Improved Hybrid Intelligent Method for Urban Road Traffic Flow Forecasting based on Chaos-PSO Optimization

Yuesheng Gu, Yanpei Liu, Jiucheng Xu

International Journal of Advancements in Computing Technology Volume 3, Number 7, August 2011

Abstract

Real time traffic flow is often difficult to predict precisely because of the nonlinear and stochastic characteristics of the traffic flow data. Intelligent prediction methods such as artificial neural network (ANN), support vector machine (SVM), etc. have been proven effective to discover the nonlinear information hidden in the traffic flow data. Nevertheless, their efficiency limits in the low recognition rate. If using them independently, the prediction accuracy would be lower than integrated use. Hence, a new hybrid intelligent prediction approach based on the combination of advanced signal processing technique and intelligent data mining method is proposed for the short time traffic flow prediction in this paper. The new method is marked as the use of empirical mode decomposition (EMD) and support vector machine (SVM) to deal with the nonlinear and stochastic characteristics of the traffic flow data. Another advantage of the proposed method is that the Chaos-particle swarm optimization (PSO) is adopted for the optimization of the combination. By doing so, the local optimization of the EMD-SVM prediction model can be avoided and the forecasting rate can be enhanced significantly. The practical traffic flow data were applied to the validation of the proposed prediction model. The analysis results show that the proposed method can extract the underlying rules of the testing data and decrease prediction error by 0.53% or better when compared with single SVM approach. Thus, the new hybrid intelligent traffic flow forecasting model can provide practical application.

Keywords: Traffic Flow Prediction, Hybrid Intelligence, EMD, SVM, Chaos-PSO

1. Introduction

Traffic flow prediction has become a popular research topic in the field of intelligent transport system (ITS). Accurate and real-time traffic flow prediction is the key fact for traffic control and traffic induction. Short time traffic flow prediction can make prediction about the next several minutes’ state of the traffic flow to provide real-time effective information for travelers, realize the dynamic route guidance, save travel time, relieve the traffic congestion, and reduce pollution, save energy and other purposes [1-2]. Therefore, it is imperative to implement short time traffic flow prediction in the practice for the traffic control and traffic induction.

Due to the fact that the traffic system is a complex nonlinear system with time-varying and high uncertainties, the prediction accuracy of existing methods cannot be satisfactory when they are used independently. The forecasting accuracy need to be improved for real practice applications. Since the integration of different analysis techniques can provide better performance than independent use, a new hybrid approach to short-term traffic flow prediction based on empirical mode decomposition (EMD) and artificial intelligence is proposed in this work. This method has been marked as the advantages of the good nonlinear signal process ability of the EMD and the powerful learning ability of the support vector machine (SVM).

Considering the fact that improper structure parameters of SVM may lead to low precision for the traffic flow prediction, in order to overcome this problem, the evolution optimization procedure is needed. In point of view that the particle swarm optimization (PSO) algorithm has good global search capability, it has been employed to optimize the ANN and SVM in wide applications and has achieved satisfactory results in improving the efficiency of the artificial intelligence [3-5]. But, the PSO will often not sensitive to fitness function, and the phenomenon of local minimum often occurs [9-10], which do hazards to the traffic flow prediction. Therefore, in order to eliminate PSO’s defects in the
optimization of the SVM, this paper puts forward a method based on Chaos-PSO and SVM. This method has been marked by chaotic technology which can improve PSO optimization process and avoid the above local minimum problem, and the SVM structure and parameters can be optimized in a better way to obtain good generalization ability of the prediction model. By using the practical dataset for experimental analysis, the experimental results show that the new method can predict short time traffic flow effectively and the prediction rate is higher than the independent use.

This paper is organized as follows. In Section 2, the proposed hybrid intelligent method for short time traffic flow forecasting based on the combination of empirical mode decomposition (EMD), Chaos-PSO and support vector machine (SVM) is described. The application of the proposed method is presented for short time traffic flow forecasting in Section 3. The performance of nonlinear signal process using EMD, as well as the traffic flow prediction performance is described. The effectiveness of the proposed method is valued by analyzing the real traffic data. Conclusions are drawn in Section 4.

2. Hybrid Intelligent Model

Due to the interference of inside and external excitations, the short-term traffic flow is a kind of typical non-stationary signal. The different signal components of short-term traffic flow exhibits various characteristics, and make different effects under the influence of change trend of traffic flow. The general trend of traffic flow is determined by deterministic signal, and the uncertain interference signals make the actual traffic flow present fluctuations nearby the general trend.

