Strategic PMU placement for stability enhancement

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 a novel method for strategic monitoring of a power system to schematically monitor power system variables that are sensitive to transients. The characteristics of a fully developed transient or power swing increase frequency slip rates, generator pole slips, rotor out-of-step etc. whose effects lead to loss of synchronism of coherent generators in a power system. When these occur, the resulting remedy could be load shedding schemes, generator tripping or controlled islanding. Failure to achieve any of these might lead to geographically extensive blackouts and/or the damage of auxiliary power system equipment.This paper looked at the Wide Area Monitoring (*WAM*) principle, consisting of collection and pre-processing of field data, using Phasor Measurement Units (*PMUs*). A data mining exercise was performed purposing to identify strategic positions for *PMU* placement using the Classification and Regression Trees (*CART*) algorithm. The logic of *CART* was therefore also discussed.The proposition of strategic *PMU* placement as implied by the Decision Tree (*DT*) model acknowledges that a few *PMUs* in the power system network are capable of achievingWide Area Protection(*WAP*)functions.

Keywords: Wide Area Monitoring, Wide Area Protection, Out-of-Step, Stability, Power Swing, Decision Tree

1. Introduction

This research was motivated to explore the capabilities of autonomous systems and its adoption to power system protection functions. These can be adapted to Wide Area Monitoring Protection and Control (*WAMPAC*) of power networks with the help of Phasor Measurement Units (*PMUs*).

The advent of Phasor Measurement Units (*PMUs*) having microsecond accuracy, Phasor Data Concentrators (*PDCs*) and fibre optic data link cables (with transmission speeds of up to 100 Terabits/second) as reported by [1], [2], [3]have enabled the adoption of the automation concept (digital relays, pilot relay schemes and artificial intelligent systems) to wide area power system protection. The aggregation of these device functions enable data and decisions to be quickly and easily transferred between databases and substations. This enables the serving of these sub-stations in an unlimited resource in terms of geographical displacement and time.

Owing to the dynamic behaviour of power systems, monitoring the grid to maintain its integrity has become a day to day activity in power stations. Having advancements in technologies to accomplish this, other power system problems still re-surface calling upon more stringent actions to be enforced. The idea and need for Artificial Intelligence (AI) and machine learning processes appear to be strikingly promising following evolving power system's dynamic behaviour, faced with various different abnormalities while at a different state. This paper thus follows to discuss issues influencing wide area monitoring to tactfully maintain power system stability.

2. Alignment of Wide Area Measurements

Conventional methods in performing power system analysis involved assessing individual measurements of a specific local power system area. These were analysed accordingly to maintain the required local area normal operating condition. However, it has been established that power system abnormalities are a combination of various power system factors, aggregated from inter connected power system areas and each being contingent, one upon the other. Having this thought, in order to curb power system abnormalities more accurately, power system wide area

measurements need to be captured for various analyses. All these measurements have to be aligned to a common reference for analysis. This process is computationally burdensome and vulnerable to erroneous calculations. If these measurements were all aligned to a common reference from their point of collection, then this burden and error vulnerability would be reduced when all local measurements are collected together in a common database and used for various functional analyses.

The convenience of Phasor Measurement Units (*PMU*) eases this burden as it pre-processes (filters electrical noise) wide area field measurements and aligns them to a common time reference. The measurements are synchronized by the Global Positioning System (*GPS)* to a common time and therefore reliably used in the analyses of various power system functions.

WAM is achieved through having *PMUs* placed optimally to observe the network activities. Monitored are voltage and current magnitudes and phasors as well as the frequency deviation from its nominal value. Optimal placements of *PMUs* are described in papers [9], [10], [11], [12], [13], [14], [15], [16], [17], [18]. These placed *PMUs* collect measurements at the bus at which they have been placed & all incident lines that terminate at these bus where the *PMU* have been placed. The *PMU* is also capable of obtaining measurements of adjacent buses that are directly connected to the *PMU* bus through a line. The latter however depends on the number of output channels that a particular manufactured *PMU* has. Thus, having a single *PMU* at an optimized location in the electrical network, a wide area of measurements of the power system can be obtained. This is wide area monitoring.

The obtained measurements are communicated to a central depository database and then to the power system control centre. At these centres, management and functional applications of these measurements are analysed and thereafter utilized. Such functions that will alter/affect the operation of this wide area are then referred to as wide area protection & control.

A *PMU* dedicated to provide measurement data for protection and control functions should be placed on the bus bar incident to the area of interest as per conventional methods. In case of topology change causing *PMU* inaccessibility to obtain or channel data, *PMU* and transmission line redundancy is encouraged. Desirable *PMU* placement locations as recommended and seconded by [5], [6] for situational awareness and for protection functions case include at:

(i) Main generating plant/source and its subsidiary sources (probably above 500 *MW*). *PMUs* will provide data relevant for generator dynamics, synchronization and generator tripping & dispatch scheduling.

(ii) High-voltage transmission systems (above 230 *kV*), including those with *HVDC* stations, *FACTS* controller installations, or large tap-changing and phase-shifting transformers: *PMUs* in these locations give useful data for voltage stability analysis (the analysis of the available transmission corridor), transmission control devices configuration (*FACTS*)

(iii) A major load centres, *PMUs* can monitor the load characteristics relevant for generator-load balancing, load-shedding schemes etc.

(iv) Pre-defined separation islands where the rate of slip between the two sources can be monitored.

System wide improvements by utilizing Synchrophasor Measuring Technologies (*SMTs*) or *PMUs* [7] include the availability of real time measurement information to perform the following (see Figure 1)
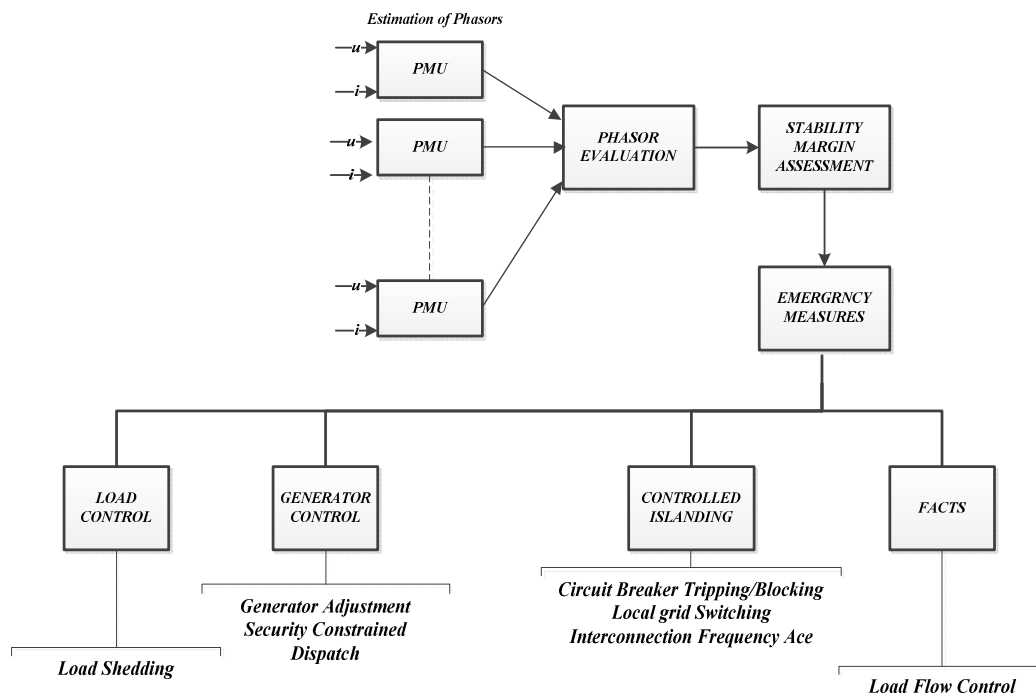

Figure 1. WAMPAC Functions in Enhancing Stability [13]

(i) Schematic visualization of the power system's state/situational awareness

(ii) Correct the conservative limits estimated (enhanced state estimation) during power system off-line planning.

