Impact of Measurement Signals on the Accuracy of On-line Identification of Power System Dynamic Signature

Impact of Measurement Signals on the Accuracy of On-line Identification of Power System Dynamic Signature

Panagiotis N. Papadopoulos, Tingyan Guo, *Graduate Student Member*, Xiangyu Wang, Jovica V. Milanović, *Fellow, IEEE*
School of Electrical and Electronic Engineering
The University of Manchester
Manchester, UK
panagiotis.papadopoulos@manchester.ac.uk, tingyan.guo@postgrad.manchester.ac.uk, jovica.milanovic@manchester.ac.uk

***Abstract*—In order to identify the dynamic signature of power systems, it is important to monitor the generator rotor angles. The rotor angles can be obtained either by direct measurements of the rotor position or by indirect calculation of rotor angle from voltage and current measurements. Considering limited availability of direct measurements of rotor angles the indirect method is used in this paper taking into account the errors in calculated rotor angles. Decision tree based algorithm is used afterwards for online identification of power system dynamic signatures. Finally, the impact of the accuracy of measurement signals on the accuracy of the assessment of system dynamic signature is discussed.**

***Index Terms*—decision tree, phasor measurement units, power system dynamic signature, predictor importance, rotor angle measurement, transient stability.**

I. Introduction

Recent market driven changes in the operation of power systems along with the introduction of new generation and load technologies are increasing the complexity of power systems. Coupled with increased complexities there is a strong drive to optimise the capital investment in the system. This leads to power systems being pushed to operate closer to the stability margins. Therefore, the need for fast and accurate assessment of the stability status of the system becomes even more important and new, advanced preventive and corrective control mechanisms need to be applied to ensure stable and secure system operation [1]-[3].

Considering transient and small signal stability, the rotor angles of the generators provide valuable information about the status of each generator and consequently of the system. The rotor angles can be used as input to several methods for fast and accurate transient stability assessment. Data mining techniques, including decision trees, have been successfully applied for both binary identification of whether the system is stable or not [1], as well as for multiclass identification of the resulting groupings of generators [4], [5].

There are two main ways to determine the rotor angles of generators. The first is to directly measure the rotor position [6] and the second is to indirectly calculate the rotor angle from voltage and current measurement data, known as electrical calculation method [7], [8]. Conventional Phasor Measurement Units (PMUs) provide direct measurements of voltages and currents only. The electrical calculation method needs to be used to calculate rotor angles from available measurement data when the actual rotor angle measurements are not available. In [8] the rotor angles are calculated using measurement data from PMUs, while in [9] PMU measurements are directly used to assess the stability of a power system using Decision Trees (DTs).

In this paper the errors of PMU measurements are considered according to [7], as well as the errors introduced by the uncertainty in the reactance of the generators. The effect of these errors and uncertainties on the calculated rotor angles is investigated using a Monte Carlo approach to define the total error distribution. Following this, the calculated rotor angles are used to determine the impact of the uncertainty/errors in estimated rotor angles on the performance of DT based algorithm for online identification of the power system dynamic signature. The importance of each measurement signal is also assessed using an appropriate sensitivity measure. The importance of each measured signal in combination with the error calculation, provide an estimation of the overall impact of the accuracy of measurement signals on the performance of DTs for on-line identification of the power system dynamic signature.

II. Methodology

The voltage and current signals directly available from PMU measurements are used to calculate the generator rotor angles and define typical error distributions for each generator. The error distributions are then added to rotor angle signals to determine their impact on online identification of dynamic signature using decision trees.

This work was supported in part by the Engineering and Physical Sciences Research Council (EPSRC) and in part by the School of Electrical and Electronic Engineering, The University of Manchester.

A. Calculation of Generator Angles

Direct rotor angle measurements from the rotor position may not be available for each generator in a power system. For this reason, voltage and current measurement data from PMUs are used to calculate the angles of the generators.

The equivalent circuit shown in Fig. 1 is used for this purpose assuming the PMU is installed at the generator terminal. The transient reactance of the generator X_d' is used since the period under study is after the fault is cleared [8]. After obtaining the voltage $\overline{V}_2$ at the generator terminal and the current $\overline{I}_{in}$ from the PMU measurements, (1) can be used to calculate the internal generator voltage $\overline{E}$. As mentioned in [7] there is a fixed difference between the angle of the internal voltage and the actual rotor angle, which is constant for a specific generator and therefore can be neglected for the purpose of this study.

