IOP Conf. Series: Materials Science and Engineering 1019 (2021) 012034 doi:10.1088/1757-899X/1019/1/012034

Application of standard ZET-algorithm for gaps filling in wide area measurement system data assets

N L Batseva and J A Foos
Power Engineering School, Tomsk Polytechnic University, Tomsk, Russia
E-mail: JuliaAlekseevna6797@gmail.com

Abstract. This article explores the application of Standard ZET-algorithm for gaps filling in Synchronized Phasor Measurement data assets received from Wide Area Measurement Systems. In spite of Synchronized Phasor Measurement data assets apparent advantages, they have gaps affecting the accuracy of solving power engineering tasks. Authors analysed the reasons of gap occurrence in data assets, provided experiments with voltage and current data assets and collected the key points about application of Standard ZET-algorithm.

1. Introduction

Synchronized Phasor Measurement (SPM) data assets received from Wide Area Measurement Systems (WAMS) are in widespread use for solving technological tasks in power systems [1]. These tasks cover the operational dispatch management and automatic emergency control:

a) improving accuracy of power system state estimation [2];
b) on-line identifying a point of power system emergency separation into parts and then their synchronization;
c) monitoring margin of power system stability;
d) choosing volume of action adjustments for emergency response.

Definitely, SPM data assets have unquestionable advantages:

a) obtained from one node, they allow calculating the voltage of neighbouring nodes by load current values of a line, increased power system observability;
b) they give an opportunity to control the phase imbalance, as the signal, when equipment tend to accident-caused failure [2];
c) high discreteness of regime parameters provides the possibility to analyse electromechanical transients for centralized and local automatics according to type of stability [3].

However, SPM data assets include noises and gaps inside, leading to errors in result data, especially, for technological tasks based on data analysis methods, and making a contribution to wrong conclusions and taking incorrect solutions [4, 5]. This fact underlines the relevance of the research.

2. The reasons of gap occurrence in SPM data assets

In order to understand the nature of gaps, we analyzed the collecting process of SPM data assets received from WAMS (Figure 1).

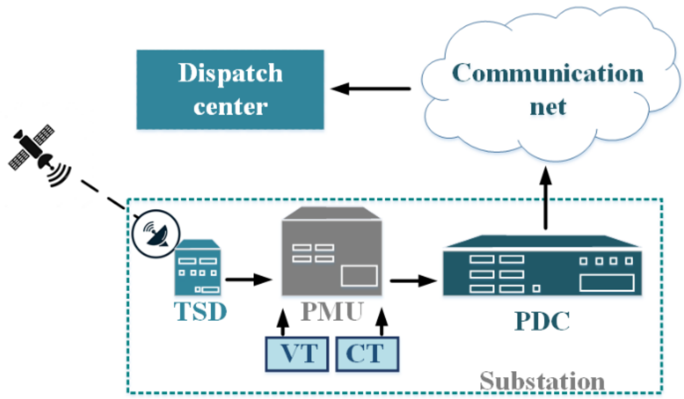

Figure 1. The collecting process of SPM data assets.

According to Figure 1, WAMS consists of three main devices: time synchronization device (TSD); phasor measurement units (PMU); and Phasor Data Concentrator (PDC).

PMU records SPM data assets received from current (CT) and voltage transformers (VT) set at a power station and/or substation. PMU assign the record time tag to every parameter in data assets. TSD generates time tags under synchronization of satellites [6]. After preprocessing SPM data assets in PDC, they are transmitted to the server of power system control center via communication net.

At the stage when data assets are recorded, gaps can be appeared due to PMU technical failures. It requires restarting PMU but it is not possible to restore PMU normal operation immediately since the additional adjustments are required. Moreover, the operating staff cannot fully complete these adjustments therefore required period of time for restoration is significantly increased.

Time synchronization malfunctions can also lead to gaps in data assets due to satellite communication channel failures. They cause by technical reasons or weather conditions in a location of power facilities. For example, low temperatures and high humidity provide icing transceiver facilities, consequently, the connection between a satellite and TSD temporary disappears. As a result the time tag shifts. Now this problem has short-term solution by integrating the special function to the program code of the system server. Unfortunately, this method of data synchronization cannot guarantee the absence of gaps in SPM data assets [4].

