Regression Tree for Stability Margin Prediction Using Synchrophasor Measurements

Ce Zheng, *Student Member, IEEE*, Vuk Malbasa, and Mladen Kezunovic, *Fellow, IEEE*

***Abstract*—A regression tree-based approach to predicting the power system stability margin and detecting impending system event is proposed. The input features of the regression tree (RT) include the synchronized voltage and current phasors. Modal analysis and continuation power flow are the tools used to build the knowledge base for offline RT training. Corresponding metrics include the damping ratio of the critical oscillation mode and MW-distance to the voltage collapse point. The robustness of the proposed predictor to measurement errors and system topology variation is analyzed. The optimal placement for the phasor measurement units (PMUs) based on the importance of RT variables is suggested.**

***Index Terms*—Decision trees, phasor measurement units, power system stability, regression analysis.**

I. Introduction

CONVENTIONAL time domain simulation based on system modeling has been used as the primary tool to analyze power system stability. This method is straightforward and accurate as long as adequate system model and measurements are used. However, two challenges have prevented the simulation method from being used for real-time applications: 1) it is computationally intensive; 2) it raises concerns over approximate analysis results when a simplified model is used. As the importance of real-time stability monitoring and early detection of system events has been increasingly emphasized recently, an alternate approach based on the decision tree (DT) [1] was applied by previous authors and encouraging results have been reported [2]–[11].

The DT method was first introduced to the field of power system by Wehenkel *et al.* [2]. It was used to conduct the online transient stability assessment in [2]–[4]. Later the DT approach was applied to the problem of real-time security assessment [5]–[9]. In [10], Diao *et al.* used DT for $N-k$ contingency analysis and security boundary identification. Teeuwsen *et al.* deployed the genetic algorithm to search for the best DT input features for oscillatory stability region prediction [11].

Manuscript received April 29, 2012; revised August 12, 2012; accepted September 23, 2012. Date of publication October 15, 2012; date of current version April 18, 2013. This work was supported in part by the Power System Engineering Research Center (PSerc) under the project S-44 titled "Data Mining to Characterize Signatures of Impending System Events or Performance Using PMU Measurements" and in part by Texas A&M University. Paper no. TPWRS-00434-2012.

The authors are with the Department of Electrical and Computer Engineering, Texas A&M University, College Station, TX 77843-3128 USA (e-mail: zhengce@neo.tamu.edu; vmalbasa@tamu.edu; kezunov@ece.tamu.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TPWRS.2012.2220988

The concept of decision tree comprises the classification tree and regression tree. While in previous works classification trees have been extensively studied to group an operating point (OP) into one of several pre-defined stability categories, the use of regression trees (RT) to predict the stability margin, i.e., how far the system is away from a possible instability event, has not yet been fully studied. With respect to its online use, the areas that remain unexplored include how fast the RT can process PMU measurements, how well the RT can deal with measurement errors, and how robust the RT is to the system topology changes.

A review of literature reveals that several other data mining tools such as multi-linear regression, neural networks, and support vector machine have been used to evaluate the system stability status [12]–[14]. In [15] and [16], Kamwa *et al.* showed that there is a trade-off between data mining model accuracy and transparency. Compared with some "black-box" tools, the RT piece-wise structure provides system operators with a clearer cause-effect relationship of how the system variables lead to the onset of an instability event. Using RTs it is possible to identify the critical variables and thresholds that need to be analyzed to gain insight into the stability margin of a system.

The objective of this paper is to examine the performance of RT in predicting power system voltage and oscillatory stability margins. It starts with the theoretical background of stability margin and instability event detection. After that the RT-based stability evaluation scheme is proposed. The robustness of the RT model to PMU measurement errors as well as the system topology variations is studied. In the end, a combined bus ranking methodology is proposed and the optimal placement for PMU installation is derived based on the importance of RT variables.

II. Problem Formulation

Two important aspects of system operational performance, namely oscillatory stability and voltage stability, are targeted for monitoring. First the definition of an instability event is revisited:

- Oscillatory stability is related to Hopf bifurcation. An instability event occurs whenever, following a small disturbance, the damping torques are insufficient to bring the system to a steady-state operating condition which is identical or close to the pre-disturbance condition.
- Voltage stability is related to saddle-node bifurcation. Voltage instability occurs when the load attempts to step beyond the capability of the combined transmission and generation system [17].

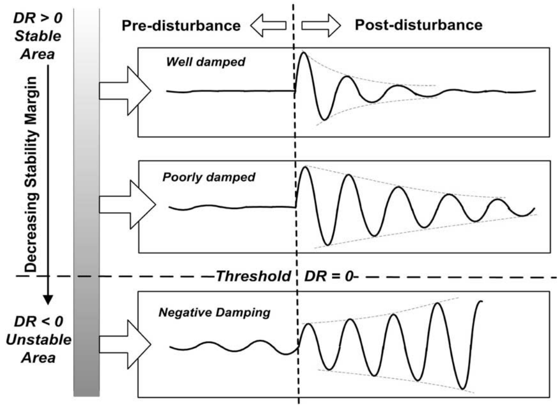

Fig. 1. Use of damping ratio as oscillatory stability margin.

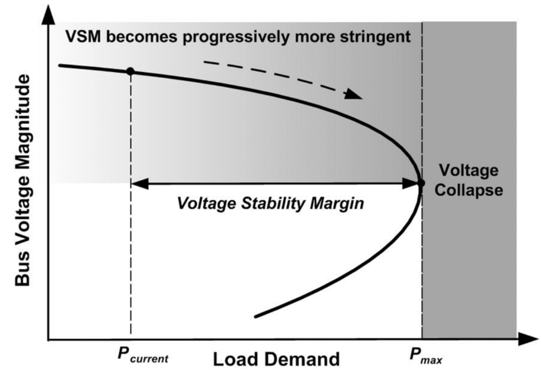

Fig. 2. Voltage stability margin.

A. Oscillatory Stability Margin (OSM)

The oscillatory stability is usually evaluated through analysis of the system nonlinear differential algebraic equations (DAE):

$$\begin{cases} \dot{x} = f(x, y, u) \\ 0 = g(x, y, u) \end{cases} \tag{1}$$

where x is the state vector, y is the output vector, and u is the control vector. The DAEs in (1) are formulated by detailed modeling of each network component.

A linearization of (1) will result in

$$\begin{cases} \Delta\dot{x} = A\Delta x + B\Delta u \\ \Delta y = C\Delta x + D\Delta u \end{cases}. \tag{2}$$

From the theory of modal analysis, each pair of complex conjugate eigenvalues of matrix A stands for an oscillation mode. Further decomposing the A matrix will get

$$A = \Phi\Lambda\Psi \tag{3}$$

where Λ represents the diagonal eigenvalue matrix, Φ and Ψ are the left and right eigenvector matrices. For the ith oscillation mode with the following conjugate pair:

$$\lambda_{\mathrm{i}} = \sigma_{\mathrm{i}} \pm \mathrm{j}\omega_{\mathrm{i}}. \tag{4}$$

The mode damping ratio (DR) is calculated as

$$\xi_{\mathrm{i}} = \frac{\sigma_{\mathrm{i}}}{\sqrt{\sigma_{\mathrm{i}}^2 + \omega_{\mathrm{i}}^2}}. \tag{5}$$

The oscillation modes that carry significant amount of energy but with insufficient DR are critical among all modes and need to be closely monitored. Occurrence of an instability event is possible when a poorly damped mode is excited by a small or large disturbance. In this work the DR of critical oscillation mode is used as the OSM indicator. As shown in Fig. 1, the OSM becomes progressively more stringent as the value of critical mode DR decreases.

