A Grey Neural Network Model Optimized by Fruit Fly Optimization Algorithm for Short-term Traffic Forecasting

Yuanyuan Pan, Yongdong Shi

Abstract—**Accurate short-term traffic forecasting can relieve traffic congestion and improve the mobility of transportation, which is very important for management modernization of transportation systems. However, it is quite difficult to predict effectively and accurately short-term traffic information since traffic information shows the strongly nonlinear, complex and uncertain characteristics. In this paper, a new hybrid model based on grey neural network and fruit fly optimization algorithm (FOA) is proposed to solve this problem. The FOA is used to select the appropriate parameter values of the grey neural network model, thereby improving the accuracy of forecasting model. The proposed hybrid model can exploit sufficiently the characteristics of grey system model requiring less data, the non-linear map of neural networks and the quick-speed convergence of FOA, and has simpler structure. The effectiveness of this proposed hybrid model is proved by experiment simulation. The experiment results show that the proposed model has better performance than the single grey model GM(1,1), the single back-propagation neural network (BPNN) model, the combined model of them, i.e., the grey neural network (GNN) model, and the GNN model with particle swarm optimization (GNN-PSO), on short-term traffic forecasting.**

Index Terms—**Short-term traffic forecasting, Grey system, Neural networks, Grey neural network, Fruit fly optimization algorithm.**

I. Introduction

SHORT-TERM traffic forecasting is an important part of management modernization of transportation systems, which has attracted more and more attentions from the academic and the practice. Short-term traffic forecasting focus on making predictions about the likely traffic parameters changes in the short-term, typically within several minutes ahead [1]. The traffic parameters in forecasting are commonly flow, occupancy and speed [2]. The accurate forecasting results for these parameters can help traffic control center in transportation systems to reduce traffic congestion and to improve the mobility of transportation [3]. However, it is quite difficult to predict effectively and accurately short-term traffic information due to the strongly nonlinear, complex and uncertain characteristics of traffic information.

Many short-term traffic forecasting methods have been developed by the scholars and practitioners to increase the forecasting accurancy in the past few years. Generally, the short-term traffic forecasting methods can be divided into three categories: statistical, artificial intelligence and hybrid methods. In the conventional statistical methods, Kalman filtering [4], autoregressive integrated moving average (ARIMA) [5] and seasonal ARIMA [6] have been applied to forecast short-term traffic flow based on past data. Although these statistical techniques can achieve reasonable prediction accuracy, they may not capture the dynamics and nonlinearities existed in traffic flow. To address this issue, the artificial intelligence methods, namely, neural networks (NNs) [7] [8] [9] [10], are widely used to the predication of short-term traffic flow due to its advantages of non-linear mapping relations [11]. The NNs methods can obtain forecasting results with better accuracies than the conventional statistical methods. However, NNs require a great deal of training data and relatively long training period for robust generalization [12] [13]. To enhance the generalization capability of NNs, several previous studies have proposed hybrid NNs models by incorporating particle swarm optimization (PSO) [14], genetic algorithms [15], fuzzy logical [16], Kalman filters [17], the ARIMA model [18], and so on. These hybrid models have been demonstrated to possess better performance than pure NNs in short-time traffic forecasting, but they use more parameters, have more complex structure and require greater computational power [19].

Grey system theory was first introduced by Deng in early 1980s [13]. The theory is simple and requires only a limited amount of data to estimate the behavior of uncertainty system. In grey system theory, the most commonly used forecasting model is GM(1,1) which can be divided into three stages [20]. In the first stage, to smooth the randomness, the primitive data obtained from the system to form the GM(1,1) is subjected to an operator, named Accumulating Generation Operator (AGO) [21]. The differential equation of GM(1,1) is then solved to obtain the n-step ahead predicted value of the system in the second stage. Finally, using the predicted value, the Inverse Accumulating Generation Operator (IAGO) is applied to find the predicted values of original data. However, the GM(1,1) model do not account of nonlinear [22]. Since neural networks possess characteristics of adaptability, nonlinearity and arbitrary function mapping capability [23], some scholars proposed a model combining grey system theory and neural networks, called GNN. The GNN model has been widely used to forecast in various fields, such as power system load forecasting, city waste water forecasting, etc. But, the shortcoming of the GNN model is that it is very difficult to select appropriate parameter values which have an effect on the performance of forecasting. The optimization technique can be used to solve this problem. Fruit fly

Manuscript received April 11, 2017; revised April 18, 2017. This work was supported by the Doctoral Scientific Research Foundation of Hubei University of Technology under Grant BSQD14036.

Y. Pan is with School of Computer Science, Hubei University of Technology, Wuhan, 430070, China (e-mail: pyy_wht@163.com).

Y. Shi is with School of Business, Macau University of Science and Technology, Macao SAR, China (corresponding author e-mail: ydshi@must.edu.mo).

