Deep-Learning Approach for Bus Arrival Time Forecasting

A study focuses on improving the forecasting capability of the bus reliability analysis system to generate bus arrival time.

WHAT IS THE NEED?

In a previous METRANS-funded project, the research team built a system that uses real-time and historical bus GPS trajectory datasets over the past two years in Los Angeles (LA) County to generate performance metrics. While the system is well-received by the California Department of Transportation (Caltrans), it only considers current and historical data and does not have the capability to forecast the performance metrics, which is important to both the transportation agencies and riders.

In this project, the researchers will study traffic flow prediction in road networks in depth; and use it to forecast the arrival time of individual public transportation vehicles and various public transportation systems performance metrics, when the second-by-second automatic vehicle location (AVL) system is not available.

The major challenges in predicting traffic flow, such as estimating future speed and volume, are due to the spatial and temporal nonlinearities when transitioning between free flow and congestion at rush hours and during non-recurring events, for example, accidents and weather conditions. The research team will conduct fundamental study in developing a novel graph convolutional recurrent network to capture such nonlinearities inherent in road networks. Furthermore, the researchers will develop a web application to show the traffic forecast results concerning real-time traffic conditions, which provides the access and visualization of the predicted performance metrics of public transportation vehicles.
WHAT ARE WE DOING?

The research team will exploit a real-world traffic sensor dataset and California Highway Patrol (CHP) accident logs collected by the Regional Integration of Intelligent Transportation Systems, which the researchers have been continuously collecting and archiving for the last six years under the Archived Traffic Data Management System (ADMS) project.

ADMS is the largest traffic sensor data warehouse built in Southern California, which includes both inventory and real-time data with update rate as high as every 30 seconds for freeway and arterial traffic sensors (14,500 loop-detectors) covering 4,300 miles and 2,000 bus and train AVL and incidents; such as accidents, traffic hazards, and road closures reported, approximately 400 per day.

It fuses and analyzes very large-scale and high-resolution (both spatial and temporal) traffic sensor data collected by different transportation authorities in Southern California, including Caltrans, CHP, LA Department of Transportation, and Long Beach Transit.

WHAT IS OUR GOAL?

The main objective of this study is to develop data mining, machine learning algorithms to effectively forecast flows in both space and time, and a deep-learning method and system that can process massive amounts of:

i. GPS trajectories from public transportation vehicles; and
ii. Real-world traffic sensor datasets archived in the data warehouse to predict traffic flow, which enables the forecasting of a variety of performance metrics of public transportation systems, including travel-time reliability, on-time performance, bus bunching and travel-time estimation.

The researchers will deliver a suite of public web services that implements the prediction algorithms. Any remote user can call these services to run the prediction models based on the data collected.

Another deliverable is an interactive web application that enables visualization, querying, and analysis of real-time and predicted traffic flow as well as performance measurements of public transportations.

WHAT IS THE BENEFIT?

The result and deliverables of this research will improve the transit system, and ultimately enhance the mobility of Californians.

WHAT IS THE PROGRESS TO DATE?

The research is expected to be executed in February 2019.