WAMS communication net needs to transmit SPM data assets from measurement and registration location to the server of power system control center. During transfer SPM data assets in on-line regime, they are often lost. These are because of connection loss between PDC and the server of power system control center; clients' reconnection; communication net reconfiguration; operation failures of net equipment – routers, hubs, modems, servers; aggregations of SPM data assets received from several PMU to one protocol packet [5].

Specialists of System Operators and Joint-stock companies identified the dependence of gaps amount from the time of day (Figure 2). They calculated the proportion of gaps in SPM data assets as percentage of total amount of received data in 10-minutes intervals per 24 hours period.

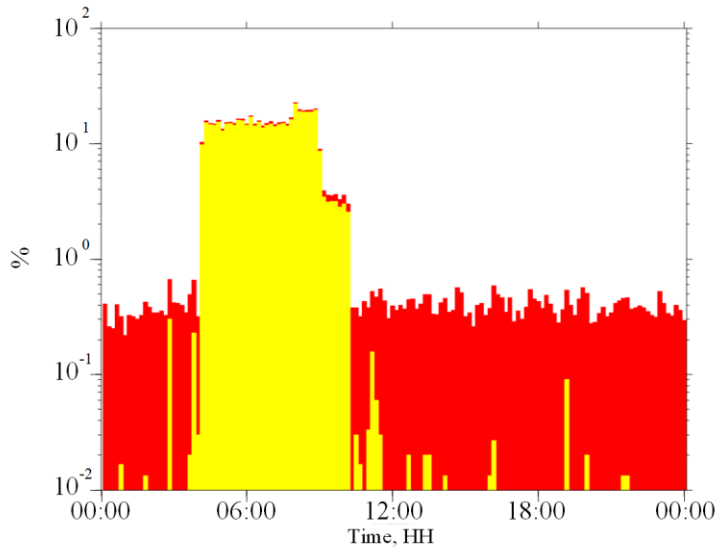

Figure 2. The proportion of gaps in SPM data assets.

Figure 2 shows total amount of gaps in assets: grouped (gaps come one after another) and single gaps (red color); only grouped gaps (yellow color).

The most part of gaps appear at the beginning of a working shift, when clients' reconnections and communication net reconfiguration take frequently place. At the peak load period about 20% of data are lost. Therefore, these data assets can be used for technological tasks only after their processing by special gaps filling algorithms [7].

3. The algorithm for gaps filling

Nowadays the algorithms for gaps filling are constantly developing [7, 8]. The simplest algorithms recommend deleting those rows and columns in data assets which have even one gap [9]. The complicated algorithms use methods of least squares and maximum likelihood [10], and EM-algorithm [11]. All of them are considered to be global ones. It means that the specified type of the dependence is relevant for the whole data assets.

Apart from global algorithms there are local algorithms which define the dependence type according to incomplete selection around a gap. These are standard ZET-algorithm and WANGA-algorithm [7, 8].

This article illustrates application of standard ZET-algorithm. The choice of ZET-algorithm is due to hypotheses which lay in the core of the algorithm and the possibility of application for value forecasting in "time-property"-type time series.

Standard ZET-algorithm consists of three main hypotheses:

a) the excessiveness: it visualizes that SPM data assets are redundant caused by conformity among rows and depending on each other columns;
b) local competence: it states that in order to fill the gap, the analysis of a whole data asset is not needed, but only "competent" part including rows similar to i-row and columns similar to j-column are taken into account;
c) linear dependency: a dependence among rows and columns must be linear.

Standard ZET-algorithm has the preliminary stage, and then three main stages are realized *l*-times.

At the preliminary stage, columns of SPM data asset are normalized according to dispersion for adjusting the properties to unified scale (1):

$$a'_{ij} = A_j^{-1} \cdot (a_{ij} - \bar{a}_j) \tag{1}$$

Where a'_{ij} is adjusted value; a_{ij} is initial value; $\bar{a}_j$ is value of mathematical expectation; A_j is mean square deviation.

Assumed that unknown gap has x and y coordinates (Figure 3).

