Research Article

Blink Number Forecasting Based on Improved Bayesian Fusion Algorithm for Fatigue Driving Detection

Wei Sun,^{1,2} Xiaorui Zhang,³ Jian Wang,⁴ Jun He,⁵ and Srinivas Peeta⁴

¹*School of Information and Control, Nanjing University of Information Science & Technology, Nanjing 210044, China*

²*Jiangsu Key Laboratory of Big Data Analysis Technology, Nanjing 210044, China*

³*School of Computer and Software, Nanjing University of Information Science & Technology, Nanjing 210044, China*

⁴*The NEXTRANS Center, Purdue University, West Lafayette, IN 47907, USA*

⁵*School of Electronic and Information Engineering, Nanjing University of Information Science & Technology, Nanjing 210044, China*

Correspondence should be addressed to Wei Sun; sunw0125@163.com

Received 9 January 2015; Accepted 1 May 2015

Academic Editor: Yakov Strelniker

Copyright © 2015 Wei Sun et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

An improved Bayesian fusion algorithm (BFA) is proposed for forecasting the blink number in a continuous video. It assumes that, at one prediction interval, the blink number is correlated with the blink numbers of only a few previous intervals. With this assumption, the weights of the component predictors in the improved BFA are calculated according to their prediction performance only from a few intervals rather than from all intervals. Therefore, compared with the conventional BFA, the improved BFA is more sensitive to the disturbed condition of the component predictors for adjusting their weights more rapidly. To determine the most relevant intervals, the grey relation entropy-based analysis (GREBA) method is proposed, which can be used to analyze the relevancy between the historical data flows of blink number and the data flow at the current interval. Three single predictors, that is, the autoregressive integrated moving average (ARIMA), radial basis function neural network (RBFNN), and Kalman filter (KF), are designed and incorporated linearly into the BFA. Experimental results demonstrate that the improved BFA obviously outperforms the conventional BFA in both accuracy and stability; also fatigue driving can be accurately warned against in advance based on the blink number forecasted by the improved BFA.

1. Introduction

Fatigue driving is one of the major causes of serious accidents in transportation. Statistics show that traffic accidents caused by driver fatigue account for about 20% of the total number of accidents and more than 40% of serious traffic accidents [1, 2]. Experts agree that the actual contributions of fatigue driving to road accidents may be much higher [3]. Frequent road traffic accidents, serious casualties, and property losses caused by fatigue driving of drivers bring a heavy burden to the society and families. Thus, the accurate and efficient detection of driver fatigue is very important [4].

Many methods have been developed for detecting the driver's fatigue recently, including the measurements of physiological features detection like EEG [5, 6], heart rate variability (HRV) [7, 8], and electrooculogram (EOG) [9],

facial features detection such as blink [10, 11] and yawning [12, 13], and vehicle behaviors detection, for instance, lane deviation [14, 15] and steering angle movement [16, 17]. Among those methods, the physiological features based methods are intrusive for drivers, because these methods need to place detection electrodes on the skin, drivers will feel uncomfortable if these electrodes are in touch with the skin for a long time. Vehicle behaviors based methods are easily affected by different sizes and shapes of vehicles, as well as different driving habits of drivers [18]. Facial characteristics based methods especially blink based detection methods using machine vision are attracting more and more researchers [18–20]. The objective of the study is to propose a new forecasting algorithm of blinking number and further verify the effectiveness of the proposed algorithm according to the forecasting results.

For the fatigue driving detection methods based on eye blinking, almost all researchers pay more attention to real time detection algorithms that can fast detect opening or closing state of the eyes and correctly make recognition of whether a driver is fatigued currently. However, the detection methods usually are not able to meet the requirement of real time processing. It is computationally costly in detecting and recognizing fatigue characters, especially for fusion detection methods based on multiple fatigue features [21] which require detecting and recognizing many more characteristic parameters [22]. Moreover, statistics show that if drivers' responses are only half a second faster when traffic accidents occurred, sixty percent of traffic accidents can be avoided. Therefore, if fatigue driving can be warned against in advance before traffic accidents happen using a certain detection instrument or forecasting method, mostly traffic accidents can be avoided [22, 23]. Furthermore, physiological researches show that fatigue is a gradual process where drivers experience from consciousness to drowsiness [24]. Therefore, if we develop a predictor that can forecast the blink number of the driver in a given time interval according to previous time intervals, then we can recognize the fatigue driving in advance and also save more time for fatigue feature extraction and detection algorithm enhancement. In a word, the advantage of using predictor method is that we can know and estimate the future driving state in advance according to the past and current states as long as we set appropriate time intervals. Meanwhile, we do not need to detect the driving state of drivers in real time, which can save more computation time for other more important processing tasks.

Currently, many successful applications of single predictor based forecasting algorithms have been reported from various fields, for instance, in the fault diagnosis [25, 26], transportation flow forecasting [27], time-series prediction [28], and so on. Based on this, a variety of methods has been put forward for goal state prediction, including the Kalman filtering (KF) model [29], nonparametric regression model [30], and autoregressive integrated moving average (ARIMA) model [31]. Generally, these prediction methods can be categorized as statistical time series analysis methods, which conduct their predictions based on historical data analysis. One benefit of statistical time series analysis methods is that they can make very good predictions when the goal state varies temporally. However, due to linear properties, these methods are inadequate for capturing the rapid variations of goal state. To overcome this problem, numerous studies have used machine learning methods such as artificial neural networks (ANN) [32, 33] and support vector machines (SVM) [34] as alternative predictors. A machine learning method is able to approximate goal state in any degree of complexity without prior knowledge of problem solving. In addition, because of its ability to learn from data, it can capture the underlying relationships of goal state even when they are not apparent.

Though the previously mentioned prediction approaches are powerful and useful methods for goal state prediction and can generate accurate results for certain patterns, they each have their own drawbacks in dealing with predictions. None of them can maintain excellent prediction performance

under all application conditions [35, 36]. Generally, driver's fatigue is related to not only the duration time of continuous driving but also the current time period. For example, drivers feel fatigued more likely at 3:00–5:00 and 14:00–16:00 than at other time periods, and drivers usually do not feel fatigued until after three hours of continuous driving. Additionally, the fluctuation of blink number of drivers from nonfatigue to fatigue is very obvious and abrupt when detected by a device or instrument. Therefore, forecasting algorithm based on a single predictor is hardly suitable for the abrupt fluctuation of blink number when drivers continuously drive for a long time. This result occurs because goal state generally exhibits a spatiotemporal behavior characterized by irregular randomness, and it is very difficult for a single prediction method to capture such a disturbed pattern.

In view of these deficiencies, some researchers have turned to multivariate modeling in which models are developed by combining multiple methods to take advantage of the merits of each algorithm. A feasible fusion method that could effectively combine the predictions of single predictors is the Bayesian fusion algorithm (BFA) proposed by Petridis et al. [37]. The BFA generates a prediction by a weighted fusion of forecasts of its all component predictors based on posterior probabilities and the Bayesian rule. However, the BFA pays no attention to the relevance between the historical data flows and the current data flow. It assumes that, at a particular interval, the weights of the component predictors in the BFA depend on the cumulative prediction performance of all past intervals. This may result in making the prediction quite impervious to the greatly fluctuated prediction accuracy of the component predictors. To overcome this problem, an improved BFA has been proposed in this paper. The underlying assumption is that the data flow of blink number at a particular prediction interval is only affected by the data flows from only few of the previous time intervals, which have a comparatively higher relevancy with it. Based on this assumption, the prediction errors of only a few intervals of the component predictors are required to be considered when calculating their weights.

