Journal of Network Intelligence ©2019 ISSN 2414-8105 (Online)
Taiwan Ubiquitous Information Volume **4**, Number **2**, May **2019**

A Traffic Prediction Method of Bicycle-sharing based on Long and Short term Memory Network

Shun-miao Zhang[1,2,4], Xin Su[1,3], Xin-hua Jiang[1,2,4], Ming-long Chen[1,2], Tsu-Yang Wu[5*]

[1]College of Information Science and Engineering
Fujian University of Technology, Fuzhou 350118, China
zshunmiao@163.com

[2]Fujian Provincial Key Laboratory of Big Data Mining and Applications &
College of Information Science and Engineering
Fujian University of Technology, Fuzhou 350118, China

[3]Fujian Key Laboratory for Automotive Electronics and Electric Drive
Fujian University of Technology, Fuzhou 350118, China
xjlw66@163.com

[4]National Demonstration Center for Experimental Electronic Information
and Electrical Technology Education, Fujian University of Technology, Fuzhou 350118, China

[5]College of Computer Science and Engineering,
Shandong University of Science and Technology, Qingdao, 266590, China
*Corresponding author: wutsuyang@gmail.com

Received December 2018; revised January 2019

Abstract. *The prediction of bicycle-sharing traffic flow plays an important role in management of the service stations for bicycle-sharing. In this paper, based on an improved long and short term memory network we propose a traffic prediction method of bicycle-sharing. Firstly, the bicycle-sharing area is divided by the probability threshold clustering algorithm. Secondly, based on the long and short term memory network the bicycle-sharing traffic model is established. In order to increase the accuracy of prediction, we optimize the related parameters by differential evolution algorithm. Finally, we adopt the traffic data set of bicycle-sharing in Yancheng city, Jiangsu province, China in our experiment. With compared to the existed methods, the experimental results are shown that our proposed method has a better effectivity and is suitable for the traffic prediction of bicycle-sharing.*
Keywords: Bicycle-sharing; Traffic prediction; Long and short memory network; Optimization

---

1. **Introduction.** Bicycle-sharing providing a convenient, low-carbon and healthy way for people is widely deployed in cities around the world. Users can rent a bicycle from a service station by scanning the code on their mobile phone. People's demand for bicycles at the bicycle-sharing service station is inconsistent, resulting in the fact that some service points have no bicycles to borrow, while others may have a large number of bicycles that are returned and cannot be accepted. Therefore, predicting the demand for rent and return bicycles at various service stations in advance (such as the number of bicycles borrowed from various service points in the next hour and the number of bicycles returned) will help to mobilize bicycles directly at different service stations in advance and achieve the balance of supply and demand. Because people's demand for bicycles is affected by

many complicated factors, such as the interaction between weather, holidays, events and service stations, it is a challenge to predict the number of bicycles in each service stations can be borrowed and returned.

Bicycle-sharing first appeared in Europe. Chen [1] and Shaheen et al. [2] analyzed past and present interests and losses, and looked forward to its future. Shared bicycle research can be classified into four categories: system design, system model analysis, system prediction and system operation. System traffic prediction is an important issue in shared bicycle research. Froehlich et al. [3] compared four simple prediction models used to predict the final availability of bicycles, historical mean, historical trend, and Bayesian networks for each station. Kaltenbrunner et al. used statistical models to predict the number of bicycles and terminals available at each service point. Borgnat et al. [5] used Lyon's shared bicycle data to propose a combined model to predict hourly traffic during the day. Vogel et al. [6, 7] used time series analysis to predict the bicycle demand in Vienna based on Dublin Shared cycle data. In [8] constructed a weighted correlation network to model the relationship among bike stations, and dynamically group neighboring stations with similar bike usage patterns into clusters. In [9, 15] proposed a hierarchical prediction model to predict the amount of bicycles rented and rented at each service point. First, the service points are clustered by two-layer clustering method, and the total bicycle traffic of the whole city is predicted by GBRT. A multi-phase similarity reasoning model was proposed to predict the ratio of rental bicycles across clusters and the transfer of bicycles across clusters. In [10], a time-space bicycle movement model based on historical bicycle sharing data is proposed, and a traffic prediction mechanism based on granularity per hour is designed. In the literature [11-14], the service points are divided into clusters according to the bicycle use situation, and the bicycle movement recognition mode is used.

Because the ride distance of bicycle-sharing is short, it is affected by weather, holidays, events, and service stations. Therefore, the rental forecast of bicycle-sharing at each service station has the characteristics of non-deterministic and random. In this paper, we based on long short-term memory network (LSTM) [16-17] to propose a traffic prediction method of bicycle-sharing. Then, we optimize the parameters by differential evolution algorithm and simulate the proposed method with a real dataset. Finally, the experimental results are shown that our method has a better effectivity and is suitable for the traffic prediction of bicycle-sharing.

The rest of this paper is organized as follows. In section 2, we briefly introduce the long short-term memory network and the traffic prediction of bicycle-sharing. Our method is proposed in section 3. In section 4, we make some experiments for our method. Meanwhile, the comparisons are shown in this section. Finally, the conclusions are drown in section 5.

