Power swing prediction for out-of-step mitigation

V. Siyoi[1], S. Kariuki[2], M. J. Saulo[2]

[1]Department of Electrical Engineering, Pan African University of Basic Science and Technology, Nairobi, Kenya
[2]Department of Electrical Engineering, Technical University of Mombasa, Mombasa, Kenya

Email address:

v.siyoi@gmail.com (V. Siyoi), kariukisamuel2004@yahoo.com (S. Kariuki), michaelsaulo@yahoo.com (M. J. Saulo)

Abstract: This paper explored the possibility of accurately predicting the classification of developing power swings. The notion of machine learning was employed, and tested the application of Decision Tree (DT) algorithms to wide area power system protection schemes. The novelty of the designed Wide Area Protection (WAP) scheme was portrayed by the WAP's ability to adaptively and accurately predict the classification of developing successive power swings. DTs being a Data Mining (DM) technique, a transient stability analysis was performed on an IEEE 39 bus test system in Dig SILENT®. The learning sample from the Phasor Measurement Unit (PMU) data was organized and stored in a data base in Microsoft Excel® 2010. The CART analysis and DT model design was done using Salford Predictive Modeller-CART® v6, trial licence. The results of this investigation were quite accurate and gave DT algorithms 'thumbs-up' in terms of classification prediction.

Keywords: Decision Trees, Power Swing, Out-of-Step, Wide Area Protection

1. Introduction

Despite the profound success of various automated industrial processes, automation capabilities were not superior enough to match up to power system dynamism and the rate at which power system changes occur. This was because power system transients, faults, power swings and other power system abnormalities develop within milliseconds, a time too fast for autonomous systems to detect and to respond to. The immediate discussion presents a non-conventional method of designing a WAP scheme that enhances the stability of a power system.

2. Decision Trees

The DT technique using the Classification and Regression Trees (CART) is employed to perform the prediction of a power swing classification. As developed in this work, DT algorithms have been used to predict power swings which are also discussed in references [4], [5], [14], [21], [22], [23], [24], [25], [26], [27], [28], [29], [30], [31].

The CART algorithm is recommended for developing DT models, the most significant traits being simplicity and speedy execution of the models. Complex hidden information is classified and simplified into binary 'yes/no' recursive statements.

The major limitation to employing DTs is that there is only a single pair of a binary output which infers the classification problems as a binary output; as either 'yes/no' answers. DTs are also unstable; a small change in the input learning sample may give a completely different decision model. The DT using the CART technique was developed as follows:

(i) The learning sample L was arranged as an $m \times n$ matrix..

(ii) Attributes were sorted in order to initialize the splitting points that maximized the splitting criterion.

(a) From the set of attributes $A = \{a_1, a_2, ..., a_n\}$ in the learning sample L, an attribute $a \in A$ was selected. If a was numeric, the splitting was as equation (1)

$$S_a(k) = \frac{x_a(k+1) - x_a(k)}{2} \qquad (1)$$

(b) If a was defined as a categorical variable of sets $S_a = \{s_1, s_2, ..., s_n\}$, then the possible splitting point was within the range of available sets of that particular attribute.

(iii) The impurity reduction level was computed from the Gini improvement function as represented in equation (2).

$$i(t) = \pi(C_j)P^2(C_j \mid t)$$

$$i(t) = 1 - \sum_j^J P^2(C_j \mid t)$$ (2)

$$\Delta i(s,t) = i(t) - \frac{n(t_L)}{n(t)}i(t_L) + \frac{n(t_R)}{n(t)}i(t_R)$$

$$\Delta i(s,t) = i(t) - [P_L i(t_L) + P_R i(t_R)]$$

(iv) A variable ranking of all attributes was performed. The measure of importance of a variable x in relation to the final tree T is the weighted sum across all splits in the tree of improvements that x has when it is used as a surrogate as shown in equation (3).

$$M(x) = \sum_{t \in T} \max_{\Delta IC_i(\tilde{S}_x,t)} \Delta I(\tilde{S}_x,t)$$ (3)

The variable importance $VI(x)$ was expressed in terms of a normalised quantity relative to the variable having the largest measure of importance, shown in equation (4).

$$VI(x) = \frac{M(x)}{M(x_{max})} \times 100$$ (4)

(v) Using the Gini purity index, the root node was identified.

(vi) On the root node of the DT, the splitting points for the resulting child nodes were located. The splitting point of the root node was determined from amongst the set of all possible splitting points of all the attribute/variables. For each splitting value $s \in S_a$ at a particular node t, the learning sample was partitioned into separate subsets t_L and t_R forming the left and right child nodes respectively.

(a) For numerical variables, then the partitioning is as shown in equation (5).

$$tL = \{x_a(k) \, if \, x_a(k) \leq s_k\}$$
$$tR = \{x_a(k) \, if \, x_a(k) > s_k\}$$ (5)

(b) For categorical variables, (have finite sets) then the partitioning is as shown in equation (6).

$$tL = \{x_a(k) \, if \, x_a(k) = s_k\}$$
$$tR = \{x_a(k) \, if \, x_a(k) \neq s_k\}$$ (6)

(vii) Optimal split $s_{optimal}$ over all possible splitting values $s \in S_a$ amongst all attributes $a \in A$ was found. Gini splitting points were computed as shown in equation (7).

$$GINI_{Split}(S) = \sum_{i \in values(s)} \frac{|n_t|}{|n|}i(t)$$ (7)

$$GINI_{Split}(S) = \frac{n_i}{n}i(t_L) + \frac{n_j}{n}i(t_R)$$ (8)

(viii) A classification decision was made from terminal nodes. A node was classified in class i if equation (9) was satisfied.

$$\frac{C(j|i)\pi(i)N_i(t)}{C(i|j)\pi(j)N_j(t)} > \frac{N_i}{N_j} \text{ for all values of } j$$ (9)

(ix) Each of the remaining predictor's best split points were defined using the Gini split criterion. The next splitting point of the subsequent node that maximizes the splitting criterion was selected and steps (viii) through (ix) were repeated.

(x) If the stopping rules had not been satisfied, steps (viii) through (x) were repeated, otherwise process stopped.

