Modeling and Predicting the Cascading Effects of Delay in Transit Systems

Aparna Oruganti, Sanjita Basak, Fangzhou Sun, Hiba Baroud, Abhishek Dubey
1Analytics Associate, Presbyterian Healthcare Services
2Department of Electrical Engineering and Computer Science, Vanderbilt University, 3Research Scientist, Facebook
4Department of Civil Engineering and Department of Earth and Environmental Science, Vanderbilt University

ABSTRACT

Research and Motivation
- Smart cities need effective transit systems planning. With the integration of AVL (automatic vehicle locators), transit agencies are able to provide arrival times estimates in real-time.
- For the study, we use data collected over several months from one such transit system and show how this data can be potentially used to learn patterns of travel time during special events such as sports games and music award ceremonies.

Goal
- Study the effect of NFL football games and other factors such as traffic and weather on the transit system delay using various machine learning algorithms.
- Predict the transit travel time and visualize the cascading effect across the transportation system using heatmaps.
- Use these models to design adaptive and transient transit schedules during scheduled events.

BACKGROUND

- As a part of the smart city initiative, Nashville is involved in the implementation of Intelligent Transportation Systems (ITS).
- Traffic congestion can be caused by
  - Regular rush hour traffic
  - Inclement weather conditions
  - Incidents and special events

- Cascading effects are defined as disruptions in one system affecting the performance of other interconnected systems.
- Traffic in downtown Nashville caused by a football game can cascade to delays in other interconnected road networks resulting in a gridlock.
- Understanding cascading effects of delay can help provide real time and reliable updates on the performance of the transit system.
- The predicted delay and cascading effect can inform an adaptive transit schedule.

METHODOLOGY

- Ensemble Tree based models
  - Random Forests
  - Extreme gradient boosted trees

RESULTS

- Extreme Gradient Boosted Method (XGBoost)
  - An ensemble method used for regression and classification.
  - In gradient boosting, trees are grown sequentially using information from previously grown trees.

- Advantages
  - Better support for multicore processing which reduces overall training time.
  - These enhancements make a big difference in speed and memory utilization.

CONCLUSIONS

- We are able to explain more than 80% of the variance in the bus travel time and we can make future travel predictions for each timepoint segment with an out-of-sample error of 2 minutes with information on bus schedule, traffic, weather and special events.
- Based on predicted delay, cascading effect of traffic in downtown can cover up to 6 mile radius.

Future Work

- Integrate predictive analytics with
  - (a) the decision framework for DelayRadar to help the transit agency develop a dynamic transit schedule during special events,
  - (b) the transit-hub application that provides the delay estimates to riders

ACKNOWLEDGEMENTS

This work is sponsored as part of the Vanderbilt Initiative on Smart City Operations and Research, a trans-institutional initiative funded by the Vanderbilt University. We acknowledge the support and suggestions provided by our partners from Nashville Metropolitan Transport Authority.

REFERENCES