BAD DATA PROCESSING FOR LOW VOLTAGE STATE ESTIMATION SYSTEMS BASED ON SMART METER DATA

Dominik WAERESCH,
Robert BRANDALIK,
Wolfram H. WELLSSOW
TU Kaiserslautern
Kaiserslautern, Germany
waeresch@eit.uni-kl.de

Joern JORDAN
IDS Gmbh
Ettlingen, Germany
joern.jordan@ids.de

Rolf BISCHLER,
Nelia SCHNEIDER
Stadtwerke Kaiserslautern
Versorgungs-AG
Kaiserslautern, Germany
rolf.bischler@swk-kl.de

ABSTRACT

The area wide usage of smart meters in low voltage grids enables the identification of the three-phase system state with linear state estimation (SE) systems. In order to localize large measurement errors also bad data detection algorithms have to be applied. But as the measurement redundancy is typically small, the probability of bad data detection is usually small, too. This paper proposes a special three-phase SE approach which enables the reliable detection of bad data on the basis of the well-known normalized residuals method. In contrast to other algorithms active and reactive currents as well as absolute current values are used as input data for a linear SE system. Despite the simplicity of the process the results gathered from simulations and a field test are promising, showing appropriate bad data detection probabilities especially for voltage and active current bad data.

INTRODUCTION

For the optimal operation and control of low voltage (LV) grids, the system state, which equals the complex voltages at each node, must be known at any time. For this case, special LV state estimation (SE) systems have to be developed and applied in practice which shall estimate the network state with sufficient accuracy. Prospective available measurement data from smart meters installed at each and every customer can be used to gather operational network variables which are used as SE input data.

Conventional SE algorithms are usually not usable for LV grids because of a lack of measurement equipment resulting in a negative measurement redundancy η. This is shown by the definition of η with the number of independent measurements M and the number of network nodes N:

$$\eta = \frac{M}{2N-1} - 1 \tag{1}$$

As experience shows, for effective compensation of measurement errors and especially bad data (BD), the measurement redundancy η should attain values of at least 0.5 [1]. In contrast to HV grids η cannot be easily increased by branch current or branch power measurements because LV grids often consist of buried cables. Another difference in LV state estimation is the need for three-phase SE due to asymmetric system states, which leads to a significantly higher computational effort.

The development of a LV SE algorithm is the purpose of the field test project SmartSCADA [2]. It includes a smart meter rollout in a semi-urban LV grid with high penetration of photovoltaic (PV) systems at 120 households and the development of a LV state estimation algorithm based on smart meter data, PV-feed-in forecasts as well as pseudo measurements for unmeasured loads. Voltage and current magnitudes as well as active and reactive currents with sign are measured for all phases. The measurement redundancy for the test grid is $\eta = 0.94$.

LINEAR LV STATE ESTIMATION BASED ON SMART METER DATA

Fundamentals of LV State Estimation

Voltage and current magnitudes, active or reactive powers and currents measured by area-wide installed smart meters can be used as input data for LV SE systems. Within the field test project, a linear SE algorithm has been developed which means that the SE is based on voltage magnitudes as well as active and reactive current measurements. The benefit is that the algorithm is fast and not prone to convergence problems. The accuracy compared to nonlinear approaches is still adequate enough. Due to usually asymmetric loads the LV SE system has to be applied for three-phase network states. In this context the three-phase optimization problem is formulated in symmetrical components (SC) with a positive ('1'), negative ('2') and a zero ('0') sequence system. For all transformations it is assumed that the angles between the line-to-ground voltages are 120° which is a permissible assumption for LV grids. The three-phase system state vector $\mathbf{x}$ in algebraic form is defined in SC as (2). Here $\mathbf{u}_{s,re}$ is the vector with all real parts, $\mathbf{u}_{s,im}$ the vector with all imaginary parts of the complex node voltages $\underline{U}_{s,i}$ at network nodes i in SC system s.

