Latency Prediction Based on Machine Learning of Real-Time Highway Data

Tsz Hong Kwok, *Yulin Xu, T. S. Choi, K. Y. Michael Wong*

Department of Physics, Hong Kong University of Science and Technology, thkwokae@connect.ust.hk

Prediction of traveling times (latencies) between different starting and ending points in the highway system is not only a matter of personal interest, but is also essential for the planning of smart cities. However, prediction uncertainty is high due to the complexity of drivers' behaviors. With the availability of a tremendous volume of traffic data collected from recently developed electronic toll collection systems, latency prediction can be achieved through machine learning methods. However, building a machine learning model without understanding the problem may result in unsatisfactory predictions and waste of computational power.

In this study, we first investigate the important features affecting latencies in the highway system, which can be used to improve the prediction. Inspecting the coefficients in linear regression, we find the most significant factors contributing to congestions, such as the influx to a certain bottleneck segment.

Second, we train different machine learning models to predict the latency. We find that their performance varies, depending on various traffic patterns. This motivates us to introduce classification methods to separate the data into different clusters for training the machine learning model. We show that a classification scheme based on the location of the system state on the fundamental diagram is relatively successful, demonstrating that pre-processing the input data with a relevant physical picture is useful.

These approaches improve the accuracy in the latency prediction and enable us to find the most significant factors contributing to traffic congestions. It also helps us in investigating the dynamics of traffic congestions.

Acknowledgements

This work is supported by research grants from the Research Grant Council of Hong Kong (grant numbers 16322616 and 16306817).