

Power System State Estimation Bad Data Detection and Identification: A Review on Issues and Alternative Formulations

Nurul Fauzana Imran Gulcharan^{*1}, Nursyarizal Mohd Nor², Taib Ibrahim³, Hanita Daud⁴

^{1,2} Department of Electrical and Electronic Engineering, Universiti Teknologi PETRONAS,
32610, Seri Iskandar, Perak, Malaysia

^{3,4} Department of Fundamental and Applied Sciences, Universiti Teknologi PETRONAS,
32610, Seri Iskandar, Perak, Malaysia

*Corresponding author, e-mail: nurulfauzana88@gmail.com

Abstract

State Estimation (SE) is the main function of power system where Energy Management System (EMS) is obliged to estimate the available states. Power system is a quasi-static system and hence changes slowly with time. Dynamic State Estimation (DSE) technique represents the time deviation nature of the system, which allows the forecasting of state vector in advance. Various techniques for DSE are available in the literature. This paper presents a review on different methodologies and developments in DSE, based on comprehensive survey of the available literature. From the survey it can be concluded that there are still areas in the developing DSE that can still be improved in terms of system computational time, redundancy and robustness of the system.

Keywords: state estimation, dynamic state estimation, distribution system state estimation, bad data detection and identification

Copyright © 2017 Institute of Advanced Engineering and Science. All rights reserved.

1. Introduction

State Estimation (SE) was introduced into the power system field by Schweppes et al, in the early 1970s, becoming a main function for the EMS by supplying database of the system in real-time state to execute other EMS functions [1-3]. SE involves in gathering real-time measurement data (line flows, injection and voltage measurements) via SCADA and calculates the vector state using the predetermined SE algorithm [4-5].

Transmission System (TS) SE (TSSE) normally takes voltage magnitude and phase angle as state variables. Since 1960s, computers are used for on-line load flow analyses of main distribution systems from approximate load models. Techniques used in TS are indirectly suitable to the distribution network from the exceptional features such as being under radial and weakly-meshed operations as well as numerous unbalanced three phase branches with high r/x ratios and unbalanced loads separated by short distances. Adaptation of generated algorithms for TSSE is necessary to suit the Distribution System State Estimation (DSSE) [6].

In distribution network, real-time measurements are limited and pseudo measurements are used to ensure network observability is achieved. Original studies on DSSE estimators carried out in 1990s in adaptation with low measurement coverage have been using the branch-current or node voltage variables either in polar or rectangular forms as state variables [6].

As the system varies, the generation as well needs to be tuned which will then change the flow and injection across the system. This will create a complete dynamic system. In order to have an uninterrupted monitoring of PS, the estimation needs to be executed at short intervals. However the expansion of PS caused the system to become exceedingly large for the static SE to be carried out at short intervals due to burdensome of computing resource consumption. The above highlighted issue leads to the development of algorithm set known as "Dynamic State Estimation" (DSE) techniques [4].

DSE holds several advantages in terms of being precise and possesses the capability to forecast which enables security analysis to be performed one step ahead and control actions can be performed by enabling the identification and rejection of measurement errors, thus

improving the estimator potential. DSE provides high quality values that avoids ill conditioning in conditions where pseudo measurements are to be implemented. DSE can be used for data validation as the states are predicted one time stamp ahead. Lastly, with assistance from the predicted state vector, any sudden changes, topological and other anomalies in the system can be identified [4].

Other than standard measurement uncertainty, supplementary measurement errors can occur, i.e. due to incorrect parameterization of the measurement device. Since these incorrect measurement values conflictingly impact the SE results, a system for bad data detection, identification (BDDI), and handling, i.e. correcting the detected error, is required which presents the issues or limitations available in the BDDI area and the available alternative formulation that has been implemented by researchers [7].

2. Bad Data Detection, Identification (BDDI) and Related Limitations

Negative measurement errors have impacts on the SE outcomes, making a must for BDDI and handling system which is an error correction method developed and tested with respect to the impact on the estimation of the system state for the distribution grid (DG). The DSSE provides insufficient real-time measurement data making the SE state changes from an overdetermined to an underdetermined system [7].