The same procedure can be used for the case when the PMU is installed at the high voltage side of the unit transformer by including also the reactance of the transformer X_t in the calculation.

$$\overline{E} = \overline{V} + jX_d^{'} \cdot \overline{I} \qquad (1)$$

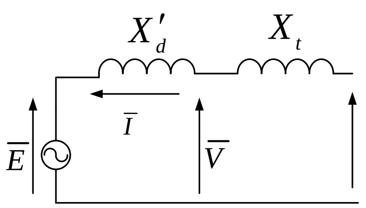

Figure 1. Equivalent circuit for the calculation of generator rotor angle.

B. Consideration of Errors

For the obtained voltage and current phasors from PMUs the Total Vector Error (TVE) defined in [7] and shown in (2) is considered. In Fig. 2, a visualization of the TVE for a value of 1% is given. This value is adopted in this paper in compliance with [7]. For each measurement provided by the PMU the point of the arrow of the phasor in Fig. 2 can be anywhere within the cycle. Therefore the maximum error in the magnitude is 1% while the maximum error in the angle is 0.573˚.

$$TVE(n) = \sqrt{\frac{(\hat{X}_r(n) - X_r(n))^2 + (\hat{X}_i(n) - X_i(n))^2}{(\hat{X}_r(n) + X_i(n))^2}} \qquad (2)$$

Where $\hat{X}_r(n)$ and $\hat{X}_i(n)$ are the estimated, i.e. the measured values, and $X_r(n)$ and $X_i(n)$ are the theoretical values for each sample n.

The voltages and currents at the terminals of each generator are calculated using detailed simulations and are considered as PMU measurement data. For each sample a random TVE within the circle shown in Fig. 2 is added to the theoretical values calculated by (1), using a Monte Carlo approach.

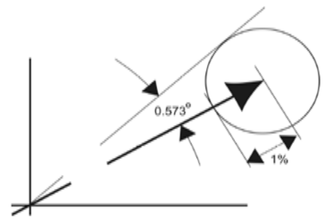

Figure 2. Visualization of TVE [7].

Apart from the TVE an error in the value of the generator transient reactance X_d' is considered as well. This error is bounded within the range of ±5% of the rated value to account for the possible differences in the generator reactance due to aging, data unavailability, etc.

By following a Monte Carlo approach, typical error distributions for each generator within a power system are defined. These error distributions are then added to simulated responses to further investigate their impact on the online identification of power system dynamic signature.

C. The effect of the measurement error

The measured rotor angle responses are used to identify very fast, within 1 sec, the dynamic behavior of the system. Two main approaches are used: binary classification, that determines if the system is stable or not, and multiclass classification that determines patterns of generator groupings, after a disturbance, i.e. the dynamic signature of the power system.

In this paper, DTs are used for both binary and multiclass classification, as in [1], [4] and [5]. In order to obtain training and testing sets for the DTs, Monte Carlo simulations are carried out in a test network, taking into consideration different fault location and duration and different system loading. Two different DTs are used for these two classification problems.

For binary classification the target is to identify whether the system is stable or not. The criterion used to detect the instability, is that the rotor angle between any two generators exceeds 360 degrees. Using this criterion a training set is obtained initially, to train the binary DT and its accuracy is evaluated using a different data set.

For multiclass classification, only the unstable cases of the above training and testing sets are used. Hierarchical clustering is initially applied to the rotor angles to determine the grouping of generators. Euclidean distance is used as the measure of similarity between the measured samples. Afterwards, the resulting groups are used as the target to train a DT as a multiclass classifier. More information on the complete process can be found in [1], [4] and [5].

In this study, errors according to the procedure described in Section II B, are added to the simulated rotor angles for both the training and testing datasets. The performance of the resulting DTs is compared to respective DTs trained directly with the simulated rotor angle responses without error. The impact on the accuracy of DTs due to the addition of error is investigated, using various algorithms such as Classification and Regression Tree (CART), C5, C5 boosting and Support Vector Machine (SVM) [4], [10].

