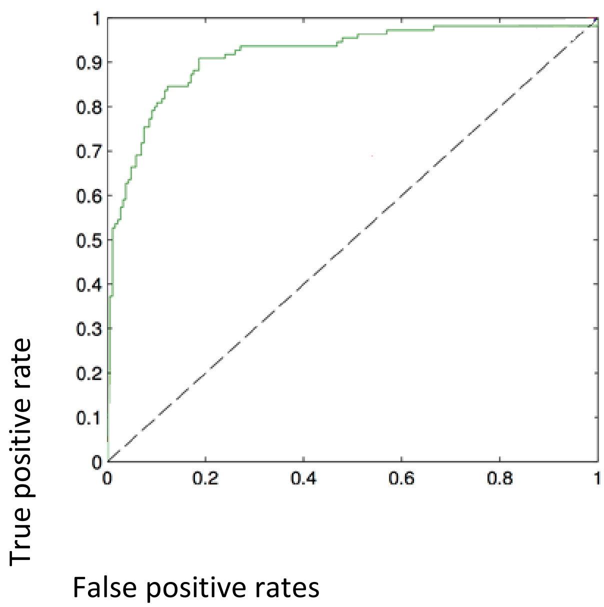

Figure 4: ROC curves of identifying texting from baseline driving

The density distribution of lane offset power was shown in Figure 4. The result shows that texting tasks led to a higher variation of lane offset when compared to the condition before the tasks, suggesting that drivers had erratic lane control right after they started texting tasks, compared to the baseline driving.

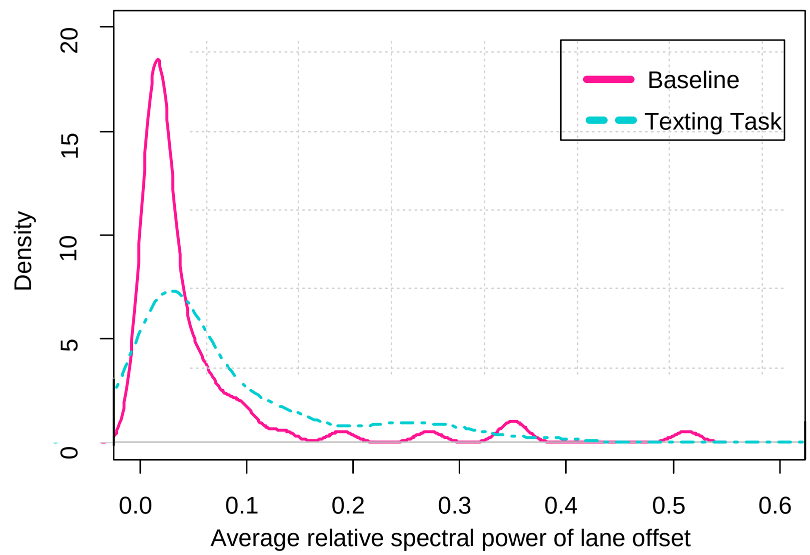

Figure 5. Density distribution of the average relative spectral power of lane offset

Conclusions

This study applied HMM techniques to predict drivers' texting behavior by using vehicle kinematic features from an existing naturalistic driving study. Both time and frequency domain analysis methods to compare driving performance during distracted driving and baseline driving were conducted. FFT analyses were performed in this study to investigate the frequency characteristics of drivers' behavior during distracted driving. The Fourier analysis can describe the variation of driving performance measurements by integrating the power over different parts of the frequency spectrum, and thus distinguish behavior that would otherwise be indistinguishable. For this analysis, the degree of variability was measured in low frequency band (i.e. 0~0.5 Hz). Driver vehicle control performances can be significantly degraded during this period. Results from the frequency domain analysis showed that drivers behave significantly different when they start engaging in texting tasks when compared to baseline driving. It was also found that in this research that stochastic modeling algorithms like HMM can be a useful technique to detect and monitor driver state. It was also found in this study that individual based models works better than generic models.

References

Bao, Shan, Flannagan, Carol A, & Sayer, Jim. (submitted). Internet Surfing and Texting While Driving: A Simulator Study of Teens and Adults. *Accident Analysis and Prevention*.

Drews, Frank, Yazdani, Hina, Godfrey, Celeste N., Cooper, Joel M., & Strayer, David. (2009). Text messaging during simulated driving. *Human Factors, 51*(5), 762-770.

Kevin Murphy, "Hidden Markov Model (HMM) Toolbox for Matlab", http://www.cs.ubc.ca/murphyk/Software/HMM/hmm.html.

Lee, John D., Caven, Brent, Haake, Steven, & Brown, Timothy L. (2001). Speech-based Interaction with In-vehicle Computers: The Effect of Speech-based E-mail on Drivers' Attention to the Roadway. Human Factors, 43, 631-640.

Liang, Yulan, & Lee, John D. (2010). Combining cognitive and visual distraction: Less than the sum of its parts. *Accident Analysis and Prevention, 42*(3), 881-890.

Nemme, Heidi E., & White, Katherine M. (2010). Texting while driving: Psychosocial influences on young people's texting intentions and behaviour. *Accident Analysis and Prevention, 42*, 1257-1265.

NHTSA. (2016). Distracted Driving2014: Traffic Safety Facts Research Note. Report No. DOT HS 812 260. Washington, DC: U.S. Department of Transportation.

National Motor Vehicle Crash Causation Survey (NMVCCS, 2008), https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/811059

Owens, JM, McLaughlin, SB, & Sudweeks, J. (2011). Driver performance while text messaging using handheld and in-vehicle systems. *Accident Analysis and Prevention, 43*, 939-947.

Redelmeier, DA, & Tibshirani, R J (1997). Association between cellular-telephone calls and motor vehicle collisions. *The New England Journal of Medicine, 336*(7), 453-458.

Sayer, J, LeBlanc, D, Bogard, S, Funkhouse, D, Bao, S, & Buonarosa, M. L. (2011). Integrated Vehicle-Based Safety Systems (IVBSS) Field Operational Test Final Program Report: University of Michigan Transportation Research Institute,Ann Arbor, MI.

Strayer, David, & Drews, Frank. (2007). Cell-phone induced driver distraction. *Current Directions in Psychological Science, 16*(3), 128-131.

Strayer, David, & Johnston, William. (2001). Driven to distraction: Dual-task studies of simulated driving and conversing on a cellular phone. *Psychological Science, 12*, 462-466.

Contacts

For more information:

Shan Bao, Ph.D.

University of Michigan Transportation Research Institute

2901 Baxter Road

Ann Arbor, MI 48109-2150

734-936-1127

shanbao@umich.edu

www.umtri.umich.edu

NEXTRANS Center
Purdue University - Discovery Park
3000 Kent Ave.
West Lafayette, IN 47906

nextrans@purdue.edu
(765) 496-9724
www.purdue.edu/dp/nextrans