To avoid unnecessary redundancy, optimization through pruning the decision model is performed. This is by removing tree branches whose cost complexities (penalty associated with misclassification of cases) reduce the reliability of the tree. For a maximal sized tree, the cost complexity $\alpha = 0$. Pruning therefore evaluates tree branches as shown in equation (10) where each subsequent branch removal $R_\alpha, T_1 > T_2 >,....,> T_t$ increases the cost complexity thus optimizing the DT.

$$R_\alpha = R(T) + \alpha \cdot L$$ (10)

Where α is the complexity function, $R(T)$ is the re-substitution error and L is the number of branch nodes.

Validation of the DT model was done through a v-fold cross-validation. Specifically, a 10-fold cross-validation was performed as follows: Let T be a tree grown using all data from the whole data set $\hbar^0$ and let $v \geq 2$ be a positive integer.

(i) Divide $\hbar^0$ into V mutually exclusive subsets $\hbar'_v$ where $v = 1,2,...,v$. Let $\hbar_v = \hbar^0 - \hbar'_v$.

(ii) For each v, consider $\hbar_v$ as a learning sample and grow a tree T_v on $\hbar_v$.

(iii) Assign $j_v^*(t)$ or $\overline{y}_v(t)$ for a node t of T_v.

(iv) Consider $\hbar'_v$ as a test sample and calculate its test sample risk estimate $R^{ts}(T_v)$.

(v) Repeat step (iv) for each value of v. The average of the test samples is used as the v-fold cross validation risk estimate of T.

The v-fold cross-validation estimate, $R^{cv}(T)$ of the risk of the tree T and its variance are estimated by equation (13) as developed by references [32], [33], [34] and [35].

$$R^{cv}(T) = \begin{cases} \sum_j \pi(j) \dfrac{1}{N^0_{f,j}} \sum_v N'_{v,f,j} R^{ts}(T_v|j) & \text{Y cat, M2} \\ \dfrac{1}{N^0_f} \sum N'_{v,f} R^{ts}(T_v) & \text{Y cont, or Y cat and M1} \end{cases} \tag{11}$$

$$\text{var}\left(R^{cv}(T)\right) = \left\{ \dfrac{1}{\left(N^0_f\right)^2} \left\{ \sum_v \sum_{t \in T'_v} \sum_{n \in \hbar'_v(t)} N'_{v,f,j}(t) C\left(j^*_v(t)|j\right)^2 - N^0_f R^{cv}(T) \right\}^2 \quad \text{Y cat, M1} \right\} \tag{12}$$

$$\text{var}\left(R^{cv}(T)\right) = \begin{cases} \sum_j \left(\dfrac{\pi(j)}{N^0_{f,j}}\right)^2 \left[\sum_v \sum_{t \in T'_v} N'_{v,f,j}(t) C\left(j^*_v(t)|j\right)^2 - \left\{ \dfrac{\sum N'_{v,f,j} R^{ts}(T_v|Y=j)}{N^0_{f,j}} \right\}^2 \right] \sum_v N'_{v,f,j} R^{ts}(T_v|j) & \text{Y cont, M1} \\ & \text{Y cont} \\ \dfrac{1}{\left(N^0_f\right)^2} \left\{ \sum_v \sum_{t \in T'_v} \sum_{n \in \hbar'_v(t)} f_n\left(y_n - \bar{y}_v(t)\right)^4 - N^0_f R^{cv}(T)^2 \right\} \end{cases} \tag{13}$$

The ROC curve represents the ability of the DT model to accurately discriminate between the stable power swings and the unstable power swings. Let x be the scale of test result variable; low values suggest a negative x_- result while a high value suggests a positive x_+ result. Area under the ROC curve is calculated as equation (14).

$$\theta = P_r\left(x_+ > x_-\right) \tag{14}$$

Successive points in the ROC curve are connected by the trapezoidal rule as expressed in equation (15).

$$W = \dfrac{1}{n_+ n_-} \sum_{x \in \{\text{all values}\}} \begin{cases} n_{-=j} \times n_{+>j} \\ \dfrac{n_{-=j} \times n_{+>j}}{2} \end{cases} \tag{15}$$

3. Results & Analysis

The aim of the transient stability simulation was to induce power disturbances/swings at the critical load centres and at the extra-high voltage lines to create generator-load imbalance. The simulation was done considering all possible power system states, until the power system was observed to be transiently unstable in each of the various states. Graphical representations of the response of the coherent generators due various contingencies induced during the simulation are shown in figure 1, figure 2 and figure 3. The figure 1 shows a successive OOS response of each of the generators after a single contingency simulation. The normal operating conditions for the transient stability study was set such that:-

(i) The voltage should be within 0.95-1.05 p.u.
(ii) The load phasor voltage angle should not advance the generator phasor voltage angle by exactly 4 pole slips.
(iii) The frequency deviation from the nominal frequency of the reference machine should not be greater than ± 4%.
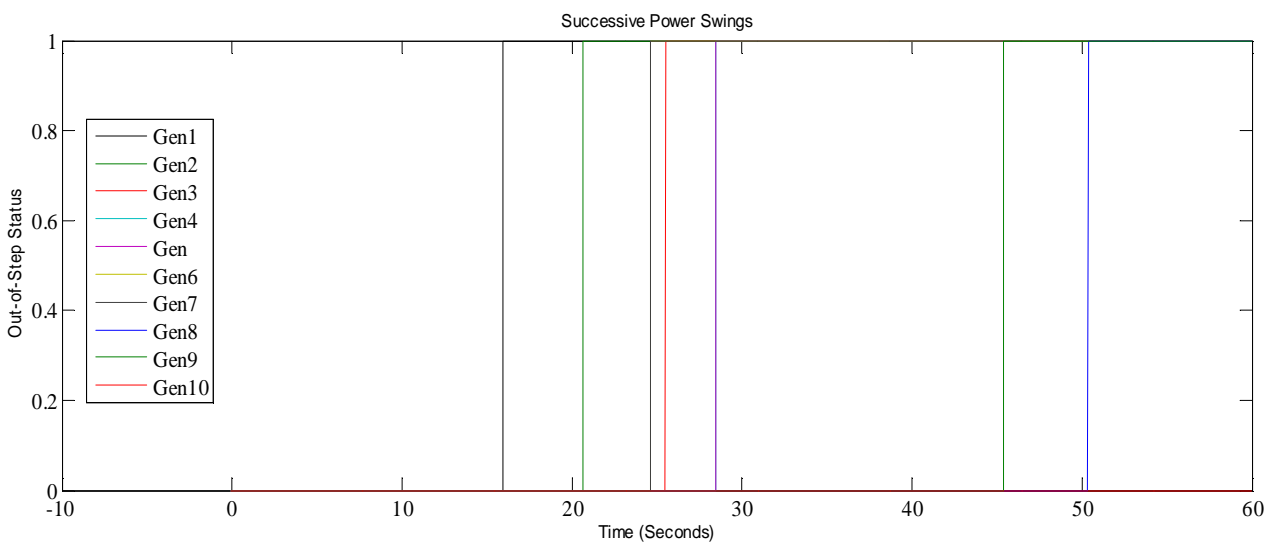
Figure 1. Simulation Responses of Successive Swings