In the analysis of short-term traffic flow, the EMD is firstly used to decompose the actual traffic flow to remove the disturbance signals. Then, the predicted mode for each IMF is established using SVM, and the Chaos-PSO is applied to the model optimization. Finally, the traffic flow is obtained by adding up the predictive value of each SVM models.

2.1. Empirical Mode Decomposition (EMD)

EMD [6] is a new approach to deal with non-stationary signal. According to different scales of fluctuation, non-stationary signal is decomposed step-by-step into some IMFs (Intrinsic Mode Functions). Each IMF includes signals of different bands from high to low, and has its unequal features. Moreover, the adaptive decomposition is presented by EMD based on the inherent characteristics of signal [6-8]. Therefore, these different modes can reflect the essence and potential rule of the traffic flow more clearly.

To extract IMFs from a signal x, all the local extrema are firstly identified. Then a cubic spline line connects all the local maxima as upper envelope and all the minima as lower envelope. The mean of upper and lower envelope is subtracted from x to obtain h1. Check h1 for the IMF conditions. If it satisfies the conditions it is an IMF, otherwise upper and lower envelopes are found for the h1 and the process is repeated till the first IMF c1 is got. Subtract c1 from x and the result is now treated as new original signal and the above process is repeated to get the second IMF. Keep continuing the process till no more IMF can be extracted. Thus, at the end of the EMD decomposition we obtain

\[ x = \sum_{i=1}^{N} c_i + r_N \]  

(1)

where, \( r_N \) is the final residue and \( c_i \) (i=1, 2, ..., N) is the ith IMF.

2.2. Least Square Support Vector Machine (LS-SVM)

Since there may be a certain correlation between current state of the traffic data and the future state, which may be difficult to describe using analytical methods, the support vector machine (SVM) [9] is applied to learn the relationship of them. The SVM, which has the ability to find the decision function from low training set sizes, has been widely used as a learning algorithm in a wide variety of applications. The concept of the kernel trick allows SVM to perform regression and prediction even for
nonlinear cases. Since the least squares support vector machine (LS-SVM) is an improved algorithm based on SVM which adopts equality constraint to replace the inequality constraints for standard support vector machine to put the SVM quadratic programming problem into linear equations and realize the simplified algorithm, the LS-SVM algorithm is used in this work. The LS-SVM is used for find a hyperplane by using maximum Euclidean distance to the nearest point [10]. The LS-SVM can map the input data into a high dimensional feature space and classify the nonseparable data in the optimal separating hyperplane.

LS-SVM theory of detailed derivation and demonstration please refer to the reference [11-12], and the LS-SVM regression model is given by

$$ f(x) = \sum_{i=1}^{l} \alpha_i k(x_i, x) + b, $$

Where $\alpha$ denotes the Lagrange multiplier, and $b$ denotes the bias constants. The RBF kernel is adopted as the kernel function

$$ K(x_i, x) = \exp[-\frac{(x-x_i)^2}{2\sigma^2}]. $$

For the RBF kernel based LS-SVM, the width of the Gaussian kernels ($\sigma$) and the regularization factor $C$ play great role for the generalization ability of the SVM [12]. It is wise to optimize these two parameters in the application of LS-SVM. So this paper adopts Chaos-PSO algorithm to optimize the LS-SVM parameters.

### 2.3. Chaos-PSO Optimization

Particle Swarm Optimization (PSO) [13] is a novel evolutionary algorithm inspired by the social behavior of bird flocks and fish schools. PSO has advantages of simplicity and effectiveness for extreme value search, and has been applied successfully to the optimization of ANN and SVM etc. However, the studies of PSO also show the problems with PSO optimization. PSO may suffer from premature convergence, leading to local minima [14-16]. Research addressing the shortcomings of PSO is ongoing. Great success has been done, but the possibility for improvement of PSO is still open. To overcome PSO’s shortcoming of local minima, a new PSO algorithm via increased particle diversity using Chaos search is adopted in this paper. The Chaos search can get all the states in the search space by the rules of itself and generate neighborhoods of near optimal solutions to maintain solution diversity [17]. The integration of Chaos and PSO can prevent the optimization process from premature.

