Bus Journey and Arrival Time Prediction based on Archived AVL/GPS data using Machine Learning

Ankit Taparia, Master of Science in Computer Science

University of Dublin, Trinity College, 2020

Supervisor: Dr. Michael Brady

With a surge in the number of vehicles, traffic congestion has increased at an alarming rate. This has led to increase in travel times and decreased accessibility and mobility. One viable solution to mitigate this issue is to promote the use of public transport. Buses are considered one of the important means of public transport. However, to encourage the use of buses, there is a need to provide reliable bus travel time and arrival information to the commuters. In this study, we propose and develop predictive models to predict bus journey and arrival times based on historical AVL/GPS data and prior bus routes and stops information. There were two parts to this study. The first was to predict overall journey times and the second was to predict bus arrival times at bus stops.

To estimate total bus journey times, three models are developed using Linear Regression, Artificial Neural Network (ANN) and Long Short Term Memory Network (LSTM). Evaluation results on ground-truth dataset show that LSTM outperformed the Linear Regression model and its performance was comparable to that of ANN.

To predict bus arrival times at bus stops, three different models, namely Historical Averaging, Linear Regression and Gradient Boosting are proposed. Experimental results show that the Gradient Boosting outperformed the other models and is more robust in predicting arrival times.

Our study also reveals that it is possible to predict bus journey time with reasonable accuracy by using only GPS observations and bus routes information. This can be particularly useful in situations where data regarding other external features affecting travel time are not available.