Energy-saving Scheme of 5G Base Station Based on LSTM Neural Network

Yuxuan Wang*
North China Electric Power University, Baoding, China
*Corresponding author e-mail: yuxuan8889999@163.com

Abstract. As China's new infrastructure, 5G has received national and social attention. 5G promotes economic growth rapidly. But, the high energy consumption caused by the massive deployment of 5G base stations cannot be ignored. The total annual power consumption is expected to reach 243 billion degrees when the 5G base station is fully built. In the tidal scene, some 5G base station in an idle state still power fully, which causes great power waste. The historical volume of base station business data is used to train LSTM model, and predict the future base station business. When the business is lower than the threshold, the base station will be closed to avoid unnecessary power waste. And the LSTM model prediction results fits the original data ideally. By implementing the power saving strategy, the energy consumption of the base station is reduced by 18.97%. A single station can save 1174 degrees of electricity yearly. It can be seen that the energy saving effect is remarkable.

Keywords: 5G; Energy Consumption; LSTM; Prediction; Energy-saving Scheme.

1. Introduction
In the 14th Five-Year Plan of 2020, the Chinese government proposed to speed up new infrastructure such as the fifth generation of mobile communications, and 5G has developed rapidly due to policy support and development needs [1]. But the further development and deployment of 5G also encountered some difficulties and problems. The high energy consumption of 5G is a major problem that troubles operators and restricts the further development of 5G. The typical power consumption of 5G base station is about 3500 watts, which is 2-3 times of 4G base station. The annual power consumption of single macro station is as high as 2.00-3.38 million kWh. In the 5G era, 5G networks will consume up to 243 billion degrees of energy yearly [2]. It is necessary to find ways to reduce the energy consumption of 5G base stations.

The volume of 5G base stations business presents an obvious tidal phenomenon. Tidal phenomenon refers to that in shopping malls, residential areas, office buildings, schools and other places, the volume of business changes according to the tidal law. In a fixed period of time each day, the volume of business is low or even nobody uses it. However, in this case, the 5G base station is still fully powered, and the unnecessary waste of electricity generated by each macro base station can reach 4000 degrees per year, while the waste of electricity generated by the micro base station is more than 1000 degrees per year. This provides a good starting point for energy saving of 5G base station [3]. In this context, this paper proposed the forecasting on the volume of 5G base stations business based on LSTM, so as to close the
base station when the service load is lower than the threshold, which can effectively reduce the energy consumption of 5G base station.

![Flow chart of power saving scheme.](image)

**Figure 1.** Flow chart of power saving scheme.

## 2. Data Preparation

### 2.1. Data Exploration

Considering that in small-scale or underdeveloped cities, the demand for network is small, and the variation of the volume of base station business is not significant. So this paper selected the base station with a large scale and active network business as the research object. With the time series data of a certain city base stations in 2020 and 2021, this paper collected the volume of more than ten base stations' 24-hour business data and trained the model to verify the feasibility and generalization of the model. In addition to the volume of business, the current, voltage, power, and the average flow of upstream and downstream of the base station are also collected to improve the reliability of the data.

### 2.2. Data Processing

Due to the signal interference and the equipment failure, the data collected are often missing or noisy. It is necessary to preprocess the original data to obtain standard and continuous data for mining and analysis.

1. **Missing Value Estimation**
   - Periodic information is used to fill missing values for common missing data in time series;
2. **Outlier Processing**
   - In order to ensure the effectiveness of the algorithm and the rationality of the index interpretation, the index values beyond the normal range are directly taken as the upper and lower bounds;
3. **Data Standardization**
   - Standardize data to reduce noise and improve model accuracy;
4. **Data set Partitioning**
Divide the data set into 9:1 training set and test set.

3. Network Construction

3.1. LSTM Instruction

LSTM is a special Recurrent Neural Network (RNN). The original hidden layer of RNN has only one state \( h \), which is sensitive to short-term input. LSTM increases a cell state \( c \) in the hidden layer, and improves the hidden layer of RNN. On the basis of learning the rules between complex data, it takes into account the timing and nonlinear relationship, and does not lead to unstable prediction results due to the passage of time. Therefore, LSTM has a good advantage in solving the data with large time series prediction span and weak linearity [4]. The volume of business forecasting of 5G base station needs to process the data with large time span, so the LSTM model has great advantages.

