International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

NxLFTNet_BGSO: NARX LSTM Forward Taylor Network Enabled Al-Biruni Group Search Optimization for Traffic Forecasting

Sathyanarayana Mangali [1]*, Jayasree Hanumantha Rao [2]

Submitted:14/03/2024 **Revised**: 29/04/2024 **Accepted**: 06/05/2024

Abstract: Traffic signals utilized to forecast are commonly produced by sensors next to the roads that may be indicated as nodes on a graph. Normally, these sensors generate common signals indicating traffic flows and abnormal signals represent unknown traffic disruptions. Graph convolution networks are broadly deployed for traffic forecasting owing to the capability for capturing the relation amid network nodes. Nevertheless, the task is difficult owing to the expected intricacy and improbability of traffic patterns. To address this shortcoming, a traffic forecasting model is established with Nonlinear Autoregressive models with exogenous inputs Long Short Term Memory Forward Taylor Network (NxLFTNet) enabled Al-Biruni Group search optimization (BGSO_NxLFTNet). The input traffic network is given to spatio-temporal embedding (STE) generator for identifying daily and weekly embedding of time corresponding to current traffic signal. The outcome of input traffic network and STE generator is subjected to traffic detection, which is executed by employing NxLFTNet trained by BGSO. Here, NxLFTNet is combined by NARX and LSTM; also BGSO is incorporated by Group Search Optimizer (GSO) and Al-Biruni Earth Radius (BER). The measures taken for BGSO_NxLFTNet such as, Mean absolute percentage error (MAPE), Root Mean square error (RMSE), and Mean Absolute Error (MAE) gained 0.001, 0.010 and 0.011.

Keywords: *Traffic forecasting, Spatio-Temporal Embedding (STE) generator, Long Short Term Memory (LSTM), Group Search Optimizer (GSO), Al-Biruni Earth Radius (BER).*

1. Introduction

Traffic plays an important role in people's daily life. In this situation, an accurate real world prediction of traffic environment is of foremost important. In this condition, precise real time prediction of traffic environment is of enormous significance for road users, governments and private fields. Broadly employed transportation services, like navigation, route planning and flow control is also relied on extreme quality traffic state environment. In the traffic survey, basic variables of traffic flow, that is density, volume, and speed are usually taken as indicators to observe the present status of traffic environment [1] and to forecast the aftertime. According to the length of detection, traffic forecast is broadly categorized into two namely, short-term and long-term. Most common statistical paradigms are capable of performing well on short interval prediction. Nevertheless, by the reason of uncertainty and complication of traffic flow, those mechanisms are less efficient for comparatively long-term predictions. Former studies on mid-and-long term traffic prediction are broadly divided such as dynamical modeling as well as data driven approaches. Dynamical modeling utilizes arithmetical tools as well as physical knowledge to generate traffic issues through computational simulation [2]. By improving the precision of traffic flow prediction, it can be able to provide advantage towards intelligent traffic management on congestion deduction as well as traffic efficacy [3] [4]. Traffic flow forecasting is somewhat a difficult technique in other words it is exaggerated by several components, like traffic patterns [5]. With the progress of intelligent transportation systems (ITSs), forecasting of traffic has attained more interest henceforth precise traffic forecasting plays a significant part in numerous extensive ITSs, comprising navigation system [6], route guidance system [7], and traffic signal control system [8].

The challenge has been considered over several decades across diverse societies starting from traffic with economics to data mining [9]. As ordinary situation traffic models are simple to forecast, an open evaluation in traffic prediction is to predict the traffic in utmost situations [9]. In latest traffic flow forecasting models, investigators have to choose these traffic aspects as well as modeling factors emerging from the acquired data in accordance with several basic presumptions implemented in the old works. In other words, the accuracy amidst these presumptions maybe significantly impact on prediction correctness. Nevertheless, as long as the proportion of data remains reasonably large, the unrevealed numerical data as well as the possible relationships will be exposed by the sets of data themselves. Thereby, in the case of a significant quantity of traffic information is taken on, it is capable of avoiding more failures created by the presumptions as well as the precision of traffic forecast could be enhanced through understanding the data as well as the relations covered across the

[1*]*Research Scholar, Computer Science and Engineering Department, Osmania University, India*
[2]*Professor, Information Technology Department, Maturi Venkata Subba Rao (MVSR) Engineering College, Osmania University, India*

information [5]. Computational forecasting methods, like Bayesian modeling [10], fuzzy logic, neural networks (NNs) [11], statistical modeling, as well as hybrid modeling [12] , has been extensively utilized in traffic flow prediction. Such methods, especially NNs, are validated to be beneficial in finding the numerical principles covered among the traffic information as well as the later traffic flow. Despite that, majority of them represent the shallow traffic patterns as well as the calculated prediction outcomes ought to be more precise. In addition to the improvements in sensing technologies, as well as widespread traffic sensing facilities, the proportion of gathered traffic flow information could be significantly improved [5].

Various Machine-Learning (ML)-based methods were brought into play in the traffic count forecast issue [13]. The notion of Deep Learning (DL) is to make use of deep architectures to take out as well as change the intrinsic characteristics across the information from the bottom to the top level, also each uninterrupted layer utilizes the former layer output as the input [14] [6]. Since traffic flow forecast is a difficult technique, DL associated techniques could symbolize traffic characteristics from the acquired information with no previous foreknowledge, which might significantly advance the prediction truthfulness [5]. DL approaches give self-regulating illustration studying from primary information, substantially decreasing the endeavor of hand-crafted feature engineering. For traffic prediction, premature efforts involves stacked denoising autoencoder, Deep belief network (DBN), as well as stacked autoencoder. Nonetheless, it failed to seize secular relations. Deep recurrent NN (DRNN), along with huge assurance in symbolizing dynamic action, have reached huge achievement in sequence modeling, particularly for speech recognition [12] as well as video segmentation [9]. In recent times, investigators designed to combine Graph Convolutional Network (GCN) or Convolutional NN (CNN) to Long Short-Term Memory (LSTM) or Recurrent NN (RNN) at the same time discovers the spatial-temporal correlations within the traffic flow information [8]. The usage of DL techniques is examined to design traffic flow forecasting. On account of the time-sequence quality of the information, RNN based techniques are utilized like LSTM as well as its variations like Convolutional LSTM to forecast traffic volumes at various junctions. It is also considered the reliance across the junctions in traffic flow forecasting. Convolution process could be utilized in numerous different jobs, where surrounding data is provided [13].

The benefaction of this traffic forecasting model is to establish an efficient technique named BGSO_NxLFTNet. Firstly, the input traffic network taken from the database is forwarded to STE generator. Secondly, the input traffic network as well as the outcome of STE generator is forwarded to traffic detection that is implemented by proposed NxLFTNet trained by BGSO, where the layers are forward Taylor network. Here, NxLFTNet is integrated using NARX and LSTM. Moreover, BGSO is amalgamated by GSO and BER.