2. Preliminaries.

2.1. **Long short-term memory (LSTM) network.** RNN is a neural network that contains loops that allow the persistence of information. It is often used to solve problems with known pre-event inferred follow-up events. Long Short Term Memory (LSTM) is a special type of RNN that can learn long-term dependency information. LSTM has achieved better results in many tasks than traditional RNN in recent years. Shared bicycle traffic forecasting is essentially a problem of predicting future traffic with known historical traffic. LSTM is good at learning long-term dependency information for shared bicycle traffic forecasts.

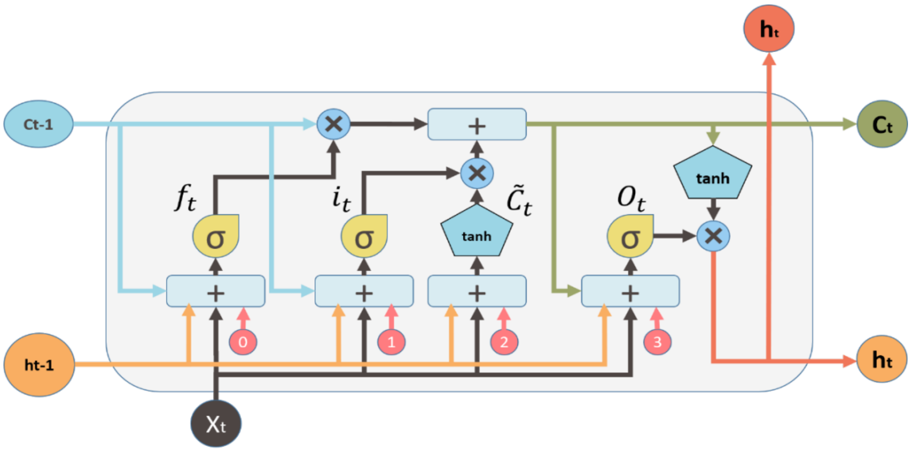

FIGURE 1. LSTM unit structure diagram

The LSTM network is made up of many LSTM units, as shown in the figure is a typical LSTM cell structure. The LSTM unit has three inputs, X_t is the current time step input, $h_{t-1}$is the output of the previous LSTM unit, and C_{t-1} is the "memory" of the previous LSTM unit. The LSTM unit has two outputs, the output h_t of the current LSTM unit and the "memory" C_t of the current LSTM unit. Therefore, the LSTM unit generates a new output and changes its memory by calculating the memory and output of the current input and the previous unit. In the figure, X represents an element multiplication operation, 0 means forgetting the old "memory", and 1 means letting the old "memory" pass. The + in the figure is a piecewise summation, which means the fusion of old and new "memory". The combination of "X+" controls the proportion of old and new memory fusion. 0, 1, 2, and 3 in the figure indicate four gate layer offsets.

Forget gate:

$$F_t = \sigma\left(W_f\left[h_{t-1}, X_t\right] + b_f\right) \tag{1}$$

The gate reads and outputs a value between 0 and 1 for each unit memory, where the number, 1 means "fully reserved" and 0 means "completely discarded".

Input gate:

$$i_t = \sigma\left(W_i\left[h_{t-1}, X_t\right] + b_i\right) \tag{2}$$

$$\tilde{C}_t = \tanh(W_C\left[h_{t-1}, X_t\right] + b_C) \tag{3}$$

Sigmoid determines what values will be updated. Then, create a new candidate value vector through the tanh layer.

Update gate:

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \tag{4}$$

This gate layer is used to update the current unit memory C_t.

Output gate:

$$O_t = \sigma\left(W_o\left[h_{t-1}, X_t\right] + b_o\right) \tag{5}$$

$$h_t = O_t * \tanh\left(C_t\right) \tag{6}$$

$$h_t = O_t * \tanh\left(C_t\right) \tag{7}$$

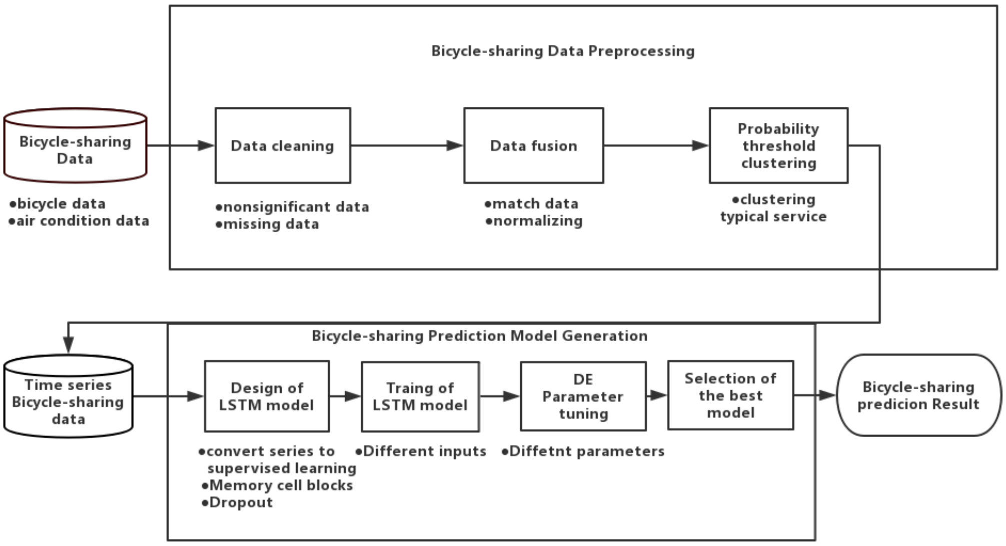

FIGURE 2. Flowchart of traffic prediction of bicycle-sharing

The sigmoid layer determines which part of the unit memory will be output. It is processed by tanh (getting a value between -1 and 1) and multiplying it by the output of the sigmoid gate to get the final output.

