DRISI
Caltrans Division of Research, Innovation and System Information

Research Results

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.

Modal

FEBRUARY 2020

Project Title:
Deep-Learning Traffic Flow Prediction for Forecasting Performance Measurement of Public Transportation Systems

Task Number: 3401

Completion Date: January 31, 2020

Task Manager:
Bradley Mizuno,
Transportation Engineer (Electrical)
bmizuno@dot.ca.gov

WHAT WAS THE NEED?

Los Angeles is ranked the most congested city in the U.S. with a typical half-hour commute taking 81% longer during evening peak periods and 60% longer during the morning peak. These traffic congestions result in a large social and economic detriment and raise serious concern for drivers and transportation agencies. Therefore, increasing ridership of public transportations and hence reducing traffic congestions has been one of the primary objectives for transportation agencies and policymakers. Previously, many researchers have worked on estimating historical performance measurements of public transportation systems. Beyond historical performance measurements, accurate predictive analysis of performance reliability helps to manage rider expectations as well as to provide a powerful tool for transportation agencies to coordinate the public transportation vehicles. For the first time, there is a unique opportunity to use data-driven approaches that analyze big datasets collected from transportation systems to understand the factors causing traffic congestions and in turn, help to forecast the performance reliability of public transportation vehicles when automatic vehicle location data is not available.

WHAT WAS 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 GPS trajectories and real-world traffic sensor data.

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

Caltrans

DRISI provides solutions and knowledge that improves California's transportation system

WHAT DID WE DO?

This research project developed a deep learning approach for traffic flow forecasting and bus arrival time estimation in Los Angeles. First, a novel Graph Convolutional Recurrent Neural Network (GCRNN) was developed to model and forecast traffic flows at different spatial and temporal resolutions. The GCRNN model considers not only the location of traffic sensors but also their relationships (i.e., topological dependency) in space, which was critical to achieving the best performance for all forecasting horizons compared to the existing methods. Next, a Geo-Convolution Long Short-Term Memory (Geo-Conv LSTM) framework was implemented to model bus Estimated Time of Arrival (ETA) by incorporating the traffic flow predictions of the GCRNN. Using the real-world traffic sensor datasets archived in a data warehouse, the research showed that the proposed bus ETA model is more accurate than the existing method, Gradient Boosted Decision Tree (GBDT), by 27% in estimating bus travel time. Lastly, both models were deployed as web applications so that users can access traffic prediction data and check bus arrival times to a destination location from a starting point.

WHAT WAS THE OUTCOME?

The University of Southern California has a data warehouse, Archived Traffic Data Management System (ADMS), with 11TB annual growth, and is the largest traffic sensor data warehouse built so far in Southern California. Using this big traffic dataset, the research had a unique opportunity to use data-driven approaches to understand the factors causing traffic congestions and in turn, help to forecast the performance reliability of public transportation vehicles. This research developed a reliability analysis system using Deep Learning (DL) techniques to forecast the future performances of the public bus system in Los Angeles.

More specifically, the researchers designed and developed a GCRNN traffic forecasting model that captures spatial and temporal dependencies as well as long-term forecasting. Then they incorporated the traffic prediction results of the GCRNN model to the Geo-Conv LSTM bus ETA model and outperformed the baseline model (GBDT) by 27% in estimating travel times. The system demonstrates the overall approach in an area near downtown Los Angeles and shows that incorporating traffic flow predictions can help to forecast short-term bus arrival times accurately (e.g., in the next few hours).

Although the experimental results show significant improvement compared to the state-of-the-art baselines methods, modeling Deep Learning (DL) techniques to forecast the future performances of the public transportations for large spatial scale (e.g., the entire Los Angeles Metropolitan Area) and long-term (e.g., days instead of hours) remains challenging. Reliable long-term forecasting of performance measurement for public transportation systems over a large area is essential for policymakers to achieve effective city planning as well as promotes ridership. For example, forecasting bus arrival time for the next day helps a rider to plan their commute early. Existing approaches typically rely on traffic simulation tools and models that require expert knowledge to execute and adjust parameters for various traffic scenarios. The researchers propose to expand the current approach and system to develop the capability for processing the entire Los Angeles Metropolitan Area for long-term forecasting of a variety of public transportation system performance metrics.

WHAT IS THE BENEFIT?

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

LEARN MORE

Website for the application:
https://infolab.usc.edu/Caltrans/

Click and hold for a second on "Set Start" to allow the computer to register the click.

IMAGES

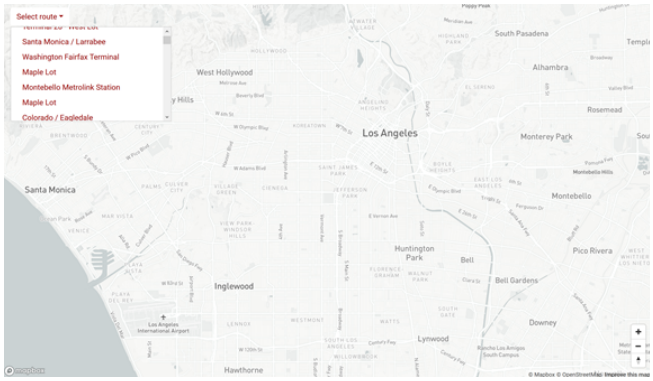

Image 1: Screenshot of the tool showing the route choices.

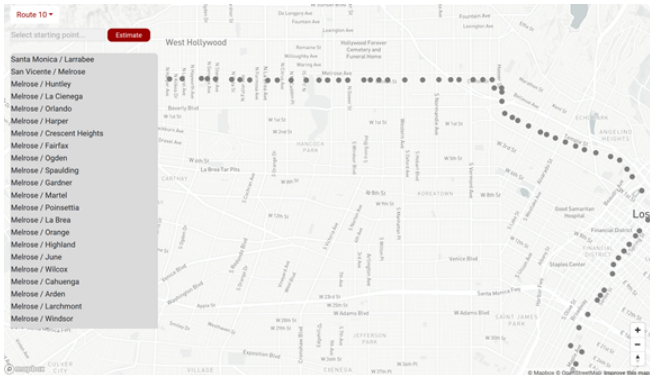

Image 2: Screenshot of the tool showing the bus stops on the chosen route.

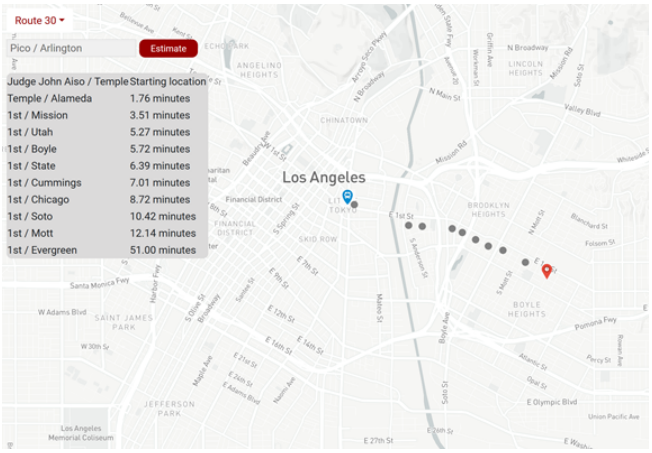

Image 3: Screenshot of the tool showing the estimated arrival times from a chosen stop.