(iii) Monitor marginal operating limits of the power system and thus be able to design early warning systems in the event operating condition s become very critical

(iv) Data mining the *SMT/PMU* measurements enables the design of adaptive protection and adaptive control systems

(v) Data patterns existing in phasor data concentrators can be used for benchmarking and validation of new designed system models

(vi) Pre-islanded portions of the network can be easily monitored and thus improve the damping of inter-area oscillations and controlled islanding.

(vii) Forensic analysis of the causes of system failure by observing power system parameter behaviour before and during fault periods.

Quasi – steady – state measurement from *PMUs* provides an almost real time data scheme for dynamic analysis of power systems. Phasor measurements give the magnitude of the internal voltage angle and the power angle of generating units for computation of the rotor position. These parameters of generating units don't change with generator operation unlike the direct-axis reactance X_d and quadratic-axis reactance X_q used in classical methods to compute the electrical internal voltage and power angle of the generator.

A *PMU* placed at a generator bus should thus measure internal and terminal phase and currents voltages of the generator; from which the power angles are derived. The *PMU* should also have a rotor position indicator. All these parameters are synchronized to a common time reference.

The rotor position can be tracked by optical or magnetic means where a 'shaft encoder' or a cyclic synchronous wave may be produced by making a provisional slot to the rotor and these compared to the reference signal. At no-load rotor angle α should equal zero, while voltage angles at terminal of generator should equal internal voltage angles of the same generator. From the reference signal, terminal voltage angle under no load is denoted β. The offset between angles α and β is γ and normally remains constant unless generators physical modification is performed (such as coil re-assembly).
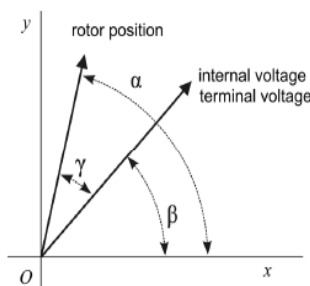

Figure 2. PMU Phasor Diagram under No-Load Conditions

Figure 2 represents the phasor locations from a common reference of generator internal and terminal voltages under no-load conditions.

At load conditions, the internal voltage angle can be calculated from the rotor position α and the calibration offset γ since these parameters remain constant. Thus the internal voltage β would be calculated as the difference between the offset and the rotor angle. (1)

$$\beta = \alpha - \gamma \qquad (1)$$

The difference between the internal voltage β and the terminal voltage gives the generator power angle δ Figure 3
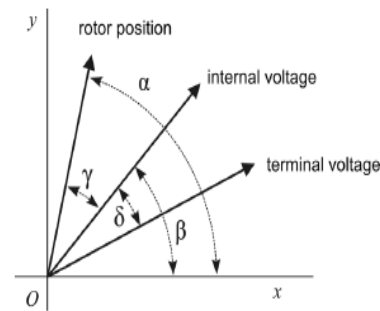

Figure 3. Phasor Diagram at Load Conditions

This paper specifically looks at variables that are thought to compromise the stability of a power system, i.e. attributes that have significant weight/influence to the decision rules of a developed *DT* model in terms of power system stability and protection. Therefore, these candidate attributes, relevant to power system protection (out-of-step relaying) will be trained (supervised) by variables when the power system's generators go Out-of-Step (*OOS*). (If predictors characteristic cause an *OOS*, the contributing attributes are classified as unstable, otherwise stable). The single developed tree should be reliable in judgement of the stability status of the power system.

The *DT* model has implications on strategic Wide Area Monitoring (*WAMs*); the splitting criterion of the optimal *DT* model identifies the critical positions for *PMU* placements. The hypothesis being tested therefore is that a few strategic *PMUs* placed on the power system are capable of achieving *WAP* and enhance the stability of the power system. The split nodes of the optimal *DT* model also identify the critical parameters/variables to monitor for power systems transients.

Wide Area Monitoring (*WAM*) for Wide Area Protection (*WAP*) is achieved from synchronised wide area measurements; through various field instruments such as *PMUs*, Digital Fault Recorders (*DFRs*) etc. which are all transmitted to a common central database. This *WAM* offer Situational Awareness (*SA*) on the operating states of the power system. The *PMU's* data acquisition is timely, ensuring time consciousness to the protection scheme. Synchronised measurements offer a more reliable data for power system analysis; as the signals are aligned to a

common time reference, are time stamped, μs accurate and filtered from electrical noise.

The transmission of the measured field data is through various telecommunication channels such as coaxial cable, fibre optic cable, wireless modems, etc. The communication channel speed needs to match up the urgency of the function to be executed in order to maintain the integrity of the power system. For power swing mitigation using out-of-step trip or block functions, the telecommunication of the synchronized wide area measurements needs to be faster than the critical clearing time of the power system so that the power system can still be stable.

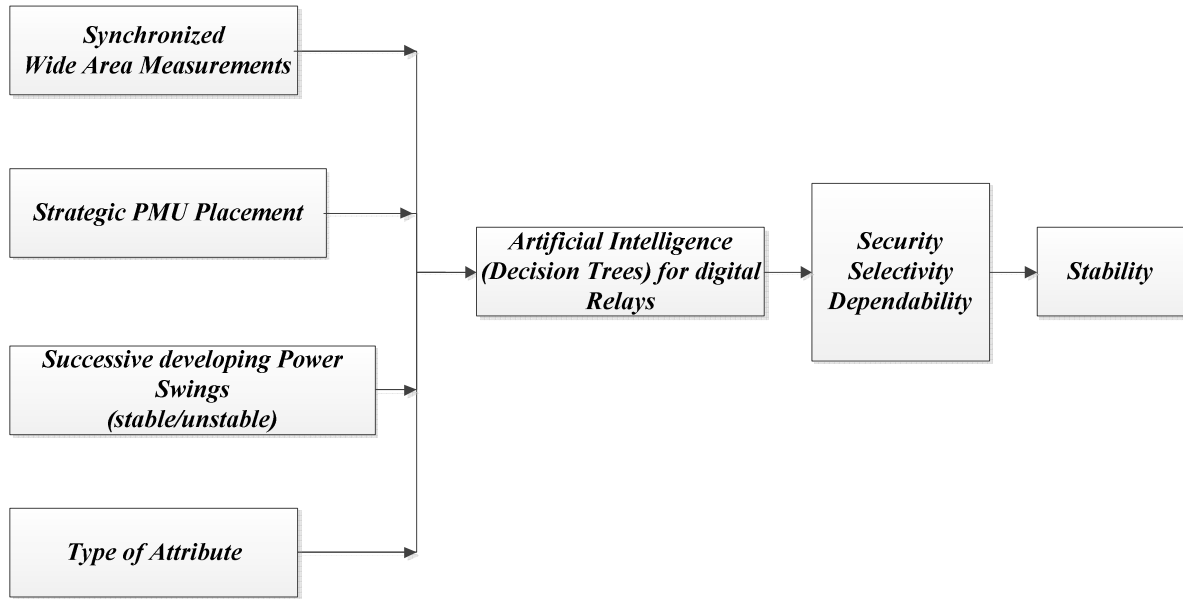*WAMs* use communication networks and devices to transmit analogue/digital data from dispersed network locations to a central analysis point. In the neo-classical supervisory control systems acquired data were measured at a common reference wave signal using clocks synchronized to Universal Time Coordinated (*UTC*). Time obtained from *UTC* is through transmitters e.g. *DCF77* transmitter in Frankfurt Germany, having accuracies of up to 10 *ms*. Greater accuracies of microsecond are obtained by satellite transmission from Global Positioning System (*GPS*). Transmitted messages from *GPS* at One Pulse per Second (*1PPS*) give calibrations for the beginning and the end of the pulse for time alignment purposes of other analogue/digital signals thus providing for synchronisation with *UTC* as the common referencing time.

