ANN/ DT Approach for Security Evaluation and Preventive Control of Power Systems

M.K.Shah
Department of Electrical Engineering
Malaviya National Institute of Technology
J.L.N.Marg, Jaipur, INDIA.

K.R.Niazi
Department of Electrical Engineering
Malaviya National Institute of Technology
J.L.N.Marg, Jaipur, INDIA.

C.M.Arora
Department of Electrical Engineering
Malaviya National Institute of Technology
J.L.N.Marg, Jaipur, INDIA.

Abstract *– Security evaluation is becoming a major concern in the real time operation of modern power systems which are large and complex. Neural Networks have shown great promise as a means of predicting security of large power systems. But because of their black box type nature, neural network based security classifiers are not able to provide information about preventive control if required following a contingency. The Decision Tree based classifiers, on the other hand, are known for their interpretability and therefore can be used to design preventive control strategy. However, the DT approach suffers from the drawback that as the size of power systems grows, the complexity of DT classifier becomes high. Building of complex DT classifier requires a prohibitively large learning set unless some compromise is made on the accuracy.*

This paper presents a hybrid approach for on-line security evaluation and preventive control of power systems, which combines ANN and DT approaches to exploit their potential while suppressing their drawbacks. It applies an ANN based classifier for security evaluation of power systems. If an operating state of power system is found to be insecure, a DT classifier is applied to infer preventive control measures. The accuracy of the DT classifier does not significantly affect the overall performance of this approach. The method has been applied on an IEEE power system and the results obtained are promising.

Keywords:* *Neural network, Decision Tree, Power system security, Feature selection, Preventive Control

I. INTRODUCTION

Security evaluation is a major concern in the real time operation of modern power systems. The present trend towards deregulation has forced modern electric utilities to operate their systems under stressed operating conditions closer to their security limits. Under such fragile conditions, any breach of security can have far reaching impact. Therefore, there is a pressing need to develop fast on-line security monitoring method which could analyze the level of security and suggests necessary preventive control measure to ensure system security, in case need arises.

A complete answer about power system security requires evaluation of transient stability of power systems following some plausible contingencies. Several methods for fast transient stability evaluation have been proposed in the past by adopting namely direct methods, pattern recognition (PR) technique, Decision Tree (DT) method and Artificial Neural Network (ANN) approach[1].

Neural Networks have shown great promise as a means of predicting security of large power systems [2,3]. But because of their black box type nature, neural network based security classifiers are not able to provide information about preventive control if required following a contingency. Though some efforts have been made to infer this information from the hidden layer of the neural network but under very simplified modeling assumption[4].

The Decision Tree based classifiers, on the other hand, are known for their interpretability and therefore can be used to design preventive control strategy. However, the DT approach suffers from the drawback that as the size of power systems grows, the complexity of DT classifier becomes high. Building of complex DT classifier requires a prohibitively large learning set unless some compromise is made on the accuracy [5].

This paper presents a hybrid approach for on-line security evaluation and preventive control of power systems, which combines ANN and DT approaches to exploit their potential while suppressing their drawbacks. It applies an ANN based classifier for security evaluation of power systems. If an operating state of power system is found to be in insecure operating state, a DT classifier is applied to infer preventive control measures. Two types of preventive control action could be inferred from the DT classifier namely generation re-dispatch and optimal load shedding. The choice of preventive control action would depend upon the feasibility and economic aspect of the preventive control solution provided by the DT classifier. The accuracy of the DT classifier does not significantly affect the overall performance of this approach. The complete algorithm of the proposed approach is outlined in section II. The ANN and DT methodology are briefly discussed in section III and IV respectively. The applicability of the method has been investigated on the IEEE 57 bus system and the test results are presented in section V. Finally conclusions are presented in section VI.

II. THE PROPOSED METHOD

The aim of the hybrid approach is to assess system security and provide operating guidelines for secure operation, in case there is violation of security constraint. The overall hybrid approach is as follows:

1. Read power system data and forecasted load data 20 minutes ahead.
2. Perform economic dispatch.
3. Apply ANN classifier to determine system security.
4. If the system is secure, dispatch generators as calculated in step 2 after 20 minutes and stop.
5. If the system is insecure apply DTC and infer the rescheduling of generators from the DTC to ensure security of power system.
6. Check the feasibility of rescheduling calculated in step 4 by performing load flow on the forecasted operating state of the power System.
7. If feasible solution exist, reschedule generation as computed in step 4 after 20 minutes and stop.
8. If feasible solution does not exist for the calculated rescheduling then evaluate optimal load shedding and carry out evaluated load shedding after 20 minutes and stop.

