Future Generation Computer Systems 89 (2018) 78–88

Contents lists available at ScienceDirect

Future Generation Computer Systems

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

A novel fuzzy deep-learning approach to traffic flow prediction with uncertain spatial–temporal data features

Weihong Chen [a,b], Jiyao An [a,*], Renfa Li [a], Li Fu [a], Guoqi Xie [a], Md Zakirul Alam Bhuiyan [c], Keqin Li [a,d]

[a] College of Computer Science and Electronic Engineering, Hunan University, China
[b] College of Information and Electronic Engineering, Hunan City University, China
[c] Department of Computer and Information Sciences, Fordham University, USA
[d] Department of Computer Science, State University of New York, New Paltz, NY, 12651, USA

HIGHLIGHTS

- The FDCN approach is proposed to lessen the impact of data uncertainty.
- A model of the fuzzy deep convolution network is designed.
- The tensor is employed to investigate temporal and spatial properties of traffic flow.
- Experiments verify the performance of the FDCN in convergence and accuracy.

ARTICLE INFO

Article history:
Received 30 March 2018
Received in revised form 16 May 2018
Accepted 10 June 2018
Available online 21 June 2018

Keywords:
Deep learning
Fuzzy representation
Residual networks
Traffic flow prediction

ABSTRACT

Predicting traffic flow is one of the fundamental needs to comfortable travel, but this task is challenging in vehicular cyber–physical systems because of ever-increasing uncertain traffic big data. Although deep-learning (DL) methods with outstanding performance recently have become popular, most existing DL models for traffic flow prediction are fully deterministic and shed no light on data uncertainty. In this study, a novel fuzzy deep-learning approach called FDCN is proposed for predicting citywide traffic flow. This approach is built on the fuzzy theory and the deep residual network model. Our key idea is to introduce the fuzzy representation into the DL model to lessen the impact of data uncertainty. A model of fuzzy deep convolutional network is established to improve traffic flow prediction while investigating the spatial and temporal correlation of traffic flow. We further propose pre-training and fine-tuning strategies that efficiently learn parameters of the FDCN. To the best of our knowledge, this is the first time that a fuzzy DL approach has been applied to represent traffic features for traffic flow prediction. Experimental results demonstrate that the proposed approach to traffic flow prediction has superior performance compared with state-of-the-art approaches.

1. Introduction

1.1. Background

In modern society, the number of vehicles has been increasing in cities and on freeways. Many problems related to this increase, such as traffic congestion, may lead to longer time wasted in travel and thus may result in money loss as well as traffic accidents [1,2]. Obtaining accurate and timely traffic flow information is necessary for individual travelers. With the current explosion of traffic flow data, predicting traffic flow using big data is crucial to ensuring safe travel and designing superefficient navigation, which may help travelers make informed travel decisions and improve public safety [3].

Traffic flow prediction on a large scale heavily depends on historical traffic data and other relevant information (i.e., weather conditions and traffic accidents) and is regarded as a key function component in the vehicular cyber–physical system (VCPS) [4]. The VCPS is a complex system with seamless integration of computation, communication and control technology, and advances in the VCPS will enable comfortability, safety and security [5,6]. Deep learning (DL), which is a new method of machine learning, learns useful features by building a multilayer model to achieve accurate

* Corresponding author.
E-mail addresses: whchen@hnu.edu.cn (W. Chen), jt_anbob@hnu.edu.cn (J. An), lirenfa@hnu.edu.cn (R. Li), fu_li@hnu.edu.cn (L. Fu), xgqman@hnu.edu.cn (G. Xie), mbhuiyan3@fordham.edu (M.Z.A. Bhuiyan), lik@newpaltz.edu (K. Li).

https://doi.org/10.1016/j.future.2018.06.021

image classification or object identification. This method can transcend conceptual learning and has ability to learn more complex knowledge [7]. DL has been applied successfully in prediction tasks, natural language processing, object detection, and motion modeling. As traffic flow prediction is of complexity in nature, deep learning algorithms can be used to represent traffic features without prior knowledge, which exhibits good performance for traffic flow prediction [8].

1.2. Motivations

As a result of traffic data containing significant noise and unpredictable uncertainty, the concept of data ambiguity emerges [9]. Such ambiguity imposes a great challenge on the ability to understand and predict traffic flow. First, the ability to represent input data is limited as variables interact in uncertain ways. Second, the convolutional neural network is not always robust when training data are disturbed by noise. To overcome these disadvantages and improve the accuracy of prediction, fuzzy logic has been introduced for solving many practical problems, including motor control, traffic prediction, and image recognition and classification [10]. Compared with conventional deterministic representations, fuzzy logic representation flexibly constructs rules for reducing the uncertainties in original data and demonstrates competitive performances in both data representation and robustness for dealing with noise.