Rotor Angle Slip (Degrees) w.r.t Reference Machine
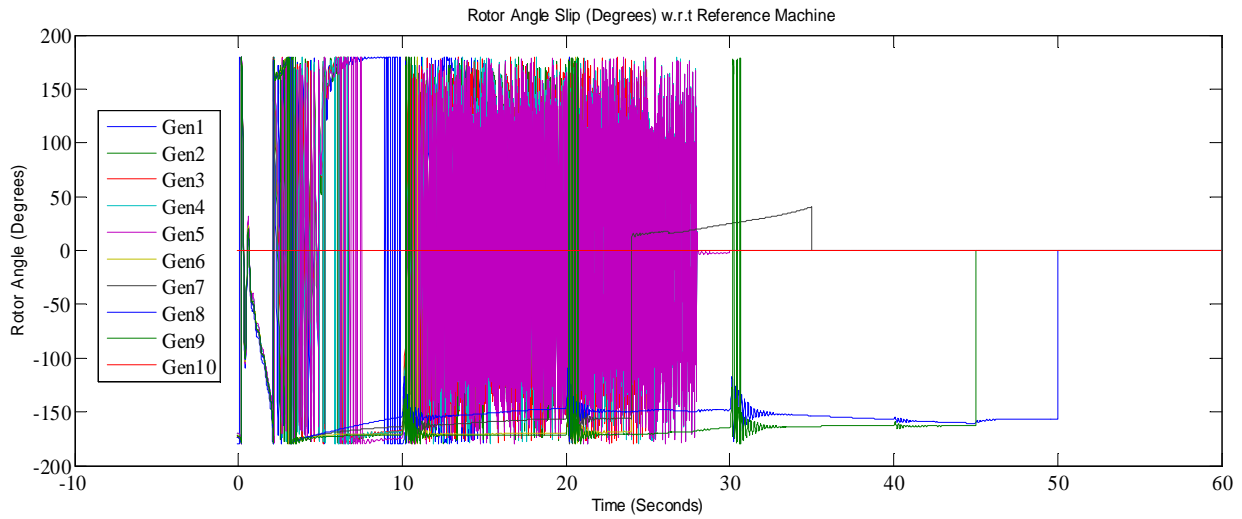
Figure 2. *Rotor Angle Slip from Reference Machine*

Generator Speed Deviation (HZ)
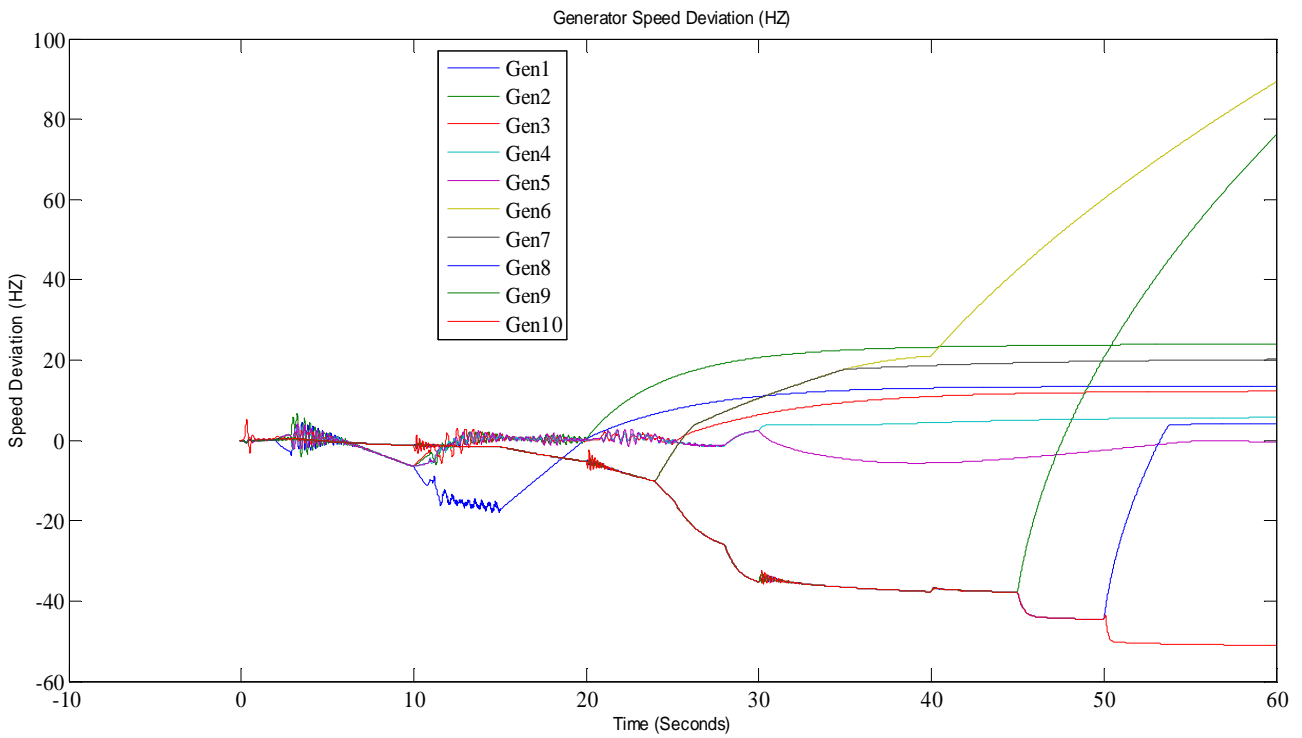
Figure 3. *Generator Speed Deviations*
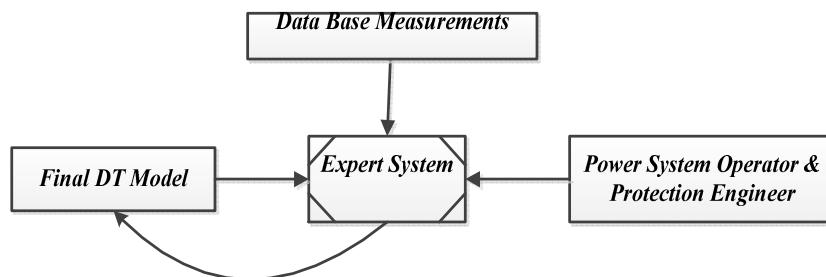
Figure 4. *Expert System DT Models*

