ENHANCEMENT OF POWER GRID MONITORING BASED ON DATA WEIGHTING

Parisa Ataeian, Abbas Rabiee
Department of Electrical Engineering, University of Zanjan, Zanjan, Iran
{ataeian.parisa, rabiee}@znu.ac.ir

Mehdi Derafshian Maram, Mohsen Ghalei Monfared Zanjani
Iran Grid Management Company, Tehran, Iran
{derafshian, monfared.m}@imc.ir

ABSTRACT
With their expansion, national power grid have had to work with huge sets of data received from a vast number of substations and power plants. Given their large volume and variety, these data can be classified as big data. Managing this massive amount of data is certainly challenging. Depending on the application, parts of these data are more important for real-time network operation. Computing a network’s observability score without assigning weights to different signals may not provide a complete picture of the received data’s validity and thus lead to incorrect assessments of the network status. Consequently, signals critical to the network operation and functions of an energy management system (EMS) should be assigned higher weights in observability calculations. The weighted observability alongside the classic non-weighted observability can serve as an indicator of each area’s condition in comparison to that of other areas and so greatly facilitate the monitoring and verification of transmitted data. For calculating a weighted observability, the current paper presents a method based on the Analytic Hierarchy Process (AHP), in which higher weights are assigned to data that are more valuable for and widely used by operators. The national electricity network of Iran was chosen for the evaluation of the proposed method’s effect on data quality and operational risk. The results indicated that the introduced method positively affects the quality of received data and also corrects erroneous data in the network.

Keywords  Observability · Operation quality · Information weighting · Information quality · Big data · AHP

1 Introduction

With the expansion of national power grids during the past decades, there has been a substantial rise in the amount of data received from stations as well as a growing dependence on these data for controlling power systems. Due to their large volume and variety, these data can be classified as big data [1]. In regard to working with such big data, the volume, variety, and retrieval speed of the data are crucial. Furthermore, the accuracy of these data cannot be overlooked, because incorrect information may cause operator errors [2]. The lack of access to reliable information was a major cause of large blackouts in North America and parts of southern Canada in 2003 as well as in Europe in 2006 [3][4].

For a power network to operate safely, it is necessary that all network variables always be within the allowable range and that the system be able to maintain its normal operation after a wide variety of potential events. These events can be predicted with expert tools, such as an energy management system (EMS) with application modules, an example of which is state estimation. Depending on the conditions of the power grid, data collection and data transfer to the control center may encounter various problems. Solving or at least alleviating these problems can greatly contribute to the quality, safety, and security of the network operation [5].

*Corresponding Author
The collection, transmission, and monitoring of data from power plants and substations are usually performed with two systems: WAMS and SCADA. WAMS offers excellent features, such as data synchronization, linearity of equations, dynamic analyzability, and the need for less equipment that results in a lower failure risk. However, due to its high costs and lack of a suitable telecommunications infrastructure, WAMS cannot completely replace SCADA [6]. Therefore, the usage of both systems is common [7, 8]. The present article focuses on SCADA.

Among all the different kinds of data received in dispatching centers, those that are widely used in the control center are a priority [9, 10]. Any interference with proper data collection and retrieval (e.g. delays in retrieval, poor condition of equipment, and incorrect measurements) can seriously disrupt the state estimation process [11, 12, 13]. The “observability” of a station (power plant or substation) or area indicates what percentage of all data sent to the control center is correct. In observability estimations, it is common to consider all of the data sent from a station or area. However, the presence of data that appear to be true, but are in fact false, can cause errors in observability estimations which may subsequently cause operator mistakes. Such apparently valid but actually invalid data can be identified through the state estimation process [14, 15, 16].

In the rest of the present article, Section 2 describes the method for identifying invalid data and classifying operationally important data. Section 3 provides an introduction to the Analytic Hierarchy Process (AHP) and explains how it is used in the current study. Section 4 presents the results of applying weighting coefficients to observability and Section 5 concludes the paper.