III. THE ANN METHODOLOGY

The complete description of ANN is not within the scope of this paper. Nevertheless, it is important to describe basic design procedure of an ANN classifier, which involves the following steps.

(1) Feature Selection.
(2) Data set generation for network training.
(3) Architecture of ANN model and its training algorithm.
(4) Performance evaluation.

Feature selection is a process of selecting a small subset of features from a large number of features (variables) that may characterize a given power system. It involves dimensionality reduction to identify most significant and useful subset of features that carries sufficient discriminating properties to perform the given classification task most accurately.

In the present method feature selection is carried out in two stages. In the first stage, an initial feature set is selected which is based on the knowledge of power system and objective of the problem to be solved. Initial feature set is a general set of features, which is independent of data set used for training. The idea behind the first stage of feature selection is to eliminate the insensitive features a priori so as to avoid exhaustive search of second stage of feature selection algorithm. The initial feature set should have the following properties [2,4]:

1. The feature set should adequately characterize an operating state of a power system from security point of view. At the same time it should be small enough to avoid unnecessary computation.
2. Features in the feature set should be independent.
3. Features should be monitorable and controllable so that control action may be exercised.
4. As far as possible, features should be independent of topology.

C.M. Arora et al [6,7] have derived that if power system is not optimally dispatched, the feature set consisting of pre-disturbance real and reactive power generation and real and reactive power demand at each system bus carry sufficient information about system security. Under certain justified assumptions, the generator currents and load currents can be expressed as a function of generator currents [8,9]. The generator currents are directly related to the real and reactive powers of the generators. Therefore, the real power and reactive power demands can be expressed as a function of real and reactive powers of generators. Thus, an attribute set (feature set) consisting of only real and reactive power generations is capable of providing sufficient discriminating information about the class of system security (secure or insecure). This fact is also supported by the outcome of research paper of C.A. Jensen et al. [2]. Therefore, the proposed initial feature set consists of pre-disturbance real and reactive power generation of each generator.

The second phase of feature selection needs special attention. One of the most important aspects of achieving good neural network performance is the proper selection of training feature (i.e. final feature subset). If too small feature subset is used it may render the ANN classifier useless on unseen patterns. On the other hand if feature subset is too large, the ANN may even fail to converge during training phase besides taking prohibitively large training time. Craig A. Jensen [2] has demonstrated that even for 50 generator IEEE power systems, only three features are sufficient to describe the operating state of the power system from ANN based security classification point of view. Therefore, as regard to ANN, feature selection should be a process to identify features, which contribute most to the discriminating ability of the ANN and discard the rest. In the literature on power system security, feature selection using Fisher values has been largely proposed. The problem with this method is that it works well with linearly separable classes because fisher discrimination algorithm basically seeks to find an optimal linear discriminate function for separating two classes [2,10]. How well this method will perform on non-linearly separable classes, is not established. Power system security evaluation is a complex non-linear problem, which may not have linear separability of classification. Therefore, the proposed method makes use of the concept of “divergence” for final stage of feature selection, which does not have this restriction. The final feature subset is specific to the power system and contingency set considered.

A. Feature Selection Using Divergence

Divergence is a measure of dissimilarly between two classes and therefore, it can be used in feature ranking and feature selection. The divergence J_{ij} between two classes, say i and j, can be expressed as [11]

$$J_{ij} = \frac{1}{2} t_r \left[\left(C_i - C_j\right)\left(C_j^{-1} - C_i^{-1}\right) \right] + \frac{1}{2} t_r \left[\left(C_i^{-1} + C_j^{-1}\right)\left(m_i - m_j\right)\left(m_i - m_j\right)^t \right]$$

Where,

t_r = trace of pattern matrix and is equal to sum of its Eigen values,

C_i = covariance matrix of class i of size [n×n],

C_j = covariance matrix of class j of size [n×n],

m_i = mean vector of class i of size [n × 1],

m_j = mean vector of class j of size [n × 1] and

n = number of features.

