Crowdsourcing and Data Analytics for Near-Miss Statistics

VERMONT AGENCY OF TRANSPORTATION RESEARCH PROGRAM

STUDY TIMELINE
11/2022 – 10/2023

INVESTIGATORS

Mazdak Tootkaboni, UMassD, PI, mtootkaboni@umassd.edu

Arghavan Louhghalam, UML, Arghavan_louhghalam@uml.edu

Franz-Josef Ulm, MIT, ulm@mit.edu

Mohammad Pourghasemi Saghand, UMassD

KEYWORDS
Near-miss, crowdsourcing, safety

FUNDING
Michelin DDI

More information about the Agency of Transportation Research Program, including additional Fact Sheets, can be found at: http://vtrans.vermont.gov/planning/research

Introduction (or Problem Statement)

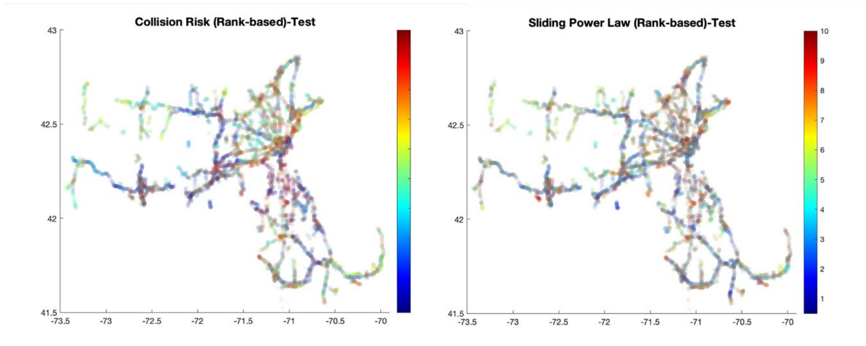

Fig: Road segments ranking based on collision risk computed from publicly available data (left) and the Sliding Power Law predictive model (right)

"Near-miss" traffic situations offer essential insights for all stakeholders in road transportation, including individual drivers, traffic planners, fleet managers, and the vehicle insurance industry. However, tracking and quantifying these events is challenging because they are represented by extreme statistics that cannot be captured through simple averaging procedures. Our framework aims at addressing this challenge by integrating the available crowd sourced data with physic-based models.

We examine the feasibility of using available crowdsourced vehicle speed data for identifying and mapping near-miss statistics. Extreme deceleration on road segments is identified as an indicator of the likelihood of near-miss events and is probabilistically linked to collision risk across various traffic densities.

Project Methodology

We relied on a model-based framework using the Nagel–Schreckenberg (NaSch) model and developed a telematics data-informed NaSch model that uses speed profiles of vehicles (at 1Hz) as input data. The model was employed to examine the internal structure of traffic and calculate the probability of deceleration to a full stop on individual road segments, as an indication of the likelihood of near-miss events. We then devised a density-dependent predictive model to effectively relate the likelihood of near-miss events to the expected collision risk, using publicly available data (geospatial distribution of crashes and AADT). The analysis was performed both in an absolute sense and from a ranking perspective.

Conclusions/Next Steps

The telematics data-driven models proved to offer great potential for providing quantitative measures of risk and near-miss events. For future, we plan to improve the proposed predictive model for low and high traffic densities, where deceleration to a full stop might not be the only indicator of near-miss events. This will be done through investigating better-parametrized models or non-parametric models to link an expanded set of internal variables to collision risk; and exploring predictive models that incorporate variables beyond those related to internal structure of traffic.

Impacts and Benefits

The availability of statistics on near-miss events has significant benefits. It will help identify the factors contributing to accidents and reveals trends and patterns in drivers' behavior that is not otherwise recognized. Such insights allow for devising proactive safety measures to effectively mitigate the risk of accidents.