2 Identification and classification of important data from the operators’ perspective

2.1 Importance of data from the operators’ perspective

In the online monitoring and control of power networks, up-to-date data from network components are crucial in addition to knowing the correct status of the operating instructions. Given the massive size of national networks, it is absolutely necessary for operators to determine the status of network operating points through the state estimation module [9]. To estimate the real-time operating point of the power system, SCADA’s state estimation module employs the raw measurement data collected from SCADA subsystems, including the voltage size and angle of all network buses [17, 18]. In order to obtain the right output from this module, appropriate inputs must be given, including network topology (presence or absence of components), the generating power of the units, the voltage of certain buses, the load of certain transformers, the load of capacitors and reactors, and the tap of transmission transformers. Obviously, the accuracy of the state estimation module’s output largely depends on the quality of these inputs and measurements [7, 19, 20].

2.2 Data tagging

Table 1 provides a list of potential status tags of data received in the control center. The received data may be either valid or invalid for various reasons.

| Definition                                                                 | Signal Tag     |
|---------------------------------------------------------------------------|----------------|
| No information on the data status                                        | Faulty (F)     |
| Data show the last status before disconnection                            | Non-current (N)|
| Valid data (these data are not tagged)                                   | Valid (V)      |
| Data appear to be valid but are invalid (these data are not tagged)       | Valid (V)      |
| Data are made invalid so as to be disregarded in the state estimation process | Invalid (I)   |
| Important data that are sometimes manually entered by operators.          |                |
| Because of the constant changes in the network operating point, these data must obviously be constantly updated. | Manually (M)   |

Apparently valid but inherently invalid data can lead to poor operating performance and decisions and, unfortunately, these data cannot be easily identified as they have no specific tag or label. In the state estimation process, one of SCADA’s functions is to identify data points when there is a significant difference between estimated values and measurements.
2.3 Selecting important data

The received data are selected from among the data that are important for different parts of the network operation, that is, according to the needs assessment and experiences of operators as well as the inputs required by the state estimation module. Table 2 presents the classification of signals by the type of quantity.

| Data                  | Tags     |
|-----------------------|----------|
| Megawatt              | MW       |
| Mega Volt*Amps Reactive| MV       |
| Kilovolt             | KV       |
| TAP                  | TAP      |
| Status               | STATUS   |

Table 3: Classification of signals by the type of component

| Signals | Components                     |
|---------|--------------------------------|
| 1       | Unit Transformer and Load Transformer |
| 2       | Transmission Transformer       |
| 3       | Generator                      |
| 4       | Transmission Line              |
| 5       | Reactor, Capacitor             |
| 6       | Busbar                         |

In this classification, signal Output MW, for example, has the same value as signal Line MW. Therefore, these signals should be differentiated. Table 3 shows the differentiation of signals by the type of component. In order to weight the signals, the set of received data must be examined from two perspectives: 1) type of quantity (unit) and 2) the component they come from. Since not all components use all the quantities listed in Table 2, the quantity related to each component is considered, as shown in Table 4.

| Component                                      | Related quantity                          |
|------------------------------------------------|-------------------------------------------|
| Unit Transformer and Load Transformer          | MW, MV, TAP, and STATUS                  |
| Transmission Transformer                      | MW, MV, TAP, and STATUS                  |
| Generator                                      | MW, MV, KV, and STATUS                   |
| Transmission Line                             | MW, MV, KV, and STATUS                   |
| Reactor, Capacitor                            | MV and STATUS                            |
| Busbar                                         | KV and STATUS                            |

3 Proposed formulation and algorithm

3.1 Observability

The classic definition of observability does not differentiate the received data in terms of their importance. Equation (1) provides the formula of the observability index.

\[ OV_i = \frac{AD_i - OB_i}{AD_i}, \]

where \( AD_i \) is the total information that is sent by state \( i \) to control center and \( OB_i \) is incorrect data. Table 5 presents another problem that calculates observability while giving the same weight to two stations (transmission substation and a power plant station).

