

Extract Voltage Stability Operating Guidelines using Decision Tree

Dengfeng Li^{1, a}, Xiangfei Meng^{2, b,*}, Ruilin Xu¹, Xiaojun Zhu¹, Ke Zhao¹, Yi He²,
Pei Zhang²

¹Chongqing Electric Power Research Institute, Chongqing Electric Power Corp., State Grid Corporation of China, Chongqing, 401120, China

²School of Electrical Engineering, Beijing Jiaotong University, Beijing, 100044, China

^aemail: 597913911@qq.com, ^bemail: 316327192@qq.com

Keywords: Voltage Stability; Operating Guideline; Decision Tree; C4.5 algorithm

Abstract. The operational planners in power grid companies perform a large number of offline voltage stability studies for a variety of operating conditions. How to effectively determine voltage stability limits from massive amount of offline data has become a challenge for operational planners. This paper applies decision tree (DT) to extract operating guidelines from off-line analysis results. Case study results on a practical power system prove that voltage stability operating guidelines, which are extracted by decision tree, can be used by system operators for online monitoring and control of electric power systems.

Introduction

Voltage stability is the focus of planning and operating electric power system. In order to ensure the power system to maintain acceptable voltage both in normal operation and in the event of severe disruption, the operational planners have done a lot of offline analyses and collected a large amount of offline data [1][2]. The operational planners developed voltage stability operating guidelines based on their experience and subjective judgement. With the development of power system, the structure of the power grid has become very complex and the scale of computational data has increased dramatically. Meanwhile, large scale integration of renewable energy resources increase complexity and uncertainty of power system operation. Once facing with the massive data, the efficiency of manual data processing will be very low and the accuracy cannot meet the needs of system operation [3][4].

Due to the advantages in extracting useful information automatically from massive data, data mining techniques have been applied into power system. Recent applications of data mining techniques, including artificial neural network (ANN), self-organizing map (SOM) and radial basis function (RBF), support vector machine (SVM) [5], Reference [6] applied a multiclass SVM for static voltage stability assessment. Reference [7] proposed a new hybrid classifier for transient stability prediction, which can forecast both TSS of the power system and synchronism state of each generator in response to a disturbance. A new approach to ANN based VCPI prediction method to be used for extracting voltage stability operating guidelines has been proposed in reference [8]. The above mentioned methods all belong to black box model, which cannot reveal the internal rules of the massive data.

Decision tree belongs to white box model. One of the advantages of decision trees is that they produce models that are relatively easy to interpret. This paper uses decision tree to extract voltage stability operating guidelines.

Decision Tree and C4.5 algorithm

A decision tree is a flowchart-like tree structure, where each internal node denotes a test on an attribute, each branch represents an outcome of the test, and each leaf node holds a class label. The topmost node in a tree is the root node [9]. A typical decision tree is shown in Figure 1.

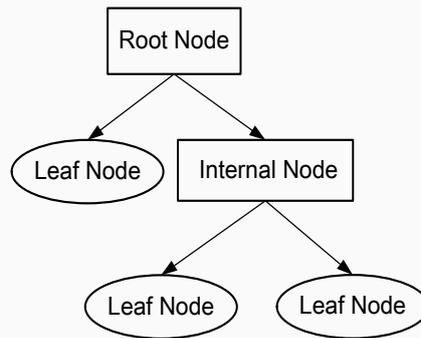


Fig. 1. Decision tree

The most advantage of decision trees is that they produce models that are relatively easy to interpret. In particular, a tree can be read as set of decision rules, with each rule's antecedent comprising the decisions on the internal nodes along a path to a leaf, and its consequent being the label of the leaf node. Further, because the regions are all disjoint and cover the entire space, the set of rules can be interpreted as a set of alternatives or disjunctions [10].

There are several decision tree algorithms, such as ID3, C4.5, and CART algorithm. All of them adopt a nonbacktracking approach in which decision trees are constructed in a top-down recursive divide-and-conquer manner. Most algorithms for decision tree induction also follow a top-down approach, which starts with a training set of tuples and their associated class labels [11].