Figure 4. Wide Area Synchronized Measurements

WAMPAC is an apprehension that mobilizes local area measurements. These local measurements are synchronized, aggregated and transmitted to various centres for various power system functions to maintain system integrity [8]. *WAMPAC's* main building blocks are the *PMUs*, phasor data concentrators (*PDCs*), application software (*AS*), and their supporting communication networks (*CNs*).

A. Strategic PMU Placement

SA of the power system is achieved through *PMU* placement. A placement to observe only critical power system parameters influencing the stability of a power system is being argued out. The identification of these parameters and therefore only monitoring or prioritizing these variables as important, influences the operation of the predictor model (adaptive *OOS* relay) in making decisions.

B. Developing Power Swings

Identifying the different types of power swings influences the relay's decisions to perform Remedial Action Schemes (*RAS*). This is the action taken by the wide area protection scheme to restore the power system back to its normal operating state or to maintain a good voltage profile having standardized electrical parameters in the power system.

The remedial action taken by the out-of-step relay is either to block a trip signal or to send a trip signal to the circuit breaker(s). These are technically controlled islanding decisions referred to as Out-of-Step Blocking (*OSB*) and Out-of-Step Tripping (*OST*) respectively. When a stable power swing is detected, a block signal is sent but when an unstable power swing is detected, a trip signal is sent to the circuit breaker. Faults are instantaneous and last for a very short time and therefore when the relay picks up a fault signal, the line experiencing the overcurrent should be isolated as soon as this overcurrent is detected. Power swings on the other hand are relatively slower than faults and therefore should be implemented considering a time delay. However, this time delay should be less than the critical clearing time of the power system to avoid loss of synchronism and stability. For power swings, they last longer and therefore if the pick-up signal by a conventional distance relay is to direct a trip, the same trip signal will be sent over and over until the power swing clears. This results in cascading trips, misbalancing the load-generator equilibrium thereby causing the system to lose synchronism.

The fact that the relay continuously sends cascading trip

signals for the wrong type of power system abnormality (power swing) means that the relays have lost their security characteristic. For the mitigation of power swings, this out-of-step function is implemented outside the third zone of the distance relay. If this function of the out-of-step relay performs properly, the relay is said to have a good selectivity. That is the outer most zone of the distance relay will be the first zone to detect the power swing and will be the first zone to act on it.

Upon classifying power swings, stable power swings will not be threatening to the power system. Attention has to be paid to both stable and unstable power swings, though unstable power swing effects are more detrimental to the power system. If these unstable power swings are left to sustain, the power system will be unstable and lose synchronism.

Nonetheless, the if power swings develop successively one after another, time consciousness to the WAP should be considered for timely prediction, classification and mitigation through the appropriate RAS.

C. Type of Attribute

The response characteristics of particular attributes to various changes in the power system are different from each other. Particular power system variables are more sensitive & vulnerable to transients than others. Therefore, monitoring parameters that show a good response to an Out-of-Step condition are crucial to detecting the stability status of a power system.

The bias to important variables will reduce the amount of data size that updates the adaptive relay enabling its execution of RAS to be faster. Upon monitoring these variables, the performance of detecting and mitigating successive power swings is also enhanced.

D. Artificial Intelligence (Decision Trees) for Digital Relays

Digital relays operate according to the ordered algorithm installed in them. Because of the changing operating states of a power system, adaptive relaying through the adoption of AI adjusts the relay's security, selectivity or dependability to achieve normal operating conditions of the power system [19].

3. Situational Awareness (SA)

For wide area protection schemes, the corresponding control devices need to be aware of what is happening in the larger power system network. Decisions made based on information gathered from the entire power network are more accurate, timely and relieve the control devices from regular adjusting and re-adjusting of various power system parameters to maintain normal operating conditions. Therefore, the need for Wide Area Monitoring (WAM) devices becomes a significant note in performing Wide Area Protection (WAP).

In tune to monitoring devices, the authors of reference [20] frontier research in Phasor Measurement Units (PMUs.) They provide the schematic design of the PMU. In their paper, they present various applications of the PMU to power systems, with little found applications to power system protection functions. Following this queue, the findings and literature of this research adds to bridge this gap on the application of PMUs to power system protection functions.

So to say, a full awareness of the potential of PMUs thus needs to be unveiled and applied to power system functions not limited to protection functions but to the entire power system functionality. This will revolutionise power system device operation and application to control thus protection functions. To this effect, the conventional Supervisory Control and Data Acquisition (SCADA) systems found in current power system control centres are paving way for the Wide Area Measurement Protection and Control (WAMPAC) paradigm.

In the context of power swing mitigation to reduce the consequence of having an out-of-step of generators, the traditional Out-of-Step (OOS) relays are locally oriented. They lack system wide awareness from the whole network. They therefore perform their actions based on this local information, mainly from the generator source which they are protecting. The impact of this local dependency of information could possibly lead to relay coordination complexities. This would be due to the fact that other relays would be similarly configured according to the information they received from their respective local surrounding. The relays would thus not be aware of contingent operations of the entire network affecting their decision making functions.

The situation is that power system states continually change, leaving the power network vulnerable to collapse in the presence of various abnormalities. The dynamism of the fault nature is brought about by these changing states. That is, a particular fault magnitude is adverse to the power system when the power system is in an emergency state. The same fault magnitude is less harmful when the power system is in a normal state. Having static settings of the traditional OOS relay, faulty operations are bound to happen as the OOS relay will not recognize the various changing power system states.

To the above notion of situational awareness, the mode of measurement, functionality and implementation of PMUs have been discussed by [21], [22], [23], [24], [25]. They give further technical attributes on the capabilities of PMUs. In the implementation process of PMUs for WAMs, it is however notable to have an optimal placement if the entire network has to be observed. The objective function of the optimal placement is to observe the entire power network having installed the least number of PMU devices. Reasons prompting optimal placement are for full network observance and to reduce the capital costs involved during purchase and implementation of PMU and wide area monitoring devices. Placement techniques have been discussed in [10], [12], [13], [15], [9].

The significance of wide area monitoring for situational awareness therefore lies in providing system wide prevailing conditions through measurements. These measurements are utilized to maintain system integrity through various 'informed' control actions. Power system operators are able

to have a clear view of operating limits and therefore make informed decisions on the whole power system network.

PMU implementations are evident in the North East United States and Italian interconnected systems after blackouts in 2003. Subsequently, the Eastern and Western Interconnection Phasor Project (*EIPP&WIPP* respectively) in the United States unraveled to curb the problem by deploying *PMUs* to monitor the system. In Europe, the Italian blackout in 2003 resulted in a number of *PMU* being deployed throughout the continent presently. Researchers are now putting more emphasis on making use of the phasor information to improve the system's security and reliability. The success of these could foster *PMU* deployment in power system networks in the African continent as well [9].

However, deployment challenges and considerations common to implementing *PMU* projects include aligning signals and time zones to the Global Positioning System (*GPS*) or a common time referencing signal. Advanced analytics are required to manage different granularities of real-time data received from *PMU*s, Phasor Data Concentrators (*PDCs*) and *SCADA* systems [14], [16].