$$\mathbf{x} = \left[\mathbf{u}_{1,re}^T \;\; \mathbf{u}_{1,im}^T \;\; \mathbf{u}_{2,re}^T \;\; \mathbf{u}_{2,im}^T \;\; \mathbf{u}_{0,re}^T \;\; \mathbf{u}_{0,im}^T \right]^T \tag{2}$$

Assuming the weighted least square (WLS) method the general objective function $J(\hat{\mathbf{x}})$ with the estimated system state $\hat{\mathbf{x}}$, the measurement value z_k and the measurement variance σ_k^2 can be described by (3). Thereby the row vector $\mathbf{h}_k^T$ relates measurement z_k to the state vector $\mathbf{x}$.

$$J(\hat{\mathbf{x}}) = \sum_{k=1}^{M} \frac{(z_k - \hat{z}_k)^2}{\sigma_k^2} = \sum_{k=1}^{M} \frac{\left(z_k - \mathbf{h}_k^T \cdot \hat{\mathbf{x}}\right)^2}{\sigma_k^2} \quad (3)$$

In matrix form the objective function results in (4), where $\mathbf{z}$ is the measurement vector in SC (5), $\mathbf{H}$ the measurement model matrix with the linear functions $\mathbf{h}_k$ (6) and where $\mathbf{R}$ contains the measurement variances (7).

$$J(\mathbf{x}) = (\mathbf{z} - \mathbf{H} \cdot \mathbf{x})^T \cdot \mathbf{R}^{-1} \cdot (\mathbf{z} - \mathbf{H} \cdot \mathbf{x}) \quad (4)$$

$$\mathbf{z} = [z_1 \ \cdots \ z_k \ \cdots \ z_M]^T \quad (5)$$

$$\mathbf{H} = [\mathbf{h}_1 \ \cdots \ \mathbf{h}_k \ \cdots \ \mathbf{h}_M]^T \quad (6)$$

$$\mathbf{R} = \mathrm{diag}\left\{\sigma_1^2 \cdots \sigma_k^2 \cdots \sigma_M^2\right\} \quad (7)$$

LV-SE based on the augmented matrix approach

The solution of the WLS optimization problem can be in general obtained by using the well-known augmented matrix approach [1]. The idea behind is to separate the virtual measurements, e.g. sum of currents at a node equals zero, from the regular measurements and write them as equality constraints. Representing virtual measurements with $\mathbf{C} \cdot \mathbf{x}$ and the estimated regular measurements with $\mathbf{H}_R \cdot \mathbf{x}$ the WLS problem can be formulated as shown in (8-10). Here, $\mathbf{r}$ represents the difference vector between actual and estimated values of regular measurements which is called the residual vector. In the end it can be formulated as a Lagrangian function as shown in (11).

$$\text{minimize} \quad J(\mathbf{x}) = \mathbf{r}^T \mathbf{R}^{-1} \mathbf{r} \quad (8)$$

$$\text{subject to} \quad \mathbf{C} \cdot \mathbf{x} = 0 \quad (9)$$

$$\text{and} \quad \mathbf{r} - \mathbf{z} + \mathbf{H}_R \cdot \mathbf{x} = 0 \quad (10)$$

$$\mathcal{L} = J(\mathbf{x}) - \boldsymbol{\lambda}^T \cdot (\mathbf{C} \cdot \mathbf{x}) - \boldsymbol{\mu}^T \cdot (\mathbf{r} - \mathbf{z} + \mathbf{H}_R \cdot \mathbf{x}) \quad (11)$$

Due to two equality constraints, (11) has two sets of Lagrangian multipliers, which are often denoted as $\boldsymbol{\lambda}$ and $\boldsymbol{\mu}$. The Lagrangian function can also be written as a linear matrix optimization problem, which is showed in (12). Here the coefficient matrix is called Hachtel's or augmented matrix. It has good mathematical properties, especially when applying an additional weighting factor α for adjusting $\mathbf{R}$ [1]. The fundamental structure of the measurement Jacobian matrix $\mathbf{H}_R$ in SC is similar to the nodal admittance matrix and described in detail in [3].