- Proposed BGSO_NxLFTNet for traffic forecasting: A Novel scheme for traffic forecasting is developed utilizing BGSO_NxLFTNet. In this model, the traffic is detection is efficaciously conducted on the basis of newly discovered technique NxLFTNet, which is merged by NARX and LSTM. Here, the proposed NxLFTNet is tuned by employing BGSO that is incorporated by GSO and BER.

The remaining portion of BGSO_NxLFTNet is: The researches of traditional methods along with the difficulties experienced in these methods are clearly in section 2. The system model is demonstrated in section 3 and its proposed determination is expressed in section 4. The performance outcomes and its discussions are deliberated in section 5.

2. Motivation

The intention of traffic forecast is to detect traffic conditions of various future time-steps provided the historic traffic data. Nevertheless, conventional techniques deploy a flexible and predetermined adjacency matrix, which does not precisely replicate real time correlations among signals. Henceforth, the scholars are interested in developing a new model by conquering this difficult task.

2.1 Literature Survey

Weng, *et al.* [15] introduced Dynamic Graph Convolutional Recurrent Network (DDGCRN). This technique reduced the dependence of prior knowledge and improved the adaptability of spatio-temporal graphs. Nevertheless, numerous traffic patterns in traffic signals did not decompose to design for enhancing the interpretability and model's performance. Gao, Y., *et al.* [16] developed LSTM module for predicting the traffic. With the help of data flow from ETC model, the speed and assess the operating condition of highways was forecasted. Nevertheless, the traffic flow gathered by some nearby detectors was not correlated. Fang, W., *et al.* [17] designed Attention Mechanism LSTM (AM-LSTM) for traffic flow prediction. Even though, this technique was superior after the stabilization process, the computational time of this module steadily higher than since it included attention model. Zhang, Y., *et al.* [14] established Multi-modal Context-based Graph Convolutional Neural Network (MCGCN). Once the incorporation of multimodal context information was performed into traffic speed prediction, MCGCN model was unable to utilize spatial and temporal context information exhibited by some states for attaining superior performance in long-term speed detection.

Nabi, S.T., *et al.* [18] presented fusion model, which comprises LSTM, Bidirectional LSTM (Bi-LSTM) and Gated Recurrent Unit (GRU). This model implemented soft factors before the eradication of quality of service (QoS). Nonetheless, it did not focus on dynamic resource allocation in heterogeneous complex networks structures. Ma, C., *et al.* [19] introduced Enhanced Information Graph Recursive Network (EIGRN). The recursive structure of the technique allowed for learning the illustrations of traffic patterns from noisy data. Yet, the training and inference times was computationally expensive, particularly for large scale databases. Ma, Y., *et al.* [20] designed CNN-LSTM model for predicting traffic flow. This approach surpassed superior results for flow and speed of traffic forecasting. But, the Deep Neural Network (DNN) did not accumulate non-linear development of traffic flow by clearly discovering local traffic patterns as well as dynamics. Shin, Y. and Yoon, Y., [21] established Progressive Graph Convolutional Network (PGCN). This technique had the ability for generalizing in numerous feature databases that proved the requirement of replicating online input data for acquiring robustness. Nevertheless, it was not capable of integrating external features of transportation structures namely, weather, Point-of-Interests (POI) and road features.

2.2 Challenges

The complexities observed in former methods are elaborated below.

- An approach designed in [15] constantly outperformed better implementation on six databases. Nonetheless, this technique failed to decompose the traffic signals into normal as well as abnormal signals to design in a proper way.
- In [16], it did not enlarge this technique for forecasting the speed of traffic for numerous detectors even a large-scale network and also did not examine the extraction technique of prior and further moment conditions of traffic flow information for increasing the assessment of detection.
- Although the established module in [17] assigned diverse weights to diverse inputs and extracting significant information, it was unable to deploy featured intelligent transportation structures.
- The module introduced in [14] explored how different contexts impacted the speed of traffic prediction. However, it neglected the correlated information of context data among numerous modalities that did not fused multimodal context data from different sources.
- Recently, Limited data availability is an important challenge in traffic prediction. More conventional traffic sensors as well as devices cannot exhibit real-time data, and new data sources can be facilitated for enhancing the precision of forecasting.

3. System Model

This model [15] describes the basic concepts of traffic networks, traffic signals, spatial and temporal properties.

3.1 Traffic Network

This network is illustrated in the form of graph as $\mathrm{A}=(\delta,\vartheta,\omega)$. Here, a set of B nodes that indicates sensors placed at subsequent positions in rod network. These sensors are accountable to record traffic information at their positions. The set of edges indicates ϑ and graph determined from pairwise distances among nodes in network. It includes dynamic adjacency matrix is represented by $\omega^{i} \in \mathrm{I}^{\mathrm{B}\times\mathrm{B}}$.

3.2 Traffic Signal

This signal is computed as $\mathrm{E}_{\mathrm{l}} \in \mathrm{I}^{\mathrm{B}\times\mathrm{H}}$ that acquired the values of overall sensors in traffic network A at time step ℓ. Here, traffic features gathered by sensors refers H.

3.3 Traffic Forecasting

By providing traffic signals $\alpha_{m}=\left[\mathrm{X}_{\mathrm{l}-m+1}....\mathrm{X}_{\mathrm{l}-1},\mathrm{X}_{\mathrm{l}}\right]\in \mathrm{I}^{m\times\mathrm{B}\times\mathrm{H}}$, traffic prediction intends for detecting further traffic signals $\beta_{n}=\left[\mathrm{Y}_{\mathrm{l}+1}.\mathrm{Y}_{\mathrm{l}-1},.....,\mathrm{Y}_{\mathrm{l}+n}\right]\in \mathrm{I}^{n\times\mathrm{B}\times\mathrm{H}}$.

3.4 Spatial and temporal properties

Assume B number of nodes in transportation network and sensor sampling data's frequency is B_{i}. The features are stored in autonomous trainable embedding matrix that indicates $\mathrm{K}\in \mathrm{I}^{\mathrm{B}\times\mathrm{N}}$, $\mathrm{T}^{\mathrm{N}}\in \mathrm{I}^{\mathrm{B}_{i}\times\mathrm{N}}$, $\mathrm{T}^{\mathrm{M}}\in \mathrm{I}^{\mathrm{B}_{\mathrm{M}}\times\mathrm{N}}$. Here, embedding dimension represents N.