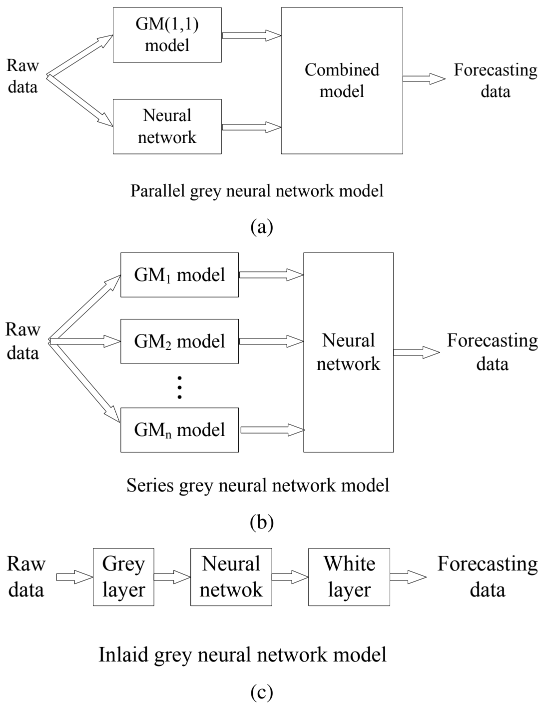

Fig. 1. The structures of three grey neural network models.

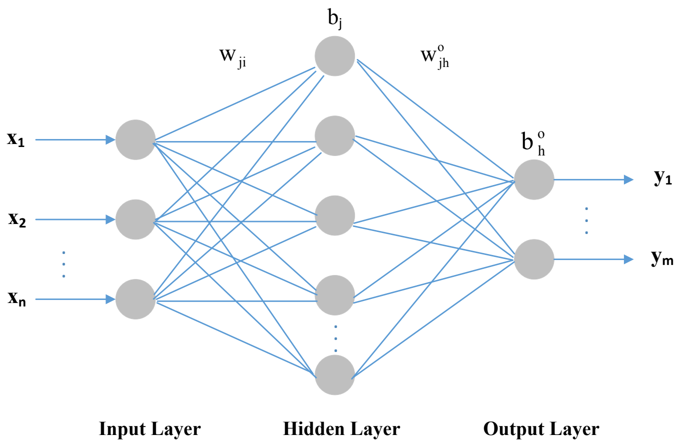

Fig. 2. The structure of BPNN.

optimization algorithm (FOA) proposed by the scholar Pan is a novel evolutionary computation and optimization technique [24]. FOA has the advantages of simple computation, being easy to be written into program code and to understand compared with other optimization algorithms, such as genetic algorithm, ant colony optimization algorithm, particle swarm optimization algorithms [25], etc. Therefore, this paper proposes a new hybrid model of grey neural network and fruit fly optimization algorithm, namely, GNN-FOA, for short-term traffic forecasting. The proposed hybrid model has the characteristics of grey system theory requiring less data, the strong non-linear map of neural networks and the quick-speed convergence of fruit fly optimization algorithm. Furthermore, the structure of GNN-FOA is relatively simpler since it only executes an accumulated generating operation (AGO) and an inversely accumulated generation operation (IAGO), which are derived from grey system theory, before and after neural network, respectively.

This rest of this paper is organized as follows. Section II introduces the models of grey neural network and FOA, then illustrates the mechanism of the GNN-FOA model in details. Section III discusses and compares the experimental results obtained by the GNN-FOA model and by other forecasting models. Section IV finally concludes this paper.

II. Grey Neural Network with Fruit Fly Optimization Algorithm

A. Grey neural network

Grey neural network combining GM(1,1) model in grey theory and neural networks can be classified into three categories: parallel, series and inlaid grey neural network [26]. Fig. 1 shows the structures of the three grey neural network models.

In parallel grey neural network model, firstly GM(1,1) and neural network are adopted separately to forecast, and then the forecasting data obtained by them are combined into a new result by some effective ways. The new result is the prediction value. In series grey neural network model, for the same series, different raw data are input into GM(1,1) models and then different output data are obtained. These output data are used as the input data for neural network. The output data of neural network is the final combined forecasting result.

The inlaid grey neural network model embeds GM(1,1) model and neural network into a single model. In this paper, the inlaid grey neural network model is utilized for short-term traffic forecasting because GM(1,1) model combines with neural network more closely in inlaid grey neural network model compared with other GNN models. This model has three basic parts: a grey layer, a general neural network, and a white layer. The grey layer before neural input nodes has accumulated generating operation (AGO) to initial input data. Let the original data sequence be denoted by $X^{(0)}$.

$$X^{(0)} = \{x^{(0)}(1), x^{(0)}(2), ..., x^{(0)}(n)\}, n \geq 4. \tag{1}$$

On the basis of $X^{(0)}$, a new data sequence $X^{(1)}$ is generated by AGO as

$$X^{(1)} = \{x^{(1)}(1), x^{(1)}(2), ..., x^{(1)}(n)\} \tag{2}$$

where $x^{(1)}(k) = \sum_{i=1}^{k} x^{(0)}(i), k = 1, 2, 3, ..., n$.