BDD involves in determining whether the measurement set carries any error which are usually categorized as single bad data (where only one measurement in the whole set of measurements have many errors), or multiple bad data (multiple large errors) which is further classified as multiple either interacting or non-interacting bad data. Furthermore, a system's state can be estimated indirectly in the existence of error using mathematical and intelligent computational approaches which process includes detecting, identifying and removing error [8–9].

Conventional approaches in solving the BDD is usually by mathematical approaches which includes chi-square distribution test, the largest normalized residual (LNR) as well as hypothesis testing identification (HTI). Meanwhile, example methods for intelligent computational includes heuristic methods that requires either intensive training under different conditions or computationally demanding methods [8]. Example of bad data includes negative voltage magnitudes, measurements with some orders of magnitude larger or smaller than presumed values, or, large differences between incoming and leaving currents at a connection node within a substation [8]. Distribution network covers a considerably large area and nodes therefore; solution from DSSE algorithm for the issues is needed to meet the requirements for near real-time applications [8]. Main issues and limitations are categorized as below.

2.1 Computation Time and Complexity

Deficiency of time measurement is a crucial issue in DSSE development that has been highlighted in several field tests due to the insufficient measurement leads to difficulty in providing the observability of the network and quality results [10].

Distribution system states have become more dynamic due to integration of intermittent distributed generation (DG). Therefore, BDD issues become more complicated. Distribution system may contain single or multiple errors. The conventional approach for SE is based on weighted least squares (WLS). This method is able to detect and identify single and multiple bad data (not interacting only) by using normalized residuals of the measurements. However, the method still fails with conforming bad data.

2.2 Measurement Redundancy

Measurement redundancy is defined as a measure of extra measurements over needed minimum number for system state estimation. With adequate redundancy, SE will be to detect, identify and remove bad data. However, the level of redundancy required may not always have attained due to the existing measurement configurations. Low local redundancy will result in several critical measurements available in the system, but data related with critical measurements cannot be detected, identified or removed [11].

2.3 Robustness Detection of Multiple Interacting and Consistent Bad Data

In some cases, interacting multiple bad data is difficult to detect. LNR based statistical criterion may have issues in accurately identifying as well as eliminating this category of bad data especially conforming type. Issues on conventional LNR for multiple interacting bad data (conforming) is where the successive elimination of the measurement may effect in eliminating true measurements instead of bad data.

Identification of the said category perchance carried out by multiple application of LNR test which is satisfactory even for multiple interacting bad data, but may cause misidentification as well as miselimination of good measurements, when errors are mutually conforming (consistent) [12].

WLS-based state estimators are only developed by using a combination of linearized measurement function with complex computations. However, the combination is also not robust in the existence of single and multiple bad data (interacting and non-interacting) especially conforming [14].

2.4 Examples of Formulation with Setbacks

Many methods were put forward to enhance the identification method but some estimator may not be able to reject the error and it may lead to the presence of leverage in the PS model and generates computational burdens by taking the unique properties of the estimation equations into consideration.

In 1992, Recursive Measurement Error Estimation Identification (RMEEI) and RMEEI for BDI is proposed where state variables, residuals and parameters can be updated subsequent to eliminating a measurement from the suspicious data set to the existing data set by implementing a set of linear recursive equations. A set of residual equations used by conventional approach can only be applied to the linear systems by dividing the raw measurements. This however may result in burden of operation calculation and complexity because each of the part is composed of some measurements [13].

The proposed Weighted Least Absolute Value Estimator (WLAV) method is able to handle multiple gross errors efficiently, but it is likely to fail in the existence of a single gross error at a Leverage Point (LP). Meanwhile, the Least Median of Squares Estimator (LMS) is essentially unaffected by the outliers in LP and is able to attend to the multiple interacting gross errors, even when conforming. However, it is in need of extreme computing time for on-line applications. Changeable weighting matrix in 2003 is proposed to recognize the errors however only applies to static SE [14].

Thus, it can be concluded that the BDDI still have room for improvements in the area of detecting multiple interacting bad data (conforming), measurement redundancy, estimation robustness and excessive computing time that leads to computation burden and complexity of the system.

3. Alternative Formulation of BDDI

Key objectives of derived formulations in BDDI is mainly to improve the computation time, having less mathematical complexity as well as reducing system redundancy.

