

---

**DISCOVERING CAUSALITY IN TRAFFIC SENSOR READINGS FOR ROAD ACCIDENTS IMPACT PREDICTION**

*Liyue Fan, University of Southern California, liyuefan@usc.edu*

*Ugur Demiryurek, University of Southern California, demiryur@usc.edu*

*Cyrus Shahabi, University of Southern California, shahabi@usc.edu*

Road accidents, a key contributor to traffic congestion, pose serious concerns to drivers, law enforcement, and transportation agencies. Reducing the impact of traffic accidents has been one of the primary objectives for transportation policy makers. The wealth of data collected from traffic sensors and accident logs offers an unprecedented opportunity to mine and understand the traffic incidents towards mitigating the consequences. In this poster, we will utilize the real-world datasets in our data warehouse and study to predict and quantify the impact (i.e., backlog and clearance-time) of road accidents on the up-stream traffic direction and in the surrounding network (e.g., arterial streets) of the accident. Our data are collected from different transportation authorities in Southern California, and include archived traffic sensor readings and accident reports. We implement our previous work on predicting impact of road accidents in upstream traffic direction, which essentially classifies traffic accidents based on their features and models the impact of each accident class on its upstream traffic by analyzing the archived traffic data. However, in reality traffic accidents may cause surges in traffic demand in their vicinity, such as adjacent arterial streets and freeways, where our previous work cannot be directly applied. To this end, we will present our newly developed methods, which investigate the underlying temporal dependencies among time series of sensor readings at different road segments. As a result, we are able to quantify the temporal-causal effect of traffic speed between sensor locations and make effective predictions by identifying local dependency structures. To mitigate the noise and random fluctuations in sensor readings, we apply a variety of time series preprocessing techniques to discover the inherent dependency between traffic speed at different locations. To eliminate spurious causation induced by unobserved confounders, we adopt Granger graphical models for the sensor time series data collected from Southern California highways. With the discovered dependency network, machine learning techniques will be applied to predict the start time and speed change for impacted road segments at the onset of an accident.