The damping ratio is not an index from the parameter space, so strictly speaking it may not be proper to term it as "margin". In this work DR is selected as the OSM indicator in the sense that it provides smooth movement trajectory, clear partition between stable/unstable states, and an explicit distance from unstable point.

In this work regression trees are trained to emulate system behavior and predict the DR values.

B. Voltage Stability Margin (VSM)

The variation of load bus voltage magnitude with different load demand is plotted as the P-V curve shown in Fig. 2. The MW-distance from the current operating point to the voltage collapse point ("Knee" point), where the load demand equals the maximum deliverable power, provides a reasonable measure of system voltage stability margin. The VSM referred here corresponds to system long-term voltage stability [18], which cannot be used to capture the short-term voltage stability.

The focus is to find the voltage collapse point. In this work the idea of continuation power flow (CPF) proposed in [19] is explored. Assuming a constant load power factor, slowly increasing load demand will push the operating point from the base case towards the collapse point along the P-V curve. The voltage collapse point is achieved when the load flow Jacobian becomes singular. System voltage stability margin is hereby expressed as [13]

$$P_{\text{distance}} = P_{\max} - P_{\text{current}} \tag{6}$$

where P_{distance} represents the MW-distance between current OP and the collapse point, $P_{\max}$ is the maximum deliverable power, and P_{current} is the current load active power demand. The proposed procedure for voltage stability margin prediction is as follows:

1) Generate n different OPs.
2) For each OP, determine the maximum deliverable power by means of the CPF technique.
3) Calculate the voltage stability margin for the ith OP using the following index:

$$\mathrm{VS}_{\text{margin}}^{i} = \frac{P_{\max}^{i} - P_{\text{current}}^{i}}{P_{\max}^{i}} \times 100\%. \tag{7}$$

4) Train the RT offline using selected features from the n OPs and their corresponding $\mathrm{VS}_{\text{margin}}$.
5) Use the trained RT to predict VSM in real time.

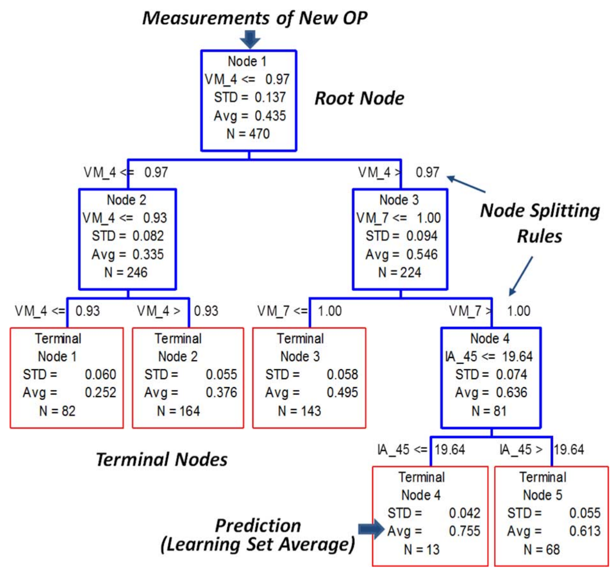

Fig. 3. Example of RT model structure.

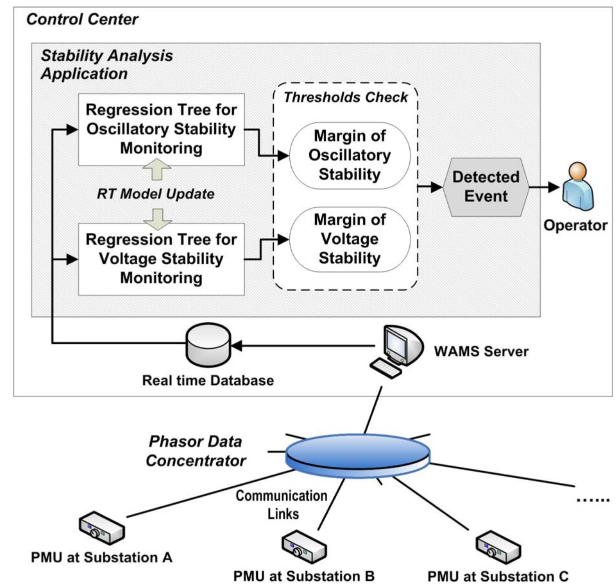

Fig. 4. Proposed framework of the RT-based stability margin prediction and event detection.

III. Predicting Stability Margins Using RT

A. Regression Tree Method

Compared with the traditional time domain simulation approach that requires full model computation each time a new OP has emerged, the advantage of RT method lies in its simplified model structure and fast OP analysis facilitated by fewer required inputs. The method is particularly appealing because the RT uses a model which makes the results easy to interpret and replicate. Fig. 3 provides a simple example of RT structure. The unfolding OP is related to its stability margin through a unique top-down path. The splitting rule at each node that belongs to a given path represents an operational threshold. Based on the combination of splitting rules along the path, preventive and corrective control strategies could be formulated and initiated.

In regression analysis, a case consists of instance (x, y) where x is the vector of attributes and y is the target. The relationship between x and y is usually described by a regression function, through which it is possible to estimate how the target y changes when x is varied. In our proposed approach, the regression function is replaced by a binary tree structure, where x is the synchrophasor measurements and y is the system stability margin, i.e., the damping ratio or MW-distance. The commercial software CART [20] is used to develop OSM-RT and VSM-RT used for evaluating oscillatory and voltage stability margins, respectively.

The approach to build a RT entails three steps: 1) tree growing using learning dataset; 2) tree pruning using test dataset or cross-validation; 3) selection of the best pruned tree. Experimental tests show that there is a trade-off between the tree complexity and its accuracy: a small-sized tree cannot capture enough system behavior, and a large-sized tree usually leads to imprecise prediction due to its over-fitting model. In this work the rule of minimum cost regardless of size to search for the best pruned RT commensurate with accuracy is adopted. The complexity cost parameter in CART has been set to equal to zero. The RT growing, node splitting, tree pruning and optimal tree selection algorithms are detailed in [1].

B. Proposed Approach

The proposed framework for RT-based stability margin prediction and event detection is shown in Fig. 4. PMU measurements from different substations are collected and time-aligned by the phasor data concentrator (PDC). The synchrophasor measurements are then delivered to the wide area measurement system (WAMS) server located at the central control facility. At the control center operator room, the RTs for monitoring OSM (OSM-RT) and VSM (VSM-RT) are trained and updated periodically. The PMU data of an upcoming OP is dropped down the respective tree until it reaches a terminal node. Then the predicted stability margin is the average value of the learning set samples falling into that terminal node. Any OP with insufficient stability margin will be detected immediately by checking corresponding thresholds. Operators are alerted with the possible event and preventive control strategies can be initiated in a timely manner.

IV. Performance Examination

The performance of the RT-based predictor is evaluated using the IEEE 9-bus [21] and 39-bus (New England) [22] test systems. These two systems are known for their realistic configurations and efficacy in testing stability-related applications.