4. Size and Type of Attributes/Variables in Hypothesis Space

In selecting data to be trained from a data base, it is important to have a bias to selecting attributes that are deemed to have weighty contribution to the problem being investigated. It is however assumed that the more information one has, the more likely the decision model will be able to make accurate decisions [26]. With limitations of digital systems taking long execution times for a large database of measurements, slower response time in actuating protection auxiliaries (circuit breakers, capacitor compensators etc.) to mitigate a fault or power swings, (lasting for periods between 3-40 cycles) it would be reasonable to employ strategies that would reduce the amount/size of executable data by using only relevant data substantial enough to give reliable judgement. This notion of reducing the hypothesis space or sample size is also suggested by the works of [27]. This therefore led to the investigation of finding out if a minimal number and a specified type of electrical variables could be identified to predict a power swing. Various studies like [28] have shown that accurate results are possible with a lesser training sample, testing its accuracy on a larger set of database information.

It is notable and evident that most analysis engineers use pre-processed measurements (using adaptive filters) to reduce the noise ratio compared to the signal. By doing so, the collected measurements give a true representation of the actual signal being produced by the power system. Some of these filters use the conventional Discrete Fourier Transform (*DFT*), Fast Fourier Transform (*FFT*) etc. while most recent filters employ a combination of algorithms to suit to the measurement environment. Relay input data and digital fault recorder data could supplement *PMU* measurements for

retrieving relevant data that the *PMU* cannot measure. This also has a great value in revealing hidden failures experienced by the *OOS* relays, though not all measurement data is used for the prediction purpose.

When developing an intelligent decision model for the purposes of power system protection, simulation cases that reflect faulty scenarios should comprise a larger majority in the hypothesis space/sample data set. These include large motor load models, network topology changes (adding/tripping of power lines etc.) to tap best measurement sets for a data mining exercise with an aim of achieving a learning function/model representing various power system parameter relations. The importance of this is for the critical learner of the model to acquaint itself with various fault levels at various power system states.

When relay threshold settings don't seem to be right, or if nuisance tripping occurs after developing a decision model it could be probable that the number of simulations having various contingencies and operating points was not wholesome to cover all possible power system situations to properly train the data. The intensity of the simulations having various operating points greatly influences the resulting decision rules. A significant simulation exercise would result in decision rules being sufficient enough to pose as a blanket policy setting for power relays to mitigate on all other power system fault types.

Best practice in setting the hypothesis space for training models involves partitioning the data into various sub-groups and using each sub-group as a training sample and similarly as a testing sample. This is commonly referred to as v-fold test/train criterion. The *v* refers to the number of partitions. Authors in [29] suggest that training should be done continuously to improve the prediction accuracy. This should not be the case as regular training will deprive the predictive model its meaning of being adaptive. However, updating of the decision models needs to be periodic. When a major topology change occurs (e.g. loss of line/generator) the updating of the important variables should be done immediately. This updating of data for developing new decision models is however influenced by the sampling rate of input measurements and communication transmission speeds. Considerable measure needs to be put when matching/selecting communication channels for updating decision models from remote databases.

Variables contributing highly in variable ranking/variable importance to performance of the model are chosen as candidate attributes and only these are used by the decision model to develop a new decision model/learning function. Moreover, these are the only attributes updated. The final decision model will therefore comprise classes of these candidate attributes each indicating cut-off values of each class. The representation of a pure class of attributes with threshold limits of contents belonging to that particular class is called a monograph. Each selected monograph therefore identifies execution rules for the decision model.

The sample size and type of attributes in the sample set for training are therefore important factors as they enhance

execution speeds of the decision models [30].

At this juncture, it is justifiable to perform classification methods of various electrical parameter characteristics. Thus, the Decision Tree (*DT*) technique using the Classification and Regression Trees (*CART*) is employed to perform this classification. The aim would be to find out how each variable responds to power system changes. The relationships of most sensitive variables to a particular power system disturbance are therefore able to be drawn from this classification. These sensitive variables can then be solely used to detect when these particular disturbances occur. *DT* methods are also discussed in references [31], [28], [32], [33], [34], [35], [36], [37], [38], [39], [40], [41], [42], [43].

DT is applicable to any data type/ data structure having a large set of input data and therefore suitable for drawing/marking decision boundaries for data having large dimensional spaces. Its major property involves classifying variables of same characteristics into classes. From these classes, rules of the developed model are easily derived from a top-down ordered sequence traversing through each terminal leaf node until the final (independent variable) terminal node infers a decision. The training of *DTs* can be done in both on-line and off-line modes thus enabling the *DT* model to be updated through whichever means.

PMUs aid in enhancing the protection function by providing timely and accurate wide area measurements to real-time protection schemes. All these wide area field measurements are aligned by the *GPS* to a clock common time reference i.e. *UTC* at *1PPS*. This *1PPS* functionality enables the *PMU* measured signals to be time stamped, enabling real-time monitoring & control of the power system as well as performing post-mortem analyses.

The standardisation of synchronised measurements also improves measurement accuracy by ensuring compatibility of:- devices, transmission protocol, and data formats. Various synchrophasor standards have been listed. The ultimate benefit of a synchrophasor standardisation is high measurement accuracy, *TVE* confined to 1%. This translates to a time accuracy of $1\mu s$ and a phase error of 0.022^0 for a 50 Hz or 0.0167^0 for a 60 Hz power system.

The *WAM* architecture as discussed with respect to Wide Area Protection &Control (*WAPC*) function. Because of the wide geographic reach of *PMUs*, communications infrastructure has to be invested on. The communications infrastructures, integrated into the power system using various standards have been viewed. Issues thought to hamper *PMU* functions have been discussed; noise/harmonics affecting frequency of measured signals, latency affecting data frame size and rate of frame delivery, and time synchronisation issues of signals to various clocks e.g. Universal Time Coordinated (*UTC*).

5. Methodology

The overview of obtaining the results of the strategic placement is as represented inFigure 5.

The Wide Area Measurement (*WAM*) concept is achieved when the aggregate or bulk of local area measurements are aligned to a common reference point (synchronized) and used for analysis of power system inter-area operating conditions. Wide Area Protection (*WAP*) is achieved through the utilization of these synchronized *WAMs* to mitigate the spread the propagation of power system abnormalities into the adjacent inter-area and subsequently to the entire network.
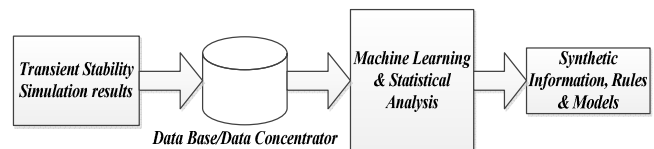
Figure 5. *Machine Learning Framework*