4. Proposed BGSO_NxLFTNet for traffic forecasting

Traffic prediction is attaining great interest owing to the broad application in ITSs. Providing challenging as well as dynamic traffic data, numerous techniques focused on how to introduce a spatial-temporal technique for expressing non-stationary traffic patterns. Thus, the fundamental intention of this potent system is to discover and explore traffic forecasting named BGSO_NxLFTNet. Initially, input traffic network in the form of graph taken from the database [22] is allowed through STE generator, which is employed for finding daily embedding and weekly embedding of time corresponding to current traffic signal. Afterwards, input traffic network and the output of STE generator is fed into detection phase. Finally, traffic forecasting is done by utilizing NxLFTNet. It is the combination of NARX [23] and LSTM [24], where the layers are modified by Taylor concept. Moreover, NxLFTNet is trained using BHSO, which is the

amalgamation of GSO [25] and BER [26]. Figure 1 articulates schematic diagram of traffic forecasting.

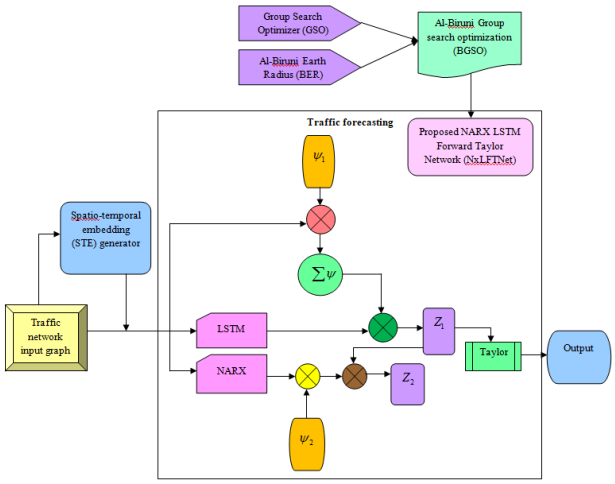


Fig. 1. Schematic diagram of traffic forecasting

4.1 Data Acquisition

Traffic data collected in representation of spatio-temporal embedding. It has more number of sensors and features that are obtained by the sensor. In this module, identifier (ID), latitude (Lat), longitude (Lng), district, country, highway (Fwy), number of lanes (Lane), type and direction are acquired. The spatio-temporal embedding indicates O with diemsnion of $[L \times H]$. Furthermore, O with $[L \times H]$ is specified by $N_{L \times H}$. Figure 2 elucidates the illustration of spatial embedding data.

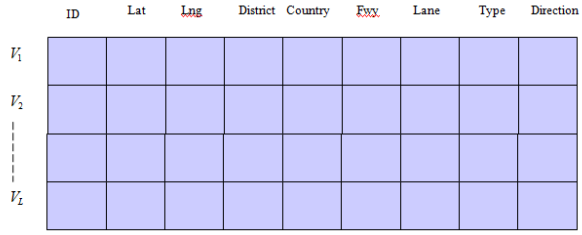


Fig. 2. Illustration of spatial embedding data

Daily embedding of time related to current traffic data represents $M_{\Re}$, that signifies $[L \times H \times \tau]$ dimension. Here, the daily embedding with $[L \times H \times \tau]$ dimension specifies $M^{\Re}_{L \times H \times \tau}$, wherein traffic data is gathered for τ periods. Figure 3 demonstrates modeled view of daily embedding.

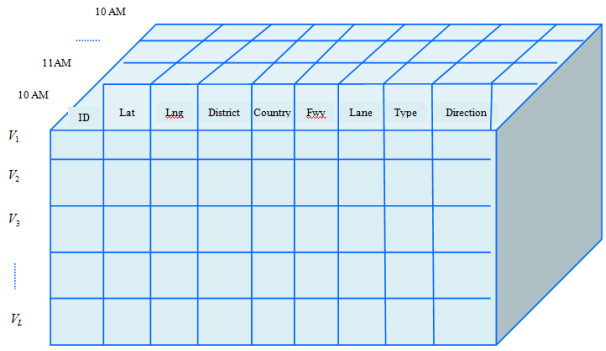


Fig. 3. Modeled illustration of daily embedding

Weekly embedding of time relevant to current data is modeled as $M^{\Im}$ with $[L \times H \times \tau]$ dimensional data. Moreover, the dimension $[L \times H \times \tau]$ of $M^{\Im}$ is signified as $M^{\Im}_{L \times H \times \tau}$. Figure 4 demonstrates modeled illustration of weekly embedding.

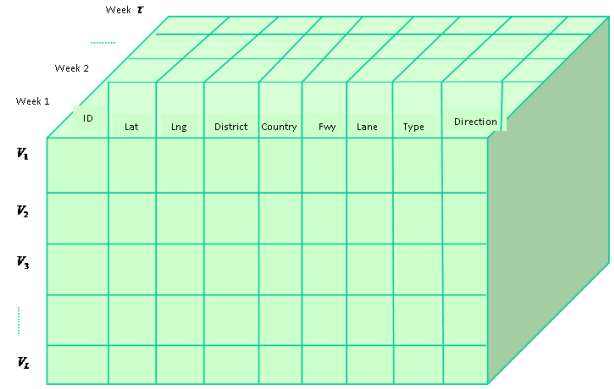


Fig. 4. Modeled illustration of weekly embedding

4.2 STE Generation

The purpose of this generator is for learning a frequent and distinct illustration of high dimensional data, which accumulates intricate relationships among the dependencies of spatial and temporal features. It is attained by integrating the input data into less dimensional space when conserving the topological of actual data and geometrical representations. This is also for embedding the data in a way which precisely observes the patterns and combinations among diverse spatial locations and time points that enables the module for efficient detection, classification and data clustering. With a compact generation and useful data representation, the traffic detection is performed.

STE determines daily embedding $M_{\Re}$ as well as weekly embedding $M_{\Im}$ at the time relevant to current traffic S_{τ}. Then, element wise product is conducted between spatial embedding N, $M_{\Re}$ and $M_{\Im}$ to acquire new spatial embedding N^{new}_{χ}, which is formulated by,

$$N^{new}_{\chi} = N \times M_{\Re} \times M_{\Im} \tag{1}$$

Figure 5 illustrates features of individual sensors.

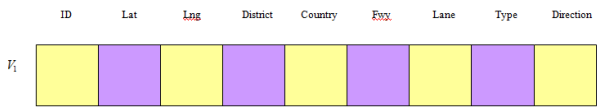


Fig. 5. Features of individual sensor

The sensor V_1 is assumed wherein it collects H features in τ number of times. Henceforth, N_{χ}^{new} of sensor illustrates $1\times(H\times\tau)$ dimension. The daily and weekly embedding is modeled in figure 6 and figure 7.

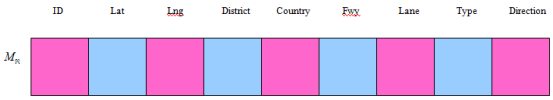


Fig. 6. Features of daily embedding

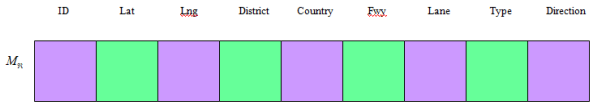


Fig. 7. Features of weekly embedding

4.3 Traffic flow prediction

Precise traffic prediction is significant in enhancing the reliability and efficacy in ITSs. Regardless of the prior investigations, precise traffic forecasting experiences beneath complexities that includes designing the dynamics of traffic data in temporal and spatial dimensions and the issue is more complex for long term forecasting approach.

