CHAOTIC GRAVITATIONAL SEARCH ALGORITHM AND LONG SHORT TERM MEMORY MODEL FOR TRAFFIC FLOW PREDICTION

Abstract

Real-time highway traffic flow prediction remains a crucial area of research in intelligent transportation, particularly when dealing with congestion, as it enables efficient control and management of traffic flow. Furthermore, the urban road network's irregular nature and large short-term variations make it difficult to estimate traffic flow effectively. Therefore, this study provides a novel approach that improves the accuracy of short-term traffic flow predictions by combining a chaotic theory, Gravitational Search Algorithm (GSA) and Long Short-Term Memory (LSTM) neural network. A modified version of the Gravitational Search Algorithm (GSA) that includes chaos theory concepts is known as the Chaotic Gravitational Search Algorithm (CGSA). The application of chaotic theory to the GSA technique aims to achieve a better balance between the exploration and exploitation phases. The chaotic GSA strategy, which reduces the impact of random hyperparameter selection on the prediction performance, is incorporated into the LSTM model in order to improve prediction accuracy. Analysis of the study's results found that, in terms of evaluation metrics, the proposed model was superior to other baseline models.

Keywords—Intelligent transportation system; Traffic flow prediction; Gravitational Search Algorithm; Long short term memory

Authors

Bharti Naheliya
Department of Mathematics
Maharshi Dayanand University
Rohtak, India
bhartinaheliya@gmail.com

Poonam Redhu
Department of Mathematics
Maharshi Dayanand University
Rohtak, India
poonamr.maths@mdurohtak.ac.in

Kranti Kumar
Department of Mathematics
Central University of Himachal Pradesh, Dharamshala
Shahpur Campus-176206, India
kranti@aud.ac.in

I. INTRODUCTION

Nowadays one of the major problems that almost all cities face is traffic congestion. The combined effect of inadequate road infrastructure and an increase in the number of registered vehicles causes traffic congestion to rise significantly [1]. Due to the inadequacy of the transportation infrastructure, particularly in India, to accommodate growth in the number of vehicles, there is traffic congestion during rush hours. The construction of more road infrastructure is a major requirement to meet the demand for traffic flow, however, this is not a practical solution to the problem, primarily due to financial constraints. Implementing an intelligent transport system (ITS) is thus the alternative strategy for alleviating traffic congestion. Research on a national and international scale is becoming interested in short-term traffic forecasting, which is a crucial component of an intelligent transportation system.

Reliable real-time traffic flow forecasting, which is crucial for the execution of guidance and traffic management, aims to increase traffic operations efficiency and decrease traffic congestion. Due to its ability to minimize travel time and prevent additional financial losses, an effective solution to traffic congestion has drawn a lot of interest from corporations, government organizations, and individual drivers. Despite this, because traffic flow is periodic and external factors have an unpredictable impact on how it behaves, it is still difficult to precisely predict. The approaches proposed for forecasting short-term traffic flow can be divided into three categories: parametric methods, non-parametric and methods deep learning methods. The parametric method is a modeling technique in which the model's parameters can be adjusted using actual data on traffic flow, but the structure of the model is pre-determined based on theoretical considerations. A parametric approach was suggested by Levin and Tsao [2] to predict the morning rush hour traffic on a motorway. According to their research, the Autoregressive Integrated Moving Average (ARIMA) (0,1,1) model presented statistical significance. By describing the fluctuating traffic flow as a linear dynamical system, the dominant dynamic model, Kalman filtering, was used to anticipate it and enable continuous monitoring of traffic changes [3]. Additionally, by utilizing complex and nonlinear features, researchers used machine learning to estimate traffic flow [4, 5]. As an example, researchers developed a two-step method using the principal component to enhance KNN's prediction ability in estimating short-term traffic volume [6]. Additionally, Cai et al. [7] proposed that the support vector machine regression model improved using the GSA technique exhibits superior efficacy to the original support vector machine for forecasting traffic flow.