DSSE is in short of real-time measurement data leading to changes in SE changes from an overdetermined to an underdetermined state. M. Cramer et al. addresses issue of inaccurate measurements (bad data) for distribution grid state estimation (DGSE) by implementing Artificial Neural Networks (ANNs) based DGSE detection method which is an extended version of error correction taken from [15] and [16]. The method has proven to be responsive towards uncorrected measurement errors and able to detect and accurately recognize single and multiple measurement error values. The proportion of varying DGSE performance is considerably minimized as well as the mean square error of the estimated system state is also lessened to a fraction by implementing the error correction process [7].

GK. Venayagamoorthy G. K. proposed a new intelligent system, Cellular Computational Network (CCN) for decentralized predictive modelling and DSE of a power system from synchrophasor data. It is shown that the CCN-based DSE is flexible towards loss of the single and multiple interacting and non-interacting bad data from one or more PMUs. Energy management system (EMS) applications could be executed in real-time with a high confidence level using the CCN-based DSE [8].

A. Primadianto et al. implements the combination of chi square (λ^2) test and LNR-test methods for BDDI as well as to test the performance of Node Voltage Polar Magnitude (NVPM) and Branch Current Rectangular Magnitude (BCRM) in combination of a proposed technique but only to handle the single and multiple non-interacting bad data. However, multiple interacting bad data is required to be solved with a more advanced method, and residual sensitivity matrix can be utilized to enhance the BDD function of DSSE [10].

Jian Chen and Ali Abur proposed a PMU placement algorithm to change the current critical measurements into repetitive estimations in order to enhance the error processing ability of state estimators by exploiting the advantage of PMU technology by a feasible numerical technique to find an ideal procedure of PMU positioning. From the executed method, the system is said to still be observable even after eliminating any one of the previously critical measurements as well as detecting the bad data in previously critical measurements. The PMU placement is much more efficient compared to conventional measurements placement. Each PMU is able to measure the bus voltage and current phasor along most or the entire lines incident to the bus [11].

LNR based statistical criterion may have issues in accurately identifying as well as eliminating the multiple interacting bad data primarily conforming. The author proposed three non-deterministic solution which are Basic Genetic algorithms (bGAs), Micro Genetic Algorithms (μ GA) and Evolution Strategy (ES). These procedures are based on Genetic Algorithms (GA) to reduce computation burden and to improve numerical efficiency. All above stated GA based techniques proved to provide satisfactory behavior in BDI. bGAs is shown to be utmost dependable in identifying the error even with high computational demand. All procedures exhibit satisfactory features yielding to a feasible, observable and bad data free solution or optimal within the first few iterations [12].

A proposed technique which worked well on BDI is known as the Extended Complex Kalman Filter Artificial Neural Network (ECKF-ANN) with The ECKF – ANN method performs with fast computational speed in two stages. Firstly, it is compared with Real Back-Propagation ANN (RBP-ANN) and secondly with the Complex Back-Propagation ANN (CBP-ANN). The stage is performed to detect and assess the performance in terms of convergence addition and noise rejection for efficiency on BDD. The method is able to look out for the appropriate and applicable training variables and able to converge with less computational time as well as better capacity of noise rejection than the conventional algorithms [13].

Bretas, N.G. Bretas, A.S. uses WLS based Geometrical Approach where the detectability of gross errors in PSSE is provided and achieved by decomposing the measurement error into undetectable and detectable parts. A method of recovering the masked errors resulting from the measurement residual estimating process was proposed based on the UI index. The total measurement residual is then composed. Proposed measurement gross error detection and identification test using the measurement composed gross error, works efficiently. However, the results indicated that the previously implemented gross error is incorrect where the existing masking effect in PSSE is not taken into account, therefore the work progresses on multiple gross errors detection or identification [14].

Shyh-Jier Huang and Jeu-Min Lin applied the gap statistic algorithm (GSA) with combination of neural networks (NN) approach for investigation of error detection in power system. From the proposed technique, abnormal data can be searched of intelligently and automatically without the specific thresholds. The proposed technique was validated with automation, simplicity and success for data debugging from computational results which can be further developed by enhancing the quality of data acquisition process, or to be use as a reference for power system control and operations [17].

