Advances in Engineering Research, volume 181
International Conference on Electronical, Mechanical and Materials Engineering (ICE2ME 2019)

# Application of the XGBOOST on the Assessment of Transient Stability of Power System

Sen Shen[1], Qunying Liu[2*], Xinchen Tao[2] and Shaojian Ni[2]
[1]dept. School of Optoelectronic Science and Engineering, University of Electronic Science and Technology of China, Sichuan, China
[2]dept. School of Automation Engineering, University of Electronic Science and Technology of China, Sichuan, China
[*]Corresponding author

***Abstract*—In the context of big data, machine learning plays an important role in many fields. With the increasing scale of power system and capacity of power grid, it becomes more and more complicated to accurately evaluate the transient stability of power system. In this paper, a power system transient stability assessment method based on XGBOOST algorithm is proposed. The XGBOOST algorithm is introduced to train the decision tree model and evaluate the transient stability of power system by converting the simulated power system operating data into the characteristic variables of power system. The results show that the training model of the algorithm can solve this kind of problem accurately and quickly.**

***Keywords—machine learning; power system; XGBOOST algorithm; transient stability***

## I. Introduction

With the sustainable growth of the world economy and the continuous development of power system, large-scale power grid is developed towards long-distance, EHV and even UHV. The power system in operation is huge in scale and complicated in physical process. The failure of components and the timing of a specific load level are stochastic; load forecasting and generator output are fuzzy; the original parameters of reliability (failure rate and repair rate of components, etc.) are also grey because of the error of statistical data and the shortage of prediction of future operating conditions and operating environment; in addition, decision-makers are unascertained due to subjective and cognitive uncertainties. It can be seen that the power system is a very complex dynamic system with randomness, suddenness, fuzziness, incomplete information uncertainty, and any accident of the system may bring great harm to the power system and even the whole society.

With the rapid development and popularization of WAMS/PMU, many scholars have studied the application of data mining and machine learning in power system on-line monitoring and security and stability assessment[1][2][3][4][5]. Reference [6] proposes a regional transient voltage stability assessment based on data mining. In [7], probabilistic neural network (PNN) and radial basis function (RBF) combined neural network (CNN) are used for transient stability assessment and margin prediction. In literature[8], the algorithm combining probability and traditional machine learning is used to realize system dynamic security assessment. In literature [9], the Bias classifier based on ensemble learning is used to evaluate the transient stability of power system. The XGBOOST is a machine learning function focused on gradient lifting algorithm, which was born in February 2014. It has attracted wide attention because of its excellent learning effect and high training speed. In this paper, a power system transient stability assessment method based on XGBOOST algorithm is proposed.

## II. Power System Transient Stability Assessment Based on XGBOOST Algorithm

### *A. Selection of Transient Characteristics in Power System*

The performance of the classifier is related not only to its own principles, but also to the selection of learning samples. A good set of learning samples can make the classifier reflect the difference between the classes of the recognition objects, but a bad one can't gain this effect. Because the information of the sample is reflected by the characteristic quantity, the quality of the learning sample set is related to the characteristic quantity. A good set of characteristic quantities can well reflect the differences between samples of different classifications and improve the performance of classifiers, so it is necessary to select the characteristic quantities of samples.

At present, there is little research focusing on the selection of power system transient stability characteristics. Because of the limitations of coverage, this paper lists the following fourteen characteristics representing power system transient stability[8].

Characteristic 1: The maximum initial acceleration of all generator rotors at the moment of failure, which reflects the instability trend of the most seriously disturbed generator.

Characteristic 2: The maximum kinetic energy of all rotors at the time of fault removal, which represents the maximum energy accumulation of the generator in the process of fault.

Characteristic 3: The initial angle of the generator with the maximum acceleration, which represents the static operation point of the most seriously disturbed generator.

Characteristic 4: The rotor angle of the maximum rotor kinetic energy generator at the time of fault removal, which reflects the deceleration ability of the maximum rotor kinetic energy generator after fault removal.

Characteristic 5: The rotor kinetic energy of the generator with the largest rotor angle at the time of fault removal, which reflects the instability trend of the generator with the largest rotor angle at the time of fault removal.

Characteristic 6: Total energy adjustment of the system.

