Heavy overload forecasting of distribution transformers based on neural network

Haining Xie¹, Yingjie Tian¹, Wei Zhu², and Zhongyu Hu^{2,*}

¹State Grid Shanghai Electric Power Research institute, Shanghai

²Shanghai University of Electric Power, Shanghai

Keywords: Heavy overload distribution transformer, Efficient forecast, Associated factor, Attention-GRU neural grid.

Abstract. The overload management is significance component in distribution network operation and maintenance to improve electricity service. According to the periodic characteristics of the electric load, this paper designs a new method to identify and predict the heavy overload states and highlight the dates where the distribution transformer most likely heavy overload through the historical load rate and meteorological data. The Attention-GRU neural network is introduced to predict electric load rate of the highlight dates to improve the prediction efficiency. In comparison with the performances traditional LSTM in prediction of distribution transformers, results show that the new method has higher accuracy and efficiency in predicting highlight dates' load rates.

1 Introduction

During power transmission, the distribution transformer is possible to heavy overload, which affected by various external factors such as the environment, and characterizations of users and devices. When the distribution transformer heavy overload frequently, there are serious damage or even failure risks for power grid safe operation. Traditionally, the method for the distribution of heavy overload is usually focused on using the power monitoring system for real-time monitoring, and then setting the threshold and other early warning methods according to work experience for emergency treatment. This type of governance is relatively passive and cannot completely avoid grid and user losses. In order to improve governance efficiency, Big Data technologies need to be introduced. The risk assessment of the heavy overload of the distribution network lays a good foundation for timely distinguishing the weak links in the network operation.

In terms of theoretical research, the paper [2,3] analyzed the relationship between meteorological indicators, electricity consumption categories, industry categories and probability of occurrence of heavy-duty overload, and discussed the possible causes of heavy overload, and finally used random forest-based the theoretically improved decision tree model predicts the state of the distribution variable overload. In paper [4], for the high-speed

* Corresponding author: huzhongyu@mail.shiep.edu.cn

development area with rapid load growth, the medium- and long-term prediction methods based on logistic regression are proposed based on the household, meteorological and historical load. In paper [5], a heavy overload prediction method based on BP neural network and Grey model is proposed for the phenomenon of heavy overload during the Spring Festival. The heavy overload prediction model obtained by this method does not applicable to the general situation, and cannot adapt to the rapid analysis of large-scale distribution network. Paper [6] proposed a trend analysis and exponential weighting model for ultra-short-term load forecasting, which can improve the prediction accuracy of load peaks and valleys, and has better adaptability to load missing values and outliers. In the research on the heavy overload of distribution network, the above papers mainly focus on the analysis and prediction of the generalization of heavy overload of distribution transformers. There is no characteristic analysis of classification of different types of distribution load characteristics, so as to establish a forecast warning system.

The heavy load of the distribution transformer needs real-time monitoring, judgment and corresponding decision-making. There are hundreds and thousands of distribution transformers in a city. Real-time modeling and calculating for each line, the calculation speed and prediction accuracy are difficult to meet the business demand standard, and tend to waste computation power unnecessarily. Therefore, it is necessary to design an efficient algorithm program. The process designed in this paper guarantees the dynamic tracking calculation, and predicts the heavy overload state of the distribution transformer quickly, efficiently and accurately.

2 The algorithm flow

At present, the determination of heavy overload events in actual business usually depends on two factors: load rate and duration.

Generally speaking, heavy load event means the load rate is between 70% and 100%, and last more than two hours. Overload means the load rate is over 100% and last more than two hours. The above principles will also be adjusted in conjunction with the operational and the personnel and on-site experience. In the actual distribution network operation, the heavy load status will be monitored in time, so it will not last for too long, so this article focuses on whether the threshold is exceeded regardless of the duration.

According to the 2018 Shanghai grid load measurement analysis report, the heavy-duty line of 10kV on the load measurement day is 1.71%, and the overall heavy load ratio is small. Since the distribution load fluctuations have obvious annual, monthly, weekly, and daytime periodicity, shown as figure 1.

Fig. 1. Typical load change of distribution transformer.

The threshold of the overload is 70%, and the warning line is 60%. Considering the prediction error, the filtering threshold is set to 45%. The indicators are as follows:

The first-level indicator --the maximum load in the past 54 weeks does not exceed 45%;

The second-level indicator --the maximum load rate on the same date of the previous year does not exceed 45% and the maximum load rate in the last 7 days does not exceed 40%.

