Accepted Manuscript

Title: Modified Bayesian data fusion model for travel time estimation considering spurious data and traffic conditions

Authors: Soknath Mil, Mongkut Piantanakulchai

Please cite this article as: Soknath M, Mongkut P, Modified Bayesian data fusion model for travel time estimation considering spurious data and traffic conditions, *Applied Soft Computing Journal* (2018), https://doi.org/10.1016/j.asoc.2018.06.046

Modified Bayesian data fusion model for travel time estimation considering spurious data and traffic conditions

Soknath Mil, Mongkut Piantanakulchai*

School of Civil Engineering and Technology, Sirindhorn International Institute of Technology, Thammasat University, P.O.Box 22, Pathum Thani 12121, Thailand

Highlights

- Data fusion and Gaussian mixture model are proposed to improve the accuracy and precision of travel time estimation at different traffic conditions on highways.
- The proposed model considering spurious data provides advantages over the basic Bayesian approach, such as being robust to noisy data, reducing biases, and producing a more precise estimation.
- The results from three case studies show significant improvement in the accuracy of travel time estimation for a range of 3.46% - 16.3% of MAPE.

Abstract

This paper presents a framework for the development of the travel time estimation model using multiple sources of data with consideration of spurious data and traffic conditions. A modified Bayesian data fusion approach, combined with the Gaussian mixture model, is used to fuse the travel time data, which are estimated from different types of sensors to improve accuracy, precision, as well as completeness of data, in terms of spatial and temporal distribution. Two additional features are added into existing models including the difference of traffic conditions classified by the Gaussian mixture model and the bias estimation from individual sensor by introducing a non-zero mean Gaussian distribution which learned from the training dataset. The methodology and computational procedure are presented. The Gaussian mixture model is used to classify states of traffic into predefined number of traffic regimes. Once a traffic condition is classified, the modified Bayesian data fusion approach is used to estimate travel time. The proposed model provides explicit advantages over the basic Bayesian approach, such as being robust to noisy data, reducing biases of an individual estimation, and producing a more precise estimation of travel time. Two different real world datasets and one simulated dataset are used to evaluate the performance of the proposed model under three different traffic regimes: free flow, transitional flow and congested flow regimes. The results when compared with the results from benchmark models show significant improvement in the accuracy of travel time estimation in terms of mean absolute percentage errors (MAPE) in the range of 3.46% to 16.3%.

Keyword: Bayesian data fusion approach, Gaussian mixture model, travel time estimation

1 Introduction

Recently, traffic congestion has become a major concern for most urbanized areas in the world. It generates various adverse effects. Perhaps one of the most important issues is its impediment to economic growth as it restrains smooth operation of transportation activities. Common practices to alleviate traffic congestion are to apply traffic policies, such as expansion of road capacity, encourage the use of public transport and use