On the other hand, deep learning has recently gained a lot of interest for its capacity to analyze complex and nonlinear traffic patterns. The stacked autoencoder (SAE) is a deep architectural model for predicting traffic flow that was developed by Lv et al. [8]. This novel model successfully isolates the nonlinear underlying spatial and temporal correlations found in traffic data. The restricted Boltzmann machines (RBMs) were used by the deep belief network (DBN) [9] to generate the input traffic data. After that, it has been demonstrated that the recurrent neural network (RNN) is incredibly successful in forecasting traffic flow [10]. In order to handle long-term dependencies and mitigate the problems associated with gradient explosion/vanishing, long short-term memory (LSTM) [11], a type of recurrent neural network, was developed. For precise traffic flow estimation, Lu et al. [12] presented a deep learning network that incorporates multi-diffusion convolution and LSTM. A novel deep learning approach, known as the Time-Dependent Attention Convolutional LSTM [13], was developed. This method incorporates a time-dependent attention mechanism, enabling it to learn the similarities of historical traffic flows across various time intervals. Deep learning models still have problems, like unpredictable parameter initialization and long processing times [14]. As a result, adjusting the parameters and improving the performance of deep learning models becomes necessary. However, manual parameter selection in these models makes it challenging to quickly identify the appropriate parameters that would produce the highest prediction performance. By applying a genetic algorithm (GA) to optimize it, a deep belief network (DBN) model was established and improved for forecasting traffic flow [15]. To improve the accuracy of short-term traffic flow forecasts, researchers created a Gravitational Search Algorithm combined with an Extreme Learning Machine (GSA-ELM) technique [16]. The PSO-Bi-LSTM model [17], an innovative approach, was developed for predicting short-term traffic flow and this model combines the capabilities of a bidirectional long short-term memory (Bi-LSTM) neural network and particle swarm optimization (PSO). The GSA algorithm usually encounters local optima, which generally do not have an efficient memory mechanism to quickly arrive at the ideal parameter combination. Researchers have looked into combining a chaotic approach with the GSA technique as a way to get around this problem [18, 19]. CGSA is an improved version of the well-known GSA technique that incorporates chaotic behavior to improve search speed. Through the use of chaotic maps, the GSA technique's velocity is modified in this revision.

In this article, the CGSA-LSTM traffic flow forecasting model is suggested. For optimization, the model incorporates a chaotic gravitational search algorithm with long short-term memory (LSTM). The accuracy of traffic flow predictions and the reduction in forecast time for the CGSA-LSTM model are both significantly improved. Our contributions to this work are outlined in the list below:

- For the purpose of predicting short-term traffic speed, a novel model based on the combination of CPSO and Bi-LSTM neural network has been presented.
- To improve the PSO algorithm's convergence performance, chaotic theory has been used.
- CGSA automatically optimizes and determines the hyperparameter values for the LSTM model.
- Appropriate experiments are conducted, comparing the GSA-LSTM model with other models, to demonstrate its superior performance.

I. MATERIAL METHODS

1. **Gravitational Search Algorithm:** In 2009, Rashedi et al. [20] initially introduced the gravitational search algorithm (GSA). In this approach, the ideal solution to the problem is considered as a collection of particles going through space. Consequently, due to the gravitational force attracting all particles to one another, they collectively move toward the objects with higher masses. The GSA exhibits superior global search capabilities, making it suitable for enhancing the model's parameters [21]. The complete procedure of the GSA is shown in Fig. 1.

   This algorithm's formulation is as follows:

The position of the k^{th} particle in a D-dimensional search space with an initial group of P particles is given by

$$U_k = u_k^1, u_k^2, ..., u_k^d, ..., u_k^D,\ k = 1, 2, ..., P \tag{1}$$

Suppose the k^{th} particle's inertial mass at time t is

$$mass_k(t) = \frac{f_k(t) - w_{wt}(t)}{b_{bt}(t) - w_{wt}(t)} \tag{2}$$

$$M_k(t) = \frac{mass_k(t)}{\sum_{l=1}^{P} mass_l(t)} \tag{3}$$

where $f_k(t)$ is the fitness value for the k^{th} particle at time t, $b_{bt}(t)$ and $w_{wt}(t)$ indicate the best and worst fitness values, respectively of the whole particle population at time t.

