

Mobile Base Station Traffic Prediction Based on Traffic Data Analysis

Mengjia Jiang^{1,a}, Jiayi Wang²

¹School of Maths and Physics, Chengdu University of Technology, Chengdu, 610059, China

²School of Labor Economics, Capital University of Economics and Business, Beijing, 100070, China

^ajiangmengjia2022@126.com

Abstract: The mobile base station is an important communication hub, which plays a very important role in the whole Internet. On the one hand, during peak traffic periods, a large number of base stations present problems of load exceeding capacity, making network speeds very slow even when signal conditions are good. On the other hand, the tidal phenomenon of base stations makes it possible for the number of users to drop significantly during certain hours, thus making the traffic load problem of base stations increasingly important. To address this problem, this paper proposes a GRU recurrent neural network based mobile communication base station traffic prediction model and improvements. Using the open-source sample dataset provided by the organising committee of the China Universities Big Data Challenge 2021, the study was carried out by conducting a smoothness test on the time series data, extracting three key indicators, namely the average number of subscribers in a cell, the PDCP traffic in a cell and the average number of activated subscribers, cutting the dataset, with 80% as the training set and 20% as the test set, and importing the GRU neural network prediction model implemented in Python, the results were unsatisfactory. Therefore, the model was improved to build a Bi-GRU-based base station traffic prediction model, and the prediction data was tested, and the error MAPE was less than 0.1, which was a good prediction. The improved base station traffic prediction solution is of great significance to the application of strategies related to the dormant energy saving of base stations based on traffic prediction.

Keywords: GRU recurrent neural network; Base station traffic prediction; Smoothness test; Bi-GRU; MAPE error test

1. Introduction

With the development of the traffic era, mobile communication base stations play an important role in the communication network and promote social and economic development, and their construction has become an important part of investment for mobile communication operators. According to the Chinese Ministry of Industry and Information Technology, the number of mobile base stations continues to grow, while at the same time increasing energy consumption. In the current global warming situation, environmental issues are slowing down the pace of China's sustainable socio-economic development, so accurate modelling and forecasting of base station traffic is needed to develop countermeasures.

The base station traffic has non-stationary chaotic characteristics. In addition to the traditional time series prediction method ARIMA model, scholars at home and abroad also proposed various models. Yong He and Yanting Li proposed a VAR-Lasso class approach ^[1] based on vector auto-regressive model (VAR) to analyse large-scale base station traffic data as a whole, transforming the multi-response variable prediction problem into a single-response model, and applying the Lasso to screen the important associated base stations of the target base stations. By comparing the ARIMA model, Jianwei Hu found that the LSTM neural network ^[2] can improve the accuracy of predicting the periodic and fluctuating changes of base station traffic, providing a prediction method to solve the base station load and tidal phenomenon in the future. Moreover, Yang Zhang and Luyi Zhang ^[3], fused ARIMA models based on LSTM models to deal with linear features as well as non-linear features in traffic sequences, which can extract traffic sequence features more comprehensively and make the fusion model more suitable for dealing with long time span prediction problems compared to single models. Jie Zhang and Guangwei Bai et al proposed a traffic prediction model based on the spatial and temporal characteristics of mobile network traffic, STFM ^[4], which uses TCN and 3DCNN to extract the temporal and spatial characteristics of mobile network traffic, respectively, and finally the fully connected layer to establish a

mapping relationship between the extracted characteristics and the actual traffic values to achieve short-term traffic prediction.

This paper applies a variant of the LSTM neural network—Gated Recurrent Unit (GRU) neural network for traffic prediction, which has a simpler structure with one less gate function compared to the LSTM, and is also a very popular network at present, making it a faster prediction method in data prediction, but the results are not good, so the model is improved. An attempt was made to build a Bi-directional Gated Recurrent Unit (Bi-GRU) model to improve the prediction accuracy. The analysis methods proposed in this paper can also be applied to other data sets to provide effective strategic reference and decision basis for the future. The improved solution for base station traffic prediction has important implications for the application of strategies related to dormant energy saving of base stations based on traffic prediction.

2. Data processing

The data in this paper comes from the open source sample dataset provided by the organising committee of the China Universities Big Data Challenge 2021. Firstly, the data indicators were extracted, data cleaning was performed using Excel, and indicators were obtained after classification and aggregation processing. The time series data of the three indicators obtained were smoothed, and outlier detection by each cell indicator required the deletion of duplicate data, which could be filtered due to the sufficient number of cell samples.

