Mobility Data Analytics Center

FINAL RESEARCH REPORT

Sean Qian (PI),
Xidong Pi, Zhangning Hu

Contract No. DTRT12GUTG11

DISCLAIMER

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated under the sponsorship of the U.S. Department of Transportation's University Transportation Centers Program, in the interest of information exchange. The U.S. Government assumes no liability for the contents or use thereof.

Problem statement

Mobility Data Analytics Center aims at building a centralized data engine to efficiently manipulate large-scale data for smart decision making. Integrating and learning the massive data are the key to the data engine. The ultimate goal of understanding massive data is to accurately estimate the historical usage of the transportation infrastructure and to forecast its future performance. To efficiently balance the infrastructure supply and demand, optimal decisions on management strategies, policies and adoption of technologies can be made.

Approach

Through MAC we propose to develop **a centralized data engine** supported by **a web application** to manage and analyze multi-jurisdictional multi-modal data for safety, mobility and sustainability, using the City of Pittsburgh and the City of Philadelphia as case studies. In the centralized data engine, massive data are stored and managed, and can be further translated into useful information for people who need it: legislators, transportation planners, engineers, researchers, travelers, and companies. Unlike the traditional single computer stand-alone software or tools for data preparation and system design, the data engine relies on web-based data sharing and browser-based human-computer interaction for it to be accessed by users. The web application visualizing data and recommending decisions serves the front end of the data engine.

Methodology

The innovative core methodology adopted in the model development is depicted in Figure 1. The essential idea is to learn the travel demand characteristics (e.g., how many trips are generated on what time of day, using what part of the transportation infrastructure and by which mode) that altogether produce the passenger and vehicle flows in the network consistent with what the multi-modal and multi-jurisdictional data measured in sampled locations and time. The resultant travel demand characteristics allow us to do two things. First, given a system intervention, we are able to predict how users will change their demand and behavior in the transportation system. The system evolution models estimate propagations of multi-modal traffic flow in the network, by which we estimate and forecast passenger and vehicle flow. Second, the passenger and vehicle flows lead to a better estimation and prediction of social cost metrics, such as system delay, emissions and energy consumption. System performance can then by optimized in order to influence travel demand and travel behavior by management strategies and/or information provision. Methodologically, this means the total expected social costs are minimized (social costs include energy use, emissions, delay, safety, etc.).

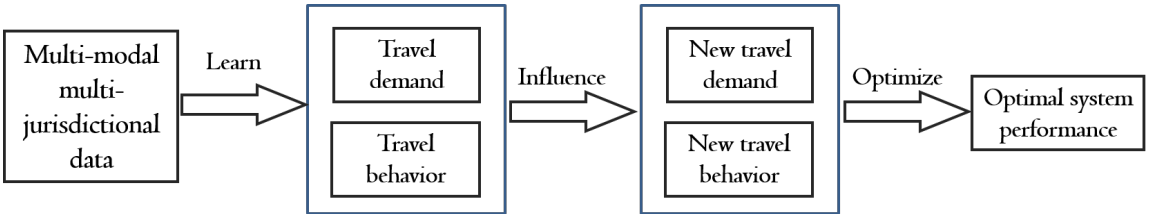

Figure 1 Core methodology of the data engine

Findings and Conclusions

In the past two years, we have started building the data engine and a prototype web application to demonstrate the feasibility of Mobility Data Analytics Center. We started from the Pittsburgh region where we have close partnerships with many local entities, and have successfully applied our data analytics tools in several case studies. The main accomplishments are summarized as follows,

1. We have gained access to various data, which include, but are not limited to, socio-demographic data, traffic speed probe data, traffic incidents, parking meters, transit APC-AVL, weather, social media data. Those data are integrated and visualized through the prototype web application.
2. We have analyzed **large-scale crash data in PA**, and **developed an online tool** to visualize and forecast crash types, frequencies and severity for each PennDOT-owned road segment.
3. We analyzed large-scale APC-AVL transit data to provide both travelers and transit agencies fine-grained customizable information regarding transit service performance (efficiency, reliability and quality), and have developed a **Transit Service Performance Information and Optimization system (TranSEPIO)**. Port Authority is the potential deployment partner. This has lead to a PITA project.
4. We have developed a tool to analyze large-scale parking meter transactions data and to provide **estimation and forecast of time-varying parking occupancies** for each block of on-street and off-street public parking. Pittsburgh Parking authority, MeterFeeder Inc. and several neighborhoods are potential partners. This has yielded a NSF three-year project regarding parking optimization.
5. We have built a **sophisticated transportation network model for the City of Pittsburgh** that describes individual travel activities on transportation systems. Operational strategies and policies can therefore be examined in the network model in terms of **system delay, crash risk, reliability, vehicle-miles traveled (VMT), fuel consumption and emissions**. It has stimulated a research project with PennDOT (Philadelphia traffic impact study) and one with the City of Pittsburgh (traffic impact study of Greenfield Bridge closure). SPC is another potential deployment partner.