Equations (4) and (5) provide the best and worst fitness values corresponding to the least optimization issue.

$$b_{bt} = \min_{l \in 1,2,...,N} f_l(t) \tag{4}$$

$$w_{wt} = \max_{l \in 1,2,...,N} f_l(t) \tag{5}$$

The gravitational force between the l^{th} and k^{th} particles in d-dimensional space at time t is given as

$$F_{kl}^d = G(t)\frac{M_k(t) \times M_l(t)}{E_{kl}(t) + \gamma}(u_l(t) - u_k(t)) \tag{6}$$

In this formula, $u_l(t)$ and $u_k(t)$ are the position of l^{th} and k^{th} particles in d-dimensional space at time t respectively, γ is a small constant value, $E_{kl}(t)$ is Euclidean distance between particle k and particle l at time t, and $G(t)$ denotes the gravitational constant at time t.

The formula for the gravitational constant is

$$G(t) = G_0 e^{-\tau\frac{t}{\Gamma}} \tag{7}$$

where τ is the mutually adjusted constant, G_0 indicates the initial value of, and Γ signifies the maximum iterations.

Assuming that in GSA, the total acting force of the k^{th} particle is given as the sum of the acting forces given in Eqn. (6) of all other particles at time t.

$$F_k^d = \sum_{l=1, l\neq k}^{P} R_l F_{kl}^d(t) \tag{8}$$

where R_l is random number from [0, 1].

When a particle accelerates in d-dimensional space at time t, it is defined as

$$acc_k^d(t) = \frac{F_k^d(t)}{M_k(t)} \tag{9}$$

The following are the descriptions of the equations for the particle's updated position and speed.

$$v_k^d(t+1) = R_k \times v_k^d(t) + acc_k^d(t) \tag{10}$$

$$u_k^d(t+1) = u_k^d(t) + v_k^d(t+1) \tag{11}$$

where R_k is a random number drawn at random from the range [0, 1] and $v_k^d(t+1)$ is the velocity of the kth particle in d-dimensional space at time t+1.

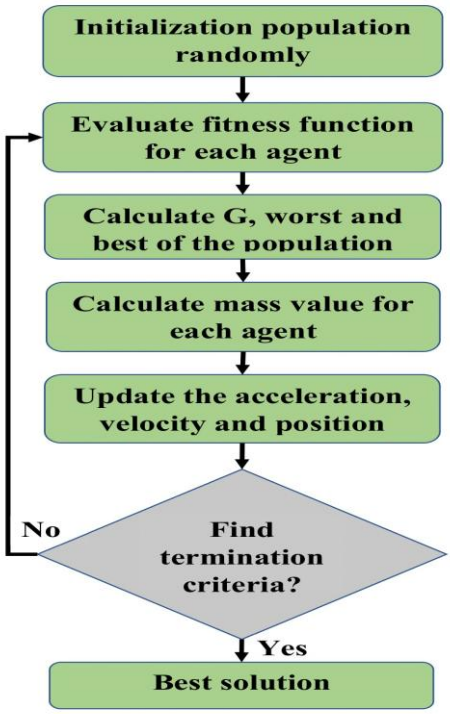

Figure 1: The Flow Chart of GSA

2. **Chaotic Gravitational Search Algorithm:** Chaos is a nonlinear dynamical system that is extremely sensitive to its initial conditions. Although chaotic systems appear to operate randomly, chaos is not always required for a system to exhibit chaotic behaviors. The effectiveness of meta-heuristic population-based optimization methods has recently been enhanced by using chaotic theory.