Transient stability analysis was performed on an *IEEE 39 bus test system* in *DIgSILENT ®* and the learning sample data was extracted using virtual instruments available in *DIgSILENT ®* platform. The Critical Clearing Time (*CCT*) criterion was used to determine the stability endurance of the system when it was subjected to transients; or rather how long the system would be able to withstand a power system abnormality without losing stability. The learning sample from the *PMU* measurement data was organized and stored in a data base in *Microsoft Excel® 2010*
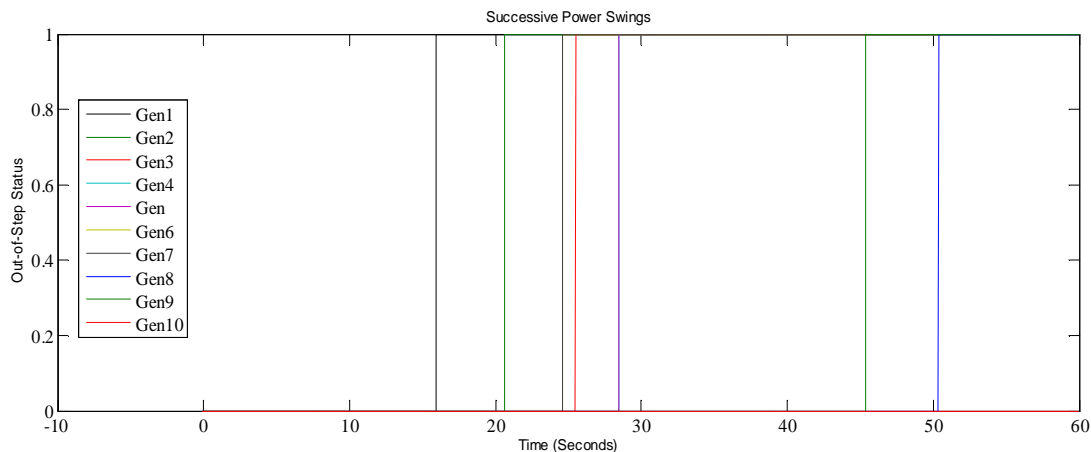
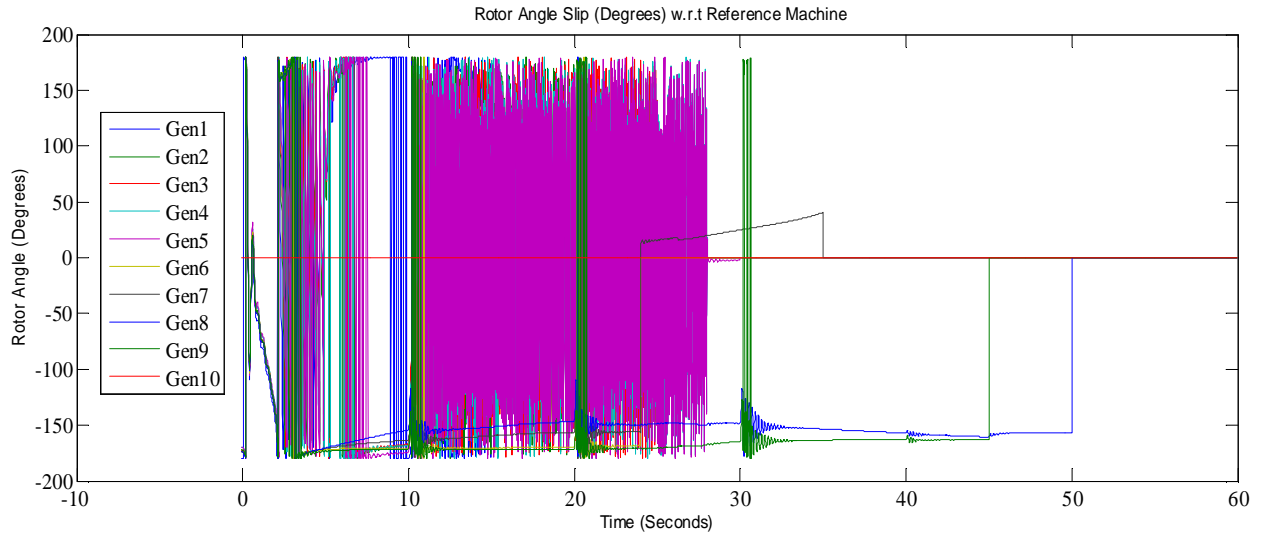**Figure 6.** *Simulation Responses of Successive Swings*

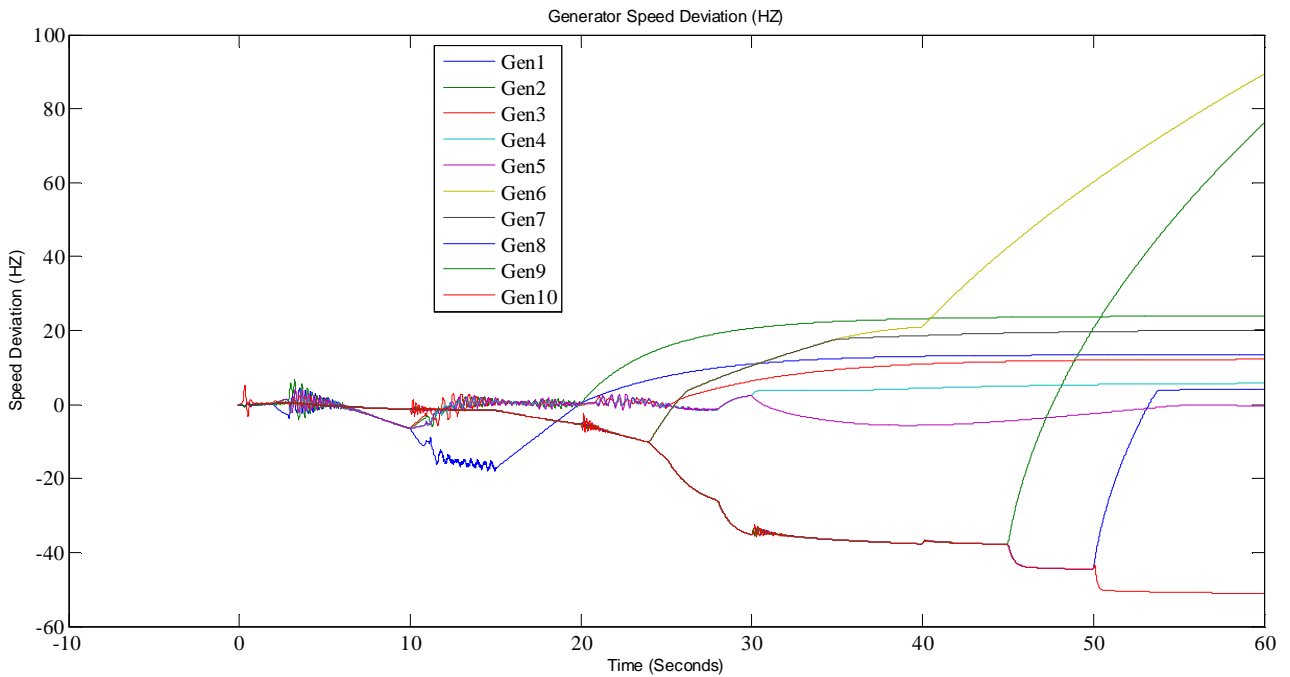**Figure 7.** *Rotor Angle Slip from Reference Machine*

Figure 8. *Generator Speed Deviations*

The *CART* analysis and *DT* model design was done using *Salford Predictive Modeller-CART® v6,* trial licence. Growing of the tree was done through the Gini splitting criterion.

The aim of the transient stability simulation was to induce power disturbances/swings at the critical load centres to create a generator-load imbalance. This was done until the power system was observed to be transiently unstable. A graphical representation of one of the simulation responses is represented as in Figure 6, Figure 7 and in Figure 8. The transient stability was bounded by a combination of the critical clearing time , frequency deviation from nominal, voltage phasor angle deviation from the reference machine and generator out-of-step.

(i) The voltage profile should be within 0.95-1.05 p.u.

(ii) The load phasor voltage angle should not advance the generator phasor voltage angle by 4 pole slips

(iii) The frequency deviation from the nominal frequency of the reference machine should not be greater than ±1%

(iv) Loss of synchronism/Out-of-Step of a generator.