Then these new data $X^{(1)}$ generated by the accumulated generating operation are feed into the network. In this paper, a three-layer back-propagation neural network (BPNN) is adopted. There are two reasons: (i) BPNN contains many interacting nonlinear neurons in multiple layers and can capture complex phenomena [27], hence it is one of the most prevalent neural networks in short-term traffic forecasting; (ii) the previous study [8], [9] have proved that the three-layer BPNN can realize non-linear map for traffic flow. The parameters of BPNN include weights and biases which are adjusted iteratively by a process of minimizing the global error or the total square error. The structure of BPNN can be seen in Fig. 2, where w_{ji} represents the connection weight from input node i to hidden node j, w^o_{jh} represents the connection weight from hidden node j to output node h, b_j stands for the bias of hidden node j and b^o_h stands for the bias of output node h.

Finally, the white layer after neural output nodes inverses accumulated generation to the output data of the network $\widehat{x}^{(1)}$. Therefore, the prediction value $\widehat{x}^{(0)}$ we need is obtained.

$$\widehat{x}^{(0)}(k+1) = \widehat{x}^{(1)}(k+1) - \widehat{x}^{(1)}(k) \tag{3}$$

where $\widehat{x}^{(1)}(k+1)$ and $\widehat{x}^{(1)}(k)$ is the output data of the network at time $k+1$ and k, respectively. $\widehat{x}^{(0)}(k+1)$ is the prediction value at time $k+1$ based on original data.

B. Fruit fly optimization algorithm

Fruit fly optimization algorithm was proposed by Pan for finding global optimization, which is a new interactive evolutionary method inspired by the food finding behavior of the fruit fly. Fruit fly is a kind of insect and is superior to other species in osphresis and vision. Fruit fly can even smell food source from 40 km away. The food finding process of fruit fly is as follows: firstly, it smells the food source by osphresis organ, and flies towards that location; then, after it gets close to the food location, the sensitive vision is also used for finding food and other fruit flies flocking location, and it flies towards that direction [28]. To describe it in details, FOA can be divided into several steps as followings:

Step 1. Initialize randomly fruit fly swarm location (X_axis, Y_axis).

Step 2. Give the random direction and distance for food finding of an individual fruit fly using osphresis.

$$X_i = X\_axis + RandomValue \tag{4}$$

$$Y_i = Y\_axis + RandomValue \tag{5}$$

Step 3. Estimate the distance of the individual fruit fly i to the origin ($Dist_i$), then calculate the smell concentration judgment value (S_i), and the value of S_i is the reciprocal of distance.

$$Dist_i = \sqrt{X_i^2 + Y_i^2} \tag{6}$$

$$S_i = 1/Dist_i \tag{7}$$

Step 4. Calculate the smell concentration ($Smell_i$) of the individual location of the fruit fly by substituting smell concentration judgment value (S_i) into smell concentration judgment function (or called Fitness function), then find out the fruit fly with maximal or minimal smell concentration (finding the maximal or minimal value) among the fruit fly swarm.

$$Smell_i = Function(S_i) \tag{8}$$

$$[bestSmell\ bestIndex] = Max/Min(Smell_i) \tag{9}$$

Step 5. Keep the best smell concentration value and x, y coordinate, and at this moment, the fruit fly swarm will use vision to fly towards that location. If the maximum number of generations is not reached, go back to Step 2; otherwise, end the algorithm.

$$Smellbest = bestSmell \tag{10}$$

$$X\_axis = X(bestIndex) \tag{11}$$

$$Y\_axis = Y(bestIndex) \tag{12}$$

C. The GNN-FOA model

In this paper, the proposed hybrid model GNN-FOA for short-term traffic forecasting not only combines grey model with neural networks to form Grey Neural Network (GNN), but also optimizes the GNN model with fruit fly optimization algorithm.

In the GNN-FOA model, there are three stages. The first stage initializes the original traffic series data by accumulating generation operator (AGO) in grey model in order to weaken the randomness of the original data, and then these data are feed into a back-propagation neural network (BPNN). The second stage focuses on the optimization of BPNN by using FOA to decide the optimal original value of parameters, such as weight and bias, in BPNN. In this stage, the structure of three-layer BPNN $N_{in} * N_{hindden} * 1$ is first determined, where N_{in} refers to the number of data input from the first stage, 1 refers to the number of output forecasting value, N_{hidden} refers to the node number of hidden layer. According to the structure, the number of fruit fly swarm N_s is then decided, where N_s is equal to the number of all weights and biases in BPNN. The second stage aims to improve the robustness and accuracy of BPNN for short-term traffic forecasting. In the third stage, to obtain the predicted values, the output data from the optimized BPNN are processed by the inverse accumulating generation operator (IAGO) in grey model. The flowchart of the GNN-FOA model is shown in Fig. 3.

next, The steps of the GNN-FOA model in the second stage are described in details as follows:

Step 1 Parameters initialization
The parameters, such as the maximum number of iterations $maxgen$, the population size of swarm $sizepop$, the initial fruit fly swarm location (X_axis, Y_axis) and the random flight distance range FR, should be decided.

