

Short-Term Traffic Flow Prediction Method Based on Dung Beetle Optimizer-Optimized LSTM Network

Zehua Zhang, Changke Shao

School of Emergency Management and Safety Engineering, North China University of Science and Technology, Tangshan 063000, China

Abstract: Short-term traffic flow is affected by various factors and usually presents strong nonlinear and stochastic characteristics, which makes accurate prediction difficult. Meanwhile, the selection of LSTM network parameters is often based on manual experience, and an inappropriate parameter setting may affect the prediction performance of the model. Therefore, this study proposes a Dung Beetle Optimizer-based LSTM (DBO-LSTM) model for short-term traffic flow prediction. The LSTM network is used to learn the temporal characteristics of traffic flow, and the DBO algorithm is applied to optimize important parameters of the LSTM network, including the learning rate and the number of hidden-layer neurons. In this way, the parameter selection process can be improved and the dependence on manual parameter adjustment can be reduced. Traffic flow data collected from a real road network are selected for model validation. The proposed DBO-LSTM model is compared with the conventional LSTM, PSO-LSTM, and WOA-LSTM models. The results show that the DBO algorithm can obtain a better parameter combination for the LSTM network. Compared with the three comparison models, the DBO-LSTM model obtains lower RMSE and MAPE values and has better overall prediction performance. The results verify the effectiveness of the proposed model for short-term traffic flow prediction and provide a reference for traffic congestion warning and traffic management in intelligent transportation systems.

Keywords: Short-term traffic flow prediction; Long Short-Term Memory (LSTM) network; Dung Beetle Optimizer (DBO); hyperparameter optimization.

1. Introduction

With the continuous acceleration of urbanization and the surge in the number of motor vehicles, urban road traffic congestion has become increasingly severe and has emerged as one of the core issues constraining the sustainable development of cities. Therefore, how to mine the dynamic evolution patterns from massive and highly volatile traffic data and construct high-precision prediction models has long been a research hotspot at the intersection of transportation engineering and artificial intelligence.

Early traffic flow prediction mainly relied on statistical models, the basic idea of which is to exploit the periodic patterns and autocorrelation characteristics of traffic flow series to infer future values through statistical analysis of historical data. Typical methods include the Historical Average (HA) method and the Autoregressive Moving Average (ARMA) model [1]. The HA method is easy to implement and has low computational cost, so it can be used for traffic conditions with relatively regular patterns. However, it is difficult for HA to describe sudden changes and random fluctuations in traffic flow. ARMA is mainly used for stationary time series, while ARIMA can also be applied to non-stationary series by means of differencing. Kumar et al. [2] used ARIMA to predict traffic flow during peak hours, while Chandra et al. [3] developed a Vector Autoregressive (VAR) model to improve the modeling of multiple traffic-related variables. To improve the adaptability of the conventional ARIMA model, Shahriari et al. [4] combined Bootstrap with ARIMA and developed an E-ARIMA model, which improved the prediction results. The Kalman filter is based on a state-space model and updates the estimated system state as new observations become available. It is therefore relatively suitable for traffic data containing noise. In general, statistical methods have a clear theoretical basis

and are relatively simple to implement. However, most of these methods focus mainly on temporal information and have difficulty describing the nonlinear and spatial characteristics of traffic flow. Their performance can therefore be limited under complicated traffic conditions [5].

With the increasing complexity of transportation systems, traditional statistical methods cannot fully describe the nonlinear characteristics of traffic flow. Machine learning methods have consequently been applied more widely because of their ability to fit complex data relationships. Representative methods include Support Vector Machine (SVM), Support Vector Regression (SVR), K-Nearest Neighbors (KNN), and Random Forest Regression (RFR) [9–10]. SVM uses kernel functions to map the original data into a higher-dimensional space and has shown good performance in small-sample and nonlinear problems. Toan [6] found that SVM performed better than several traditional methods under different traffic conditions. SVR extends SVM to regression problems and has been widely used in traffic flow prediction. Li et al. [7] demonstrated the applicability of SVR to traffic flow prediction with limited samples. Zou Zongmin et al. [8] introduced the Particle Swarm Optimization (PSO) algorithm to determine the parameters of SVR, resulting in improved prediction accuracy. Ju Dongxiang et al. [9] incorporated weather and traffic incident information into the prediction model and obtained better results for freeway traffic flow prediction. These studies indicate that machine learning methods have a stronger ability to deal with nonlinear relationships than traditional statistical models. However, their performance still depends to some extent on feature selection and feature construction. In addition, it is not easy for these methods to extract the spatiotemporal relationships and long-term changes directly from traffic flow data, which may limit their performance in more complicated and rapidly changing traffic conditions [10].