2.2. **Traffic prediction of bicycle-sharing.** The flowchart of traffic prediction of bicycle-sharing is depicted in Fig. 2. Firstly, the data preprocessing part cleans the original shared bicycle monitoring data to make up for the missing data and remove the non-significant data. Then the weather data is matched and fused with the shared bike monitoring data, and the data is standardized. Finally, the hierarchical probabilistic clustering algorithm is used to divide the traffic characteristics into similar regions for classification and prediction.

Shared bicycle flow forecasting model converts the pre-processed shared bicycle time series data into supervised learning sequence, which is convenient for LSTM model to predict and set the relevant parameters of LSTM. Then the LSTM model is trained by dividing the training set and the test set, and the differential evolution algorithm is invoked to iteratively optimize the LSTM parameters of the optimal prediction effect. Finally, the output shared bicycle traffic prediction results and corresponding LSTM model parameters are applied.

3. **Proposed prediction method of bicycle-sharing.**

3.1. **Data preprocessing.** Processing missing data: There are two types of missing data in the data set, one is missing site number and the other is missing date. There are two kinds of missing site numbers, one is due to the cancellation of service points caused by missing and the other is due to sensor failure and failed to collect. In the two cases above, we will eliminate these missing data. The other is a dynamic mean substitution method. Extract the missing date closest to the two day before and after the average compensation deficiency. The mean interpolation method can be expressed as:

$$\hat{y}_k = \text{average}\left(y_{k-2} + y_{k-1} + y_{k+1} + y_{k+2}\right) \tag{8}$$

The concrete implementation is as follows:

Algorithm 1 Dynamic mean replacement method

Input: service point original data D; service point original data size N; Date series Dt; Date series size dt;
Output: service point translate data S;

```
for i = 1 → dt do
    if Dt(i) in D then
        S(i) ← D(Dt(i))
    else
        S(i) ← sum(D(Dt(i−2)) + D(Dt(i−1)) + D(Dt(i+1)) + D(Dt(i+2)))/4
    end if
end for
return service point translate data S
```

3.2. **Probabilistic threshold clustering approach.** Bicycle-sharing is affected by weather, season, date, commuting time, holidays, service location, and other factors. The service stations in different regions have great influence on the bicycle-sharing traffic. Within a certain region, the lending and returning traffic of the service stations has obvious trend characteristics. In this paper, a probability threshold clustering approach is used to cluster multiple service stations into several service areas by sharing bike probability threshold clustering algorithm. It is helpful for effective classification and prediction of service points in different service areas. The specific process of shared bicycle probability threshold clustering algorithm is as follows:

Step 1: To generate shared bicycle and return traffic matrix F. The shared bicycle traffic data contains the travel data of N service points, that is, the start riding time, stop riding time, start riding service point, stop riding service point and so on. In this paper, we share the data of sharing bicycle journey from 1 to July. The construction of shared bicycle return flow matrix F is constructed as follows: The bicycle-sharing loan and return flow matrix *F(N, N)* has N starting service points in abscissa and N ending service points in ordinate. Take the first row element of the matrix as an example to show the number of journey data generated by starting point 1 to ending point 1 within the range from January to July. By sharing the bicycle return flow matrix F, we can better characterize the relationship between the service points and the traffic flow.

$$F = \begin{bmatrix} f_{11} & \cdots & f_{1n} \\ \vdots & \ddots & \vdots \\ f_{n1} & \cdots & f_{nn} \end{bmatrix} \tag{9}$$

Algorithm 2 generate traffic matrix

Input: Start riding service point data S ; End riding service point data E; Clustering service point number N ;
Output: traffic matrix F;

```
for i = 1 → N do
    F(i, :) ← S(i)
    for j = 1 → N do
        F(i, j) ← E(i)
    end for
end for
return generate traffic matrix F
```

Step 2: To generate shared bicycle service point transition probability matrix *B(N, N)*. In order to further digitize the relationship between service points, this paper calculates the transfer probability matrix B based on the shared bike lending and returning flow matrix F. The elements

$$\alpha_{ij} = \frac{f_{ij}}{f_{i1} + f_{i2} + .. + f_{in}} and\ \alpha_{i1} + \alpha_{i2} + \ldots + \alpha_{in} = 1 \tag{10}$$

The α_{ij} indicates that the starting point is I service point, and the termination point is the probability of J service station.

$$\mathrm{B} = \begin{bmatrix} \alpha_{11} & \cdots & \alpha_{1n} \\ \vdots & \ddots & \vdots \\ \alpha_{n1} & \cdots & \alpha_{nn} \end{bmatrix} \tag{11}$$

Algorithm 3 generate transition probability matrix

Input: traffic matrix F; Clustering service point number N ;
Output: transition probability matrix B;

```
for i = 1 → N do
    sum = 0 ;
    for h = 1 → N do
        sum = sum + F(i,h)
    end for
    for j = 1 → N do
        B(i, j) = F(i,j) / sum
    end for
end for
return transition probability matrix B
```

Step 3: To generate the clustering probability set of each service point $X = \gamma_1, \ldots, \gamma_n$. The probability set of service point clustering is generated by the class probability transfer matrix. Take No. 1 service point as an example, $\gamma_1 = \alpha_{11}, \gamma_2 = sum(\alpha_{12} + \alpha_{21})/2, \ldots,$ $\gamma_{n-1} = sum(\alpha_{1,n-1} + \alpha_{n-1,1})/2$, $\gamma_n = sum(\alpha_{1,n} + \alpha_{n,1})/2$. Finally, N clustering points of service points are generated.

