Journal of ELECTRICAL ENGINEERING, VOL. 63, NO. 4, 2012, 255–260

FAULT IDENTIFICATION BASED ON NLPCA IN COMPLEX ELECTRICAL ENGINEERING

Yagang Zhang — Zengping Wang — Jinfang Zhang *

The fault is inevitable in any complex systems engineering. Electric power system is essentially a typically nonlinear system. It is also one of the most complex artificial systems in this world. In our researches, based on the real-time measurements of phasor measurement unit, under the influence of white Gaussian noise (suppose the standard deviation is 0.01, and the mean error is 0), we used mainly nonlinear principal component analysis theory (NLPCA) to resolve fault identification problem in complex electrical engineering. The simulation results show that the fault in complex electrical engineering is usually corresponding to the variable with the maximum absolute value coefficient in the first principal component. These researches will have significant theoretical value and engineering practical significance.

K e y w o r d s: fault identification, nonlinear principal component analysis, NLPCA, complex system theory, electrical engineering

1 INTRODUCTION

The fault is inevitable in any complex systems engineering. In general, a fault is a deviation from the normal behavior in the equipment or its components, and it is a process abnormality or symptom. It can also be defined as a departure from an acceptable range of an observed variable or calculated parameter associated with the equipment. The faults may arise in the basic technological equipment or in its measurement and control instruments, and may represent performance deterioration, partial malfunctions or total breakdowns [1–4]. The detection procedure locates the process or unit malfunction that caused the symptoms.

The goal of fault detection is to ensure the success of the planned operations by recognizing anomalies of system behavior. Fault detection is a well established concept in many areas of applied systems engineering. It implies the capability of determining, either actively or passively, whether a system is functioning as intended or as modeled. A system with faults does not necessarily imply that the system is not functioning. Detecting a fault involves identifying a characteristic of the system, when a fault occurs, which can be distinguished from other characteristics of the system. Generally speaking, the process of fault detection can be divided into three main steps: [5–8]

- Alarm: In this process, information on current processing status will be extracted from the signals measured by internal sensors;
- Identification: This process will determine the location of a failure;
- Evaluation: The evaluation process involves the determination of the extent or severity of a failure.

Electric power system is essentially a typically nonlinear system. It is also one of the most complex artificial systems in this world. As we know, the safe, steady, economical and reliable operation of electric power system plays a very important part in guaranteeing socioeconomic development, even in safeguarding social stability. The complexity of electric power system is determined by its characteristics about constitution, configuration, operation, organization, *etc*, which has caused many disastrous accidents, such as the large-scale blackout of America-Canada electric power system on August 14, 2003, the large-scale blackout of Italy electric power system on September 28, 2003. In this paper, based on the real-time measurements of phasor measurement unit (PMU) [9–12], under the influence of white Gaussian noise, we used mainly nonlinear principal component analysis theory (NLPCA) to resolve fault identification problem in complex electrical engineering.

2 WIDE AREA MEASUREMENT SYSTEM

Phasor Measurement Units (PMU) is the remote measurement devices of the Wide Area Measurement System, which is the product of the wide application of Global Position System (GPS) in the world. The first PMU equipment is born in Virginia Tech in USA by Professor Arun G. Phadke and James S. Throp in 1980s [13]. The basic structure and principle of PMUs is similar with that of a computer relay, excluding the GPS receiver. By using the synchronized clock signals from GPS, the PMUs dispersedly equipped in the electric network could obtain the

* Department of Electrical Engineering, North China Electric Power University, Key Laboratory of Power System Protection and Dynamic Security Monitoring and Control under Ministry of Education, Box 205, Baoding, Hebei 071003, PR China, yagangzhang@gmail.com