D. Calculation of generator importance

For each specific created DT, the predictors used can be ranked according to their importance by computing a sensitivity measure S_i, as defined in (3). The predictor importance VI_i, can be then computed as the normalized sensitivity, as shown in (4). More information on the procedure is available in [11], [12].

$$S_i = \frac{V_i}{V(Y)} \tag{3}$$

$$VI_i = \frac{S_i}{\sum_{j=1}^{k} S_j} \tag{4}$$

where X_i is the predictor for which the sensitivity measure is calculated, V is the output variance considering all predictors, V_i is the variance without considering predictor X_i and k is the number of predictors used.

As explained above, the predictors are measurement samples of the rotor angles of all generators in the system. By determining the most important predictors, which are related to specific generator rotor angles, the most important/influential measured responses can be identified. By adding the predictor importance VI_i of all predictors corresponding to a specific generator, which are actually different samples within the measurement time frame, the generators are ranked according to how important the measured rotor angle is for determining the system dynamic signature. Two separate importance lists are generated for binary and multiclass classification, respectively. Moreover, a combined importance list by adding the predictor importance factors of the binary and multiclass classification DTs for each generator, is defined. This is an overall measure of the importance of the measured rotor angle of each generator [9].

III. Test Network and Simulations

A. Test system

The test system used in this study is the 16 machine, 68 bus reduced order equivalent model of the New England Test System and the New York Power System (NETS – NYPS), shown in Fig.3 and adopted from [1]. The network is simulated using Matlab/Simulink to obtain the necessary rotor angle responses along with the voltage and current data assumed to be available as PMU measurement data. G13 is the reference machine and therefore its rotor angle is not used in the following studies, since it is considered constant.

B. Monte Carlo Simulations to account for system uncertainties

To obtain the datasets required for the training and testing of the DTs, Monte Carlo simulations are carried out by varying the fault location, the fault duration and the system loading. Three phase self clearing faults are simulated in all cases. The fault location is changing randomly along all lines in the network following a uniform distribution. The fault duration follows a normal distribution with mean value of 13 cycles and standard deviation 0.667 cycles. The system load variation follows a normal distribution with mean value 1 pu based on [1] and standard deviation 0.033 pu and all the individual loads in the system are scaled accordingly. The simulations last 20 sec, i.e., 20 sec of rotor angle responses are captured for further analysis. For the training dataset, 5000 cases are simulated, while for the test dataset the number of simulated cases was 2000. There were 438 and 167 unstable cases (system loses stability) in the training and test data sets, respectively.

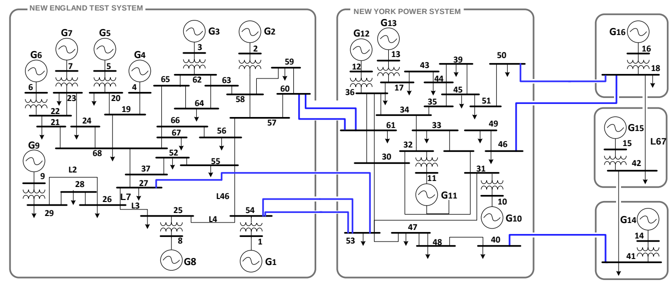

Figure 3. 68 bus test network.

IV. Impact of Measurement Error on Online Identification of Dynamic Signature

A. Calculation of rotor angle signals from PMU measurements

The calculated rotor angles, with the added error according to the method described in Section II, of all generators following 8-cycle self clearing three phase fault at Bus 1 are shown in Fig. 4. The absolute error ε for each sample and for each generator is calculated by (5). The maximum error observed is 1.5˚ for generator G12.

$$\varepsilon = \delta_e - \delta \tag{5}$$

where δ_e is the rotor angle value with the added error and δ is the theoretical value calculated by (1).

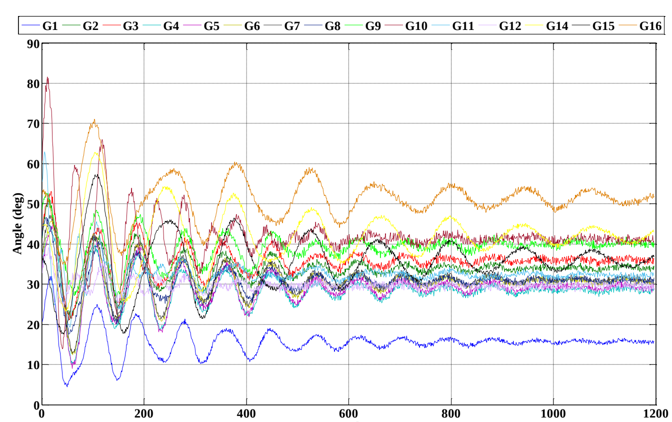

Figure 4. Generator rotor angles with error.