The *DT* using the *CART* technique was developed as follows:

(i) The learning sample L was arranged as an $m \times n + 1$; where m = number of variable cases and $n+1$ = number of attributes + actual/known output from simulation. Because of space limitations, the learning sample could not be appended.

(ii) From the learning sample L, the minimum number of

elements $n_{\min}$ to make up a terminal node was set as 1. In this training, the *DT* model was proposed to be grown until its maximal i.e. having only one case in the final terminal node.

(iii) Attributes were sorted in order to initialize splitting points

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 is numeric, then splitting value is at the midpoint of adjacent measurements i.e.

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

Where

$x_a(k)$ = variables of attribute a having values $k = 1, ..., m$

b) Define a as a categorical variable of sets $S_a = \{s_1, s_2, ..., s_n\}$. If a is categorical, the possible splitting points fall within the range of available sets of that particular attribute.

(i) The impurity reduction level was computed, achieved by the improvement from (3).

$$\Delta i(s, t) = i(t) - [p_L \cdot i(t_L) + p_R \cdot i(t_R)] \tag{3}$$

Where

a) $i(t)$ = Gini index $i(t) = 1 - \sum_j P^2(C_j | t)$. The *Gini index* computes the impurity levels at both subsets t_L and t_R as (4).

$$\begin{aligned} P_L &= \frac{n(t_L)}{n(t)} \\ P_R &= \frac{n(t_R)}{n(t)} \end{aligned} \tag{4}$$

Where

$n(t)$ = 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

$p(C_j | t)$ = estimated possibility that a case falls in node t and belongs to class C_j

$$p(C_j | t) = \frac{p(C_j, t)}{p(t)} \tag{5}$$

Where

b) $p(C_j, t)$ = re-substitution estimator of the probability that a case falls in node t and belongs to class C_j

$$p(C_j, t) = \pi(C_j) \cdot \frac{n_t(C_j)}{n(C_j)} \tag{6}$$

Where

$n_t(C_j)$ = the actual number of cases of class C_j at node t

$n(C_j)$ = the total number of cases that belong to class C_j

$\pi(C_j)$ = Prior probability provided by the trainer of the data

$$\pi(C_j) = \frac{n(C_j)}{n} \tag{7}$$

c) $p(t)$ = estimator of the possibility that a case falls in node t

$$p(t) = \sum_j^J p(C_j) \tag{8}$$

(ii) After improvement of each attribute was computed, 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.

a) The measure of importance of a variable was expressed as $M(x) = \sum_{t \in T} \Delta I(\tilde{S}_x, t)$

Where

$M(x)$ = Measure of importance of a variable x in relation to the final tree T

$\Delta I(\tilde{S}_x, t) = \max \Delta IC_i(\tilde{S}_x, t)$ = maximal decrease in node impurity for division of a parent node t into child nodes C_1 and C_2 guided by surrogate splits $\tilde{S}_x$

b) The variable importance $VI(x)$ of x is expressed in terms of a normalised quantity relative to the variable having the largest measure of importance. This was calculated as (9).

$$VI(x) = \frac{M(x)}{M(x_{\max})} \times 100 \tag{9}$$

(iii) Using the Gini purity index, the root node was identified and selected; node having the greatest variance hence highest Gini impurity value.

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

a) For numerical variables, then

$$\begin{aligned} tL &= \{x_a(k) \, if \, x_a(k) \leq s_k\} \\ tR &= \{x_a(k) \, if \, x_a(k) > s_k\} \end{aligned} \tag{10}$$

b) For categorical variables, (have finite sets) then

$$tL = \{x_a(k) \, if \, x_a(k) = s_k\}$$
$$tR = \{x_a(k) \, if \, x_a(k) \neq s_k\} \qquad (11)$$

(v) The next step involved finding the optimal split $s_{optimal}$ over all possible splitting values $s \in S_a$ amongst all attributes $a \in A$ until no more splitting can be achieved. For Gini splitting, we recall the Gini Index for node purity as (13).

$$GINI_{Node}(Node) = 1 - \sum [p(j)]^2 \qquad (12)$$

$$GINI_{split}(S) = \sum_{i \in values(S)} \frac{|n_i|}{|n|} GINI(N_i) \qquad (13)$$

Where

n_i = Number of cases at node i

n = Number of cases at node p

$GINI(N_i)$ = the Gini Index of each node $GINI_{Node}(Node) = 1 - \sum [p(j)]^2$

$$GINI_{split}(S) = \frac{n_i}{n}(N_1) + \frac{n_j}{n}(N_2) \qquad (14)$$

(vi) A classification decision was made from terminal nodes. A node was classified in class i if

$$\frac{C(j|i)\pi(i)N_i(t)}{C(i|j)\pi(j)N_j(t)} > \frac{N_i}{N_j} \text{ for all values of } j \qquad (15)$$

Where

$C(j|i)$ = cost of classifying i as j

$\pi(i)$ = prior probability of i

N_i = number of class i in dataset

10-fold cross validation was used because of its robust learning and testing technique; it tests the decision tree model eleven times. The reliability of each sample will be measured by a misclassification rate given by the *Gini Index*. This is defined by $i(t)$ in (16).

$$i(t) = 1 - \sum_j^J P^2(C_j|t) \qquad (16)$$

Where

$P(C_j|t)$ = the probability that a scenario belongs to class C_j given that it falls into t.

The cross-validation gives a best fit for the value of the cost complexity α. This is was represented as (17)

$N_i(t)$ = number of class i in node

(vii) 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.

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

5.1. V-fold Cross Validation

Testing of the decision tree model is performed against the whole learning sample. 10–fold cross validation technique was employed. 10-fold cross validation was performed by dividing the learning sample into ten portions. The decision tree rules are tested using only one portion (1/10) of the ten portions at a time. The remaining nine portions (9/10) are then used to grow another tree and the error rate of the first and second trees are computed. This is repeated until testing of the model is done on all ten strata of the partitioned learning sample. This was diagrammatically represented as Figure 9.
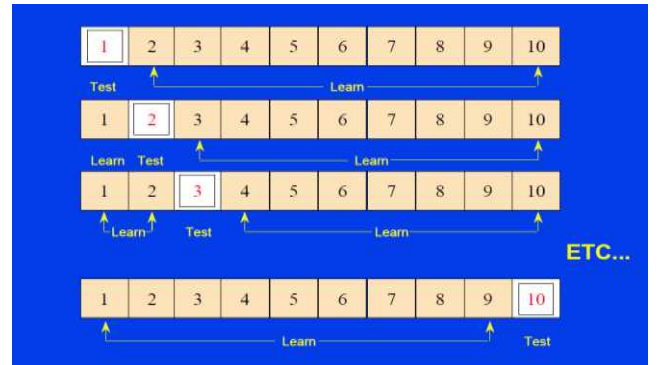

Figure 9. *10-fold cross validation (Adopted from [65])*

Set $\beta_1 = 0$

$$\beta_2 = \sqrt{\alpha_1\alpha_2}$$
$$\beta_2 = \sqrt{\alpha_1\alpha_2} \qquad (17)$$
$$\vdots$$
$$\beta_{m-1} = \sqrt{\alpha_{m-2}\alpha_{m-1}}$$
$$\beta_m = \infty$$