$$\begin{bmatrix} \alpha \cdot \mathbf{R} & \mathbf{H}_R & 0 \\ \mathbf{H}_R^T & 0 & \mathbf{C}^T \\ 0 & \mathbf{C} & 0 \end{bmatrix} \cdot \begin{bmatrix} \boldsymbol{\mu} \\ \mathbf{x} \\ \boldsymbol{\lambda} \end{bmatrix} = \begin{bmatrix} \mathbf{z} \\ 0 \\ 0 \end{bmatrix} \quad (12)$$

Consideration of absolute current measurements

By using only voltage magnitude as well as active and reactive current measurements as SE input data, current BDs are not identifiable as there exist no dominating elements in the residual sensitivity matrix for current measurements. Hence, also measured current magnitudes are considered. But as the SE equations have to be linear an approximation has to be used. A complex current $|\underline{I}| \cdot e^{j\varphi_I}$ with absolute value $|\underline{I}|$ and current angle φ_I is shown in (13). Here, the exact absolute value is calculated by the active and reactive currents I_{active} and $I_{reactive}$. It can be shown that $|\underline{I}|$ can be approximately calculated as a linear combination by (14) depending on I_{active}, $I_{reactive}$ and a selectable angle γ.

$$|\underline{I}| \cdot e^{j\varphi_I} = \sqrt{I_{active}^2 + I_{reactive}^2} \cdot e^{j\varphi_I} \quad (13)$$

$$|\underline{I}'| \cdot e^{j\varphi_I} = (\cos\gamma \cdot I_{active} + \sin\gamma \cdot I_{reactive}) \cdot e^{j\varphi_I} \quad (14)$$

The relative approximation error e_{rel} is calculated by (15) and equals zero when $\gamma = \varphi_I$ is chosen. A deviation between γ and φ_I of 30° results in a relative error of 10 %,

$$e_{rel} = \frac{|\underline{I}| - |\underline{I}'|}{|\underline{I}|} = 1 - \cos(\gamma - \varphi_I) \quad (15)$$

With this approximation the SE equations for absolute current values can be formulated in SC by (16) where s represents the SC system. Here it is assumed that the voltage $\underline{U}_{1,k}$ in the positive sequence system is placed on the real axis of the complex plane.

$$\left|\underline{I}'_{s,k}\right| = \cos\varphi_{I,s,k} \cdot \mathrm{Re}\left(\underline{I}_{s,k}\right) + \sin\varphi_{I,s,k} \cdot \mathrm{Im}\left(\underline{I}_{s,k}\right) \quad (16)$$

The measured three-phase absolute currents $\left|\underline{I}_{Lx,k}\right|$, Lx representing the phase, also have to be transformed in SC by (17) where $\underline{T}$ is the well-known SC transformation matrix. For this, also the current angle $\varphi_{I,Lx,k}$ is used which is calculated by the measured active and reactive current. In this context, voltage angles are assumed as zero which is a permissible assumption in LV grids.

$$\begin{bmatrix} \left|\underline{I}'_{1,k}\right| \cdot e^{j\varphi_{I,1,k}} \\ \left|\underline{I}'_{2,k}\right| \cdot e^{j\varphi_{I,2,k}} \\ \left|\underline{I}'_{0,k}\right| \cdot e^{j\varphi_{I,0,k}} \end{bmatrix} = \underline{T} \cdot \begin{bmatrix} \left|\underline{I}_{L1,k}\right| \cdot e^{j(\varphi_{I,L1,k} + \Delta\varphi_{I,L1,k})} \\ \left|\underline{I}_{L2,k}\right| \cdot e^{j(\varphi_{I,L2,k} + \Delta\varphi_{I,L2,k} + 240°)} \\ \left|\underline{I}_{L3,k}\right| \cdot e^{j(\varphi_{I,L3,k} + \Delta\varphi_{I,L3,k} + 120°)} \end{bmatrix} \quad (17)$$