The data is checked as follows: Extending Grubbs' Test to k outlier detection requires gradually removing the values that deviate most from the mean (as the maximum or minimum) in the dataset, and simultaneously updating the corresponding t-distribution critical values to test whether the original hypothesis holds. The exact flow of the algorithm is shown below. Calculate the residuals that deviate furthest from the mean. Note that the data series used to calculate the mean should be the data from the sample with the largest residuals removed from the previous round.

$$R_j = \frac{\max_i |Y_i - \bar{Y}|}{s}, 1 \leq j \leq k \quad (1)$$

Calculate the critical value, test the original hypothesis, compare the test statistic with the critical value, if $R_i > \lambda_j$, then the original hypothesis H_0 is not valid and the sample point is an outlier, repeat the above steps k times until the end of the algorithm.

$$\lambda_j = \frac{(n-j) * t_{p, n-j-1}}{\sqrt{(n-j-1 + t_{p, n-j-1}^2)(n-j+1)}}, 1 \leq j \leq k \quad (2)$$

The paper then performs a data smoothness test on the time series to prevent the occurrence of pseudo-regressions. After entering the data into "Eviews", click on "View - Unit Root Test" in the top left corner, after which you can select whether the series has a unit root after first and second order differences, and if it does, the series is not smooth. The series of mobile base station traffic is non-smooth and the series after first order differencing of the data after outlier processing is smooth.

3. Model building

3.1. Selection of indicators

We consider three key indicators, as well as some public opinion indicators and the overall three key indicator values of the base stations belonging to the cell as inputs in the prediction model, which includes the average number of subscribers in the cell, the PDCP traffic in the cell and the average number of activated subscribers.

3.2. Introduction to the model

GRU is a variant of LSTM networks, which can solve the long dependency problem in RNN networks, and is also very popular at present because of its simpler structure and good results compared to LSTM networks [5, 6].

Three gate functions are introduced in the LSTM, the input gate, the forget gate and the output gate.

these threshold gates can be opened or closed and are used to control the memory state of the state of the previous network of the LSTM model and whether the output gate value reaches the threshold at that layer and is thus merged into that layer for calculation [7].

In contrast, there are only two gates in the GRU model, the update gate(r_t) and the reset gate(z_t), eliminating the output gate in the LSTM. The update gate is used to control how much information from the previous state is brought into the current state, and a larger value of the update gate indicates that more information from the previous state is brought in.

$$r_t = \sigma(W_r \cdot [h_{t-1}, x_t]) \quad (3)$$

The reset gate z_t controls how much information from the previous state is written to the current candidate set \tilde{h}_t ; the smaller the reset gate, the less information is written to the previous state.

$$z_t = \sigma(W_z \cdot [h_{t-1}, x_t]) \quad (4)$$

$$\tilde{h}_t = \tanh(W \cdot [r_t \otimes h_{t-1}, x_t]) \quad (5)$$

The update gate and the reset gate together control how the new hidden state h_t is calculated from the previous hidden state h_{t-1} .

$$h_t = z_t \otimes \tilde{h}_t + (1 - z_t) \otimes h_{t-1} \quad (6)$$

Where, r_t refers to the update gate. z_t refers to the reset gat. σ refers to the sigmoid function, and, W_z, W_r, W are the parameters of the individual gate neurons to be obtained during the training process [8].

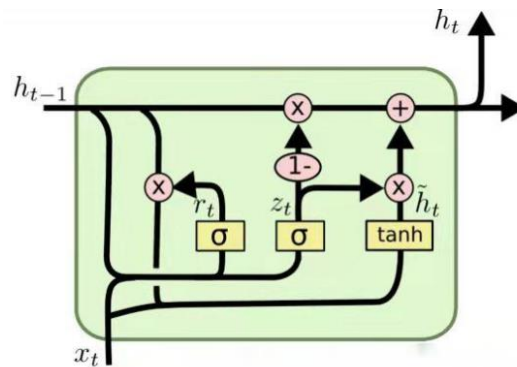


Figure 1: Schematic diagram of the GRU neural network

3.3. Bi-GRU

The hidden state of the GRU network can only read the information of the history, in order to obtain the future information on the current input at the same time, this paper adopts the Bi-GRU network with the structure shown in Figure 2. Bi-GRU consists of forward-propagating GRU units and backward-propagating GRU units, using the forward GRU to calculate the information before the current moment, and calculating the information after the current moment by the backward GRU. The two are linearly superimposed to obtain the output, which improves the accuracy of the network. The main calculation equations are as follows.