Characteristic 7: The minimum initial acceleration of all rotors, which reflects the instability mode of the hysteretic generator relative to the inertial center.

Characteristic 8: Root mean square error of all initial accelerations, which represents the dispersion degree of the acceleration motion of the generator rotor. Generally, the larger the value is, the more unbalanced the disturbance to the generator in the system is, and the more likely the system will be unstable.

Characteristic 9: The average kinetic energy of all rotors at the time of fault removal, which reflects the total kinetic energy increment of all generators.

Characteristic 10: The difference of the maximum power angle at the time of fault removal, which reflects the electrical distance between the front and rear engine.

Characteristic 11: The maximum kinetic energy difference of all generators at the time of fault removal.

Characteristic 12: The difference between the maximum angular velocity at the time of fault removal.

Characteristic 13: The average initial acceleration of all generators, which reflects the average disturbance to all generators in the system.

Characteristic 14: The sum of mechanical input power of all generators before faults, which reflects the generation level of the generators in the system before faults. The generation level of the system in the power system corresponds to the load level from time to time and it directly reflects the level of the load level.

### *B. Two Screening of Characteristic Quantities by XGBOOST*

In order to improve the accuracy and generalization ability of the classifier, it is necessary to reduce the number of features as much as possible, and select features that have significant impact on the results as the classification criteria. The built-in function xgb.plot.importance of the XGBOOST algorithm can calculate the importance of the features and present the results in graph form. By gradually eliminating the features with low importance and preserving the features that have significant impact on the results, 19 characteristic quantities that can characterize the transient stability of power system are extracted. The screening process is illustrated in Figure I.

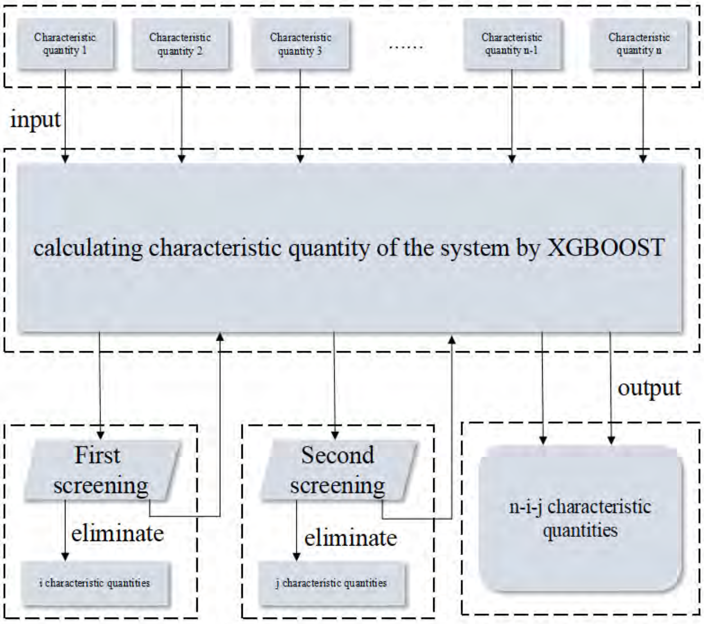


FIGURE I. SCREENING FLOW CHART OF CHARACTERISTIC QUANTITIES BY XGBOOST

## III. EXAMPLE ANALYSIS

### *A. Data Generation*

1200 samples were generated in the New England 39-node system, and all of them were the running state of the system within 1 s after a branch short circuit. Stable and unstable samples accounted for fifty percent of the total. 80% of them were selected randomly as training samples and 20% as test samples.

### *B. Full Feature Simulation and Simulation after the Screening Feature*

Twenty-three characteristic quantities are used as input, and XGBOOST is used for training and testing. The classification accuracy is obtained and compared with other traditional machine learning algorithms in Table I. The importance of the characteristic quantity is shown in Figure II. As can be seen from Figure II, the importance of these three characteristic quantities which are the maximum initial acceleration of all generator rotors at fault moment, the performance index representing the distance between the system and the stable equilibrium point after fault removal, and the difference of the maximum kinetic energy between all generator at the time of removal time to the result is 1. After removing these three characteristic quantities and retesting, we got the simulation after the first screening feature. The results and the importance of the feature are represented by Table II and Figure III respectively. According to the simulation result after the first screening feature, the importance of the minimum active power impact of a single generator is 1 at the beginning of the characteristic fault. After removing this characteristic quantity, the results and the importance of the feature are represented by Table III and Figure IV respectively.

