Computing, Performance and Communication Systems (2021) Vol. 5: 52-55
Clausius Scientific Press, Canada

DOI: 10.23977/cpcs.2021.51009
ISSN 2371-8889

Mobile communication base station traffic forecast

Dongxiao Liu, Wei Li

North China University of Technology, Beijing, 100043

Keywords: LSTM model, ARIMA model, long-term forecasting

Abstract: This paper studies the prediction problem based on the historical time series data of the base station. Established **LSTM and ARIMA models** to train and predict traffic data respectively. This paper choose to put the model species into each time series, and put the forecast results into the data table according to the best. At the same time, we found that the traffic data of the cell showed three patterns of rising, steady, and falling over time, instead of rising all the time. Our LSTM model uses the 'elu' function as the activation function, the number of training rounds is 50, the number of neurons is 1000, and the observation and output sequence is 120, which is closer to the true value in short-term prediction. The ARIMA model retains the trend of the original model and has achieved better results in **long-term forecasting**.

1. Introduction

The rapid development of the mobile Internet has brought great convenience to people. At the same time, mobile traffic has exploded, and the traffic load of base stations has become more and more important. On the one hand, during peak traffic periods, a large number of base stations present a problem that the load exceeds the capacity. On the other hand, due to the tidal phenomenon of the base station, the number of users will be greatly reduced in certain periods. Since physical expansion involves issues such as procurement, cost, and overall layout, planning requires a very long time [1]. We need to build a model for analysis and predict the time required for physical expansion of the base station in advance, so that planning and design can be carried out earlier.

2. Model establishment and solve

2.1 Data set preprocessing

Data segmentation: We divide the complete base station traffic data set according to the cell number to reduce the size of a single training data, which is conducive to the training of the subsequent model.

Data cleaning: Find and correct errors in the data set, including checking data consistency, processing invalid and missing values, and deleting duplicate data.

Data integration: Integrate the hourly flow data in the data set into daily flow data for processing. And send it as a new data set to the model for training.

Data exchange: Different evaluation indicators often have different dimensions and dimensional units, which will affect the results of data analysis. In order to eliminate the dimensional influence between indicators, data standardization is required [2].

2.2 LSTM Model

The internal structure of the memory module is shown in Figure 1. There are three gates in the picture that function like valves. The opening and closing of the valve affects the transmission of neuron information, and determines how much base station traffic information participates in the calculation of the current neuron and how much base station traffic information participates in the calculation of the next neuron [3]. The activation function of the three gates generally uses the sigmoid function, so after the data is input to the activation function of the three gates, the output value range is [0,1]. Represents the input function of the memory module and represents the output function of the memory module. Generally, $tanh$ functions or sigmoid functions are used:

$$sigmoid(x) = \frac{1}{1+e^{-x}} \tag{1}$$

$$tanh(x) = \frac{e^{x}-e^{-x}}{e^{x}+e^{-x}} \tag{2}$$

In this model, we used the activation function as an activation function. The average value of the activation value input by the previous unit can be controlled to 0, and the negative part of the activation function can also be used. Its expression is:

$$elu(x) = \begin{cases} x & x > 0 \\ \alpha(e^{x} - 1) & x \leq 0 \end{cases} \tag{3}$$

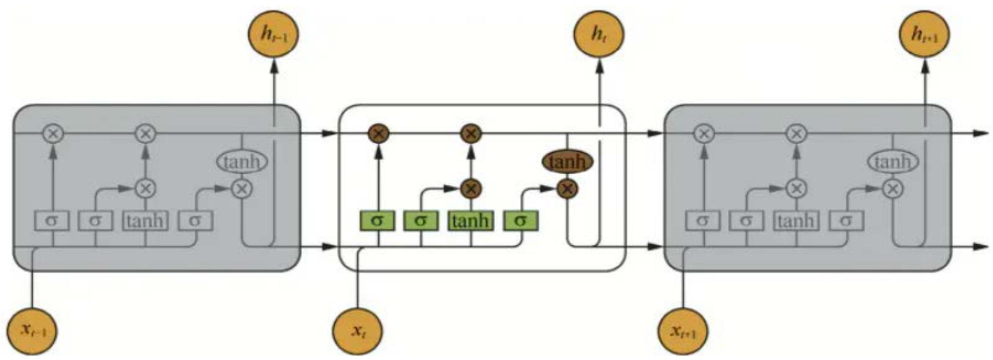

Figure 1: LSTM Neural network schematic

The state of the input gate at time t is:

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{4}$$

Among them is the activation function, the W represents the weight matrix, the b is the bias, the h_{t-1} is the state when the modulation gate is $t-1$ is output, and the state of the output modulation gate at the moment is:

$$h_t = o_t \cdot \phi(C_t) \tag{5}$$

Among them, $\phi(\cdot)$ is the hyperbolic tangent function, o_t is the output gate, C_t is the storage unit. The states at time t are:

$$o_t = \sigma(W_0 \cdot [h_{t-1}, x_t] + b_0) \tag{6}$$

$$C_t = f_t \cdot C_{t-1} + i_t \cdot g_t \tag{7}$$

In the above formula, f_t is the forget gate, g_t is the input modulation gate, and the states at time t are:

$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] + b_f\right) \tag{8}$$

$$g_t = \phi(W_c \cdot [h_{t-1}, x_t] + b_c) \tag{9}$$

2.3 ARIMA Model

For a time series $\{\chi_t\}$, a model of the following form can be fitted:

$$\Phi(B)\Delta^d x_t = \Theta(B)\delta_t \tag{10}$$

The $\{\delta_t\}$ is a white noise sequence:

$$\begin{cases} \Phi(B) = 1\text{-}\phi_1 B - \cdots - \phi_p B^p \\ \Theta(B) = 1\text{-}\theta_1 B - \cdots - \theta_q B^q \end{cases} \tag{11}$$

The roots of the above two operator polynomials are outside the unit circle. This model is called the summation autoregressive moving average model, denoted as ARIMA (p, d, q). Here, the ARIMA model has three parameters:

1) P: refers to the autoregressive order of the ARMA model fitted by the difference sequence, that is, the number of lags in the time series data itself used in the prediction model.

2) D: refers to the order of difference, indicating that the time series data needs to be differentiated in several orders to be stable.

3) Q: refers to the order of the moving average of the ARMA model fitted by the difference sequence, that is, the number of lags in the prediction error used in the prediction model.

For:

$$\Delta^d x_t = (1-B)^d x_t = \sum_{i=0}^{d}(-1)^i C_d^i x_{t-i} \tag{12}$$

The sequence after the difference is the weighted sum of several sequence values of the original sequence, and can fit the ARMA model, so this model is called the summation autoregressive moving average model. The meaning of summation is: ARIMA process can be expressed as the sum of ARMA process.

When p, q, and d are known, ARIMA is expressed in mathematical form as:

$$\hat{y}_t = \mu + \varphi_1 * y_{t-1} + \cdots + \varphi_p * y_{t-p} + \theta_1 * e_{t-1} + \cdots + \theta_q * e_{t-q} \tag{13}$$

The φ represents the coefficient of AR, the θ represents the coefficient of MA.

2.4 Model Results

In hourly traffic data, LSTM has a relatively good performance [4]. We randomly select a training result graph of a cell. Figure 2 is the training result of its test set. It can be seen that the training has achieved a good effect, In Figure 3, you can see that the training and test losses are declining. The specific parameters of the model are: the number of neurons is 1000, the forgetting factor is 0.5, and the epoch is 50.