Simulation experiments of fatigue driving find that the improved BFA is more sensitive to the fluctuation of the component predictors and can adjust their weights more rapidly compared to the conventional BFA and also can obtain better forecasting accuracy and stability. Therefore, it can be used to judge in advance whether drivers are fatigued by the blink number forecasted in next time period.

2. Shortcoming of Conventional BFA

The conventional Bayesian fusion algorithm was originally proposed by Petridis et al. [37]. Its general idea is summarized as follows.

Let y_t denote the actual blink number at time interval t . Then, we have

$$y_t = y_t^n(y_{t-1}, y_{t-2}, \dots, y_1) + e_t^n, \quad (1)$$

where y_t^n is the n th predictor, $y_t^n(y_{t-1}, y_{t-2}, \dots, y_1)$ is the predicted blink number at time interval t by the n th predictor, and e_t^n is the corresponding prediction error.

For a certain time interval, we can use Z to denote the uncertainty. The posterior probability of the n th predictor to be the best model at time interval t is defined as

$$p_t^n = \text{prob}(Z = n / y_t, y_{t-1}, \dots, y_1). \quad (2)$$

Then, according to Bayesian rule, we can obtain

$$p_t^n = \frac{\text{prob}(y_t, Z = n / y_{t-1}, \dots, y_1)}{\sum_{m=1}^N \text{prob}(y_t, Z = m / y_{t-1}, \dots, y_1)}, \quad (3)$$

where N is the total number of component predictors; notice that

$$\begin{aligned} & \text{prob}(y_t, Z = n / y_{t-1}, \dots, y_1) \\ &= \frac{\text{prob}(y_t, Z = n, y_{t-1}, \dots, y_1)}{\text{prob}(y_{t-1}, \dots, y_1)} \\ &= \frac{\text{prob}(y_t, Z = n, y_{t-1}, \dots, y_1)}{\text{prob}(y_{t-1}, \dots, y_1, Z = n)} \\ &\quad \times \frac{\text{prob}(y_{t-1}, \dots, y_1, Z = n)}{\sum_{m=1}^N \text{prob}(y_{t-1}, Z = m / y_{t-1}, \dots, y_1)} \\ &= \text{prob}(y_t / y_{t-1}, \dots, y_1, Z = n) p_{t-1}^n. \end{aligned} \quad (4)$$

Assuming that the prediction error $e_t^n = y_t - y_t^n$ is a Gaussian white noise time series with zero mean and standard deviation σ_n , we have

$$\begin{aligned} & \text{prob}(y_t / y_{t-1}, \dots, y_1, Z = n) \\ &= \text{prob}(e_t^n = y_t - y_t^n / y_{t-1}, \dots, y_1, Z = n) \\ &= \frac{1}{\sqrt{2\pi}\sigma_n} \exp\left(-\left[\frac{(y_t - y_t^n)}{\sqrt{2}\sigma_n}\right]^2\right). \end{aligned} \quad (5)$$

Combining (3), (4), and (5) yields

$$\begin{aligned} & p_t^n \\ &= \frac{(1/\sqrt{2\pi}\sigma_n) p_{t-1}^n \exp\left(-\left[(y_t - y_t^n)/\sqrt{2}\sigma_n\right]^2\right)}{\sum_{m=1}^N (1/\sqrt{2\pi}\sigma_m) p_{t-1}^m \exp\left(-\left[(y_t - y_t^m)/\sqrt{2}\sigma_m\right]^2\right)}. \end{aligned} \quad (6)$$

Then, at time interval $t + 1$, the prediction of the conventional BFA can be written as the linear combination of the output of all the predictors as follows:

$$\hat{y}_{t+1} = \sum_{n=1}^N p_t^n y_{t+1}^n, \quad (7)$$

where $\hat{y}_{t+1}$ is the prediction generated by the conventional BFA and y_{t+1}^n is the result estimated by the n th predictor.

Similar to (6), $p_{t-1}^n, n = (1, 2, \dots, N)$ is formulated as

$$\begin{aligned} & p_{t-1}^n \\ &= \frac{(1/\sqrt{2\pi}\sigma_n) p_{t-2}^n \exp\left(-\left[(y_{t-1} - y_{t-1}^n)/\sqrt{2}\sigma_n\right]^2\right)}{\sum_{m=1}^N (1/\sqrt{2\pi}\sigma_m) p_{t-2}^m \exp\left(-\left[(y_{t-1} - y_{t-1}^m)/\sqrt{2}\sigma_m\right]^2\right)}. \end{aligned} \quad (8)$$

Combine (6) and (8) and we have

$$\begin{aligned} & p_t^n \\ &= \frac{(1/\sqrt{2\pi}\sigma_n)^2 p_{t-2}^n \exp\left(\sum_{i=0}^1 -\left[(y_{t-i} - y_{t-i}^n)/\sqrt{2}\sigma_n\right]^2\right)}{\sum_{m=1}^N (1/\sqrt{2\pi}\sigma_m)^2 p_{t-2}^m \exp\left(\sum_{i=0}^1 -\left[(y_{t-i} - y_{t-i}^m)/\sqrt{2}\sigma_m\right]^2\right)}. \end{aligned} \quad (9)$$

Similarly, we can formulate explicit expressions for $p_{t-2}^n, p_{t-3}^n, \dots, p_1^n$. Substituting them into (9), we have

$$\begin{aligned} & p_t^n \\ &= \frac{(1/\sqrt{2\pi}\sigma_n)^{t-1} \exp\left(\sum_{i=0}^{t-1} -\left[(y_{t-i} - y_{t-i}^n)/\sqrt{2}\sigma_n\right]^2\right)}{\sum_{m=1}^N (1/\sqrt{2\pi}\sigma_m)^{t-1} \exp\left(\sum_{i=0}^{t-1} -\left[(y_{t-i} - y_{t-i}^m)/\sqrt{2}\sigma_m\right]^2\right)}. \end{aligned} \quad (10)$$

From (10), we can see that, rather than only depending on the prediction error at prediction interval t , the weight p_t^n depends on the prediction errors of all past intervals. This characteristic makes the conventional BFA very inert to the fluctuating accuracy of component predictors. If the dominant predictor is no longer the most accurate, it will take many intervals to reduce the dominant status of that predictor, thus imposing a negative impact on the predictions of the conventional BFA.

This problem arises because the conventional BFA does not consider the correlation between the historical data flows and the data flow at the prediction interval. Generally, only the data flows of blink number at the latest few intervals may strongly correlate with the data flow at a given prediction interval. For the data flows of blink number at earlier time intervals, they may have less impact on the current data flow.

3. Improved Bayesian Fusion Algorithm

3.1. Selection of Correlative Data Flow. For most of the forecasting methods, the selected blink data flows from only a few intervals are considered. Inspired by this idea, this paper hypothesizes that, at a particular prediction interval, the data flow of blink number is only affected by the past few intervals' data flows which have comparatively higher relevance with it. With this assumption, the weight of each component predictor in the BFA at a given prediction interval would just depend on the accumulative prediction performance of a few selected intervals rather than all intervals. Based on this assumption, (1) can be rewritten as

$$y_t = y_t^n(y_k) + e_t^n; \quad \forall k \in K, \quad (11)$$

where K represents the set of the previous intervals at which the data flows of blink number have a comparatively higher correlation with the data flow at prediction interval t .

