Voltage Security Events Classification of Power System Using Machine Learning Techniques

Sanjiv Kumar Jain*, N. P. Patidar**, Yogendra Kumar***
*Electrical Engg. Dept., Medi-Caps University, Indore. sanjivkj@gmail.com
** Electrical Engg. Dept. M.A.N.I.T., Bhopal. nppatidar@yahoo.com
*** Electrical Engg. Dept. M.A.N.I.T., Bhopal. ykmact@yahoo.co.in

ABSTRACT

Voltage insecurity is a most vital problem that the modern power systems are imposed off. Timely and accurate assessment of voltage security is necessary to detect post-contingency voltage problems in order to prevent a large scale blackout. The paper presents machine learning techniques based on CART algorithm and probabilistic decision tree based strategy for estimation of voltage security. To reduce redundancy and to improve efficiency certain critical contingencies are considered. The data for inputs are fuzzified and outputs target class are utilized for decision tree learning method. Probabilistic classification trees are most suitable for estimating stability problems because the objective is secure or insecure subsequent a credible contingency. The proposed method is tested using IEEE 30 bus test system. The advantage of fuzzy decision tree based exposure of voltage security is the robust classification of future samples. Also the performances of obtained results are shown through accuracy tests. The fuzzy decision tree method is also compared with CART algorithm.

Key-words: *CART, classification, fuzzy decision tree (FDT), power system security, voltage collapse.*

1. INTRODUCTION

The current power systems operation is closer to their voltage limits are dictated by the needs of deregulated electricity markets. Many collapses have occurred as a result of voltage instability. Hence the voltage stability has been a subject of severe concern for system operators. It requires extensive investigation because it plays an important role to enhance the security of the system and the quality of the power [1].

As the size and complexity of power system grows, the list of contingencies that should be taken into count during voltage security [2] evaluation also grows. In practice it is very infeasible to evaluate real time power flow with large number of outages. In contingency screening we can determine, only the contingencies, which may responsible for voltage insatiability. It will reduce the number of contingencies that should be analyzed by full AC load flow.

Contingencies are ranked in order of their severity for the help of the operator. A suitable scalar voltage performance index is used for ranking. The operator uses this information to take corrective action in order to move the system towards the more secure operating state as necessary.

Voltage collapse is one of the most critical problems that threaten system secure operations. It is usually initiated by either a continuous load increase and/or a major change in network topology resulting from a critical contingency. During a voltage collapse event, bus voltages in a localized area decrease below an acceptable level. Without timely control actions, the low voltages may spread throughout adjacent areas of the power system and may eventually cause a large scale blackout instead of a localized outage. The time span for voltage instability ranges from 0.1 s to 1 h, representing transient and long-term voltage instability, respectively [3], [4]. For the fast transients, detailed voltage security analysis for online applications is a great challenge because of the high computational burden. As a result, an accurate assessment tool to determine the voltage security in real time becomes a necessity for system operators to determine remedial controls for preventing voltage collapse. Moreover, such a disturbance can further result in Cascading outages may resulted and cause a lasting power interruption to clients in a large area[5]. Therefore, the development of a tool to accurately assess security issues and provide sufficiently fast prediction results in case of critical contingencies is an important requirement for real-time operation of modern power systems.

Many machine learning techniques such as DT [6], Binary SVM based pattern recognition [7], Artificial Neural Networks [8] and genetic algorithms, Multi-Objective Genetic Algorithm[9] have been proposed.

The paper focuses voltage security assessment based on loadability margin of IEEE 30 bus system using trained decision trees and further classification of

system based on new samples. The probabilistic model is used to cover the range from light load to overloads efficiently. The projected method is tested on 30 bus IEEE power systems. The results show that the method gives highly efficient result to estimate the voltage security in verification test phase.
The performance is compared in terms of percentage accuracy of classification for FDT with CART [10-12] algorithm.

2. VOLTAGE SECURITY ASSESSMENT

In a large size power system the possible number of contingencies including multiple ones will be an astronomical number. However, the probability of occurrence of a large number of these contingencies is quite low.
P-Q and P-V curve based methods use repetitive load flow simulations for various load conditions to calculate the static voltage stability margin during the post disturbance period [4]. The main drawback of these methods is that the Jacobian matrix becomes singular near the maximum loading point, so it is numerically difficult to obtain the power flow solution. Also, these methods are computationally expensive and could not be used to detect short-term voltage instability
When applied to power system security assessment, classification trees are quite suitable for evaluating stability problems because the prediction objective is secure or insecure following a contingency.
Hence, keeping in view the real-time requirement of the security function the aim of contingency definition is to reduce this list to a shorter one containing only credible contingencies. The probability of occurrence of credible contingencies is very high. Hence the credible contingencies [6] are first ranked in order of their severity and a few most severe contingencies are short-listed in Table1 for evaluation.

