Analysis of Smart Grid Stability and Security Management Based on Data Mining

Chao Zhang¹ *, Zhihang Yuan², Pei Yan³

¹School of Mechanical and Engineering, Wuhan Donghu University, Wuhan, Hubei, 430200
²Anhui Normal University, Wuhu, Anhui, 241000
³School of Computer Science, South China Normal University, Guangzhou, Guangdong, 510630

*Corresponding author: zhangchao@wdu.edu.cn

Abstract. In the process of the operation and construction of the smart grid, the development of science and technology and the improvement of social productive forces have put forward higher requirements for the development of the power system. In this process, the scale and form of the power system have also changed to a certain extent, and power security has also become the most concerned issue. In the information society, through the application of the Internet of things and big data technology in power security management and control. It can simplify the operation process of the staff and reduce the occurrence of accidents. Distributed smart grid is a new technology proposed for power networks with elastic nodes, which can realize dynamic electricity price demand response without large-scale transformation of infrastructure. In order to analyze the system stability of DSG, six representative machine learning classification models are applied to analyze the stability data of 10000 samples of 4-node system. Combined with the requirements of power system security, stability and economy, the effect of each classification model on the prediction of DSG system stability is tested.

Keywords: Big data technology, Electric power safety, Smart grid, Grid stability, Machine learning.

1. Introduction
According to the data in the "Analysis and Forecast report of the National Power supply and demand situation in the first half of 2020" released by the China Federation of Electric Power Enterprises, the installed capacity of non-fossil energy power generation in China is expected to reach about 930 GW by the end of 2020, accounting for 43.6% of the total installed capacity, which is about 1.6% higher than that at the end of 2019[1]. The emergence of a large number of new energy nodes will bring many new problems to the power system structure. Distributed smart grid (decentralized smart grid, DSG) operates in a balanced manner of supply and demand based on frequency [2]. By sampling the grid frequency reasonably, the smart grid with new energy nodes can effectively enhance the economy of the overall operation. With the continuous maturity of new energy technologies represented by photovoltaic and solar energy in the power system, how to integrate these elastic nodes
into the power grid and power market, and real-time and effectively maintain the stability of the system has gradually become a hot issue. In addition, the information and control system of the power system can be integrated with various detection equipment and safety devices to realize the purpose of monitoring automation and dispatching automation, which is conducive to the stable development of all aspects. Under the application of Internet of Things and big data technology, this aspect has certain research value.

At present, the idea of machine learning has been adopted in power system to study the stability of power system. Literature [3] adopts the neural network model to analyze the stability of the power system, and studies the system behavior of the traditional large-scale power system. However, because the traditional differential equation is adopted to build the system, the general evaluation of different inputs is not carried out. In literature [5], stability analysis was carried out for small and medium-sized systems with 39 nodes and 470 nodes. After the input simplification was completed by polynomial series, ridge regression algorithm was adopted to conduct stability judgment modeling. Literature uses support vector machine and neural network for model analysis, focusing on the analysis of the rationality of characteristic engineering of network input, and finds out the influence of dimension specification on the stability judgment of power system, but does not explain the significance of the specification. In literature the decision tree algorithm is adopted to analyze the stability of The Iranian power grid, and is compared with artificial neural network and support vector machine. It is proved that the decision tree algorithm has high stability, but the conclusion can only be reflected by the accuracy, so the model evaluation cannot be carried out comprehensively.

2. Smart grid control technology

DSG is a new kind of technology for dynamic electricity price of smart grid. The ideal electricity price strategy of the traditional smart grid is to conduct electricity price auction with a cycle of 15 minutes, while DSG can realize dynamic demand response without large-scale transformation of infrastructure through binding electricity price and frequency.

2.1. DSG and system model

Smart grid has certain requirements for the stable operation of the system, and the influence of elastic nodes on the stability of the system also exists in DSG system. When using traditional differential equations to describe DSG system, there are many limitations in the study of stability problems due to too many modeling constraints, such as considering the fixed input problem caused by a single transformation and the problem of equal opportunity caused by too idealization. Therefore, SYS equations can be used to describe DSG system and SYS model.

\[
\langle \tau, P, \gamma, \alpha, K, T \rangle
\]

The results were determined based on reaction time, mechanical power \( P \), economic elasticity parameter, motor loss parameter, transmission parameter \( K \), and system running time \( T \).