At the first main stage, from the parent SPM data asset where columns have been normalized according to dispersion by equation (1), for an unknown gap (Saxon blue color), a sub multitude of "competent" rows (blue color) is selected. Then, for "competent" rows, the "competent" columns (blue color) are selected to produce a "competent" matrix. The matrix order is S_{ij}, when $i \in [1; p]$ and $j \in [1; q]$. We must select (p-1) "competent" rows for y-row with an unknown gap and (q-1) "competent" columns for x-column with an unknown gap.

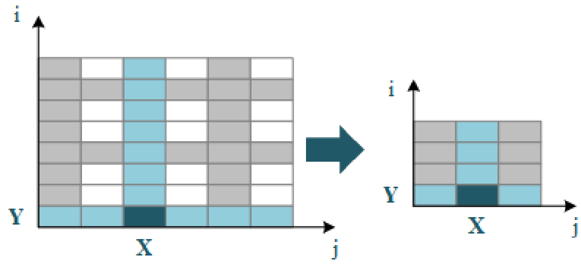

Figure 3. The first stage of ZET-algorithm: asset normalized according to dispersion and "competent" matrix.

The i-row "competence" L_{iy} in relation to the y-row is defined by (2):

$$L_{iy} = t_{iy} / r_{iy} \tag{2}$$

where i is "competent" row number; t_{iy} is completeness that is the number of values known for i- and y-rows; r_{iy} is Descartes' distance between rows.

The "competent" y-row must not contain the gap on the x-position.

The j-column "competence" L_{iy} in relation to the x-column is defined as (3):

$$L_{jy} = |k_{jx}| \cdot t_{jx} \tag{3}$$

where t_{jx} is completeness that is the number of values known for the columns j and x; k_{jx} is correlation coefficient for j- and x-columns. It should be emphasized that for k_{jx} calculation we must use those column values which belong to competent rows.

The "competent" column must not contain the gap on the y- position.

At the second stage, reasonable coefficients σ_p and σ_q are defined for the goal of row and column forecasting gap. This procedure is similar as for rows as for columns. Given coefficient a (a=σ_p for rows and a=σ_q for columns within stated limits and with stated step size, we minimize the function (4):

$$\sum |a_{ik} - b_{ik}| \rightarrow \min, \tag{4}$$

where a_{ik} is i-component real value of k-row (k-column) with gap; b_{ik} is i-component forecasting value calculated by (5):

$$b_{ik} = \sum_{j=1}^{p-1} bl_{jk} \cdot L_{ij}^{\sigma_p} \Big/ \sum_{j=1}^{q-1} L_{ij}^{\sigma_q}, \tag{5}$$

where bl_{jk} is k-row (k-column) forecasting value defined as simple regression by the least squares method.

IOP Conf. Series: Materials Science and Engineering 1019 (2021) 012034 doi:10.1088/1757-899X/1019/1/012034

At the last stage, we forecast the gap value: for "competent" rows by (6):

$$b_y = \sum_{i=1}^{p-1} bl_{iy} \cdot L_{iy}^{\sigma_p} \Big/ \sum_{i=1}^{p-1} L_{iy}^{\sigma_p}, \qquad (6)$$

and for "competent" columns by (7):

$$b_x = \sum_{i=1}^{q-1} bl_{ix} \cdot L_{ix}^{\sigma_q} \Big/ \sum_{i=1}^{q-1} L_{ix}^{\sigma_q}. \qquad (7)$$

The restored gap value is calculated by (8):

$$b_{xy} = 0.5 \cdot (b'_{xy} + b''_{xy}). \qquad (8)$$

The error of the gap value is defined by (9):

$$e_{xy} = \left|B_{xy} - b_{xy}\right| / \left|b_{xy}\right|, \qquad (8)$$

where B_{xy} is initially known value.

4. The application of standard ZET-algorithm for gap filling in SPM data assets

We researched the application of standard ZET-algorithm by SPM data assets of current and voltage modules and phases. It should be noted that the data of current and voltage phases play a key role when the task of power system state estimation is decided. Meanwhile, it is impossible to understand, whether the current and voltage phase values are trusted or not, since a phase values can dynamically change from +π to –π during the time (Figure 4).