The figure 2 shows the response of each of the generator's rotor position. A pole slip at the onset of the fourth pole slip reflects an oscillating response on the graph figure 2. The figure 3 shows the speed response due contingencies simulated. The response curve shows the speed deviation of the generators due loss of synchronism and therefore deviate

from their normal synchronism speed.

```
                    ┌────────────────────────┐
                    │    GENERATE DATABASE    │
                    └────────────────────────┘
                                │
                    ┌────────────────────────┐
                    │ EXTRACT LEARNING SAMPLE │
                    │         FROM DB         │
                    └────────────────────────┘
                                │
                    ┌────────────────────────┐
                    │   DT MODEL FROM ADAPTIVE │
                    │        OOS RELAY         │
                    └────────────────────────┘
                                │
                         UNKNOWN/UNTRAINED
                              CASE
```

GENERATE DATABASE

EXTRACT LEARNING SAMPLE FROM DB

DT MODEL FROM ADAPTIVE OOS RELAY

UNKNOWN/UNTRAINED CASE

No   Yes

GENERATE NEW CASE

DID DT MODEL FROM ADAPTIVE OOS RELAY DT CLASSIFY NEW CASE WELL?

Yes   No

IS DT MODEL FROM ADAPTIVE OOS RELAY POWERED ?

Yes

No

EXTRACT DT FROM ES (figure 4

DID A NEW SECOND SUCCESSIVE CASE EMERGE IN<0.1SECONDS?

Yes

No

DID ES-DT CLASSIFY NEW CASE WELL?

Yes   No

FINAL DECISION TREE

PMU ONLINE APPLICATION

UPDATE DECISION TREE

COMPARE WITH DECISION TREE THRESHOLD RULES

IS POWER SWING STABLE?

No   Yes

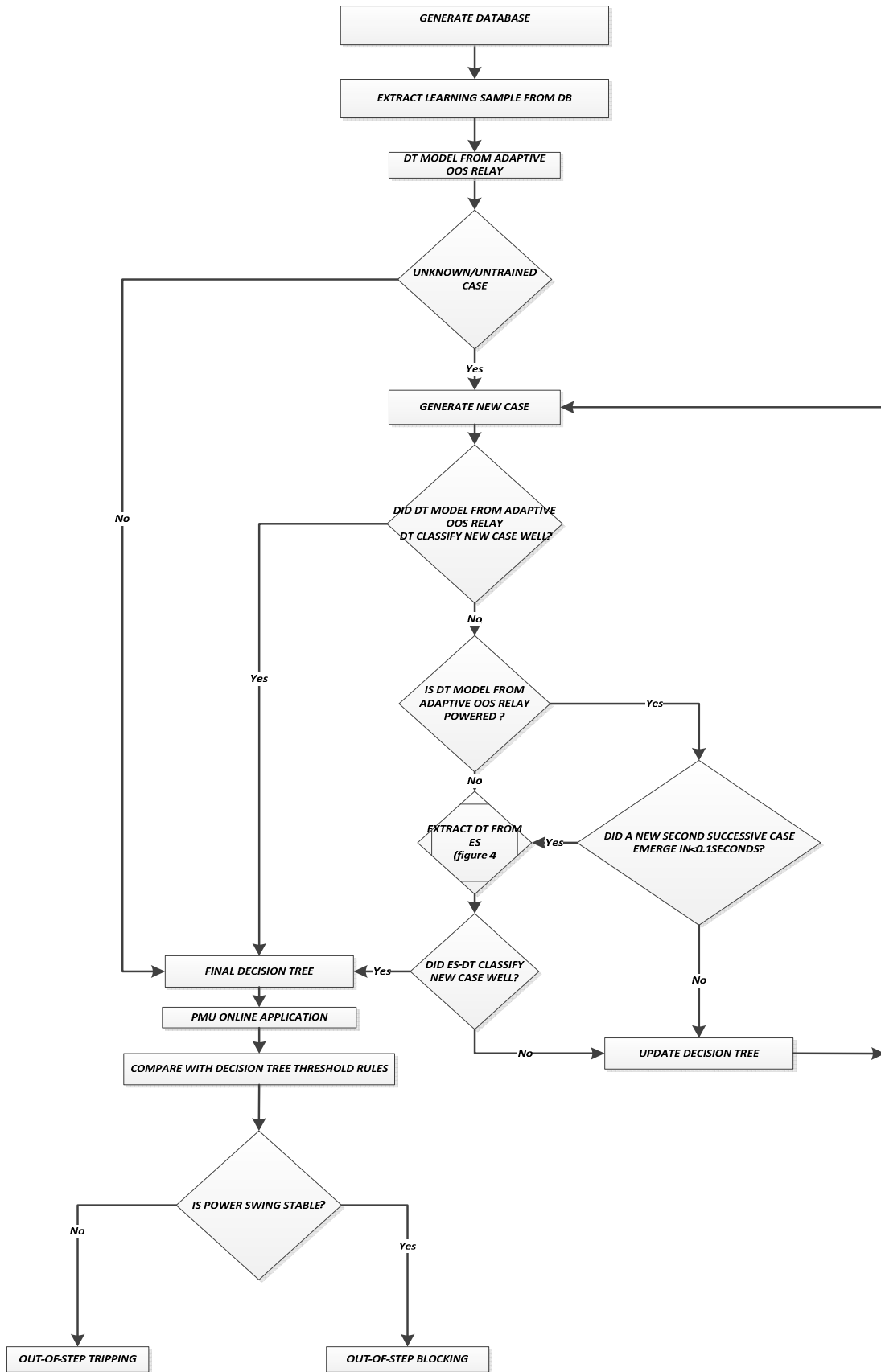**OUT-OF-STEP TRIPPING**

OUT-OF-STEP BLOCKING

Figure 5. *DT Model Execution Process Flow*

The process of executing the designed DT model involved a procedure proposed by this paper illustrated in

figure **5**. An Expert System (ES) was chosen as the secondary engine for executing the DT model for the following reasons:

(i)   The proposed ES as shown in figure 4 has the ability to learn from a wider base of experience than conventional decision support systems.