A. Knowledge Base Generation

The offline training of a RT using empirical data from a knowledge base is the first and most important step. The knowledge database is composed of known system OPs and corresponding system stability margins.

Both the voltage and oscillatory stability are closely related to the load/generation composition of a power system, and their increase/decrease trend at a certain system snapshot [23]. If the load/generation composition varies, different OPs are formed.

The change in the load demand and generation output can be described as

$$P_G = P_G^0 + \Delta P_G \quad Q_G = Q_G^0 + \Delta Q_G$$
$$P_L = P_L^0 + \Delta P_L \quad Q_L = Q_L^0 + \Delta P_L \times Q_L^0 / P_L^0 \qquad (8)$$

where P_G and Q_G are active/reactive power outputs of all the generators except the slack bus generator, P_L and Q_L are vectors of active/reactive power delivered to the loads. Superscript 0 represents base case OP. The vectors ΔP_G, ΔQ_G, ΔP_L, and ΔQ_L stand for the variations in power.

In this work, the commercial software PSS/E is used for iteratively solving load flows, and deriving characteristic matrix A at different OPs through numerical perturbation. Python and MATLAB programs are developed to automate the PSS/E simulations, perform modal analysis, conduct the CPF-based voltage stability analysis, compute stability margins, and establish the knowledge base. The pseudo-code for knowledge base creation is illustrated below.

Pseudo-code for Knowledge Base Generation

---

1. Initialize PSS/E in Python. Import system model parameters:

   $\mathrm{Number\ of\ Generation\ Buses} = \mathrm{i}$,
   $\mathrm{Number\ of\ Load\ Buses} = \mathrm{j}$

   $\mathrm{Number\ of\ buses\ with\ shunt\ capacitor} = \mathrm{k}$

2. Let $\mathrm{u}(\mathrm{u} \in \mathrm{N})$ be the iteration index with a step change of $C_{G/L/S}\%$

   Suppose G_1 is slack bus. Repeat:

   for $\mathrm{A}_2 = 0 \rightarrow \mathrm{u}_2$ do

   Scale the output of G_2 to: $P_{G2} = P_{G2}^0(1 + A_2 \times C_{G2}\%)$

   ...

   for $\mathrm{A_i} = 0 \rightarrow \mathrm{u_i}$ do

   Scale the output of G_i to: $P_{Gi} = P_{Gi}^0(1 + A_i \times C_{Gi}\%)$

   for $\mathrm{A_{i+1}} = 0 \rightarrow \mathrm{u_{i+1}}$ do

   Scale load 1 to: $P_{L1} = P_{L1}^0(1 + A_{(i+1)} \times C_{L1}\%)$

   ...

   for $\mathrm{A_{i+j}} = 0 \rightarrow \mathrm{u_{i+j}}$ do

   Scale load j to: $P_{Lj} = P_{Lj}^0(1 + A_{(i+j)} \times C_{Lj}\%)$

   for $\mathrm{A_{i+j+1}} = 0 \rightarrow \mathrm{u_{i+j+1}}$ do

   Scale shunt 1 to: $Q_{S1} = Q_{S1}^0(1 + A_{(i+j+1)} \times C_{S1}\%)$

   ...

   for $\mathrm{A_{i+j+k}} = 0 \rightarrow \mathrm{u_{i+j+k}}$ do

   Scale shunt k to: $Q_{Sk} = Q_{Sk}^0(1 + A_{(i+j+k)} \times C_{Sk}\%)$

   Solve the load flow at: $\{P_{G2}, \ldots, P_{\mathrm{Gi}}, P_{L1}, \ldots, P_{\mathrm{Lj}}, Q_{S1}, \ldots, Q_{\mathrm{Sk}}\}$

   If this OP is unsolvable: eliminate

   Oscillatory Stability Analysis:

   Import system model dynamic data. Derive the A matrix.

   Voltage Stability Analysis:

   Derive the voltage collapse point via continuation-based method

   Export computed features of current OP

   End Loops

3. Repeat: for i $= 0 \rightarrow$ number of OPs do

   Modal analysis of A matrix using (3)–(5): $DR(\zeta_i)$

   Compute voltage stability index using (6)–(7): $\mathrm{VS}_{\mathrm{margin}}^i$

   Export computed stability margins

   End Loop

The power supply at generation buses, demand at load buses, and the output of shunt capacitors were systematically varied. A total of 1071 OPs with corresponding OSMs, and 1153 OPs with corresponding VSMs have been produced for the 9-bus system. The number of records generated for the 39-bus system knowledge base is 4276 and 3664, respectively.

In addition, in this work the generator active/reactive power limits have been taken into account to reflect the practical stability margin. This has significant impact on the computation of VSM: when the load demand increases, a feasible load flow solution may not exist due to the limited generation capacity, even before the maximum loadability of the transmission system is reached. Therefore the derived $P_{\max}$ may be somewhere on the top half of the PV curve before the "Knee" point shown in Fig. 2.

In order to build a sufficiently large knowledge base, in this work two stopping criteria are followed:

a) Each generator/load/shunt should be varied at least 4 times ($u \geq 4$) and the total variation should be at least 30% of the base value ($u \times C_{G/L/S} \geq 30$). The goal is to capture the most system behavior from the problem space.
b) The RT training and testing accuracy converges. The R^2 is used to measure the prediction accuracy and will be detailed in next section.

The trajectory of the 39-bus system stability margin is shown in Fig. 5. Corresponding stability thresholds are shown as the flat planes dividing each margin space into two halves: an instability event will be immediately identified in the top half. For this power system the voltage stability threshold is put at $\mathrm{VS}_{\mathrm{margin}} = 30\%$. This value can be further adjusted according to the real-time operational needs.

As it can be observed from Fig. 5, a high imbalance in size between the stable and unstable cases exists. This is a very practical issue in power system operation since most of the time the system is in its stable state. From the classification point of view,

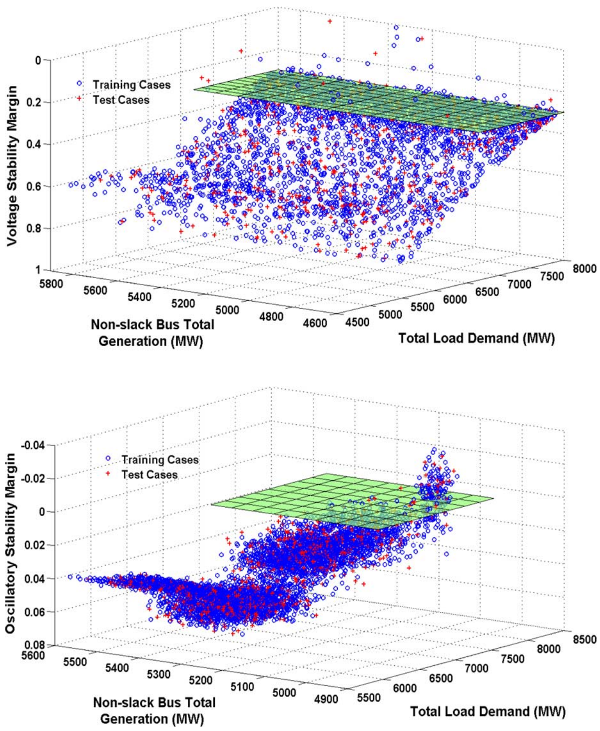

Fig. 5. Trajectory of voltage and oscillatory stability margins of the IEEE 39-bus (New England) test system.

compared with some other data mining tools that do not perform well when dealing with unbalanced data, the decision tree implemented in the CART software has the property of assuring that every class is treated equally regardless of its size. This is achieved by specifying the *Prior* of each class. From the regression point of view, there is no need to set *Priors* because each case will be treated as an equal point on the continuous stability margin space. Because of the least squares loss function for regression, as implemented in CART, large mistakes are penalized more than smaller ones, thus large errors at any OP are emphasized be they on the stable or unstable part of the stability margin space. Once the relationship between input and output is identified, the regression model defines a mapping of an OP to its stability margins regardless of the state/class to which the OP belongs.