Complexity parameter β with smallest risk is selected as the optimal pruned tree. (For a reliable decision tree, the complexity parameter should be as low as possible.)
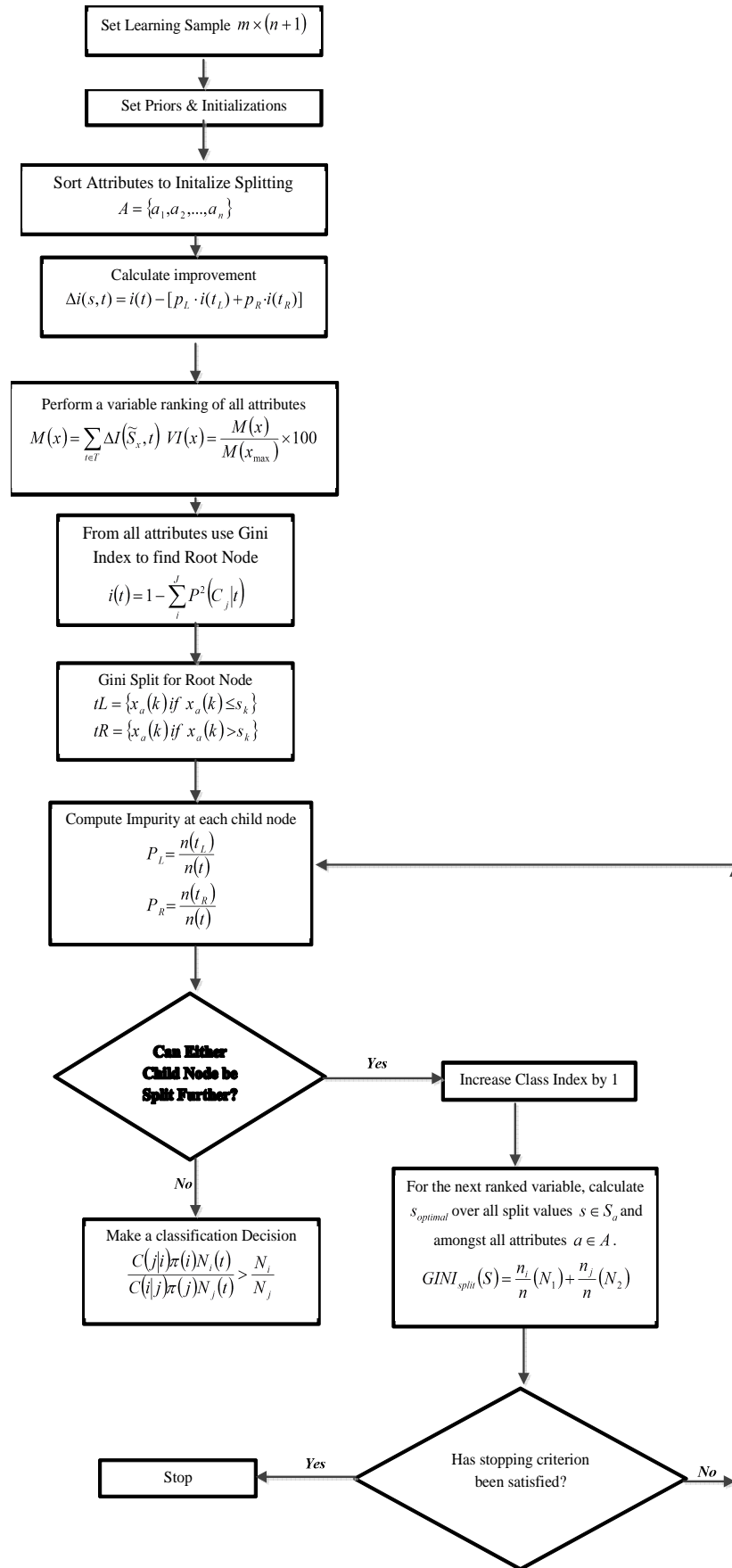
Figure 10. *Decision Tree Growing Algorithm (Courtesy of [44], [45], [46])*

6. Results & Analysis

The hypothesis being presented here is that strategic *PMU* placements would be sufficient to monitor an entire power system network for the purpose of Wide Area Protection & Control (*WAPC*). The split nodes of the *DT* model identify the important variables that predict a power swing. Therefore, placement would be towards monitoring these specific important variables. The primary splitters present the best *PMU* placement location while the surrogate splitters present the second best alternative. The latter is so because the surrogate splitters imitate the primary splitter's characteristics. Moreover, since surrogate splitters handle missing data, in application to a power network they would thus make a provision for *PMU* placements with topology change (line/generator outage) considerations.

The execution time of the maximal *DT* model was 0.1 seconds. Though this is desirable, lesser execution time would be more efficient. Thus the need for pruning to have an optimal *DT* model would reduce this execution time. This is because the optimal tree has fewer nodes and lesser variables. Although the relative error of the optimal *DT* model increases from 0.000 to 0.003, the misclassification error of cases is less than 0.1%. Though this error is minimal, terminal nodes of a pure classification are selected for the prediction. This is to ensure that the misclassification error doesn't have a role in the prediction of the power swings.

According to the splitters presented in Figure 11 optimal placement of *PMUs* for the purpose of Wide Area Protection & Control (*WAPC*) would be on buses 16, 22, 23, 31 and 35. Placement on these buses would monitor bus parameters and incident lines terminating on these buses. The *EHV* lines 6-22, 7-23, 10-31 and 2-35 link the generators 6, 7, 10 and 2 respectively. Therefore, these 5 buses are the critical observation points for monitoring developing transients in the *IEEE 39* bus system as tabulated inTable 1. The electrical parameters to be monitored as represented by the splitter nodes in Figure 11are generator rotor angles, generator speed deviation, voltage phasor angle, positive sequence current magnitude and the active & reactive current magnitudes. Therefore, only these aforementioned parameters will be updated in the *DT* model.
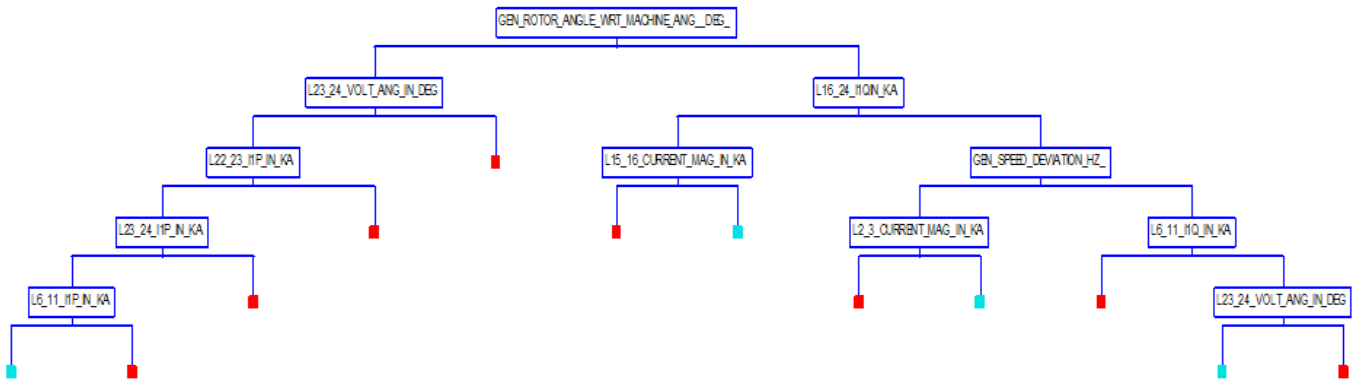

Figure 11. *Optimal DT Tree Splitters*

Table 1.*Strategic PMU Placement Positions on the IEEE 39 Bus System*

PMU Placement Bus Number	Incident Lines	EHV Line	Generator	Load Bus
16	16-15, 16-17, 16-21 and 16-24	Nil	Nil	15, 16, 21 and 24
22	22-6, 22-21 and 22-23	22-6	6	21 and 23
23	23-7, 23-22 and 23-24	23-7	7	23
31	31-10, 31-25, 31-32 and 31-39	31-10	10	25, 31 and 39
35	35-2, 35-13, 35-34 and 35-36	35-2	2	36

However, the disadvantage emanating from the aforementioned limitation is that in case a transmission line of the important splitter variable is tripped, the *PMU* would not have information about this line. Because of this tripping/removal of a critical line is a major change in the network's topology, updating of the *DT* model may fail. The performance accuracy of identifying the splitting criterion for stability enhancement after training and testing were presented in Table 1.