Following the same inferential procedures from (2) to (6) and from (9) to (10), we have

$$\begin{aligned} & p_t^n \\ &= \frac{(1/\sqrt{2\pi}\sigma_n)^{R(K)} \exp\left(\sum_{k \in K} -\left[(y_k - y_k^n)/\sqrt{2}\sigma_n\right]^2\right)}{\sum_{m=1}^N (1/\sqrt{2\pi}\sigma_m)^{R(K)} \exp\left(\sum_{k \in K} -\left[(y_k - y_k^m)/\sqrt{2}\sigma_m\right]^2\right)}, \end{aligned} \quad (12)$$

where $R(K)$ represents the dimension of set K . From (12), we can see that the weights of the component predictors just depend on their prediction errors at the previous $R(K)$ intervals. Therefore, theoretically, the weights calculated by (12) are more sensitive to the fluctuated accuracy of component predictors, and the number of data flows to be used to calculate is also reduced greatly. Substituting it into (7), we can obtain the prediction of the improved BFA.

3.2. Grey Relation Entropy-Based Analysis. According to the discussion in the previous section, one challenge to implementing the improved BFA is how to identify the time interval set K where the data flows of blink number have comparatively higher correlation with the data flows at the prediction interval. Because (12) includes a nonlinear, complex exponential function, the regression method is not applicable here. An alternative method is the grey relation entropy-based analysis (GREBA) [38, 39]. It has been extensively applied for relevance analysis in various disciplines, such as logistics [40], economics [41], and engineering [42]. In this paper, we use this method for relevance analysis between current and historical data flows of blink number.

Assuming that, at the prediction interval t , blink number y_t is affected by its previous w intervals' data flows, listed as $\mathbf{Y}(t) = (y_{t-1}, y_{t-2}, \dots, y_{t-w+1})$, w should be large enough to cover most of the important data flows. We use $\mathbf{Y}_t(j) = (y_{t-1}, y_{t-2}, y_{t-j}, \dots, y_{t-N})$ and $\mathbf{Y}_{t-i}(j) = (y_{t-i-1}, y_{t-i-2}, y_{t-i-j}, \dots, y_{t-i-N})$ express the target data flow time sequence at the t th interval and the alternative data flow time sequence at the $(t-i)$ th interval, respectively; $i = \{1, 2, \dots, M\}$, M is the amount of the alternatives data flow time sequence; $j = \{1, 2, \dots, N\}$, N is the length of the target and alternative data flow time sequences; $\mathbf{Y}_t(j) \subset \mathbf{Y}(t)$ and $\mathbf{Y}_{t-i}(j) \subset \mathbf{Y}(t)$. Here, the sequences of $\mathbf{Y}_t(j)$ and $\mathbf{Y}_{t-i}(j)$ are comparable because they are obtained from the same data sequence $\mathbf{Y}(t)$; therefore, it is unnecessary to normalize the original data. The grey relation coefficient between $\mathbf{Y}_t(j)$ and $\mathbf{Y}_{t-i}(j)$ is then calculated as

$$\begin{aligned} & \psi(\mathbf{Y}_t(j), \mathbf{Y}_{t-i}(j)) \\ &= \frac{\min_i \min_j |\mathbf{Y}_t(j) - \mathbf{Y}_{t-i}(j)| + \rho \max_i \max_j |\mathbf{Y}_t(j) - \mathbf{Y}_{t-i}(j)|}{|\mathbf{Y}_t(j) - \mathbf{Y}_{t-i}(j)| + \rho \max_i \max_j |\mathbf{Y}_t(j) - \mathbf{Y}_{t-i}(j)|}, \end{aligned} \quad (13)$$

where ρ is the distinguishing coefficient in the range of $[0, 1]$ and it can be adjusted to make a better distinction between the target sequence and the alternative sequence.

To satisfy the rule of grey entropy, the grey relation coefficients should be transformed into a grey relation density $\gamma(t, t-i)$, calculated by

$$\begin{aligned} \gamma(t, t-i) &= \frac{\psi(\mathbf{Y}_t(j), \mathbf{Y}_{t-i}(j))}{\sum_{j=1}^N \psi(\mathbf{Y}_t(j), \mathbf{Y}_{t-i}(j))}; \\ i &= (1, 2, \dots, M); \quad j = (1, 2, \dots, N). \end{aligned} \quad (14)$$

Once the grey relation density $\gamma(t, t-i)$ is obtained, the entropy of the grey relation coefficient of each alternative

$\mathbf{Y}_{t-i}(j)$, $i = (1, 2, \dots, M)$, is then computable, which represents the relevance degree of the grey relation coefficient. The calculation is shown below as

$$E(t-i) = \frac{-\sum_{j=1}^N \gamma(t, t-i) \ln(\gamma(t, t-i))}{\ln(N)}; \quad i = (1, 2, \dots, M), \quad (15)$$

where $-\sum_{j=1}^N \gamma(t, t-i) \ln(\gamma(t, t-i))$ is the grey entropy between the target data flow time sequence $\mathbf{Y}_t(j)$ and the alternative data flow time sequences $\mathbf{Y}_{t-i}(j)$ and $\ln(N)$ is the maximum grey entropy, which guarantees $0 \leq E(t-i) \leq 1$.

Finally, by multiplying the entropy $E(t-i)$ and the average grey relation coefficient of alternative $\mathbf{Y}_{t-i}(j)$, we can obtain the grey relevancy grade (GRG). In this study, the GRG is defined as the numerical measurement of the relevancy between the alternative data flow time sequence $\mathbf{Y}_{t-i}(j)$ and the target data flow time sequence $\mathbf{Y}_t(j)$, which is calculated as

$$D(t-i) = \frac{E(t-i)}{N} \sum_{j=1}^N \psi(\mathbf{Y}_t(j), \mathbf{Y}_{t-i}(j)), \quad (16)$$

where $D(t-i)$ is the GRG of the alternative time sequence $\mathbf{Y}_{t-i}(j)$ with respect to the target sequence $\mathbf{Y}_t(j)$. It is distributed between 0 and 1. Based on the definition of the GRG, it can be seen that the higher the $D(t-i)$ is, the more relevance there is between the alternative data flow time sequence $\mathbf{Y}_{t-i}(j)$ and the target data flow time sequence $\mathbf{Y}_t(j)$.

3.3. Procedures of the Improved BFA. Based on the above discussions, the procedures of the improved BFA are described in Figure 1 and detailed steps are as follows.

Step 1 (predictor calibration and validation). Use the historical data flow of blink number to calibrate the coefficients for each component predictor and then forecast the data flows that will be tested with the trained predictors.

Step 2 (set K identification). Assume that, at the prediction interval, the data flow is potentially correlated with the latest w th intervals' data flows. Then, use the GREBA method to calculate the GRG of each alternative data flow time sequence $\mathbf{Y}_{t-i}(j)$. Ranking the alternative data flow time sequence $\mathbf{Y}_{t-i}(j)$ according to their GRGs in decreasing order, the intervals of the first $R(K)$ alternative data flow time sequences are identified as the set K .