In recent years, deep learning methods have been increasingly used in traffic flow prediction. One important reason is that these models can learn useful features from traffic data with less reliance on manually constructed features. Compared with traditional methods, deep learning models can learn temporal and spatial information from the input data and are more suitable for traffic flow sequences with strong nonlinearity and non-stationarity [11]. LSTM and GRU introduce gating mechanisms into recurrent neural networks, which help preserve useful information over longer time intervals and reduce the influence of gradient vanishing. Harrou et al. [12] compared LSTM and GRU with traditional BP neural networks and found that both recurrent models achieved better prediction results. CNN-based models use convolution operations to extract local patterns in traffic sequences and can also be applied to multi-step prediction. In addition, models such as Graph WaveNet combine graph-based modeling with dilated convolution to describe the relationships between different traffic locations and their temporal changes, thereby improving the modeling of spatiotemporal dependencies [13].

Based on the existing studies, deep learning methods have shown good performance in short-term traffic flow prediction, but some issues remain to be considered. First, the performance of deep learning models is affected by the choice of hyperparameters, such as the learning rate and the number of hidden-layer neurons. In many studies, these parameters are still selected mainly through manual adjustment or repeated experiments, which can be time-consuming and may not result in the best parameter combination. Second, optimization algorithms such as PSO and WOA have been used to improve hyperparameter selection, but their optimization results may be affected by the algorithm settings and search process. Different parameter settings can lead to different optimization results, and the algorithms may also converge to local solutions. The Dung Beetle Optimizer (DBO) is a relatively recent population-based optimization algorithm with several search mechanisms and a balance between exploration and exploitation. It has been applied to different optimization problems and has shown good search performance. Therefore, this study introduces DBO into the LSTM model to optimize its key hyperparameters and develops a DBO-LSTM model for short-term traffic flow prediction. The aim is to obtain a more suitable parameter combination for LSTM and further improve its prediction performance under complex traffic conditions.

2. LSTM Neural Network

The Long Short-Term Memory (LSTM) network is a recurrent neural network developed to better handle sequential data. It introduces memory cells and gating mechanisms to control the storage and updating of information from previous time steps. Compared with the traditional RNN, LSTM can preserve useful historical information for a longer period and is better able to capture dependencies in sequential data. For this reason, LSTM has been widely applied to time series prediction and other sequence-related problems.

2.1. Basic Principles of LSTM

The key to LSTM lies in its memory cell, which regulates the cell state with the aid of a "gating" mechanism. The LSTM hidden layer typically contains three types of gate structures: the forget gate, the input gate, and the output gate, which are

respectively responsible for the selective forgetting, writing, and output of information, thereby achieving purposeful screening of the memory content.

3. Dung Beetle Optimizer

The Dung Beetle Optimizer (DBO) features few parameters, a simple structure, ease of implementation, strong global search capability, and fast convergence, making it suitable for continuous function optimization and engineering parameter optimization.

3.1. Basic Principles

DBO represents each candidate solution to the problem under consideration as a "dung beetle" in the search space. Different roles correspond to different position-update strategies, and global optimization is achieved through cooperative search. Among them, the ball-rolling behavior is mainly responsible for the global search of the population, the breeding behavior enhances population diversity, the foraging behavior improves local exploitation capability, and the stealing behavior is used to prevent the algorithm from falling into local optima. These various behaviors coordinate with one another, enabling DBO to balance global exploration and local exploitation during optimization, thereby improving the algorithm's optimization efficiency and stability.

3.2. DBO-LSTM Prediction Model

The prediction model is shown in Fig. 1.