NxLFTNet is utilized in forecasting traffc flow, in which NxLFTNet has NARX and LSTM. The spatial embedding N_{χ}^{new} is subjected to both modules. Firstly, N_{χ}^{new} is concatenated with weight ψ_1, and then weights are summed up. Then, N_{χ}^{new} from LSTM is concatenated with summarized value for acquiring outcome Z_1. On the other hand, N_{χ}^{new} from NARX is concatenated by weight ψ_2, and it is concatenated by Z_1 for the resultant Z_2. Lastly, Z_1 and Z_2 are applied to Taylor concept for detecting the traffic that is signified by T. The overall outlook of NxLFTNet is illustrated in figure 8.

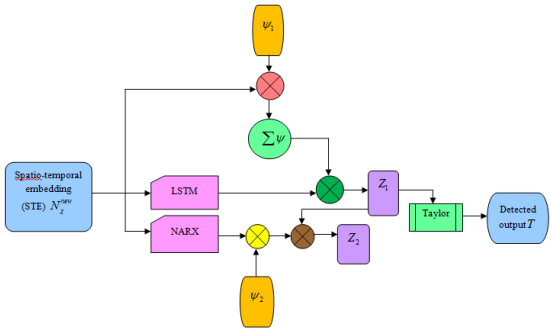


Fig. 8. Schematic depiction of NxLFTNet

4.3.1 Structure of LSTM

LSTM [24] comprises two blocks and dense layer which also comprises cell state, and it implements on entire chain with the ability of including as well as eliminating the data to cell state. LSTM has point-wise multiplication function as well as sigmoid layer. Here, N_{χ}^{new} and hidden state of former time step (C_{g-1}) are included in LSTM block. This predicts the type of data is neglected from cell. The outcome is expressed as,

$$Z_1 = \mu\left(N_{\chi}^{new}J^{\varepsilon} + C_{g-1}K^{\varepsilon} + j_{\varepsilon}\right) * \sum\psi_1 N_{\chi}^{new} \tag{2}$$

where, Z_1 specifies LSTM output, J implies input weight, K refers biased weight, and j specifies bias. Figure 9 displays architecture of LSTM.

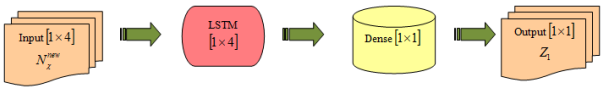


Fig. 9. Structure of the LSTM model

4.3.2 Structure of NARX

NARX [23] exhibits a feedback connection which accumulates more layers of network for acquiring nonlinear time series prediction. In addition to that, it deploys the memory for processing former actual or detected time series values. The learning process is effectual and network convergence is speeder for generating extreme outcomes. Thus, the resultant is enumerated as,

$$Z_2 = \left[d_2\left(\left[\sum_{o=1}^{P_k} F_{\phi o}E_o(z) + N_{\chi}^{new}\right] * \psi_2\right)\right] * \mu\left(N_{\chi}^{new}J^{\varepsilon} + C_{g-1}K^{\varepsilon} + j_{\varepsilon}\right) * \sum\psi_1 N_{\chi}^{new} \tag{3}$$

where, connection weight indicates $F_{\phi o}$ amid o^{th} hidden neurons and ϕ^{th} detected resultant. The entire hidden neurons refer P_k and activation function for resultant layer articulates d_2, bias explicates j_{ε}, hidden layer at time $\boldsymbol{z}$ implies $E_o(z)$. Figure 10 demonstrates architecture of NARX.

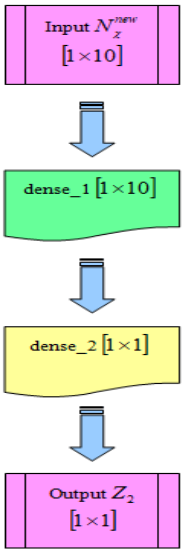


Fig. 10. Structure of NARX

By using Taylor concept, the final output expresses,

$$T=\mu\left(N_{\chi}^{new}J^{\varepsilon}+C_{g-1}K^{\varepsilon}+j_{\varepsilon}\right)*\sum\psi_1 N_{\chi}^{new}-2\left[d_2\left(\left[\sum_{o=1}^{p_0}F_{\psi_0}E_o(z)+N_{\chi}^{new}\right]*\psi_2\right)\right] \quad (4)$$

Here, the detected traffic is symbolized as T .

4.3.3 Training Algorithm of BGSO

BGSO is introduced by the amalgamation of GSO [25] and BER [26], which is used to train the network named NxLFTNet for showing the effectualness of this new hybrid optimization algorithm. GSO [25] is a nature-inspired optimization algorithm and it is motivated by searching behavior of animal in real life. This optimizer is deployed for exploring an extreme solution in excess of candidate solution for resolving the optimization issues by analyzing minimal or maximal intent function for a certain issue. Every group of agents has producers, scroungers, and dispersed members. BER [26] is motivated by determination of earth radius for examining the search space surrounding the solutions with swarm members' behavior for fulfilling their global goals. It enhances the performance of exploitation, strikes a balance among exploitation and exploration, enlarging the search space. By using this effectiveness, GSO and BER are incorporated for showing the efficaciousness of the model. The steps followed in this model are demonstrated as,

a) Solution Encoding

It enumerates the illustration of optimization problem and mapped to a set of potential solutions. Moreover, it is utilized for simplifying the search space (γ) that makes it easier to identify fine solutions that is computed as,

$$\gamma=[1\times\eta] \quad (5)$$

Here, learning parameter of NxLFTNet is η .

b) Fitness function

This objective function is applied for finding an optimal solution that examines how close a given solution is to the optimum solution of desired issue that is determined by,

$$F=\frac{1}{p}\sum_{l=1}^{p}\left[v_{out}-T_{out}\right]^2 \quad (6)$$

Here, p indicates total number of data, l represents input data, v_{out} implies targeted output and detected output symbolizes T_{out} .

c) Algorithmic steps

The algorithmic phases of NxLFTNet are derived in the following stages.

Step 1: Initialization

This optimizer is conducted utilizing a set of populations named group, where every agent as member. This is the initial step considered for analyzing the random solutions with a given search space that is computed in the form of,

$$G=\{G_1,G_2,.....G_r,.....G_{\mathrm{M}}\} \quad (7)$$

Here, G_r implies the candidate solution considered for the below evaluation, G_{M} signifies total number of populations.