B. Features Available to RT for Prediction

With respect to the inputs of a decision tree, different feature combinations may result in different data mining accuracies. In order to accelerate the prediction process, it is desirable to use the least features as RT inputs and meanwhile keep an acceptable level of overall prediction accuracy. This requires the selected features to capture as much system behavior as possible.

The objective of this work is to explore a way that directly makes use of synchrophasor measurements for stability analysis. Therefore only the basic measurements from a PMU are considered. Assuming all buses are installed with PMUs, the involved input features are as follows:

- VM_i and VA_i: positive sequence voltage magnitude and phase angle at Bus i;
- IM_i_j and IA_i_j: positive sequence current magnitude and phase angle from Bus i to Bus j.

C. Offline Training and New Case Testing

Each knowledge base is split into two independent data sets: 80% of the records are randomly selected for training of OSM-RT and VSM-RT; the remaining 20% of the records will serve the purpose of RT testing. The 10-fold *cross validation* method is adopted to grow the RT in CART. In experiments because of the random nature of the splitting process, slight differences may occur between the performances of each derived RT. Therefore in this work, the process of knowledge base splitting, tree training and testing has been replicated 10 times until the mean and standard deviation of RT accuracy become stable.

In contrast with a classification tree for which the accuracy could be directly derived from the misclassification rate, the performance of a regression tree is measured through a statistical index, termed *Residuals Squared Error* (R^2) [24]. We report the accuracy of a RT model as follows:

$$R^2 = 1 - \frac{\sum_{(x_i,y_i)\in TS}[y_i - d(x_i)]^2}{\sum_{TS}(y_i - \bar{y}_{\text{root}})^2} \tag{9}$$

where TS is the set of training samples, x_i is input, y_i is the actual stability margin, $d(x_i)$ is the RT predicted value, and y_{root} is the mean of y_i in the tree root node.

In general the closer the value of R^2 is to 1, the better the prediction is. However in practice, how good an R^2 is depends on the particular application and the way it is measured [25]. Experimental results from this work show that a quite acceptable value of $R^2 > 0.90$ can be achieved.

Sometimes the R^2 alone may not be sufficient, especially in the case when the typical difference between values predicted by RT and the actual stability margins is desired. Therefore another measure, the *root-mean-square* (RMS), is utilized:

$$\text{RMS} = \sqrt{\frac{\sum_{i=1}^{n}[y_i - d(x_i)]^2}{n}} \tag{10}$$

where n is the number of test cases. The numerator stands for the sum of squared deviations of the actual stability margins around the RT predictions. The value of RMS error depends on the base magnitude of the target stability margin to be predicted. In the proposed scheme, a typical value of OSM is in the range of -0.01 to 0.1, and the VSM is usually ranging from 0.05 to 1.0. Hence the RMS errors of VSM-RT are usually several times larger than that of the OSM-RTs.

Once the training is complete, the derived RTs were evaluated using the unseen test cases. Much more emphasis must be put on the accuracy of new case testing because, for real-time applications, a predictive model which cannot fit the unseen system behavior well is unacceptable, even if high accuracy is obtained during the offline training, as it lacks generalization power. The corresponding training and new case testing accuracy is summarized in Table I. In addition, the results of new case testing were reported separately in terms of *Security Test* and *Reliability Test*. While the security test examines how well the stable OPs are predicted, the reliability test checks if all unstable OPs are correctly identified.

The prediction for 300 new OPs of the 39-bus system is shown in Fig. 6. The RT-based approach has exhibited encouraging capability for system stability margin prediction.

TABLE I
PERFORMANCE OF THE REGRESSION TREES

System	Oscillatory Stability Margin (OSM-RT)				
	Train R^2	Unseen OPs Overall Accuracy		Reliability and Security Test (RMS)	
		R^2	RMS	Reliability	Security
9-bus	**0.9984**	**0.9858**	**0.0023**	**0.00083**	**0.00235**
39-bus	**0.9617**	**0.9519**	**0.0034**	**0.00386**	**0.00328**
System	Voltage Stability Margin (VSM-RT)				
	Train R^2	Unseen OPs Overall Accuracy		Reliability and Security Test (RMS)	
		R^2	RMS	Reliability	Security
9-bus	**0.9928**	**0.9791**	**0.0184**	**0.03357**	**0.01480**
39-bus	**0.9941**	**0.9694**	**0.0211**	**0.02736**	**0.01965**

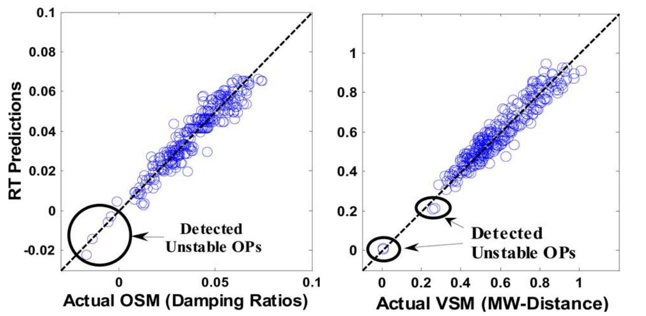

Fig. 6. RT predicted margins versus the actual stability margins of the IEEE 39-bus system. Left: OSM-RT performance; Right: VSM-RT performance.

V. Application to a Larger System

The RT-based predictive model has been applied to the Western Electric Coordinating Council (WECC) equivalent system shown in Fig. 7 [26]. This network consists of 179 buses, 29 generators, 42 shunts, and 104 loads.

The same methodology of creating the knowledge base for the 9-bus and 39-bus systems is adopted. In addition, two practical issues have been considered: 1) the real/reactive power output limit of each generator is more stringent in this larger system and should be complied with strictly; 2) it is computationally too expensive to generate the database by varying only one component each time. For instance, if the iteration index u is set to be 4, a total of 4^{175} OPs will need to be analyzed. It may be more practical to group the loads and generators according to their geographical locations. Seven areas are formed and it is assumed that the loads/generators within each area will increase/decrease at the same rate.

A total of 12 572 records have been generated for the OSM-RT and 15 303 records for the VSM-RT. The impact of the size of training set on the performance of resulted RT is examined: 100%, 50%, 20%, 10%, 5%, and 2% of the training cases are used to derive RT for each task. All experiments have been replicated 10 times and the Mean of new case prediction accuracy is summarized in Fig. 8. It clearly shows that the prediction accuracy increased when more cases were used to train the RTs.