In present study, we used a chaotic map with gravitational constant to improve the performance of GSA. The gravitational constant (G) determines the strength of the overall gravitational forces between the search agents, as can be demonstrated in Eqs. 6 and 7. This variable, which serves as the primary regulating parameter, decides how the search agents will travel. To better balance exploration and exploitation in each iteration, from the beginning to the last, our fundamental aim is to introduce systematic adjustments to this main regulating parameter of GSA. The algorithm now alternates between exploration and exploitation actions as the value of G gradually decreases over the iteration period. It should be identified that a novel approach's adaptability is essential because metaheuristics are stochastic in nature. We are not changing the gravitational constant's adaptive behavior, which is the primary determinant of the robustness of the chaotic approach used. We induced chaotic changes when the parameter G decreased adaptively. Because of this, search agents must continue looking over the search until the last iteration.

3. **Long Short-Term Memory:** Long Short-Term Memory (LSTM) is a particular type of Recurrent Neural Network (RNN) constructed to overcome some of the drawbacks of conventional RNN. Regular RNNs have issues with exploding and vanishing gradients, which makes it challenging to learn long-range relations in sequential data. By including memory units, LSTM networks solve this kind of problem and the network learns when to update memories and when to ignore previous ones.

We represent a multivariate time series containing K variables with a length of L as $X = (x_1, ..., x_L)$, where x_l denotes the set of observations for all variables at time step t. The memory cell c_t in the LSTM model, which retains data at time step t on the observations obtained in this step, is its primary contribution. Multiple gates operate on the cell, which can either reset the value or keep it depending on the gate's state. Mainly, three gates are used to determine whether the current cell value is to be forgotten (forget gate f_t), read from its input (input gate i_t), and output the new cell value (output gate o_t); in addition, an input modulation gate with the name $\hat{c}_t$ is included. Figure 2 depicts the LSTM network's structural architecture. The following equations correspond to the gates, cell update, and output.

$$i_t = \sigma(U_i \otimes h_{t-1} + V_i \otimes x_t + b_i) \tag{12}$$

$$f_t = \sigma(U_f \otimes h_{t-1} + V_f \otimes x_t + b_f) \tag{13}$$

$$o_t = \sigma(U_o \otimes h_{t-1} + V_o \otimes x_t + b_o) \tag{14}$$

$$\hat{c}_t = \psi(U_c \otimes h_{t-1} + V_c \otimes x_t + b_c) \tag{15}$$

$$c_t = f_t \otimes c_{t-1} + i_t \otimes \hat{c}_t \tag{16}$$

$$h_t = o_t \otimes \psi(c_t) \tag{17}$$

where $\otimes$ is used to denote the product operation and the V and b matrices are used to represent the network parameters. The LSTM networks are well-trained because their gates can handle exploding/vanishing gradients. The sigmoid $\sigma(\cdot)$ and hyperbolic tangent $\psi(\cdot)$ represent the nonlinearities, and h_t is the hidden state.

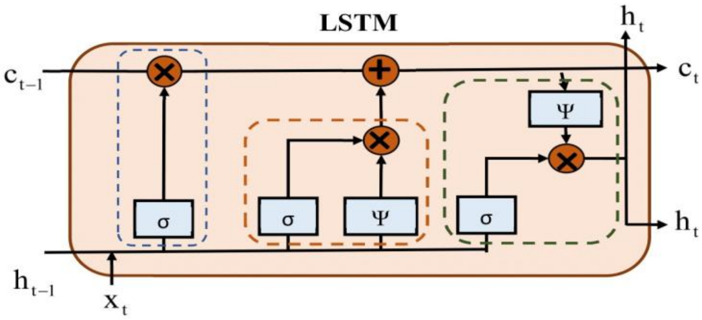

Figure 2: The Structure of the LSTM Neural Network

4. **Chaotic GSA-LSTM:** To improve the predictability of traffic flow data, the "Chaotic GSA-LSTM" model for neural network architecture incorporates the Long Short-Term Memory (LSTM) model with the Gravitational Search Algorithm (GSA) and chaotic dynamics. The inclusion of chaotic dynamics promotes more exploration during training and the application of GSA optimizes the LSTM's parameters, both of which are potentially helpful for capturing complicated and unpredictable patterns in traffic flow. Figure 3 illustrates the proposed model's process.

