A trade-off analysis between penetration rate and sampling frequency of mobile sensors in traffic state estimation

Bucknell , C., Herrera, J. (2014). A trade-off analysis between penetration rate and sampling frequency of mobile sensors in traffic state estimation. *Transportation Research Part C: Emerging Technologies, 46*, 132-150.
https://doi.org/10.1016/j.trc.2014.05.007

Abstract

The rapid-growth of smartphones with embedded navigation systems such as GPS modules provides new ways of monitoring traffic. These devices can register and send a great amount of traffic related data, which can be used for traffic state estimation. In such a case, the amount of data collected depends on two variables: the penetration rate of devices in traffic flow (P) and their data sampling frequency (z). Referring to data composition as the way certain number of observations is collected, in terms of P and z, we need to understand the relation between the amount and composition of data collected, and the accuracy achieved in traffic state estimation. This was accomplished through an in-depth analysis of two datasets of vehicle trajectories on freeways. The first dataset consists of trajectories over a real freeway, while the second dataset is obtained through microsimulation. Hypothetical scenarios of data sent by equipped vehicles were created, based on the composition of data collected. Different values of P and z were used, and each unique combination defined a specific scenario. Traffic states were estimated through two simple methods, and a more advanced one that incorporates traffic flow theory. A measure to quantify data to be collected was proposed, based on travel time, number of vehicles, penetration rate and sampling frequency. The error was below 6% for every scenario in each dataset. Also, increasing data reduced variability in data count estimation. The performance of the different estimation methods varied through each dataset and scenario. Since the same number of observations can be gathered

with different combinations of P and z, the effect of data composition was analyzed (a trade-off between penetration rate and sampling frequency). Different situations were found. In some, an increase in penetration rate is more effective to reduce estimation error than an increase in sampling frequency, considering an equal increase in observations. In other areas, the opposite relationship was found. Between these areas, an indifference curve was found. In fact, this curve is the solution to the optimization problem of minimizing the error given any fixed number of observations. As a general result, increasing sampling frequency (penetration rate) is more beneficial when the current sampling frequency (penetration rate) is low, independent of the penetration rate (sampling frequency).