$$\vec{h}_t = f(x_t, \vec{h}_{t-1}) \quad (7)$$

$$\overleftarrow{h}_t = f(x_t, \overleftarrow{h}_{t-1}) \quad (8)$$

$$h_t = w_{11} \vec{h}_t + w_{12} \overleftarrow{h}_t + b_t \quad (9)$$

Where h_t is the implied layer state at time t ; w_{11} and w_{12} are the forward and backward propagated GRU implied layer output weights, respectively.

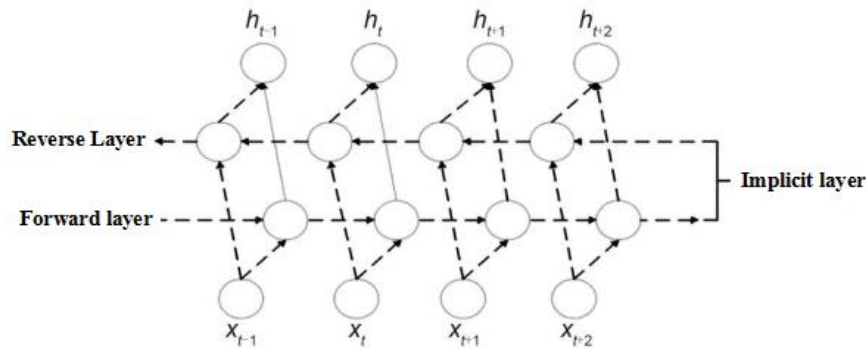


Figure 2: Bi-GRU schematic

Bi-LSTM and Bi-GRU are formed by replacing the RNN in the bidirectional RNN with LSTM or GRU structures. as an improvement, a prediction model based on Bi-GUU neural network is constructed.

4. Model results

The sample data are normalized, and then 80% of the sample data are randomly selected as training samples and 20% as test samples. The excitation functions of the hidden layer and output layer are set as relu and sigmoid functions respectively, the training function is traingdx, the number of neurons in the input layer is 3, the number of neurons in the hidden layer is initially set as 10, and the number of neurons in the output layer is 1. Set the network parameters. The number of network iterations epochs is 1000, and batch_size means only one sample is trained at a time. After setting the parameters, we start training the network. The network completes learning after reaching the desired error through 1000 iterations.

The Python toolkit was chosen to train the network, and the final confusion matrix was generated and visualised based on the predicted results. Figure 3 shows that the predicted classifications were anomalous. The following ROC curve was used to evaluate the model, as shown in the Figure 4.

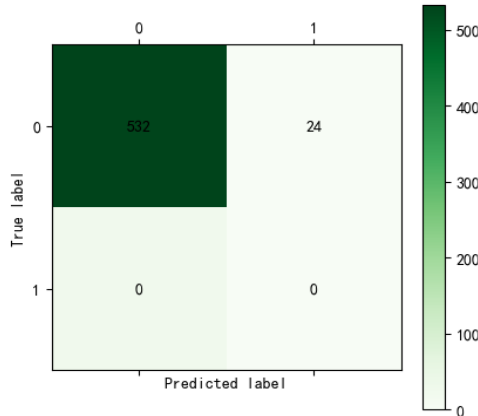


Figure 3: Confusion matrix

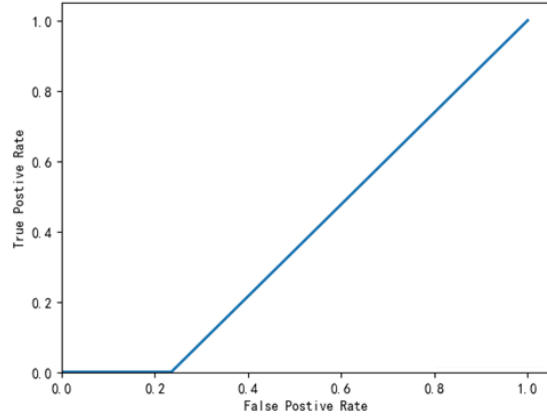


Figure 4: ROC curve

It can be seen that the area occupied by the ROC curve is small, the prediction is not good and the model is chosen to be improved or replaced.