The membership function and rule base are crucial to the design of the fuzzy logic system. The most straightforward approach is to define membership functions and rules in advance based on research experience and adjust them according to test results. Another way to design the fuzzy logic is to make the system learn from experience. Most of the existing methods are based on the experience of system research to directly define the membership function and then test the output performance. If the test results are not satisfactory, the membership function and rules should be adjusted. This approach heavily depends on human experience and requires human intervention. For pattern classification, the algorithm in [11] first inputs the original data to the hidden layer space in the DL model and then fuzzifies the deep representation at the output layer of the DL model. DL exploits the brain's cognitive mechanism of hierarchical processing of information from layer to layer, which is a breakthrough in machine learning. In practice, task-driven feature learning allows knowledge to be propagated sequentially from the lower layers to the upper layers so that an intelligent way of automatically discovering informative features from data is provided. Another exploration of the fuzzy logic system is to design the membership parameters and fuzzy rules by learning from the actual data [12,13].

Large-scale traffic flow data is complex, which is mainly affected by three complex factors: spatial dependencies, temporal dependencies, and external factors. Given the uncertainty of these factors, traffic flow prediction becomes quite challenging. Thus, various methods have been proposed for traffic flow prediction [12,14]. From a data representation perspective, no-fuzzy and fuzzy methods are used. From a model structure perspective, shallow and deep methods are proposed. Zhang et al. designed an end-to-end structure of ST-ResNet based on the unique properties of spatial–temporal data [15]. This structure employs the framework of the residual networks to model the temporal closeness, period, and trend properties of traffic flow. The proposed method integrates the residual network and convolutional network with traffic flow and external factors to predict traffic flow in each region. However, the fuzzy method can be introduced to DL, which can further reduce the effect of data uncertainty on system performance.

1.3. Our contributions

In this study, we focus on traffic flow prediction of spatial–temporal data with uncertainty to improve the accuracy of prediction by integrating fuzzy logic and deep learning. The FDCN approach is proposed for traffic flow prediction. The proposed approach constructs the FDCN model, which integrates DL with fuzzy representation to alleviate the limitations of shallow methods in traffic flow prediction. Then, the FDCN model is trained in a layered manner to learn general features of traffic flow. In this model, the fuzzy rules are adaptively learned, and the spatial as well as temporal correlations of traffic flow are inherently considered. In addition, the proposed approach demonstrates good performance for traffic flow prediction. The contributions of this study are as follows:

(1) A fuzzy deep-learning approach called FDCN is proposed for traffic flow prediction. The FDCN approach integrates the fuzzy theory with the deep residual network and the fuzzy rules are generated adaptively using the learning algorithm. It explores the fuzzy as well as deep representations to construct features of traffic flow while relatively solving the problem of uncertainty.

(2) A model of the fuzzy deep convolution network is designed for traffic flow prediction, which employs tensor data representation to investigate temporal and spatial properties of traffic flow. The optimized structure of the FDCN is obtained by exploring the number of layers in the model and the regression functions. The FDCN exhibits powerful prediction capability.

(3) The proposed approach is evaluated on the real Beijing taxicab trajectory data set. The results validate the performance of the proposed method compared with state-of-the-art methods.

The rest of this paper is organized as follows. Section 2 reviews related work. Section 3 presents preliminaries. Section 4 proposes an FDCN model and its learning algorithm. Section 5 discusses experimental results. Section 6 concludes this study and provides direction for future research.

2. Related work

An efficient algorithm to predict traffic flow is crucial for travelers when they are planning their travel. As early as the 1970s, the autoregressive integrated moving average (ARIMA) model was designed for short-term freeway traffic-flow prediction [16]. Since then, many studies have investigated traffic flow prediction. These approaches can be classified into three categories: parametric, nonparametric, and hybrid approaches [17]. The parameter techniques are ARIMA-based models and Kalman-filtering models, which are based on time-series approaches [18]. Kong et al. used one parameter (i.e., velocity) to efficiently predict traffic state, in which information is collected by the global positioning system (GPS) and the curve-fitting and vehicle-tracking mechanism is used [19]. Because of the complexity of the traffic network, however, traffic flow shows a stochastic and nonlinear quality. Time-series approaches tend to be inefficient in predicting traffic flow as well.

Research has paid much attention to nonparametric approaches from the perspective of the traffic network. The widely used nonparametric approaches include k-nearest neighbor (k-NN) methods, the Bayesian network approach, the online learning weighted support vector regression (SVR), and artificial neural networks (ANNs) [20–23]. Chang et al. presented a dynamic multi-interval traffic volume model, which is based on the k-NN non-parametric regression (KNN-NPR) [20]. Kumar et al. proposed a Bayesian network approach for traffic flow forecasting [21]. Jeong et al. presented an online learning weighted support-vector regression method for short-term traffic flow predictions [22]. Kumar et al. applied the artificial neural network (ANN) model for short term