4. Experiments and Result Analysis

4.1. Experimental Environment and Parameter Settings

In this paper, the LSTM neural network is adopted as the basic prediction architecture. Data from five consecutive historical time steps are selected as the model input to predict the traffic flow at the next time step; that is, the data from the preceding five time steps are used as input features and the flow at the next time step as the output target. Therefore, the numbers of neurons in the input and output layers are set to 5 and 1, respectively. In the parameter settings of the baseline LSTM model, the default number of hidden layer neurons is set to 35; the maximum number of training epochs is set to 500; and the initial learning rate is set to 0.001. To ensure stable convergence in the later stage of training, a learning rate decay strategy is introduced: after every 150 iterations, the learning rate is decayed by multiplying it by a scaling factor of 0.2.

To further overcome the subjectivity of manual parameter tuning and improve prediction accuracy, this paper introduces the Dung Beetle Optimizer (DBO) to perform global optimization of three core hyperparameters of the LSTM model, namely the number of hidden layer neurons, the maximum number of training epochs, and the initial learning rate. The optimization ranges of the three hyperparameters are set to [10, 50], [300, 1000], and [0.001, 0.01], respectively. The DBO population size is set to 30, and the maximum number of iterations is set to 50.

To verify the effectiveness of the proposed model, three models—LSTM, PSO-LSTM, and WOA-LSTM—are selected as comparison models, and traffic flow prediction experiments are conducted for each model.

