Power System Transient Stability Based on Data Mining Theory

Zhen Cui 1, Jia Shi 2, Runsheng Wu 1, Dan Lu 1, Mingde Cui 1

1State Grid Sichuan Tianfu new district power supply company, Chengdu 610000, China
2College of Electrical Engineering and New Energy, China Three Gorges University, Yichang 44300, China

*Corresponding author e-mail: cuizhen19881022@163.com

Abstract. In order to study the stability of power system, a power system transient stability based on data mining theory is designed. By introducing association rules analysis in data mining theory, an association classification method for transient stability assessment is presented. A mathematical model of transient stability assessment based on data mining technology is established. Meanwhile, combining rule reasoning with classification prediction, the method of association classification is proposed to perform transient stability assessment. The transient stability index is used to identify the samples that cannot be correctly classified in association classification. Then, according to the critical stability of each sample, the time domain simulation method is used to determine the state, so as to ensure the accuracy of the final results. The results show that this stability assessment system can improve the speed of operation under the premise that the analysis result is completely correct, and the improved algorithm can find out the inherent relation between the change of power system operation mode and the change of transient stability degree.

1. Introduction

With the rapid development of computer and network technology, human beings are stepping into the information society. The application of information technology in power industry is relatively early. So far, it has gone through three stages of development. In the power system, with the wide application of digital technology (such as switching equipment, digital relay protection and real-time control), power plant, substation and dispatching center time record a large number of electrical components and electrical energy information data [1]. With the deepening of the power system reform and the continuous improvement of the power market, each power plant should regularly offer power generation and power prices to the power companies, and participate in bidding for the internet [2]. It is necessary to integrate each information island into an auxiliary decision-making information system to ensure the safety, stability, quality and economic operation of the power system [3].

The scale of power system is increasing gradually, and the amount of data is increasing. However, the current computing technology lacks the description of the overall characteristics of data, and cannot make accurate and sufficient decisions according to the existing data. Power system calculation needs the combination of scientific theory and practical experience [4]. The purpose of this paper is to introduce
the data mining technology in the field of computer to the transient stability assessment of power system. The relationship between the operation mode and the transient stability is studied based on large data processing, and the calculation level of power system is improved.

2. Establishment of mathematical model

2.1. Selection of input feature

In the transient stability assessment of power system, the first problem is to determine the variable set and the variable domain of the input characteristics [5]. Generally, the following principles should be observed in the determination of input variables:

First, the number of input variables should not increase proportionally with the scale of the system, otherwise it is not suitable for the stability analysis of large-scale power systems [6]. Second, the state variables which are strongly related to transient stability should be chosen as an input variable. Third, although the system state of fault removal time is strongly related to stability, it needs to be obtained by time domain simulation, while it will increase computation and evaluation time. Such states should be chosen as input features to avoid losing speed advantage.

The commonly used input characteristic variables are divided into two categories: Single machine characteristic: It includes the power angle, speed and voltage of the single generator. If the single machine feature is selected, the number of the original input characteristics will increase proportionally with the increase of the system size, and the "Curse of dimensionality" problem will emerge [7]. Therefore, this method is not suitable for the stability analysis of large scale systems. System characteristic: By extracting or combining the state variables of the system, the system characteristics do not change significantly with the increase of the system size, so it is suitable for the stability analysis of large scale systems.

2.2. Network partition

Electrical network partitioning is the analysis of a problem in which an electrical network is divided into several sub-regions consisting of several buses or nodes. The origin of electrical network partitioning: Power system is a typical large system, and in most cases, the object of study is only part of it [8]. If the problem can be defined in a small range, it will simplify the scale of the problem and improve the efficiency of calculation. As far as the final purpose is concerned, the function of the network division is the same as that of the network equivalence, and the network is simplified. But as for the method, the former embodies the idea of divide and conquer, and the latter represents a whole equivalent relation.

