Big Data in the transportation arena, enabled by the rapid popularization of Intelligent Transportation Systems (ITS) in the past few decades, could considerably enhance experts' understanding of their system due to the real-time nature. This research explored how to efficiently use Big Data for urban expressway safety and efficiency evaluation and improvement. The urban expressway system operated by Central Florida Expressway Authority (CFX) is equipped with Automatic Vehicle Identification (AVI) and Microwave Vehicle Detection System (MVDS) for traffic along with multiple Dynamic Message Signs (DMS), reaching 109 miles. The traffic congestion and safety on these expressways were evaluated.

Three congestion measures based on the ITS systems were introduced to evaluate the expressway mainline and ramp congestion in real-time and lead to consistent conclusions. Among them, the Congestion Index and occupancy derived from MVDS data outperformed Travel Time Index from AVI data. They can accurately reflect the congestion segments and locations, congestion time and duration, and congestion intensity in spatio-temporal dimensions. The high density of deployment of MVDS system and the detailed traffic flow information collected by MVDS system could be responsible for the good performance of these two congestion measures. Based on the congestion evaluation, improvement strategies were proposed through the DMS.

Substantial efforts have also been dedicated to Big Data applications in safety evaluation and improvement. Bayesian multilevel models proved that the uncapped AVI data retained the original speed information and performed better in safety analysis compared with the capped AVI data. Geometric factors such as inside shoulder width, horizontal curves and existence of auxiliary lanes were identified to significantly affect crash frequency on expressways along with traffic volume, speed and speed variation.

With the introduction of MVDS system and more data collected, a more thorough safety analysis using both AVI and MVDS system was conducted. For an interpretation of congestion's effects without interference, multicollinearity was checked and addressed by introducing ridge regression. The correlated random effects model distinguishing congestion's effects by peak and non-peak hour proved congestion is a major concern for safety during the peak hours.

While the Big Data could be used in crash frequency analyses, their true power lies in real-time traffic safety evaluation. Data mining techniques and Bayesian logistic regression were used to identify the significant traffic variables contributing to crashes in real-time. Their results were further leveraged in reliability analysis to identify the appropriate time to trigger the safety warnings. Close relationship between congestion and rear-end crashes were found and improvement strategies were suggested to simultaneously improve operation efficiency and safety using the real-time traffic data.

Big Data could also be used for emergency response. Using real-time traffic data both before and after crash, the effects of crashes on congestion were classified. It was found that not all crashes had a direct impact on traffic congestion. Bayesian logistic models found that peak hour, number of lanes, weather and crash severity could significantly alter the effects of crashes on operation.

In the end, future extensions of Big Data applications in traffic studies were raised.

Major: Civil Engineering

Educational Career:
Bachelor's of Transportation Engineering, BS, 2011, Tongji University, China
Master's of Civil Engineering, MS, 2013, University of Central Florida