DOI: 10.2478/v10187-012-0036-4, ISSN 1335-3632

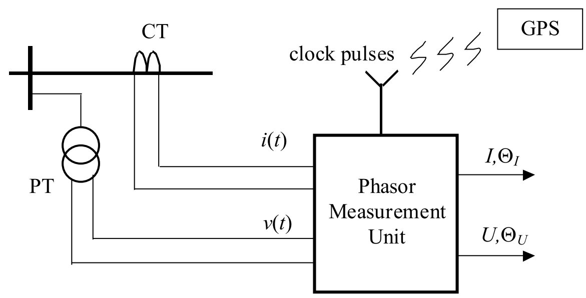

Fig. 1. A simple depiction of the PMUs

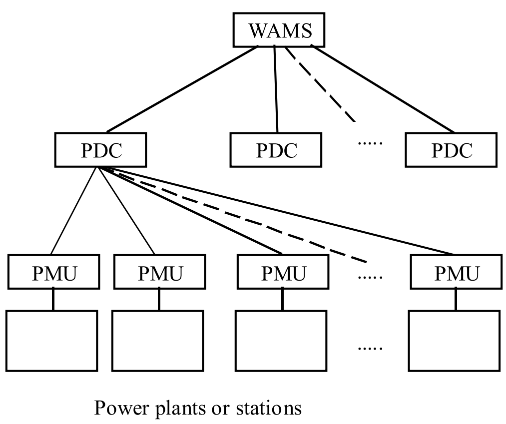

Fig. 2. The typical hierarchical structure of the WAMS

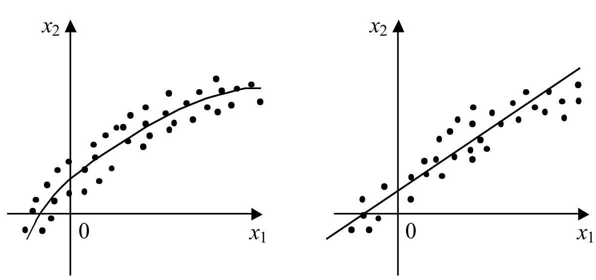

Fig. 3. The geometry principle curves of NLPCA and PCA

same sampling clock. Then the corresponding input signals (consisting of voltages at buses and feeder currents) will be sampled and converted into positive sequence quantities (if necessary, negative and zero sequence also can be obtained). Consequently, the operation condition of the power system in one snapshot is to be depicted with mutil-points synchronized phasors indicating in the same coordinate. The refreshed rate of the synchronized phasors provided by the PMUs is as frequently as one every cycle, which also can be recommended as 25 frames/sec, 50 frames/sec or 100 frames/sec in China.

One typical structure in PMU could be represented in Fig. 1 [13, 14]. The nodal voltage and feeder currents analog signals are sampled and converted into the needed sequence phasor, then the synchronized data could be uploaded to the Phasor Data Concentrator (PDC) with the certain refreshed rate.

Compare to the traditional measurement system such as SCADA/EMS with Remote Terminal Unit (RTU), the PMU/WAMS, defined as the modern measurement system, not only could finish the functions required in the conventional one, but also will or has brought profound impact on state estimation, dynamic monitoring and system protection and so on. A lot of documents have explored the wide application of WAMS and PMUs in current or future power system.

The above mentioned PMU, PDC and communication links are the main devices to realize the full benefit of the PMU measurement. The architecture of the WAMS could be divided as different levels, and in each level the PDCs could match the time tags of data received from the various PMUs so that the phasor data stream is created for application, and communicated to upper levels (as well as PMUs) [13, 14]. In this way, different level will service for the various functions. Especially, in the researches of the authors, the regional or central control centers will be the appropriate target levels to implement the wide area backup protection, which requires the phasor data from much wider areas, even the whole system, with the longer time delay.

A classical architecture of the WAMS/PMU could be shown in Fig. 2.

3 NONLINEAR PRINCIPAL COMPONENT ANALYSIS THEORY

Nonlinear principal component analysis is a novel technique for multivariate data analysis, similar to the well-known method of principal component analysis (PCA) [15–17]. The classic linear PCA method assumes that the transformed features of the process are linear functions of the observed variables. In industrial engineering, however, this assumption may not be true when the observations are from highly nonlinear processes. In such cases, it may be more appropriate to assume that the feature subspace is defined by nonlinear functions of the process variables. Figure 3 presents the geometry principle curves of NLPCA and PCA. Now let us illuminate the concrete principle of nonlinear principal component analysis.