(ii) Ability to respond quickly and successfully to new situations.

(iii) Utilizes reasoning to solve problems at perplexing situations.

(iv) Recognizes the relative importance of different elements in a situation.

(v)   Ease of duplication of decisions and dissemination of the same [1], [36].

The ES manipulates DT models from three different sources, all of which are stored within the memory of the ES. The sources of these DT models are:

(i)   Developed by the ES independently from the main population database of measurements.

(ii) Knowledge induced to the ES by the control centre operator and protection engineer.

(iii) A replica copy of the final DT model developed by the main Intelligent Decision Support System IDSS (adaptive OOS digital-relay).

The management and timing functions are important when successive swings develop. If an instantaneous swing or successive swings develop within a duration of >0.1 seconds, then the DT model in the IDSS is given first priority to execute. If the swings develop within a duration of <0.1 seconds or when the DT model from the main IDSS fails, then the DT model from the ES is executed. The main IDSSmay fail if its window cycles are not complete amongst other time factors. Both the IDSS and ES models are updated to learn of new cases. The chosen DT model to execute compares its decision rules with that of an online PMU to initiate Out-of-Step Trip (OST) or Out-of-Step Block (OSB) functions.

The DT model therefore gives an insight on relay algorithms in mitigating various power system faults without over depending on impedance transfer methods. The hypothesis thus tested was that unlike conventional distance relays which use impedance tracking, WAP schemes can use selected important variables for OOS detection. For real time applications, these important variables are the only parameters updated to keep the model attuned to prevailing power system conditions. Updating only these selected variables reduces the digital relay execution time and is thus able to perform with speed.

The implementation of these DT models is achieved through a top-down induction of the DT rules. The DT rules from the optimum DT (figure 8) for predicting power swings are shown in figure 6.

```
/*Terminal Node 10*/
if
(
GEN_ROT_ANGLE_WRT_MACHINE_ANG__DEG_ > -43.9803 &&
L16_24_I1QIN_KA> -0.0643745 &&
GEN_SPEED_DEVIATION_HZ_ > -29.9187 &&
L11_35_I1Q_IN_KA <= 0.168518
)
{
terminalNode = 10;
class = Unstable_Power_Swing;
}
```

```
/*Terminal Node 1*/
if
(
GEN_ROT_ANGLE_WRT_MACHINE_ANG__DEG_ <= -43.9803 &&
L23_24_VOLT_ANG_IN_DEG <= 126.119 &&
L22_23_I1P_IN_KA <= 0.885491 &&
L23_24_I1P_IN_KA <= 0.529068 &&
L11_35_I1P_IN_KA <= 0.246655
)
{
terminalNode = 1;
class = Stable_Power_Swing;
}
```

Figure 6. DT model representing rules for Predicting Power Swings

The variable ranking of individual attributes in predicting a

TABLE 1. The reliability index of the performance of the DT model in making an accurate decision to a predicted power swing was given in terms of the relative cost. Figure 7 shows the quantitative graph representing the relative cost.

power swing is shown in

The relative cost is the penalty assigned (as a numeric quantity) due to wrong classification made by the DT model. The relative cost as observed is quite low implying that the DT model generally made the right decisions.
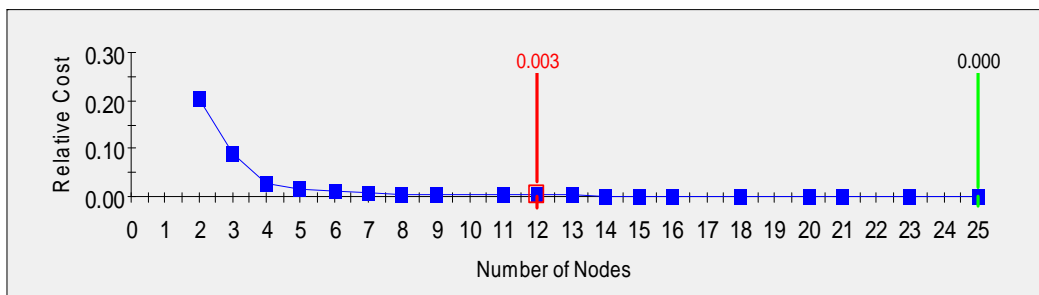

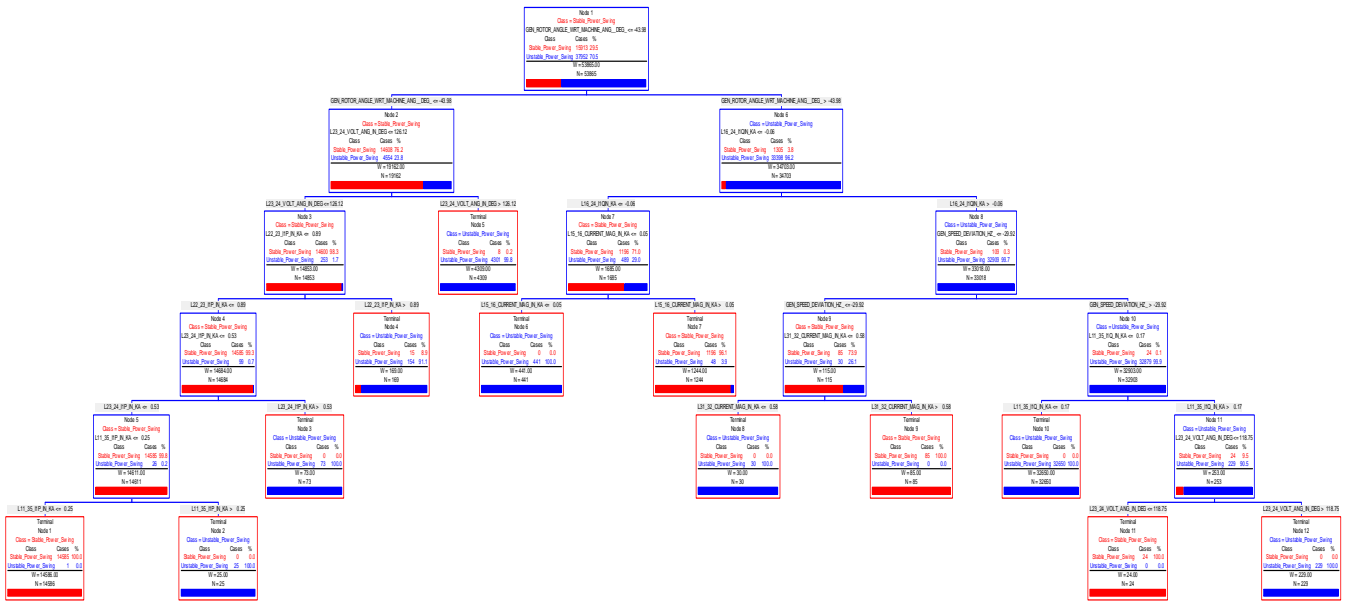*Figure 7.* Optimal Tree's Relative Cost Performance

