Transportation Research Part C 17 (2009) 498–509

Contents lists available at ScienceDirect

Transportation Research Part C

journal homepage: www.elsevier.com/locate/trc

Bayesian committee of neural networks to predict travel times with confidence intervals

C.P.IJ. van Hinsbergen [a,b,*], J.W.C. van Lint [a,1], H.J. van Zuylen [a,b,1]

[a] Department of Transport and Planning, Faculty of Civil Engineering and Geosciences, Delft University of Technology, Stevinweg 1, P.O. Box 5048, 2600 GA, Delft, The Netherlands
[b] TRAIL Research School, The Netherlands

ARTICLE INFO

Article history:
Received 29 April 2008
Received in revised form 26 January 2009
Accepted 14 April 2009

Keywords:
Neural network
Bayesian inference
Committee
Travel time prediction
Confidence interval
Evidence framework

ABSTRACT

Short-term prediction of travel time is one of the central topics in current transportation research and practice. Among the more successful travel time prediction approaches are neural networks and combined prediction models (a 'committee'). However, both approaches have disadvantages. Usually many candidate neural networks are trained and the best performing one is selected. However, it is difficult and arbitrary to select the optimal network. In committee approaches a principled and mathematically sound framework to combine travel time predictions is lacking. This paper overcomes the drawbacks of both approaches by combining neural networks in a committee using Bayesian inference theory. An 'evidence' factor can be calculated for each model, which can be used as a stopping criterion during training, and as a tool to select and combine different neural networks. Along with higher prediction accuracy, this approach allows for accurate estimation of confidence intervals for the predictions. When comparing the committee predictions to single neural network predictions on the A12 motorway in the Netherlands it is concluded that the approach indeed leads to improved travel time prediction accuracy.

1. Introduction

The widely acknowledge potential of traffic information to alleviate congestion and to decrease negative environmental and societal side effects has led to a surge of research into reliable and accurate traffic and travel time prediction models in the past few decades (van Lint et al., 2005).

Among the most applied types of traffic prediction models are ARIMA-like time series approaches (Nihan, 1980; Lee and Fambro, 1999), Kalman filtering (Okutani and Stephanedes, 1984; Yang, 2005), linear weighted regression (Zhong et al., 2005; Nikovski et al., 2005), nearest neighbor techniques (Clark, 2003; Smith and Demetsky, 1996), neural networks (van Lint et al., 2005; Dougherty and Cobbett, 1997; Mark et al., 2004; Zhang, 2000; Innamaa, 2005; Dharia and Adeli, 2003) and so-called *committee* or *ensemble* approaches, in which multiple model-predictions are combined (Petridis et al., 2001; Kuchipudi and Chien, 2003; Zheng et al., 2006). The last two approaches, neural networks and committees, have shown a high accuracy for prediction of traffic conditions (van Hinsbergen et al., 2007). However, these two approaches exhibit some imperfections when applied in real-time applications, as will be shown in Sections 1.1 and 1.2.

* Corresponding author. Address: Department of Transport and Planning, Faculty of Civil Engineering and Geosciences, Delft University of Technology, Stevinweg 1, P.O. Box 5048, 2600 GA, Delft, The Netherlands. Tel.: +31 15 2784066; fax: +31 15 2783179.

E-mail addresses: c.p.i.j.vanhinsbergen@tudelft.nl (C.P.IJ. van Hinsbergen), j.w.c.vanlint@tudelft.nl (J.W.C. van Lint), h.j.vanzuylen@tudelft.nl (H.J. van Zuylen).

[1] Tel.: +31 15 2781681; fax: +31 15 278 3179.

0968-090X/$ - see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.trc.2009.04.007

One valuable and objective piece of traffic information is the *travel time*. Real-time travel time predictions can be used in dynamic traffic management applications and in commercial applications, such as pre-trip planning or en-route navigation. This paper presents a neural network-based committee approach as an alternative for online travel time prediction.