Step 2 Preliminary calculations
Give the random direction and distance for food finding of an individual fruit fly based on FR, and then calculate the flight distance of the fruit fly i to the origin $Dist_i$ and the smell concentration judgment value S_i by using Eqs. (6) and (7). The value of S_i is input into the BPNN model for short-term traffic forecasting. According to the results of forecasting, calculate the smell concentration $Smell_i$ which is employed the root-mean-square error (RMSE). RMSE measures the deviation between the forecasting value and the actual value, which is expressed as follows.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \widehat{x}_i)^2} \tag{13}$$

where n is the total number of the forecasting value, x_i and $\widehat{x}_i$ is the actual value and the forecasting value at the ith time interval, respectively. Finally, find out the minimum value of $Smell_i$.

Step 3 Iterations
Update fruit fly swarm location (X_axis, Y_axis) according to the minimal $Smell_i$. And calculate

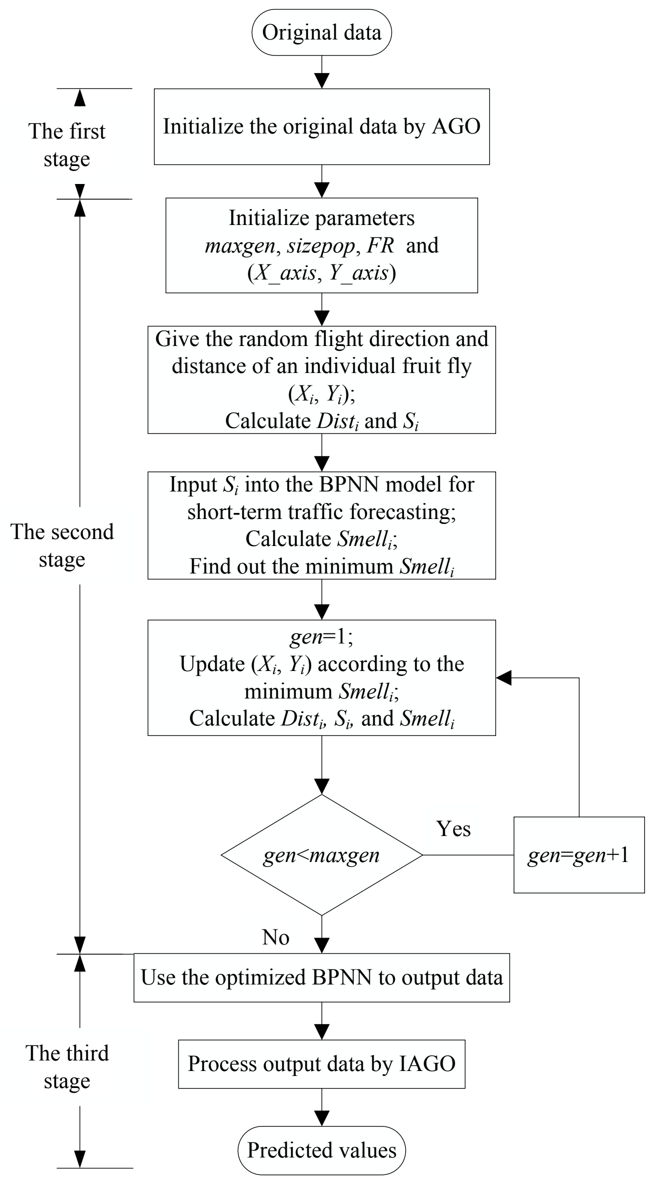

Fig. 3. The flowchart of the GNN-FOA model.

$Dist_i$, S_i and $Smell_i$ by using Eqs. (6-8) again. Then, set $gen = gen + 1$.

Step 4 Circulation stops

If gen reaches the maximum number of iterations, stop circulation and obtain the optimal parameters value of the BPNN model. Otherwise, repeat Step 3.

III. EXPERIMENT RESULTS

In this section, the proposed hybrid model GNN-FOA is used to forecast the future vehicular speed on Barbosa road in Macao during the period of PM 14:00-15:00, 19th August. The forecasting results are compared with that obtained from the single models, such as GM(1,1) and BPNN, and the combined model, such as grey neural network (GNN) [29] and grey neural network with particle swarm optimization (GNN-PSO) [30].

A. Traffic Data

The speed data used for forecasting was collected by the real-time traffic information system which has been developed by intelligent transportation systems (ITS) research laboratory of Macau University of Science and Technology. This system adopts dynamic data-collecting technology. The GPS terminals in 140 vehicles send data of the position and the speed of vehicles to the lab server in every 1 minute. The received data are saved in MySQL database. By calculating the arithmetic mean of the speed data in every last 3 minutes, we get the average real-time speed of vehicles in the road. So, we get 60 speed data totally that are calculated in every 1 minute between 14:00-15:00.