Suppose there are m variables $\xi_1, \xi_2, \dots, \xi_m$, each variable has n components, that is $\{\xi_{ij}\}_{n\times m}$. Now let us adopt *Centralized Logarithm Transformation* to carry out nonlinear principal component analysis.

First step, one transforms the original data by centralized logarithm transformation,

$$\eta_{ij} = \lg \xi_{ij} - \frac{1}{m}\sum_{t=1}^{m} \lg \xi_{it}\,, \qquad (i = 1, 2, \dots, n\,;\ j = 1, 2, \dots, m)\,. \quad (1)$$

Then, one can calculate the covariance matrix,

$$S = (s_{ij})_{m\times m} \quad (2)$$

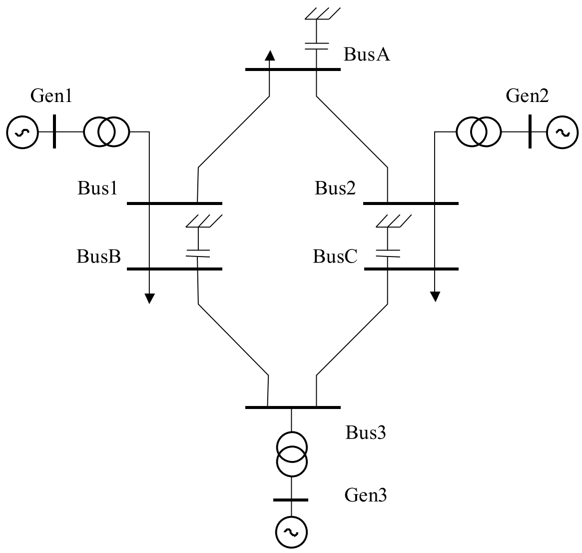

Fig. 4. Electric diagram of IEEE 9-Bus system

where

$$s_{ij} = \frac{1}{n-1}\sum_{t=1}^{n}(\eta_{ti}-\bar{\eta}_i)(\eta_{tj}-\bar{\eta}_j)\,, \qquad \bar{\eta}_i = \frac{1}{n}\sum_{t=1}^{n}\eta_{ti}\,, \quad \bar{\eta}_j = \frac{1}{n}\sum_{t=1}^{n}\eta_{tj}\,. \tag{3}$$

Let $\lambda_1 \ge \lambda_2 \ge \cdots \ge \lambda_m \ge 0$ be the characteristic roots of S, and $t_1, t_2, \ldots, t_m$ be their corresponding unitization characteristic vectors. So, the k-th nonlinear principal component can be expressed as

$$Y_k = \sum_{j=1}^{m} t_{kj} \lg \xi_{kj}\,, \ (k = 1, 2, \ldots, m)\,. \tag{4}$$

Generally speaking, the k-th principal component should satisfy $t'_k t_k = 1$, $t'_k t_i = 0$, $t'_i t_k = 0\,(i < k)$. Therefore, we can construct an objective function

$$\varphi_k(t_k, \lambda, \rho_i) = t'_k S t_k - \lambda(t'_k t_k - 1) - 2\sum_{i=1}^{k-1}\rho_i(t'_i t_k)\,. \tag{5}$$

Let us differentiate it,

$$\frac{\partial \varphi_k}{\partial t_k} = 2St_k - 2\lambda t_k - 2\sum_{i=1}^{k-1}\rho_i t_i = 0\,. \tag{6}$$

Left multiplication t'_i,

$$t'_i S t_k - \lambda t'_i t_k - t'_i \sum_{i=1}^{k-1}\rho_i t_i = 0\,, \tag{7}$$

namely $\rho_i t'_i t_i = 0$, one can get $\rho_i = 0 \ (i = 1, 2, \ldots, k-1)$. Then,

$$(S - \lambda I)t_k = 0 \tag{8}$$

and

$$t'_k S t_k = \lambda\,. \tag{9}$$

So, the k-th principal component has been solved.