   The following are the main steps in developing and utilizing the Chaotic GSA-LSTM model for traffic flow prediction:

   - From the data that has been collected, make two sets: one for testing and the other for training.
   - Initialize the LSTM model and CGSA algorithm's settings.
   - The initial fitness function of each particle is then calculated.
   - Particles are continuously updated to their own and global optimum locations.
   - Test or validate the trained Chaotic GSA-LSTM model using a separate dataset.

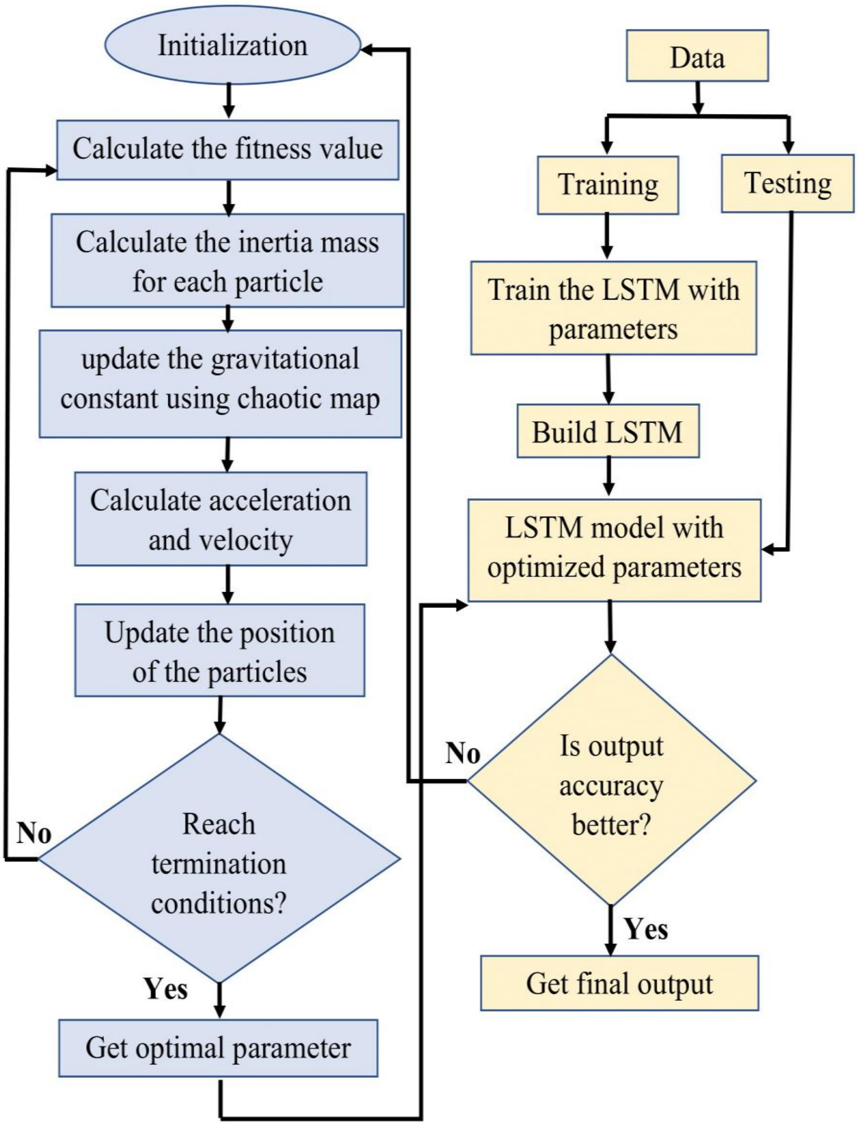

Figure 3: The Structure of the Proposed Model.