Table 1: Credible contingencies

Rank (critical)	CPF method	
	Line outage no.	From bus-to bus
1	1	(1-2)
2	5	(2-5)
3	36	(28-27)
4	2	(1-3)
5	4	(3-4)

3. PROBABILISTIC FUZZY DECISION TREE [13]

Machine Learning is a branch of artificial intelligence. Decision tress are one of the methods which comes under machine learning. A decision tree[14] in its basic form is a collection of 'test nodes' and 'terminal nodes', structured in a tree, upside down. The test node is connected with the testing of some 'attribute' with a threshold value. An attribute is a factor relevant to the concerned classification problem. Each descendant of a test node corresponds to a possible result of the test. The class is indicated by terminal nodes to which it belongs.

The basis of Decision Tree is Knowledge Base (KB) having number of functional points. The KB is divided in Learning Set (LS) and the Testing Set (TS) for performance evolution of new operating points. Proceeding from the root node, an attribute is chosen, say A_i, and let the threshold value of the A_i, is t, then dichotomy test T is defined as

$$T : A_i = t \qquad (1)$$

This test decomposes the training cases into two subsets corresponding to the two successors of the root node. Attributes and their threshold (β) value are chosen in such a way so as to obtain the maximum information gain about the classification.
Database generation consist of large number of operating states of the power system. Database is generated by performing full AC load flow under varying real and reactive loads at all the buses randomly in a wide range (50% to 150%) and under each line outage to monitor the situations relevant. For the fuzzification[15] of data samples trapezoidal membership function is employed. Each attribute is partitioned in three categories.
Decision based clustering is done of whole data samples after fuzzification. Membership

value is added attribute wise for each clustering. Entropy and information gain is calculated based on calculated data for selection of best attribute for classification.
Shannon entropy is used where p_i is the *priori* probability of class *i* defined as

$$\text{Entropy } H(s) = \sum_{i=1}^{N} (-p_i)\log_2(-p_i) \qquad (2)$$

Where p_i is the ratio of class *i* defined as $\frac{N_i}{N}$

After the selection of best attribute, classification is done at each parent node until the stopping criteria is achieved if the proportion of the data set of a class is greater than or equal to a value. If the criteria is fulfilled the tree is terminated as leaf. Hence a tree taking into account the complete data set for each credible contingency. Three methods are used for estimation of accuracy of a Decision Tree. For the testing data sets a test of precision is carried out. The Overoptimistic, the first category is the re-substitution of estimates, in which same data is used to train the DT for further computation. Second which is used in the paper is test sample estimates. The database cases are dived randomly into a learning set (LS) and testing set (TS). To train a DT LS is used, while TS is utilized to evaluate the accuracy of the formed DT. Cross-validation the third parsimonious with data and favourite for databases with low size. The accuracy is given by Eq.

$$Accuracy = \left(\frac{Number\_of\_well\_classified\_classes}{Total\_no.\_of\_cases} \right)$$

3.1 Fuzzy decision tree algorithm:

Data sets obtained from load flow solutions due to random variations of loads are fuzzified as shown in Fig.1 according to following trapezoidal membership function. Each attribute is partitioned into three categories.

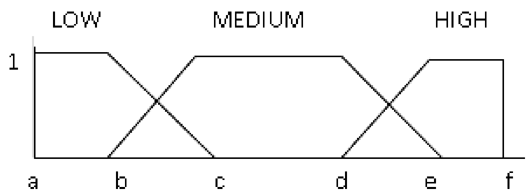

Fig. 1: Fuzzification of attributes

The point of discourse 'a' is minimum and 'f' is maximum values of attributes (P, Q) respectively.
Point b – minimum value + 20% (max-min)
Point c – minimum value + 40% (max-min)
Point d – minimum value + 60% (max-min)
Point e – minimum value + 80% (max-min)

4. CLASSIFICATION ALGORITHM

In the proposed strategy firstly fuzzified training data samples are generated and then decision tree is constructed. The continuation power flow is run and the training and target class (i.e., secure or insecure) are generated.
The scheme for classification prediction algorithm is adopted through the following steps.