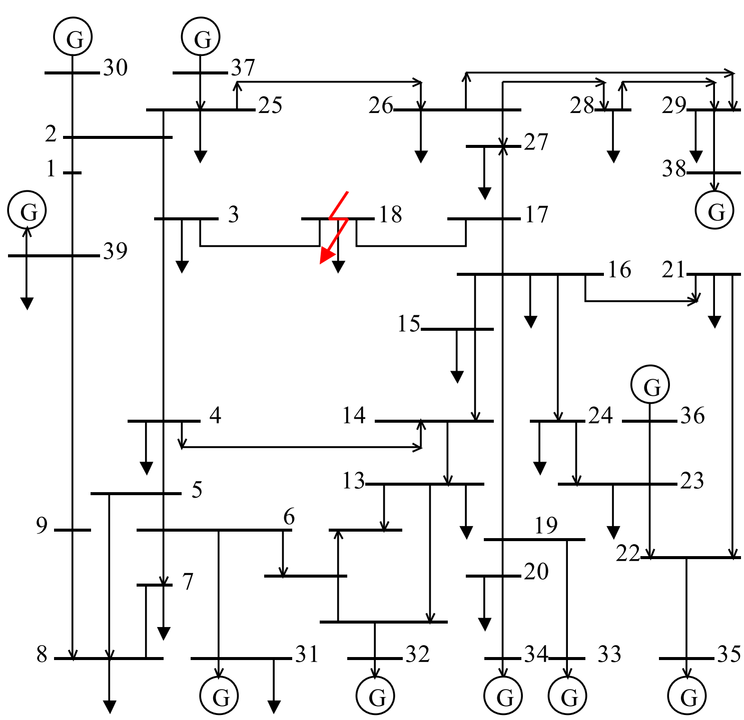

Fig. 5. Electric diagram of IEEE 39-Bus system

Table 1. The covariance matrix of node negative sequence voltages in IEEE9-Bus system

Bus	Gen1	Gen2	Gen3	Bus1	Bus2	Bus3	BusA	BusB	BusC
Gen1	0.004327	0.008537	0.003883	-0.01545	0.002752	0.000273	0.003435	-0.00536	-0.0024
Gen2	0.008537	0.036649	0.020056	-0.06197	0.012424	0.002112	0.016915	-0.02203	-0.01269
Gen3	0.003883	0.020056	0.014078	-0.03622	0.0071	0.001579	0.009804	-0.01313	-0.00715
Bus1	-0.01545	-0.06197	-0.03622	0.110178	-0.02152	-0.00494	-0.03107	0.039918	0.021068
Bus2	0.002752	0.012424	0.0071	-0.02152	0.004289	0.000915	0.006116	-0.00776	-0.00431
Bus3	0.000273	0.002112	0.001579	-0.00494	0.000915	0.000705	0.002136	-0.00207	-0.00071
BusA	0.003435	0.016915	0.009804	-0.03107	0.006116	0.002136	0.010159	-0.01167	-0.00582
BusB	-0.00536	-0.02203	-0.01313	0.039918	-0.00776	-0.00207	-0.01167	0.014628	0.007482
BusC	-0.0024	-0.01269	-0.00715	0.021068	-0.00431	-0.00071	-0.00582	0.007482	0.004536

4 FAULT IDENTIFICATION BASED ON NLPCA IN COMPLEX POWER SYSTEM

NLPCA is used to identify and remove correlations among objective variables as an aid to dimensionality reduction, visualization, and exploratory data analysis. While PCA identifies only linear correlations between variables, NLPCA uncovers both linear and nonlinear correlations, without restriction on the character of the nonlinearities present in the data.

Now let us consider IEEE9-Bus system, Figure 4 is its electric diagram. In the structure of electricity grid, Bus-1 appears single-phase to ground fault. By BPA simulations, the vector-valued of corresponding variables is only exported one times in each period. Considering the influence of white Gaussian noise, suppose the standard deviation is 0.01, and the mean error is 0. Using these actual measurement data of corresponding node negative sequence voltages, we will carry through nonlinear principal component analysis of fault component and non-fault component.

4.1 Fault Identification of IEEE9-Bus system based on node negative sequence voltage

After computing IEEE9-Bus system, we can get node negative sequence voltages at T_{-1}, T_0 (Fault), T_1, T_2 and T_3 five times. First of all, after centralized logarithm transformation, the covariance matrix of node negative sequence voltages in IEEE9-Bus system can be calculated, see Table 1.