Step 3 (BFA prediction). Predict the data flows using the improved BFA where the weights of the component predictors are calculated by (12).

4. Component Predictors Selection

4.1. ARIMA Model. In theory, ARIMA model is the most common models for time series forecasting. It formulates its mathematical model to fit the time series based on the Box-Jenkins autoregressive moving average (ARMA) model [43].

FIGURE 1: Flowchart of improved BFA.

It is often referred to as ARIMA(p, d, q) model, where p and q are the factors of AR (autoregressive) and MA (moving average), respectively, while d is the difference frequency to make time series stationary [44–46]. The model can be written as

$$\Phi(B)(1-B)^d \hat{y}_t = \Theta(B)\varepsilon_t, \quad (17)$$

where $\hat{y}_t$ denotes the predicted data flow at t th interval. B is the backshift operator defined by $B(\hat{y}_t) = \hat{y}_{t-1}$; $\Phi(B) = 1 - \varphi_1(B) - \varphi_2(B^2) - \dots - \varphi_p(B^p)$ is the AR operator of order p , where $\varphi_1, \varphi_2, \dots$, and φ_p are the autoregressive coefficients; $\Theta(B) = 1 - \theta_1(B) - \theta_2(B^2) - \dots - \theta_q(B^q)$ is the MA operator of order q , where $\theta_1, \theta_2, \dots$, and θ_q are the moving average coefficients; ε_t is the random error which meets normal distribution with mean zero, variance σ^2 , and $\text{cov}(\varepsilon_t, \varepsilon_{t-m}) = 0$, for all $m \neq 0$.

The procedures to formulate ARIMA model are summarized in Figure 2: firstly, AR and MA operator are used to identify whether the time series is a stationary sequence. If not, we difference the data to make it stable. Secondly, the maximum likelihood estimation method is used to evaluate the parameters in ARIMA. Then the resulting model is tested. We diagnose whether the residual error of the time series is white noise or not. If it is not, we reestimate the parameters until the generated residual error turns out to be white noise. Finally, with the suitable parameters, the ARIMA could be applied for prediction [47].

4.2. Radial Basis Function Neural Network Model. ANN models are formulated by emulating the processes of the human neurological system. Among numerous ANN models, the radial basis function neural network (RBFNN) is perhaps the best model, which is often considered as an alternative for data flow forecasting [48]. The RBFNN is a feed-forward neural network with strong local nonlinear approximation ability and fast learning rate. It contains three layers, including the input layer, hidden layer, and output layer.

FIGURE 2: Flowchart of ARIMA modeling.

FIGURE 3: Structure of RBF neural network.

Its structure is presented in Figure 3 [49]. There are k neurons in the input layer; therefore, input data $\mathbf{X}$ can be expressed by a k -dimension column vector, $\mathbf{X} = \{x_t, x_{t-1}, \dots, x_{t-(k-1)}\}$, where x_t is the blink number at current interval t . There are i neurons in the hidden layer, which can be expressed using radial basis function $R(\mathbf{X})$, $R(\mathbf{X}) = \{R_1(\mathbf{X}), R_2(\mathbf{X}), \dots, R_i(\mathbf{X})\}$. There is one neuron in the output layer, which is used to express the prediction value x_{t+1} at next interval $t+1$. $x_{t+1} = f(\mathbf{X}) = \sum_{p=1}^i w_{p,q} R_p(\mathbf{X})$, where $p = \{1, 2, \dots, i\}$, $q = 1$, $w_{p,q}$ is the weight value between the p th neuron in the hidden layer and the neuron in the output layer.

4.3. Kalman Filtering Model. KF is an estimator that provides a set of mathematical equations to estimate the state of a process by minimizing the mean of the squared error. The filter is very powerful in estimating past, present, and even future states. With its high computational efficiency, it has been applied to numerous fields such as process control, flood prediction, tracking, and navigation [50–52]. Assuming

that t is the time interval index, $\hat{y}_t$ and y_t represent the predicted and observed data flow at the current interval, respectively. Let $\mathbf{X}_t = [y_t, y_{t-1}, y_{t-2}, \dots, y_{t-(h-1)}]^T$. According to KF theory, the predicted data flow $\hat{y}_t$ can be formulated as measurement equation:

$$\hat{y}_t = \mathbf{H}_t \mathbf{X}_t + \mathbf{v}_t, \quad (18)$$

where $\mathbf{v}_t$ represents the measurement noise with normal probability distributions $p(\mathbf{v}) \sim N(0, \mathbf{R})$ and $\mathbf{H}_t$ is measurement matrix. $\mathbf{X}_t$ is the state vector at interval t which is updated as

$$\mathbf{X}_t = \Phi_t \mathbf{X}_{t-1} + \omega_{t-1}, \quad (19)$$

where $\mathbf{X}_{t-1}$ is the state vector at interval $t - 1$, ω_{t-1} is the process noise with normal probability distributions of $p(\omega) = N(0, \mathbf{Q})$, and Φ_t is the transition matrix.

According to KF theory, state equations can be expressed as follows:

$$\begin{aligned} \hat{\mathbf{X}}_t &= \mathbf{H}_t \hat{\mathbf{X}}_{t-1} + \mathbf{G}_t [y_t - \mathbf{H}_t \Phi_t \hat{\mathbf{X}}_{t-1}], \\ \mathbf{G}_t &= \mathbf{p}(t | t-1) \mathbf{H}_t^T [\mathbf{H}_t \cdot \mathbf{p}(t | t-1) \cdot \mathbf{H}_t^T + \mathbf{R}]^{-1}, \\ \mathbf{p}(t | t-1) &= \Phi_t \mathbf{p}_{t-1} \Phi_t^T + \mathbf{Q}, \\ \mathbf{p}_t &= [\mathbf{I} - \mathbf{G}_t \cdot \mathbf{H}_t] \mathbf{p}(t | t-1). \end{aligned} \quad (20)$$

The prediction equation is

$$\hat{y}_{t+1} = \mathbf{H}_{t+1} \Phi_{t+1} \hat{\mathbf{X}}_t. \quad (21)$$

To simplify the forecasting procedures and improve the accuracy of KF method, the following tips are adopted when it is applied for data flow prediction: (1) the dimension of state matrix is decided by the ARIMA model, and it is set to be the number of the history data flows that is most correlative with the current data flow; (2) the transition matrix Φ_t is set to be identity matrix; (3) let $\mathbf{H}_t = [1, 0, 0, \dots, 0]^T$; (4) the measurement noise covariance matrices $\mathbf{R}$ can change with each time step or measurement, but for simplicity it is assumed to be a constant decided by sample data of blink number; (5) let $\hat{\mathbf{X}}_0 = E[\mathbf{X}_0]$ and $\mathbf{p}_0 = \text{var}[\mathbf{X}_0]$.

5. Experiments and Analysis

5.1. Experiment Setup. It is very dangerous for subjects to drive in real highway when they feel fatigued; therefore, the experimental data of blink number are acquired by driving simulation experiments in laboratory. The platform of our driving simulation experiment includes hardware and software. The hardware consists of driving simulation equipment, computer, scene display device, video collection card, CCD camera, and so forth. The software is comprised of virtual driving scene and visual detection program of blink number. Experiments are carried out in bright lighting condition. Twenty subjects without any diseases or drinking before the experiment, including 12 male subjects and 8 female subjects who all have more than three years of driving experience, are allowed to take part in this experiment.