Figure 8. *Optimal Decision Tree Model*

The area under ROC curve strengthens the validity of the designed DT model. The area under the ROC curve evaluates the accuracy of discrimination between two decisions. As the area value tends towards 1, then the more accurate the choice of decision made by the DT model. Performance of the DT model as valued by the area under ROC curve is represented in TABLE 4.

The response statistics of each of the terminal node of the optimum DT model are shown in TABLE 2. The overall test performance of 99.82% as shown in TABLE 3 was quite accurate and therefore suggested a reliable DT model.

Table 1. *Variable Ranking*

Variable	Percentage score
GEN_ROTOR_ANGLE_WRT_MACHINE_ANG__DEG	100.00
GEN_I1P_KA	79.95
GEN_SPEED_DEVIATION_HZ	75.66
GEN_ACTIVE_PWR_MW	73.51
GEN_CURRENT_MAG_KA	73.44
GEN_ELECTRICAL_TORQUE_IN_P_U	72.35
L23_24_VOLT_ANG_IN_DEG	31.41
L38_39_VOLT_ANG_IN_DEG	24.57

Table 2. *Response Statistics Of Optimal Tree's Terminal Nodes*

Node	Cases	Percent Score Data	Percent Train Data	Node Class	Percent Correct	Train Pct.Stable Power Swing	Train Pct.Unstable Power Swing	Score Pct.Stable Power Swing	Score Pct.Unstable Power Swing
1	14586	24.18	27.08	Stable Swing	99.99	99.99	0.01	99.99	0.01
2	25	0.04	0.05	Unstable Swing	100.00	0.00	100.00	0.00	100.00
3	73	0.12	0.14	Unstable Swing	100.00	0.00	100.00	0.00	100.00
4	169	0.28	0.31	Unstable Swing	91.12	8.88	91.12	8.88	91.12
5	4309	7.14	8.00	Unstable Swing	99.81	0.19	99.81	0.19	99.81
6	441	0.73	0.82	Unstable Swing	100.00	0.00	100.00	0.00	100.00
7	3637	6.03	2.31	Stable Swing	98.68	96.14	3.86	98.68	1.32
8	996	1.65	0.06	Unstable Swing	3.01	0.00	100.00	96.99	3.01
9	2286	3.79	0.16	Stable Swing	100.00	100.00	0.00	100.00	0.00
10	33492	55.51	60.61	Unstable Swing	98.78	0.00	100.00	1.22	98.78
11	87	0.14	0.04	Stable Swing	100.00	100.00	0.00	100.00	0.00
12	229	0.38	0.43	Unstable Swing	100.00	0.00	100.00	0.00	100.00

Table 3. *Test Prediction Success*

Actual Class	Total Class	Percent Correct	Unstable Swing N=37899	Stable Swing N=15966
Unstable_Power_Swing	37952	99.80	37877	75
Stable_Power_Swing	15913	99.86	22	15891
Total:	53865.00			
Average:		99.83		
Overall % Correct:		99.82		

Table 4. *ROC &Error Profiles*

No. of Nodes	5-foldRel. Error	10-fold Rel. Error	20-fold Rel. Error	Average Rel. Error	Min Rel. Error	Max Rel. Error	5-fold ROC	10-fold ROC	20-fold ROC	Average ROC	Min ROC	Max ROC
2	0.2023	0.2024	0.2026	0.2025	0.2023	0.2026	0.8988	0.8988	0.8987	0.8988	0.8987	0.8988
3	0.0895	0.0893	0.0896	0.0894	0.0893	0.0896	0.9597	0.9598	0.9597	0.9597	0.9597	0.9598
4	0.0269	0.0269	0.0269	0.0269	0.0269	0.0269	0.9924	0.9924	0.9924	0.9924	0.9924	0.9924
5	0.0153	0.0153	0.0153	0.0153	0.0153	0.0153	0.9929	0.9929	0.9929	0.9929	0.9929	0.9929
6	0.0108	0.0108	0.0108	0.0108	0.0108	0.0108	0.9955	0.9954	0.9954	0.9954	0.9954	0.9955
7	0.0077	0.0077	0.0077	0.0077	0.0077	0.0077	0.9976	0.9976	0.9976	0.9976	0.9976	0.9976
8	0.0058	0.0058	0.0058	0.0058	0.0058	0.0058	0.9987	0.9987	0.9987	0.9987	0.9987	0.9987
9	0.0054	0.0055	0.0055	0.0055	0.0054	0.0055	0.9990	0.9987	0.9987	0.9988	0.9987	0.9990
11	0.0044	0.0040	0.0040	0.0041	0.0040	0.0044	0.9990	0.9992	0.9992	0.9992	0.9990	0.9992
12	0.0037	0.0034	0.0034	0.0035	0.0034	0.0037	0.9994	0.9997	0.9997	0.9996	0.9994	0.9997
13	0.0030	0.0027	0.0030	0.0029	0.0027	0.0030	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998
14	0.0023	0.0018	0.0018	0.0020	0.0018	0.0023	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998
15	0.0019	0.0017	0.0018	0.0018	0.0017	0.0019	0.9997	0.9998	0.9998	0.9998	0.9997	0.9998
16	0.0015	0.0015	0.0015	0.0015	0.0015	0.0015	0.9996	0.9998	0.9997	0.9997	0.9996	0.9998
18	0.0012	0.0010	0.0011	0.0011	0.0010	0.0012	0.9997	0.9998	0.9997	0.9997	0.9997	0.9998
20	0.0007	0.0005	0.0006	0.0006	0.0005	0.0007	0.9997	0.9998	0.9997	0.9997	0.9997	0.9998
21	0.0007	0.0005	0.0006	0.0006	0.0005	0.0007	0.9997	0.9998	0.9997	0.9997	0.9997	0.9998
23	0.0006	0.0005	0.0006	0.0006	0.0005	0.0006	0.9997	0.9998	0.9997	0.9997	0.9997	0.9998
25	0.0006	0.0004	0.0005	0.0005	0.0004	0.0006	0.9997	0.9998	0.9997	0.9997	0.9997	0.9998

4. Conclusion

This paper investigated the suitability of DTs in enhancing WAP schemes. DT models enable fast execution and present a simplified interpretation of rules to the task involved. Upon testing of the optimal DT model, it was found to be 99.82% accurate in predicting power swings as presented in TABLE 3.

