Transportation Research Part C 43 (2014) 95–111

Contents lists available at ScienceDirect

Transportation Research Part C

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

A Hidden Markov Model for short term prediction of traffic conditions on freeways

Yan Qi [a,*], Sherif Ishak [b]

[a] Department of General Engineering, Montana Tech of the University of Montana, 1300 W. Park St., Butte, MT 59701, USA
[b] Department of Civil and Environmental Engineering, Louisiana State University, Baton Rouge, LA 70803, USA

ARTICLE INFO

Article history:
Received 11 October 2013
Received in revised form 3 February 2014
Accepted 13 February 2014

Keywords:
Short-term traffic prediction
Freeway traffic
Hidden Markov Models
Stochastic modeling
Transition probability
Clustering analysis
Second-order statistics

ABSTRACT

Accurate short-term prediction of traffic conditions on freeways and major arterials has recently become increasingly important because of its vital role in the basic traffic management functions and trip decision making processes. Given the dynamic and stochastic nature of freeway traffic, this study proposes a stochastic approach, Hidden Markov Model (HMM), for short-term freeway traffic prediction during peak periods. The data used in the study was gathered from real-time traffic monitoring devices over six years on a 60.8-km (38-mile) corridor of Interstate-4 in Orlando, Florida. The HMM defines traffic states in a two-dimensional space using first-order statistics (Mean) and second-order statistics (Contrast) of speed observations. The dynamic changes of freeway traffic conditions are addressed with state transition probabilities. For a sequence of traffic speed observations, HMMs estimate the most likely sequence of traffic states. The model performance was evaluated using prediction errors, which are measured by the relative length of the distance between the predicted state and the observed state in the two-dimensional space. Reasonable prediction errors lower than or around 10% were obtained from HMMs. Also, the model performance was not remarkably affected by location, travel direction, and peak period time. The HMMs were compared to two naïve predication methods. The results showed that HMMs perform better and are more robust than the naïve methods. Therefore, the study concludes that the HMM approach was successful in modeling short-term traffic condition prediction during peak periods and in accounting for the inherent stochastic nature of traffic conditions.

Published by Elsevier Ltd.

1. Introduction

As a result of the increased travel demands, traffic congestion has become a major problem for urban freeways, especially during peak periods. Capacity expansion projects cannot solely address the problem due to land and cost constraints, as well as the hidden demand generated from capacity improvements. Therefore, effective utilization of existing freeway facilities and transportation network resources offers a viable alternative for improving the level of service, which is the primary goal of traffic management agencies. With the advancement in communication technologies, applications in Intelligent Transportation Systems (ITS) continue to offer innovative solutions for many traffic problems through dissemination of real time information on traffic conditions to travelers. Short-term traffic prediction, which is the focus of this research, is one example

* Corresponding author. Tel.: +1 406 496 4449.
E-mail addresses: yqi@mtech.edu (Y. Qi), sishak@lsu.edu (S. Ishak).

http://dx.doi.org/10.1016/j.trc.2014.02.007
0968-090X/Published by Elsevier Ltd.

of several user services that identify with Advanced Traveler Information Systems (ATIS) and Advanced Traffic Management Systems (ATMS).

Using traffic monitoring data as input to mathematical models, short-term traffic prediction algorithms can be developed to predict traffic conditions over a relatively short period of time, usually ranging from a few minutes to 30 min into the future. An accurate short-term prediction of traffic parameters can help improve the transportation system functions with respect to real-time control strategies, advance warning in monitoring systems, as well as reduction of congestion, delay, and energy consumption. In the past few years, more emphasis has been placed on active traffic management strategies such as traffic prediction. Short term predictive information is essential to both transportation system users and providers for better decision making that could improve the productivity of the transportation system and reduce the direct and indirect cost of both the system users and providers. However, given the dynamic nature of traffic flow, such task has been challenging to transportation practitioners and researchers. To tackle this issue, researchers continue to seek more robust and accurate prediction algorithms. This study explores a probabilistic approach for freeway traffic condition prediction during peak periods based on the observed stochastic variations in freeway traffic conditions during those periods.

