

Research on Nonintrusive Load Decomposition of Enterprises Based on Bidirectional LSTM

Xiangqian Yu¹, Xiaoming Li¹, Jianhua Chen¹, Zhiguo Wang¹ and Shiyuan Zhang¹

¹Marketing Service Center of State Grid Gansu Electric Power Company, Gansu,730000, China

Abstract. To detect the operating condition of equipment and understand the environmental management situation of enterprises in real-time, this paper studies the non-intrusive load decomposition of enterprises based on bidirectional LSTM. In this paper, we first obtain the load characteristic parameters of different equipment in different states, and then obtain the electrical power measured from the main power meter of decontamination equipment in TOP-5 through the softmax layer of bidirectional LSTM, and then change the softmax layer to decompose the load data from the main power meter.

1 Background

Load identification and monitoring can help electric utilities to make reasonable electricity consumption plans for their customers and contribute to the planning and scheduling of power systems, thereby enabling customers to reduce energy consumption. Load identification can be used to measure the current energy consumption of equipment in an enterprise and helps the enterprise to formulate energy-saving plans to relieve its economic pressure and energy shortage[1].

Load identification techniques can generally be grouped into intrusive and non-intrusive load identification. Intrusive load identification requires installing sensors and would generate harmonics, which affects the data collection of meters and the steady-state operation of power grids[2]. Non-intrusive load identification does not require installing sensors. Based on wavelet energy feature extraction, the load identification technique proposes to extract the wavelet energy value as a new feature value through wavelet transform, and then the fuzzy C-means clustering algorithm can be used for load identification. One problem is that the wavelet transform has a higher frequency for load collection. For the load identification based on supervised learning, four algorithms, which are MLP, KNN, LR, and SVM, are used for load identification and comparison and analysis. But the problem is that supervised learning requires the setting of label values, which is difficult to obtain. For the load identification based on genetic algorithm optimization, the third harmonic is added as the load characteristic and the genetic algorithm is used to solve the multi-feature objective function with high decomposition accuracy. But the problem is that the problem-solving speed of the genetic algorithm is slow, which is difficult to apply to online analysis and monitoring [3,4,5].

In this paper, we aim to learn the load characteristic parameters of decontamination equipment of enterprises. By combining the long-term memory feature of LSTM, the operating data of the decontamination equipment can be saved in the long term so that the operating condition of the equipment of enterprises can be accurately analyzed. In doing so, the load identification and monitoring of electric power can be achieved.

2 Overall Design

The enterprise equipment load identification model is divided into the following four main parts.

Data Processing: Preliminary pre-processing is carried out after data acquisition, and then a clustering algorithm is used to screen out the explicit features of decontamination equipment with the same operating rules for subsequent load identification and decomposition.

Model Design: In view of the superiority of LSTM on time-series data analysis, a bidirectional LSTM network architecture is used, where the activation functions are sigmoid and Tanh.

Model Training: AdaGrad optimizer is used for loss function optimization, and the model is trained on GPU.

Model Saving: After model training, the load identification model of the enterprise equipment can be saved. As model training is time-consuming, migration learning can be considered.

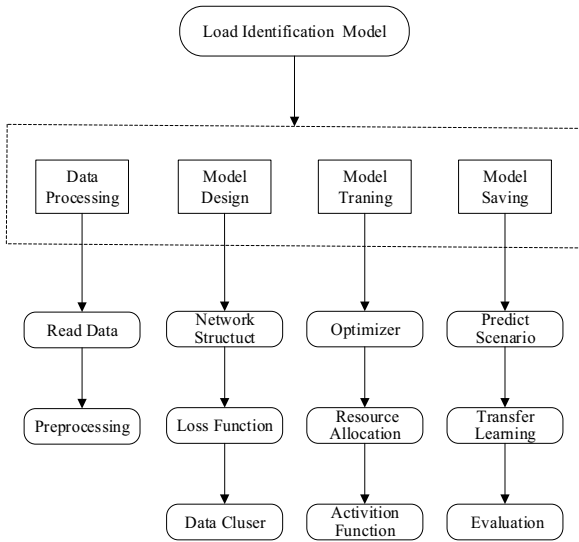


Figure 1. The Flow Diagram of Load Identification Model

3 Algorithm

3.1 Model Design

Long Short-Term Memory (LSTM) is a modified version of the recurrent neural network (RNN) that solves the problem of vanishing and exploding gradients. Compared to RNNs, LSTMs have memory units, input gates, output gates, and forget gates, which greatly enhance the ability of LSTMs to process long sequences of data. Input gates control how much new information from the current cell state should be added into the memory cell, while forget gates control how much information from the previous cell state should be forgotten, and output gates control how much information should be used to compute the output activation the current memory cell. The structure of LSTM is shown in the figure[6].