While the two stations have the same observability according to the control center’s experts, Station B is in better condition because MW of unit and shunt reactor condition are more important than bus frequency and kV of line.
Table 5: Difference between two stations with the same observability index

| Station | Uncorrect signals | #Uncorrect signals | #Total signals | Observability indicator |
|---------|-------------------|--------------------|----------------|-------------------------|
| A       | MW of unit and breaker of head line reactor | 2                  | 100            | %98                     |
| B       | Bus frequency and kV of line               | 2                  | 100            | %98                     |

3.2 Proposed weighting approach

To solve the above problems, the current paper proposes weighting the data according to their importance to the network operation and the state estimation process. The present study conducts this weighting through a procedure based on Analytic Hierarchy Process (AHP).

3.3 Analytic Hierarchy Process (AHP)

Individuals may make decisions without considering all aspects of the issue at hand or when influenced by others or their own personal background. It is, therefore, advisable that the decision-making process employ methods which are more accurate, comprehensive, and rational. One of the most commonly used is the Analytic Hierarchy Process (AHP). In AHP, available options are classified according to several criteria and decisions are made by comparing pairs of options in terms of each criterion [21, 22]. With the assistance of several experts in the field, the present study utilized AHP to determine the operational importance of each datum. Since not all components have all of the considered quantities, AHP was applied separately for each component and each quantity. After obtaining the evaluations of each expert, the current work produced the final evaluations by averaging. The obtained averages were considered as the basis for determining the weight assigned to the signals.

3.4 Questionnaire

A questionnaire was prepared for the pairwise comparison of signals. Considering the number of operators in the studied network, the present study provided its questionnaire to nine experienced operators as well as experts in network operations and state estimation modules. Also examined was the issue of real-time operation control from the perspectives of the frequency control, voltage control, and reactive power of the network.

4 Results and analysis

For its evaluation, the proposed weighted observability method was applied to the Iranian electricity network. The network consists of ten areas, from A to J, each of which may be more or less critical depending on geographical factors (important exchange points), political factors, economic factors (industry), and the size of the network (the number of stations and signals). The current study assigned the same importance to all of these areas. Since the signals of critical regional stations are usually considered in regional exchange line operating instructions, these signals were given higher weights than other signals.

4.1 Signal weighting based on expert inputs

The results of each expert completed questionnaire were imported into the AHP software to determine the signal weights recommended by that expert. For example, Table 6 reports an expert’s opinion on the importance of signals received from a generation unit from the viewpoint of the state estimation module.

| Measure | Weight |
|---------|--------|
| MW      | 41%    |
| MV      | 23.3%  |
| KV      | 40%    |
| Status  | 31.7%  |

For this component, the examined signals are MW, MVAR, kV, and connected/disconnected status. As seen, the unit’s MW signal was more important than its kV for the state estimation module. The connected/disconnected status of the unit and its MVAR were not as important. Obviously, the expert may not have given the same ratings to these
Enhancement of Power Grid Monitoring Based on Data Weighting

signals because of other reasons. Similar pairwise comparisons were also made between the components from the quantity point of view. After collecting the results of the questionnaires and importing the data into the software, the current work averaged the resulting ratings, the results of which are listed in Tables 7 and 8. In the averaging of the ratings, all of the expert inputs were considered to be of equal importance. The product of values given in Tables 7 and 8 was considered as the final weight of operationally important signals. Table 2 and 3 and 4 reports the importance of each quantity for each component. For example, the most important quantity for a transmission transformer was MW, followed by MVAR, TAP, and Status. As seen, because of the normalization of the geometric means’ results, the weights in each column summed up to 100.

Table 7: Resulting rating based on information type

| M | Unit Transformer and Load Transformer | Transmission Transformer | Generator | Transmission Line | Reactor, Capacitor | Busbar |
|---|-------------------------------------|---------------------------|-----------|-------------------|-------------------|-------|
| MW | 57.5                                | 45.07                     | 31.5      | 59.97             | -                 | -     |
| KV | 17.63                               | 21.8                      | 23.27     | 24.53             | 61.1              | -     |
| TAP | 4.87                               | 18.37                     | 7.37      | 6.27              | -                 | 71.1  |
| STATUS | 19.83                          | 14.8                      | 17.9      | 9.27              | 38.87             | 28.9  |