2. Background

Since the 1970s, many researchers have explored the area of short-term traffic prediction and a considerable number of publications can be found in the open literature. Numerous tools and algorithms have been applied, most of which focused on developing new or improving existing prediction models. For instance, the Kalman Filtering (KF) method was first used for traffic volume prediction by Okutani and Stephanedes (1984). Recently, in a study by Yang et al. (2004), a Recursive Least Square (RLS) approach was proposed for short-term traffic speed prediction by means of KF to adapt to changing patterns quickly, based on the maximum likelihood method and Bayesian rule. Their model was evaluated using 5-min real-time aggregated loop detector data collected from I-405, California. The results showed that after the adaptive learning to update the model parameters when new observations were available, most of the true speed values fell in the 95% confidence interval in the on-line process.

Xie et al. (2006) conducted a study to combine the Wavelet decomposition with the KF method for short-term traffic speed prediction. A wavelet KF model with various data decomposition levels and an original KF model were compared using 5-min aggregate data collected on weekdays. The study showed that the wavelet KF model consistently outperformed the KF model in terms of both accuracy and stability and that a higher data decomposition level was more advantageous for non-stationary data prediction. In another study by Xia and Chen (2009), a dynamic short-term corridor travel time prediction model was developed using the KF method, which involves a multi-step-ahead prediction of traffic condition with a seasonal autoregressive integrated moving average model. In conjunction with KF, Wang and Papageorgiou (2008) proposed an approach for real-time adaptive estimation of traffic flow variables based on stochastic macroscopic traffic flow modeling. Recently, Guo and Williams (2010) proposed an autoregressive moving average plus generalized autoregressive conditional heteroscedasticity structure for modeling the station-by-station traffic speed series to forecast the short-term traffic condition level and uncertainty. They employed an online algorithm based on layered KF for processing the heteroscedasticity structure in real time.

Another group of short-term traffic prediction models are based on Time Series models, which were used for short term traffic forecasting in late 1970s and became popular in the late 1990s. Various Time Series models and algorithms have been reported in the literature to predict different types of traffic parameters. For instance, Hamed et al. (1995) developed a time-series model to predict future traffic volumes on urban arterials using the Box–Jenkins approach. Lee and Fambro (1999) applied the subset Autoregressive Integrated Moving Average (ARIMA) model for short-term freeway traffic volume prediction. Similarly, Williams and Hoel (2003) modeled univariate traffic condition data streams as seasonal Autoregressive Integrated Moving Average processes. In other another study, Smith et al. (2002) compared parametric modeling approach of ARIMA and non-parametric regression models for short term traffic flow forecasting. It is concluded that heuristic forecast generation methods did significantly improve the performance of nonparametric regression, but they did not equal the performance of seasonal ARIMA models and traffic condition data is characteristically stochastic rather than chaotic. More recently, Shekhar and Williams (2007) proposed an adaptive seasonal models for univariate traffic flow forecasting through the use of three well-known filtering techniques: the Kalman filter, recursive least squares, and least mean squares.

Time Series models were also used for short-term traffic speed prediction. Using two algorithms, Expectation Maximization (EM) and the Cumulative Sum (CUSUM) algorithms, into ARIMA Time Series model, Cetin and Comert (2006) proposed an adaptive approach for traffic speed prediction. The proposed approaches were tested on a public available loop dataset collected by the California PATH with detailed records of all incidents. The study showed that compared to the ARIMA model, the two adaptive techniques provide more accurate results when the data generation process is not stable. Chandra and Al-Deek (2009) developed a vector auto regressive Time Series model to predict traffic speed and volume of a 4-km (2.5-mile) segment of I-4 Orlando, Florida. The effect of upstream and downstream location information was incorporated in the method based on cross-correlation analysis of the data between different locations. Farokhi et al. (2010) evaluated the performance of three moving average techniques (one simple moving average method with constant weights and two adaptive moving average methods) in predicting average travel speeds up to 10 min ahead of time. It was found that the method using optimized weights produced slightly better predictions at a higher computational cost.