In order to handle continuous attributes, this paper uses C4.5 algorithm to extract voltage stability operating guidelines. C4.5 algorithm is the most common algorithm of decision tree, which developed from ID3 algorithm. It creates a threshold and then splits the list into those whose attribute value is above the threshold and those that are less than or equal to it [12]. The core of C4.5 algorithm is the entropy principle. Entropy is used to measure the amount of disorder or uncertainty in a system. In the classification setting, a partition has lower entropy means it is relatively pure, that is, if most of the points have the same label [13].

Suppose set S is a training samples set, C represents attributes of decision tree and m represents class labels (voltage stable or voltage unstable).

The entropy of set S is defined as follows:

$$Entropy(S) = -\sum_{i=1}^m p_i \log_2 P_i \quad (1)$$

Where $p_i (i=1,2,\dots,m)$ is the frequency of the class attribute C associated with the m class labels.

Select an attribute A then calculate the information entropy of partition according to attribute A :

$$Entropy_A(S) = \frac{|S_L|}{|S|} Entropy(S_L) + \frac{|S_R|}{|S|} Entropy(S_R) \quad (2)$$

S_L and S_R correspond to different molecular sets which are split by the division point.

The information gain of attribute A to training set is obtained using the following equation:

$$Gain(S, A) = Entropy(S) - Entropy_A(S) \quad (3)$$

C4.5 introduces the attribute's split information $SplitE(A)$. The attribute's split information can be calculated using the following equation.

$$SplitE(A) = -\sum_{i=1}^k \frac{|S_i|}{|S|} \log_2 \frac{|S_i|}{|S|} \quad (4)$$

The information gain ratio of the attribute A can be calculated using the following equation.

$$GainRatio(A) = \frac{Gain(A)}{SplitE(A)} \quad (5)$$

The attributes are sorted by the information gain ratio. The attribute with the maximum information gain ratio is selected as the node.

Select a feature as the node of this level, after testing all the branches, calculate the maximum information gain ratio, then construct the next level of decision tree.

In order to extract voltage stability operating guidelines for dispatchers, decision tree is constructed by the massive offline data. The flow chart of applying decision tree for extracting guidelines is shown in Figure 2.

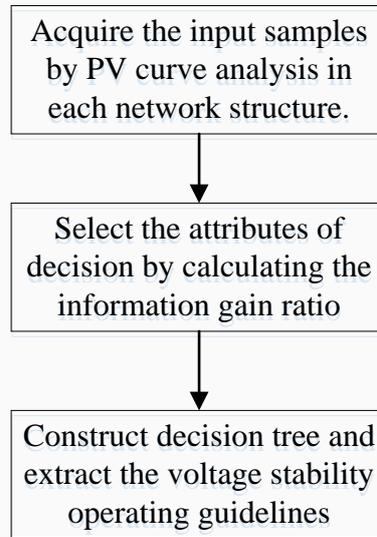


Fig. 2. The flow chart

First, in order to use the decision tree to extract voltage stability operating guidelines, the operating points before the system reaches voltage stability critical points and the voltage stability critical points which obtained by PV curve analysis in all of generators, transmission lines and transformers ‘N-1’ contingencies are selected as the input sample of decision.

Second, the attributes of decision tree include voltage magnitude and angle of buses, the reactive power of branches and the reactive power generated by generators to construct the decision tree, the information gain ratio of these four attributes in every input samples need to be calculated and select the maximum information gain ratio as the root node of decision tree.

Third, after obtaining the root node, the sample split into two subsamples. The leaf node can be obtained by calculating the maximum information gain ratio in each subsample. Decision tree will repeat this process automatically until the sample of the leaf node meet the stopping criteria of decision tree. Once decision tree is constructed, a set of voltage stability operating guidelines can be generated which are used for dispatchers to monitor voltage stability conditions in real-time.

Case study

In order to verify the feasibility of the proposed method, case study is conducted using an actual power system data. The test system includes 781 buses, 40 500kV transmission lines, 145 loads, 40 generators and 271 transformers. The total generator capacity is 12478MW and the total load is 12259.7MW.

P-V curve analysis is performed to generate samples for decision tree. ‘N-1’ contingency analysis is conducted on all of generators, transmission lines and transformers to obtain different operating conditions. In total, there are 314 contingencies. Figure 3 shows part of P-V curve analysis results.