Step 2: Compute fitness

It is used to examine the fine solution for attaining a superior outcome that is already formulated in Eq. (6).

Step 3: Producers phase

In GSO, at the course number that is s^{th} iteration, the producer G_a is implemented as follows:

It determine at zero and with stochastic testing three positions in the validating position: one instance at zero rate is formulated as,

$$G_v=G_a^s+w_1u_{\max}W_a^s\left(\xi^s\right) \quad (8)$$

A point in right hand faction hypercube is illustrated by,

$$G_e=G_a^s+w_1u_{\max}W_a^s\left(\xi^s+w_2\Theta_{\max}\setminus 2\right) \quad (9)$$

A point in left hand faction hypercube is computed by,

$$G_u=G_a^s+w_1u_{\max}W_a^s\left(\xi^s+w_2\Theta_{\max}\setminus 2\right) \quad (10)$$

Here, commonly distributed stochastic value with average value 0 symbolizes $w_1\in\aleph^1$, the value of standard deviation as 1 and consistently distributed stochastic values in the interval of $[0,1]$ refers $w_2\in\aleph^{z-1}$.

Then, the producer processed for attaining the nearby extreme solution with better function. When the extreme

solutions have objective value than new solution, it transmits towards this position. Otherwise, it remains in its position and given its caption to a diverse arbitrarily generated location.

$$\xi^{s+1} = \xi^{s} + w_2 \lambda_{\max} \tag{11}$$

Here, maximal adjusting location indicates $\lambda_{amx} \in \aleph^1$.

When the producer failed to attain an extreme search space after λ count of iterations, the leader back to zero degree is used that is computed by,

$$\xi^{s+1} = \xi_s \tag{12}$$

Here, constant value represents $\lambda \in \aleph^1$.

Step 4: Scroungers phase

The scroungers may frequent the searching process for extreme objective value for attaining the objective value analyzed by the producer. Here, only the space copying that is the prime scrounging behavior in sparrows is employed. At s^{th} diffusion, the behavior of space copying of r^{th} scrounger is illustrated as stochastic walk near producer is computed by,

$$G_r^{s+1} = G_r^{s} + w_3 \, o\left(G_a^{s} - G_r^{s}\right) \tag{13}$$

Here, parallel stochastic sequence value represents $\left(w_3 \in \aleph^x\right)$ in the interval of $[0,1]$ and product signifies $\circ$ that computes the multiplication of two vectors. Through this process, the r^{th} scrounger keep discovering for other probabilities to attain, this is designed by deploying r^{th} scrounger's start to an original randomly created position utilizing Eq. (12).

Step 5: Dispersion phase

Several illustrations of dispersions are obtained that ranges from effortless insects to human being. Scattered creatures deploy ranging behavior for searching and finding various environments. Ranging initiates external pointing to a specific device. Here, the r^{th} group agent is scattered for classification process. Normally, ranging beings create searching process that includes stochastic walks as well as methodical exploration techniques for identifying the resources. The arbitrary walks are regarded to be the normal efficient searching technique for scholastic allocated values employed by rangers. It generates a scholastic front position ξ_r at s^{th} search and considers random distance as,

$$u_r = \lambda \cdot u_{\max} \tag{14}$$

The update solution from GSO is expressed as,

$$G_r^{s+1} = G_r^{s} + u_r W_r^{s}\left(\xi^{s+1}\right) \tag{15}$$

From BER, the solution is employed to incorporate with GSO's solution that is enumerated as,

$$\vec{Q}(q+1) = h^2\left(\vec{Q}(q) + \vec{R}\right) \tag{16}$$

Here,

$$\vec{R} = \vec{h}_3\left(\vec{U}(q) - \vec{Q}(q)\right) \tag{17}$$

$$\vec{Q}(q+1) = h^2\left(\vec{Q}(q) + \vec{h}_3\left(\vec{U}(q) - \vec{Q}(q)\right)\right) \tag{18}$$

$$\vec{Q}(q+1) = h^2\vec{Q}(q) + h^2\vec{h}_3\vec{U}(q) - \vec{h}_3\vec{Q}(q) \tag{19}$$

$$\vec{Q}(q+1) = \vec{Q}(q)\left(h^2 - \vec{h}_3\right) + h^2\vec{h}_3\vec{U}(q) \tag{20}$$

Assume,

$$\vec{Q}(q+1) = G_r^{s+1} \tag{21}$$

$$\vec{Q}(q) = G_r^{s} \tag{22}$$

$$\vec{U}(q) = U_r^{s} \tag{23}$$

Substituting Eq. (21), (22), (23) in Eq. (20),

$$G_r^{s+1} = G_r^{s}\left(h^2 - \vec{h}_3\right) + h^2\vec{h}_3 U_r^{s} \tag{24}$$

$$G_r^{s} = \frac{G_r^{s+1} - h^2\vec{h}_3 U_r^{s}}{\left(h^2 - \vec{h}_3\right)} \tag{25}$$

Applying Eq. (25) in Eq. (15),

$$G_r^{s+1} = \left[\frac{G_r^{s+1} - h^2\vec{h}_3 U_r^{s}}{\left(h^2 - \vec{h}_3\right)}\right] + u_r W_r^{s}\left(\xi^{s+1}\right) \tag{26}$$

$$G_r^{s+1} - \frac{G_r^{s+1}}{\left(h^2 - \vec{h}_3\right)} = \frac{u_r W_r^{s}\left(\xi^{s+1}\right)\left(h^2 - \vec{h}_3\right) - h^2\vec{h}_3 U_r^{s}}{\left(h^2 - \vec{h}_3\right)} \tag{27}$$

$$\frac{\left(h^2 - \vec{h}_3 - 1\right)G_r^{s+1}}{\left(h^2 - \vec{h}_3\right)} = \frac{u_r W_r^{s}\left(\xi^{s+1}\right)\left(h^2 - \vec{h}_3\right) - h^2\vec{h}_3 U_r^{s}}{\left(h^2 - \vec{h}_3\right)} \tag{28}$$

$$G_r^{s+1} = u_r W_r^{s}\left(\xi^{s+1}\right)\left(h^2 - \vec{h}_3\right) - h^2\vec{h}_3 U_r^{s} \tag{29}$$

Here, G_r^{s+1} is the update solution, $\vec{Q}(q)$ signifies solution vector at iteration q, $\vec{R}$ symbolizes distance vector, $\vec{h}_3$ indicates random vector and best solution vector refers $\vec{U}$.

Step 6: Termination

This kind of process will be continuously takes place till it acquires the supreme solution. Algorithm 1 articulates pseudo code of BGSO.