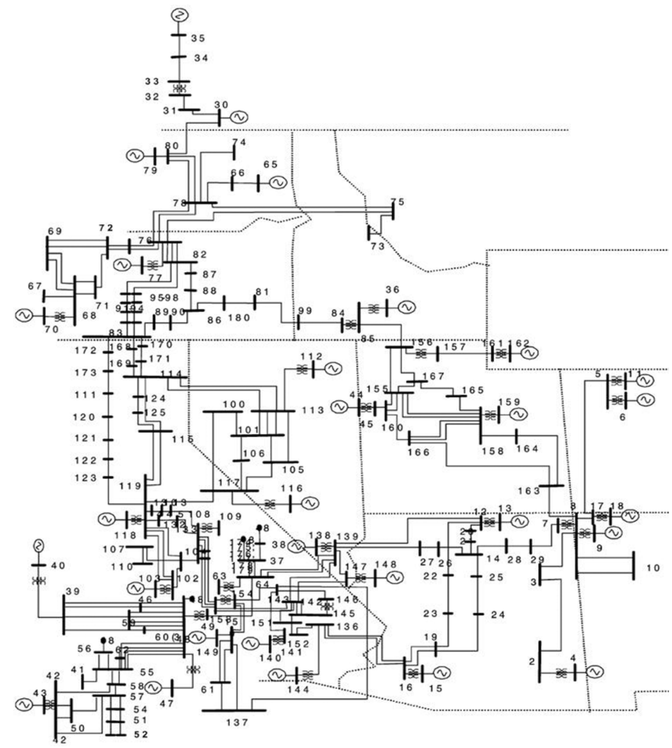

Fig. 7. One-line diagram of the WECC 179-bus equivalent system.

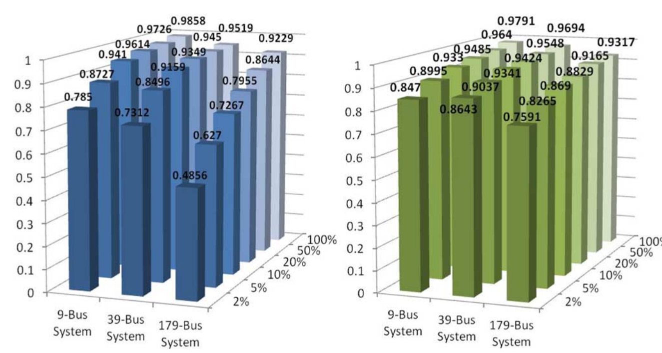

Fig. 8. New case prediction accuracy of RTs trained with differently sized data sets (Left: OSM-RT; Right: VSM-RT).

In order to embed the RT model into an actual online process, three aspects need to be examined and corresponding requirements must be satisfied: 1) eligibility for high speed analysis; 2) robustness to measurement error; 3) capability to accommodate topology change.

A. Data Processing Speed

Traditionally the data used for the stability analysis in electrical utilities are obtained from the Supervisory Control and Data Acquisition (SCADA) system or state estimation functions, which are refreshed on a time scale from several seconds to several minutes. These slowly updated data can only provide limited decision making support under the new situation where fast variations are present at both demand side and supply side. The capability to take advantage of the fast updated PMU data is critical in real-time applications.

TABLE II
COMPUTATIONAL SPEED OF REGRESSION TREES

Type of Regression Models	IEEE 39-bus System		WECC 179-bus System	
	Off-line Training	New Case Prediction	Off-line Training	New Case Prediction
OSM-RT	**36.01 s** (3421 cases)	**about 3 s** (855 cases)	**164.97 s** (10058 cases)	**about 5 s** (2514 cases)
VSM-RT	**31.38 s** (2931 cases)	**about 2 s** (733 cases)	**195.45 s** (12242 cases)	**about 7 s** (3061 cases)

In practice, the PMU measurements are updated very fast, most likely at least 30 times per second. In order to evaluate the system stability status at each snapshot, the processing of PMU data must be less than $1/30 = 0.033$ second.

The data processing speed of RTs is summarized in Table II. The computational time is estimated using the built-in clock of CART executed on an Intel Pentium IV 3.00-GHz CPU with 2 GB of RAM. It can be seen that the derived OSM-RT or VSM-RT can assess 1000 new OPs in less than 4 s for the 39-bus system, and 3000 new OPs in less than 8 s for the WECC 179-bus system. According to the results, the RTs satisfy the speed requirement of real-time applications.

B. Impact of Measurement Errors

The phasor estimation process may introduce errors. PMUs manufactured by multiple vendors can also yield inaccurate readings. In real-time application, the PMU measurement errors of the ith OP can be expressed as

$$\begin{aligned} \mathrm{VM}_i^{\mathrm{meas}} &= \mathrm{VM}_i^{\mathrm{real}} + \Delta\varepsilon_{\mathrm{VM}i} \quad \mathrm{VA}_i^{\mathrm{meas}} = \mathrm{VA}_i^{\mathrm{real}} + \Delta\varepsilon_{\mathrm{VA}i} \\ \mathrm{IM}_i^{\mathrm{meas}} &= \mathrm{IM}_i^{\mathrm{real}} + \Delta\varepsilon_{\mathrm{IM}i} \quad \mathrm{IA}_i^{\mathrm{meas}} = \mathrm{IA}_i^{\mathrm{real}} + \Delta\varepsilon_{\mathrm{IA}i} \end{aligned} \tag{11}$$

where the superscript *real* means actual values of the phasor, and *meas* stands for measured values.

According to the IEEE C37.118 "*Standard for Synchrophasors for Power Systems*" [27], PMUs that are Level 1 compliant with the standard should provide a total vector error (TVE) less than 1%. This implies that the following constraints must be satisfied:

$$\begin{aligned} 1\% &> \left| \frac{\mathrm{VM}_i^{\mathrm{meas}} \angle \mathrm{VA}_i^{\mathrm{meas}} - \mathrm{VM}_i^{\mathrm{real}} \angle \mathrm{VA}_i^{\mathrm{real}}}{\mathrm{VM}_i^{\mathrm{real}} \angle \mathrm{VA}_i^{\mathrm{real}}} \right| \\ 1\% &> \left| \frac{\mathrm{IM}_i^{\mathrm{meas}} \angle \mathrm{IA}_i^{\mathrm{meas}} - \mathrm{IM}_i^{\mathrm{real}} \angle \mathrm{IA}_i^{\mathrm{real}}}{\mathrm{IM}_i^{\mathrm{real}} \angle \mathrm{IA}_i^{\mathrm{real}}} \right|. \end{aligned} \tag{12}$$

Considering (11) and (12), random noise $\Delta\varepsilon$ has been added to the original phasor magnitudes and angles of the WECC 179-bus system knowledge base. In Table III two scenarios were tested. While in both scenarios errors were added to the test cases, it is shown that the RTs trained with measurement error had much better performance than the ones without the error taken into account in the training data set.

C. Impact of Topology Variation

In this paper the robustness of RT to certain system topology changes was examined. The scenarios that were evaluated and RT performances are summarized in Table IV.