Largest studentized residual (LSR) was proposed for detection and identifying the gross error utilizes the statistical properties of calculated residuals and modified where it uses the mean square error (MSE) observations in its calculation. For the treatment of multiple non-interacting bad measurements, a sequential application of the LSR test is proposed, bad measurement is compensated for error but not deleted and the bad value is treated one at a time. Proposed scheme has greatest power for treating the bad measurements as compared to joint implementation of conventional tests. To save system from observability problem and information loss, the given technique was applied in sequential way by replacing the grossly measurements one at a time by their estimates rather than their deletion from

measurements [18]. However, this method is only proven efficient for only single and multiple non-interacting bad data and was not implemented for the multiple interacting bad data.

A pre-estimation for BDD based on the Principal Component Analysis (PCA), reducing the time and cost in comparison with ANN approaches, that could be very tedious and expensive is presented. The PCA method find relations between variables that describe the main data behavior, and builds a statistical model that represents the system that uses statistical indexes to detect and identify bad data or anomalous data. It is concluded that PCA algorithm detects and identifies the fundamental structure of gross and historic errors before SE, allowing a faster and more reliable convergence. The PCA algorithm works satisfactorily with both types of bad data, single and multiple, and results show how single and multiple bad data are detected even when the deviation of the bad data is lower than $\pm 4\sigma$. Given an added advantage, the technique does not require information about the topology and parameters of the system, where it can work in real time, but the results show how the detection sensitivity depends obviously on the existing data used to build the model [19].

A. Alamin et al. conducted a comparison between the WLS and Iterative Extended Kalman Filter (IEKF) to investigate the advantages between of static and dynamic state estimators. The operational robustness of IEKF was extended by completing it with Largest Normalized Residual Test (LNRT). Based on the simulation results, IEKF showed slightly higher errors than WLS eventhough both methods achieved reasonable estimation accuracy. When the system was only concerned with minimum interruption due to repetitive Kalman function, IEKF could be more accurate than WLS. The IEKF algorithm to enhance its bad data detection capability of the IEKF algorithm is enhanced when LRNT was applied. IEKF was unable to accurately identify the location of bad data eventhough the detection scheme was able to determine faulty measurements [20].

Different from [6], this author employs Projected Unscented Kalman Filter (PUKF) technique based on unscented transformation (UT) to achieve the goal of considering zero injection restriction to deal with nonlinearities in DSE. It is stated that the PUKF manage to obtain better approximation to the nonlinearity of PS by unscented transformation and able to take zero injection constraints into account by including estimate projection. The PUKF based DSE is able to detect different kinds of bad data, sudden load changes and topology errors by implementing the weighted innovation test. However, supplementary investigation for the proposed technique is required for bad data identification [21].

A geometrical analysis (GA) based WLS demonstrated that the measurement residue is the composed measurement error (CME) where it is a portion of the measurement error. The GA method is considered where it is proposed to detect, identify and correct gross based on statistical analysis of CME. The GA method was demonstrated being more reliable than Residual Analysis (RA) mainly in identification of multiple interacting gross errors especially conforming that existed in leverage points. By implementing the GA and RA methods with given load models, the study aims to investigate the bad data analysis performance comparison in systems with AMI and pseudo-measurements. Considering measurements and pseudo-measurements is obtained at a Global Redundancy Level (GRL) of 1.72. This is lower than the redundancy levels obtained from TS which intefers with the bad data analysis. In proposed assessment, GA approach performed better than RA in BDDI process of gross errors where the gross error level influences BDD whereas the load models influences the gross error correction. GA Concluded to performed well and provides a potential improvement to BD analysis of smart distribution systems, even with low GRL [22].

The Chi-Squares test is a common post-estimation bad data detection test which commonly carried out by almost all WLS estimators. However, Chi-Squares test fails to detect available bad data in the measurement set in some cases. Therefore, the author proposed a basic modification that will enhance the BDD capability where the modification requires calculation of residual covariance matrix. Comparison between proposed method and conventional Chi Squares method is conducted in terms of computational performance and bad measurement. The results showed the proposed metric obtained a superior performance in comparison to the conventional test in detecting presence of bad data. However, for identification and elimination of measurements errors will still need to be conducted by means such as normalized residuals test [23].