5.2. Blink Number Calculation. In this paper we select R_E as an indicator of open-close state of eyes and it can be calculated by (22). Whether driver's eyes are open or closed is estimated according to the size of R_E and the set threshold T_{RE} . If $R_E \geq T_{RE}$, then it will be inferred that the driver is conscious; otherwise, the driver is deduced to be blinking. Consider

$$R_E = \frac{H_E}{W_E}. \quad (22)$$

Physiological studies show that blinking number increases and the duration of the blinking is longer when a driver feels fatigued. Therefore, it is very reasonable that fatigue driving is deduced by the blinking number N_B in given time intervals, where N_B is the frame number of the images where driver's eyes are recognized as being closed. If N_B exceeds the threshold set, then fatigue driving is determined [53].

5.3. Data Collection. In this study, eight hours of blink data each day were collected on driving simulator from 1:00 to 5:00 and from 12:00 to 16:00. Whether drivers are blinking or not is detected and judged by visual detection algorithms based on eye state recognition every 100 milliseconds. A time interval of prediction is set to be five minutes. Four-day data set is separated into two parts. The first part including the data from the first three days is used as historical data. The second part including the data from the fourth day is used for model testing and comparison. The blink data collected in five-minute interval from 1:00 to 5:00 and from 12:00 to 16:00 on the first day are plotted in Figures 4(a) and 4(b), respectively. Figures 4(c) and 4(d) show blink data from 1:00 to 5:00 and from 12:00 to 16:00 on the second day. From Figure 4, we can see that the blink numbers from 3:00 to 5:00 and from 14:00 to 16:00 are much higher than those of other periods. In addition, blink data in midnight and noon contains two peak value, including the midnight peak value (about from 3:30 to 4:20) and the noon peak value (about from 14:40 to 15:20). The observed blink data from 1:00 to 2:30 and from 12:00 to 13:30, however, exhibits no apparently peak value. The testing data is collected from the fourth day, including both peak value data and nonpeak value data. Thus we can compare the prediction performance between the conventional and improved BFA under different data flow pattern.

5.4. Construction of Three Component Predictors

5.4.1. ARIMA. The first three-day data from the first part are used for model identification, which is to determine the orders of AR, difference, and MA. The procedure is discussed in Section 4.1. First of all, the autocorrelation test shows that the partial correlations of the original sample data decay very slowly. Then we difference the sample data to make the sample data become a stationary time sequence. The corresponding test shows the data differenced is stable. Following that, the differenced data is used to identify the orders of AR and MA. With parameter estimation; finally, ARIMA(4, 1, 0) model is identified to be the most suitable model with statistical packages.

FIGURE 4: Two-day blink data in 5-minute interval: (a) blink data collected in five-minute interval from 1:00 to 5:00 on the first day; (b) blink data collected in five-minute interval from 12:00 to 16:00 on the first day; (c) blink data collected in five-minute interval from 1:00 to 5:00 on the second day; (d) blink data collected in five-minute interval from 12:00 to 16:00 on the second day.

5.4.2. RBFNN. Blink data collected in the first three days is used for network training. RBFNN is selected with the following architecture: five neurons in the input layer, one hidden layer with p neurons, and one output neuron. At a prediction interval, the latest five blink data, x_t , x_{t-1} , x_{t-2} , x_{t-3} , and x_{t-4} are used for the input data. Gaussian function is chosen as the radial basis function in the hidden layer; that is, $R_p(\mathbf{X}) = \exp(-\|\mathbf{X} - \mathbf{c}_p\|^2 / 2(\sigma_p)^2)$, where $\mathbf{c}_p$ is the center of Gaussian function and σ_p is the width. During the training RBFNN, $\mathbf{c}_p$ and p are calculated by using k -means clustering algorithm [54] and $\sigma_p = d_{\max} / \sqrt{2l}$, where $d_{\max}$ is the largest distance between each two $\mathbf{c}_p$ that has been obtained.

5.4.3. KF. KF was used for forecasting by extracting the most recent blink data measured. As is mentioned before, the orders of AR are regarded as the number of previous blink data that the KF is based on. Therefore, the dimension of $\mathbf{H}_t$ in (18) is 1×4 .

5.5. Evaluation Criterion of Prediction Performance. We select and employ three evaluation criterions, that is, the mean absolute error (MSA), the mean relative error (MRE), and variance of relative error (VRE), to evaluate the prediction performance of predictors. Evaluation criterions are calculated by (23). Among these criterions, MSA and MRE are

TABLE 1: Prediction performance of the three component predictors.

	ARIMA			RBFNN			KF		
	MSA	MRE (%)	VRE (%)	MSA	MRE (%)	VRE (%)	MSA	MRE (%)	VRE (%)
1:00–2:00	4.25	7.51	8.20	5.32	7.82	8.53	6.26	6.35	7.18
2:00–3:00	19.60	8.23	9.07	10.42	7.51	8.02	15.66	8.17	9.08
3:00–4:00	15.53	8.04	9.12	10.07	7.62	7.99	17.72	8.28	9.24
4:00–5:00	22.71	7.93	8.11	15.87	7.49	7.93	23.72	8.32	8.77
Total	16.13	8.09	9.14	12.17	7.71	8.29	16.33	8.21	9.19

used to measure the prediction accuracy, while VRE is used to present the prediction stability. Consider

$$\begin{aligned}
 \text{MSA} &= \frac{1}{S} \sum_{t=1}^S |\hat{y}_t - y_t|, \\
 \text{MRE} &= \frac{1}{S} \sum_{t=1}^S \left| \frac{\hat{y}_t - y_t}{y_t} \right|, \\
 \text{VRE} &= \sqrt{\frac{S \sum_{t=1}^S (|y_t - \hat{y}_t| / y_t)^2 - \left[\sum_{t=1}^S (|y_t - \hat{y}_t| / y_t) \right]^2}{S(S-1)}},
 \end{aligned} \quad (23)$$

where y_t is the observed blink number in time interval t , $\hat{y}_t$ is the predicted blink number in time interval t , and S is the number of prediction intervals.

5.6. Predictions of Component Predictors. After being trained with the blink data of the first three days, the three component predictors are used to predict the blink number of next observation period on the fourth day. Statistics show that drivers usually feel the most fatigued in the period from 1:00 to 5:00 on each day, even though they have good sleep on the last day, and the blink data flow in the period has an obvious fluctuation. Therefore, we select the blink data from 1:00 to 5:00 to test prediction performance of each predictor. Figure 5 describes the outputs of the three component predictors for blink number in the period from 1:00 to 5:00. The corresponding prediction performances are summarized in Table 1. Table 1 indicates that, despite all component predictors successfully making predictions as a whole, no single predictor can retain the best prediction performance all the time. It can be observed from Table 1 that while the ARIMA predictor made relatively better predictions in the period from 1:00 to 2:30 than RBFNN and KF, it delivers the poorest prediction performance in the period from 2:30 to 3:30. Therefore, it is necessary to find ways to integrate the merits of the single predictors to improve the prediction accuracy and stability.