TABLE III
PERFORMANCE OF THE 179-BUS REGRESSION TREES CONSIDERING PMU MEASUREMENT ERROR

Type of Regression Models	Add Noise Only to the Test Cases			
	Security Test		Reliability Test	
	R^2	RMS	R^2	RMS
OSM-RT	**0.7906**	**0.00106**	**0.7403**	**0.00121**
VSM-RT	**0.8091**	**0.02785**	**0.7629**	**0.03010**
Type of Regression Models	**Add Noise to Both Training and Test Cases**			
	Security Test		Reliability Test	
	R^2	RMS	R^2	RMS
OSM-RT	**0.9170**	**0.00068**	**0.8994**	**0.00071**
VSM-RT	**0.9266**	**0.01789**	**0.9045**	**0.01940**

TABLE IV
REGRESSION TREE PERFORMANCE UNDER SYSTEM TOPOLOGICAL VARIATIONS

Scenarios of Topology Change	Type	RMS Error of OSM-RT	RMS Error of VSM-RT
Line 8-9 taken out	9 BUS N-1	**0.00880**	$\boxed{\mathbf{0.15481}}$
G10 out of service	39 BUS N-1	**0.00417**	**0.04089**
G10 and Line 26-28 taken out	39 BUS N-2	**0.00726**	$\boxed{\mathbf{0.20702}}$
Line 1 of 76-82 out of service	179 BUS N-1	**0.00337**	**0.03046**
Line 1 of 90-156 out of service	179 BUS N-1	**0.00421**	**0.02654**
Line 1 of 95-98 out of service	179 BUS N-1	**0.00385**	**0.03198**
Line 81-180 out of service	179 BUS N-1	**0.00552**	$\boxed{\mathbf{0.08325}}$
Line 1 of 90-156 and Line 1 of 76-82 out	179 BUS N-2	**0.00473**	**0.04830**
G63-1 and Line 1 of 95-98 out of service	179 BUS N-2	**0.00574**	**0.03792**
G63-1 and Line 81-180 out of service	179 BUS N-2	**0.00588**	$\boxed{\mathbf{0.10736}}$

It can be seen that OSM-RTs were able to provide somewhat acceptable predictions with low RMS errors, even under situations the network topology had changed. On the other hand, VSM-RTs appear to be less robust and the performance varied case by case: the $N-1$ test in the 9-bus system had a significant impact on the VSM prediction due to the small size of the system; acceptable predictions were achieved for the case of generator outage in the 39-bus system; the $N-2$ scenario in the 39-bus system was too severe for the VSM-RT to handle. More case studies were conducted on the 179-bus system VSM-RT: low RMS errors were observed in experiments where slight topology changes are made, such as one of the double-circuit transmission lines out of service.

D. Discussion

1) Ability of RTs to Handle Evolving System Conditions: The problem of how to sustain the prediction accuracy of RT under the evolving system operating conditions is critical for its online

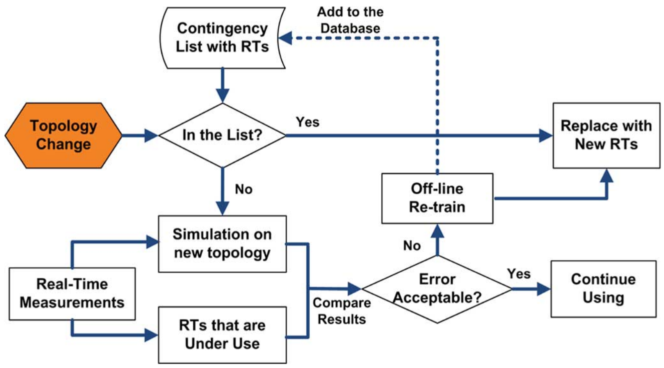

Fig. 9. Scheme for RTs to handle system topology change.

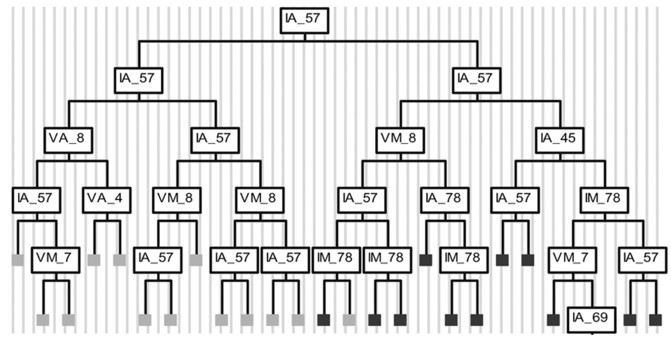

Fig. 10. OSM-RT topology and node splitters of the 9-bus system.

implementation. In fact this is also a problem in all data mining tools. In general, the change of system operating conditions can be categorized into two types:

- the variation of system load/generation patterns;
- the variation of system topology due to contingencies, scheduled maintenance, and system dispatch.

The work reported in Section IV and part of Section V tackles the first type of variation. As illustrated in the knowledge base creation process, the generator/load/shunt has been widely varied in a systematical way to capture the most system behavior from the problem space.

The change in system topology is a major reason that causes a data mining tool to fail in real-time applications. The results shown in Table IV indicate that the RT sensitivity to topology changes becomes less distinct in large sized network and under mild changes in topology. This obviously helps in making RTs useful even under topology changes. It is also observed that RTs are not able to accommodate certain severe contingencies, e.g., the line 81–180 out of service. In the field of data mining and machine learning, the so-called "concept change" describes methodology for dealing with such type of topology variation. A literature search reveals that there is not a generally effective way for the data mining tool to cope with the concept change incrementally, although some work has shown results [28]. Most of the time a re-train using the updated knowledge base is necessary to reflect new topology condition.

2) When and How to Update the RTs: To re-train an RT model whenever it is obsolete is time-consuming and may not satisfy the requirement of seamless on-line monitoring. An effective solution may be to prepare a knowledge base for each of the credible contingencies beforehand, and train a series of candidate RTs accordingly. Fig. 9 shows the proposed scheme. The list of credible contingencies is usually readily available at utility companies. If in online application an unseen contingency occurs and RT fails to provide accurate predictions, a new RT will be trained and deployed. The new contingency scenario and RTs will be added to the historical database. With the increase of contingency scenarios accumulated in database, fewer unseen topology conditions will be encountered. The obsolete models can be quickly replaced by the candidate RTs corresponding to the post-contingency condition.

VI. Optimal Location of PMUs

In previous sections, the RTs were fed with voltage and current phasors measured at all buses. An underlying assumption is that almost every substation is equipped with a PMU. In practice, this is not economically feasible since the installation of PMUs and corresponding telecommunication path is very costly. A reasonable approach may be to install only a limited number of PMUs at the most critical substations. The problem of finding the optimal PMU location is equivalent to selecting the best reduced set of RT input features without a significant degradation in RT performance.

A. Combined Bus Ranking

Ideally, the optimal solution could be obtained through an exhaustive trial and comparison of all possible feature combinations. However it is computationally too involved to do so. In this work we are proposing a different approach and the idea comes from a unique property of the RT model structure. The topology of the OSM-RT derived in Section IV-C is shown in Fig. 10. It is interesting to track the actions of the tree. Each node has been split by an input variable, and the variable is selected as the splitter because it is the most powerful variable among all candidate features that can best split the node. The variables gain credit towards their importance by serving as primary splitters that actually split a node, or as back-up splitters (surrogates) to be used when the primary splitter is missing. By summarizing the variables' contribution to the overall tree when all nodes are examined, the *variable importance (VI)* can be obtained.