II. RESULTS AND DISCUSSION

1. **Data Description:** Video cameras were used to collect the data for this study at the Inner Ring Road, South Extension-II, Delhi, India from February 24 through February 28, 2020. Data from both sides of the road was captured between the hours of 8 am and 11 am and 4:30 pm and 7 pm. 678 distinct sets of traffic data were extracted from the recorded videos in total. Training and testing data sets were selected from these 678 data sets at random.

2. **Evaluation Criteria:** The performance of the proposed model's predictions is assessed in this article using three performance metrics. The following is a description of these assessments:

$$\text{RMSE (Root Mean Square Error)} = \sqrt{\frac{1}{n}\sum (Y_P - Y_A)^2} \tag{18}$$

$$\text{MAPE (Mean Absolute Percentage Error)} = \frac{1}{n}\sum \frac{|Y_P - Y_A|}{Y_A} \tag{19}$$

$$\text{r (Correlation Coeficient)} = \frac{n\sum (Y_P * Y_A) - \sum (Y_A) * \sum (Y_P)}{\sqrt{(n\sum (Y_A)^2 - (\sum (Y_A))^2)} * \sqrt{(n\sum (Y_P)^2 - (\sum (Y_P))^2)}} \tag{20}$$

where Y_P and Y_A represent predicted and actual values respectively and n is the number of data sets.

3. **Analysis of Compared Models:** The effectiveness of the proposed CGSA-LSTM model is assessed using a number of methodologies, including machine learning and deep learning techniques, as benchmarks. For predicting short-term traffic flow, the designed CGSA-LSTM model has been compared to various neural network-based models such as GSA-LSTM, LSTM, and WNN. The evaluation indices for each prediction model are calculated and the results are displayed in Table 1 to evaluate the model's prediction accuracy. The results of the formulas (18), (19), and (20) show that the performance of the forecasting model improves when RMSE, and MAPE decreases and MAPE increases, along with r.

   Moreover, Figure 4 graphically displays the discrepancy between the short-term traffic flow forecast model's output and the actual traffic flow. All of the graphs in this figure were created using MATLAB R2022a. When compared to other chosen models, Figure 4 demonstrates how well the proposed CGSA-LSTM model performs. The proposed model's predicted traffic flow is more reliable, close to the real traffic flow, and more accurate at fitting nonlinear curves and it also converges more quickly. As a conclusion, we can conclude that when compared to existing models, this proposed CGSA-LSTM model shows improved prediction accuracy in projecting short-term traffic flow.

Table 1: Analysis of the Evaluation Observations for the Various Models.

Models	RMSE	MAPE	r
WNN	102.7652	0.1799	0.9749
LSTM	96.881	0.1639	0.9766
GSA-LSTM	86.7497	0.1596	0.9811
CGSA-LSTM	**85.2154**	**0.1503**	**0.9823**