The proposed feature selection algorithm makes use of the divergence in backward sequential manner. The backward sequential search guarantees the optimal solution if the criteria function satisfies the monotonicity condition. The monotonicity condition requires that values of the criterion function be non-decreasing when additional features are added. The divergence satisfies the condition of monotonicity [11].The feature selection method proceeds as follows: Determine the divergence J_{ij} (n) of initial feature subset. Now selectively remove one feature at a time till all the features are considered and determine the divergences corresponding to all the n subset of (n-1) features. The feature that result in smallest decrease in the divergence, at each iteration, is then removed. This process is then repeated for remaining features.

The size of feature subset is an important ANN design consideration. So far, in the literature, the size of feature set has been chosen using heuristics or trial and error method. However to solve this problem in systematic manner it is proposed to define and use two terms namely "maximum permissible percentage change in divergence", ΔJ_{max} and "minimum number of features required by the classifier", n_{min}. The parameter, ΔJ_{max} is a measure of maximum permissible reduction in the discriminating properties of the feature subset, whereas the parameter n_{min} is the minimum size of feature subset required by the ANN classifier. Thus, at any intermediate stage of feature selection, if it is found that further reduction in dimensionality is causing a decrease in divergence more than ΔJ_{max}, then the process of feature selection is stopped

To generate a data set, initially a large numbers of load samples are randomly generated in the typical range of 50 to 150 percent of their base case values. For each load sample (load combination) optimal power flow (OPF) study is performed to obtain steady operating state. A disturbance (fault), from a predefined set of contingency, is simulated for a specified duration of time. Using dynamic stability studies, load angle trajectories of all generators is computed and plotted over a period long enough to ascertain system stability under the specified disturbance. Similarly for each of the disturbances from the contingency set dynamic simulation is performed to ascertain system stability under the corresponding disturbance. For carrying out dynamic simulation, numerical integration techniques is used as it has the flexibility to include all kind of modeling sophistication and thus is able to provide desired degree of accuracy. If a steady state operating point is found to be stable, for all disturbances of the contingency set, the operating state is assigned "secure (0)" class label else it is assigned "insecure (1)" class label.

The ANN model selected for on-line security evaluation is a multi-layer perceptron (MLP) as shown in Fig. 1. It consists of an output layer with one neuron specifying the security class. The number of inputs to the network is equal to the number of training features. The number of hidden layer is one or more. The network is trained using resilient error back propagation algorithm [10,12]. During the training the network performance is closely monitored to prevent network memorization. The trained network is tested for its performance on a test set of unseen patterns.

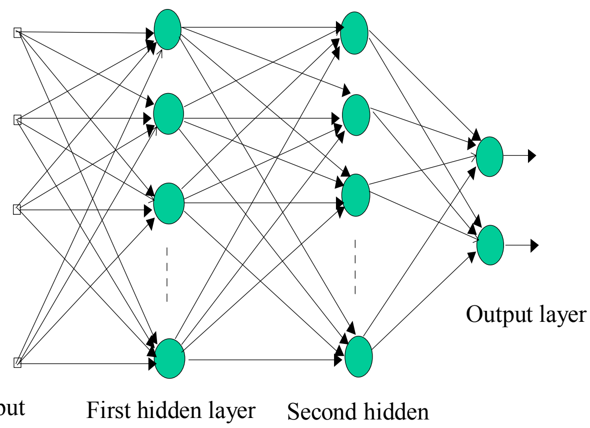

Fig. 1 Architecture of a MLP with two hidden layers

IV. THE DECISION TREE METHODOLOGY

The decision tree (DT) approach is one of the possible approaches to multistage decision making. It is a non-parametric learning technique, which generates a classifier in the form of multistage decision rules expressible in the form of a tree. The overall design of the DT classifier involves problem of tree structure design, feature selection and formulation of decision rules. A large number of Decision Tree Classifier (DTC) design algorithms are available in the literature

[13]. The present paper makes use of inductive inference (II) based method for DTC design. The complete description of this approach is given in [14]. The II method infers decision rules from a pre-analyzed learning set (LS) and expresses them in the form of a tree. The tree is structured in top-down fashion consisting of various test and terminal nodes. Each test node is associated with an optimal splitting rule and a subset of the LS. A built tree is, thus, a hierarchical organization of the LS into a collection of subsets. The most general subset is the LS itself and corresponds to the top (root) node of the DT.