To calculate the *VI*, search all splits $s \in S$ on variable x_m at each tree node $t \in T$, and find the split s_m^* that gives the largest decrease in regression R [1], [15]:

$$\Delta R(s_m^*, t) = \max_{s \in S} \Delta R(s, t). \tag{13}$$

Suppose s^* is the best of s_m^*, and $\tilde{s}_m$ is the split on variable x_m that has the best agreement with s^* in terms of partitioning cases, the measure of importance of variable x_m is defined as

$$VI(x_m) = \sum_{t \in T} \Delta R(\tilde{s}_m, t). \tag{14}$$

Fig. 11 shows the computed *VI* for the OSM-RT and VSM-RT of the 9-bus system derived in Section IV-C. The actual measures of importance have been normalized so that the most important variable has a *VI* of 100.

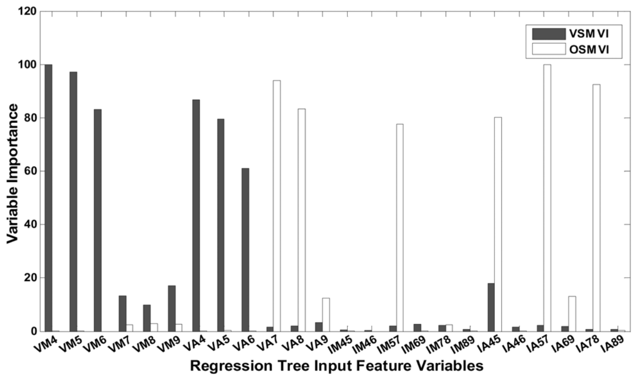

Fig. 11. Variable importance for VSM-RT and OSM-RT.

TABLE V
COMBINED BUS RANKING OF THE 179-BUS SYSTEM

Top Ranked Buses			Lowest Ranked Buses		
Rank	Location	CBR	Rank	Location	CBR
# 1	**Bus 90**	**100.82**	**# 170**	**Bus 162**	**0.31**
# 2	**Bus 100**	**100.23**	**# 171**	**Bus 163**	**0.28**
# 3	**Bus 95**	**38.27**	**# 172**	**Bus 172**	**0.24**
# 4	**Bus 96**	**18.47**	**# 173**	**Bus 168**	**0.12**
# 5	**Bus 97**	**13.99**	**# 174**	**Bus 85**	**0.11**
# 6	**Bus 67**	**12.73**	**# 175**	**Bus 50**	**0.02**
# 7	**Bus 12**	**12.52**	**# 176**	**Bus 92**	**0.02**
# 8	**Bus 11**	**8.48**	**# 177**	**Bus 94**	**0.01**
# 9	**Bus 9**	**8.44**	**# 178**	**Bus 165**	**0.01**
# 10	**Bus 20**	**8.24**	**# 179**	**Bus 171**	**0.00**

The idea of *combined bus ranking* (CBR) is as follows: The overall contribution of each bus to the oscillatory and voltage stability evaluation can be quantified by combining the importance of variables measured at this bus.

Mathematically the CBR of Bus i can be expressed as

$$\mathrm{CBR}_i = \sum_{x\in i} VI_{\mathrm{OSM-RT}}(x) + \sum_{x\in i} VI_{\mathrm{VSM-RT}}(x) \tag{15}$$

where X is the vector of RT input variables, x is the individual variable belong to X, and $VI(x)$ is its importance. By specifying $x \in i$, only the variables measured at Bus i will be counted.

B. Optimal PMU Locations

A ranking list of the bus contributions can be obtained by sorting the CBR values from high to low. The optimal PMU locations will be suggested by selecting the top ranked buses from the list. In this work the CBR of top ranked buses were computed by considering only the primary splitters, because the surrogate variables that appear to be important but rarely split nodes are almost certainly highly correlated with the primary splitters and contain similar information. Once the top ranked buses were selected, the standard *VI* considering both primary and surrogate splitters were used to rank the remaining buses. In Table V, the CBR for the WECC 179-bus system was calculated and top 10 buses are listed. Also shown in the table are the 10 buses with the lowest CBR.

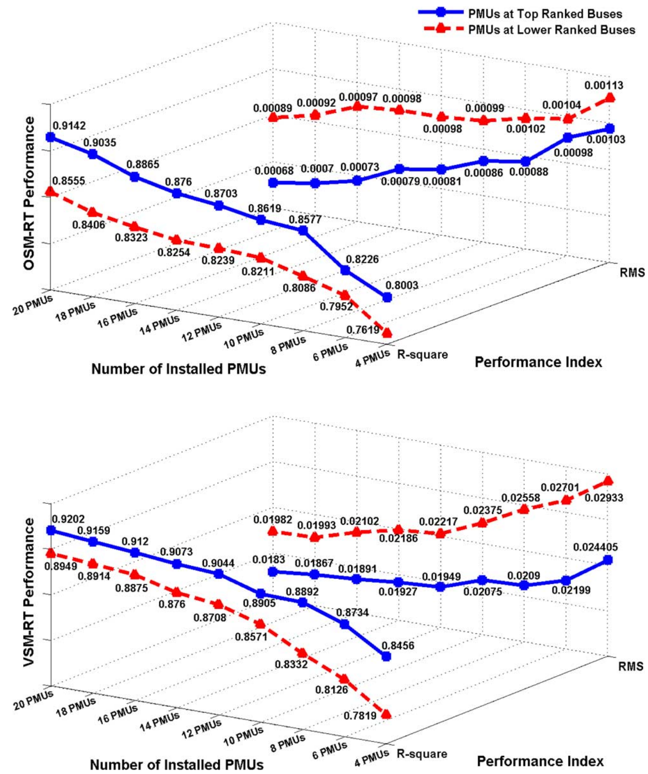

Fig. 12. RT performance considering different PMU placements in the 179-bus system.

Suppose that a number of 4 to 20 PMUs will be installed in the WECC system. By placing them at the top ranked buses of Table V, the resulted RT prediction accuracy for unseen OPs are summarized in Fig. 12. The RT performance using the measurements from the lowest ranked buses is also presented for the purpose of comparison.

As shown in Fig. 12, in contrast with the RTs fed with measurements from the lowest ranked buses, those constructed using the measurements from top ranked buses have exhibited better performances. Another conclusion could be made by comparing the R^2 of Fig. 10 with Fig. 8: almost identical RT prediction R^2 was achieved by using the reduced set of measurements from the PMU locations suggested by CBR. Last but not least, there is a huge decrease of the complexity in RT training since much fewer features are used. The training time of the 179-bus RTs has been reduced from about 3 min to less than 30 s.

VII. CONCLUSION

In this paper the approach of using regression tree to predict power system stability margins is explored and the following conclusions have been reached:

- Synchronized voltage and current phasors have been used as RT input feature. With a sufficiently large knowledge base, the RT model can predict the system oscillatory and voltage stability behavior with high accuracy.
- According to the test results, the RT model is fast enough to process PMU measurements, and it is robust to handle measurement errors that are within 1% TVE.

- The RT sensitivity to system topology variation becomes less distinct in large sized network and under mild changes in topology.
- The combined bus ranking derived from RT variable importance is used to suggest optimal PMU locations. Test results show that the measurements from reduced locations can still lead to satisfactory RT prediction accuracy.

References

[1] L. Breiman, J. Friedman, R. A. Olshen, and C. J. Stone, *Classification and Regression Trees*. Pacific Grove, CA: Wadsworth, 1984.

[2] L. Wehenkel, T. V. Cutsem, and M. Ribbens-Pavella, "An artificial intelligence framework for on-line transient stability assessment of power systems," *IEEE Trans. Power Syst.*, vol. 4, no. 2, pp. 789–800, May 1989.