The application of DT models shows significance in digital relay configuration settings. The splitting point values of the optimal DT model mark the boundary between the stable and unstable cases, therefore the threshold digital relay settings. The violation of these threshold limits would actuate the digital distance relay to perform the RAS, specifically the OST and OSB. The RAS is to mitigate the impact of OOS of generators, pole slip/frequency deviation of the power system and the loss of stability of the power system network due to power swings/transients.

In performing the RAS it is recommended that circuit breaker locations for OST should be at the electrical centre where the voltage is zero. The electrical centre is found at $\delta = 180^0$. Further work could be investigated on methods of islanding location. The identified islands should reduce areas cut out of power supply by employing smart dispatch programs.

On studying DT suitability to enhancing WAP schemes, the author's specific contributions presented in this paper are thus:

(i) Designed an adaptive OOS relay using a DT model, which illustrated how a reliable WAP scheme could be developed. The designed model exhibited novelty in its ability to predict successive power swings in a timely fashion. The DT model had a high accuracy in discriminating between the various power swing types.

(ii) Proposed a novel execution procedure for the designed DT model. The procedure was to ensure timely execution of the right RAS.

The beneficiaries of the findings of this paper include power system protection engineers and system operators.

Acronyms and Notation

$\hbar^0$	Whole data set.
$i(t)$	Gini index.
T	Final tree.
$n(t)$	The total number of vector measurements at node t.
$n(t_L)$ and $n(t_R)$	Total number of vectors falling into the left and right subsets respectively.
$n_t(C_j)$	The actual number of cases of class C_j at node t.
$C(j\|i)$	Cost of classifying i as j.
P_L and P_R	Impurity levels at both subsets t_L and t_R respectively.

$p(C_j, t)$	Re-substitution estimator of the probability that a case falls in node t and belongs to class C_j.
n_+, n_-	Number of cases with positive and negative actual states respectively.
$n_{-=j}$	Number of true negative cases with test results equal to j.
$n_{+=j}$	Number of true positive cases with test results equal to j.
$n_{+>j}$	Number of true positive cases with test results less than j.
$n_{+<j}$	Number of true positive cases with test results greater than j.
$M1$	For categorical Y denotes the empirical prior situation.
$M2$	For categorical Y denotes the non-empirical prior situation.
N_f	$\sum_{n \in \hbar} f_n$; number of cases in data set in test sample.
$N_{f,j}$	$\sum_{n \in \hbar'_v(t)} f_n I(y_n = j)$; number of class j in $\hbar$.
$M\overline{y}(t)$	$\frac{1}{N_f(t)} \sum_{n \in \hbar(t)} f_n y_n$ Mean dependent variable in $\hbar(t)$
$\Delta I(\widetilde{S}_x, t)$	Maximal decrease in node impurity for division of a parent node t into child nodes C_1 and C_2 guided by surrogate splits.
$p(t)$	$\sum_j^J p(C_j)$; estimator of the probability that a case falls in node t.
$\pi(C_j)$	$\frac{n(C_j)}{n}$; prior probability provided by the trainer of the data.
$p(C_j \| t)$	$\frac{p(C_j, t)}{p(t)}$; estimated probability that a case falls in node t and belongs to class C_j.

References

[1] M. Enns, L. Budler, T. W. Cease, A. Elneweihi, E. Guro, M. Kezunovic, J. Linders, P. Leblanc, J. Postforoosh, R. Ramaswami, F. Soudi, R. Taylor, H. Ungrad, S. S. Venkata, and J. Zipp, "Potential applications of expert systems to power system protection," *IEEE Transactions on Power Delivery*, vol. 9, no. 2, pp. 720–728, Apr. 1994.

[2] K. Yabe, J. Koda, K. Yoshida, K. H. Chiang, P. S. Khedkar, D. J. Leonard, and N. W. Miller, "Conceptual designs of AI-based systems for local prediction of voltage collapse," *IEEE Transactions on Power Systems*, vol. 11, no. 1, pp. 137–145, Feb. 1996.

[3] I. H. Witten, E. Frank, and M. A. Hall, *Data Mining: Practical Machine Learning Tools and Techniques: Practical Machine Learning Tools and Techniques*. Elsevier, 2011.

[4] T. M. Mitchell, "Machine learning and data mining," *Commun.ACM*, vol. 42, no. 11, pp. 30–36, Nov. 1999.

[5] E. Bernabeu, "Methodology for a Security-Dependability Adaptive Protection Scheme based on Data Mining," Virginia Polytechnic Institute and State University, Blacksburg, Virginia U.S.A, 2009.

[6] D. Novosel and R. L. King, "Identification of power system emergency actions using neural networks," in *Proceedings of the First International Forum on Applications of Neural Networks to Power Systems*, pp. 205–209, Seattle, WA, 1991,.

[7] R. Zivanovic and C. Cairns, "Implementation of PMU technology in state estimation: an overview," *4th IEEE AFRICON*, vol. 2, pp. 1006–1011, 1996.

[8] Y. V. Makarov, P. Du, S. Lu, T. B. Nguyen, X. Guo, J. W. Burns, J. F. Gronquist, and M. A. Pai, "PMU-Based Wide-Area Security Assessment: Concept, Method, and Implementation," *IEEE Transactions on Smart Grid*, vol. 3, no. 3, pp. 1325–1332, Sep. 2012.

[9] D. G. Hart and V. Gharpure, "PMUs – A new approach to power network monitoring," Review 1 1/2001, 2001.

[10] D. Novosel, "Final Project Report Phasor Measurement Application Study," University of California, Prepared for CIEE, Jun. 2007.