Table 8: Resulting rating based on type of component

| N | MW | MV | KV | TAP | STATUS |
|---|----|----|----|-----|--------|
| Unit Transformer and Load Transformer | 13.07 | 8.53 | -   | 32.2 | 12.03  |
| Transmission Transformer | 15.33 | 19.43 | -   | 67.77 | 16.93  |
| Generator | 54.57 | 49.97 | 39.93 | - | 40.97 |
| Transmission Line | 17.03 | 8.6 | 14.53 | - | 16.5 |
| Reactor, Capacitor | - | 13.47 | - | - | 9.07 |
| Busbar, Capacitor | - | - | 45.53 | - | 4.6 |

Table 2, 3, and 4 provides the importance of each component from the perspective of each quantity. To simultaneously consider the importance of components and quantities, the weight of each signal in the observability calculations was defined as \( M \times N \). For example, the weight of KV for the transmission lines in all stations was determined as \( 14.53 \times 6.27 = 91.10 \). After calculating all weights in the same way, the present study observed that the KV of the busbar and the MW of the generator were the most important signals. As the voltage levels of all substations were displayed on a dedicated screen, it was obvious that these signals were important for operation. The megawatt output of the units was also critical for frequency control. The signal with the lowest importance was found to be the KV of the transmission lines; this was predictable given the redundancy of this data in each station.

4.2 Observability index

After calculating all the signal weights, the present study obtained the weighted observability by inserting the final weights in Equation 2:

\[
OV_{i_{new}} = \frac{AD_{i_{new}} - OB_{i_{new}}}{AD_{i_{new}}} \times 100, \tag{2}
\]

where \( OV_{i_{new}} \) is the observability indicator and \( AD_{i_{new}} \) calculated as follows:

\[
AD_{i_{new}} = M_i \times N_i \times AD_i \times 2 \tag{3}
\]

In Equation 2, \( AD_{i_{new}} \) sum of all weighted signals (obtained from Equation 3). Also, \( OB_{i_{new}} \) in Equation 2 is the sum of all weighted invalid data:

\[
OB_{i_{new}} = M_i \times N_i \times OB_i \times 2 \tag{4}
\]

The multiplier of 2 is applied when the signal is present in the instruction so as to ensure that its importance is duly noted. During the calculation of the weighted observability index, it should be noted that data considered unimportant from an operating perspective are disregarded. Aside from this reason, such signals are also discounted in part because of their large number which makes pairwise comparisons impossible.

4.3 Observability results for the Iran power grid

Both weighted observability and non-weighted observability were calculated for the Iran power grid, which consists of 16 areas. The following compares these two observability values to demonstrate the effect of differentiating important
Enhancement of Power Grid Monitoring Based on Data Weighting

signals. Table 9 presents the observability of the considered network with and without the weight coefficients. It can be seen that, in the non-weighted approach, area A ranked first and area P ranked last in terms of observability. In any case, computing observability without weighting the signals is impractical and does not meet the needs of operators in the safety area. In the current work, applying the weights to observability significantly decreased the observability of areas N, M, L, J, D, and O. In other words, the non-weighted observability of these areas was not useful from an operational point of view and many of their invalid data were actually important data. In contrast, there was no change in area B’s observability, which indicates that the non-weighted observability of this area was adequate from an operational perspective. Applying the weights also changed the ranking of the areas. As shown in Table 9, Areas C and D both had a non-weighted observability of 97%, while Area C had more favorable conditions from an operational standpoint. Illustrated in Figure 1, the results of Table 9 show a significant decline in the observability of many areas after the weights were applied to the observability formula.

| Area | Without weighting | With weighting |
|------|-------------------|---------------|
| A    | %99               | B             |
| B    | %98               | A             |
| C    | %97               | C             |
| D    | %97               | F             |
| E    | %96               | K             |
| F    | %95               | E             |
| G    | %95               | H             |
| H    | %93               | G             |
| I    | %92               | D             |
| J    | %90               | I             |
| K    | %88               | J             |
| L    | %88               | L             |
| M    | %87               | P             |
| N    | %86               | M             |
| O    | %84               | N             |
| P    | %80               | O             |