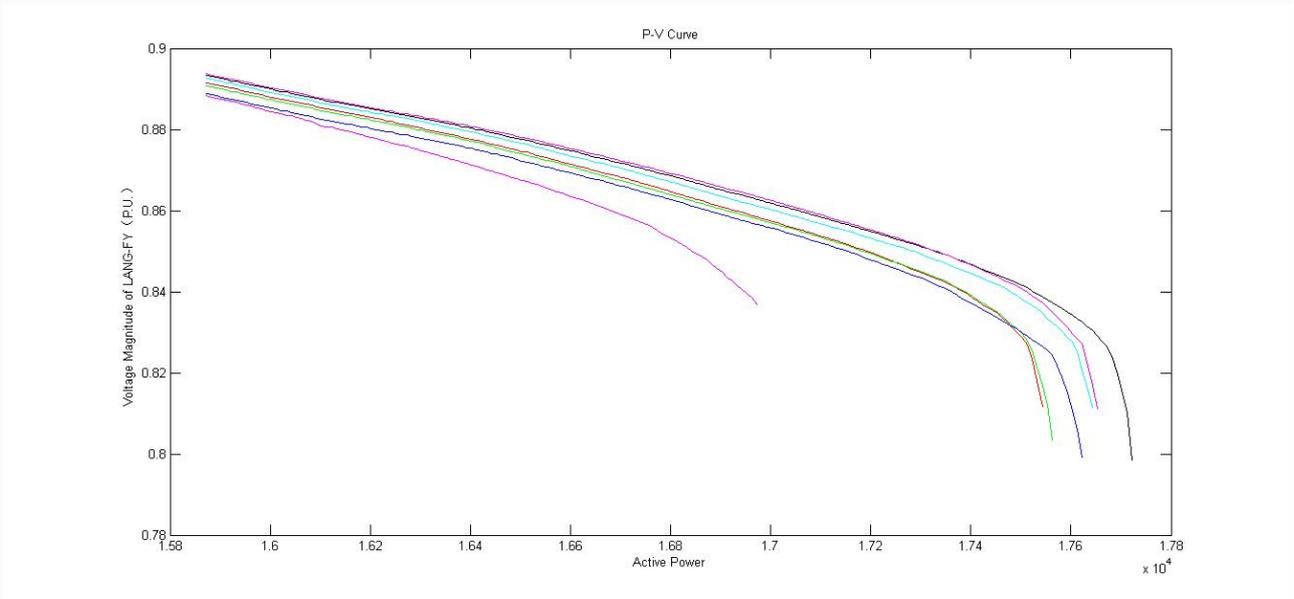


Fig. 3. P-V curve analysis for different contingencies.

314 critical voltage stability limit points can be obtained in 314 contingencies. These points are composed of the voltage unstable sample. In each P-V curves, the last operating point before the system reaches voltage stability limit is used to construct voltage stable sample. Each sample includes 75 attributes (11 critical branches transmission power and 64 critical buses node voltage magnitude and phase angle) and a target attribute (voltage stability / voltage instability). C4.5 algorithm is used to construct decision tree using 75 attributes.10-fold cross-validation approach is used to select the training set and validation set of decision tree.

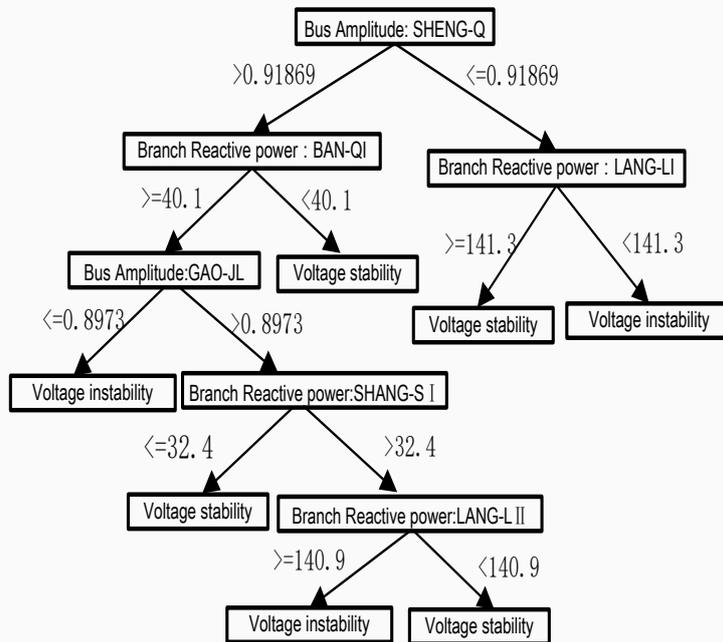


Fig. 4. Decision rules

Figure 4 shows the decision tree constructed based on C4.5 algorithm. Voltage stability rules can be extracted from this decision tree. According to decision tree, the following rules can be extracted:

(1) If voltage magnitude of Bus SHENG-Q is above 0.91869 and reactive power of Branch BAN-Q is equal or above 40.1Mvar as well as voltage magnitude of Bus GAO-JL is equal or less than 0.8973, the system operators can judge that the system voltage is unstable.