As active and reactive currents are affected by measurement errors e_I, also $\varphi_{I,Lx,k}$ has an additional error angle. It is denoted as $\Delta\varphi_{I,Lx,k}$ and calculated by (18). The maximum value $\left|\Delta\varphi_{I,Lx,k}\right|_{max}$ depending on the standard deviation of current measurement errors is achieved when

$e_{I,active} = -e_{I,reactive}$ or vice versa. Fig. 1 shows the maximum absolute current angle error as a function of the absolute current.

$$\Delta\varphi_I = \tan^{-1}\left(\frac{I_{reactive} + e_{I_{reactive}}}{I_{active} + e_{I_{active}}}\right) - \tan^{-1}\left(\frac{I_{reactive}}{I_{active}}\right) \quad (18)$$

The effect of current angle errors in three-phase system onto the absolute current values in SC is shown in Fig. 2. It can be stated that the maximum relative error $\left|e_{I_{120}}\right|_{max}$ related to the maximum current value of SC increases with $\left|\Delta\varphi_{I,Lx,k}\right|_{max}$. Analytically $\left|e_{I_{120}}\right|_{max}$ can be approximately calculated by (19) which is the result of a curve fitting process on the basis of simulation data.

$$\left|e_{I_{120}}\right|_{max} = \sin\left|\Delta\varphi\right|_{max} + 0.04 \cdot \sin\left(4 \cdot \left|\Delta\varphi\right|_{max}\right) \quad (19)$$

Bad Data Detection and Localization

After solving the linear state estimation equations BD detection and localization algorithms have to be applied in order to check if BD exist. A common way for BD detection is the χ^2-test which determines the probability of an existing BD on the basis of the estimated state [1]. If a BD exists it can be identified by analyzing the normalized residuals (NR), the weighted difference between measured and estimated values. If a NR is greater than a specified limit the related measurement is denoted as BD and can be replaced by a pseudo measurement.

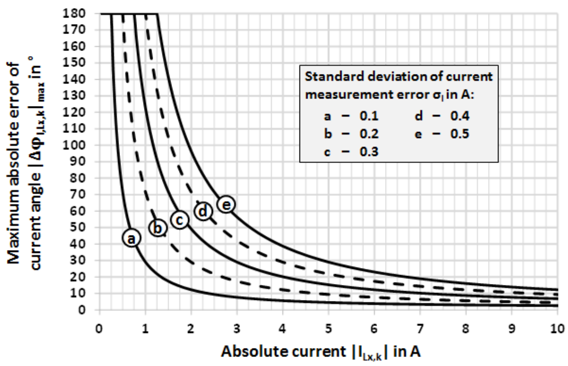

Fig. 1: Maximum absolute error of current angles in phase quantities

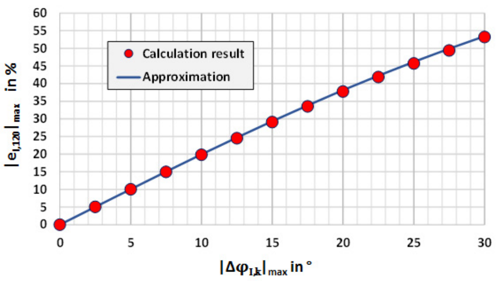

Fig. 2: Maximum relative error of absolute current values in sequence components

In the first step of the BD detection all residuals r_k are normalized to their respective variance as shown in (20). These variances are obtained by the so called residual covariance matrix $\mathbf{\Omega}$. It includes all variances of the unnormalized residuals and can be derived outgoing from the augmented matrix and the measurement variance matrix $\mathbf{R}$ by (21). The scalar Ω_{kk} which is used for the normalization of the residual r_k by (20) is the diagonal element at position kk of $\mathbf{\Omega}$. All r_k^N combined into a vector form the normalized residual vector $\mathbf{r}^N$. The elements of it are normally distributed with a mean value of zero and a variance of one as shown in (23) which enables the correct comparison of the normalized residuals against each other. By comparing the r_k^N against a specified limit ε_{120} voltage and current BD can be localized in the SC system. However, statements on the phases affected by BD cannot be taken.