Sys model combines the physical model of the rotating motor in the system with the economic model of energy cost, when the phase angular velocity is far less than the periodic phase Angle of power frequency and the phase angular acceleration is far less than the frictional rotational potential

\[
\frac{d^2 \theta_j}{dt^2} = P_j - \alpha_j \frac{d \theta_j}{dt} + \sum_k K_{j,k} \sin(\theta_k - \theta_j)
\]

DSG bundles the electricity charge and frequency, and allocates the elastic proportional factor CP, so that nodes can adjust their load state through the change of electricity price. There are four response states, and the following are obtained:

\[
P_j(pi) = P_j + c_j(p_j + pw)
\]
2.2. Criterion of system stability analysis
The local linear stability criterion was used to analyze the stability of the system. The value \( SR \) of the root real part of the system characteristic equation (1) represented whether the system was unstable: when \( SR > 0 \), the system was unstable; When \( SR < 0 \), the system remains stable. When \( SR = 0 \), it is a critical state.

To sum up, a series of operation data and network stability can be obtained by changing different input states of the system equation. UCI data set parameters are adopted here (see Table 1), and the system stability data simulation model adopts the four-node star grid.

2.3. Machine learning method for power system stability analysis
The stability of the power system is different from that of the general data set. According to the topological structure and operation state of the system, the balance of samples in the data set is quite different. In the DSG system in this paper, the proportion of positive and negative samples is about 3:7. this basic covers the current mainstream, strong generalization ability of machine learning algorithm, achieved through super parameter optimization aiming at the accuracy of the model, the optimal state, and through the multi-index comprehensive evaluation of algorithm performance.

Accuracy is the most basic model evaluation method, and accuracy (the definition quantity symbol is \( bACC \)) is the most basic index to describe the model accuracy, and the calculation method is

\[
b_{ACC} = \frac{T_P + T_N}{T_P + T_N + F_P + F_N}
\]  

In the dichotomy problem, the Cohen Kappa index can be used to complete the accuracy assessment of the algorithm. Compared with the accuracy rate, the Cohen Kappa coefficient can complete the accuracy assessment of the asymmetric sample data set.

Subjects operations positive rate and false positive rate of curve, the area under the curve can be quantified as offline area. The AUC index is the relative area with a range of, and the classification model with a high AUC has better accuracy.

Sensitivity and specificity are concepts used to measure decision bias, which can be measured by recall rate (defined quantity symbol \( bR \)) and precision rate (defined quantity symbol \( bP \)) respectively.

\[
b_R = \frac{T_P}{T_P + F_N}, \quad b_P = \frac{T_N}{T_N + F_P}
\]  

Sensitivity and specificity represent the rigor of the model decision result. High sensitivity can be quantified as high recall rate. At this time, the model tends to make the test data decision as positive samples. High specificity can be quantified as high precision, so the model is more cautious about the decision of positive samples. The data set was divided into training set and test set according to the ratio of 7:3, the model hyperparameter training was carried out by cross validation of training set, and the configuration was 50% crossover. The optimization evaluation index is accuracy, and the specific method is 2-step search. First, the random search method is used to find the rough high-performance parameters in a large range, and then the grid search method is used to find the precise optimal parameters.
With the continuous development of smart grid, electric power monitoring has become the main direction of future development of electric power system. First of all, the big data technology can be combined with cloud computing together, realize on the improvement of the power equipment monitoring method, the data to improve the existing distributed storage strategy, USES the many kinds of signal parallel analysis and feature extraction method, so as to look for in the large data sets of multi-source data form, and carry on the analysis of the correlations. The Internet of Things can realize the effective transmission of information and form an integrated development system of monitoring and control.

The accuracy of the algorithm applied to the test set is taken as the measurement standard. The performance of the six parameterized classification models is in terms of accuracy and AUC values, so that the difference in classification accuracy of each model can be more clearly compared. Among the 6 models, the RBFSVM model has the best performance, with an accuracy of about 97%. Its ROC image is also the most consistent with the reasonable high-performance mode, and its AUC is close to 1. The accuracy of the other five models is above 80%, among which LD model has the lowest accuracy (80.47%). However, LD model is still qualified for the actual situation. LD, GNB stability of these two models the judgment only expressed as "moderate" right, kNN, DT and ADA, the three kinds of model can show the stability judgment on Cohen Kappa coefficient, the real work of the SVM is a few kinds of machine learning model is the most outstanding performance, can achieve "almost perfect". The evaluation index of comprehensive accuracy and consistency shows that all the 6 models show qualified classification accuracy. For the consistency of classification decision, they all show qualified performance except LD model. RBFSVM model has the best performance. Among the other 5 models, DT model has high consistency and accuracy, but the performance gap between the other models is not big.