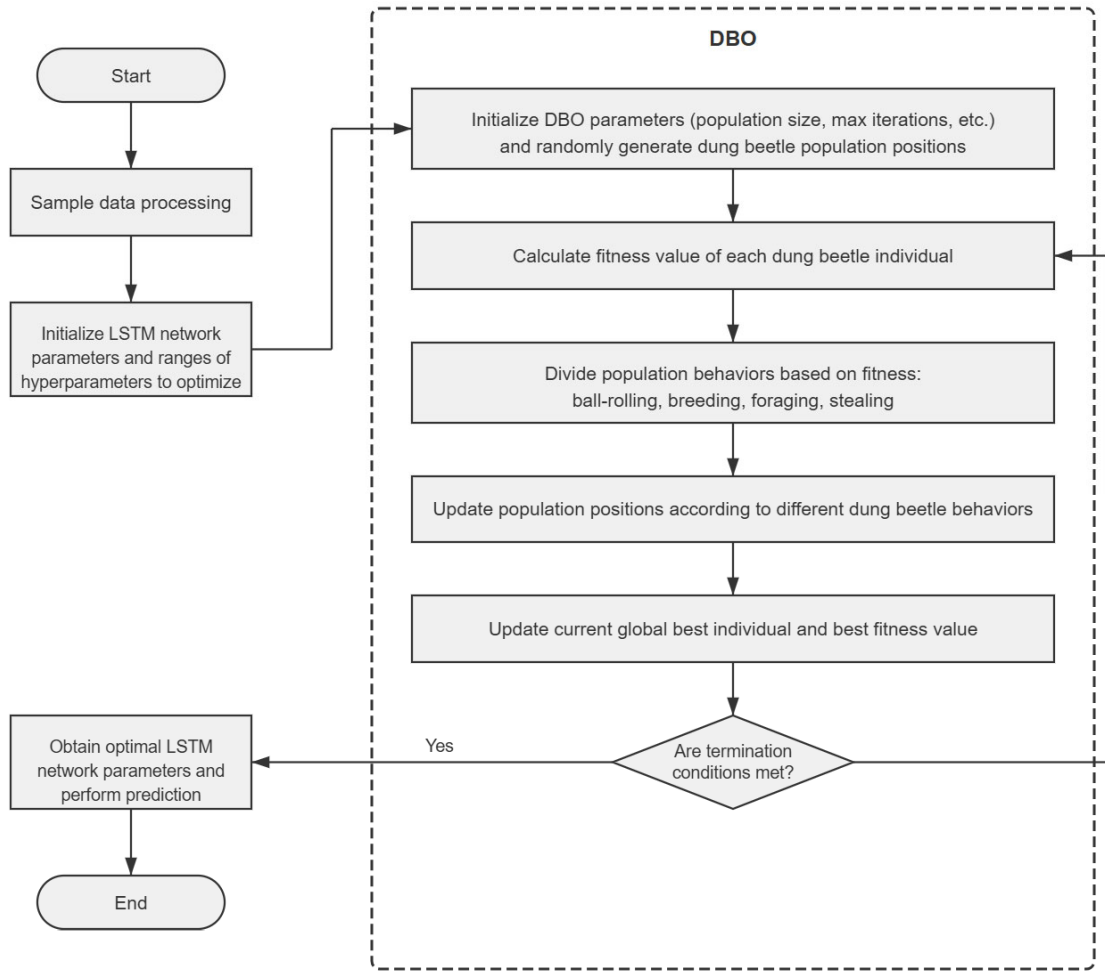


Figure 1. Flow chart of traffic flow prediction based on DBO-LSTM

4.2. Prediction Result Analysis

LSTM, PSO-LSTM, and SSA-LSTM are selected for comparative testing against the proposed model, and the

comparison curves of the prediction results of the four models against the actual monitored traffic flow values are finally obtained, as detailed in Fig. 2.

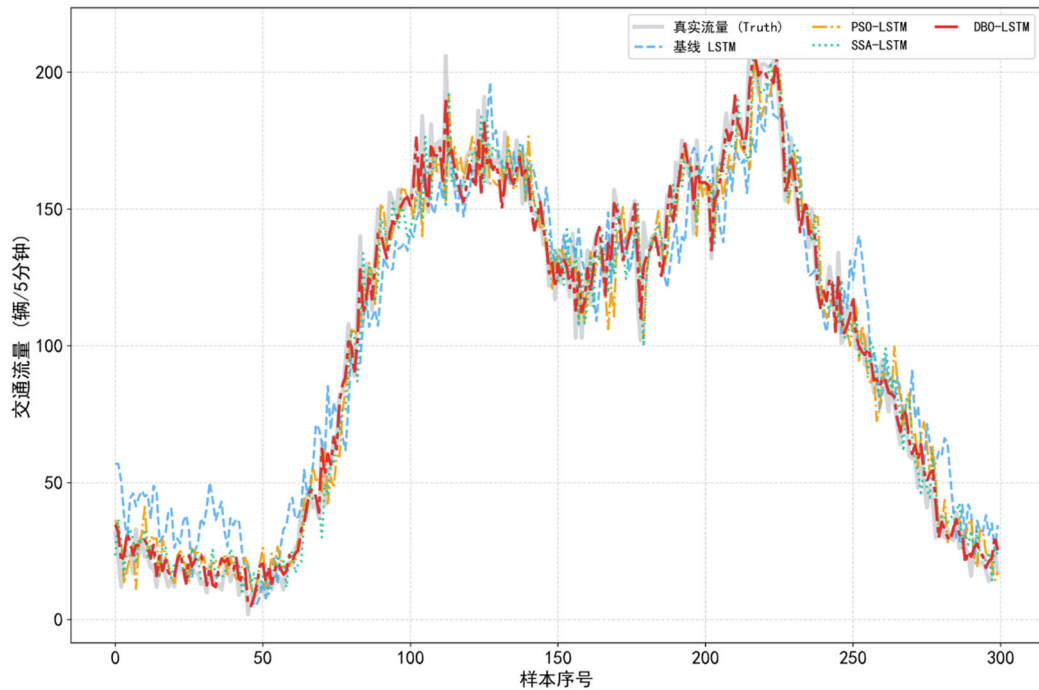


Figure 2. Comparison of prediction results of DBO-LSTM

From the analysis results in Fig. 2, it can be observed that both the standalone baseline LSTM model and the three hybrid models incorporating swarm intelligence optimization algorithms can satisfactorily fit the variation patterns of the traffic flow data; however, the prediction performance of the

three hybrid models is markedly superior to that of the standalone baseline LSTM model. Among the three hybrid models, the proposed DBO-LSTM model achieves the best prediction performance.

Table 3. Comparison of Evaluation Metrics for DBO-LSTM Prediction Results

Network Model	MAE	RMSE	MAPE (%)
LSTM	9.42	12.93	21.37
PSO-LSTM	7.68	10.45	17.25
SSA-LSTM	6.21	9.66	15.33

Table 3 details the various error evaluation metrics between the predicted and actually monitored values for the four models in the traffic flow prediction task. The analysis results show that, compared with the original LSTM, PSO-LSTM, and SSA-LSTM models, the proposed DBO-LSTM model achieves the best results across all three evaluation metrics—MAE, RMSE, and MAPE—indicating that the DBO algorithm can effectively enhance the prediction performance of the LSTM model.