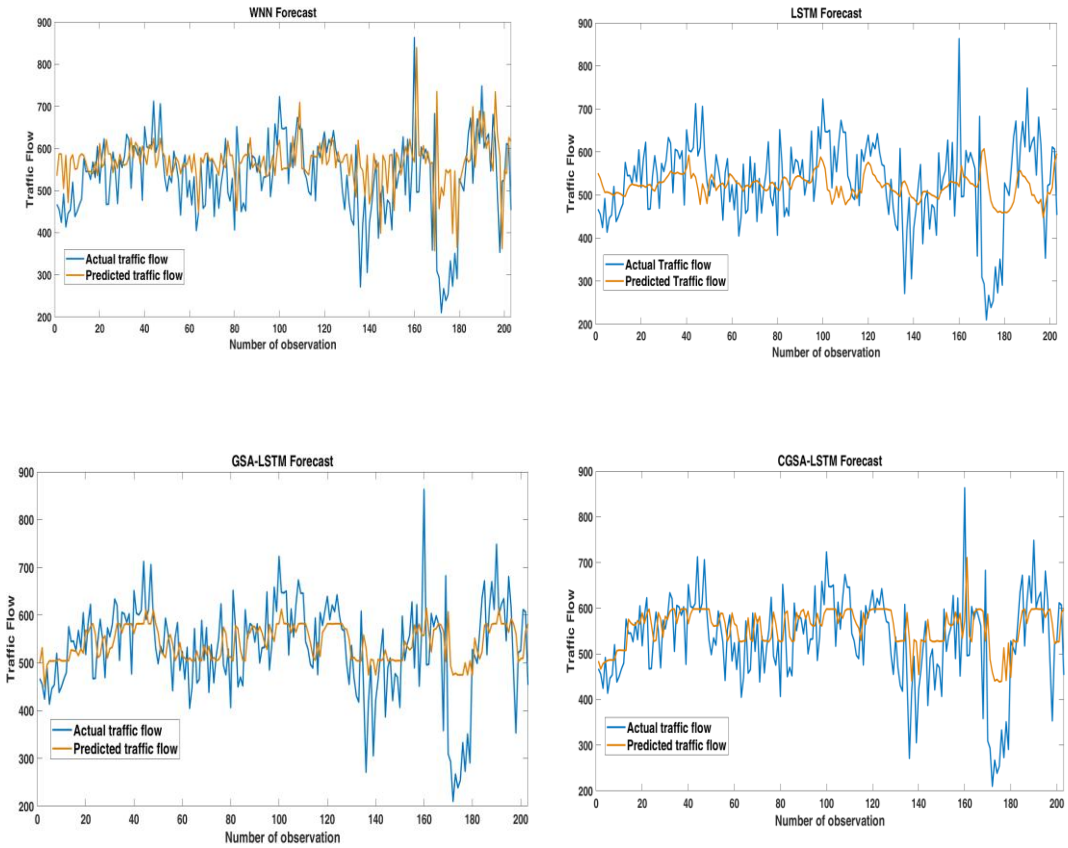

Figure 4: Prediction Performance of the Compared Models.

III. CONCLUSION

For estimating short-term traffic flow, a chaotic gravitational search algorithm technique-optimized LSTM structure is developed. The potential to better balance searching locally and globally is discovered to make the CGSA algorithm converge more precisely and fast than GSA. The CGSA approach is an optimization method that combines the GSA's gravitational principles with chaotic dynamics to increase exploration potential and perhaps raise the conventional quality of the discovered solutions. Mainly, CGSA technique is used to optimize the parameters of the LSTM neural network in this study. In terms of performance measures (MAPE, RMSE, and r), the proposed CGSA-LSTM model outperforms other selected baseline models and has more significant predictive abilities. Future research will demonstrate the feasibility of the proposed approach, which combines chaos theory with more effective optimization methods than GSA, and can be applied to new relevant application domains.

REFERENCES

[1] S. A., Gamel, A. I., Saleh, and H. A., Ali, "Machine learning-based traffic management techniques for intelligent transportation system," Nile journal of communication and computer science, vol. 1, pp. 9-18, 2021.
[2] M., Levin, and Y. D., Tsao. "On forecasting freeway occupancies and volumes (abridgment)," Transportation Research Record, 1980.