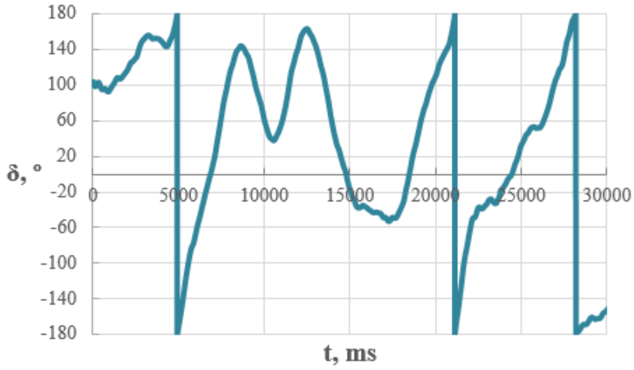

Figure 4. Dynamical changes of phase values during the time.

All researches are provided by SQL Server Management Studio.

SPM data assets are received from PMU installed at 500kV substation 1 in 500 kV power lines Substation 1 – Substation 2; Substation 1 – Substation 3, and in 220 kV power line Substation 1 – Substation 4 (Figure 5).

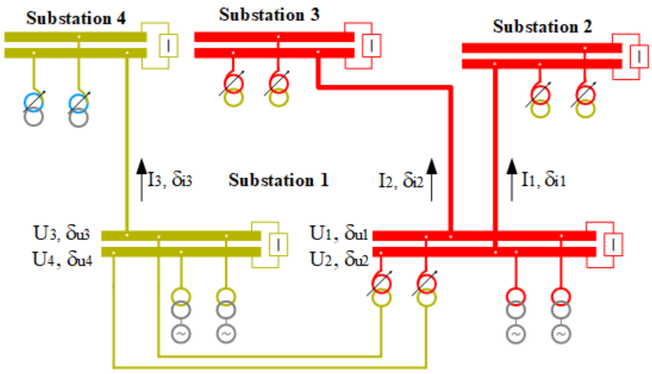

Figure 5. The scheme of power grid.

SPM data assets include 42 000 values of synchronized voltage and current vectors U_A, U_B, U_C, δ_{Ua}, δ_{Ub}, δ_{Uc}, I_A, I_B, I_C, δ_{Ia}, δ_{Ib}, δ_{Ic} per 10 minutes with sample spacing 5 values per a second.

Two test assets are created based on the real SPM data assets. The first test asset consists of 25% of single gaps and 5% of grouped gaps, and the second one includes 5% of single gaps and 25% of grouped gaps (in following tables see asset I and asset II respectively).

The values of the gaps are calculated by (1–8). The error for the gap value is defined by (9). Tables 1 and 2 demonstrate the results of calculations.

Table 1. Result of calculations for gap filling in assets of voltage and current modules.

Test asset	Type of gaps	Error, %
Asset I	single	0.078
	grouped	0.002
Asset II	single	0.033
	grouped	0.046

Table 2. Result of calculations for gap filling in assets of voltage and current phases.

Test asset	Type of gaps	Error, %
Asset I	single	0.335
	grouped	0.798
Asset II	single	0.280
	grouped	0.485

Tables 3 – 6 illustrate the real and calculated phase values for assets I and II with the lowest error of calculation for both single and grouped gaps.

Table 3. Result of voltage and current phase values calculations for asset I. Single gaps.

Real δ°	**ZET-algorithm δ°**	**Δ δ°**	**Error, %**
0.67252	0.57999	0.09253	15.953
-1.71186	-2.07783	0.36597	17.613
-1.54104	-2.01156	0.47052	23.39
2.04425	1.64436	0.39989	24.3188

Table 4. Result of voltage and current phase values calculations for asset I. Group gaps.

Real δº	ZET-algorithm δº	Δ δº	Error, %
3.06494	1.92793	1.13701	58.975
3.22679	1.92827	1.29852	67.341
3.38853	1.93140	1.45713	75.444
-3.20492	-4.27932	1.0744	25.106
-3.03814	-4.24894	1.2108	28.496
-2.87093	-4.23185	1.36092	32.159
-35.77520	-36.79924	1.02404	2.782
-2.71039	-4.21641	1.50602	35.718

Table 5. Result of voltage and current phase values calculations for asset II. Single gaps.