TABLE I. COMPARISON BETWEEN XGBOOST ALGORITHM AND TRADITIONAL ALGORITHM

| Training algorithm | Misjudgment samples | Accuracy |
|---|---|---|
| XGBOOST | 7 | 96.25% |
| Logistic | 15 | 93.75% |
| C4.5 | 13 | 94.58% |

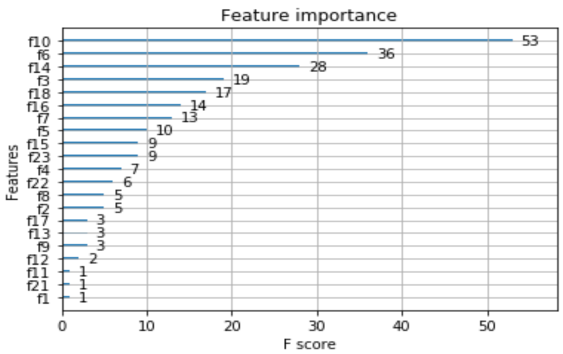


FIGURE II. IMPORTANCE OF CATEGORIZATION QUANTITIES

TABLE II. COMPARISON BETWEEN XGBOOST ALGORITHM AND TRADITIONAL ALGORITHM

| | Feature before screening | Feature after screening |
|---|---|---|
| Accuracy | 96.25% | 96.67% |

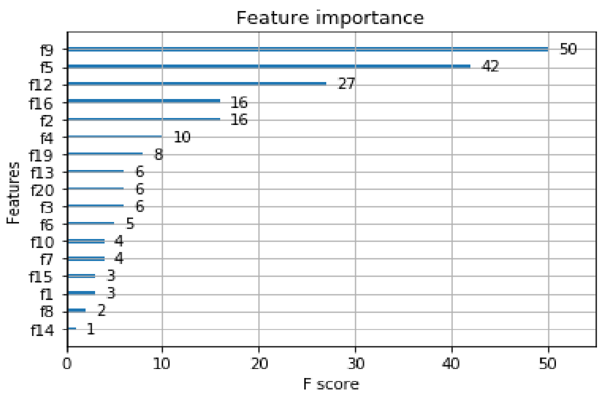


FIGURE III. IMPORTANCE OF CATEGORIZATION QUANTITIES

TABLE III. COMPARISON BETWEEN XGBOOST ALGORITHM AND TRADITIONAL ALGORITHM

| | Feature before screening | Feature after screening |
|---|---|---|
| Accuracy | 96.67% | 97.08% |

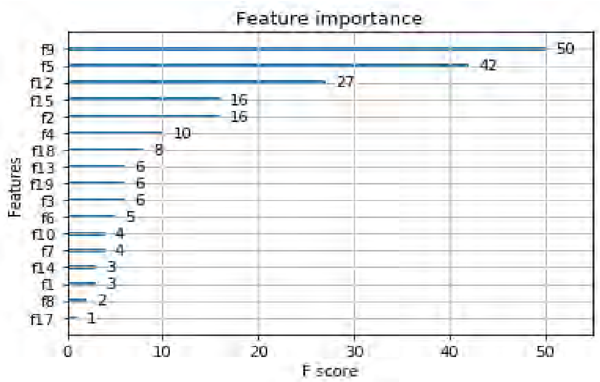


FIGURE IV. IMPORTANCE OF CATEGORIZATION QUANTITIES

## *C. Example Verification*

The decision tree model trained by XGBOOST algorithm in the previous section is validated in this part.

By simulating the running state of the new England 39-node system, the location of the fault is set up artificially. When the stability of the system is unknown, the 19 characteristic quantities are calculated which represent the stability of the system given by the XGBOOST algorithm and then these characteristic quantities are put into the decision tree model to judge whether the system is unstable or not. The performance of the decision tree model is judged by comparing with the simulation results.

Table IV shows the system monitoring data when branch 5 has three phase short circuit fault.

Table V gives the system characteristic quantities calculated from the above monitoring data.

We judged that the system is unstable by bringing the characteristic quantities in Table IV into the known model.