4.4 Analysis of 3-month observability

The weighted and non-weighted observability results were sent to the areas with invalid signals and, after three months, the observability of these areas was reexamined. Table 10 reports the areas’ weighted and non-weighted observability indexes after this period. As the results demonstrate, most areas showed significant growth in observability. As illustrated in Figure 2, Areas F, B, and L experienced decreased observability. This may be related to long periods of disconnection in these areas because of some maneuvers occurring in those areas. observability remained unchanged in Areas H, D, C, and A, while it improved in other areas. Since this diagram alone cannot indicate changes in important data (signals), the changes in weighted observability are plotted in Figure 3. Figure 3 shows that, except for Areas B and H, where observability decreased, and areas F and L, where it remained unchanged, the rest of the areas earned better observability scores than they had three months earlier.

5 Conclusion

The present paper presents a weighted observability calculation method which places more emphasis on receiving valid data for signals considered more critical by operators and the state estimation module. To identify these important signals, nine experts in real-time operation control were asked to rate signals while taking into account the inputs required in the state estimation module. The experts completed questionnaires prepared for the pairwise comparisons of important signals, the results of which were analyzed by AHP software for determining each expert’s signal weights. The final signal weights were obtained by averaging these individual weights and then were applied to the observability formula. With the performance of these steps, periodic reports of important invalid signals can be sent to their origins.
so that the data may be corrected. The result of this process will be the continuous correction of erroneous data that produces a network with high observability and more reliability from an operator’s point of view. Furthermore, while the classic non-weighted approach to observability ignores the possibility of apparently valid but actually invalid data, the weighted approach ensures that such data are properly considered in observability calculations. Another advantage of the proposed approach is its attention to correcting certain data over others. It is crucial to assign a higher priority to more important data, especially in DCS substations where data transmission is prioritized. By improving network observability, these modifications can reduce faults and errors due to faulty operator performance. Providing more accurate input data to the state estimation module also lowers the convergence error and accelerates the process of real-time assessment and control.
Enhancement of Power Grid Monitoring Based on Data Weighting

Figure 3: Comparison of two weights observability after three months

Table 10: Observability of the network (after 3 months)

| Area | Information quality indicator | Area | Information quality indicator |
|------|-------------------------------|------|-------------------------------|
| A    | %99                           | A    | %98                           |
| B    | %97                           | B    | %97                           |
| C    | %97                           | C    | %96                           |
| D    | %97                           | E    | %95                           |
| E    | %97                           | K    | %92                           |
| G    | %96                           | G    | %91                           |
| F    | %94                           | F    | %90                           |
| H    | %93                           | D    | %89                           |
| I    | %93                           | I    | %89                           |
| J    | %90                           | H    | %88                           |
| K    | %90                           | J    | %84                           |
| M    | %88                           | L    | %79                           |
| N    | %88                           | N    | %79                           |
| L    | %86                           | O    | %79                           |
| P    | %85                           | M    | %78                           |
| O    | %82                           | P    | %78                           |

References

[1] Houda Daki, Asmaa El Hannani, Abdelhak Aqqa, Abdelfattah Haidine, and Aziz Dahbi. Big data management in smart grid: concepts, requirements and implementation. *Journal of Big Data*, 4(1):1–19, 2017.

[2] Edmund Handschin, Fred C Schweppe, Jurg Kohlas, and AAFA Fiechter. Bad data analysis for power system state estimation. *IEEE Transactions on Power Apparatus and Systems*, 94(2):329–337, 1975.

[3] John F Hauer, Navin B Bhatt, Kirit Shah, and Sharma Kolluri. Performance of “wams east” in providing dynamic information for the north east blackout of August 14, 2003. In *IEEE Power Engineering Society General Meeting*, 2004., pages 1685–1690. IEEE, 2004.