1.1. Committees of prediction models

One way of improving prediction accuracy and reliability is to combine multiple prediction models in a *committee*, where the outcomes are a weighted combination of the outcomes of its members. It has been shown (Bishop, 1995) that a committee cannot produce an increase in the expected error, even if just the average of the predictions is taken and the weights are not optimized.

Previous attempts to combine traffic prediction models typically use the errors the models make in the previous time intervals (Petridis et al., 2001; Kuchipudi and Chien, 2003; Zheng et al., 2006). However, when applied to predicting *travel time*, one major complication occurs: it takes time (in fact the travel time) for the actual trip to be realized and consequently for a travel time to become available. Therefore, in most practical situations the actual travel time is not available within one discrete time step, especially in congested situations where accurate travel time prediction is most valuable. Using the error in the previous intervals to combine travel time prediction models must thus be considered a theoretical exercise and inapplicable to most real-time applications (Van Lint, 2008).

In (van Hinsbergen and van Lint, 2008) an alternative committee approach using Bayesian inference theory was applied to the travel time prediction problem. In this theory, a model's prediction as well as the *probability* that a model predicts the travel time correctly (the *evidence* for a model) is used. The relative probabilities of the models are then used to select the committee members. This approach does not involve evaluating the prediction error of the last prediction(s) made, which makes it appropriate for online applications. In (van Hinsbergen and van Lint, 2008) it is demonstrated that prediction accuracy can be improved using this approach.

1.2. Artificial neural networks

It is common practice in the application of (artificial) neural networks for travel time prediction to train many different candidate networks and then to select the best, based on the performance on an independent validation set, to make predictions. Although this might intuitively make sense, there are a number of serious drawbacks to this approach. In the first place, this implies that much effort involved in training networks is wasted. More seriously, the fact that one neural network model outperforms all other models on one particular validation data set does not guarantee that this neural network model indeed contains the "optimal" weights and structure, nor that this model has the best generalization capabilities. This completely depends on the statistical properties of the training and validation set (e.g. the amount of noise in the data), the complexity of the problem at hand and most importantly on the degree to which the training and validation set are representative for the true underlying process which is modeled. The network performing best on the validation set may therefore not be the one with the best performance on new data (Bishop, 1995).

These drawbacks can be overcome by combining all (or a representative selection of) trained neural network models in a committee. The Bayesian framework that is applied in (van Hinsbergen and van Lint, 2008) can be used for this purpose. The theory of Bayesian inference to train and combine a committee of feed-forward neural networks has been described in (Bishop, 1995; MacKay, 1992b, 1995) and has been applied in various fields of study (MacKay, 1994; Thodberg, 1993; Chua and Goh, 2003; Lisboa et al., 2003; Baesens et al., 2002). To the authors' knowledge this approach has not yet been applied to travel time prediction or traffic prediction in general.

1.3. Objective of this study

In this study the abovementioned Bayesian approach for neural network based travel time prediction will be used and its workings will be demonstrated on real data from the A12 motorway in The Netherlands. In this approach two *intrinsic* and informative quantities are calculated, which allow for real time model comparison and combination, First, during training, the so-called model-*evidence* is calculated, which ranks the models on the basis of the fit on the training data taking into account the degree of over-fitting (inducing variance) or under-fitting (resulting in bias). Second, in actual operation the approach also allows the analyst to estimate errors (error bars) on each prediction, which indicate the degree in which the currently presented input pattern matches with the input patterns "seen" during training. The committee approach is compared to individual neural networks to show that the committee provides a more accurate prediction of travel times and has better generalization performance.

As traffic systems are highly dynamic, it is expected that in order to make highly accurate travel time predictions, neural networks that are able incorporate these dynamics, such as feed-forward neural networks with multiple layers, recurrent neural networks or state-space neural networks (van Lint et al., 2005), are needed. However, to maintain focus on the workings and powerful properties of the Bayesian framework, relatively simple feed-forward neural networks are used in this study; the principles applied in this study can also be applied to more complex neural networks structures in future studies.