This topic is "practice ready."   ☒ **Yes**   ☐ **No**

Implementing Data Quality Assurance with Visualization Tools

Tingting Huang[1], Skylar Knickerbocker[2], and Anuj Sharma[3]

Abstract

The ability to respond and make effective traffic management decisions requires a data driven approach of which data quality must be the primary focus. It is common for agencies to have multiple traffic data sources each requiring different procedures to identify and establish quality. For example, a traffic sensor may be working as expected but the data stream is of no value (sensor is pointing at the sky). The use of probe data presents another type of challenge as agencies have to consider the amount of real-time data that is available across the network (interstate may have great probe coverage and collectors may have no coverage). This study applies a self-health-monitoring technique to identify sensor faults and data errors based on over 400 traffic sensors installed across the State of Iowa. The study also developed a tool to determine the amount of real time data available by roadway from probe data sources. Data collected by sensors include average vehicle speed, volume, and occupancy in 20 second intervals while the probe data includes speed and travel time collected every 1 minute statewide for roadway segment lengths that vary between .2 to 10+ miles. These high resolution data sources necessitate a big data computing and visualizing approach.

For sensor data, several performance measures were established to evaluate the sensor data, including error rate by type, downtime, vehicle volumes and speed distributions, etc. In the interactive tool for visualizing these measures, the measure selector is used to dynamically change between the calculated performance measures and to update the display panel. There are multiple ways to select the sensors of interest, which helps facilitate the use by engineers from different jurisdictions. Also, a speed heat map could be obtained from this tool to help visualize the speed data and check out how the sensor error

[1]Graduate Research Assistant; InTrans, Iowa State University; 2711 S. Loop Dr., Suite 4700, Ames, IA, 50010; phone (515) 686-0925; email: thuang1@iastate.edu

[2]Research Engineer; InTrans, Iowa State University; 2711 S. Loop Dr., Suite 4700, Ames, IA, 50010; phone (515) 294-2238; email: sknick@iastate.edu

[3]Associate Professor; Department of Civil, Construction and Environmental Engineering, Iowa State University; 2711 S. Loop Dr., Suite 4700, Ames, IA, 50010; phone (515) 686-1471; email: anujs@iastate.edu

reflects on the speed profile. This tool serves to visualize Iowa DOT sensor performance for assessing re-calibration, troubleshooting, and maintenance scheduling.

For probe data, the amount of real-time probe data can vary by roadway and time of day across the State of Iowa. The visualization tool shows the amount of real-time data available on any roadway segment. The thematic map shows the percentage of real-time data by roadway and time period. A user can further refine the results to a specific year, month, or hour of the day. This information is guiding researchers and the Iowa DOT on the availability of data to accurately calculate performance measures or use the data as part of an analysis.

Keywords: Quality Assurance—Sensor Data—Probe Data—Visualization—Big Data Application