5.7. Comparison of Prediction Performance of the Improved BFA and the Conventional BFA. Prior to applying the improved BFA for blink number forecasting, the set K should be determined by using the GREBA method. We assumed that the blink number at the prediction interval ($Y_t(j)$) is

FIGURE 5: Observed and predicted blink number of the three component predictors on the fourth day from 1:00 to 5:00.

potentially affected by the blink number from the last 10 intervals (50 minutes) ($Y_{t-1}(j), Y_{t-2}(j), \dots, Y_{t-10}(j)$). Our aim is to identify a few intervals at which the blink number has comparatively higher relevance to the blink number at the prediction intervals. Once determined, the prediction errors of each of the component predictors at these intervals will be used to decide their weights in the improved BFA. The GRG between each alternative ($Y_{t-1}(j), Y_{t-2}(j), \dots, Y_{t-10}(j)$) and the target blink number time sequences $Y_t(j)$ is described in Figure 6. It demonstrates that although there are some perturbations between adjacent intervals, the GRG decreases with respect to decreasing time intervals ($t-1, t-2, \dots$) in a longer period, which implies that the further the time gets away from the current prediction intervals, the less relevance there is between the corresponding blink number and the current blink number. This result coincides with our hypothesis that the blink number is very complicated and usually exhibits irregular randomness. It is affected by subject's sleep situation, continuous driving time, road traffic conditions, and subject's psychological or physiological conditions, and only a few of the latest blink data are competent enough to dynamically observe the blink number in prediction interval. Note that, in Figure 6, the GRG ceased to decrease dramatically after three intervals, the blink number at the past 3 intervals is thus identified to be the value that mostly

TABLE 2: Prediction performance of the conventional BFA and the improved BFA.

	Conventional BFA			Improved BFA ($K = \{1, 2, 3\}$)		
	MSA	MRE (%)	VRE (%)	MSA	MRE (%)	VRE (%)
1:00–2:00	5.03	5.63	5.03	3.28	4.16	3.52
2:00–3:00	16.33	8.05	9.14	12.11	7.99	8.75
3:00–4:00	13.21	7.03	8.77	9.03	6.87	7.89
4:00–5:00	15.94	8.09	9.26	10.04	7.01	8.37

FIGURE 6: The GRG between the alternative data flow time sequences ($Y_{t-1}, Y_{t-2}, \dots, Y_{t-10}$) and the target data flow time sequences Y_t of blink number.

correlates with the blink number at forecasting interval; that is, $K = \{t-1, t-2, t-3\}$.

The prediction performances of the improved BFA and the conventional BFA are listed in Table 2. Tables 1 and 2 imply that while the conventional BFA improves the prediction performance of single predictors to some extent, it does not compare to the improved BFA. Table 2 indicates that the improved BFA considerably outperforms the conventional BFA in both accuracy and stability. For example, the improved BFA has a performance of 4.16% MRE and 3.52% VRE in the period from 1:00 to 2:00 and a 26.11% improvement for MRE and a 30.01% improvement for VRE over the conventional BFA, which has a MRE of 5.63% and a VRE of 5.03%. Even the improved BFA has its worst performance in the period from 2:00 to 3:00, with 7.99% MRE and 12.11 MSA; it is still better, however, than the corresponding performance of the conventional BFA. Another insight gained from the comparison of conventional BFA and the improved BFA in Tables 1 and 2 is that the performances of both methods depend heavily on the performance of the component predictors that each method combines. Thus, both BFAs perform better if each component predictor is more accurate.

Note that the prediction performances of both the conventional BFA and the improved BFA rely heavily on the weights of the component predictors calculated by (10) and (12). Because the weight in the BFA characterizes the estimated possibility that a particular predictor best forecasts the observed blink number, whether the BFA can assign more weight to the “actual” best predictor at each prediction interval is thus very crucial to its prediction accuracy.

Figure 7(a) shows that the weight of the ARIMA predictor assigned by the conventional BFA is the highest from 1:00 to 2:30 and then decreases slowly and finally approaches 0 after 2:30. The weight of the RBFNN predictor increases monotonically over time and eventually dominates the predictions of the conventional BFA. It happens because the predictor with the best average performance changed during the day, from the ARIMA predictor to the KF predictor. However, upon inspection, we find that the weight calculated by the conventional BFA is unable to dynamically monitor the disturbed accuracy of each component predictor. For example, Figure 5 shows that the RBFNN predictor delivers the best predictions in the periods from 2:30 to 3:30 and from 4:20 to 5:00. However, during the periods, the weight of the RBFNN assigned by the conventional BFA is not the highest of the three, as shown in Figure 7(a). Nevertheless, the extraordinary prediction performance of the RBFNN predictor is indeed detected by the conventional BFA, given that the weight of the RBFNN predictor increases monotonously during this period. However, because the weights of the component predictors at every prediction interval in the conventional BFA are accounted for by their prediction errors from all past intervals, the increased scope is very limited. Once the weight of RBFNN predictor finally has a leading role, it is also difficult to make a change even if there are comparatively larger prediction errors for RBFNN predictor during a period. This effect could be justified by the predictions in Figure 5 in the period from 3:30 to 4:20, during the period the RBFNN predictor makes much worse predictions than the KF predictor. However, Figure 7(a) shows that the weight of the RBFNN predictor assigned by the conventional BFA is dramatically larger than that of the KF predictor, which may lead to inaccurate predictions. For our improved BFA, there is no such deficiency because the weight of the component predictors at the current interval is only affected by their prediction errors from the latest three intervals. Therefore, the improved BFA is more sensitive to the fluctuated accuracy of the component predictors and can more quickly adjust their weight. Figure 7(b) shows that, in the periods from 2:30 to 3:30 and from 4:20 to 5:00, when

FIGURE 7: Forecasted results of the conventional BFA and the improved BFA in the period from 1:00 to 5:00: (a) evolution of the weights assigned to each component predictor by the conventional BFA; (b) evolution of the weights assigned to each component predictor by the improved BFA; (c) comparison of the predictions of the conventional BFA and the improved BFA.

the RBFNN predictor makes the best predictions, its weight assigned by the improved BFA is always the highest of the three, and when it is no longer the best predictor in the period from 3:30 to 4:20, its corresponding weight decreases dramatically; thus, better predictions are generated compared to the conventional BFA (Figure 7(c)).

Figure 7(a) also reveals that, in most cases, none of single predictors can achieve the best performance during all prediction periods. However, our improved BFA is sensitive to the disturbed accuracy of each component predictor and can assign more weight to the best predictor. Therefore, it is definitely more preferable than the conventional BFA.

Based on the forecasted blink number, further experiments that can be used to identify whether drivers are

fatigued or not show that correct rate of fatigue driving detection can reach 98.4%; false-alarm rate is only 1.9%. Therefore, fatigue driving can be accurately warned against in advance based on the blink number forecasted by the improved BFA.