B. Traffic forecasting

To evaluate the performance of the GNN-FOA model for short-term traffic forecasting, the time series data from 14:00 to 15:00 were divided into two sub-sets. The first sub-set of the time series data, namely training data, collected from 14:00-14:50, were used for training the GNN-FOA model, the second sub-set of that, namely test data, collected from 14:51:15:00, were used to evaluate the generalization capability of this model. Whether training or evaluating, this model selects a time series data with equal dimension to predict the next data. The so-called equal dimension means, for a time series data, after predicting one traffic speed data, a new datum is added to the sequence at the end, meanwhile the oldest datum from the head of the sequence is take out, which results a new time series data with same dimension is generated to forecast the next traffic speed data. For example, using a time series data $\{x^{(0)}(k), x^{(0)}(k+1), ..., x^{(0)}(k+4)\}$, the GNN-FOA model predicts the value after one sampling times (i.e., 1 minutes) $x^{(0)}(k+5)$. In the next steps, the first data is always shifted to the second. It means that the model uses a time series data $\{x^{(0)}(k+1), x^{(0)}(k+2), ..., x^{(0)}(k+5)\}$ to forecast the value of $x^{(0)}(k+6)$. In this way, the new superseding the old, forecasting one by one, all need prediction results can be obtained. In this paper, the dimension of the time series data is set to 5. This means there are 45 and 10 pieces of speed data to train and validate the GNN-FOA model, respectively.

In the GNN-FOA model, a time series data with 5 data are first selected to generate a new sequence by accumulated generating operation (AGO). The new sequence is used as the input data of BPNN. According to the number of input and output data, the structure of BPNN is decided as $5 * N_{hidden} * 1$. Since the node number of hidden layer N_{hidden} is not the goal of this paper, we use one method recommended in [31], where $N_{hidden} \approx \log_2(N_{tr})$ with N_{tr} be the number of training data. As mentioned above, N_{tr} is equal to 45, so N_{hidden} is 6 in this case, i.e., $\log_2(45) \approx 6$. Based on the structure of BPNN, the number of fruit fly swarm in FOA is set to 43. The maximum number of iterations is predefined as 100. The population size of swarm is set to 20. The initial swarm location range is $[-1, 1]$. The random fly distance and direction range is $[-10, 10]$. Fig. 4 shows the iterative RMSE trend of the GNN-FOA model for searching optimal parameters. After 100 times of iterative evolution, the convergence can be seen in generation 22 and the RMSE value is 0.6159. We use the optimum obtained by FOA as the initial parameters of BPNN, and then train again the optimized BPNN based on training data. The trained BPNN is used to forecast the last 10 data.

TABLE I
THE COMPARISONS OF FORECASTING RESULTS

Time	Actual data	GNN-FOA		GM		BPNN		GNN		GNN-PSO	
		Result	Error(%)	Result	Error(%)	Result	Error(%)	Result	Error(%)	Result	Error(%)
14:51	21.245	18.76	11.70	19.049	10.34	18.474	13.04	20.785	2.17	20.039	5.68
14:52	29.275	23.236	20.63	21.789	25.57	24.376	16.74	22.196	24.18	23.429	19.97
14:53	30.92	228.795	6.87	29.248	5.41	33.057	6.91	27.791	10.12	28.928	6.44
14:54	27.95	26.03	6.87	28.209	0.93	38.172	36.57	30.374	8.67	29.215	4.52
14:55	21.75	21.844	0.43	22.035	1.31	32.752	50.58	25.243	16.06	23.128	6.34
14:56	14.34	12.837	10.48	18.908	31.86	21.957	53.12	13.391	6.62	13.649	4.82
14:57	10.88	11.347	4.29	14.394	32.30	13.123	20.61	14.354	31.93	12.784	17.50
14:58	15.80	11.067	29.96	12.244	22.51	7.9807	49.49	8.387	46.92	10.055	36.36
14:59	10.94	11.735	7.27	14.144	29.29	10.336	5.52	12.428	13.60	14.057	28.49
15:00	10.76	10.799	0.36	11.893	10.53	11.748	9.18	12.351	14.78	12.469	15.89

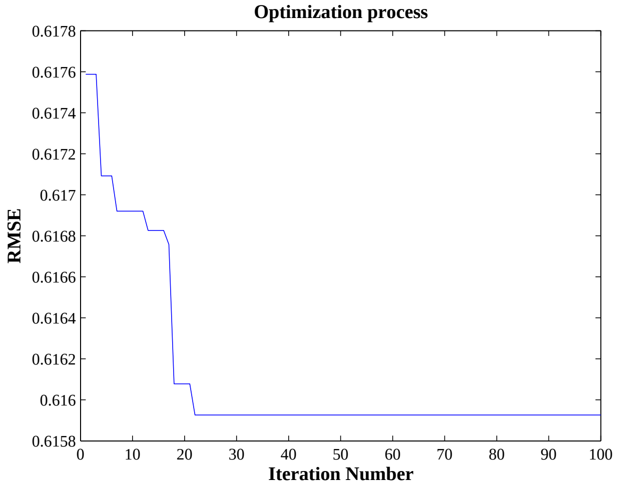

Fig. 4. The iterative RMSE trend of the GNN-FOA model for searching optimal parameters.