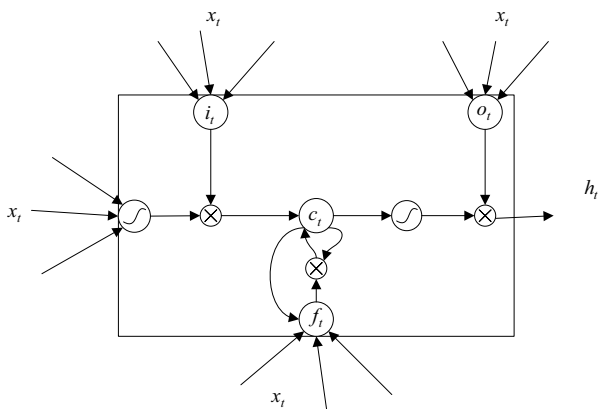


Figure 2. The cell unit of the LSTM at time T

The update formula of LSTM in step T is

$$i_t = \sigma(W_i x_t + U_i h_{t-1} + b_i) \quad (1)$$

$$f_t = \sigma(W_f x_t + U_f h_{t-1} + b_f) \quad (2)$$

$$o_t = \sigma(W_o x_t + U_o h_{t-1} + b_o) \quad (3)$$

$$\tilde{c}_t = \text{Tanh}(W_c x_t + U_c h_{t-1}) \quad (4)$$

The input gates i_t is obtained by linear transformation of the input x_t and the output of the hidden layer h_{t-1} in the previous step, and then by activation function σ . The two matrices (W_i, U_i) and vectors (b_i) are the parameters of the input gate. The calculation method of the forgetting gate and the output gate is similar to the input gate, and they also have their own parameters W, U and B. The state of the current memory unit is not completely determined by the activation function Tanh, and is controlled jointly by the input gate and the forgetting gate. The mathematical expression of the activation function Tanh is:

$$f(z) = \text{Tanh}(z) = \frac{e^z - e^{-z}}{e^z + e^{-z}} \quad (5)$$

For the normal-order RNN, h_t contains the input information up to moment t, that is, the information above, but for load identification, all the time-series data of the operation of equipment is required. Therefore, on the basis of LSTM, the input can be processed in reverse order and a bidirectional LSTM based on LSTM can be constructed using LSTM in a reversed direction. The odd-numbered LSTM layers are forward, while the even-numbered LSTM layers are backward. The higher LSTM layer uses the information from the lower LSTM layer and all previous layers as input. The final representation can be obtained using max pooling in the time dimension for the highest-level LSTM sequences, as shown in Figure 1.

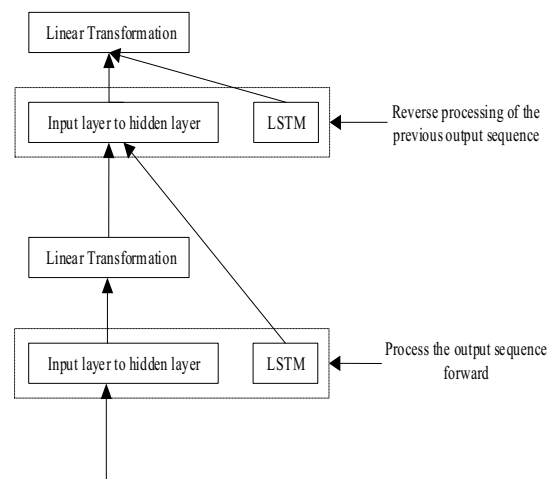


Figure3. Network Structure Diagram of Bi-LSTM

3.2 Model Training

When the enterprise equipment load model is trained, the bidirectional LSTM abstracts the high-level features and

maps them to vectors of the same size as the number of classes. Also, an AdaGrad optimizer is defined, which can adjust the learning rate of each dimension according to the gradient value of the independent variables in each dimension, thus avoiding the problem that a uniform dimension is difficult to adapt to all dimensions. To return results faster, training is performed on the GPU and the range of the loss value or the number of training steps is adjusted to reduce the training time at the cost of accuracy[7].

The model is evaluated using a ROC curve, which is a curve drawn on a two-dimensional plane with the False Positive Rate (FPR) on the x-axis and the True Positive Rate (TPR) on the y-axis, where the TPR measures the fraction of positive examples that are correctly labeled and the FPR measures the fraction of negative examples that are misclassified as positive. They are calculated using the following formula[8].

$$TPR = \frac{TP}{TP + FN} \quad (6)$$

$$FPR = \frac{FP}{FP + TN} \quad (7)$$

TP(True Positive):The positive class is predicted to be the positive class, and the true number is 0, and the prediction is also 0.

FN(False Negative):The positive class is predicted to be the negative class, and the true number is 0, and the prediction is also 1.

FP(False Positive):The negative class is predicted to be the positive class, and the true number is 1, and the prediction is also 0.

TN(True Negative):The negative class is predicted to be the negative class, and the true number is 1, and the prediction is also 1.

For a certain classifier, the FPR and TPR obtained from the test samples are mapped to a point on the ROC plane. Through adjusting the positive and negative sample classification thresholds, we can obtain a curve that passes through the origin and (1,1), which is the ROC curve of the classifier. The area under the ROC curve is the value of AUC (Area Under ROC Curve), which is usually between 0.5 and 1.0. The larger the AUC value, the better the classification performance.