The electrical network is divided to establish a new model of the electrical network. Each area is considered as a node, and the whole power system is regarded as a simplified system consisting of three parts: power output, power input and transmission line. After that, the statistics of the running data of the nodes in each region represent the operation of the area, and the input space of power system transient stability assessment is constructed. This avoids the curse of dimensionality caused by the use of node data as input characteristic variables. Generally, the division of an electrical network should follow three principles: First, the lines to be divided should be as few as possible. Second, the direction of power flow between the adjacent transmission lines should be consistent. Third, there is a high degree of electrical connection (strong coupling) between the buses in the same area, and the electrical connection between the buses in different regions is relatively weak (weak coupling).

3. Transient stability assessment based on data mining technology

The essence of the transient stability assessment method based on pattern recognition theory is to establish the mapping relation between the characteristic variables and the transient stability results, and construct a fast classifier for the state of the power system.
3.1. Actual situation of power system
As far as transient stability problem is concerned, it is closely related to the operation mode of power system. When the operation mode of power system changes, the experienced field dispatcher can roughly predict some state variables, such as bus voltage and line current. Because of the long-time field work, the dispatcher has grasped the relation between the state of the key parts that represent the safe running level of the power system and other control variables. They can predict the states of these critical parts according to the changing trend of control and disturbance in the power system. This phenomenon is largely due to the accumulation of long-term work experience of dispatchers.

The premise of the transient stability analysis based on the statistical method is that there is a sufficient amount of data that can comprehensively reflect the various possible operational conditions in the system. On the one hand, the demand for mass data processing in power system is continuously improving. On the other hand, it is necessary to solve the problem. These requirements make it ripe to use the historical data directly and analyze the system running status by pattern recognition method at the present stage.

3.2. Transient stability assessment index
There are two main methods to analyze the transient stability of power system: numerical simulation method and direct method. Numerical simulation is a method for solving numerical solutions of differential algebraic equations describing the dynamic behavior of a system. It has strong adaptability to the model of power system, and can get detailed information of electromechanical transient process of power system. The calculation staff can make a qualitative evaluation of the stability of the system according to the swing curve of the system. This method can only answer the question whether the system is stable, and cannot give the quantitative index of the stability margin of the system. The direct method is a method to evaluate the transient stability of power system by analyzing transient energy. The transient stability of the system can be evaluated quantitatively according to the transient energy margin.

As shown in the figure, in order to study the transient stability of power system, when the generator i is unstable, there are two kinds of changes in the kinetic energy of the generator. In the figure, $t_c$ means failure removal time, and $t_{b1}$ refers to the moment when $V_{KE1}$ reaches its minimum. $t_{a1}$ refers to the moment when $V_{KE1}$ reaches its maximum, and $t_{end}$ means the end of simulation computation.

![Figure 1. The variation curve of kinetic energy $V_{KE1}$](image)

4. Association classification
4.1. Hybrid method with time domain simulation
Transient stability of power system is a typical dynamic nonlinear problem, which determines that it is not enough to solve the problem of relying on the initial state of the problem. Most of the existing
solutions based on pattern recognition theory consider transient stability as a known initial condition. It selects the state of the system at a specific moment (fault occurrence time and fault removal time) as input characteristic variable to infer the dynamic development trend of the system. From the existing research results, it is not difficult to see that this method tries to predict the dynamic process of the system with stationary point data, and adjusts the training quality by adding and dropping characteristic variables and optimizing classifier structure design. However, no matter how it is improved, the nature of the dynamic characteristics of the power system cannot be neglected (or partly neglected) in the analysis process.

The characteristics of transient stability problems and pattern recognition methods determine the unreliability of the final evaluation results, which cannot be solved from the perfect algorithm alone. Because of the complexity of transient stability problems, it is difficult to solve them quickly and accurately only using one method. Therefore, in this section, combining the rapidity of pattern recognition method with the accuracy of time domain simulation method, a hybrid solution method is proposed to improve the computational accuracy and speed, so as to meet the needs of practical applications.

4.2. Association classification analysis with time series
In the practice of power system, there is a kind of data related to time series. For example, the voltage, current and circuit breaker action of the observation point are recorded at regular intervals so that a data set with time dimension is formed. The commonly used classification mining methods can only obtain the classification rules between the conditional attribute (characteristic variable) set and the decision attribute (class variable) set at a certain time, and establish the point to point mapping between the two data sets. Although the rules of the system can be judged according to the classification rules, the inherent relationship between the change of the conditional attributes and the change of the system state cannot be fully excavated. For example, it is not possible to determine which or what attributes change the state of the system.

