Intersection Delay Estimation from Floating Car Data via Stacked Generalization: A Case Study on Beijing's Road Networks

Xiliang Liu, Feng Lu, Hengcai Zhang

State Key Lab of Resources and Environmental Information System,

Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences,

11A, Datun Road, Chaoyang District, Beijing 100101, P. R. China

{liuxl, luf, zhanghc}@lreis.ac.cn

Abstract

Modeling the time cost on driving through the intersections among road networks can significantly help improve the reliability of car navigation systems or web map systems, as well as other real traffic related applications (Liu et al., 2013). Some researchers rely on parametric approaches including the historical mean (Sun 2007; Zhao et al. 2013), linear interpolation (Ban et al. 2009), auto regression moving average (ARMA) (Zhang and Liu 2009a), etc. These methods perform well in stationary traffic flow scenario, but may fail facing with the dynamics of the urban networks. Some researchers notice the instability of traffic flow and employ sophisticated non-parametric models such as artificial neural network (ANN) (Zhang et al. 2011), support vector machine (SVM) (Zhang and Liu 2009b) and so on. These methods show advantages in over saturated conditions, but require prior knowledge about the models and extra specific parameter tuning, which restrict the generalization of these models. To date, it seems still hard to get a satisfying result just from one model facing with the complexity of traffic phenomena due to the complexity of transport system. In view of the limitations of traditional approaches, in this paper, we propose a pervasive framework to estimate the intersection delays from floating car data (FCD) based on stacked generalization ensemble learning theory (Wolpert, 2002), trying to integrated the advantages of different models and to get rid of extra artificial interference. Firstly we analyze the optimal linear combination based on error-ambiguity decomposition and redesign the learning strategy in the level-1 phrase of stacked generalization; Secondly we integrate six classical and frequently used approaches into this framework, including linear least squares regression (LLSR), autoregressive moving average (ARMA), historical mean (HM), Artificial Neural Network (ANN), Radical Basis Function Neural Network (RBF-NN), Support Vector Machine (SVM), and

In: S. Winter and C. Rizos (Eds.): Research@Locate'14, Canberra, Australia, 07-09 April 2014, published at http://ceur-ws.org

conduct experiments with a real floating car dataset collected with more 20,000 taxies in a four-month period of 2011 in Beijing city. In order to testify the effectiveness of the proposed model, we further compare our estimation result with other four linear combination methods, namely the equal weights method (EW), optimal weights method (OW), minimum error method (ME) and minimum variance method (MV). Final results on 400 main intersections among Beijing's road networks confirm that the proposed stacked generalization approach behaves more robust and accurate than any single approach, and outperforms those classical linear combination strategies both in variance and bias. We give possible explanations based on mathematical analysis and spatial-temporal distribution of floating car data. In addition, the proposed stacked generalization approach provides a scalable, open-ending framework for the intersection delay estimation, which means any new promising models can be easily incorporated in our future work.

References

Ban, X. J., Herring, R., Hao, P., Bayen, A. M. (2009). Delay pattern estimation for signalized intersections using sampled travel times. *Transportation Research Record*, (2130), 109-119.

Liu, X., Lu, F., Zhang, H. and Qiu P. (2013). Intersection delay estimation from floating car data via principal curves: a case study on Beijing's road network. *Frontiers of Earth Science*, 7(2), 206-216.

Sun, L. (2007). An approach for intersection delay estimate based on floating vehicles. *Ph.D thesis*. Beijing: Beijing University of Technology (in Chinese).

Wolpert, D.H. (2002). Stacked generalization, *Neural Networks,* 5(2), 241-260.

Zhang, H., Lu, F., Zhou, L., Duan, Y. (2011) Computing turn delay in city road network with GPS collected trajectories. In: *Proceedings of the 2011 International Workshop on Trajectory Data Mining and Analysis*, September 17-21, Beijing, China, ACM, 45-52.

Zhang, Y., and Liu, Y. (2009a) Comparison of Parametric and Nonparametric Techniques for Non-peak Traffic Forecasting. In: *Proceedings of World Academy of Science, Engineering and Technology*, 8-14.

Zhang, Y., and Liu, Y. (2009b) Traffic forecasting using least squares support vector machines. *Transportmetrica*, 5(3):193-213.

Zhao, M., Li, X. (2013). Deriving Average Delay of Traffic Flow around Intersections from Vehicle Trajectory Data, *Frontiers of Earth Science*, 7(1), 28-33.