PSO involves two important parameters, i.e. the individual extreme $P_i$, the global extreme $G$ experienced for the whole populations is the optimal solution at present. The algorithm for updating speed $\dot{v}$ and position $P$ is given by

$$ \dot{v}_i(k+1) = \omega \dot{v}_i(k) + c_1 r_1 [P_e - P_i(k)] + c_2 r_2 [G - P_i(k)], $$

$$ P_i(k+1) = P_i(k) + \dot{v}_i(k+1), $$

where, $\omega_i$ denotes the inertial weights, $r_1$ and $r_2$ denote the random coefficients among [0,1], $c_1$ and $c_2$ denote the speeding constants.
The chaotic algorithm is adopted for the secondary optimization to overcome the premature of standard particle swarm algorithm. The most used Chaos algorithm is Logistic equation. The mapping expression of Logistic can be expressed as [18]

\[ P_i(k+1) = \mu P_i(k)(1 - P_i(k)), \]  

(6)

where \( \mu \) is the control parameter. The Chaos optimization process is as follows.

Firstly, the flying position \( P \) is mapped to chaos variables in \([0, 1]\) by the carrier given by Eq. (7)

\[ Q_i(k) = (P_i(k) - a_i)/b_i, \]  

(7)

where \( a_i \) and \( b_i \) need to be specified. Then Eq. (6) is adopted to calculate the chaos behavior:

\[ Q_i(k+1) = \mu Q_i(k)(1 - Q_i(k)), \]  

(8)

After several iterative, the optimal value \( f^* \) is find for the best \( Q_i \), and the second carrier is used to map the chaos back to its original space:

\[ P_i(k) = a_i + b_i Q_i(k), \]  

(9)

Through this process, the chaos algorithm can effectively find reasonable global extreme \( G \), help PSO jump out of local optimal solution.

The Henon Map is proposed to enhance the chaos search ability. The Henon map is given by [19]

\[ \begin{align*}
  P_i(k+1) &= 1 - c P_i(k)^2 + d Q_i(k) \\
  Q_i(k+1) &= P_i(k)
\end{align*} \]  

(10)

2.4. The Proposed Prediction Model

In this paper the SVM with RBF kernel is used for the traffic condition forecasting. Moreover, to improve the prediction model robust, the Chaos-particle swarm optimization (PSO) algorithm is adopted to optimize the LS-SVM regularization factor \( C \) and kernel width \( \sigma \). The proposed forecasting processes are given as follows:

Step 1: Pre-treat the original traffic data to standardized data format.

Step 2: Extract nonlinear features from the input traffic data in the form of IMF by EMD.

Step 3: Train LS-SVM by each IMF, optimize the regularization factor \( C \) and kernel width \( \sigma \) by Chaos-PSO, and sum each SVM model output to obtain the prediction result.

Step 4: Test the performance of the LS-SVM prediction model, and provide the test result as the base for a valid traffic management decision. A flow chart of the proposed prediction method for short-term traffic flow is illustrated in Fig. 1.

![Figure 1. The short time traffic flow prediction system based on EMD-Chaos-PSO-SVM.](image-url)
Improved Hybrid Intelligent Method for Urban Road Traffic Flow Forecasting based on Chaos-PSO Optimization
Yuesheng Gu, Yanpei Liu, Jiucheng Xu
International Journal of Advancements in Computing Technology Volume 3, Number 7, August 2011

The following indexes are selected to evaluate the performance of the traffic flow prediction:

1. Mean Absolute Error (MAE)

\[
MAE = \frac{1}{N} \sum_{t=1}^{N} |y_t - \hat{y}_t|
\]

2. Mean Square Error (MSE)

\[
MSE = \frac{1}{N} \sum_{t=1}^{N} \left( y_t - \hat{y}_t \right)^2
\]

3. Mean Absolute Percent Error (MAPE)

\[
MAPE = \frac{1}{N} \sum_{t=1}^{N} \left| \frac{y_t - \hat{y}_t}{y_t} \right|
\]

4. Mean Square Percent Error (MSPE)

\[
MSPE = \frac{1}{N} \sum_{t=1}^{N} \left( \frac{y_t - \hat{y}_t}{y_t} \right)^2
\]

3. Experimental Analysis

In order to validate the performance of the proposed algorithm, the traffic information is recorded for 5 days in real practice application in this paper. The experimental urban crossroad is shown in Fig. 2. 720 data sets are prepared for the traffic forecasting procedure. 576 sample data of the first 4 days are used to train the prediction model, and the rest 144 samples are for test. A portion of the traffic flow time series is shown in Fig. 3. It can be seen for Fig. 3 that the traffic flow trend for each day is similar, and the external noise makes a slight variation in the day traffic change. Thus, to ensure the prediction accuracy, it is essential to figure out the influence of the nonlinear components caused by noise signals.