Step1: Create the random load samples between 50% to 150% of base load of all 21 load buses.
Step2: Fuzzification of each input (Pi, Qi) samples.
Step3: Run CPF for each input samples.
Step4: Measure the voltage for each load node.
Step5: Construct input-output pairs such as (Input Features, Target Class)
Input Features: [Pi, Qi]
Target Class: [secure, insecure]
Step6: Construct Root node: Entropy computations for all samples for each credible outage cases.
Step7: Develop decision tree.
Step8: Based on adjustment the optimal decision tree is approached.
Step9: Test samples are examined in the testing phase.
Step10: Classification of samples into secure and insecure through the best decision tree.
Step11: End.

5. SIMULATION AND RESULTS

The IEEE 30 bus system is the test bench study system used in this paper. The training and test data set is obtained by CPF method at several operating conditions. Load variations between 50% to 150% of base case are considered. A set of credible five contingencies due to line outages are considered for data preparation. Each case is characterized by 42 attributes which shows the status of real and reactive power loads at 21 load buses of the IEEE 30 bus systems.

Each attribute is the fuzzified in the categories using trapezoidal membership function so that uncertainty may also be considered in the training as well as in the testing patterns. A data set of 250 was used at the training stage and 50 are used for testing set of 42 attributes.
Performance of model trees for IEEE-30 bus system is shown by Table 2. It shows the number of false alarms and missed cases. When secure case has been classified as insecure case it is termed as false alarm and when an insecure case has been classified as a secure case it is termed as missed case. False alarms do not harm any to power system operations, because they may be further calculated through a conventional algorithm for contingency evaluation.

Table 2: Performance of FDT Vs CART:

Line out no.	CART		FDT	
	% False alarms	% Missed cases	% False alarms	% Missed cases
1	4	2	8	8
5	6	8	0	10
36	8	8	6	4
2	4	10	0	12
4	4	8	0	8

Table 3 shows the comparison of percentage accuracy of FDT against CART method for the IEEE-30 bus system.

Table 3: Percentage accuracy:

Line out no.	Percentage accuracy	
	CART	FDT
1	94	84
5	86	90
36	84	90
2	86	88
4	88	92

The proposed fuzzy decision tree used with 250 samples of learning and 50 samples of testing patterns for the load bus attributes gives better percentage accuracy with respect to CART algorithm which is shown by Fig.2.

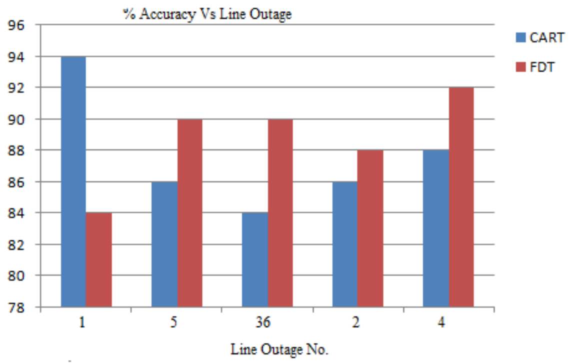

Fig. 2: Percentage Accuracy of classification.

6. CONCLUSION

In the paper a novel probabilistic decision tree based method considering the fuzzy inference in the data sets are presented for voltage security decision. Also the load variations considered was in wide range for all critical contingencies. However for online problem of voltage security assessment the proposed methodology should work healthy for all kinds of unexpected operating conditions. After voltage security assessment, the proposed FDT methodology shall be used for real-time optimal load shedding plan for insecure operating states of the power system to alleviate the voltage collapse during emergencies.
All the decision trees showed a good performance on all five critical line outage conditions on IEEE 30 bus system. Fuzzy logic further simplifies and also speed ups the complex algorithms. Hence the proposed method allows the power system operator to estimate the secure or insecure status for online voltage security operations and further decisions to be taken.

After security assessment, detecting the insecure cases is not enough for practical purposes; hence effective and practical solutions must be proposed in real-time to mitigate the voltage collapse is proposed for further improvement.

