Summary of Final Report, Contract BC355-08
November 2004

LINKING CRASH PATTERNS TO ITS-RELATED ARCHIVED DATA

PROBLEM STATEMENT

There are considerable amounts of data that are collected and stored for ITS applications. These data include speed, volume and lane-occupancy provided by loop detectors. Most of these variables are known to be related to crash occurrence and patterns. Various traffic safety studies have shown variables such as speed and volume to be related to crashes. However, this volume is measured as ADT (Annual Daily Traffic) or hourly volumes that happen to be a static and aggregate measure of traffic flow. Instantaneous measures of traffic flow variables obtained from loop detector data would be useful for detecting crash prone conditions on freeways (i.e., to anticipate crashes in real-time). However, the relationships between these instantaneous variables and crash patterns must be established before such a capability will be possible. Investigating the relationships would improve current understanding of safety problems and would also help to predict crashes in real-time.

OBJECTIVES

The main objectives of this research are identified as follows:

1. To prepare a detailed database with relevant characteristics of the all crashes.
2. To identify the location and time of occurrence for all crashes and extract the corresponding loop detector data from the archived ITS database; to inspect the trends in crash patterns over the period of analysis.
3. To identify a methodology for linking the ITS-related archived data to crashes in order to identify the spatio-temporal patterns for the variation of crash risk.
4. Based on the statistical methodology, to develop a classification model to differentiate between crash prone and normal conditions on the freeway.
5. To demonstrate the online applicability of the methodology for generating the patterns of variation in crash risk at freeway segments along with the classification accuracy.

FINDINGS AND CONCLUSIONS

This project involves the development of real-time crash prediction models for the Interstate-4 corridor in the Central Florida area. Crash data (collected from 1999 through 2002) for a 36.25-mile stretch of the freeway has been used to link crash occurrences with real-time traffic patterns observed through loop detector data.

The analysis technique adopted for this study is within stratum matched case-control logistic regression. The purpose of matched case-control analysis is to explore the effects of independent variables of interest on the binary outcome while controlling for other confounding variables through the design of the study. In the context of this project, crash or non-crash is the binary outcome, and traffic parameters are the independent variables. The design of the study allows controlling for external factors such as geometric design of the freeway, time of the day, and day of the week.

Following the aforementioned analysis technique, a series of crash prediction models were estimated based on the statistical link between crash occurrence and the turbulence in the traffic flow observed through the loop detectors. First, the simple models (involving one co-variate) were estimated, following the exploratory analysis. Next, a multi-variate logistic regression model was estimated. For the final model, 5-minute average occupancy and 5-minute standard deviation of volume observed at the downstream station, during the slice of 5-10 minutes prior to the crash along with the 5-minute coefficient of variation in speed at the station closest to the location of the crash during the same time slice, were found to affect the crash occurrence most significantly. This model was used to calculate the log-odds ratio of observing a crash versus not observing one. A threshold value for this ratio could then be set in order to determine whether the location has to be flagged as a potential "crash location." It was shown that using 1.0 as the threshold for the log odds ratio, over 62% crash identification was achieved from the final model. This threshold value could be changed to achieve better crash prediction accuracy. However, false alarms might occur more frequently, i.e., there is a tradeoff between crash and non-crash prediction accuracy.

Researchers observed that although the simple models have the advantage of being tolerant in their data requirements, their classification accuracy is poorer than that of the final multi-variate model. Hence, the simple models were used to deduce spatio-temporal patterns of the variation in crash risk. Simple model output may be recommended for preliminary assessment of the crash risk. If there is an indication of high crash risk, however, the multi-variate model is preferred to explicitly classify the data patterns as leading to or not-leading to crash occurrence.

BENEFITS

The extensive review of the literature shows that very limited work is starting in this area of research, most of which is theoretical and without a solid implementation plan. The experience gained in this project is vital for developing a fully functional proactive traffic management system. The models and implementation plan would contribute to the incident detection capabilities of traffic management authorities. The research findings have resulted in an improved understanding of traffic crashes and preventative measures.

To reap the full benefits from this research, however, requires that more work be done (i.e., to incorporate other data elements that are related to traffic safety, to achieve better real-time prediction capabilities). Such issues were not adequately understood at the outset of this project. With a fully functional proactive traffic management system, driver information provided by TMCs could be tailored to alleviate situations that may be expected to lead to crashes. The type of information currently provided by TMCs would change to address safety (crash prevention). Currently, the information provided addresses delay and congestion issues, without much direct concern to traffic safety. Finally, the results of this research could be very helpful to current FDOT efforts to introduce variable speed limits (VSL) to Florida Freeways.

This research project was conducted by Dr. M. A. Abdel-Aty, of the University of Central Florida. For more information, contact Mr. Liang Hsia, Project Manager, at Ph. (850) 410-5615, liang.hsia@dot.state.fl.us