Algorithm 4 generate Clustering probability set

Input: transition probability matrix B; Clustering service point number N ;
Output: Clustering probability set X;

```
for i = 1 → N do
    for h = 1 → N do
        X(i,h) = sum(B(i,h)+B(h,i))/2
    end for
end for
return Clustering probability set X
```

Step 4: to determine the association probability of service points. Traversing N service point cluster probability sets, when there is a certain service point cluster probability density set gamma value greater than the probability threshold β (probability threshold is chosen as the main empirical value, the probability threshold $\beta = 0.8$), it is judged that the two service points have strong correlation, suitable for clustering. Take the 1 service point as an example, the clustering probability set of point 1 is $X_1 = \gamma_1, ...\gamma_n$. When the probability threshold value is higher than the probability threshold value, it is judged that the No. 1 service station has a strong correlation with the No. n service station and is suitable for clustering.

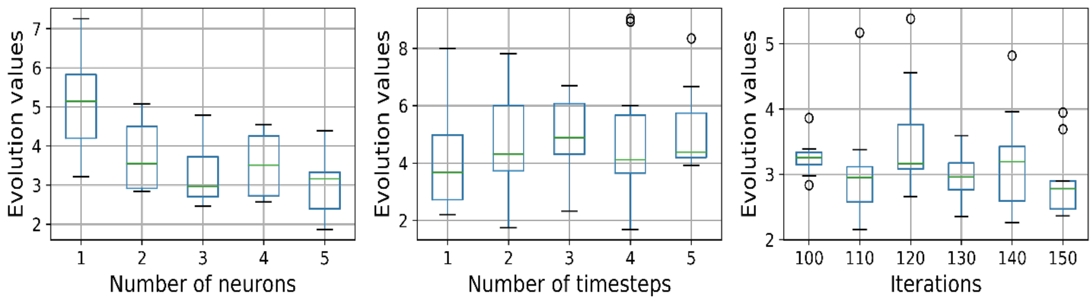

FIGURE 3. Experimental box diagram of prediction model parameters

Step 5: service station clustering. Two service points suitable for clustering are judged by Step 4, and the two service points are regarded as a large service point. Continue iterating the Step 1 to Step 5 steps, at which point the generated shared bike loan and return flow matrix is $F(N-1, N-1)$. With the continuous iteration process, the final clustering results are obtained.

Algorithm 5 judgment process

Input: Clustering probability set X Clustering service point number N ; Clustering probability threshold β;
Output: Clustering service point number N ; Clustering service point list L;

for $i = 1 \to N$ **do**
    for $h = 1 \to N$ **do**
        if $X(i,h) > \beta$ **then**
            $L \leftarrow i, h$
            N = N - 1
        end if
    end for
end for
return N , L

3.3. **Parameters optimization.** Because the prediction performance of LSTM model is also related to the selection of parameters, it is necessary to train the test to select the appropriate parameters of LSTM. In the training process, the recurrent neural network with LSTM link [17] is trained by error back propagation algorithm. The training objective function is the cost function of LSTM prediction model. The loss function shows the prediction performance of the algorithm. This section studies the effects of neuron number, time step and number of iterations on training effect.

1) According to the number of neurons box diagram, the value of the loss function decreases as a whole with the increase of the number of neurons. If the number of neurons increases too much, the speed of the algorithm will decrease, and the problem of overfitting is easy to occur. Shared bicycle flow is the result of multiple factors. Multi-neuron LSTM [18] can improve the nonlinear fitting ability of shared bicycle LSTM model.

Algorithm 6 DE-LSTM

Input: Population size M ; Scaling factor F; Crossover rate CR ; Dimension D; Genetation T; Series S; F(x) : LSTM evaluation function;
Output: The best vector(solution)- γ;

function EVALUATE(γ):
    fitness = F(γ,S);
return fitness
end function
for $i = 1 \to M$ **do**
    for $j = 1 \to D$ **do**
        $x_{i,t}^{(j)} = int(x_{min}^{(j)} + rand(0,1) \cdot (x_{max}^{(j)} - x_{min}^{(j)}))$
    end for
end for
while $(|EVALUATE(\gamma)| \geq \varepsilon) or (t \leq T)$ **do**
    for $i = 1 \to M$ **do**
        ▶ $(Mutation \quad and \quad Crossover)$
        for $j = 1 \to D$ **do**
            $v_{i,t}^{(j)} = Mutation(x_{i,t}^{(j)})$;
            $u_{i,t}^{(j)} = Crossover(x_{i,t}^{(j)}, v_{i,t}^{(j)})$;
            if $u_{i,t} < x_{min}^{(j)}$ **then**
                $u_{i,t} \leftarrow x_{min}^{(j)}$
            else if $v_{i,t} < x_{min}^{(j)}$ **then**
                $u_{i,t} \leftarrow x_{min}^{(j)}$
            end if
        end for
        ▶ $(Greedy \quad Selection)$
        if $f(u_{i,t}) < f(x_{i,t})$ **then**
            $x_{i,t} \leftarrow u_{i,t}$;
            if $f(x_{i,t}) < f(\gamma)$ **then**
                $\gamma \leftarrow u_{i,t}$;
            end if
        else
            $x_{i,t} \leftarrow x_{i,t}$;
        end if
    end for
    $t \leftarrow t + 1$;
end while
return the best vector vector γ ;