REFERENCES

[1] T. Cutsem, L. Wehenkel, M. Pavella, B. Heilbronn, and M.Goubin, Decision tree approaches to voltage security assessment, *Proc. Inst.Elect. Eng.*, vol. 140, no. 3, May 1993, 189–198,
[2] L. Wehenkel and M. Pavella, Decision Tree Approach to Power System Security Assessment, *International J. Electrical Power and Energy Systems,* Vol. 15, No. Feb. 1993,13-36.
[3] V. jjarapu and C. Christy, The continuation power flow: A tool for steady state voltage stability analysis, *IEEE Trans. Power Systems.,* vol. 7, no. 1, Feb. 1992, 416–423.
[4] T. V. Cutsem and C. Vournas, *Voltage Stability of Electric Power Systems*. Norwell, MA: Kluwer, 1998.
[5] A. J. Wood and B. F. Wollenberg, *Power Generation, Operation and Control*. New York: Wiley, 1996.
[6] N.P. Patidar, Jaydev Sharma**,** *Loadability margin estimation of power system using model trees,*Power India Conference, IEEE, 2006.
[7] S Kalyani and K Shanti Swarup Power System Security Assessment using Binary SVM Based Pattern Recognition *International Journal of Electrical and Electronics Engineering* 2009.
[8] Grillo, S. Static Security Assessment of Electrical Power Systems Using Neural Classification Techniques, *Mathematical Problems in Engineering, Aerospace and Sciences*, 2008.
[9] J. Preetha Roselyna, D. Devarajb, , Subhransu Sekhar Dasha, Multi-Objective Genetic Algorithm for voltage stability enhancement using rescheduling and FACTSdevices, *Ain Shams Engineering Journal Volume* 5, Issue 3,September 2014, , 789–801.
[10] L Breiman, J Friedman, RA Olshen, CJ Stone.,*Classification and Regression Trees,* Belmont:Wadsworth, 1984.
[11] S. Gey and E. Nedelec, Model selection for CART regression trees, *IEEE Trans. on Information Theory*, Vol. 51, No. 2, 658-670, Feb. 2005.
[12] R. J. Lewis, An Introduction to *Classification & Regression Tree (CART) Analysis,* Proc. Annual meeting of society for academic emergency medicine, San Francisco, California, USA, 2000,1-14.

[13] Cristina Olaru , Louis Wehenkel A complete fuzzy decision tree technique Elsevier journal of Fuzzy Sets and Systems vol. 138 ,2003, 221–254.

[14] Ruisheng Diao, EEE, Kai Sun, Vijay Vittal, Robert J.O'Keefe, Michael R. Richardson, Navin Bhatt, Dwayne, and Sanjoy K. Sarawgi, Decision Tree-Based Online Voltage Security Assessment Using PMU Measurements *IEEE transactions on power systems,* 2009, vol. 24, no. 2.

[15] Mohammad Abdul Mannan, Asif Islam, Mohammad Nasir Uddin, Mohammad KamrulHassan, Toshiaki Murata, Junji Tamura, Fuzzy- Logic Based Speed Control of Induction Motor Considering Core Loss into Account, *journal of Intelligent Control and Automation,* vol. 3,2012,.229-235.

Author Profile:

Sanjiv Kumar Jain is working as Assistant Professor in the department of Electrical Engineering, Medi-Caps University, Indore, India. He received M.E. degree in 2008 from SGSITS, Indore and B.E (Electrical Engg.) from RGPV in 2000. He is currently pursing PhD from M.A. National Institute of Technology, Bhopal India. His research interests include Power System control and Planning, Load Frequency Control.

Dr N. P. Patidar received his bachelor degree in Electrical Engineering from SGSITS Indore, India, in 1993 and the M. Tech. degree in integrated power system from Vishveshvaraya Regional college of Engg. Nagpur, India in 1995. Currently, he is Associate Professor in the Department of Electrical Engineering, M.A. National Institute of Technology Bhopal, and India. He is Ph.D. in power system from Indian Institute of Technology, Roorkee, India. His research interests are in the area of application of artificial intelligence for power system voltage security assessment.

Dr Yogendra kumar received B.Sc. Engg. in Electrical (1991) from Jamia Millia Islamia New Delhi, M.Tech in Heavy Electrical Eng. In 1998 from MACT, Bhopal, and PhD degree (2006) from IIT Roorkee in power system. Since 1992 he is with M.A. National Institute of Technology, Bhopal. Presently he is Professor in the Electrical Engineering Dept. His research area includes Power system optimization and distribution automation.