To investigate further the effect of the calculated rotor angle error a Monte Carlo approach is applied. Random error values are added multiple times for each sample within the 1% TVE and 5% impedance limit, as described in Section II B. One thousand cases considering random, uniformly distributed phasor values and impedance errors are simulated for each measurement point. Initially, the mean value and standard deviation of the 1000 errors for each single measurement point is calculated. Afterwards, the mean value of the individually calculated mean values and standard deviations for all measurement points and for each generator are calculated. Typical rotor angle error distributions for each generator, shown in Table I, are identified in order to investigate the impact of the added error during the calculation of rotor angles from voltage and current measurements. Identified normal distributions are all mean value 0 and standard deviation as shown in Table I.

TABLE I. TYPICAL ERROR DISTRIBUTION FOR EACH GENERATOR

Generator	**1**	**2**	**3**	**4**	**5**
St. Dev. (°)	0.315	0.482	0.539	0.435	0.464
Generator	**6**	**7**	**8**	**9**	**10**
St. Dev. (°)	0.426	0.416	0.489	0.597	0.717
Generator	**11**	**12**	**14**	**15**	**16**
St. Dev. (°)	0.441	0.619	0.298	0.290	0.451

B. Decision tree performance with added error

Calculated errors with normal distributions with mean value 0 and standard deviations as shown in Table I are added to the rotor angles obtained from the Monte Carlo simulations as described in Section II. Signals with added errors are then used for training as well as testing of the DTs. The IBM SPSS Modeler software is used for the purpose of training and testing DTs [13].

Considering multiclass classification, as mentioned in Section II C, hierarchical clustering is applied initially to determine the generator groupings. Only the last sample of the rotor angles at the end of the simulation, i.e. at 20 sec, is used in this study as in [5]. There is a total of 12 patterns observed in the training set and 9 in the test set. Seven are common between the training and test set, while two are only observed in the test set. All the occurring patterns are considered during the training and testing process. However, only the four most important patterns (>95% of the cases) observed in both the training and test datasets, are presented in Table II. The rest of the patterns are only appearing only one or two times in each dataset.

In Table III, the accuracy of binary and multiclass DTs is shown for the case where the rotor angles from the simulation are directly used as well as for the case when error is added. The length of data sets of the rotor angles used, is considered to be either 10 or 60 cycles, in order to determine if this has also any effect on the performance. The performance of the DTs without adding error is shown in brackets in Table III.

It can be seen form Table III that the performance of the DTs is only slightly affected by added measurement/calculation errors. The influence of errors is more pronounced in case of mainly multiclass classification, but the effect is not significant. More specifically, for binary classification the accuracy with and without error remains practically unaffected.

TABLE II. MOST IMPORTANT GROUPING PATTERNS

Pattern No	**No of cases**	**Generator grouping**	**No of groups**
1	92	(G1-G9,G12,G14-G16)/(G10)/(G11)	3
2	40	(G1-G8,G10-G12,G14-G16)/(G9)	2
3	9	(G1-G9)/(G10)/(G11)/(G12)/(G14-G15)/(G16)	6
4	20	(G1-G9)/(G10)/(G11)/(G12)/(G14-G16)	5

TABLE III. ACCURACY OF DTS WITH ADDED ERROR TO THE SIGNALS

DT algorithm	**Binary**		**Multiclass**	
	10 cycles	**60 cycles**	**10 cycles**	**60 cycles**
CART	97.2 (97.15)	99.5 (99.6)	85.63 (83.83)	79.04 (88.62)
C5	98.15 (98.45)	99.45 (99.75)	86.23 (83.23)	90.42 (90.42)
C5 boosting	98.1 (98.8)	99.75 (99.75)	86.83 (88.62)	91.62 (89.82)
SVM	98.7 (98.8)	98.05 (98.65)	82.63 (85.63)	86.63 (89.82)

Considering multiclass classification, the largest drop in accuracy when error is added, around 9.5%, is observed for CART algorithm when 60 cycles are used. Small variations in the accuracy are noticed for all algorithms for both 10 and 60 cycles. It should be noted that only a few misclassified cases can cause differences of around 2% for multiclass classification since the number of the testing group is not very large. This is also the reason for an increase in accuracy noticed in some cases when the error is added. Since different DTs are built for each of the cases shown in Table III, it is possible that in some cases (CART and C5 for 10 cycles and C5 boosting for 60 cycles) the DT trained with measurement errors in the rotor angles is “more robust” and can identify correctly a few cases more.

