TECHSUMMARY October 2017

State Project No. DOTLT1000053 | LTRC Project No. 15-1SA

Exploring Naturalistic Driving Data for Distracted Driving Measures

INTRODUCTION

Distracted driving is a dangerous epidemic that continues to cause deaths and injuries in related crashes throughout the U.S. According to the National Highway Traffic Safety Administration, 3,328 people (including 540 non-occupants) were killed and an estimated additional 421,000 were injured in 2012 from distraction-affected crashes. In Louisiana, a recent study funded by LTRC and UTC, "Distracted Driving and Associated Crash Risks," concluded that texting and talking to passengers while driving impaired driving performance but failed to find any significant effects for cellphone conversation. The study was, however, unable to make any statistical findings on the driving performance based on demographics and road facility type because of the limited sample utilized. With the recent availability of data from the Strategic Highway Research Program (SHRP 2) Naturalistic Driving Studies (NDS), there is ample opportunity to utilize a bigger sample size in a further study that will allow statistical conclusions to be drawn on various strata including gender, road facility type, age, and time of day. Therefore, this study aimed to perform a comprehensive exploration of the SHRP 2 NDS data with the view of identifying if it can provide the data required for an enhanced study on the crash risks of distracted driving. This study also included an outline for the development of a Crash Risk Index to evaluate potential risk associated with drivers based on their socioeconomic characteristics and secondary task involvement.

OBJECTIVE

The study achieved the following objectives: (1) conduct a thorough literature review of nationwide laws regulating distracted driving with particular emphasis on cellphone conversation (handheld and hands-free) and texting; (2) thoroughly explore the SHRP 2 NDS database; (3) identify appropriate performance measures that can be used as surrogate measures of distraction; (4) outline a methodology of developing a crash risk index; and (5) make recommendations for enhanced research on distracted driving based on the SHRP 2 NDS data.

SCOPE

The research was restricted to the use of the SHRP2 NDS data to achieve the study objectives. Traditional statistical analysis and artificial intelligence techniques were conducted on the driver behavior (time-series data), events, and demographic data obtained from SHRP2 NDS database.

METHODOLOGY

The methodology to achieve the research objective included (a) performing a comprehensive review of the NDS data to identify the appropriate sample that can potentially represent Louisiana drivers, (b) reviewing the available performance variables in the SHRP 2 NDS, and (c) conducting a statistical assessment on each variable's appropriateness as a surrogate measure to quantify distractions. For the surrogate measure selection, statistical analysis and artificial intelligence were performed. For each type of modeling, the data went through several steps of data cleaning and reduction. Finally, the NDS data were explored to develop an outline for a crash risk index.

LTRC Report 580

Read online summary or final report:
www.ltrc.lsu.edu/publications.html

PRINCIPAL INVESTIGATOR:
Sherif Ishak, Ph.D., P.E.
225.578.4467

LTRC CONTACT:
Kirk Zeringue, P.E.
225.767.9169

FUNDING:
SPR: TT-Fed/TT-Reg

Louisiana Transportation Research Center

4101 Gourrier Ave
Baton Rouge, LA 70808-4443

www.ltrc.lsu.edu

The time series variables were analyzed to identify the surrogate measures that can be used to detect distracted driving behavior. To achieve this, a statistical model was developed using multiple logistic regression (MLR) analysis to detect distracted driving behavior using the five time-series performance variables: GPS Speed, Lateral Acceleration, Longitudinal Acceleration, Throttle Position, and Yaw Rate. In other words, five driving performance measures were used to detect the secondary task in which a driver was engaged. Three separate MLR tests were completed to compare the overall statistical output between a control group (no secondary task) and three individual cellphone based groups of secondary tasks (Group 1: cellphone talking/listening, Group 2: cellphone texting/dialing, and Group3: interaction with adjacent passenger). The three models were developed using (1) all data, (2) the data broken down by driver's age, and (3) the data broken down by driver's gender.

For the neural networks modeling, a supervised feed-forward network with backward propagation (FFBP) was used to develop the detection models. FFBP architecture is well-known for its ability in solving pattern recognition problems. A sigmoid function was used as an internal transfer function. Three hidden layers were selected because of the large size of the data (10,000 observations for each secondary tasks) such that a reasonable number of neurons can be selected in each layer. The Levenberg-Marquardt algorithm was selected for the performance (optimization) function. The model output defines whether a secondary task was associated with the driving behavior or not. Therefore, a binary outcome of 0 or 1 was used, where 1 indicates association with a secondary task and 0 otherwise. Three ANN models were developed individually for each type of secondary task. The input layer included five neurons to represent the selected five driving performance attributes. After a preliminary analysis to improve the model accuracy, the number of neurons in the hidden layers was selected as 9, 13, and 7 for the first, second, and third hidden layer, respectively. To develop the three models, the dataset for each secondary task was randomly divided into 70% for training, 15% for validation, and 15% for testing.

CONCLUSIONS

This research showed how useful the SHRP2 NDS data could be for distracted driving studies. Although the statistical analysis results of this research cannot be taken as credible in most cases, they showed that the high-quality and high-resolution data available on the SHRP 2 NDS database can provide useful insight on detecting distracted driving. Identifying the right surrogate measures and the use of a more suitable analysis tool that can recognize nonlinear patterns in driving behavior can help in detection of distracted driving behavior. Therefore, the neural network modeling was deployed to analyze the five performance measures. Unlike the multiple logistic regression, the neural network analysis identified the five-time series measures as important surrogate measures of distracted driving behavior. The developed neural network models also showed high accuracy in detecting drivers' engagement in secondary tasks. The proposed crash index outline can also provide an insight on quantifying the crash risk associated with distracted driving behavior.

RECOMMENDATIONS

While the objectives of the proposal were met, the SHRP2 NDS database has further data that can be used to expand the state of knowledge in several areas of distracted research. Thus, recommendations were made to (1) perform statistical analysis on the SHRP 2 NDS and socioeconomic characteristics data to provide a clear insight on the proposed crash index; (2) identify whether further recommendations can be made about the available state regulations of cellphone use; (3) integrate SHRP 2 NDS with the Roadway Information Database to evaluate the crash risk associated with distracted driving behavior at different roadway facility types; (4) investigate whether it is possible to identify the type of secondary task using artificial intelligence; and (5) investigate changes in driving pattern before crashes and near crashes take place when specific secondary tasks are performed.