5. Model evaluation

5.1. Advantages and disadvantages of model

Advantages: GRU can effectively suppress gradient disappearance or explosion when capturing association before and after long sequences, fast training and less computational complexity than LSTM.

Disadvantages: GRU still cannot completely solve the gradient disappearance problem, and as a variant of RNN, it has a major disadvantage of the RNN structure itself, i.e. it cannot be computed in parallel, which is a key bottleneck for the development of RNN in the future when the volume of data and the volume of models are gradually increasing.

5.2. Model improvements

(1) Stability check

We use STL time series decomposition to obtain the trend components, using the time series decomposition module of Python's statsmodels library to obtain the following trend plots and observe the smoothness of the series.

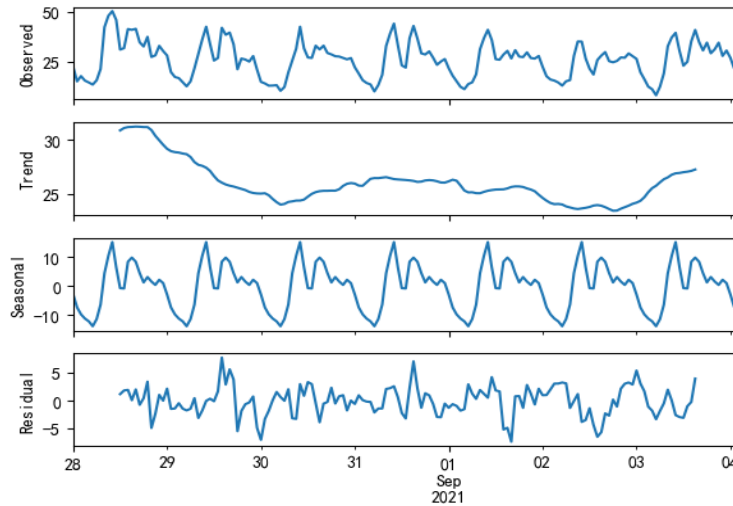


Figure 5: Time series trend plots

The graph shows that the trend component of the original series has a clear monotonic trend and can be judged as a non-stationary series.

For some non-stationary time series $\{y(t)\}$, the general form is

$$\Phi(B)(1 - B)^d y(t) = v(B)a(t) \tag{10}$$

d is the order of the difference operation, where

$$\Phi(B) = 1 - \Phi_1 B - \dots - \Phi_n B^n, \Theta(B) = 1 - \theta_1 B - \dots - \theta_m B^m \tag{11}$$

If $(1 - B)^d y(t)$ is denoted as $z(t)$, the incremental sequence can be found by the difference method:

$$Deta y(t) = y(t) - y(t - 1) (t = 1, 2, \dots, N) \tag{12}$$

After one difference, if this incremental sequence $\{Deta y(t)\}$ is smooth, then model denoted as:

$$\varphi(B)\nabla y(t) = v(B)a(t) \tag{13}$$

According to the relation between the difference operator and the back-shift operator ($\nabla = 1 - B$)

$$(1 - B)y_t = y_t - y_{t-1} \tag{14}$$

The relation $\Phi(B)(1 - B)y(t) = v(B)a(t)$ is obtained. That is, the reduction obtains the differential integrated moving average autoregressive model of $\{y(t)\}$, and the above differencing of the non-stationary time series $\{Deta y(t)\}$ is called first-order differencing.

The original series was first-order differenced and tested for smoothness and white noise, and the series after first-order differencing was judged to be smooth. The results obtained are shown below.

Table 1: Unit root test for series after first-order differencing

adf	cValue			P-value
	1%	5%	10%	
-3.567	-3.954	-2.967	-2.078	0.023

The results show that the time series plot after first order difference fluctuates relatively smoothly around the mean and the p-value of the unit root test is less than 0.05, judging that the series after first

order difference is a smooth series.