In this covariance matrix, a remarkable characteristic is present that the covariance of Bus1 is 0.110178, which is the biggest one. So, one can analyze preliminarily that the Bus1 is a probable fault component.

Let us solve the eigenvalues of the covariance matrix, these results have been listed in Table 2. Finally, the first principal component can be obtained, its expression is

$$\begin{aligned} Y_1 = 0.103555Z_1 + 0.429839Z_2 + 0.250647Z_3 \\ - 0.758725Z_4 + 0.149019Z_5 + 0.033330Z_6 \\ + 0.213893Z_7 - 0.274438Z_8 - 0.147119Z_9 \end{aligned} \tag{10}$$

To sum up the above NLPCA results, although there exists the influence of white Gaussian noise, from the feature of the first principal component, Bus1 corresponds with variable Z_4, and the coefficient absolute value of Z_4 is 0.758725, which is also the biggest one. So, Bus1 is just the fault component. This result is entirely identical with the fault set in advance.

Now let us further consider a more complicated IEEE 39-Bus system, Figure 5 is its electric diagram. In the structure of electricity grid, Bus-18 appears single-phase to ground fault. By BPA simulations, the vector-valued of corresponding variables is only exported one times in each period. Considering the influence of white Gaussian noise, suppose the standard deviation is 0.01, and the mean error is 0. Using these actual measurement data of corresponding variables, we will carry through nonlinear principal component analysis of fault component and non-fault component.

4.2 Fault Identification of IEEE39-Bus system based on node negative sequence voltage

Similarly, we calculate the node negative sequence voltages at T_{-1}, T_0 (Fault), T_1, T_2 and T_3 five times. After centralized logarithm transformation, the covariance matrix of node negative sequence voltages in IEEE39-Bus system can be obtained, see Table 3. In this place, we only intercept Bus18 section.

In Table 3, a remarkable characteristic is in esse, the covariance of Bus18 is 0.410435, which is not only the biggest one in Table 3 (only intercept Bus18 section), but also the biggest one in the complete covariance matrix based on node negative sequence voltages in IEEE 39-Bus system. So, one can analyze preliminarily that the Bus18 is a probable fault component.

Let us further solve the eigenvalues of this covariance matrix, see Table 4. Finally, the first principal component

Table 3. The covariance matrix of node negative sequence voltages in IEEE39-Bus system (Only intercept Bus18 section)

Bus	Bus1	Bus2	Bus3	Bus4	Bus5	Bus6	Bus7	Bus8	Bus9	Bus10
Bus18	-0.10175	-0.04901	0.134574	0.095032	0.062569	-0.04465	-0.18953	-0.07488	-0.22995	-0.0542
Bus	Bus11	Bus12	Bus13	Bus14	Bus15	Bus16	Bus17	Bus18	Bus19	Bus20
Bus18	0.08282	-0.07464	-0.14773	0.231508	0.194423	0.129879	0.198885	0.410435	0.049834	-0.12922
Bus	Bus21	Bus22	Bus23	Bus24	Bus25	Bus26	Bus27	Bus28	Bus29	Bus30
Bus18	0.192603	-0.09827	-0.07417	0.143789	-0.00576	0.001598	0.111723	-0.05032	-0.0396	-0.06794
Bus	Bus31	Bus32	Bus33	Bus34	Bus35	Bus36	Bus37	Bus38	Bus39	
Bus18	-0.05089	0.066787	-0.00325	-0.10006	-0.12987	-0.00146	-0.12064	0.014432	-0.28311	

Table 2. The eigenvalues of covariance matrix in IEEE9-Bus system

No.	Eigenvalues	Proportion	Cumulative
1	0.191085	0.9576	0.9576
2	0.003637	0.0182	0.9758
3	0.002833	0.0142	0.9900
4	0.001993	0.0100	1.0000