6. Conclusions

For forecasting fatigue state of drivers in advance, we propose an improved BFA to enhance the accuracy and stability of blink number forecasted in next time interval. By analyzing the correlation between historical data flows and current data flows, the improved BFA only considers the prediction errors from the latest few intervals when calculating the weights

of each component predictor. Thus, it is more sensitive to the disturbed accuracy of the component predictors and can adjust their weights more quickly than the conventional BFA. To make full use of the advantages of various methods, ARIMA, BPNN, and KF are designed and fused linearly into the BFA. Simulating experiments demonstrate that the improved BFA outperforms the conventional BFA in both accuracy and stability considerably. We found that the conventional BFA is quite impervious to the disturbed accuracy of the component predictors. Most of the time, it assigns nearly all weight to one component predictor. Compared with the conventional BFA, our improved BFA can more rapidly adjust the weights of the component predictors and generate better predictions. Therefore, the improved BFA is more preferable for blink number prediction than the conventional BFA.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

The research has been supported in part by the National Nature Science Foundation of China (nos. 61304205, 61203273, 61103086, and 41301037), Nature Science Foundation of Jiangsu Province (BK20141002), the Open Funding Project of State Key Laboratory of Virtual Reality Technology and Systems, Beihang University (no. BUAA-VR-13KF-04), Jiangsu Ordinary University Science Research Project (no. 13KJB120007), and Innovation and Entrepreneurship Training Project of College Students (nos. 201510300228, 201510300276).

References

- [1] S. Y. Hu, G. T. Zheng, and B. Peters, "Driver fatigue detection from electroencephalogram spectrum after electrooculography artefact removal," *IET Intelligent Transport Systems*, vol. 7, no. 1, pp. 105–113, 2013.
- [2] S. K. L. Lal, A. Craig, P. Boord, L. Kirkup, and H. Nguyen, "Development of an algorithm for an EEG-based driver fatigue countermeasure," *Journal of Safety Research*, vol. 34, no. 3, pp. 321–328, 2003.
- [3] S. K. L. Lal and A. Craig, "Driver fatigue: electroencephalography and psychological assessment," *Psychophysiology*, vol. 39, no. 3, pp. 313–321, 2002.
- [4] C. Ahlstrom, M. Nyström, K. Holmqvist et al., "Fit-for-duty test for estimation of drivers' sleepiness level: eye movements improve the sleep/wake predictor," *Transportation Research Part C: Emerging Technologies*, vol. 26, pp. 20–32, 2013.
- [5] M. Simon, E. A. Schmidt, W. E. Kincses et al., "Eeg alpha spindle measures as indicators of driver fatigue under real traffic conditions," *Clinical Neurophysiology*, vol. 122, no. 6, pp. 1168–1178, 2011.
- [6] B. T. Jap, S. Lal, P. Fischer, and E. Bekiaris, "Using EEG spectral components to assess algorithms for detecting fatigue," *Expert Systems with Applications*, vol. 36, no. 2, pp. 2352–2359, 2009.
- [7] G. Li and W.-Y. Chung, "Detection of driver drowsiness using wavelet analysis of heart rate variability and a support vector machine classifier," *Sensors*, vol. 13, no. 12, pp. 16494–16511, 2013.
- [8] M. Patel, S. K. L. Lal, D. Kavanagh, and P. Rossiter, "Applying neural network analysis on heart rate variability data to assess driver fatigue," *Expert Systems with Applications*, vol. 38, no. 6, pp. 7235–7242, 2011.
- [9] C.-S. Hsieh and C.-C. Tai, "An improved and portable eye-blink duration detection system to warn of driver fatigue," *Instrumentation Science and Technology*, vol. 41, no. 5, pp. 429–444, 2013.
- [10] J. Jo, S. J. Lee, K. R. Park, I.-J. Kim, and J. Kim, "Detecting driver drowsiness using feature-level fusion and user-specific classification," *Expert Systems with Applications*, vol. 41, no. 4, pp. 1139–1152, 2014.
- [11] T. D'Orazio, M. Leo, C. Guaragnella, and A. Distanto, "A visual approach for driver inattention detection," *Pattern Recognition*, vol. 40, no. 8, pp. 2341–2355, 2007.
- [12] Y. Du, Q.-H. Hu, D.-G. Chen, and P.-J. Ma, "Kernelized fuzzy rough sets based yawn detection for driver fatigue monitoring," *Fundamenta Informaticae*, vol. 111, no. 1, pp. 65–79, 2011.
- [13] T. S. Wang and P. F. Shi, "Yawning detection for determining driver drowsiness," in *Proceedings of the IEEE International Workshop on VLSI Design and Video Technology*, pp. 385–388, May 2005.
- [14] Ö. Tunçer, L. Güvenç, F. Coşkun, and E. Karsligil, "Vision based lane keeping assistance control triggered by a driver inattention monitor," in *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics (SMC '10)*, pp. 289–297, October 2010.
- [15] T. Pilutti and A. Galip Ulsoy, "Identification of driver state for lane-keeping tasks," *IEEE Transactions on Systems, Man, and Cybernetics, Part A: Systems and Humans*, vol. 29, no. 5, pp. 486–502, 1999.
- [16] Q. He, W. Li, and X. Fan, "Estimation of driver's fatigue based on steering wheel angle," in *Engineering Psychology and Cognitive Ergonomics: 9th International Conference, EPCE 2011, Held as Part of HCI International 2011, Orlando, FL, USA, July 9-14, 2011. Proceedings*, vol. 6781 of *Lecture Notes in Computer Science*, pp. 145–155, Springer, Berlin, Germany, 2011.
- [17] R. Sayed and A. Eskandarian, "Unobtrusive drowsiness detection by neural network learning of driver steering," *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering*, vol. 215, no. 9, pp. 969–975, 2001.
- [18] A. Sahayadhas, K. Sundaraj, and M. Murugappan, "Detecting driver drowsiness based on sensors: a review," *Sensors*, vol. 12, no. 12, pp. 16937–16953, 2012.
- [19] R. Schleicher, N. Galley, S. Briest, and L. Galley, "Blinks and saccades as indicators of fatigue in sleepiness warnings: looking tired?" *Ergonomics*, vol. 51, no. 7, pp. 982–1010, 2008.
- [20] M. Eriksson and N. P. Papanikolopoulos, "Driver fatigue: a vision-based approach to automatic diagnosis," *Transportation Research Part C: Emerging Technologies*, vol. 9, no. 6, pp. 399–413, 2001.
- [21] B.-G. Lee and W.-Y. Chung, "Driver alertness monitoring using fusion of facial features and bio-signals," *IEEE Sensors Journal*, vol. 12, no. 7, pp. 2416–2422, 2012.
- [22] Y.-C. Dong, Z.-C. Hu, K. Uchimura, and N. Murayama, "Driver inattention monitoring system for intelligent vehicles: a review," *IEEE Transactions on Intelligent Transportation Systems*, vol. 12, no. 2, pp. 596–614, 2011.