Figure 2. The traffic information record experiment.
As mentioned above, the EMD-Chaos-PSO-SVM is proposed to forecast the traffic flow. The original data is firstly decomposed into 6 IMFs by EMD. The time spectra of the IMFs are shown in Fig. 4. One can note from Fig. 4 that the first three IMF signals present the overall trend of the original traffic flow, and several other IMFs represent the uncertainty inference. The EMD decomposition can well identify the different characteristics from the original data, and hence benefit the traffic flow prediction through different SVM models.

Then, the Chaos-PSO-SVM is used to get the prediction component of each IMF, and thus their sum indicates the final short time traffic flow prediction value. In the optimization process, set $c_1 = c_2 = 1.5$, $\omega = 0.725$ for PSO, and the size of particle swarm selects 30. For Henon chaos, set, $c = 1.3$ and $d = 0.35$. The Chaos search performance of Henon is shown as Fig. 5. The Henon can enhance the efficiency of the Chaos optimization, and provide more stable optimal fitness value for the PSO optimization.

---

**Figure 3.** The original traffic flow time series.

**Figure 4.** The six IMFs of the original traffic data.

**Figure 5.** The chaotic sequences in [0 1] mapped by Henon.
Fig. 6 gives the performance of the proposed method for traffic flow forecasting. We have compared the performance of the proposed method with the independent use of SVM. One can note that the proposed method has higher prediction accuracy compared with the SVM.

The prediction performance of the proposed model and the EMD-PSO-SVM and EMD-SVM models is compared in Table 1. The comparison results show that the proposed method for short time traffic flow prediction is more effective than the EMD-PSO-SVM and EMD-SVM models. By the Chaos search processing, the local minimum is avoided and thus the forecasting error is decreased by 0.47% or better. One can note that the Chaos plays an effective role in the improvement of short time traffic flow prediction.

<table>
<thead>
<tr>
<th>Prediction model</th>
<th>Prediction performance (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MAE</td>
</tr>
<tr>
<td>EMD-Chaos-PSO-SVM</td>
<td>0.82</td>
</tr>
<tr>
<td>EMD-PSO-SVM</td>
<td>1.29</td>
</tr>
<tr>
<td>EMD-SVM</td>
<td>2.51</td>
</tr>
</tbody>
</table>

The prediction performance of the integrated model and independent models is compared in Table 2. It can be seen from Table 2 that by the EMD processing, the nonlinear elements are depressed and thus the forecasting error is decreased by 0.53% or better. Hence, the hybrid intelligent model can provide more efficient prediction rate for the traffic flow than the separated used of EMD and SVM. One can also note that the PSO optimization plays an important role in the improvement of the short time traffic flow prediction.

<table>
<thead>
<tr>
<th>Prediction model</th>
<th>Prediction performance (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>MAE</td>
</tr>
<tr>
<td>EMD-Chaos-PSO-SVM</td>
<td>0.82</td>
</tr>
<tr>
<td>Chaos-PSO-SVM</td>
<td>1.38</td>
</tr>
<tr>
<td>PSO-SVM</td>
<td>2.29</td>
</tr>
<tr>
<td>SVM</td>
<td>2.94</td>
</tr>
</tbody>
</table>

4. Conclusions

Intelligent Transportation management relies on precise traffic flow forecasting. It is necessary to employ advanced data mining approaches to excavate the hidden knowledge of the traffic data. This paper presents a new hybrid intelligent model for the short time traffic flow forecasting. This new
method combines the advantages of the nonlinear analysis of EMD and supervised learning of SVM to mine distinct and potential patterns of the traffic data. Moreover, the Chaos-PSO algorithm is applied to optimize the SVM parameters. The experimental test results have proven that the presented prediction approach is feasible and efficient for short time traffic flow forecasting. The prediction rate of the proposed EMD-Chaos-PSO-SVM is much better than the model without EMD processing and PSO processing. Thus, the proposed method has application importance.

5. Acknowledgement

This work was supported by the Natural Science Foundation of Henan Educational Committee (2011A520013), Education Reform Project of HIST (2010YB20). And we would like to thank Mr. Li for his good advice for this paper.

6. References


Yuesheng Gu: birth 1973, male, vice professor of Henan institute of science and technology. His current research area is computer network technology and artificial intelligence.