Development of Cost-Effective Sensing Systems and Analytics (CeSSA) to Monitor Roadway Conditions and Mobility Safety

PI: Chun-Hsing Ho
Northern Arizona University
Chun-hsing.ho@nau.edu

Project Objective

The objectives of the project are to (1) determine promising computing models for prediction of pavement distresses (2) evaluate the promising/optimum placements/locations of sensor loggers within a vehicle, and (3) provide an affordable method that will benefit state, city, county governments, as well as local communities who have an immediate need but with limited budgets to evaluate the road quality and prioritize repair needs.

Problem Statement

Pavement condition surveys normally involve data acquisition, interpretation, and documentation. Automated pavement condition survey is considered one of most commonly used methods for pavement condition assessments. However, as of today, the automated pavement condition surveys have not yet been well adopted by highway agencies as a promising system due to the fact that a well-established dynamic vibration and sensor models that can accurately capture the signatures of pavement surface condition have not been proposed and widely adopted by transportation authorities due to its costly expenses on the equipment and software that an agency has to invest at front. Therefore, there is a need to address this issue by developing a pavement sensing technology affordable and applicable to the community.

Research Methodology

A sensor logger consisting of ADXL 335 triple-axis accelerometers, Arduino MKR1000 computer boards, GPS, and a battery was designed (Figure 1) by the research team for field data collection. Vibration data were obtained from a yearlong field test. Before collection, four sensor loggers (named as M1, M2, M3, and M4) developed by the research team were placed on the top of the control arm with another one (named as M5) being placed inside of the vehicle to gather vibration data simultaneously. During testing, the pavement temperature was recorded by an infrared thermometer and the driving speed was maintained at 60 miles per hour except the traffic congestion occurred in road testing. To select the most suitable machine learning method for the determination of pavement conditions, we implemented four commonly used machine learning models: Logistic Regression (LR), Support Vector Machine (SVM), Random Forest (RF), and Neural Network (NN). Additionally, statistical analyses were used to determine thresholds that can be used to identify major pavement distress known as point of interest (POI).

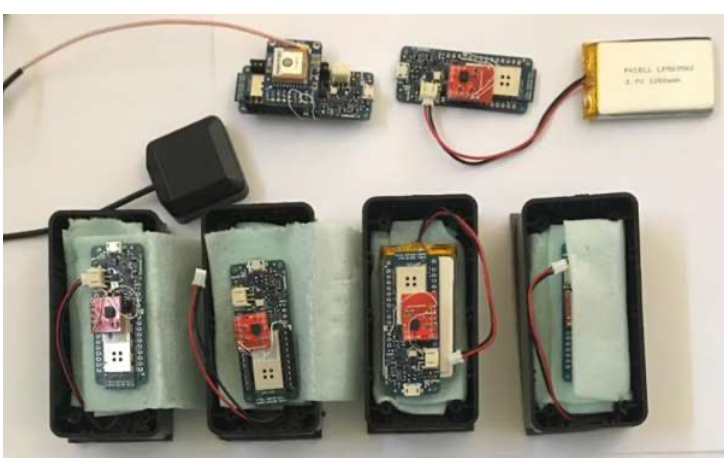

Figure 1. Components of sensor logger

Results

Based on machine learning analysis, it is determined that RF has the highest accuracy and the best performance than the other three algorithms to predict pavement conditions (good, fair, and poor) as shown in Table 2. The Tukey's Test shows that the mean of magnitudes from sensor 5 (M5), which was placed inside of the vehicle is significantly different from either of other four sensors (M1 to M4).

Table 1. The performance of the Random Forest Model

MCC: 0.7654		**Accuracy:** 0.98	
	Precision	**Recall**	**F1-score**
Good	0.99	1.00	0.99
Fair	0.86	0.63	0.73
Poor	1.00	0.67	0.80
Average	0.95	0.77	0.84

Figure 2. Pavement temperature and selected POIs

The result highly recommends excluding the use of fifth sensor (inside a vehicle) for data collection as its magnitudes of vibration are substantially different from the other four sensors placed on the control arms of a vehicle, or M5 should be used individually for data collection to avoid any misinformation. Furthermore, the computation from statistical analysis indicated that a value of 1.7g can be used as a threshold based on the 99% confidence level from the statistical results.

To further understand the effect of temperatures on the pavement conditions, an ANOVA test was conducted to evaluate if the pavement temperature has a significant impact on the number of POIs. The ANOVA results show that there is an association between two variables at the significance level 0.05 in both sections which indicates the pavement temperature does play an important role in controlling pavement conditions as expressed in Figure 2 and Table 2. Based on Time Series function, the Autoregressive Integrated Moving Average (ARIMA) model was used to fit the selected number of POIs to crease a forecasting plot as displayed Figure 3. Clearly, after the period of 11 months, the predicted number of POIs has an increasing trend pattern that is within the 95% prediction intervals in the following two years (up to a total of 35 months) which indicated that the POIs of pavements will be substantially increased by means of accelerated deterioration in two years if the maintenance and rehabilitation will not be scheduled. In general, the sensing technologies and algorithms developed in the project are cable of capturing pavement vibration patterns and determining a level of pavement distress.

Table 2. ANOVA test results of pavement temperature and POIs

Analysis of Variance Table: Section 2 – Baseline Rd. through Chandler Blvd.					
	df	Sum Square	Mean Square	F value	P-value
Sensor	5	83.417	16.683	335.64	$< 2.2 \times 10^{-16}$
Temperature	1	0.112	0.112	2.286	0.0331
Residuals	49	2.436	0.049		

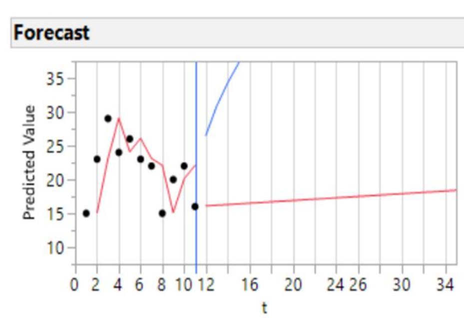

Figure 3. The forecasting plots of number of POIs