2) There is fluctuation uncertainty in the rule of time step based on the time step box diagram. Different time step will affect the effect of loss function. The appropriate time step helps to track information over a longer period of time. Shared bicycle traffic forecasting is to predict the shared bicycle traffic in August by the historical traffic data from January to July. Different time steps represent the memory cycle of LSTM network. Choosing the right memory cycle is conducive to the accurate prediction of shared bicycle traffic.

3) From the iteration number box diagram, it can be seen that increasing the number of iterations in a certain range can obtain better loss function results. But it may also lead to larger variance. Shared bicycle flow forecasting is highly nonlinear and uncertain. Short-term shared bicycle flow is vulnerable to random disturbances. By choosing appropriate iteration times, the algorithm can avoid falling into local optimum and predict short-term shared bicycle traffic better. In this paper, multi-objective differential evolution algorithm [19] is used to optimize the main performance parameters of LSTM: the number

Table 1. Comparison of performance index of full pile average error

Prediction method	RMSE	MAE	MedSE	MSLE
Proposed method (DE-LSTM)	**1.748930544**	**1.4775**	**1.33622**	**0.0223**
ARIMA	10.04354959	7.612839	6.03343	0.227147
GBDT	9.960228736	8.060933	7.093302	0.22822
RandomForest	13.86189303	11.92812	11.49633	0.331383

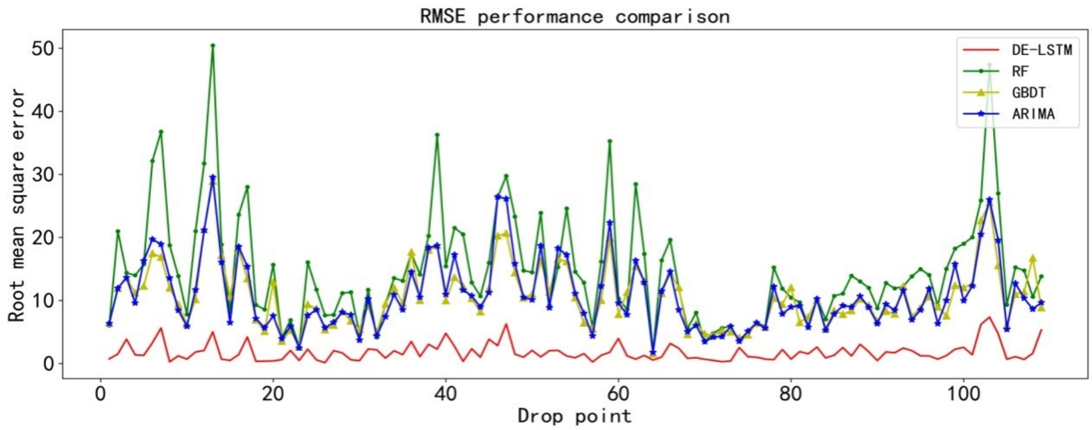

Figure 4. RMSE performance comparison

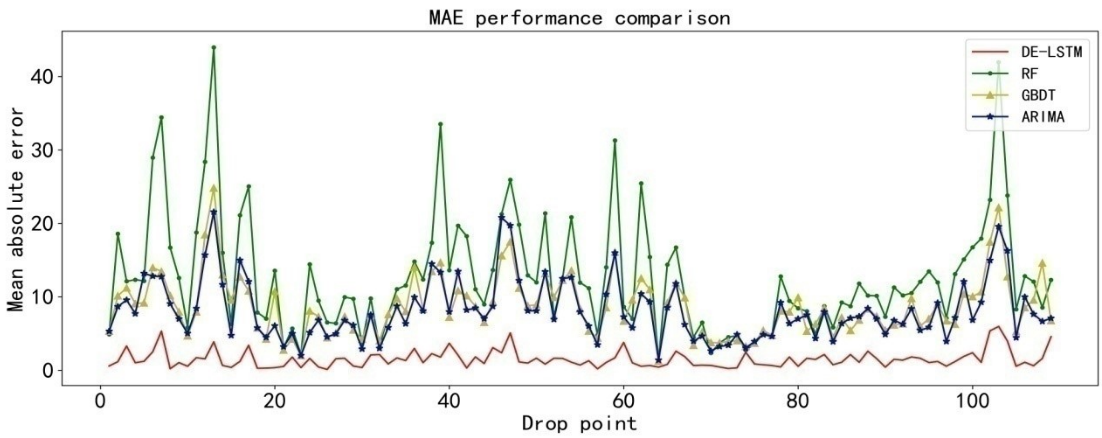

Figure 5. MAE performance comparison

of neurons, the time step, the number of iterations. The cost function of LSTM prediction model is called to evaluate the optimized parameters. The specific pseudo code and flowchart are as follows:

4. **Experimental results and analysis.** The data set was used to extract the data of Yancheng City in Jiangsu Province from the order data of Yongan Xing Company from January to August 2015. The LSTM regression prediction model was established by using the data from the first 1 to 7 months, and the data from the last 8 months were used to

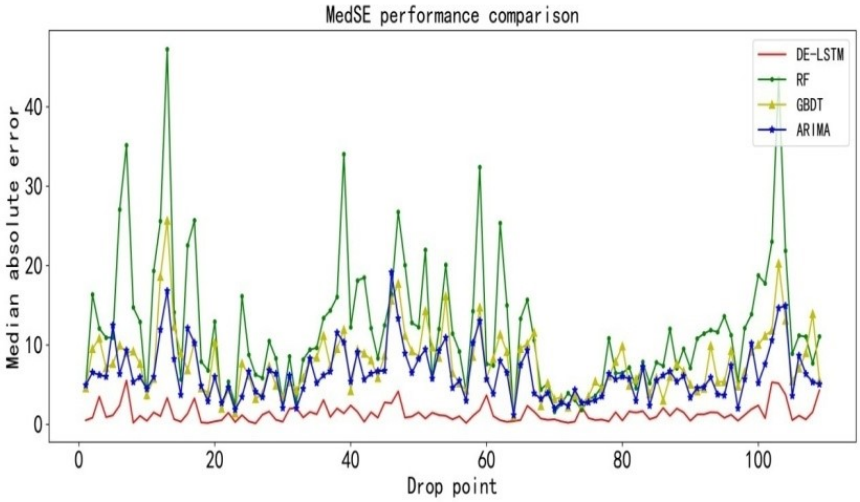

FIGURE 6. MedSE performance comparison

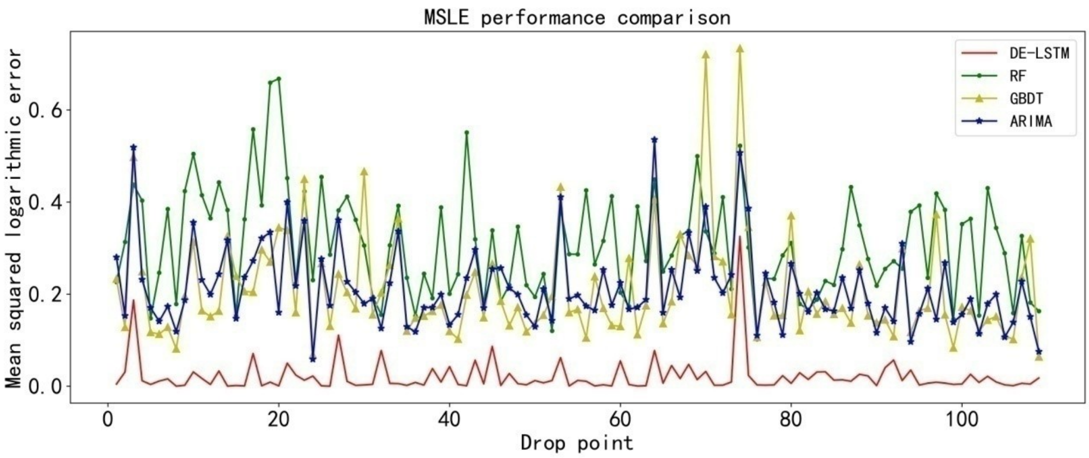

FIGURE 7. MSLE performance comparison

verify the accuracy of the prediction model. In this paper, the model is used to predict the total pile shared bicycle flow in August, and the Random Forest [21], ARIMA [22], GBDT [23] algorithms are compared. In order to compare the prediction results, this paper introduces the following 4 performance indicators:

1) RMSE (Root mean square error)

$$\mathrm{RMSEy}, \hat{y} = \sqrt{\frac{1}{n_{samples}} \sum_{i=0}^{n_{samples}-1} (y_i - \hat{y}_i)^2} \tag{12}$$

2)MAE (Mean absolute error)

$$\mathrm{MAEy}, \hat{y} = \frac{1}{n_{samples}} \sum_{i=0}^{n_{samples}-1} |y_i - \hat{y}_i| \tag{13}$$

3) MedSE(Median absolute error)

$$\mathrm{MedAEy}, \hat{y} = \mathrm{median}\left(|y_1 - \hat{y}_1|, \ldots, |y_n - \hat{y}_n|\right) \tag{14}$$

4) MSLE(Mean square logarithmic error)