The cell unit of LSTM has three gates: input gate, forgetting gate and output gate. The specific structure is shown in the figure 2.

![Figure 2: The specific structure of LSTM.](image)

3.2. Loss Function Definition

In order to illustrate the effectiveness of LSTM model in predicting the volume of business of 5G base station, this paper parameterized the error of target classification. Loss function is used, which reflects the error between the target output and the actual output. This paper selected the mean square error as the loss function. The calculation formula of mean square error is shown in Equation (1).

\[
E(W) = \frac{1}{2} \sum_{d \in D} (t_d - o_d)^2
\]

\( d \) is the training sample, \( D \) is the training sample set, \( t_d \) is the target output, and \( o_d \) is the actual output.

For multivariate function \( o = f(x) = f(x_0, x_1, \ldots) \), its gradient at \( X' = [x_0', x_1', \ldots, x_n']^T \) is \( \nabla f(x_0', x_1', \ldots, x_n') = [\frac{\partial f}{\partial x_0}, \frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n}]^T |_{x=X'} \). The direction of the gradient vector points to the fastest growing direction of the function. Therefore, the negative gradient vector \(-\nabla f\) points to the fastest descent direction of the function.

The idea of gradient descent is to let the loss function follow the square of the negative gradient search and iteratively update the parameters to minimize the loss function. The minimization of loss function can make the LSTM model better simulate and predict the volume of base station business. In this paper, Mini-Batch SGD (MBGD) is used as gradient descent algorithm. MBGD is special for using a small batch of fixed size samples to calculate \( \Delta w_j \) (\( w_j \) is the weight of each unit, \( \Delta w_j \) is used to adjust
$w_j$) each time, and then updating the weight. The calculation formula of $\Delta w_j$ is shown in Equation (2)–(3).

$$
\delta_j = \begin{cases} 
- (e_j - d_k) & \text{if } j \in \text{outputs} \\
   o_j (1 - o_j) \sum_{k \in D(j)} \delta_k w_{kj} & \text{otherwise}
\end{cases}
$$

(2)

$$
\Delta w_{ji} = -\eta \delta_j x_{ji}
$$

(3)

3.3. Optimizer Selection

MBGD is used to update each parameter with the same learning rate, but in the actual training process, the learning rate is expected to get slower and slower. At the initial stage of the learning rate, it is still far from the optimal solution of the loss function. With the increase in the number of updates, it is closer to the optimal solution, so the learning rate will also slow down. Therefore, this paper uses Adagrad as the optimizer to set different learning rates for different parameters. The calculation formula of Adagrad is shown in Equation (4)–(7).

$$
g_i \leftarrow + \frac{1}{m} \nabla_{\theta} \sum_i L(f(x_i, w), y_i)
$$

(4)

$$
r \leftarrow r + g_i^2
$$

(5)

$$
\Delta w = - \frac{\eta}{\varepsilon + \sqrt{r}} g_i
$$

(6)

$$
w \leftarrow w + \Delta w
$$

(7)

$g_i$ is the t-th gradient, $r$ is the cumulative gradient variable, whose initial value is 0, which will keep increasing. $\eta$ is the global learning rate, which needs to be set manually. $\varepsilon$ is a small constant, this paper sets it to about $10^{-7}$ for numerical stability. The calculation formula (4) calculates the gradient, the calculation formula (5) accumulates square gradient, the calculation formula (6) calculates updated values, the calculation formula (7) applies the updated values.

3.4. Model Evaluation Index Definition

MAE and RMSE are used to evaluate the prediction results of the model. The smaller the error value, the smaller the gap between the predicted value and the real value, that is, the better the prediction effect. The calculation formulas of MAE and RMSE are shown in Equation (9)–(10).

$$
MAE = \frac{1}{N} \sum_{i=1}^{N} |y_i - \hat{y}_i|
$$

(9)

$$
RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (y_i - \hat{y}_i)^2}
$$

(10)

4. Model Training

Then, this paper entered the data from the processed training set mentioned above into the model for iterative training with 500 iterations.