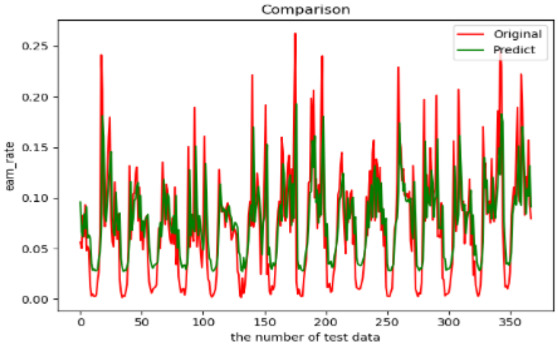

Figure 2: LSTM test set result graph

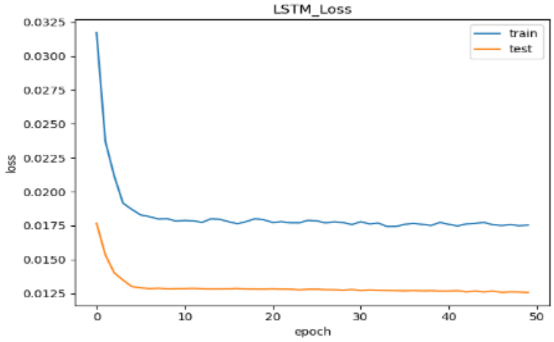

Figure 3: LSTM loss graph

The traditional ARIMA model can process time series data well, and discover the periodicity, trend and other information in it, and find that not all cells (base stations) are on the rise. The traffic sequence in Figure 4 is relatively stable. And the flow sequence of 5 shows a downward trend. In Figure 6, we can see that the model performs well in the mid- and long-term predictions.

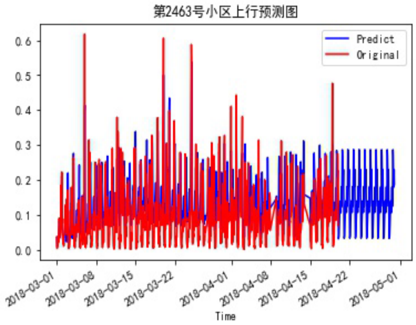

Figure 4: Steady trend

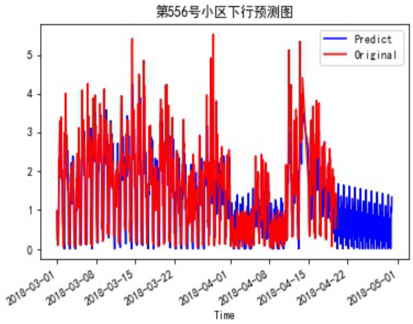

Figure 5: Trend down

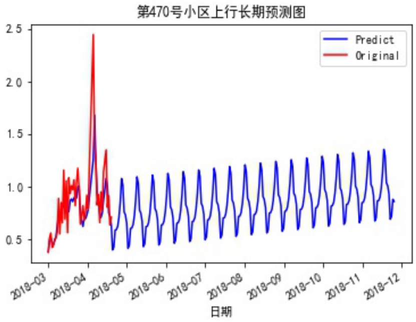

Figure 6: Long-term forecast

3. Evaluation of Model

We used two models to train and predict the time series of traffic, namely LSTM and ARIMA models.

A recursive neural network such as LSTM can be competent for short-term time series forecasting tasks, and its results are closer to the real value. The disadvantage is that some parameters may be insufficient in processing, so that it can further predict the future based on the predicted value. The value is biased, and the ARIMA model retains this trend, so it performs well in long-term forecasts. In addition to the base station traffic data, the model is also suitable for other time series prediction tasks

References

[1] Paparrizos J, Gravano L. k-shape: Effcient and accurate clustering of time series [C]// Proceedings of the 2015 ACM SIGMOD International Conference on

[2] Gullo F, Ponti G Tagarelli A, et al. A time series representation model for accurate and fast similarity detection [J]. Pattern Recognition. 2009, 42(11): 2998-3014.

[3] Dickey, D.A. and W. A. Fuller (1979), "Distribution of the Estimators for Autoregressive Time Series with a Unit Root, "Journal of the American Statistical Association, 74, p 427–431

[4] Luo, Yonghong, et al. "Multivariate time series imputation with generative adversarial networks."Advances in Neural Information Processing Systems. 2018.