Table 4. The eigenvalues of covariance matrix in IEEE39-Bus system

No.	Eigenvalues	Proportion	Cumulative
1	1.807145	0.6928	0.6928
2	0.352246	0.1350	0.8279
3	0.326595	0.1252	0.9531
4	0.122387	0.0469	1.0000

is obtained, which can be expressed as

$$\begin{aligned} Y_1 = &-0.124846Z_1 - 0.056522Z_2 + 0.159715Z_3 \\ &+ 0.112571Z_4 + 0.082188Z_5 - 0.046110Z_6 - 0.214863Z_7 \\ &- 0.095761Z_8 - 0.270454Z_9 - 0.065628Z_{10} + 0.091790Z_{11} \\ &- 0.093987Z_{12} - 0.154935Z_{13} + 0.270549Z_{14} + 0.227069Z_{15} \\ &+ 0.162749Z_{16} + 0.231255Z_{17} + 0.474260Z_{18} + 0.052699Z_{19} \\ &- 0.136666Z_{20} + 0.220112Z_{21} - 0.121749Z_{22} - 0.090534Z_{23} \\ &+ 0.159917Z_{24} - 0.006500Z_{25} + 0.006073Z_{26} + 0.133852Z_{27} \\ &- 0.064502Z_{28} - 0.037887Z_{29} - 0.081580Z_{30} - 0.050878Z_{31} \\ &+ 0.067623Z_{32} - 0.006795Z_{33} - 0.124308Z_{34} - 0.145257Z_{35} \\ &- 0.000107Z_{36} - 0.139279Z_{37} + 0.020443Z_{38} - 0.343712Z_{39} \end{aligned} \tag{11}$$

Based on comprehensive analysis of these NLPCA results, although there exists the influence of white Gaussian noise, from the feature of the first principal component, Bus18 corresponds with variable Z_{18}, and the coefficient of Z_{18} is 0.474260, which is also the biggest one. Consequently, Bus18 is just the fault component. This conclusion is also entirely identical with the fault set in advance.

These instances have fully proven that fault identification of fault component and non-fault component in complex electrical engineering can be performed by nonlinear principal component analysis and calculation. The results of nonlinear principal component analysis are accurate and reliable.

5 CONCLUSIONS

Noise is random in nature and has different attributes depending on its origin. In most of the works on communications, the transmitted data is assumed to be corrupted by Gaussian noise. The Gaussian model is successful in modeling some important random phenomena such as thermal noise and leads to tractable equations [18–20]. The fault is inevitable in any complex systems engineering. In general, a fault is a deviation from the normal behavior in the equipment or its components, and it is a process abnormality or symptom. A system with faults does not necessarily imply that the system is not functioning. Detecting a fault involves identifying a characteristic of the system, when a fault occurs, which can be distinguished from other characteristics of the system.

NLPCA is a novel technique for multivariate data analysis, similar to the PCA. In many industrial engineerings, it may be more appropriate to assume that the feature subspace is defined by nonlinear functions of the process variables. Electric power system is essentially a typically nonlinear system. In this paper, based on the real-time measurements of phasor measurement unit, under the influence of white Gaussian noise (suppose the standard deviation is 0.01, and the mean error is 0), we used mainly NLPCA to resolve fault identification problem in complex electrical engineering. The simulation results show that the fault in complex electrical engineering is usually corresponding to the variable with the maximum absolute value coefficient in the first principal component. These researches will have significant theoretical value and engineering practical significance.

Acknowledgements

This research was supported partly by the National Natural Science Foundation of China (50837002, 50777016), the Science Foundation for the Outstanding Doctors of NCEPU and Fundamental Research Funds for the Central Universities.

References

[1] WANG, Z. P.—ZHANG, Y. G.—ZHANG, J. F.—MA, J. : Recent Research Progress in Fault Analysis of Complex Electric Power Systems, Advances in Electrical and Computer Engineering **10** (2010), 28.

[2] ZHANG, Y. G.—WANG, Z. P.—ZHANG, J. F.—MA, J. : PCA Fault Feature Extraction in Complex Electric Power Systems, Advances in Electrical and Computer Engineering **10** (2010), 102.

[3] ZHANG, Y. G.—WANG, Z. P.—ZHANG, J. F.—MA, J. : Fault Localization in Electrical Power Systems: A Pattern Recognition Approach, International Journal of Electrical Power and Energy Systems **33** (2011), 791, DOI:10.1016/j.ijepes.2011.01.018.