Therefore, this paper demonstrated the capabilities and gratification of synchronized phasor measurements (using *PMUs*) in power system Wide Area Monitoring Protection and Control (*WAMPAC*). It brings forth the relative worth of *PMU* data for the purposes of a data mining exercise in that the *PMUs* will provide a comprehensive acquisition of synchronized wide area phasor measurements. Geographically dispersed local measurements of the power system are synchronized to a common time reference and transmitted to a central processing unit in the control centre. The *PMU* measurements are also μs accurate, therefore more reliable data for various analyses.

7. Conclusion

The theory behind Wide Area Monitoring Protection and Control (*WAMPAC*) was brought into concept. Wide Area

Monitoring (*WAM*) forms the basis of the Wide Area Protection & Control. Synchronized wide area measurements enhance the integrity of measurement data for various power system analyses. The Phasor Measurement Unit (*PMU*) was advocated for in this paper for the purposes of *WAMs*. This is because of its effortless capabilities in synchronizing wide area measurements, high measurement sensitivity (μs accurate), high sampling rate, time stamping and the filtering of electrical noise capabilities.

The utility of the *PMUs* in a data mining is therefore realised through its high sampling rate and μs calibration of measurement data. Strategic placement of *PMUs* therefore achieves *WAM* for quasi or real-time analyses of the power system's functions.

The applications of data mining techniques are explored, where power system analysers will be able to realise the benefits of learning from past recorded data and therefore adopt this method in discovering meaningful information from their archives of recorded data. A more meaningful relationship of events can be traced from a larger set of data, portraying relationships of outcomes of events.

Noteworthy is that some power system variables which are thought to be major contributing factors to detecting a particular power system event may not unfold to be so. *DT* models are able to show the ranking of various variables. This focuses the attention of protection engineers and power system operators to these attributes and scenarios and thereby helps to develop efficient and effective solutions.

The contribution of this paper is that it provided a methodology for identifying strategic placement locations for *WAM* devices for the purpose of designing a *WAP* scheme. The placements strategies to enhance power system stability, monitoring a few identified important variables that will predict a power swing.

References

[1] R. E. Bratton, "Transfer-Trip Relaying Over A Digitally Multiplexed Fiber Optic Link," *IEEE Transactions on Power Apparatus and Systems*, vol. 103, no. 2, pp. 403 –406, Feb. 1984.

[2] S. Ward, T. Dahlin, and B. Ince, "Pilot protection communication channel requirements," in *Protective Relay Engineers, 2004 57th Annual Conference for*, 2004, pp. 350 – 391.

[3] C. Dillow, "Record-Breaking New Fiber Optic Cables Transmit 100 Terabits Per Second," *Popular Science*, 29-Apr-2011. [Online]. Available: http://www.popsci.com/technology/article/2011-04/two-different-fiber-optic-technologes-top-100-terabit-second-speeds-fastest-ever.

[4] M. Afzali and A. Esmaeilian, "A novel algorithm to identify power swing based on superimposed measurements," in *11th International Conference on Environment and Electrical Engineering*, Venice Italy, 2012, pp. 1109–1113.

[5] J. H. Chow, L. Beard, M. Patel, P. Quinn, A. Silverstein, D. Sobajic, and L. Vanfretti, "Guidelines for Siting Phasor Measurement Units," *North American SynchroPhasor Initiative Research Initiative Task Team (RITT) Report*, no. NASPI RITT Report Version 8, Jun. 2011.

[6] Y. Wang, W. Li, P. Zhang, B. Wang, and J. Lu, "Reliability Analysis of Phasor Measurement Unit Considering Data Uncertainty," *IEEE Transactions on Power Systems*, vol. 27, no. 3, pp. 1503 –1510, Aug. 2012.

[7] B. V. Iii, M. Ieee, A. Apostolov, and F. Steinhauser, "Testing of PMU Based Wide Area Monitoring and Recording Systems," in *Communication*, pp. 1–5.

[8] E. Price, "Practical Considerations for Implementing Wide Area Monitoring , Protection and Control," *System*, pp. 36–47.

[9] 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.

[10] C. Rakpenthai, S. Premrudeepreechacharn, S. Uatrongjit, and N. R. Watson, "An Optimal PMU Placement Method Against Measurement Loss and Branch Outage," *IEEE Transactions on Power Delivery*, vol. 22, no. 1, pp. 101–107, Jan. 2007.

[11] F. Aminifar, A. Khodaei, M. Fotuhi-Firuzabad, and M. Shahidehpour, "Contingency-Constrained PMU Placement in Power Networks," *IEEE Transactions on Power Systems*, vol. 25, no. 1, pp. 516–523, Feb. 2010.

[12] B. Gou, "Generalized Integer Linear Programming Formulation for Optimal PMU Placement," *IEEE Transactions on Power Systems*, vol. 23, no. 3, pp. 1099–1104, Aug. 2008.

[13] 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.

[14] R. Sodhi, S. C. Srivastava, and S. N. Singh, "Multi-criteria decision-making approach for multi-stage optimal placement of phasor measurement units," *IET Generation, Transmission & Distribution*, vol. 5, no. 2, p. 181, 2011.

[15] 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.

[16] S. Chakrabarti and E. Kyriakides, "Optimal Placement of Phasor Measurement Units for Power System Observability," *IEEE Transactions on Power Systems*, vol. 23, no. 3, pp. 1433–1440, Aug. 2008.

[17] S. Chakrabarti, E. Kyriakides, and D. G. Eliades, "Placement of Synchronized Measurements for Power System Observability," *IEEE Transactions on Power Delivery*, vol. 24, no. 1, pp. 12–19, Jan. 2009.

[18] N. M. Manousakis, G. N. Korres, and P. S. Georgilakis, "Taxonomy of PMU Placement Methodologies," *IEEE Transactions on Power Systems*, vol. 27, no. 2, pp. 1070–1077, May 2012.

[19] 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.

[20] J. S. Thorp, A. G. Phadke, S. H. Horowitz, and M. M. Begovic, "Some applications of phasor measurements to adaptive protection," *IEEE Transactions on Power Systems*, vol. 3, no. 2, pp. 791–798, May 1988.

[21] B. Kasztenny and M. Adamiak, "Implementation and Performance of Synchrophasor Function within Microprocessor Based Relays." General electric Multilin, Sep-2005.

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

[23] 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.

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

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

[26] 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.

[27] G. K. Venayagamoorthy, "Intelligent sense-making for smart grid stability," in *Power and Energy Society General Meeting*, Detroit, Michigan, USA, 2011, pp. 1–3.

[28] 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.

[29] 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.

[30] 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.

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

[32] 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.

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

[34] 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.

[35] 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, 1993. APSCOM-93*, 1993, pp. 47 –53 vol.1.

[36] 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*, Pittsburgh, PA, 2008, vol. 22, pp. 1–1.

[37] 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.

[38] 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.

[39] 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].

[40] 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*, Birmingham, UK, 2012, pp. 134–134.

[41] 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)*, Manchester, 2009, pp. 358–361.

[42] 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*, Singapore, 2000, vol. 2, pp. 1075 –1080 vol.2.

[43] 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.

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

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

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