Due to the excellent accuracy, the recall and precision of RBFSVM models perform well and are approximately the same. The classification models with higher precision and more likely to be positive samples include DT model and ADA model. However, GNB model and kNN model have higher recall rate and are more inclined to reduce false positives of instability.

3. Application of stability control mode in power system

Power system stability control mode is a kind of control technology in the process of power system operation, which can guarantee the transient and dynamic stability. Under the application of big data technology and functions such as the Internet of Things, the fault can be limited to a certain area, and the analysis method of data characteristic value can be realized to increase the efficiency of the
regulation system, effectively avoid the expansion of accidents, and control the load power in the operation of the power system.

Big data technology and Internet of Things can promote the development process of dispatching automation in power network dispatching security management. First of all, in the process of the operation of the power system, they can realize the data collection and monitoring, and display the running status of each part of the staff passed through information data analysis of the status of the evaluation in time, can also be achieved by the application of the big data technology solutions for intelligent search process, reduce maintenance time, improve work efficiency. On the other hand, the application of the Internet of Things can make the real-time communication and sharing of various information, and strengthen the application of wireless network in this aspect.

The stability prediction of DSG system still needs to pay more attention to the sensitive classification model while meeting the accuracy, that is, more consideration should be given to $F$, especially the evaluation index of classification model when $\lambda = 2.0$ or even $> 2.0$. At this time, DT, ADA, RBFSVM and other more sensitive models may be more feasible for further research. RBFSVM has a strong performance, but it may not meet the requirements of the system in practice. The essential reason is the "combination explosion" caused by the failure of dimensionality specification for high-dimensional data. According to the above comparison, GNB model or DT model may be more applicable in practical application. In data analysis, RBFSVM model and ADA model may be more applicable.

4. Conclusions

In this paper, by analyzing the stability data of DSG system and combining with the demand of power system safety, stability and economy, six kinds of prediction models are obtained through training optimization, and the multi-angle evaluation indexes of each model for system stability prediction are proposed. RBFSVM model is suitable for data analysis and its accuracy can reach 97.10%. DT model is suitable for real-time forecasting of power system, and its accuracy can reach 84.90%. Both adopt high sensitivity strategy to ensure the security of DSG system. In terms of computational performance, THE DT model is more excellent. Its modeling and predicted computing time are only 0.98% and 1.59% of that of the RBFSVM model, respectively, with higher real-time performance.

To sum up, the power security is in the process of running our country industry foundation, the relevant personnel should find electricity problems existing in the safety control, apply the function such as big data technology and the Internet of things to it, through methods such as information integration and sensory perception in stable electric power monitoring and control mode, power grid scheduling security, and the application of plant safety control and so on, so as to realize the stable development of the power system.

References

[1] Song Yaqi, Zhou Guoliang, Zhu Yongli. Current situation and Challenge of big data processing Technology in Smart Grid [J]. Grid Technology, 2013 (04).
[2] Peng Xiaosheng, Deng Di Yuan, Cheng Shijie, Wen Jinyu, Li Zhaohui, Niu Lin. Key Technology of Electric Power big data for Smart Grid Application [J]. Chinese Proceedings of Electrical Engineering, 2015 (03).
[3] Liu Scientific Research, Sheng Wanxing, Zhang Dongxia, Jia Dongli, Hu Lijuan, he Kaiyuan. Application requirements and scenario Analysis of big data in Intelligent Distribution Network [J]. Chinese Proceedings of Electrical Engineering, 2015 (02).
[4] Chen Qingming, Xu Huan, Ye Peishan. Big data Analysis Technology of Smart Grid based on Cloud Computing [J]. Automation Technology and Application, 2020 (11).
[5] Sun Xiaohu, Qin Hao, Zhang Yaping, Wang Chao. Research on big data quality Calibration method of Power Grid based on Association rules [J]. Electronic Design Engineering, 2020 (21).