$$r_k^N = \frac{r_k}{\sqrt{\Omega_{kk}}} = \frac{z_k - \hat{z}_k}{\sqrt{\Omega_{kk}}} \quad (20)$$

$$\mathbf{\Omega} = \frac{1}{\alpha} \cdot \mathbf{R} \cdot \mathbf{A}_1 \cdot \mathbf{R} \quad (21)$$

$$\begin{bmatrix} \alpha \cdot \mathbf{R} & \mathbf{H}_R & \mathbf{0} \\ \mathbf{H}_R^T & \mathbf{0} & \mathbf{C}^T \\ \mathbf{0} & \mathbf{C} & \mathbf{0} \end{bmatrix}^{-1} = \begin{bmatrix} \mathbf{A}_1 & \mathbf{A}_2^T & \mathbf{A}_3^T \\ \mathbf{A}_2 & \mathbf{A}_4 & \mathbf{A}_5^T \\ \mathbf{A}_3 & \mathbf{A}_5 & \mathbf{A}_6 \end{bmatrix} \quad (22)$$

$$\mathbf{r}^N \sim N\left(\mu = 0, \sigma_{120}^2 = 1\right) \quad (23)$$

The BD localization in the three-wire system is done for each measurement k by (24) by retransforming the complex normalized residuals $\underline{r}_{k,(1)}^N$, $\underline{r}_{k,(2)}^N$ and $\underline{r}_{k,(0)}^N$ in SC into the complex residuals $\underline{r}_{k,L1}$, $\underline{r}_{k,L2}$, $\underline{r}_{k,L3}$ in the three-wire system.

$$\begin{bmatrix} \underline{r}_{k,L1} \\ \underline{r}_{k,L2} \\ \underline{r}_{k,L3} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ \underline{a}^2 & \underline{a} & 1 \\ \underline{a} & \underline{a}^2 & 1 \end{bmatrix} \cdot \begin{bmatrix} \underline{r}_{k,(1)}^N \\ \underline{r}_{k,(2)}^N \\ \underline{r}_{k,(0)}^N \end{bmatrix} \quad (24)$$

Experience shows that it is appropriate to define a detection limit ε_{L123} in the three-wire system as $\varepsilon_{L123} = 4$ for voltage residuals, $\varepsilon_{L123} = 3$ for active current residuals and $\varepsilon_{L123} = 2$ for reactive current residuals [4]. Then residuals above the specific limit are denoted as BD and can be replaced by pseudo values. However, as each and every residual is also more or less dependent from each other which is described by the residual covariance matrix $\mathbf{\Omega}$ it is recommended to replace always only the measurement value which results in the largest residual. After that the optimization problem has to be solved once again.

SIMULATIVE SYSTEM VERIFICATION

The analysis of the BD localization process is done outgoing from synthetic measurement data sets. Various BD scenarios have been defined on the basis of smart meter data which cover a large part of the BD cases occurring in reality. Regarding this, also single and multiple BD are considered with different bad data values. The evaluation of the test data shows that voltage magnitude BD can be reliably detected by the developed SE process depending on the chosen detection limit and the ratio between bad data absolute value $|e_{BD}|$ and the applied standard deviation for the voltage measurement error σ_U. Within the project σ_U was specified to 0.2 V. Regarding this, Fig. 3. shows the detection probability functions. It indicates that a detection limit related to the normalized residuals of $\varepsilon_{L123} = 4$ is appropriate. The localization of active and reactive current BD is much more difficult as the current residuals are more or less dependent from each other. But due to the consideration of absolute current values for each and every phase the detection probabilities for active and reactive current BD are sufficient. In detail they are shown in Fig. 4 and Fig. 5. Here the standard deviation of the current measurement error σ_I was specified to 0.1 A. It can be shown that the current BD detection is not as reliable as for voltage magnitude BD. However, active current BD with $|e_{BD}|/\sigma_I$ greater 15 and reactive current BD with $|e_{BD}|/\sigma_I$ greater 20 can be detected with a probability of 95 % when $\varepsilon_{L123} = 3.5$ respectively $\varepsilon_{L123} = 2$ has been chosen. The taken assumption of $\sigma_I = 0.1$ A implies that active current BDs with values over 1.5 A and reactive current BDs with values over 2 A can probably be detected.