Algorithm 1. Pseudo code of BGSO

SL.NO	Pseudo code of BGSO
1	**Input:** $s=0$, iterations, number of populations M
2	**Output:** G_r^{s+1}
3	Begin
4	Initialize agent G_r and head angles ξ_r
5	Determine fitness by Eq. (6)
6	While max$-s$ is not attain do
7	For every agents r in group do
8	Examine the Producer phase by Eq. (8), (9), (10), (11), (12)
9	Examine the scrounging phase by Eq. (13)
10	Examine the dispersion phase by Eq. (29)
11	End for
12	End while
13	**Return**
14	**Terminate**

5. Results and Discussions

The valuation of BGSO_NxLFTNet is clearly deliberated in the following phases and description of databases and implementation setup.

5.1 Experimental Setup

The experimental setup of BGSO_NxLFTNet is successfully executed in PYTHON tool.

5.2 Database Description

The database of BGSO_NxLFTNet is described in this phase:

5.2.1 LargeST Database

This database [22] has four sub-databases where it is elaborated utilizing diverse sensors. Amongst these four sub- databases, California (CA) is the largest one with 8,600 sensors. Furthermore, three subsets of CA are Greater Los Angeles (GLA), Greater Bay Area (GBA), and San Diego (SD).

5.3 Evaluation Metrics

The measures of BGSO_NxLFTNet are elaborated here.

5.3.1 MAPE

It [27] defines the detection error as a percentile of actual value, which is indicated by,

$$\Phi_1 = \frac{1}{\varsigma}\sum_{\varpi=1}^{\varsigma}\left|\frac{\rho_\varpi - \kappa_\varpi}{\kappa_\varpi}\right| \tag{30}$$

5.3.2 RMSE

MSE illustrates the mean value of square of difference amongst actual and detected values, where its square root indicates RMSE [27].

$$\Phi_2 = \sqrt{\frac{1}{\varsigma}\left(\kappa_\varpi - \rho_\varpi\right)^2} \tag{31}$$

5.3.3 MAE

Absolute error is enumerated as degree of dissimilarity among determined value and actual value and its average value is illustrated as [27],

$$\Phi_3 = \frac{1}{\varsigma}\left|\left(\kappa_\varpi - \rho_\varpi\right)\right| \tag{32}$$

In this section, the overall samples symbolize ς , κ_ϖ specifies original value and ρ_ϖ implies detected value.

5.4 Comparative Methods

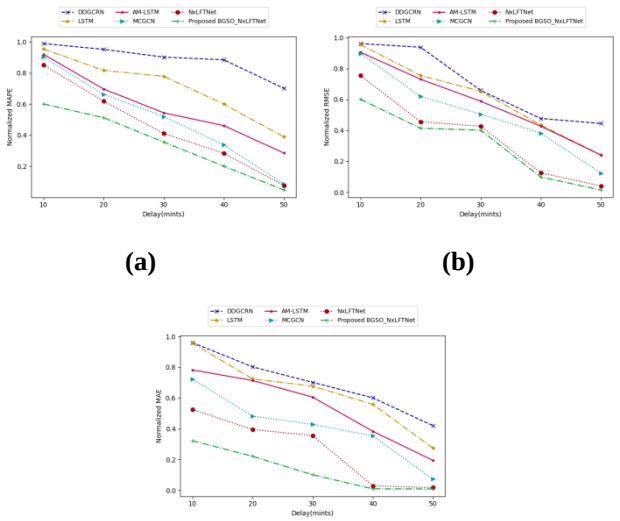
DDGCRN [15], LSTM [16], AM-LSTM [17], MCGCN [14] and NxLFTNet are the traditional modules of Proposed BGSO_NxLFTNet.

5.5 Comparative Analysis

The estimation of BGSO_NxLFTNet is altered by delay based on three databases namely, Alameda, Contra Costa, and El Dorado.

5.5.1 Analysis of BGSO_NxLFTNet using Alameda

Figure 11 describes BGSO_NxLFTNet's assessment utilizing Alameda varying delay. In figure 11 a), the BGSO_NxLFTNet evaluation with relevance of MAPE is designed. When delay is 50 mints, the estimation of BGSO_NxLFTNet achieved MAPE with 0.047, while prior schemes attained 0.701, 0.389, 0.285, 0.086 and 0.077. Figure 11 b) displays BGSO_NxLFTNet valuation concerning RMSE. If delay is 50 mints, then BGSO_NxLFTNet observed RMSE of 0.015, wherein conventional approaches acquired 0.445, 0.242, 0.241, 0.124 and 0.040. Figure 11 c) explicates BGSO_NxLFTNet examination regarding MAE. Assuming delay as 50 mints, the BGSO_NxLFTNet in terms of MAE obtained 0.009, whereas the traditional techniques accomplished 0.421, 0.275, 0.195, 0.073 and 0.017.



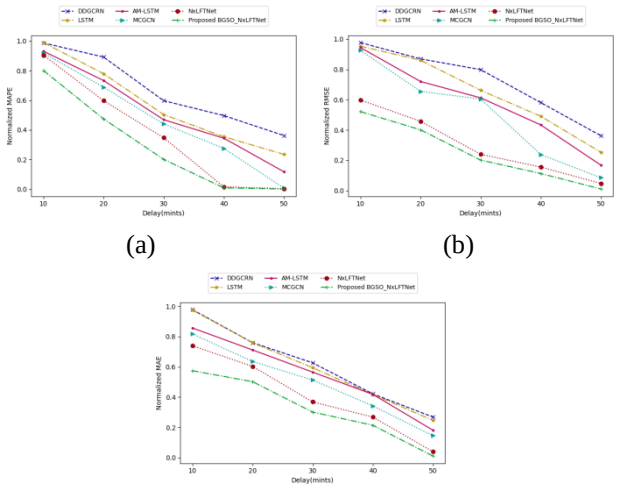
(a) **(b)**

(c)

Fig. 11. Analysis of BGSO_NxLFTNet using Alameda, a) MAPE, b) RMSE, c) MAE

5.5.2 Analysis of BGSO_NxLFTNet using GBA

Figure 12 elucidates BGSO_NxLFTNet's estimation employing Contra Costa altering delay. In figure 12 a), the BGSO_NxLFTNet assessment in relating to MAPE is represented. While delay is taken as 50 mints, the analytic values of BGSO_NxLFTNet reached MAPE with 0.001, while preceding strategies obtained 0.362, 0.235, 0.118, 0.007 and 0.003. Figure 12 b) expresses BGSO_NxLFTNet estimation according to RMSE. Considering delay as 50 mints, then BGSO_NxLFTNet achieved RMSE of 0.010, whereas other modules observed 0.362, 0.253, 0.168, 0.085 and 0.047. Figure 12 c) designs BGSO_NxLFTNet assessment relevant to MAE. With delay as 50 mints, the BGSO_NxLFTNet with MAE achieved 0.011, wherein the former models gained 0.269, 0.247, 0.182, 0.146 and 0.040.