Bad data detection for smart grids is mainly given only at the central level due to limitations in legacy technologies employed in many substations. The authors proposed a

substation level BDD algorithm that upholds the advantages featured by the IEC61850 standard. The proposed algorithm is based on automatically detecting the substation topology by determining the standard substation description files and online state of circuit breakers and disconnectors. The BDD employs the linear WLS based LNR state estimation algorithm (identification by elimination method) where bad data from failing Current Transformers (CT) can be detected. For the case of bad data identification, the bad data were replaced by the estimated states, however, requires excessive computational time since the dimension of the matrices used for calculation will be changed. Before using the estimated state, the redundant protective Current Transformers (CTs) data were used to replace the bad data in the bad data handling part, thereby the capability of the bad data detection can be maintained. The proposed BDD algorithm is verified by using different bad data scenarios where the results present the output of the proposed algorithm provides the smallest error compared with using either measurement CT or protective CT output in both static and dynamic situations. Upon the incapability to identify the bad data sources in some cases, the algorithm still managed to notify the presence of bad data to users [24].

Authors Zakerian A. et al. adapts the Chi-square method as well as the Probabilistic Neural Network (PNN) and Decision Tree (DT) which are used to detect BDD in state estimation. The combination of proposed method was simulated on IEEE30 bus network and applied to multiple scenarios. It is shown that the developed PNN and DT are able to detect bad data in SE. Different combinations of residual measurements were utilized as input features for PNN and DT training. The best bad data detection accuracy was found using the PNN method by four types of input features. The accuracy of these methods is compared with the Chi-square method. Results indicated that the traditional Chi-Square method may fail to give accurate detection whereas the PNN and DT method were able to be utilized for bad data detection. The obtained results verified that by implementing the data mining techniques, the accuracy of bad data detection could be promoted [25].

4. Conclusion

Real-time monitoring and control of power systems is highly crucial for efficient and reliable operation of a power system. SE forms the main function for real-time monitoring and control functions. The operator has to be exceptionally alert in obtaining decisions on real time since PS changes continuously. A DSE technique provides the predicted state vector at the next time instant to the operator, in which the operator will be able to take suitable control actions. Various DSE techniques proposed in the literature were primarily divided into two types which are mathematical and intelligent computational approach. The paper discussed methods such as Chi-square and LNR test, Modified Chi-Square test, ANN based DGSE, CCN based DSE, PMU based algorithm, Basic Genetic Algorithm, many forms of Kalman Filter technique (Iterative Extended based LNR, Extended Complex based ANN Projected unscented based UT), WLS based Geometrical Approach, Largest Studentized Residual, Principle Component Analysis and many more. The advantages, disadvantages and specialties of the discussed methods have been briefly described in this paper to achieve the objective in obtaining a solution method for issues available in BDDI. Thus, concluding that the BDDI still have room for improvements in the area of detecting multiple interacting bad data (conforming), measurement redundancy, estimation robustness and excessive computing time that leads to computation burden and complexity of the system.

References

- [1] Schweppe FC, Wildes J, Rom DB. Power System Static State Estimation. *Power Syst. Eng. Group*, M.I.T. Report number: 10. 1968.
- [2] Schweppe FC and Wildes J. Power System Static-State Estimation, part I: Exact model. *IEEE Transactions on Power Apparatus and Systems*. 1970; PAS (89): 120-125.
- [3] Schweppe FC, Rom DB. Power System Static-State Estimation, part II: Approximate model. *IEEE Transactions on Power Apparatus and Systems*. 1970; PAS (89): 125-1330.
- [4] Shivakumar NR, Jain A. *A Review of Power System Dynamic State Estimation Techniques*. Joint International Conference on Power System Technology and IEEE Power India Conference, New Delhi, 2008:1-6.