In general, considering binary classification, it can be concluded that the performance of DTs will not be significantly affected by adding realistic values of measurement errors. A difference of a few degrees in rotor angles due to adding the measurement error does not have significant impact as the rotor angles of unstable generators will attain large values anyway. For multiclass classification though, even a few degrees difference in rotor angle samples that are important predictors for the DT can change the classification result.

V. IDENTIFICATION OF IMPORTANT GENERATORS

A. List of important generators

In Table IV and V, the predictor importance of the most important predictors for binary and multiclass classification using the C5 and C5 boosting algorithm is shown. The results for two algorithms are presented to highlight that the important predictors can vary according to the algorithm used to create the DTs. The case where 60 cycles of the rotor angles are used to train and test the DTs, is investigated for both algorithms. In a similar manner, for different duration of rotor angles, the importance list would be different.

When the tree is created, the most important predictors are identified according to the methodology described in Section II. These important predictors are actually samples of rotor angles of specific generators during the 60 cycle period. For example, the last sample of the rotor angle of a generator for the 60th cycle can be an important predictor considering binary classification, as the case is for G10 in Table V. More than one sample corresponding to one generator might be included in the list. This happens for example for G3 in multiclass classification as shown in Table IV.

The predictor importance values for each generator are added to calculate the importance of the specific generator. Moreover, the importance of both binary and multiclass classification methods can be further added to come up with an overall importance list of generators, as shown in Table VI. This combined importance list is used to provide an insight as to which signals are more important to be used in the online identification of the power systems dynamic signature.

TABLE IV. MOST IMPORTANT PREDICTORS FOR C5 ALGORITHM

Binary			Multiclass		
Generator	*Cycle*	*Predictor Importance*	*Generator*	*Cycle*	*Predictor Importance*
G11	42	0.5	G3	60	0.46
G9	53	0.26	G9	18	0.35
G10	59	0.23	G11	46	0.15
G2	60	0.01	G5	58	0.02
G3	25	0.01	G3	26	0.01
-	-	-	G6	1	0.01

TABLE V. MOST IMPORTANT PREDICTORS FOR C5 BOOSTING ALGORITHM

Binary			Multiclass		
Generator	*Cycle*	*Predictor Importance*	*Generator*	*Cycle*	*Predictor Importance*
G10	60	0.03	G9	18	0.19
G10	33	0.03	G3	60	0.16
G9	27	0.03	G5	58	0.12
G3	9	0.03	G6	1	0.12
G14	42	0.03	G9	52	0.1
G8	16	0.03	G5	1	0.1
G10	41	0.03	G3	26	0.1
G3	60	0.03	G11	46	0.1
G6	19	0.03	-	-	-
G10	25	0.03	-	-	-

B. *Accuracy of DTs using important generator signals*

The accuracy of DTs using generators from the importance list is compared to that of the original trees using all generator data as well as to DTs built using signals from the rest of the generators that are not included in the importance list of Table VI. The procedure is applied to the importance list obtained from using the C5 boosting algorithm with 60 cycles duration of rotor angles. There are 8 generators in the combined importance list and 15 in total in the network, excluding G13 which is the reference generator. Initially all the important generators are used to train a DT and test it using the dataset described previously. Afterwards, starting from the last generator of the combined importance list (G8), one generator at a time is removed until only the most important remains and DTs are trained and tested using the same dataset. Furthermore, the rotor angles of the 7 generators that are not included in the importance list (G1, G2, G4, G7, G12, G15, G16) are used to train and test DTs. Finally, rotor angle signals of the generators not included in the importance list are removed randomly until the rotor angle of one random generator remains. The results for binary and multiclass classification are shown in Fig. 5 and 6 respectively.