For a continuous time-varying system such as power systems, the steady accidents in recent years show that most of the stability problems are not paid, but have undergone a process from quantitative change to qualitative change. More often, people want to understand the impact of dynamic changes in control volume on system stability, which can be seen as a mapping between data sequences. In this section, an improved algorithm is introduced. The time factor of variables is introduced into them, and an association analysis method based on time series is proposed.

4.3. Analysis results of association classification
The association classification method based on data mining is used for transient stability assessment. Simulation results show that this method has the advantages of high precision and short computation time. Moreover, the rules produced by the algorithm are easy to understand and have less human experience and practical value. In view of the nature of transient stability problems and the subjective limitations, the transient stability assessment based on pattern recognition method inevitably results in the deviation. Therefore, according to the stability index, the state of the system is divided into three parts: stability, instability and fuzzy stability. The first two parts are analyzed by using the fast calculation method of classified Association method. According to the critical stability, the accuracy of the algorithm and the speed of execution, the time domain simulation method is used to calculate the samples of fuzzy region. The simulation results show that the use of hybrid method ensures the accuracy of results and improves the computation speed.

5. Conclusion
Taking the transient stability problem of power system as the object of study, the architecture of transient stability assessment based on data mining is presented. In view of the different research stages of the problem, a practical mining algorithm is proposed. A mathematical model of transient stability assessment based on data mining technology is established. A power system is divided into zones based
on the principle of minimum cut. Taking the region as a unit, the statistical data of each region node is extracted as the input characteristic variable of the transient stability assessment problem. The curse of dimensionality is effectively avoided from the dimensionality reduction of spatial data. In addition, according to the different degree of the critical stability, the correlation classification method and the time domain simulation method are adopted to confirm the status of different stability levels. This stability assessment system can improve the speed of operation under the premise that the analysis result is completely correct. As a new interdisciplinary technology, data mining has great potential for application in power system. From the point of view of power data management, the calculation and analysis level of power system is improved, so as to open up a new field for the development of information technology in the future.

References
[1] Ayan, K., Kılıç, U., & Baraklı, B. (2015). Chaotic artificial bee colony algorithm based solution of security and transient stability constrained optimal power flow. International Journal of Electrical Power & Energy Systems, 64(64), 136-147.
[2] Shao, Y., Tang, F., Liu, D., Ma, Z., & Bian, C. (2015). An approach of transient angle stability assessment in power system for wams measured data. Dianli Xitong Baohu Yu Kongzhi/power System Protection & Control, 43(6), 33-39.
[3] Wang, B., Fang, B., Wang, Y., Liu, H., & Liu, Y. (2016). Power system transient stability assessment based on big data and the core vector machine. IEEE Transactions on Smart Grid, 7(5), 2561-2570.
[4] Omi, T., Kakisaka, H., & Iwamoto, S. (2016). Transient stability multi swing step-out prediction with online data mining. Ieej Transactions on Power & Energy, 136(2), 137-144.
[5] Ji, Luyu, Wu, Junyong, Zhou, & Yanzhen, et al. (2016). Energies, vol. 9, pages 898: using trajectory clusters to define the most relevant features for transient stability prediction based on machine learning method. Energies, 9(11), 898.
[6] Liu, Z., Liu, C., Li, G., Liu, Y., & Liu, Y. (2015). Impact study of pmsg-based wind power penetration on power system transient stability using ecac theory. Energies, 8(12), 13419-13441.
[7] Sharifian, A., & Sharifian, S. (2015). A new power system transient stability assessment method based on type-2 fuzzy neural network estimation. International Journal of Electrical Power & Energy Systems, 64(64), 71-87.
[8] Ji, L., Wu, J., Zhou, Y., Hao, L., Yu, Z., & Yan, J. (2015). Transient stability prediction of power system based on wams characteristic of perturbed voltage trajectory clusters. Gaodianya Jishu/high Voltage Engineering, 41(3), 807-814.