Starting from the root node, at each level of the DT, the corresponding subsets are partitioned on the basis of some optimal splitting rules. These rules are in the form of "if- then- else" rules. The lower is the level, the more refined is the corresponding partition. Therefore, generating successors of a given non-terminal (test) node amounts to reducing the uncertainty about the classification. For given LS, DT is built according to the following procedure:

(1) Starting at the top node, splitting of a subset En requires a dichotomic test, defined as

$$T : a_I \leq v_{ik} ?$$

Where, a_I is the ith attribute (feature) and v_{ik} is the kth threshold value of the ith attribute. The possible threshold values of an attribute are obtained via a quantization procedure [5,14]. The test is applied to each attribute and each operating state of the LS with all possible threshold value so as to obtain the best split i.e. "optimal splitting rule".

(2) Using the optimal splitting rule, decompose the corresponding subset E_n into two successors subset namely left successor; which satisfy the rule and right successor; which does not satisfy the rules.

(3) Check whether the obtained successors are terminal node or test node. This testing requires a "stop splitting" rule.

(4) Repeat the above procedure at every non-terminal node.

The formulation of optimal and stop splitting rules makes use of entropy functions, information measures and statistical hypothesis test. Ideally a terminal node should be a class pure terminal node. In practice, however, a terminal node may contain statistically insignificant information about the goal classification. Such terminal nodes are called dead end and are assigned uncertain class label. The desired degree of purity and desired degree of statistical significance of the tree are governed by DTC design parameters such as H_{min}, called minimal residual entropy and α, called the risk factor of the optimal splitting rule. The values of DTC design parameters are chosen so as to obtain a near optimal DTC specific to the task.

The attribute (feature) selection is an important design consideration. Since the intended use of DTC is preventive control, its attribute set should exclude features, which are not controllable. In each power system there is at least one swing generator, which makes up the difference between the scheduled load and generated power that are caused by losses in the power system and so active and reactive power generations of swing generator are not controllable [9]. Therefore, attribute set consists of real and reactive power generations of generators except real and reactive power generation of swing generator.

By applying DT design algorithm on the training set a DT classifier is generated for the given set of contingency. The values of different DTC design parameters are chosen in a manner to obtain near optimal DT most suitable for preventive control purposes. The DT can provide two types of preventive control measures namely generation rescheduling and optimal load shedding, which are discussed below.

A. Generation Rescheduling Using DT

One of the important properties of the proposed DTC is that it provides security classification rules in the form of constraints on the system attributes i.e., active and / or reactive power generation of some or all generators. That is for secure operation

$$P_{Gi} < P^{s}_{Gi} \quad (1)$$

$$Q_{Gi} < Q^{s}_{Gi} \quad (2)$$

Where P^{s}_{Gi} and Q^{s}_{Gi} are the upper limits of active and reactive generation of ith generating unit, imposed by DTC for secure operation. Therefore, when the operating state of the power system is insecure, it violates some or all of the above constraints. To bring the system back into secure operating state, the generator are re-dispatched optimally in a manner to satisfy the security constraints (Eq. (1) and Eq. (2)).

The DT provides several re-dispatch alternatives. The choice of an appropriate re-dispatch alternative would depend upon the feasibility of the solution and nearness to secure leaf.

B. Optimal Load Shedding using DT

When rescheduling of generators fails to provide a feasible solution for secure operation, optimal load shedding is resorted. Let the system be operating in an insecure state with loads

S_{Li}^{o} , i = 1, 2, ……….. N. Where, N is the number of loads.

Let after applying load shedding system moves into a secure state with loads S_{Li}^{C} , i = 1, 2, ….. N. Thus, the values of loads after load shedding are constrained by the following inequalities.

$$S^{c}_{Li} \leq S^{o}_{Li} \quad (3)$$

Since $S_{Li} = P_{Li} + j\, Q_{Li}$ therefore,

$$P^{c}_{Li} \leq P^{o}_{Li} \quad (4)$$

$$Q_{Li}^{c} \leq Q_{Li}^{o} \quad (5)$$

where, i = 1, 2, ….. N.