TABLE VI. COMBINED IMPORTANCE LIST

	C5			**C5 boosting**	
Order	**Generator**	**Predictor Importance**	**Order**	**Generator**	**Predictor Importance**
1	G11	0.65	1	G3	0.23
2	G9	0.61	2	G9	0.23
3	G3	0.48	3	G6	0.15
4	G10	0.23	4	G5	0.13
5	G2	0.01	5	G10	0.12
6	G6	0.01	6	G11	0.1
-	-	-	7	G14	0.03
-	-	-	8	G8	0.03

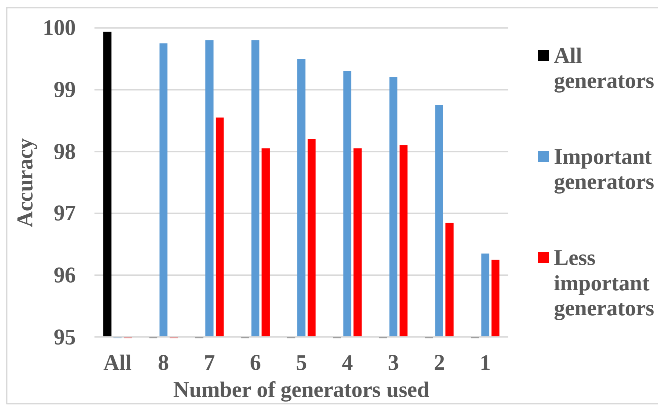

Figure 5. Accuracy of binary classification DTs with reduced number of used signals.

For binary classification, the accuracy of the DTs for all cases is relatively high, above 96%. When using the rotor angles of at least two important generators the accuracy is very close to the case when all rotor angles are used (98.75% compared to 99.9%). The accuracy drops though (about 96.25%) when only one generator is used, whether the most important one or any other generator in the system. This means, that the transient stability status of the system can be identified with relatively high accuracy even with rotor angle measurements of one generator, whether it is the most important one or any random generator in the system. However, at least two signals are recommended to be used to achieve the highest possible accuracy.

Considering multiclass classification, a similar behavior is observed but in this case the accuracy of DTs is significantly higher when rotor angles of generators within the combined importance list are used. Even the rotor angle of one generator (G3) from the list is enough to identify with relatively high accuracy the power system dynamic signature compared to the case when all rotor angles are used (88% compared to 91% when signals from eight generators are used and 96% when signals from all generators are used). However, when random rotor angles of a small number of generators outside the importance list are used, the accuracy drops significantly. This means that identifying important generators is essential when a small number of signals needs to be used.

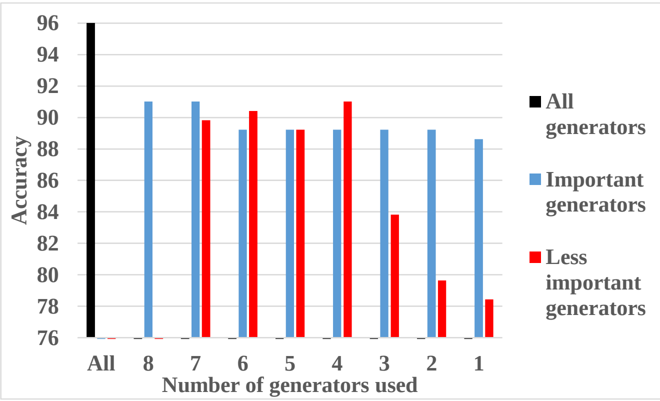

Figure 6. Accuracy of multiclass classification DTs with reduced number of used signals.

To further explain the reason why the most important generator rotor angle (G3) is still sufficient for the online identification of the generator groupings, the rotor angle of G3 for 50 of the 167 cases of the test dataset are presented in Fig. 7. The most important predictors used when the DT is trained using only the rotor angle of G3 are also highlighted in Fig. 7. It is shown that by observing the responses of G3 that belong to the 3 most common patterns (1, 2, 4) of Table II, the DT can identify the grouping pattern of all the generators. This means that G3 has a specific response when it belongs to one of the most common patterns, that provides the ability to the DT to clearly identify the grouping pattern it belongs to. However, in the case when a larger number of patterns and more complex patterns appear in the power system dynamic signature, more rotor angle signals might be required.