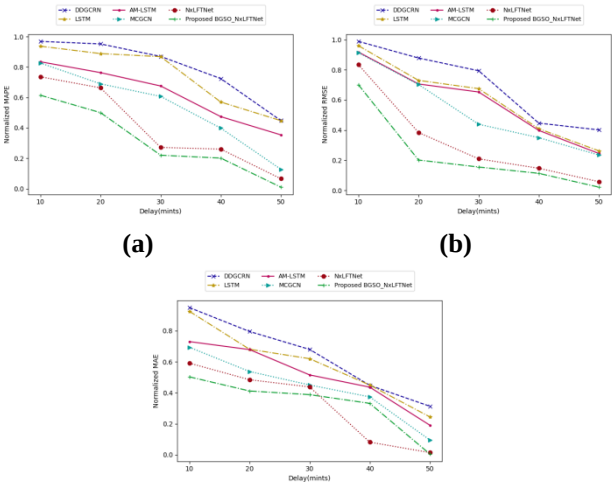
(a) (b)

(c)

Fig. 12. Analysis of BGSO_NxLFTNet using Contra Costa, a) MAPE, b) RMSE, c) MAE

5.5.3 Analysis of BGSO_NxLFTNet using EL dorado

Figure 13 represents BGSO_NxLFTNet's computation using employing EL dorado differing delay. In figure 13 a), the BGSO_NxLFTNet determination based on MAPE is illustrated. Assume delay 50 mints, the evaluation values of BGSO_NxLFTNet accomplished MAPE with 0.010, where former techniques achieved 0.448, 0.448, 0.354, 0.128 and 0.066. Figure 13 b) deliberates BGSO_NxLFTNet estimation relating to RMSE. Consider delay as 50 mints, and then BGSO_NxLFTNet attained RMSE of 0.021, whereas existing modules gained 0.401, 0.262, 0.245, 0.234 and 0.057. Figure 13 c) designs BGSO_NxLFTNet assessment relevant to MAE. With delay as 50 mints, the BGSO_NxLFTNet with MAE achieved 0.004, wherein the conventional modules acquired 0.313, 0.244, 0.190, 0.095 and 0.014.



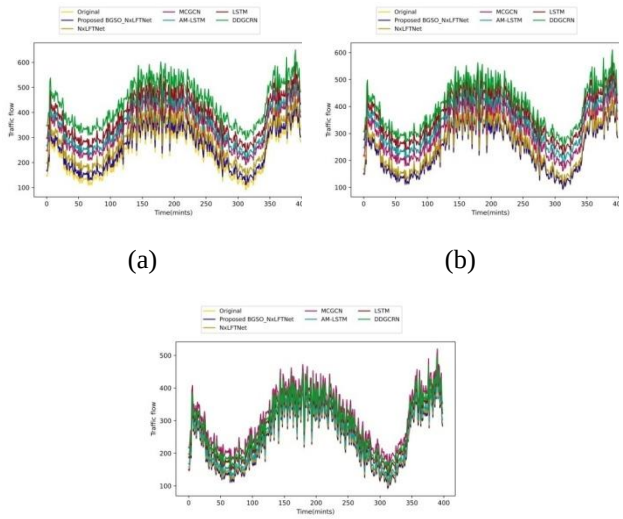
(a) **(b)**

(c)

Fig. 13. Analysis of BGSO_NxLFTNet using EL dorado, a) MAPE, b) RMSE, c) MAE

5.6 Prediction Graph

The prediction graph of this module is illustrated in Figure 14 in terms of Alameda, Contra Costa, and El Dorado varying time (mints). In figure 14 a), the Alameda of prediction graph is designed. When the time is considered as 250 mints, the original traffic flow observed 453 and BGSO_NxLFTNet acquired 403, where the other former technologies obtained 253, 273, 303, 353 and 373. Figure 14 b) represents Contra Costa of prediction graph. Assume time as 250 mints, original flow attained 413 and BGSO_NxLFTNet achieved 383, while other techniques accomplished 253, 256, 273, 323 and 353. In figure 14c), the El Dorado of prediction graph is demonstrated. With time as 250 mints, the original traffic flow observed 303, BGSO_NxLFTNet gained 293, whereas the traditional models attained 253, 255, 263, 323 and 273.



(a) (b)

(c)

Fig. 14. Analysis of BGSO_NxLFTNet, a) MAPE, b) RMSE, c) MAE

5.7 Comparative Discussions

The values acquired from proposed BGSO_NxLFTNet is discussed in table 1. Here, the estimation takes place with the preceding techniques. The developed approach BGSO_NxLFTNet achieved MAPE with the value of 0.001 that is sensitive to higher errors, making it helpful for modules required for forecasting values. The RMSE of BGSO_NxLFTNet gained 0.010 that is required for capturing delicate variations in the data. Moreover, MAE of discovered technique observed 0.011 which provided a direct measure of mean error magnitude. Thus, the overall performance of the proposed technique BGSO_NxLFTNet shows the consistent, reliability, generalizability by evaluating with comparative modules.

Table 1. Comparative discussion

Alterations based on		Metrics/Methods	DDGCRN	LSTM	AM-LSTM	MC GCN	NxLFTNet	Proposed BGSO_NxLFT Net
Alameda	Delay (mints) =50	MAPE	0.701	0.389	0.285	0.086	0.077	0.047
		RMSE	0.445	0.242	0.241	0.124	0.040	0.015
		MAE	0.421	0.275	0.195	0.073	0.017	0.009
Contra Costa		MAPE	0.362	0.235	0.118	0.007	0.003	0.001
		RMSE	0.362	0.253	0.168	0.085	0.047	0.010
		MAE	0.269	0.247	0.182	0.146	0.040	0.011
EL dorado		MAPE	0.448	0.448	0.354	0.128	0.066	0.010
		RMSE	0.401	0.262	0.245	0.234	0.057	0.021
		MAE	0.313	0.244	0.190	0.095	0.014	0.004

6. Conclusion

Traffic prediction has been active investigation domain for the past few years and with the advancement of DL methodologies, investigators are demanding to apply DL for excellent enhancements. The reliable and precise traffic prediction may help in intelligent route guidance, proactive and dynamic traffic control that might assist for exhibiting extreme congestion issue in the model. In this research, a traffic forecasting module is designed utilizing BGSO_NxLFTNet. The input traffic network is forwarded to STE generator to recognize the daily as well as weekly embeddings of time in context of present traffic signal. The output of STE generator and input traffic network is given for traffic detection phase that is successfully forecasted using NxLFTNet, which is trained by BGSO. The measures employed for the analysis of proposed BGSO_NxLFTNet are MAPE, RMSE and MAE acquired uttermost results with better values of 0.001, 0.010 and 0.011. The future exploration will explore a system by taking dynamic complexity of road network and weather's disruption and more data sources will consider in detecting traffic.

Conflicts of Interest:

The authors declare no conflict of interest.