(2) If voltage magnitude of Bus SHENG-Q is above 0.91869. The reactive power of Branch BAN-Q is equal or above 40.1Mvar. The voltage magnitude of Bus GAO-JL is above 0.8973 and

the reactive power of Branch LANG-L II is equal or above 140.9Mvar, the system operators can judge that the system voltage is unstable.

(3) If voltage magnitude of Bus SHENG-Q is equal or less than 0.91869 and reactive power of Branch LANG-L II is less than 141.3Mvar, the system operators can judge that the system voltage is unstable.

These rules can be used by system operators as operating guide to assess system voltage stability conditions.

Conclusions

Operational planners make a lot of studies on voltage stability to keep the stability and safety of power system. This paper applies decision tree to extract voltage stability operating guidelines based on massive offline data.

With the development of power system and the widely use of PUMs, the attributes of decision tree can be obtained by massive offline data, which used to monitor system voltage stability conditions in real-time.

Study on a practical power system indicates that decision tree can extract voltage stability operating guidelines and therefore help dispatchers assess system voltage stability in real-time.

References

- [1] P.Kundur, *Power System Stability and Control*, 1st ed., McGraw-Hill, 1994
- [2] Taylor C.W. *Power System Voltage Stability*. New York: McGraw 2Hill Inc, 1994
- [3] M. K. Jalboub, A. M. Ithal, H. S. Rajamtani, R. A. Abd-Alhameed and A. M. Ithal, "Determination of static voltage stability-margin of the power system prior to voltage collapse," *Systems, Signals and Devices (SSD)*, 2011 8th International Multi-Conference on, Sousse, 2011, pp. 1-6.
- [4] J. Min, H. j. Wang and J. Ma, "A comprehensive approach for static voltage stability preventive control using immune algorithm," *Information Technology and Artificial Intelligence Conference (ITAIC)*, 2011 6th IEEE Joint International, Chongqing, 2011, pp. 438-441.
- [5] Bishop, C. (2006). *Pattern Recognition and Machine Learning*. Information Science and Statistics. New York: Springer Science+Business Media.
- [6] Kalyani S, Shanti S S. Classification and assessment of power system security using multiclass SVM [J]. *IEEE Transactions on Systems, Man, and Cybernetics-Part C: Applications and Reviews*, 2011, 41(5): 753-758.
- [7] Amjady N, Banihashemi S A. Transient stability prediction of power systems by a new synchronism status index and hybrid classifier [J]. *IET Generation, Transmission and Distribution*, 2010, 4(4): 509-518.
- [8] B. Suthar and R. Balasubramanian, "A new approach to ANN-based real time voltage stability monitoring and reactive power management," *TENCON 2008 - 2008 IEEE Region 10 Conference*, Hyderabad, 2008, pp. 1-6.
- [9] Feng Shaorong: Research and Improvement of Decision Trees Algorithm. *Journal of Xiamen University(Natural Science)*, 2007, pp.498-500.
- [10] Quinlan, J. R. (1986). "Induction of decision trees." *Machine Learning*, 1 (1): 81–106.
- [11] Zhu Xiaoliang, Yan Hongcan, Wang Jian and Wu Shangzhuo, "Research and application of the improved algorithm C4.5 on Decision tree," *2009 International Conference on Test and Measurement*, Hong Kong, 2009, pp. 184-187.

[12] Fucan Chen, Xiaowei Li and Lixiong Liu, "Improved C4.5 decision tree algorithm based on sample selection," *Software Engineering and Service Science (ICSESS)*, 2013 4th IEEE International Conference on, Beijing, 2013, pp. 779-782.

[13] Quinlan, J. R. (1993). *C4.5: Programs for Machine Learning*. New York: Morgan Kaufmann.