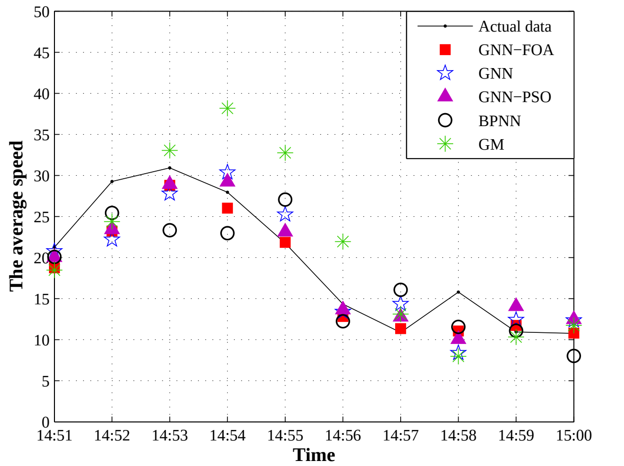

Fig. 5. The forecasting results of GNN-FON, GM, BPNN, GNN, GNN-PSO models.

Finally, we deal with the output 10 data of neural network by inverse accumulated generation operation (IAGO) to get the prediction value we need.

C. Evaluation and comparison

The performance of the GNN-FOA model is compared with the following models.

1) The single GM(1,1) model. Like the GNN-FOA model, the GM(1,1) model is built by using a time series data with 5 data and applying the equal dimension. But, due to no requirement of training neural network, the GM(1,1) model takes a time series data start from 14:46 to build grey model and forecast one data. This model iterates 10 times, then gives the last 10 speed forecasting results.
2) The single BPNN model. The structure and parameters of BPNN are identical to those of GNN-FOA. For instance, adopting $5*6*1$ neural network, using Tansig as the transfer function between input layer and hidden layer, and applying Pureline for output layer. However, unlike GNN-FOA, the initial neurons connection weights of the single BPNN model are set to stochastic real number belonging to [-1, 1].
3) The grey neural network (GNN) model. As mentioned above, the GNN model has three basic parts: a grey layer, a back-propagation neural network (BPNN), and a white layer [26]. The grey layer executes accumulated generating operation (AGO) to initialize input data. These new data generated by AGO are then feed into BPNN. Finally, the white layer inverses accumulated generation to the output data of BPNN. Therefore, the prediction value we need is obtained. The parameters used in GNN are identical to those used in GNN-FOA, except that there is no optimization in GNN by FOA.
4) The grey neural network with particle swarm optimization (GNN-PSO) model. Its structure is the same as that of the GNN-FOA model. The parameters are used in the GNN-PSO model as follows: the dimension of search space in PSO is 43, the population size of swarm is 20, the maximum number of iterations is 100, ω is set to 0.8 and the training errors are below 10%.

All the five models for short-term traffic forecasting execute in the same environment which includes Matlab 7.0 and the computer with the Intel(R) Core(TM)i5-4200U 1.6 GHz CPU, 4 GB RAM and Windows 7 professional system. The training time of the five models, i.e., GNN-FOA, GM, BPNN, GNN, GNN-PSO, is 36s, 2s, 11s, 62s, 47s, respectively. The GNN-FOA and GNN-PSO models use longer time than the GM and GNN models since they need to determine the parameters in the each generation. But, the GNN-FOA and GNN-PSO models use shorter time than the BPNN model because the parameters are not optimized and training data are not processed by AGO in the BPNN model. Moreover, the training time of the GNN-FOA model is shorter than that of the GNN-PSO model. The forecasting results and relative errors of the five models are as shown in Table I and Fig. 5.

To investigating the viability of short-term traffic forecasting models, several performance measures have been applied in the previous studies. In this paper, the mean absolute percentage error (MAPE) and the mean square error (MSE) [19] [32], shown as Eqs. (14) and (15), are used as the measures for comparison in these forecasting models. MAPE and MSE respectively reflect the mean prediction accuracy and stability. The smaller MAPE is, the more accurate the prediction is. Similarly, the smaller MSE is, the more stable the prediction is.

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}|\frac{\widehat{y}(i) - y(i)}{y(i)}| \tag{14}$$

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(\widehat{y}(i) - y(i))^2 \tag{15}$$

where $y(i)$ is the actual value at the ith time interval, $\widehat{y}(i)$ is the forecasted value at the ith time interval, and n is the total number of the forecasted value.

Table II summarizes the results of MAPE and MSE values obtained by the five models, i.e., GNN-FOA, GM(1,1), BPNN, GNN and GNN-PSO. It can be seen that the MAPE value of GNN-FOA (9.89%) is the lowest, and decreases 62.23%, 50.89%, 43.52%, 32.28% compared to that of GM(1,1), BPNN, GNN and GNN-PSO, respectively. This implies that the GNN-FOA model can improve the accuracy of forecasting short-term traffic speed. Furthermore, as shown in Table II, the MSE value of GNN-FOA, GM(1,1), BPNN, GNN and GNN-PSO, is 7.64, 38.73, 18.34, 15.09, 9.28, respectively. This indicates that GNN-FOA is more stable than other four models in short-term traffic forecasting. In a word, the results of MAPE and MSE show that the proposed hybrid model GNN-FOA has the higher prediction accuracy and stability among the five models.