5. Conclusion

In response to the problems of strong nonlinearity and high randomness in short-term traffic flow, as well as the substantial blindness of manual hyperparameter tuning and the tendency to fall into local optima in traditional LSTM networks, this paper proposes a DBO-LSTM short-term traffic flow prediction model based on the Dung Beetle Optimizer. The results show that, compared with the standalone LSTM, PSO-LSTM, and SSA-LSTM models, the proposed DBO-LSTM model exhibits superior fitting capability in terms of high- and low-frequency fluctuations, sudden peak-valley changes, and long-term trend tracking of traffic flow, and achieves optimal levels in error evaluation metrics such as RMSE and MAPE. Meanwhile, the DBO-LSTM model features faster convergence speed, higher prediction accuracy, and stronger robustness, providing reliable technical support for congestion early warning and refined traffic management and control in urban intelligent transportation systems.

References

- [1] Zhang, X., Zheng, J., Shen, Q., et al. (2023). Multi-channel spatio-temporal traffic flow prediction model based on hybrid graph convolution. *Telecommunications Science*, 39(9), 97-110.
- [2] Arunkumar, K., Kalaga, D. V., Kumar, C. M. S., et al. (2021). Forecasting the dynamics of cumulative COVID-19 cases (confirmed, recovered and deaths) for top-16 countries using statistical machine learning models: Auto-Regressive Integrated Moving Average (ARIMA) and Seasonal Auto-Regressive Integrated Moving Average (SARIMA). *Applied Soft Computing*, 103, 107161. <https://doi.org/10.1016/j.asoc.2021.107161>
- [3] Chandra, S. R., & Al-Deek, H. (2009). Predictions of freeway traffic speeds and volumes using vector autoregressive models. *Journal of Intelligent Transportation Systems*, 13(2), 53-72.
- [4] Shahriari, S., Ghasri, M., Sisson, S. A., et al. (2020). Ensemble of ARIMA: combining parametric and bootstrapping technique for traffic flow prediction. *Transportmetrica A: Transport Science*, 16(3), 1552-1573. <https://doi.org/10.1080/23249935.2020.1764662>
- [5] Zhou, F., Wang, S., Zhang, K., et al. (2024). Short-term traffic flow prediction model of urban road network based on graph Transformer network. *Science Technology and Engineering*, 24(10), 4307-4316.
- [6] Toan, T. D., & Truong, V. H. (2021). Support vector machine for short-term traffic flow prediction and improvement of its model training using nearest neighbor approach. *Transportation Research Record*, 2675(4), 362-373.
- [7] Li, C., & Xu, P. (2021). Application on traffic flow prediction of machine learning in intelligent transportation. *Neural Computing and Applications*, 33, 613-624. <https://doi.org/10.1007/s00521-020-05042-0>
- [8] Zou, Z., Hao, L., Li, Q., et al. (2021). Short-term highway traffic flow prediction based on particle swarm optimization-support vector regression. *Science Technology and Engineering*, 21(12), 5118-5123.
- [9] Ju, D., Yuan, H., Dong, W., et al. (2025). Construction of short-term traffic flow prediction platform based on random forest algorithm. *China ITS Journal*, (3), 94-97+115.
- [10] Deng, H., Chen, H., Xiao, Q., et al. (2024). Traffic flow prediction based on multi-head attention dynamic graph convolutional network. *Journal of Taiyuan University of Technology*, 55(1), 172-183.
- [11] Yao, J., He, R., Shi, T., et al. (2023). Review of traffic flow prediction methods based on machine learning. *Journal of Traffic and Transportation Engineering*, 23(3), 44-67.
- [12] Harrou, F., Zeroual, A., Kadri, F., et al. (2024). Enhancing road traffic flow prediction with improved deep learning using wavelet transforms. *Results in Engineering*, 23, 102342. <https://doi.org/10.1016/j.rineng.2024.102342>
- [13] Tian, C., & Chan, W. K. (2021). Spatial-temporal attention wavenet: A deep learning framework for traffic prediction considering spatial-temporal dependencies. *IET Intelligent Transport Systems*, 15(4), 549-561. <https://doi.org/10.1049/itr2.12074>