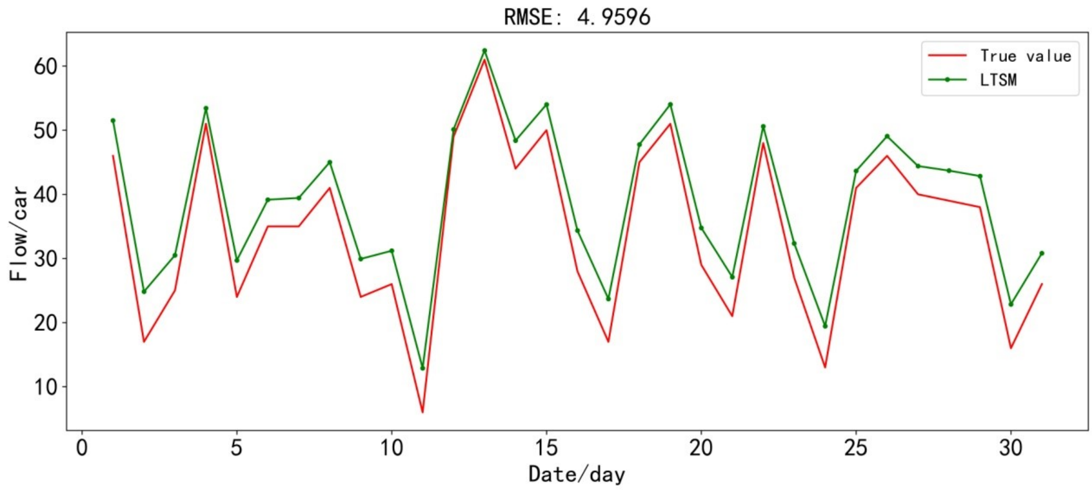

FIGURE 8. Prediction effect of traditional LTSM service points

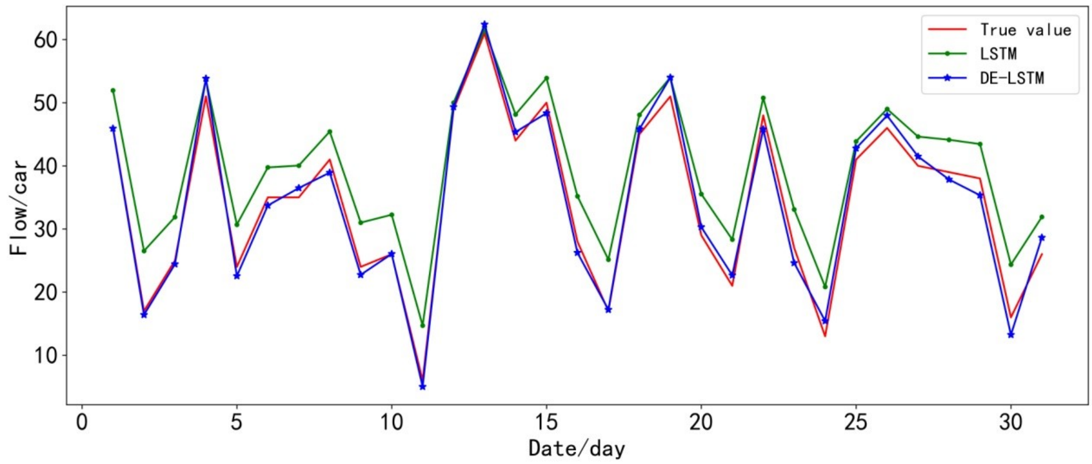

FIGURE 9. DE-LSTM service point prediction effect

$$\text{MSLEy}, \hat{y} = \frac{1}{n_{samples}} \sum_{i=0}^{n_{samples}-1} \left(\log_e\left(1+y_i\right) - \log_e\left(1+\hat{y}\right)\right)^2 \tag{15}$$

Formula:$n_{samples}$ is the number of service points, y_i is the actual value of network traffic, $\hat{y}_i$ is the prediction model of network traffic prediction value.

Table 1 shows the comparison of the average performance of RMSE, MAE, MedSE and MSLE in the same service area. Figures 4, 5, 6 and 7 show the performance of each pile and the RMSE, MAE, MedSE and MSLE in this method. Energy index is superior to other prediction methods.

In order to better display the performance of the algorithm, this paper also makes a monthly prediction of a service point in the same area, and compares Random Forest, ARIMA, GBDT algorithm through RMSE performance indicators. The effect is shown in Figure 8 and 9.

Illustrated by figures 8, 9 and 10, this method is superior to other methods in predicting service points. It can better predict the trend of shared bicycle traffic in the region.