TABLE II
THE MAPE AND MSE VALUES OF DIFFERENT MODELS

Criteria	GNN-FOA	GM	BPNN	GNN	GNN-PSO
MAPE (%)	9.89	26.18	20.13	17.5	14.6
MSE	7.64	38.73	18.34	15.09	9.28

IV. CONCLUSION

An accurate and stable short-term traffic forecasting model is very important for transportation systems. However, after the investigation on most of short-term traffic forecasting models, we find four issues: (i) a great deal of history data; (ii) the non-linear characteristic of traffic data; (iii) the slow speed of convergence; (iv) great parameters and complex structure. To address these issues, this paper proposed a new hybrid model combining grey neural network and fruit fly optimization algorithm, namely, GNN-FOA, for short-term traffic forecasting. In this proposed model, the FOA is used to select the appropriate parameter value of the GNN model, thereby improving the accuracy and stability of forecasting model. The proposed model can be divided into three stages. In the first stage, the original traffic data are initialized by AGO in grey model to generate a new data sequence which is input into a BPNN. In the second stage, the BPNN is optimized by the FOA algorithm. In the third stage, the data output from the optimized BPNN are processed by IAGO in grey model, thus the predicted values are obtained. To evaluate the performance of the GNN-FOA model, it is used to predict short-term traffic speed on Babosa road in Macao. The experimental results show: (i) the optimized GNN models, such as GNN-FOA and GNN-PSO, outperform the single GM(1,1) model, the single BPNN model, the combined model of them, i.e., the GNN model, in short-term traffic forecasting; (ii) compared with the GNN-PSO model, the training time of the GNN-FOA model is shorter and the values of MAPE and MSE are both smaller, which means the GNN-FOA model has better performance than the GNN-PSO model on short-term traffic forecasting. Of course, like most studies, the structure of GNN-FOA, such as the dimension of the time series data input into BPNN and the number of hidden nodes, is required to be pre-defined and fixed, which can not guarantee the optimal one will be obtained. How to optimize the structure of GNN-FOA with respect to time is left for our future work.

REFERENCES

[1] K. Y. Chan, T. S. Dillon, and E. Chang, "An intelligent particle swarm optimization for short-term traffic flow forecasting using on-road sensor systems," *IEEE Transactions on Industrial Electronics*, vol. 60, no. 10, pp. 4714–4725, Oct. 2013.

[2] E. I. Vlahogianni, J. C. Golias, and M. G. Karlaftis, "Short-term traffic forecasting: Overview of objectives and methods," *Transport Reviews*, vol. 24, no. 5, pp. 533–557, Sep. 2004.

[3] Y. Wang, M. Papageorgiou, and A. Messmer, "A real-time freeway network traffic surveillance tool," *IEEE Transactions on Control Systems Technology*, vol. 14, no. 1, pp. 18–32, Jan. 2006.

[4] I. Okutani and Y. J. Stephanedes, "Dynamic prediction of traffic volume through kalman filtering theory," *Transportation Research Part B: Methodological*, vol. 18, no. 1, pp. 1–11, Feb. 1984.

[5] B. L. Smith, B. M. Williams, and R. K. Oswald, "Comparison of parametric and nonparametric models for traffic flow forecasting," *Transportation Research Part C: Emerging Technologies*, vol. 10, no. 4, pp. 303–321, Aug. 2002.

[6] B. M. Williams, P. K. Durvasula, and D. E. Brown, "Urban freeway traffic flow prediction: Application of seasonal autoregressive integrated moving average and exponential smoothing models," *Transportation Research Record: Journal of the Transportation Research Board*, 1998.

[7] M. Dougherty, "A review of neural networks applied to transport," *Transportation Research Part C: Emerging Technologies*, vol. 3, no. 4, pp. 247–260, Aug. 1995.

[8] S. C. Chang, R. S. Kim, S. J. Kim, and M. H. Ahn, "Traffic flow forecasting using a 3-stage model," in *Proc. the IEEE Intelligent Vehicles Symposium 2000*, Dearborn, USA, 2000, pp. 451–456.

[9] H. Dia, "An object-oriented neural network approach to short-term traffic forecasting," *European Journal of Operational Research*, vol. 131, no. 2, pp. 253–261, Jun. 2001.

[10] C. Ledoux, "An urban traffic flow model integrating neural network," *Transportation Research Part C: Emerging Technologies*, vol. 5, no. 5, pp. 287–300, Oct. 1997.

[11] S. Mao and X. Xiao, "Short-term traffic flow grey forecasting model gm(1,1 vertical bar tan(k-tau)p, sin(k-tau)p) of single-cross-section and its particle swarm optimization," *JOURNAL OF GREY SYSTEM*, vol. 22, no. 4, pp. 383–394, 2010.