[4] Göran Andersson, Peter Donalek, Richard Farmer, Nikos Hatzigiayiou, Innocent Kamwa, Prabhashankar Kundur, Nelson Martins, John Paserba, Pouyan Pourbeik, Juan Sanchez-Gasca, et al. Causes of the 2003 major grid blackouts in North America and Europe, and recommended means to improve system dynamic performance. *IEEE Transactions on Power Systems*, 20(4):1922–1928, 2005.
Enhancement of Power Grid Monitoring Based on Data Weighting

[5] Felix F Wu, Khosrow Moslehi, and Anjan Bose. Power system control centers: Past, present, and future. Proceedings of the IEEE, 93(11):1890–1908, 2005.

[6] Arun G Phadke. Synchronized phasor measurements in power systems. IEEE Computer Applications in Power, 6(2):10–15, 1993.

[7] Dalu Li, Rui Li, Yuanzhang Sun, and Han Chen. Recognizing and correcting the wrong parameters in state estimation considering the wams measurements [j]. Automation of Electric Power Systems, 14, 2008.

[8] Mao Anjia, Yu Jiaxi, and Guo Zhizhong. Pmu placement and data processing in wams that complements scada. In IEEE Power Engineering Society General Meeting, 2005, pages 780–783. IEEE, 2005.

[9] Kate L Morrow, Erich Heine, Katherine M Rogers, Rakesh B Bobba, and Thomas J Overbye. Topology perturbation for detecting malicious data injection. In 2012 45th Hawaii International Conference on System Sciences, pages 2104–2113. IEEE, 2012.

[10] Hyde M Merrill and Fred C Schweppe. Bad data suppression in power system static state estimation. IEEE Transactions on Power Apparatus and Systems, (6):2718–2725, 1971.

[11] Willy W Kotiuga and M Vidyasagar. Bad data rejection properties of weighted least absolute value techniques applied to state static estimation. IEEE Transactions on Power Apparatus and Systems, (4):844–853, 1982.

[12] Tian Hongxun, Wang Honggang, Zhou Kun, Shi Mingtai, Li Haosong, Xu Zhongping, Kang Taifeng, Li Jin, and Cai Yaqi. Data quality assessment for on-line monitoring and measuring system of power quality based on big data and data provenance theory. In 2018 IEEE 3rd International Conference on Cloud Computing and Big Data Analysis (ICCCBDA), pages 248–252. IEEE, 2018.

[13] DM Falcao, PA Cooke, and A Brameller. Power system tracking state estimation and bad data processing. IEEE Transactions on Power Apparatus and Systems, (2):325–333, 1982.

[14] A Monticelli and A Garcia. Reliable bad data processing for real-time state estimation. IEEE Transactions on Power Apparatus and Systems, (5):1126–1139, 1983.

[15] William L Peterson and AA Girgis. Multiple bad data detection in power system state estimation using linear programming. In The Twentieth Southeastern Symposium on System Theory, pages 405–406. IEEE Computer Society, 1988.

[16] Felix F Wu, W-HE Liu, and S-M Lun. Observability analysis and bad data processing for state estimation with equality constraints. IEEE Transactions on Power Systems, 3(2):541–548, 1988.

[17] Ali Abur. A bad data identification method for linear programming state estimation. IEEE Transactions on Power Systems, 5(3):94–901, 1990.

[18] Yih-Fang Huang, Stefan Werner, Jing Huang, Neelabh Kashyap, and Vijay Gupta. State estimation in electric power grids: Meeting new challenges presented by the requirements of the future grid. IEEE Signal Processing Magazine, 29(5):33–43, 2012.

[19] Savu Crivat Savulescu. Real-time stability assessment in modern power system control centers, volume 42. John Wiley & Sons, 2009.

[20] A Mazhabjafari, M Kabiri, Kh Khangholi, and S Ghiami. Feasibility study of running state estimation software in a large power system with a very weak rtu coverage condition. In 2010 Asia-Pacific Power and Energy Engineering Conference, pages 1–4. IEEE, 2010.

[21] Bruce L Golden, Edward A Wasil, and Patrick T Harker. The analytic hierarchy process. Applications and Studies, Berlin, Heidelberg, 2, 1989.

[22] Thomas L Saaty. Group decision making and the ahp. In The analytic hierarchy process, pages 59–67. Springer, 1989.