- [23] G. Yang, Y. Lin, and P. Bhattacharya, "Multimodality inferring of human cognitive states based on integration of neuro-fuzzy network and information fusion techniques," *EURASIP Journal on Advances in Signal Processing*, vol. 2008, Article ID 371621, 2008.
- [24] L. Wang and Y. Pei, "The impact of continuous driving time and rest time on commercial drivers' driving performance and recovery," *Journal of Safety Research*, vol. 50, pp. 11–15, 2014.
- [25] S. A. Gadsden, Y. Song, and S. R. Habibi, "Novel model-based estimators for the purposes of fault detection and diagnosis," *IEEE/ASME Transactions on Mechatronics*, vol. 18, no. 4, pp. 1237–1249, 2013.
- [26] D. Kim, P. Kang, and S. Cho, "Machine learning-based novelty detection for faulty wafer detection in semiconductor manufacturing," *Engineering Applications of Artificial Intelligence*, vol. 25, no. 4, pp. 814–823, 2012.
- [27] Y.-S. Jeong, Y.-J. Byon, M. M. Castro-Neto, and S. M. Easa, "Supervised weighting-online learning algorithm for short-term traffic flow prediction," *IEEE Transactions on Intelligent Transportation Systems*, vol. 14, no. 4, pp. 1700–1707, 2013.
- [28] Y. Ouyang and H. J. Yin, "A neural gas mixture autoregressive network for modelling and forecasting FX time series," *Neurocomputing*, vol. 135, pp. 171–179, 2014.
- [29] H. Tamura, P. Bacopoulos, D. B. Wang, S. C. Hagen, and E. J. Kubatko, "State estimation of tidal hydrodynamics using ensemble Kalman filter," *Advances in Water Resources*, vol. 63, pp. 45–56, 2014.
- [30] P. Lorca, M. Landajo, and J. de Andrés, "Nonparametric quantile regression-based classifiers for bankruptcy forecasting," *Journal of Forecasting*, vol. 33, no. 2, pp. 124–133, 2014.
- [31] J. E. Contreras-Reyes and W. Palma, "Statistical analysis of autoregressive fractionally integrated moving average models in R," *Computational Statistics*, vol. 28, no. 5, pp. 2309–2331, 2013.
- [32] E. Ebrahimi, H. Bayat, M. R. Neyshaburi, and H. Z. Abyaneh, "Prediction capability of different soil water retention curve models using artificial neural networks," *Archives of Agronomy and Soil Science*, vol. 60, no. 6, pp. 859–879, 2014.
- [33] A. Azadeh, S. F. Ghaderi, M. Sheikhalishahi, and B. P. Nokhandan, "Optimization of short load forecasting in electricity market of iran using artificial neural networks," *Optimization and Engineering*, vol. 15, no. 2, pp. 485–508, 2014.
- [34] X. Yan and N. A. Chowdhury, "Mid-term electricity market clearing price forecasting: a multiple SVM approach," *International Journal of Electrical Power & Energy Systems*, vol. 58, pp. 206–214, 2014.
- [35] H. R. Kirby, S. M. Watson, and M. S. Dougherty, "Should we use neural networks or statistical models for short-term motorway traffic forecasting?" *International Journal of Forecasting*, vol. 13, no. 1, pp. 43–50, 1997.
- [36] B. L. Smith, B. M. Williams, and R. K. Oswald, "Comparison of parametric and nonparametric models for traffic flow forecasting," *Transportation Research Part C: Emerging Technologies*, vol. 10, no. 4, pp. 303–321, 2002.
- [37] V. Petridis, A. Kehagias, L. Petrou et al., "A Bayesian multiple models combination method for time series prediction," *Journal of Intelligent and Robotic Systems: Theory and Applications*, vol. 31, no. 1–3, pp. 69–89, 2001.
- [38] Q. S. Zhang, J. L. Deng, and Y. Shao, "Grey correlation analysis by the method of degree of balance and approach," *Journal of Huazhong University of Science and Technology*, vol. 23, no. 11, pp. 94–98, 1995.
- [39] Z. Wang, Q. Wang, and T. Ai, "Comparative study on effects of binders and curing ages on properties of cement emulsified asphalt mixture using gray correlation entropy analysis," *Construction and Building Materials*, vol. 54, pp. 615–622, 2014.
- [40] Y. Zhang and R.-C. Zhang, "Study on the third party logistics service providers' performance evaluation based on the weighted entropy and analysis process of grey relation," in *Proceedings of the 17th International Conference on Management Science and Engineering (ICMSE '10)*, pp. 582–587, IEEE, Melbourne, Australia, November 2010.
- [41] C. F. Liu and G. F. Dang, "Related analysis of the evolvement of land utilize structural and social economic development based on the information entropy," in *Proceedings of the 8th International Conference on Information and Management Sciences*, pp. 351–354, July 2009.
- [42] X. Xia and F. Meng, "Grey relational analysis of measure for uncertainty of rolling bearing friction torque as time series," *Journal of Grey System*, vol. 23, no. 2, pp. 135–144, 2011.
- [43] Y.-S. Lee and L.-I. Tong, "Forecasting time series using a methodology based on autoregressive integrated moving average and genetic programming," *Knowledge-Based Systems*, vol. 24, no. 1, pp. 66–72, 2011.
- [44] W. H. K. Tsui, H. Ozer Balli, A. Gilbey, and H. Gow, "Forecasting of hong kong airport's passenger throughput," *Tourism Management*, vol. 42, pp. 62–76, 2014.
- [45] T. M. Wyatt, W. Dejong, and E. Dixon, "Population-level administration of alcoholedu for college: an ARIMA time-series analysis," *Journal of Health Communication*, vol. 18, no. 8, pp. 898–912, 2013.
- [46] H.-W. Kim, J.-H. Lee, Y.-H. Choi, Y.-U. Chung, and H. Lee, "Dynamic bandwidth provisioning using ARIMA-based traffic forecasting for Mobile WiMAX," *Computer Communications*, vol. 34, no. 1, pp. 99–106, 2011.
- [47] P. Han, P. X. Wang, S. Y. Zhang, and D. H. Zhu, "Drought forecasting based on the remote sensing data using ARIMA models," *Mathematical and Computer Modelling*, vol. 51, no. 11–12, pp. 1398–1403, 2010.
- [48] H. L. Niu and J. Wang, "Financial time series prediction by a random data-time effective RBF neural network," *Soft Computing*, vol. 18, no. 3, pp. 497–508, 2014.
- [49] A. J. Telmoudi, H. Tlijani, L. Nabli, M. Ali, and R. M'Hiri, "A New rbf neural network for prediction in industrial control," *International Journal of Information Technology and Decision Making*, vol. 11, no. 4, pp. 749–775, 2012.
- [50] P. Baraldi, F. Mangili, and E. Zio, "A Kalman filter-based ensemble approach with application to turbine creep prognostics," *IEEE Transactions on Reliability*, vol. 61, no. 4, pp. 966–977, 2012.
- [51] C. Papadimitriou, C.-P. Fritzen, P. Kraemer, and E. Ntotsios, "Fatigue predictions in entire body of metallic structures from a limited number of vibration sensors using Kalman filtering," *Structural Control and Health Monitoring*, vol. 18, no. 5, pp. 554–573, 2011.
- [52] P. Louka, G. Galanis, N. Siebert et al., "Improvements in wind speed forecasts for wind power prediction purposes using Kalman filtering," *Journal of Wind Engineering and Industrial Aerodynamics*, vol. 96, no. 12, pp. 2348–2362, 2008.
- [53] W. Sun, X. R. Zhang, W. Zhuang, and H. Tang, "Driver fatigue driving detection based on eye state," *International Journal of Digital Content Technology and its Applications*, vol. 5, no. 10, pp. 307–314, 2011.

- [54] S. K. Oh, W. D. Kim, W. Pedrycz, and S.-C. Joo, "Design of K-means clustering-based polynomial radial basis function neural networks (pRBF NNs) realized with the aid of particle swarm optimization and differential evolution," *Neurocomputing*, vol. 78, no. 1, pp. 121–132, 2012.