CONCLUSION AND OUTLOOK

The results of the developed SE algorithm tested by various simulations are very promising. They show that the future rollout of smart meters at every household can provide a way for estimating the state of LV grids. The increased measurement redundancy allows the BD analysis to work great for voltage measurements. Additionally, it works appropriate for active and reactive current BD even in cases of asymmetric system states. This is especially due to the linear consideration of absolute current measurements. Further investigations, simulations and tests will be done concerning the quality of pseudo measurements as well as an advanced SE process for systems with lower measurement redundancies. Also, additional investigations will be done concerning an improved BD detection of small multi-phase current bad data. Finally, the algorithm will be implemented on a SCADA server at the local DSO in spring 2016 and tested under real grid operation conditions.

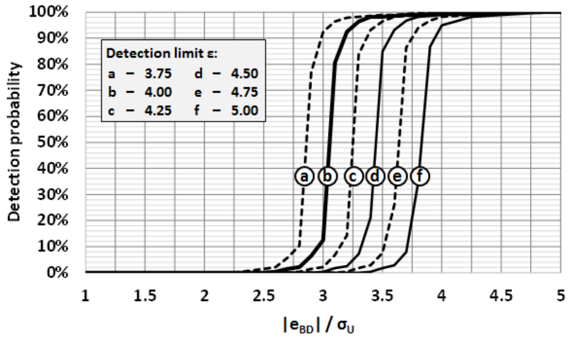

Fig. 3: Voltage bad data detection probability depending on the detection limit

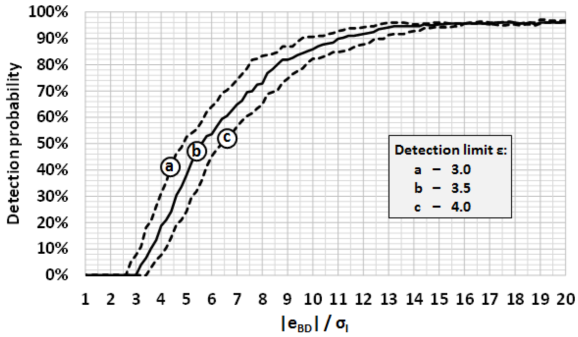

Fig. 4: Active current bad data detection probability depending on the detection limit

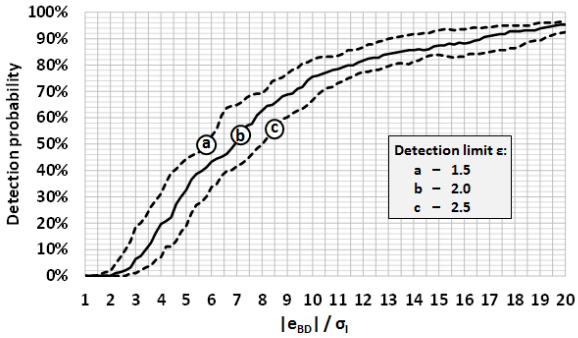

Fig. 5: Reactive current bad data detection probability depending on the detection limit

REFERENCES

[1] Abur, A.: Power System State Estimation. CRC Press, Boca Raton, USA, 2004
[2] Waeresch, D. et.al.: "State estimation in low voltage grids based on smart meter data and photovoltaic-feed-in-forecast", 23rd International Conference on Electricity Distribution, June 2015, Lyon
[3] Waeresch, D.; Brandalik, R.; Wellssow, W.H.; Jordan, J.; Bischler, R.; Schneider, N.: "Linear state estimation in low voltage grids based on smart meter data", IEEE Powertech, July 2015, Eindhoven
[4] Waeresch, D.; Brandalik, R.; Wellssow, W.H.; et.al.: "Linear state estimation in low voltage grids based on smart meter data", August 2016, July 2016, Paris