For every operating point, system loads are also required to satisfy the following equalities

$$\Sigma P_{Li}^{c} = \Sigma P_{Gi} - P_{LOSS} \quad (6)$$

$$\Sigma Q_{Li}^{c} = \Sigma Q_{Gi} - Q_{LOSS} \quad (7)$$

With loads S_{Li}^{c}, the operating state of the power will be secure, if it satisfies the constraints imposed by the DT classifier

i.e. $P_{Gi} \leq P_{Gi}^{s}$ (8)

$Q_{Gi} \leq Q_{Gj}^{s}$ (9)

Where i and j may take one or more values between 1 to m depending upon the DT classifier.

Therefore, the optimal value of load after load shedding can be evaluated by solving the following optimization problem:

Minimizing objective function

$$U = \sum_{i=1}^{n} \left(S_{Li}^{o} - S_{Li}^{c} \right) \quad (10)$$

Subject to the constraints imposed by Eqs. 3 to 9

Some additional constraints may also be included to account for minimum generation limits, feasible values of loads etc. To determine the choice between preventive control alternatives, cost analysis of generation re-dispatch and load shedding could be carried out. Thus overall preventive control strategy brings an insecure operating state into a secure operating state with minimum loss of economy.

IV. SIMULATION AND RESULTS

To illustrate the applicability of the hybrid method a case study was performed on the IEEE – 57bus system. The system consists of 7 generators, 67 transmission lines, 18 transformers and 42 loads. It is assumed that contingency set contains only one disturbance, which is a 3-phase fault on 400kv line connecting bus 8 and 9 near bus 9. Duration of disturbance has been assumed to be 210 ms., which is cleared by opening the line from both the ends. The application of the divergence based feature selection algorithm on the training set of 500 operating state gave on optimal feature subset consists of only three features namely P_{G1}, Q_{G5} and Q_{G7}. Using the selected features and the training set, an MLP has been trained using resilient back prop algorithm. The network architecture used in the study consists of two hidden layers h1 and h2 and one output layer, 0.The output layer consists of one neuron, which specifies the security class. The test results of three training runs are shown in table 1. The test results highlight the effectiveness of the ANN classifier to accurately predict security of power system

By applying DTC design algorithm on the training set, 11 different DTCs have been generated for different combination of DTC design parameters. Their test results are shown in table 2. The test results show that there is a trade off between accuracy and confidence level on one side and number of dead end samples and complexity on the other sides. If a DTC is designed with a high accuracy and confidence level, the resulting classifier would be simple but with a large number of samples remain unclassified (fall in dead end). On the other hand if the classifier is designed with minimum number of samples in the dead end, the resulting classifier is complex with poor accuracy rate. Therefore, DT classifiers are not suitable for security evaluation purpose. However, the aim of DTCs in the proposed method is preventive control and so the essential requirements of DTC are :

- Minimum number of false dismissal (FD) on unseen operating condition i.e. accuracy.
- Simplicity of DTC, which is judged by the number of nodes in the classifier. The requirement emerges from the fact that a simple DTC provides a simple preventive control strategy with minimum number of constraints.

The number of dead end samples are not very important. The DTC is used only when the ANN classifier finds an operating state as insecure. Therefore, if an operating state falls in dead end terminal (uncertain class) it will be in insecure state. So a dead end terminal can be assigned insecure class label. Keeping in mind the above-mentioned requirements, the most suitable DTC for preventive control action purposes is the one shown in Fig. 2. By assigning insecure class to the dead end terminal nodes, this DTC can be redrawn as shown in Fig 3.

The classifier is simple with a false dismissal rate of 0.01%. It gives two re-dispatch alternatives to ensure system security. To test the effectiveness of the DTC, the generators were re-dispatching for all the 492 insecure operating states of the data set as obtained by ANN classifier. It has been found that after rescheduling of generators all the insecure operating state moved into secure operating states except one. In the present case cost analysis has not been carried. Load shedding option is needed if it is cost effective or if the generation re-scheduling fails to provide feasible solution.