Real δº	ZET-algorithm δº	Δ δº	Error, %
-0.39773	-0.47312	0.07539	15.934
1.02544	0.92375	0.10169	11.008
1.34082	1.22206	0.11876	9.718
2.23318	2.11331	0.11987	5.672
-5.47433	-5.81228	0.33795	5.814

Table 6. Result of voltage and current phase values calculations for asset II. Group gaps.

Real δº	ZET-algorithm δº	Δ δº	Error, %
9.16659	8.14266	1.02393	12.574
9.33588	8.15297	1.18291	14.509
9.51735	8.15548	1.36187	16.698
27.58634	26.22689	1.35945	5.183
27.71748	26.22881	1.48867	5.675
27.85212	26.22936	1.62276	6.186
9.16659	8.14266	1.02393	12.574
9.33588	8.15297	1.18291	14.509

Table 7 shows the results of gap filling for voltage and current phases if the phase values variate at the boundary of [-180º; +180 º] diapason.

Table 7. Results of phase values calculations. Phase values are changing at the boundary of [-180°; +180°] diapason.

Real δº	ZET-algorithm δº	Error, %
-179.77574	**144.92745**	**224.045**
-179.25549	-179.32759	0.0402
-179.15950	-179.27198	0.0627
-179.02856	-179.10581	0.0431
-179.70389	**89.97538**	**299.725**
-179.77574	**144.92745**	**224.045**
179.85089	179.77194	0.043
179.81882	179.66501	0.085
179.12233	179.03520	0.0486
179.41821	179.41525	0.001
179.31870	179.13311	0.103
179.37088	179.35614	0.0082
179.50395	179.43094	0.04

5. Conclusion

Proposed researches and results of experiments demonstrate that:

a) the excessiveness: it visualizes that SPM data assets are redundant caused by conformity among rows and depending on each other columns;
b) the main advantage of standard ZET-algorithm is the robustness against occurrence of grouped gaps in SPM data assets;
c) results of experiments from tables 1 and 2 illustrate that in assets when grouped gaps are dominant, they can be adequately filled by standard ZET-algorithm (Asset II);
d) the average error for gap filling in assets of voltage and current phases is higher than for assets of voltage and current modules;
e) results of experiments from tables 3 – 6 prove the high calculation accuracy for both single and grouped gaps as in asset of phases as in asset of modules. The difference between real and calculated phase values is less than 2º;
f) in case of phase value is changing near the boundary of [-180°; +180°] diapason, the gap value is calculated with low accuracy. Therefore, standard ZET-algorithm needs to be sophisticated by the algorithm for a “competent” matrix size correction.

References

[1] Phadke A G and Thorp J S 2008 *Synchronized phasor measurements and their applications* (New York: Springer) p 81
[2] Novosel D *et al* 2007 *Proceeding of IEEE 40th Annual Hawaii Intern. Conf. on System Sciences* 118.
[3] Phadke A and Kasztenny B 2009 *IEEE transactions on power delivery* **24** 89–95
[4] Crotti G *et al* 2019 *IEEE Transactions on Instrumentation and Measurement* **68** 1724–1731
[5] Chen Y *et al* 2016 *Dianli Zidonghua Shebei/Electric Power Automation Equipment* **36** 95–101
[6] Zhang Y *et al* 2010 *IEEE Transactions on smart grid* **1** 159–167
[7] Little R, Rubin D 1989 *Sociological Methods & Research* **18** 292–326
[8] Beale E and Little R 1975 *Journal of the Royal Statistical Society* **37** 129–145
[9] Gleason T, Staelin R 1975 *Psychometrica* ***40*** 229–252

[10] Dempster A, Laird N and Rubin D 1977 *Journal of the Royal Statistical Society* **39** 1–38
[11] Kaminskyi R, N. Kunanets N *et al* 2018 *Paper presented at the CEUR Workshop Proceedings* 108–118