NCHRP 03-138: Application of Big Data Approaches for Traffic Incident Management (TIM)

SSP Idea Sharing Network Show and Tell

March 10, 2021

Overview of Presentation

- Overview of the NCHRP 03-138 project

- What is meant by "big data"?

- Current TIM data challenges

- Possible big data use cases and pipelines

- SSP use case

- Agency involvement

Overview of NCHRP 03-138

- Objective
  - Demonstrate the feasibility and practical value of big data approaches to improve TIM
- Tasks
  - Gather information and data
  - Store and assess a wide variety of data to support the TIM use cases
  - Define potential big data use cases
  - Develop big data pipelines, analytics, and data products for each use case
  - Refine and enhance the guidelines presented in NCHRP Research Report 904 Leveraging Big Data to Improve TIM

What is Meant by "Big Data"

- The term "big data" represents more than just "a lot of data."

- Big data refers to not only the <u>volume</u> or amount of data available, but also the <u>velocity</u> or speed at which the data are available, the <u>variety</u> or diversity of datasets available, the <u>veracity</u> or trustworthiness of the data, and the <u>value</u> that data bring to an organization.

- More importantly, big data represents a <u>fundamental change in how data are collected, managed, analyzed, and used</u> by organizations both to support real-time operations and to uncover important trends and relationships that could improve transportation operations and safety.

Current TIM Data Challenges

- Making use of existing data sources (in house and partner agencies)

- Applying new types of data (e.g., probe data, crowdsourced data) to improve TIM

- Data, technical, and institutional readiness for big data

Potential Big Data Use Cases

1. TIM timeline & performance measures

2. Improving incident detection

3. Predicting the duration of incidents

4. Understanding secondary crashes

5. Improving safety service patrol (SSP) operations

Potential Use Case – Improving SSP Operations

- Question: Leveraging big data, can SSP operations be improved (and updated more frequently based on changing conditions)?

- Current practice / problem or limitation: Many mature SSP programs have established operations that are modified occasionally as necessary. These programs operate using relatively "fixed" criteria that are not often changed to reflect changes in traffic, incident trends, or expected weather or special events. Driven more by funding than data.

- Big data opportunity: Develop a data pipeline that ingests big data sources, integrates them with various DOT/partner data sources, and applies big data techniques to optimize SSP operations on a more frequent basis to improve operational efficiencies.

- Potential data sources: Waze, historical ATMS; roadway inventory; SSP planned operations, assists, and AVL; weather; and special events.

What is the Ask?

- Understand current SSP operational challenges/limitations that could be improved by a more data-driven / big data approach.

- Share data to support the development of a modern data pipeline.

- Follow-up call to involve agencies interested in sharing information/data.

Big Data Pipelines

- Apply a process to refine, transform, and clean the raw data.

- Apply descriptive, inferential, predictive, prescriptive, or causal analyses to the transformed and cleaned data to generate a resulting data product.

- Data exiting a modern data pipeline can be routed directly to an application such as a visualization tool, to a messaging application, or even another data pipeline.

A modern, big data pipeline is a sequence of processing and analysis steps applied to data for a specific purpose.

Data → Collect → Enrich → Aggregate → Filter → Format → Alert

Aggregate → Store → Analyze → Visualize