[3] L., Cai, Z., Zhang, J., Yang, Y., Yu, T., Zhou, and J., Qin, "A noise-immune Kalman filter for short-term traffic flow forecasting," Physica A: Statistical Mechanics and its Applications, vol. 536, pp. 122601, 2019.
[4] I. O., Olayode, A., Severino, T., Campisi, and L. K., Tartibu, "Prediction of vehicular traffic flow using levenberg-marquardt artificial neural network model: italy road transportation system," Komunikácie, vol. 24, 2022.
[5] A., Navarro-Espinoza, O. R., López-Bonilla, E. E., García-Guerrero, E., Tlelo-Cuautle, D., López-Mancilla, C., Hernández-Mejía, and E., Inzunza-González, "Traffic flow prediction for smart traffic lights using machine learning algorithms," Technologies, vol. 10, pp. 5, 2022.
[6] Z., Zheng, and D., Su, "Short-term traffic volume forecasting: A k-nearest neighbor approach enhanced by constrained linearly sewing principle component algorithm," Transportation Research Part C: Emerging Technologies, vol. 43, pp. 143-157, 2014.
[7] L., Cai, Q., Chen, W., Cai, X., Xu, T., Zhou, and J., Qin, "SVRGSA: a hybrid learning-based model for short-term traffic flow forecasting," IET Intelligent Transport Systems, vol. 13, pp. 1348-1355, 2019.
[8] Y., Lv, Y., Duan, W., Kang, Z., Li, and F. Y., Wang, "Traffic flow prediction with big data: A deep learning approach," IEEE Transactions on Intelligent Transportation Systems, vol. 16, pp. 865-873, 2014.
[9] W., Huang, G., Song, H., Hong, and K., Xie, "Deep architecture for traffic flow prediction: Deep belief networks with multitask learning," IEEE Transactions on Intelligent Transportation Systems, vol. 15, pp. 2191-2201, 2014.
[10] H., Zhu, Y., Xie, W., He, C., Sun, K., Zhu, G., Zhou, and N., Ma, "A novel traffic flow forecasting method based on RNN-GCN and BRB." Journal of Advanced Transportation, vol. 2020, pp. 1-11, 2020.
[11] B., Yang, S., Sun, J., Li, X., Lin, and Y., Tian, "Traffic flow prediction using LSTM with feature enhancement," Neurocomputing, vol. 332, pp. 320-327, 2019.
[12] H., Lu, D., Huang, Y., Song, D., Jiang, T., Zhou, and J., Qin, "St-trafficnet: A spatial-temporal deep learning network for traffic forecasting," Electronics, vol. 9, pp. 1474, 2020.
[13] X., Huang, J., Tang, X., Yang, and L., Xiong, "A time-dependent attention convolutional LSTM method for traffic flow prediction," Applied Intelligence, vol. 52, pp. 17371-17386, 2022.
[14] Z., Tian, and H., Chen, "A novel decomposition-ensemble prediction model for ultra-short-term wind speed," Energy Conversion and Management, vol. 248, pp. 114775, 2021.
[15] Y., Zhang, and G., Huang, "Traffic flow prediction model based on deep belief network and genetic algorithm," IET Intelligent Transport Systems, vol. 12, pp. 533-541, 2018.
[16] Z., Cui, B., Huang, H., Dou, G., Tan, S., Zheng, and T., Zhou, "GSA-ELM: A hybrid learning model for short-term traffic flow forecasting," IET Intelligent Transport Systems, vol. 16, pp. 41-52, 2022.
[17] Bharti., P., Redhu, and K., Kumar, "Short-term traffic flow prediction based on optimized deep learning neural network: PSO-Bi-LSTM," Physica A: Statistical Mechanics and its Applications, pp. 129001, 2023.
[18] S., Mirjalili, and A. H., Gandomi, "Chaotic gravitational constants for the gravitational search algorithm," Applied soft computing, vol. 53, pp. 407-419, 2017.
[19] J., Ji, S., Gao, S., Wang, Y., Tang, H., Yu, and Y., Todo, "Self-adaptive gravitational search algorithm with a modified chaotic local search," IEEE Access, vol. 5, pp. 17881-17895, 2017.
[20] E., Rashedi, H., Nezamabadi-Pour, and S., Saryazdi, "GSA: a gravitational search algorithm," Information sciences, vol. 179, pp. 2232-2248, 2009.
[21] D., Pelusi, R., Mascella, L., Tallini, J., Nayak, B., Naik, and Y., Deng, "Improving exploration and exploitation via a hyperbolic gravitational search algorithm" Knowledge-Based Systems, vol. 193, pp. 105404, 2020.
[22] S. H., Kellert, "In the wake of chaos: Unpredictable order in dynamical systems" University of Chicago Press, 1993.