(2) Model testing

The data for the next three days for the three key indicators are predicted and the predictions are tested on request, here using the MAPE error test [8], calculated as shown below.

$$MAPE = \frac{100\%}{n} \sum_{i=1}^n \left| \frac{\hat{y}_i - y_i}{y_i} \right| \quad (15)$$

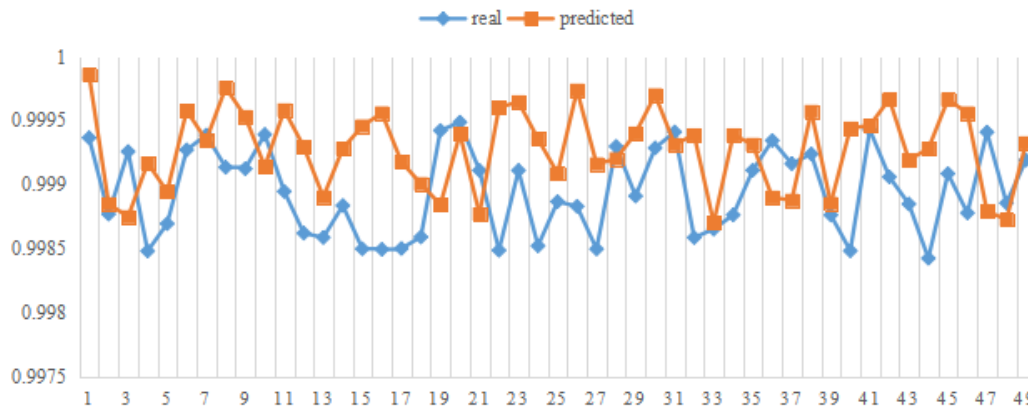


Figure 6: Predicted results

The MAPE error test is used here, and the error MAPE obtained from the calculation of the data is less than 0.1. Therefore, it is a good prediction.

6. Conclusion

The trend is to achieve flexible allocation of communication base station resources and to achieve appropriate service, so it is important to analyse and forecast communication base station traffic with accurate modelling. Traditional time series forecasting methods such as the Arima model, which focuses on individual base stations, can predict future traffic trends to a certain degree of accuracy.

And this paper applies its variant GRU on the basis of LSTM for mobile communication base station traffic prediction with simple structure and faster speed by training set, but the prediction result is not good, so the model is improved to establish the traffic prediction model based on Bi-GRU neural network, and the prediction result with higher accuracy is obtained, and using MAPE error test, the prediction effect is obtained better. This is of practical significance for the improvement and promotion of future flow prediction methods.

Mobile base station traffic forecasting can help operators to understand the realities of their subscribers and therefore strategically keep their investment costs to a minimum, meet their customers' network service needs and improve their network coverage, as well as improve their optimal network service quality [9].

Biographical notes

Mengjia Jiang is an undergraduate student of Chengdu University of Technology, China. Her research interest focuses on Applied Mathematics. Email: jiangmengjia2022@126.com.

Jiayi Wang is an undergraduate student of Capital University of Economics and Business, China. Her research interest focuses on Business Analysis. Email: ebaodou@126.com.

References

- [1] He Y, Li YT. Mobile communication base station traffic prediction based on vector autoregressive model [J]. *Industrial Engineering and Management*, 2017, 22(04): 79-84.
- [2] Hu JW. LSTM-based traffic prediction for mobile communication base stations [J]. *Information*

Technology and Informatization, 2021(05): 84-86.

[3] Zhang Yang, Zhang leucistic. A mobile communication base station traffic prediction method based on LSTM and ARIMA models [J]. *Information and Computer (Theoretical Edition)*, 2021, 33(07): 100-102.

[4] Zhang J, Bai GW, Sha XL, Zhao WT, Shen H. A mobile network traffic prediction model based on spatio-temporal characteristics [J]. *Computer Science*, 2019, 46(12): 108-113.

[5] Fugang Liu, Ziwei Zhang, Ruolin Zhou. Automatic Modulation Recognition Based on CNN and GRU [J]. *Tsinghua Science and Technology*, 2022, 27(02): 422- 431.

[6] Zhang Jie, Zhang Zipeng. Research on Fermented Grains Steam Running Condition Prediction Model of Make Chinese Liquor Robot Based on GRU [C]// *Proceedings of 2021 3rd International Conference on Advanced Information Science and System (AISS 2021)*., 2021: 250-253.

[7] Luo Junxia, Ding Bangxu. Research on the analysis and forecasting method of CPI economic indicators based on ARMA and LSTM [J]. *Journal of Huangshan College*, 2021, 23(05): 14-18.

[8] JIN Yu, ZHAO Bingwen, ZHENG Hanyu, LI Wan. Research on heating load prediction based on GRU neural network [J]. *Science and Technology Bulletin*, 2022, 38(01): 68-72.

[9] Zhang Lei, Tian Guo. Strategic analysis of mobile communication base station construction [J]. *Electronic Technology*, 2021, 50(11): 40-41.