Table 1

S.No	Network Architecture used or h1- h2-0	% Error on training set (500)	% Error on test set (500)
1	20-10-1	1.2	1.2
2	20-10-1	1	1.6
3	20-10-1	1	1.2
	Average	**1.06**	**1.33**

Table 2

S.No	DTC Parameters			DTC Performance					
	$(1-\alpha)\times 100$	HMIN	SCMIN	Nodes	Dead End Samples	Total Error	Error on LS	Error on TS	% of False dismissals
1.	99.99	0.045	0.14	14	592	1	0	1	0.01
2.	99.90	0.045	0.14	18	563	2	1	1	0.02
3.	99.00	0.045	0.14	18	563	2	1	1	0.02
4.	95.00	0.045	0.14	22	514	4	1	3	0.03
5.	90.00	0.045	0.14	32	490	6	2	4	0.03
6.	99.99	0.015	0.10	26	464	3	0	5	0.03
7.	99.90	0.015	0.10	24	394	3	0	5	0.02
8.	99.00	0.015	0.10	28	231	1	0	1	0.01
9.	99.99	0.045	0.10	22	106	4	1	5	0.03
10.	90.00	0.045	0.10	42	0	9	2	7	0.05
11.	90.00	0.015	0.10	46	0	7	0	7	0.04