Author Contributions: Conceptualization, methodology, software, validation, formal analysis, investigation, resources, writing—original draft preparation, writing—review and editing: Sathyanarayana Mangali; supervision-Dr. Jayasree H.

References

[1] B. N. Kamath R. Fernandes, A. P. Rodrigues, M. Mahmud, P. Vijaya, T. R. Gadekallu, and M. S. Kaiser, "TAKEN: a traffic knowledge-based navigation system for connected and autonomous vehicles", *Sensors,* vol. 23, no. 2, pp.653, 2023.

[2] B. Yu, H. Yin, and Z. Zhu, "Spatio-temporal graph convolutional networks: A deep learning framework

for traffic forecasting", *arXiv preprint arXiv: 1709.04875*, 2017.

[3] M. E. Ben-Akiva, S. Gao, Z. Wei, and Y. Wen, "A dynamic traffic assignment model for highly congested urban networks", *Transportation research part C: emerging technologies*, vol. 24, pp.62-82, 2012.

[4] A. Hamilton, B. Waterson, T. Cherrett, A. Robinson, and I. Snell, "The evolution of urban traffic control: changing policy and technology", *Transportation planning and technology*, vol. 36, no. 1, pp.24-43, 2013.

[5] H. F. Yang, T. S. Dillon, and Y. P. P. Chen, "Optimized structure of the traffic flow forecasting model with a deep learning approach", *IEEE transactions on neural networks and learning systems*, vol. 28, no. 10, pp.2371-2381, 2016.

[6] K. Boriboonsomsin, M. J. Barth, W. Zhu, and A. Vu, "Eco-routing navigation system based on multisource historical and real-time traffic information", *IEEE Transactions on Intelligent Transportation Systems*, vol. 13, no. 4, pp.1694-1704, 2012.

[7] M. Veres, and M. Moussa, "Deep learning for intelligent transportation systems: A survey of emerging trends", *IEEE Transactions on Intelligent transportation systems*, vol. 21, no. 8, pp.3152-3168, 2019.

[8] 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, no. 9, pp.1474, 2020.

[9] R. Yu, Y. Li, C. Shahabi, U. Demiryurek, and Y. Liu, "Deep learning: A generic approach for extreme condition traffic forecasting", *In Proceedings of the 2017 SIAM international Conference on Data Mining, Society for Industrial and Applied Mathematics*, pp. 777-785, June 2017.

[10] J. Nahar, Y. P. P. Chen, and S. Ali, "Kernel-based naive bayes classifier for breast cancer prediction", *Journal of biological systems*, vol. 15, no. 01, pp.17-25. 2007.

[11] K. Y. Chan, S. Khadem, T. S. Dillon, V. Palade, J. Singh, and E. Chang, "Selection of significant on-road sensor data for short-term traffic flow forecasting using the Taguchi method", *IEEE Transactions on Industrial Informatics*, vol. 8, no. 2, pp.255-266, 2011.

[12] K. Y. Chan, T. Dillon, E. Chang, and J. Singh, "Prediction of short-term traffic variables using intelligent swarm-based neural networks", *IEEE Transactions on Control Systems Technology*, vol. 21, no. 1, pp.263-274, 2012.

[13] S. Kundu, M. S. Desarkar, and P. K. Srijith, "Traffic forecasting with deep learning", *In proceedings of 2020 IEEE Region 10 Symposium (TENSYMP)*, pp. 1074-1077, June 2020.

[14] Y. Zhang, T. Zhao, S. Gao, and M. Raubal, "Incorporating multimodal context information into traffic speed forecasting through graph deep learning", *International Journal of Geographical Information Science*, vol. 37, no. 9, pp. 1909-1935, 2023.

[15] Weng, Wenchao, J. Fan, H. Wu, Y. Hu, H. Tian, F. Zhu, and Jia Wu., "A Decomposition Dynamic graph convolutional recurrent network for traffic forecasting", *Pattern Recognition*, vol. 142, 2023.

[16] Y. Gao, C. Zhou, J. Rong, Y. Wang, and S. Liu, "Short-term traffic speed forecasting using a deep learning method based on multitemporal traffic flow volume", *IEEE Access*, vol. 10, pp. 82384-82395, 2022.

[17] W. Fang, W. Zhuo, J. Yan, Y. Song, D. Jiang, and T. Zhou, "Attention meets long short-term memory: A deep learning network for traffic flow forecasting", *Physica A: Statistical Mechanics and its Applications*, vol. 587, pp. 126485, 2022.

[18] S. T. Nabi, M. R. Islam, M. G. R. Alam, M. M. Hassan, S. A. AlQahtani, G. Aloi, and G. Fortino, "Deep learning based fusion model for multivariate LTE traffic forecasting and optimized radio parameter estimation", *IEEE Access*, vol. 11, pp.14533-14549, 2023.

[19] C. Ma, K. Sun, L. Chang, and Z. Qu, "Enhanced Information Graph Recursive Network for Traffic Forecasting" *Electronics*, vol. 12, no. 11, pp.2519, 2023.

[20] Y. Ma, Z. Zhang, and A. Ihler, "Multi-lane short-term traffic forecasting with convolutional LSTM network", *IEEE Access*, vol. 8, pp.34629-34643, 2020.

[21] Y. Shin, and Y. Yoon, "PGCN: Progressive graph convolutional networks for spatial–temporal traffic forecasting", *IEEE Transactions on Intelligent Transportation Systems*, 2024.

[22] Large ST dataset will be taken from "https://github.com/liuxu77/LargeST?tab=readme-ov-file", accessed on May 2024.

[23] Q. Liu, W. Chen, H. Hu, Q. Zhu, and Z. Xie, "An optimal NARX neural network identification model for a magnetorheological damper with force-distortion behavior", *Frontiers in Materials*, vol.7, pp.10, 2020.

[24] A. Sagheer, and M. Kotb, "Time series forecasting of petroleum production using deep LSTM recurrent networks", *Neurocomputing,* vol. 323, pp.203-213, 2019.

[25] L. Abualigah, "Group search optimizer: a nature-inspired meta-heuristic optimization algorithm with its results, variants, and applications", *Neural Computing and Applications,* vol. 33, no. 7, pp.2949-2972, 2021.

[26] E. S. M. El-Kenawy, A. A. Abdelhamid, A. Ibrahim, S. Mirjalili, N. Khodadad, M. A. Alduailij, A. A. Alhussan, and D. S. Khafaga, "Al-Biruni Earth Radius (BER) Metaheuristic Search Optimization Algorithm", *Comput. Syst. Sci. Eng.,* vol. 45, no. 2, pp.1917-1934, 2023.

[27] S. H. Bhojani, and N. Bhatt, "Wheat crop yield prediction using new activation functions in neural network", *Neural Computing and Applications*, vol. 32, no. 17, pp.13941-13951, 2020.