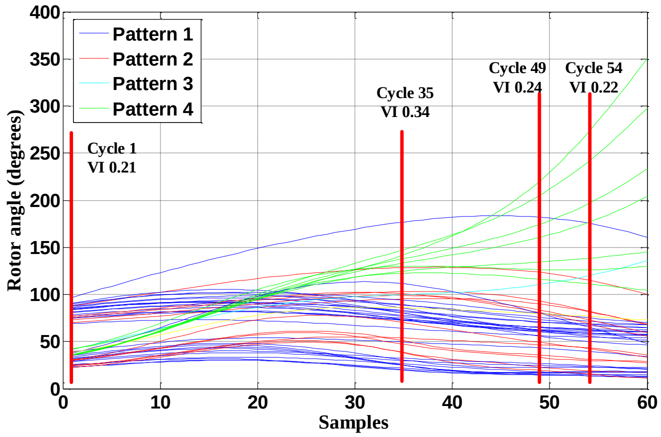

Figure 7. Rotor angle of G3 for 50 representative cases.

VI. Conclusions

The impact of measurement signals and procedures required to obtain rotor angle measurements on the performance of methods for online identification of the power system dynamic signature are investigated.

Due to the fact that measurements of rotor angle may not be available for each generator an electrical calculation method is followed in this paper, using voltage and current measurements from PMUs. A Monte Carlo approach is used to define typical error distributions which are then used to study the impact of measurement errors on the performance of binary and multiclass DTs. The performance of DTs is not significantly affected, especially for binary classification.

The importance of each measured rotor angle is also identified, using an appropriate sensitivity measure and a combined importance list is defined. DTs are created using only some of the generators rotor angles as predictors. While the performance can be affected when not all generator rotor angles are used, even a small number of the signals from important generators can lead to high accuracy, in excess of 88%, of prediction of system dynamic signature.

References

[1] T. Guo and J. V. Milanovic, "Probabilistic Framework for Assessing the Accuracy of Data Mining Tool for Online Prediction of Transient Stability," *IEEE Trans. Power Syst.*, vol. 29, pp. 377-385, 2014.

[2] M. He, V. Vittal, and J. Zhang, "Online Dynamic Security Assessment With Missing PMU Measurements: A Data Mining Approach," *IEEE Trans. Power Syst.*, vol. 28, no. 2, pp. 1969-1977, 2013.

[3] 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, 1994.

[4] T. Guo, Jiachen He, Zhengyou Li and J. V. Milanović, "Evaluation of Classification Methods for On-line Identification of Power System Dynamic Signature", *18th Power Systems Computation Conference*, Wroclaw, Poland, 18-22 August 2014.

[5] T. Guo and J. V. Milanović, "Identification of Power System Dynamic Signature Using Hierarchical Clustering", *IEEE Power & Energy Society General Meeting 2014*, Washington, USA, 27-31 July 2014.

[6] Zhou Jie, Zhang Daonong, Yu Yuehai, Wu Jingtao, Zhou Ji, Bi Tianshu, Xiong Ming, "Rotor position measurement method for generator power angles", *International Conference on Advanced Power System Automation and Protection (APAP) 2011*, Vol.3, pp. 2196 – 2200, 2011.

[7] "IEEE Standard for Synchrophasor Measurements for Power Systems," *IEEE Std C37.118.1-2011 (Revision of IEEE Std C37.118-2005)*, pp. 1-61.

[8] P. K. Olulope, G. K. Venayagamoorthy, K. A. Folly, "Wide area monitoring of rotor angle based on synchrophasor measurements", *Power Engineering Society Conference and Exposition in Africa (PowerAfrica), 2012 IEEE*, pp. 1 – 7, 2012.

[9] Ce Zheng, V. Malbasa, M. Kezunovic, "Regression tree for stability margin prediction using synchrophasor measurements", *IEEE Trans. Power Syst.*, vol. 28, no. 2, pp. 1978 – 1987, 2013.

[10] J. Han, M. Kamber, and J. Pei, *Data Mining : concepts and techniques*: Morgan Kaufmann, 2012.

[11] IBM SPSS Modeler Algorithms guide, IBM Corporation 1994, 2012.

[12] A. Saltelli, S. Tarantola, F. Campolongo, and M. Ratto, *Sensitivity Analysis in Practice – A Guide to Assessing Scientific Models*, Wiley, 2004.

[13] IBM SPSS Modeler v14.2, http://www-01.ibm.com/software/analytics/spss/products/modeler/