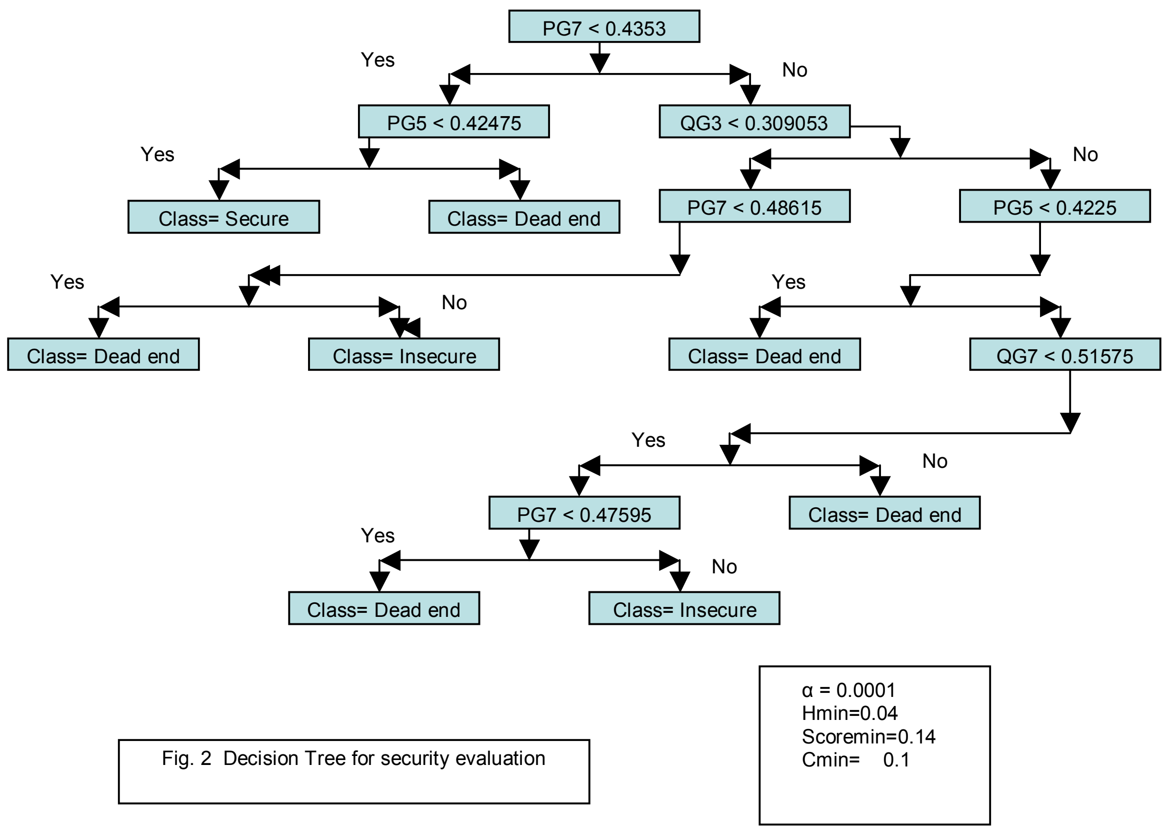

Fig. 2 Decision Tree for security evaluation

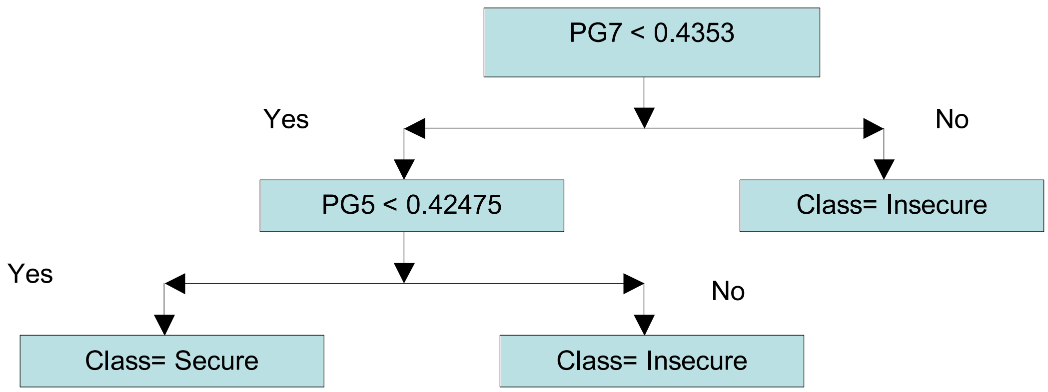

Fig. 3 Decision Tree for Preventive Control

VI. REFERENCES

[1] V. Miranda, J.N. Fidalgo, J.A.P. Lopes, "Real time preventive actions for transient stability enhancement with a hybrid neural network-Optimization approach", IEEE Trans. on Power Systems, Vol. 10, No.2, May 1995, pp. 1029-1034.

[2] C.A. Jensen, M.A. El-Sharkawi, R.J. Marks II, "Power systems security assessment using neural networks: Feature selection using Fisher discrimination", *IEEE Trans. on Power Systems*, Vol. 16, No. 4, Nov., 2001, pp. 757-763.

[3] K. Santi Swarup, P. Britto Corthis, "ANN approach assesses system security", *IEEE Computer Applications in Power*, July 2002 pp. 31 –38.

[4] J.N. Fidalgo, V. Miranda, J.A. P. Lapes, "Neural networks applied to preventive control measures for the dynamic security of isolated power systems with renewables", IEEE Trans. on Power Systems Vol. 11, No. 4, Nov., 1996, pp. 1811-1816.

[5] L. Wehenkal, Th. Van Cutsem, M. Ribbens-Pavella, "Inductive inference applied to on-line transient stability assessment of electric power systems", Automatica, Vol. 25, No. 3, 1989, pp. 445-451.

[6] C.M. Arora, "On-line transient security evaluation using pattern recognition technique", Ph.D. Thesis, Deptt. of Elect. Engg., JNV Univ., Jodhpur, India, 1991.

[7] C. M. Arora, S. L. Surana, "Transient security evaluation and preventive control of power systems using PR techniques", IE(I) Journal-EL, Vol. 76, Feb/March 1996, pp199-203.

[8] A.R. Burgen, "Power System Analysis", Prentice Hall, 1986.

[9] I.J. Nagrath , D.P. Kothari, "Modern power system analysis", Tata McGraw Hill, 1994.

[10] Simon Haykin, "Neural networks– A comprehensive foundation", Pearson Education Asia, 2001.

[11] J.T. Tau, R.C. Ganzalez, "Pattern recognition principles", Addison-Wesley, 1974.

[12] H. Demuth, M. Beale, "Neural network tool box for use with MATLAB", User's Guide, Version 4, The Math works Inc, 2001.

[13] S.R. Safavin, D. Landgrebe, "A survey of decision tree classifier methodology", IEEE Trans. on SMC, Vol. 21, No. 3 May/ June 1991, pp. 660-673.

[14] L. Wehenkal, Th. Van Custsem, M. Ribbens-Pavella, "An artificial intelligence frame work for on-line transient stability assessment of power systems", IEEE Trans. on Power Systems, Vol. 4, No.2,May, 1989, pp. 789-800.