[3] L. Wehenkel, M. Pavella, E. Euxibie, and B. Heilronn, "Decision tree based transient stability method a case study," *IEEE Trans. Power Syst.*, vol. 9, no. 1, pp. 459–469, Feb. 1994.

[4] S. Rovnyak, S. Kretsinger, J. Thorp, and D. Brown, "Decision trees for real-time transient stability prediction," *IEEE Trans. Power Syst.*, vol. 9, no. 3, pp. 1417–1426, Aug. 1994.

[5] K. Sun, S. Likhate, V. Vittal, V. S. Kolluri, and S. Mandal, "An online dynamic security assessment scheme using phasor measurements and decision trees," *IEEE Trans. Power Syst.*, vol. 9, no. 1, pp. 459–469, Feb. 2007.

[6] T. V. Cutsem, L. Wehenkel, M. Pavella, B. Heilbronn, and M. Goubin, "Decision tree approaches to voltage security assessment," *Proc. Inst. Elect. Eng.*, vol. 140, no. 3, pp. 189–198, May 1993.

[7] R. Diao *et al.*, "Decision tree-based online voltage security assessment using PMU measurements," *IEEE Trans. Power Syst.*, vol. 24, no. 2, pp. 832–839, May 2009.

[8] I. Kamwa, S. R. Samantaray, and G. Joos, "Catastrophe predictors from ensemble decision-tree learning of wide-area severity indices," *IEEE Trans. Smart Grid*, vol. 1, no. 2, pp. 144–158, Sep. 2010.

[9] R. F. Nuqui, A. G. Phadke, R. P. Schulz, and N. Bhatt, "Fast on-line voltage security monitoring using synchronized phasor measurements and decision trees," in *Proc. IEEE Power Engineering Society Winter Meeting*, 2001, vol. 3, pp. 1347–1352.

[10] R. Diao, V. Vittal, and N. Logic, "Design of a real-time security assessment tool for situational awareness enhancement in modern power systems," *IEEE Trans. Power Syst.*, vol. 25, no. 2, pp. 957–965, May 2010.

[11] S. P. Teeuwsen, I. Erlich, M. A. El-Sharkawi, and U. Bachmann, "Genetic algorithm and decision tree-based oscillatory stability assessment," *IEEE Trans. Power Syst.*, vol. 21, no. 2, pp. 746–753, May 2006.

[12] L. Bruno and V. Ajjarapu, "Development of multi-linear regression models for online voltage stability margin estimation," *IEEE Trans. Power Syst.*, vol. 26, no. 1, pp. 374–383, Feb. 2011.

[13] D. Q. Zhou, U. D. Annakkage, and A. D. Rajapakse, "Online monitoring of voltage stability margin using an artificial neural network," *IEEE Trans. Power Syst.*, vol. 25, no. 3, pp. 1566–1574, Aug. 2010.

[14] F. R. Gomez, A. D. Rajapakse, U. D. Annakkage, and I. T. Fernando, "Support vector machine-based algorithm for post-fault transient stability status prediction using synchronized measurements," *IEEE Trans. Power Syst.*, vol. 26, no. 3, pp. 1474–1483, Aug. 2011.

[15] I. Kamwa, S. R. Samantaray, and G. Joos, "Development of rule-based classifiers for rapid stability assessment of wide-area post-disturbance records," *IEEE Trans. Power Syst.*, vol. 24, no. 1, pp. 258–270, Feb. 2009.

[16] I. Kamwa, S. R. Samantaray, and G. Joos, "On the accuracy versus transparency trade-off of data-mining models for fast-response PMU-based catastrophe predictors," *IEEE Trans. Smart Grid*, vol. 3, no. 1, pp. 152–161, 2012.

[17] T. Van Cutsem and C. Vournas, *Voltage Stability of Electric Power Systems*. Boston, MA: Kluwer, 1998.

[18] P. Kundur *et al.*, "Definition and classification of power system stability IEEE/CIGRE joint task force on stability terms and definitions," *IEEE Trans. Power Syst.*, vol. 19, no. 3, pp. 1387–1401, Aug. 2004.

[19] V. Ajjarapu and C. Christy, "The continuation power flow: A tool for steady state voltage stability analysis," *IEEE Trans. Power Syst.*, vol. 7, no. 1, pp. 416–423, Feb. 1992.

[20] S. Dan and M. Golovnya, *CART 6.0 User's Manual*. San Diego, CA: Salford Systems, 2006.

[21] P. M. Anderson and A. A. Fouad, *Power System Control and Stability*. Ames, IA: The Iowa State Univ. Press, 1977, p. 38.

[22] M. A. Pai, *Energy Function Analysis for Power System Stability*. Boston, MA: Kluwer, 1989.

[23] Y. Zhou and V. Ajjarapu, "A fast algorithm for identification and tracing of voltage and oscillatory stability margin boundaries," *Proc. IEEE*, vol. 93, no. 5, pp. 934–946, May 2005.

[24] R. G. D. Steel and J. H. Torrie, *Principles and Procedures of Statistics*. New York: McGraw-Hill, 1960.

[25] R. F. Nau, Forecasting, Feb. 2005. [Online]. Available: http://www.duke.edu/~rnau/rsquared.htm.

[26] Electric Power Research Institute, DC Multi-Infeed Study, EPRI TR-104586s, Projects 2675-04-05, Final Report, 1994.

[27] *IEEE Standard for Synchrophasors for Power Systems*, IEEE Std. C37.118-2005, 2005.

[28] Y. Yang, X. Wu, and X. Zhu, "Mining in anticipation for concept change: Proactive-reactive prediction in data streams," *Data Mining Knowl. Discov.*, no. 13, pp. 261–289, 2006.

Ce Zheng (S'07) received the B.S. and M.S. degrees from North China Electric Power University, Beijing, China, in 2005 and 2007, respectively, both in electrical engineering. He is now pursuing the Ph.D. degree at Texas A&M University, College Station.

His research interests include application of data mining techniques to power system monitoring and protection, power system stability analysis based on PMU measurements, and the impact analysis of grid integration of renewable energy.

Vuk Malbasa received the B.S. degree from PMF, University of Novi Sad, Novi Sad, Serbia, in 2006 and the Ph.D. degree from Temple University, Philadelphia, PA, in 2012.

He is now with Texas A&M University, College Station, as a post doctoral researcher. His research interests include regularized spatial models, intelligent sensor applications, graphical models, and cognitive communications in networking.

Mladen Kezunovic (S'77–M'80–SM'85–F'99) received the Dipl.Ing., M.S., and Ph.D. degrees in electrical engineering in 1974, 1977, and 1980, respectively.

Currently, he is the Eugene E. Webb Professor, Director of the Smart Grid Center, Site Director of NSF I/UCRC "Power Engineering Research Center, PSerc", and Deputy Director of another NSF I/UCRC "Electrical Vehicles: Transportation and Electricity Convergence, EV-TEC". His main research interests are digital simulators and simulation methods for relay testing, as well as application of intelligent methods to power system monitoring, control, and protection. He has published over 450 papers, given over 100 seminars, invited lectures and short courses, and consulted for over 50 companies worldwide. He is the Principal of XpertPower™ Associates, a consulting firm specializing in power systems data analytics.

Dr. Kezunovic is a member of CIGRE and a Registered Professional Engineer in Texas.