It can be seen from figure 3 that with the increase of iterations, the accuracy of the LSTM model is significantly improved. At the same time, in order to determine whether or not to reach the tidal migration period, the tidal identification bit needs to be introduced. And its output 1 indicates that the tidal migration period is about to reach. If the output is 0, it indicates that the tidal migration period has not yet reached.
Figure 3. Relationship between Training Frequency and Accuracy.

This paper tested with test set to ensure the generalization ability of the model. Generalization ability refers to the ability to measure the accuracy of the model in the actual business data. In the test process, the model has a strong ability to apply to new samples, and there is no under-fitting or over-fitting, indicating that the model learns the characteristics of 5G base station service load changing with time, and has a good generalization ability.

5. Model Evaluation

After the training, this paper predicted the volume of base station business and compared the predicted value with the data of the test set. The RMSE between the predicted results and the actual business volume is 7.34, and the MAE value is 3.42.

It can be seen that the fitting between the LSTM model prediction results and the original data is ideal, which can basically reflect the change trend of the volume of business. At the same time, the LSTM model does not decrease the prediction accuracy due to the nonlinearity of energy consumption data and the large prediction time span [5].

6. Energy-saving Analysis

The trained LSTM model is used to predict the volume of business of a base station from 5.17 to 5.24 in 2020, and the base station is simulated to close when the predicted volume of business is lower than the threshold [6]. Comparing to the average daily power consumption of the base station of 5.18, 5.20 and 5.23, the average daily power consumption of the whole station is 3.217 degrees after the implementation of the energy saving strategy, and the energy consumption of the base station is reduced by 18.97 %. The annual single station can save 1174 degrees of electricity. It can be seen that the energy saving effect is remarkable.

Table 1. Energy-saving Effect.

| Date  | Energy Saving State | Electric Quantity | Average Electricity | Daily Energy Saving | Energy Saving Effect |
|-------|---------------------|-------------------|----------------------|---------------------|---------------------|
| 5.17  | Turn on Energy Saving | 13.97             |                      |                     |                     |
| 5.19  | Turn on Energy Saving | 13.56             |                      |                     |                     |
| 5.21  | Turn on Energy Saving | 13.86             |                      |                     |                     |
| 5.24  | Turn on Energy Saving | 13.57             |                      |                     |                     |
| 5.18  | Close Energy Saving  | 17.02             | 16.957               | 3.217               | 18.97               |
| 5.20  | Close Energy Saving  | 17.03             |                      |                     |                     |
| 5.23  | Close Energy Saving  | 16.82             |                      |                     |                     |
References

[1] Liu Kun. China Manufacturing New Leap Forward to 'Create' Information Communication [N]. Guangming Daily, 2021-09-14 (010).

[2] Jiao 'an Qun, Chen Chao, Gao Hongliang, Song Lichen. Strategies for large base station power consumption in 5G era [J]. Electronic technology and software engineering, 2021 (04): 3-4.W. Strunk Jr., E.B. White, The Elements of Style, third ed., Macmillan, New York, 1979.

[3] Liu Hongjia, Wang Xinyan, Zhao Wei, Shi Hong. Research and application of 5G terminal power consumption under NSA networking mode [J]. Post and telecommunication design technology, 2020 (11): 53-57.

[4] Fu, Chen, Yang, Yang, Yang Guoyin, Yang Zhiqiang, Cao Zihan. The energy consumption prediction of air conditioning system using long-term and short-term memory neural network [J/OL]. Refrigeration and air conditioning :1 – 4 [2021 - 09 - 19].http://kns.cnki.net/kcms/detail/11.4519. TB. 20210906.1244.006.html.

[5] Shang Guanwei, Meng Yueyue, Yang Jiaming, Cai Bogen. The fault diagnosis of train control vehicle equipment based on LSTM-BP cascade network [J]. Journal of Beijing Jiaotong University, 2019, 43 (01): 54-62.

[6] Zheng Jiahuan, Xiang Yong. 5G base station energy saving method based on personalized scenarios [J]. Mobile communication, 2021, 45 (03): 91-96.