4 Data Processing and Performance Analysis

The load data of a certain area for one year is selected together with the operating parameters and configuration parameters of the equipment of enterprises. Due to different seasons, weather, and production plan, the operating cycle of the equipment shows a certain pattern. The dataset is as follows:

Table 1. Indicators of load classification of the enterprise.

Column Name	Meaning	Type	Instance
DEVICE_ID	Device number	object	08089927957

DATA_DATE	The data of time	date	2019-01-03
DATE_TYPE	The data of type	object	Weekday
P1	Load data	float	198.3
P2	Load data	float	278.2
...

Firstly, the cluster algorithm is used to preliminarily filter the data and screen out the obvious characteristics of the decontrol equipment with the same operation rule for subsequent load identification and decomposition.

The steps of clustering algorithm are as follows:

Step 1: Randomly select a few days from the sample data as the initial load mean vector.

Step 2: Calculate the distance between the sample and each mean vector, divide the sample according to the distance, and recalculate the mean vector.

Step 3: Repeat the above steps until the center of mass does not change or the change is less than the set threshold.

Secondly, The data is initially filtered by a clustering algorithm, and then the explicit features of the decontamination equipment with the same operation pattern are selected for subsequent load identification and decomposition.

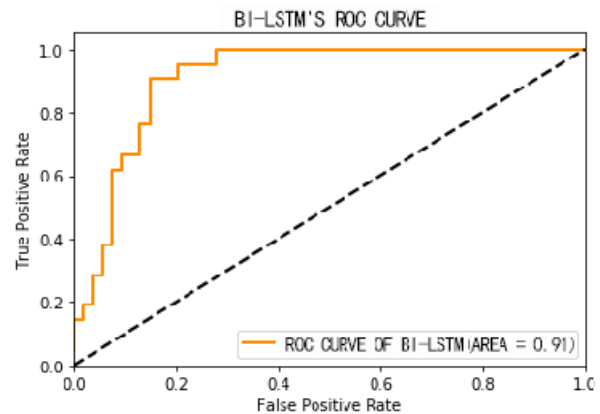


Figure4. AUC area diagram of enterprise equipment Load identification model

The clustering results are passed into a bidirectional LSTM architecture for model training and model evaluation through the AUC, which is shown in the figure.

The bidirectional LSTM performs well in the training of the enterprise equipment load identification model and still shows good classification performance when new data is added.

5 Conclusion

In this paper, we propose a load identification and monitoring system that can save the operation data in a long term and accurately analyze the operating status of the equipment of enterprises by learning the load characteristics of decontamination equipment of enterprises and combining the long-term memory characteristics of LSTM. This enables electric utilities to monitor the electricity usage of enterprises in real-time

so as to give energy use advice and enables customers to save electricity. Also, it effectively solves the problem of energy consumption monitoring, which is helpful for saving energy and reducing carbon emissions. In addition, the electricity consumption of enterprises can be managed appropriately. It not only protects the interests of the electricity utilities but also encourages the enterprises to use electricity safely, reducing the occurrence of electrical accidents, and thus improving the safety and reliability of grid operation.

References

1. Bo Jiang. Non-invasive Residential power Load decomposition Method based on Deep learning [D]. Hefei University of Technology,2017.
2. Xiaohuan Wang,Ruyi Li,Dongguo Zhou, Hong Zhou, Wenshan Hu. Noninvasive power load identification method based on decision fusion and its application [J]. Power system protection and control,2016,44(07):115-121.
3. Luyang Guo, Shouxiang Wang, Haiwen Chen, Haiyue Yang, Jianzhen Han. Noninvasive Load decomposition Algorithm based on bidirectional GRNN and Time series translation Model [J].Power supply, 2019, 4 (10) : 9 to 15 + 86.
4. Ke Wang, Haiwang Zhong, Nanpeng Yu, Qing Xia. Non-intrusive load decomposition of residential users based on SEQ2SEQ and Attention mechanism [J].Chinese journal of electrical engineering,2019,39(01): 75-83+322.
5. Siyun Chen, Feng Gao, Jing Liu, Qiaozhu Zhai, Xiaohong Guan. Load decomposition method and sensitivity analysis based on factor hidden markov model [J]. Power system automation,2016,40(21):128-136.
6. Zhankui Wang, Junying Wu, Rui Xin, Tao Bai, Jianbin Zhao, Minglei Wei, Jingquan Li, Lei Zhuang. The Power of Short-term Load Algorithm Based on LSTM[A]. Advanced Science and Industry Research Center. Proceedings of 2019 5th International Conference on Green Materials and Environmental Engineering(GMEE 2019)[C].Advanced Science and Industry Research Center: Science and Engineering Research Center,2019:5.
7. Lei Guo, Linyu Wang, Hao Chen. Electrical Load Forecasting Based on LSTM Neural Networks[A]. Advanced Science and Industry Research Center. Proceedings of 2019 International Conference on Big Data, Electronics and Communication Engineering (BDECE 2019)[C].Advanced Science and Industry Research Center: Science and Engineering Research Center,2019:5.
8. Biometrics, A closed max-t test for multiple comparisons of areas under the ROC curve.[J]. 2020.