Fig. 2. Algorithm flowchart.

3 GRU neural network based on attention mechanism

3.1 Attention mechanism

The attention mechanism is a solution to the problem that mimics human attention. Simply put, it quickly filters out high-value information from a large amount of information. It is mainly used to solve the problem that the LSTM/RNN model input sequence is long and it is difficult to obtain the final reasonable vector representation. The method is to retain the intermediate result of LSTM, learn it with the new model, and associate it with the output. In order to achieve the purpose of information., Attention mechanism is used in the GRU model here^[7]

3.2 GRU

GRU is a variant of the LSTM network. Compare with the LSTM network, it reduces the complexity of the network by reducing the parameters of the entire neural network model, to improve the ability to prevent over-fitting, and to converge faster, shown as figure 3.

Fig. 3. A GRU unit in the 1st layer for time step t.

The GRU's compute node consists of an update gate and a reset gate. The update gate determines whether the current state is to be combined with the previous information, and how much memory is required to reset the gate definition. The GRU calculation formula is as follows:

$$\gamma_t = \sigma(W_\gamma \cdot [h_{t-1}, x_t]) \quad (1)$$

$$z_t = \sigma(W_z \cdot [h_{t-1}, x_t]) \quad (2)$$

$$\tilde{h}_t = \tanh(W_{\tilde{h}} \cdot [\gamma_t * h_{t-1}, x_t]) \quad (3)$$

$$h_t = (1 - z_t) * h_{t-1} + z_t * \tilde{h}_t \quad (4)$$

where 1 is for all information to pass the cell state; 0 is for no information is allowed to pass the cell state; σ is the sigmoid function [8].

3.3 Attention-GRU

In this paper, the attention mechanism is used as the interface of two GRU networks. Firstly, the input sequence is processed by a GRU network to learn high-level feature. Next, rationally assign attention weights base on attention mechanism and get a new sequence. Finally, running another GRU network to predict load rate.

Fig. 4. Structure of GRU network based on Attention mechanism.

In the figure, a^t represents the characteristics of the input sequence x obtained by learning; w^t is the attention weight of each feature [9].

4 Forecast result evaluation

In order to evaluate the predicted performance, the effects of model selection, data fitting, and data prediction were evaluated using RMSE, which is used to measure the deviation between the observed value and the predicted value. The formula is as follows:

$$\sqrt{\frac{1}{m} \sum_{i=1}^m (y_i - \hat{y}_i)^2} \quad (5)$$

Because the heavy overload event is a small probability event, most of the time the distribution is still in a light load state. In order to more validly express the ability of the overload identification early warning system in early warning of heavy overload events, the $C_{overload}$ is set to the system's early warning and alarm accuracy under the condition of heavy overload. The formula is as follows:

$$C_{\text{overload}} = \frac{T_{\text{pre}}}{T_{\text{fact}}} \quad (6)$$

where T_{pre} is the number of time points in the early warning and alarm state among results of prediction. T_{fact} is the actual number of time points in the early warning and alarm state.

In this paper, it is considered acceptable to predict the security status close to warning line as an early warning state. So, when this happens, the point time will be counted as a successful prediction.

5 Correlation factor analysis

The data sample is all the distribution transformer load data and meteorological data (temperature, air pressure, humidity, wind speed) of a certain area of Shanghai. The time span is from January 1, 2017 to June 30, 2018, and the frequency is one point every 15 minutes which means 96 points a day. The sample data is relatively complete and representative.

A typical residential type distribution transformer and a typical commercial type distribution transformer were selected for comparative analysis. It was found that the change law of residential type distribution and commercial type distribution load rate has the following characteristics: commercial distribution has a weekend effect, and the weekend load is significantly less than the working day, while there is no weekend effect in residential distribution. The daily load change of residential distribution has a complementary effect with commercial distribution, which is consistent with regularity of work and rest.

According to the analysis of the correlation between the meteorological data and the load rate in 2017, it is found that the load rate has the strongest correlation with temperature, and the correlation is 0.6. The correlation between load rate and air pressure is second, which is -0.58. The correlation between load factor and humidity is weak, which is -0.13.

After removing the weekend impact, it was found that the correlations between the maximum load rate and external factors were enhanced. If weekend effects are excluded at data preprocessing stage, the enhanced correlation between load factor and meteorological data will improve the prediction accuracy.