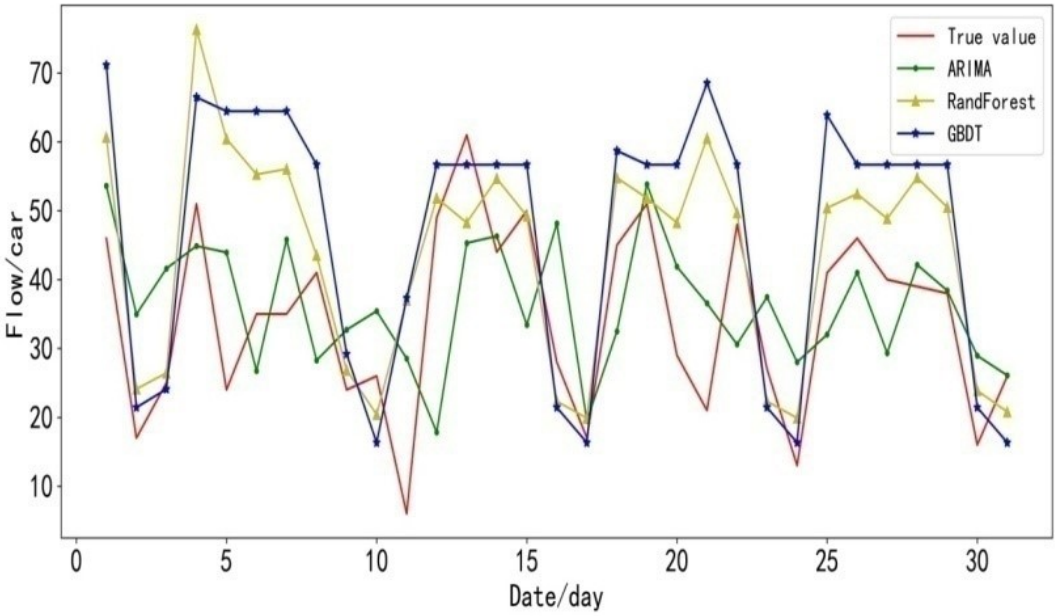

FIGURE 10. Other methods service points prediction results

5. **Conclusions.** In this paper, a LTSM-based bicycle-sharing traffic prediction method based on differential evolution algorithm is proposed. Differential evolution algorithm is used to optimize the parameters of LSTM model, and the optimized LSTM model is used to predict the shared traffic of service points. The simulation results show that the LTSM shared bicycle flow forecasting method optimized by the differential evolution algorithm has a good forecasting effect and is suitable for the service point shared bicycle flow forecasting.

Acknowledgment. This work was supported by the Fujian Young Teacher Education Research Foundation under Grant No. JA1417. The science and technology projects of the Ministry of Education under Grant No. 2017A13025. The work of Tsu-Yang Wu was supported in part by the Natural Science Foundation of Fujian Province under Grant No. 2018J01636.

REFERENCES

[1] S.Y.Chen, C.C.Lu,A Model of Green Acceptance and Intentions to Use BikeSharing: YouBike Users in Taiwan, j*ournal of Networks and Spatial Economics*, vol.16, no.4, pp. 1103-1124, 2016.

[2] S.A.Shaheen,S.Guzman,H.Zhang,Bikesharing in Europe,the Americas,and Asia,*journal of the Transportation Research Board*, vol.2143, no.1, pp. 159-167, 2010.

[3] J.Froehlich,J.Neumann,N.Oliver,Sensing and Predicting the Pulse of the City through Shared Bicycling, *Proceedings of IJCAI International Joint Conference on Artificial Intelligence*, pp.1420-1426,2009.

[4] A. Kaltenbrunner, R. Meza, J. Grivolla, J. Codina, R. Banchs, Urban cycles and mobility patterns: Exploring and predicting trends in a bicycle-based public transport system, *journal of Pervasive and Mobile Computing*, vol.6, no.4, pp. 455-466,2010.

[5] P. Borgnat, C. Robardet, P. Abry , et al., A Dynamical Network View of Lyon's Vélo'v Shared Bicycle System, *journal of Modeling and Simulation in Science, Engineering and Technology*, vol. 55, pp. 267-284,2013.

[6] P. Vogel, B. A. Neumann Saavedra, D. C. Mattfeld, A Hybrid Metaheuristic to Solve the Resource Allocation Problem in Bike Sharing Systems, *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics*), vol. 8457, pp. 16-29, 2014.

[7] P. Vogel, D. C. Mattfeld, Strategic and operational planning of bike-sharing systems by data mining: a case study, *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* , vol. 6971, pp. 127-141,2011.

[8] L. Chen, D. Zhang, L. Wang ,et al., Dynamic cluster-based over-demand prediction in bike sharing systems, *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing. ACM*, pp. 841-852,2016.

[9] Y. Li, Y. Zheng, H. Zhang, L. Chen, Traffic prediction in a bike-sharing system, *Proceedings of the 23rd SIGSPATIAL International Conference on Advances in Geographic Information Systems. ACM*, pp.33, 2015.

[10] Z. Yang, J. Hu,Y. Shu,et al.,Mobility Modeling and Prediction in Bike-Sharing Systems, *Proceedings of the 14th annual international conference on mobile systems*, applications, and services. ACM, pp.165-178, 2016.

[11] J. Froehlich, J. Neumann, N. Oliver. Sensing and Predicting the Pulse of the City through Shared Bicycling, *Proceedings of Twenty-First International Joint Conference on Artificial Intelligence*, pp. 1420-1426,2009.

[12] E. O'Mahony, D. B. Shmoys, Data Analysis and Optimization for (Citi) Bike Sharing, *Proceedings of the 29th AAAI Conference on Artificial Intelligence*, vol.1, pp. 687-694, 2015.

[13] P. Borgnat, C. Robardet, P. Abry, P. Flandrin, J. Rouquier, and N. Tremblay, A Dynamical Network View of Lyon's Vélo'v Shared Bicycle System, *journal of Modeling and Simulation in Science, Engineering and Technology*, vol. 55, pp. 267-284,2013.

[14] Y. Zheng, L. Capra, O. Wolfson, H. Yang, Urban Computing: concepts, methodologies, and applications, *journal of ACM Transactions on Intelligent Systems and Technology*, vol.5,no.3, pp. 1-55, 2014.

[15] S. Pal ,S. Ghosh , A.Nag ,Sentiment Analysis in the Light of LSTM Recurrent Neural Networks, journal of International Journal of Synthetic Emotions, vol.9, no.1,pp. 33-39, 2018.

[16] J. Wang, J. Zhang and X. Wang, Bilateral LSTM: A Two-Dimensional Long Short-Term Memory Model With Multiply Memory Units for Short-Term Cycle Time Forecasting in Re-entrant Manufacturing Systems, *journal of IEEE Transactions on Industrial Informatics*, vol.14, no.2, pp. 748-758, 2018.