- [5] Mehrdad Tarafdar Hagh, Seyyed Mehdi Mahaei, Kazem Zare. Improving Bad Data Detection in State Estimation of Power Systems. *International Journal of Electrical and Computer Engineering (IJECE)*. 2011; 1(2): 85-92.
- [6] Primadianto A, Lu CN. A Review on Distribution System State Estimation. *IEEE Transactions on Power Systems*. 2016; PP(99):1-1
- [7] Cramer M, Goergens P, Schnettler A. *Bad Data Detection and Handling in Distribution Grid State Estimation Using Artificial Neural Networks*. IEEE Eindhoven PowerTech. Eindhoven. 2015: 1-6
- [8] Venayagamoorthy GK. Computational Approaches for Bad Data Handling in Power System Synchrophasor Networks. *IFAC Proceedings Volumes*. 2014; 47 (3): 11269-11274.
- [9] Seyed Mahdi Mahaei, Mohammad Reza Navayi. Power System State Estimation with Weighted Linear Least Square. *International Journal of Electrical and Computer Engineering*. 2014; 4(2): 169-178.
- [10] Primadianto A, Lin WT, Lu CN. *Performance Comparison of Distribution System State Estimation Methods*. IEEE Innovative Smart Grid Technologies-Asia (ISGT-Asia). Melbourne, VIC, 2016: 1121-1126
- [11] Jian Chen, Abur A. Improved bad data processing via strategic placement of PMUs. *IEEE Power Engineering Society General Meeting*. 2005; 1: 509-513.
- [12] Gastoni S, GP Granelli GP, Montagna M. Multiple Bad Data Processing By Genetic Algorithms. *IEEE Bologna Power Tech Conference Proceedings*. 2003; 1. 6.
- [13] Huang CH, Lee CH, Shih KR, Wang YJ. *Extended Complex Kalman Filter Artificial Neural Network for Bad-Data Detection in Power System State Estimation*. International Conference on Intelligent Systems Applications to Power Systems. Toki Messe, Niigata. 2007: 1-7.
- [14] Bretas NG, Bretas AS. *Bad Data Analysis Using the Composed Measurements Errors for Power System State Estimation*. IREP Symposium Bulk Power System Dynamics and Control – VIII. Rio de Janeiro. 2010: 1-7.
- [15] Salehfar H, Zhao R. A Neural Network Preestimation Filter For Bad-Data Detection and Identification In Power System State Estimation. *Electric Power Systems Research*. 1995; 34:127–134.
- [16] Shyh-Jier Huang and Jeu-Min Lin. Artificial Neural Network Enhanced by Gap Statistic Algorithm Applied for Bad Data Detection of a Power System. *IEEE/PES Transmission and Distribution Conference and Exhibition*. 2002; 2. 764-768.
- [17] Souza JCS, da Silva AML, da Silva APA. Online Topology Determination and Bad Data Suppression In Power System Operation Using Artificial Neural Networks. *IEEE Trans. Power Systems*. 1998; 13. 796–803.
- [18] Khan Z, Razali R, Daud H, Nor NM, Firuzabad MF. The Largest Studentized Residual (LSR) Test for Bad Data Identification in State Estimation of A Power System. *ARPN Journal of Engineering and Applied Sciences*. 2015; 10(21). 10184-10191.
- [19] Gerard Vancells, Joaquim Meléndez, Sergio Herraiz, Juan Prieto, Guillermo Bravo. *Bad Data Detection and Identification in Distribution Power Systems By Means Of Principal Component Analysis*. 21st International Conference on Electricity Distribution. Frankfurt. 2011: 1-4.
- [20] Alamin A, Khalid HM, Peng JCH. *Power System State Estimation Based On Iterative Extended Kalman Filtering And Bad Data Detection Using Normalized Residual Test*. IEEE Power and Energy Conference. Champaign, IL. 2015: 1-5
- [21] Yang Y, Hu W, Min Y. *Projected Unscented Kalman Filter For Dynamic State Estimation And Bad Data Detection In Power System*. 12th IET International Conference on Developments in Power System Protection (DPSP 2014). Copenhagen. 2014: 1-6.
- [22] Braunstein SH, Bretas NG, Rossoni A, Bretas A. *Bad data Analysis In Distribution State Estimation Considering Load Models*. IEEE Power & Energy Society General Meeting. Denver, CO. 2015: 1-5.
- [23] Göl M, Abur A. *A Modified Chi-Squares Test for Improved Bad Data Detection*. IEEE Eindhoven PowerTech, Eindhoven. 2015: 1-5.
- [24] Y Wu, Y Xiao, F Hohn, L Nordstrom, J Wang, W Zhao. Bad Data Detection Using Linear WLS and Sampled Values in Digital Substations. *IEEE Transactions on Power Delivery*. 2017; PP(99): 1-8.
- [25] Zakerian A, Maleki A, Mohammadnian Y and Amraee T. *Bad Data Detection in State Estimation Using Decision Tree Technique*. Iranian Conference on Electrical Engineering (ICEE), Tehran. 2017: 1037-1042.