[11] F. J. Marín, F. García-Lagos, G. Joya, and F. Sandoval, "Genetic algorithms for optimal placement of phasor measurement units in electrical networks," *Electronics Letters*, vol. 39, no. 19, p. 1403, 2003.

[12] D. Dua, S. Dambhare, R. K. Gajbhiye, and S. A. Soman, "Optimal Multistage Scheduling of PMU Placement: An ILP Approach," *IEEE Transactions on Power Delivery*, vol. 23, no. 4, pp. 1812–1820, Oct. 2008.

[13] N. H. Abbasy and H. M. Ismail, "A Unified Approach for the Optimal PMU Location for Power System State Estimation," *IEEE Transactions on Power Systems*, vol. 24, no. 2, pp. 806–813, May 2009.

[14] S. Rovnyak and Y. Sheng, "Using measurements and decision tree processing for response-based discrete-event control," in *IEEE Transactions on Power Systems*, vol. 24, pp. 10–15.

[15] W. C. Morris, "One Slip Cycle Out-of-Step Relay Equipment," *Transactions of the American Institute of Electrical Engineers*, vol. 68, no. 2, pp. 1246–1248, Jul. 1949.

[16] B. Kasztenny and M. Kezunovic, "Digital relays improve protection of large transformers," *IEEE Computer Applications in Power*, vol. 11, no. 4, pp. 39–45, Oct. 1998.

[17] M. L. Othman, I. Aris, S. M. Abdullah, M. L. Ali, and M. R. Othman, "Knowledge Discovery in Distance Relay Event Report: A Comparative Data-Mining Strategy of Rough Set Theory With Decision Tree," *IEEE Transactions on Power Delivery*, vol. 25, no. 4, pp. 2264–2287, Oct. 2010.

[18] V. Centeno, A. G. Phadke, A. Edris, J. Benton, M. Gaudi, and G. Michel, "An adaptive out-of-step relay [for power system protection]," *IEEE Transactions on Power Delivery*, vol. 12, no. 1, pp. 61–71, Jan. 1997.

[19] V. Centeno, A. G. Phadke, A. Edris, J. Benton, and G. Michel, "An Adaptive Out-of-Step Relay," *IEEE Power Engineering Review*, vol. 17, no. 1, pp. 39–40, Jan. 1997.

[20] D. Tholomier, S. Richards, and A. Apostolov, "Advanced distance protection applications for dynamic loading and out-of step condition," *Power and Energy Society General Meeting - Conversion and Delivery of Electrical Energy*, pp. 1–8 2007.

[21] R. Tiako, D. Jayaweera, and S. Islam, "A class of intelligent algorithms for on-line dynamic security assessment of power systems," in *20th Australasian Universities Power Engineering Conference (AUPEC)*, pp. 1 –6, 2010.

[22] N. D. Hatziargyriou, G. C. Contaxis, and N. C. Sideris, "A decision tree method for on-line steady state security assessment," *IEEE Transactions on Power Systems*, vol. 9, no. 2, pp. 1052–1061, May 1994.

[23] L. Wehenkel and M. Pavella, "Advances in decision trees applied to power system security assessment," in *, 2nd International Conference on Advances in Power System Control, Operation and Management*, vol.1, pp. 47 –53, 1993.

[24] Kai Sun, S. Likhate, V. Vittal, S. Kolluri, and S. Mandal, "An online dynamic security assessment scheme using phasor measurements and decision trees," in *Power and Energy Society General Meeting - Conversion and Delivery of Electrical Energy in the 21st Century*, vol. 22, pp. 1–6, Pittsburgh, PA, 2008.

[25] T. Van Cutsem, L. Wehenkel, M. Pavella, B. Heilbronn, and M. Goubin, "Decision tree approaches to voltage security assessment," *Generation, Transmission and Distribution, IEE Proceedings C*, vol. 140, no. 3, pp. 189 –198, May 1993.

[26] R. Diao, K. Sun, V. Vittal, R. J. O'Keefe, M. R. Richardson, N. Bhatt, D. Stradford, and S. K. Sarawgi, "Decision Tree-Based Online Voltage Security Assessment Using PMU Measurements," *IEEE Transactions on Power Systems*, vol. 24, no. 2, pp. 832 –839, May 2009.

[27] L. Wehenkel and M. Pavella, "Decision Trees and Transient

Stability of Electric Power Systems," 1991. [Online]. Available: http://orbi.ulg.ac.be/handle/2268/80412. [Accessed: 26-Jul-2012].

[28] O. Ozgonenel, D. W. P. Thomas, and T. Yalcin, "Superiority of decision tree classifier on complicated cases for power system protection," in *11th International Conference on Developments in Power Systems Protection*, pp. 134–134, Birmingham, UK, 2012.

[29] Z. Li and W. Wu, "Phasor Measurements-Aided Decision Trees for Power System Security Assessment," in *2nd International Conference on Information and Computing Science(ICIC '09)*, pp. 358–361, Manchester, 2009.

[30] J. A. Pecas Lopes and M. H. Vasconcelos, "On-line dynamic security assessment based on kernel regression trees," in *IEEE Power Engineering Society Winter Meeting*, vol. 2, pp. 1075 – 1080 v2, Singapore, 2000.

[31] E. E. Bernabeu, J. S. Thorp, and V. Centeno, "Methodology for a Security/Dependability Adaptive Protection Scheme Based on Data Mining," *IEEE Transactions on Power Delivery*, vol. 27, no. 1, pp. 104–111, Jan. 2012.

[32] D. Steinberg and M. Golovnya, *CART v 6.0 User's Manual*, vol. 1. San Diego USA: Salford Systems, 2002.

[33] D. Steinberg and C. Phillip, *CART-Classification and Regression Trees*, vol. 1. San Diego, USA: Salford Systems, 1997.

[34] L. Brieman, J. Friedman, and R. Olshen, *Classification and Regression Trees*. Pacific Groove, Wadsworth: Salford Systems, 1984.

[35] IBM Corporation, *IBM SPSS Statistics 20 Command Syntax Reference*, 1st ed., vol. 1. USA: IBM Corporation, 2011.

[36] A. Chaouachi, R. M. Kamel, R. Andoulsi, and K. Nagasaka, "Multiobjective Intelligent Energy Management for a Microgrid," *IEEE Transactions on Industrial Electronics*, vol. 60, no. 4, pp. 1688 –1699, Apr. 2013.