[4] CAO, J. : Principal Component Analysis Based Fault Detection and Isolation, PhD thesis of George Mason University of Virginia, 2004.

[5] TONG, C. Q.—LU, D. : Detection of Multiple Faults using SSFTS in CMOS Logic Circuits, Computers & Electrical Engineering **21** (1995), 271.

[6] MUENCHHOF, M.—BECK, M.—ISERMANN, R. : Fault-Tolerant Actuators and DrivesStructures, Fault Detection Principles and Applications, Annual Reviews in Control **33** (2009), 136.

[7] MESEGUER, J.—PUIG, V.—ESCOBET, T.—SALUDES, J. : Observer Gain Effect in Linear Interval Observer-Based Fault Detection, Journal of Process Control **20** (2010), 944.

[8] LEE, Y. K. : A Fault Diagnosis Technique for Complex Systems using Bayesian Data Analysis, PhD thesis of Georgia Institute of Technology, 2008.

[9] WANG, Y.—LI, W. Y.—LU, J. P. : Reliability Analysis of Phasor Measurement Unit using Hierarchical Markov Modeling, Electric Power Components and Systems **37** (2009), 517.

[10] SODHI, R.—SRIVASTAVA, S. C.—SINGH, S. N. : Phasor-Assisted Hybrid State Estimator, Electric Power Components and Systems **38** (2010), 533.

[11] SODHI, R.—SRIVASTAVA, S. C.—SINGH, S. N. : Optimal PMU Placement Method for Complete Topological and Numerical Observability of Power System, Electric Power Systems Research **80** (2010), 1154.

[12] PENG, C. H.—SUN, H. J.—GUO, J. F. : Multi-Objective Optimal PMU Placement using a Non-Dominated Sorting Differential Evolution Algorithm, International Journal of Electrical Power & Energy Systems **32** (2010), 886.

[13] PHADKE, A. G.—THORP, J. S. : Synchronized Phasor Measurements and their Applications, Springer Verlag, 2008.

[14] PHADKE, A. G.—THORP, J. S. : Computer Relaying for Power System, Second edition, John Wiley & Sons Ltd, Chichester, 2009.

[15] GLASCOCK, M. D. : Characterization of Archaeological Ceramics at MURR by Neutron Activation Analysis and Multivariate Statistics, Prehistory Press, Madison, 1992.

[16] JACKSON, J. E. : A Users Guide to Principal Components, John Wiley and Sons, New York, 1991.

[17] ZHU, J. P. : Applied Multivariate Statistical Analysis, Science Press, Beijing, 2006.

[18] LI, T. : Signal Detection in Unknown Interference Channels, PhD thesis of Hong Kong University of Science and Technology, 2006.

[19] BHATIA, V.—MULGREW, B. : Non-Parametric Likelihood based Channel Estimator for Gaussian Mixture Noise, Signal Processing **87** (2007), 2569.

[20] WANG, Y. G. : Nonlinear Statistics to Improve Signal Detection in Generalized Gaussian Noise, Digital Signal Processing **18** (2008), 444.

Received 16 February 2011

Yagang Zhang (Assoc Prof, PhD) received his BS degree in Department of Mathematics from Yunnan University, Kunming, China in 2001 and the MS degree in Center for Nonlinear Complex Systems from Yunnan University, Kunming, China in 2004. He received the PhD degree from North China Electric Power University, Beijing in 2011. His current research topics include relay protection of power system and nonlinear complex system theory.

Zengping Wang (Prof, PhD) received the BS and MS degree in electric engineering from North China Electric Power University, Baoding, China in 1985 and 1988, respectively, and the PhD degree from Harbin Institute of Technology, Harbin, China in 1997. He is a Professor and the Dean of the School of Electrical and Electronic Engineering, North China Electric Power University. His special fields of interest include power system equipment protection, fault analysis and wide-area protection.

Jinfang Zhang (PhD) received his BS degree in Department of Electrical Engineering from North China Electric Power University, Baoding, China in 2007, and the PhD degree from North China Electric Power University, Beijing, China in 2012. His current research topics include power system relay protection.