TABLE IV. SYSTEM MONITORING DATA FOR THREE PHASE SHORT CIRCUIT FAULT

| Time (second) | Generator _BUS30δ (deg.) | Generator _BUS30ω (p.u.) | …… | Generator _BUS39Q (p.u.) | Generator _BUS39Pm (p.u.) |
|---|---|---|---|---|---|
| **0** | 28.98189 | 1 | | 0.13851 | 8.3 |
| **0.01** | 29.01086 | 1.00032 | | 3.33001 | 8.3 |
| **0.02** | 29.09779 | 1.00064 | | 3.33054 | 8.3 |
| **0.03** | 29.24267 | 1.00097 | | 3.33143 | 8.3 |
| **0.04** | 29.44553 | 1.00129 | | 3.3327 | 8.3 |
| **0.05** | 29.70639 | 1.00161 | | 3.33435 | 8.3 |
| **……** | | | | | |
| **0.93** | 178.86524 | 1.01599 | | 3.09883 | 8.3 |
| **0.94** | 181.76023 | 1.01618 | | 3.14148 | 8.3 |
| **0.95** | 184.69409 | 1.01642 | | 3.16073 | 8.3 |
| **0.96** | 187.67452 | 1.0167 | | 3.1561 | 8.3 |
| **0.97** | 190.70936 | 1.01702 | | 3.12834 | 8.3 |
| **0.98** | 193.80624 | 1.01739 | | 3.07958 | 8.3 |
| **0.99** | 196.97216 | 1.01779 | | 3.01353 | 8.3 |
| **1** | 200.21315 | 1.01822 | | 2.93554 | 8.3 |

TABLE V. SYSTEM CHARACTERISTIC QUANTITIES

| | |
|---|---|
| The maximum kinetic energy of all rotors at the time of fault removal | 18.21290125 |
| The initial angle of the generator with the maximum acceleration | -5.086638508 |
| The rotor angle of the maximum rotor kinetic energy generator at the time of fault removal | 201.3471015 |
| The rotor kinetic energy of the generator with the largest rotor angle at the time of fault removal | 0.911084639 |
| Total energy adjustment of the system | 9114.207487 |
| The minimum initial acceleration of all rotors | -0.073954431 |
| Root mean square error of all initial accelerations | 0.034476623 |
| The average kinetic energy of all rotors at the time of fault removal | 2.426366469 |
| The difference of the maximum power angle at the time of fault removal | 663.90647 |
| The difference between the maximum angular velocity at the time of fault removal | 0.08909 |
| The average initial acceleration of all generators | 0.03273895 |
| The sum of mechanical input power of all generators before faults | 45.04865 |
| The maximum active power impact on a single generator at the start of the fault | 3.57215 |
| The maximum reactive power impact on a single generator at the start of the fault. | 6.16331 |
| The minimum reactive power impact on a single generator at the start of the fault | 1.30852 |
| The maximum normalized active power impact on a single generator at the start of the fault | 0.042025294 |
| The minimum normalized active power impact on a single generator at the start of the fault | -0.044477059 |
| The performance index representing the instability trend and structural change of the most seriously disturbed generator | 0.06728 |
| The performance index representing the impact of the fault on the system (deceleration power) | 0.0199988 |

Figure V shows the running state curve of the system obtained by simulation. It can be seen that the system is unstable after three-phase-short-circuit fault occurs in line 5. The results obtained by XGBOOST algorithm are in agreement with the facts.

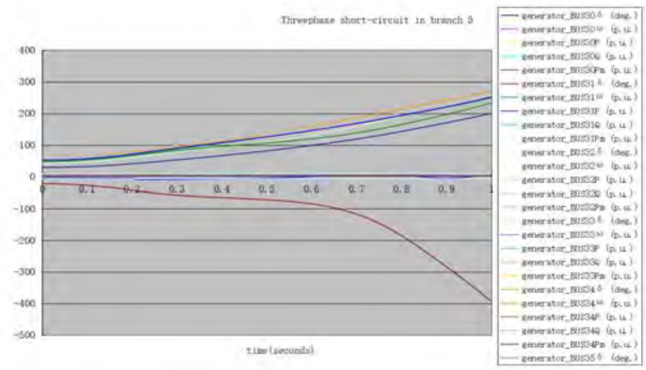


FIGURE V. OPERATION CURVE OF THE SYSTEM AFTER THREE PHASE SHORT CIRCUIT

Then, the different fault locations are set up and the above experiments are repeated ten times. The results are shown in Table V:

TABLE VI. EXPERIMENTAL RESULTS OBTAINED FROM REPEATED TRIALS 10 TIMES

| Fault locations | The stability of the system | Is the result of the decision tree model consistent with the facts |
|---|---|---|
| Three phase short circuit in branch 9 | unstable | Yes |
| Three phase short circuit in branch 10 | unstable | Yes |
| Three phase short circuit in branch 11 | unstable | Yes |
| Three phase short circuit in branch 12 | unstable | Yes |
| Three phase short circuit in branch 21 | unstable | Yes |
| Three phase short circuit in branch 22 | stable | Yes |
| Three phase short circuit in branch 16 | stable | Yes |
| Three phase short circuit in branch 23 | stable | Yes |
| Three phase short circuit in branch 1 | stable | Yes |
| Three phase short circuit in branch 6 | stable | Yes |

## IV. SUMMARY

For the characteristics of power system such as vulnerability, non-linearity, randomness, fuzziness and suddenness, it is difficult to evaluate the transient stability of power system quickly and accurately by using traditional methods. In this paper, an evaluation method of power system transient stability based on XGBOOST is proposed. The XGBOOST algorithm can automatically learn the splitting direction of a sample with missing eigenvalues, and can process the table data well. Its model can be interpreted because it uses some methods such as adding regular terms to the cost function, reducing the learning rate to prevent the model from over-fitting, using parallel approximate histogram algorithm to improve the computational efficiency, etc. A good and built-in feature importance function can eliminate the features that have little effect on the classification results and reduce the difficulty of data processing. So it has more advantages than traditional methods.

The above example shows that the XGBOOST algorithm is effective in solving the transient stability problems of power system, and is obviously superior to the traditional machine learning algorithm in evaluation accuracy and the computation speed.

## ACKNOWLEDGMENT

Our deepest gratitude goes to the anonymous reviewers for their careful work and thoughtful suggestions that have helped improve this paper substantially. This work was supported by Natural Science Foundation of China under Grant NSFC51677020, the Provincial Key Laboratory of power electronics energy saving technology and equipment (No.szjj2016-093) , the 2017 Fundamental Research Funds for the Central Universities and the Key research and development projects in Sichuan province (No. 2019YFG0129).

## REFERENCES


[1] Sun K, Likhate S, Vittal V, et al. An online dynamic security assessment scheme using phasor measurements and decision trees[J]. IEEE Transactions on Power Systems, 2007, 22(4): 1935-1943.

[2] Zhou D Q, Annakkage U D, Rajapakse A D. Online monitoring of voltage stability margin using an artificial neural network[J]. IEEE Transactions on Power Systems, 2010, 25(3): 1566-1574.

[3] Li Dahu, Cao Yijia. Power system transient stability analysis based on PMU and hybrid support vector machine[J]. Power System Technology, 2006, 30(9): 46-52(in Chinese).

[4] Wu Jiekang, Deng Song, Liang Zhiwu, et al. Evaluation of power system voltage stability based on fuzzy neural network decision tree[J]. Power System Technology, 2008, 32(14): 25-29(in Chinese).

[5] Zhu Lipeng, Lu Chao, Sun Yuanzhang, et al. Data mining based regional transient voltage stability assessment[J]. Power System Technology, 2015, 39(4): 1026-1032(in Chinese).

[6] Yao Dequan, Jia Hongjie, Zhao Shuai. Power system transient stability assessment and margin prediction based on composite neural network[J]. Power system automation, 2013, 37(20): 41-46(in Chinese).

[7] Ye Shengyong. Research on power system transient stability assessment based on machine learning[D]. Southwest JiaoTong University, 2010(in Chinese).

[8] Li Hongwei. Power system transient stability assessment based on pattern recognition and ensemble learning[D]. North China Electric Power University (Hebei province), 2010.

[9] Chen Tianqi, HE Tong, BENESTY M, etal. XGBOOST: Extreme gradient Bosting[EB / OL]. [2017-01-05]. http: / / github. com / dmlc / XGBOOST.