[12] T. C. Jo, "The effect of virtual term generation on the neural based approaches to time series prediction," in *Proc. 4th International Conference on Control and Automation (ICCA '03)*, Montreal, Que., Canada, 2003, pp. 516–520.

[13] E. Kayacan, B. Ulutas, and O. Kaynak, "Grey system theory-based models in time series prediction," *Expert Systems with Applications*, vol. 37, no. 2, pp. 1784–1789, Mar. 2010.

[14] J. Hu and X. Zeng, "A hybrid pso-bp algorithm and its application," in *Proc. 2010 Sixth International Conference on Natural Computation (ICNC 2010)*, vol. 5, Yantai, Shandong, China, 2010, pp. 2520–2523.

[15] E. I. Vlahogianni, M. G. Karlaftis, and J. C. Golias, "Optimized and meta-optimized neural networks for short-term traffic flow prediction: a genetic approach," *Transportation Research Part C: Emerging Technologies*, vol. 13, no. 3, pp. 211–234, Jun. 2005.

[16] C. Quek, M. Pasquier, and B. B. S. Lim, "Pop-traffic: A novel fuzzy neural approach to road traffic analysis and prediction," *IEEE Transactions on Intelligent Transportation Systems*, vol. 7, no. 2, p. 133C146, Jun. 2006.
[17] D. Srinivasan, C. W. Chan, and P. G. Balaji, "Computational intelligence based congestion prediction for a dynamic urban street network," *Neurocomputing*, vol. 72, no. 10, pp. 2710–2716, 2009.
[18] M.-C. Tan, S. C. Wong, J.-M. Xu, Z.-R. Guan, and P. Zhang, "An aggregation approach to short term traffic flow prediction," *IEEE Transactions on Intelligent Transportation Systems*, vol. 10, no. 1, p. 60C69, Feb. 2009.
[19] 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, Jan. 2013.
[20] C. Dai, D. Pi, Z. Fang, and H. Peng, "Wavelet transform-based residual modified gm(1,1) for hemispherical resonator gyroscope lifetime prediction," *IAENG International Journal of Computer Science*, vol. 40, no. 4, pp. 250–256, 2013.
[21] J. L. Deng, "Introduction to grey system theory," *The Journal of Grey System*, vol. 1, 1989.
[22] H. Guo, X. Xiao, and F. Jeffrey, "Urban road short-term traffic flow forecasting based on the delay and nonlinear grey model," *Journal of Transportation Systems Engineering & Information Technology*, vol. 13, no. 6, pp. 60–66, 2013.
[23] Y. Wei and M. C. Chen, "Forecasting the short-term metro passenger flow with empirical mode decomposition and neural networks," *Transportation Research Part C Emerging Technologies*, vol. 21, no. 1, pp. 148–162, 2012.
[24] W.-T. Pan, "A new fruit fly optimization algorithm: Taking the financial distress model as an example," *Knowledge-Based Systems*, vol. 26, pp. 69–74, Feb. 2012.
[25] Z. Zhong and D. Pi, "Forecasting satellite attitude volatility using support vector regression with particle swarm optimization," *IAENG International Journal of Computer Science*, vol. 41, no. 3, pp. 153–162, 2014.
[26] S.-Y. Chen, G.-F. Qu, X.-H. Wang, and H.-Z. Zhang, "Traffic flow forecasting based on grey neural network model," in *Proc. 2003 International Conference on Machine Learning and Cybernetics*, vol. 2, Xi'an, China, 2003, pp. 1275–1278.
[27] Y. Wei and M.-C. Chen, "Forecasting the short-term metro passenger flow with empirical mode decomposition and neural networks," *Transportation Research Part C*, vol. 21, 2012.
[28] H. Z. Li, S. Guo, C. J. Li, and J. Q. Sun, "A hybrid annual power load forecasting model based on generalized regression neural network with fruit fly optimization algorithm," *Knowledge-Based Systems*, vol. 37, no. 2, pp. 378–387, 2013.
[29] Y. dong Shi, Y.-Y. Pan, and J. qing Li, "Urban short-term traffic forecasting based on grey neural network combined model: Macao experience," in *Proc. 2009 International Workshop on Intelligent Systems and Applications (ISA 2009)*, Wuhan, China, 2009, pp. 1–4.
[30] Y. Pan and Y. Shi, "Short-term traffic forecasting based on grey neural network with particle swarm optimization," in *Lecture Notes in Engineering and Computer Science: Proceeding of The World Congress on Engineering and Computer Science 2016, WCECS 2016*, San Francisco, USA, 19-21 October 2016, pp. 752–757.
[31] G. Mirchandani and W. Cao, "On hidden nodes for neural nets," *IEEE Transactions on Circuits System*, vol. 36, no. 5, pp. 661–664, May 1989.
[32] Y. Cong, J. Wang, and X. Li, "Traffic flow forecasting by a least squares support vector machine with a fruit fly optimization algorithm," *Procedia Engineering*, vol. 137, pp. 59–68, 2016.