6 Establish model

The software framework is a TensorFlow framework based on the Keras deep learning tool. The Keras design uses a minimalist principle and is a highly modular library of neural network architectures. Keras is easy to use and supports free combination and layer stacking of model layers to reduce duplication of code implementation^[10].

6.1 Data preprocessing

According to analysis above, removing the commercial disturbance transformers' weekend impact to improve the correlation between the historical load rate and meteorological data.

The historical load rate at 96 time points of the previous day and the weather forecast data for the next day were used to predict the load at 96 time points the day after. The input feature quantity is divided into 96 rows and 5 columns of matrices by date, and the time step is 96.

Dividing the data in the way that set the data of the first 54 weeks as the training group, and the next 100 days is set as the validation group, and the random extraction into the second-level attention state as the prediction object.

Experiments have found that increasing the depth of the model by increasing the number of GRU network layers helps to improve the predictive power and speed of the model with a smaller number of neurons in each layer. In this paper, the three-layer GRU network is used within the acceptable range of model training time, and the number of neurons is set to 32,64,128. Finally, the vector of the specified format is output through the Dense layer.

6.2 Analysis of test results

From the data of residential distribution transformer and commercial distribution transformer, randomly extract a certain period of time into the second-level attention state, as the prediction object, and compare with the actual value, as shown in Figure 5.

Fig. 5. A prediction of a second-level attention period base on the Attention-GRU.

The results show that, for a random period of two interest rates load prediction, the prediction accuracy of the Attention-GRU is higher than LSTM, as shown in Table 1.

Table 1. Comparison of the Attention-GRU and LSTM prediction results.

Prediction method	RMSE	$C_{overload}$
Attention-GRU	0.04	0.86
LSTM	0.06	0.54

7 Summary

With the preliminary statistical analysis of data, the cyclical characteristics of the load ratio of residential distribution and commercial distribution are different, and the weekend effect of commercial distribution is obvious. In the data preprocessing stage, the weekend impact in commercial distribution data can be eliminated, which can improve the prediction accuracy of the model. The attention GRU neural network performed better than traditional LSTM in predicting heavy overload state of distribute transformer, whose $C_{overload}$ increase by 32%.

The main innovation of this paper is that, according to the characteristics of small probability events of heavy overload. With the two-level filtering principle, the time points when no overload occurs can be excluded, and people can only focus on the time when accurate prediction is needed, and then forecast in advance and give the corresponding warning level. Throughout the process most of the points are lightly calculated, and just a few proportions of time periods enabled to be calculated by accurate model, so that the state of dynamic monitoring can be maintained, and the calculation power can be saved to achieve

accurate prediction results. A large amount of computing power depends on continuous energy, and the algorithmic process is also in line with the sustainable green concept.

The Project Supported by Shanghai Science and Technology Program (18DZ1203200).

References

1. WANG Jiye, Big Data of Smart Grid, third ed., China Electric Power Press, Beijing, 2017.
2. HE Jianzhang, WANG Haibo, JI Zhixiang, Analysis of Factors Affecting Distribution Transformer Overload in Smart Grid, Power System Technology.41(1)(2017):279-284.
3. HE Jianzhang, WANG Haibo, JI Zhixiang, Heavy Overload Forecasting of Distribution Transformers Based on Random Forest Theory, Power System Technology.41(8)(2017):2594-2597.
4. LI M, ZHOU Q. Distribution transformer mid-term heavy and over load pre-warning based on logistic regression,2015 IEEE Eindhoven Powertech, Netherlands, 2015:1-5.
5. SHI Chang-kai, YAN Wenqi, ZHANG Xiaohui, Heavy Overload Forecasting of Distribution Transformer During the Spring Festival Based on BP Network and Grey model, Journal of Electric Power Science and Technology.31(3)(2016):140-145.
6. Ngo V C, Wu W, Zhang B, Ultra-short-term load forecasting using robust exponentially weighted method in distribution networks,Power & Energy Society General Meeting, 2015.
7. PENG Wen, WANG Jinrui, YIN Shanqing, Short-term Load Forecasting Model Based on Attention-LSTM in Electricity Market, Power System Technology.43(5)(2019):1745-1751.
8. Wang Weia, Sun Yuxiab, Qi Qingjie, Text sentiment classification model based on BiGRU-Attention neural network, Application Research of Computers.32(12)(2019).
9. XUE Yang, WANG Lin, Wang Shu, An ultra-short-term wind power prediction model combined with CNN and GRU networks, Renewable Energy Resources.37(3)(2019):456-462.
10. Information on <https://keras.io/>