Cyber Security-Oriented Smart Grid State Estimation

Irina Kolosok[1], *Liudmila* Gurina[1*]

[1]Melentiev Energy Systems Institute, Lermontov str. 130, Irkutsk, Russia

Abstract. Development of Smart Grid involves the introduction of Wide Area Measurement System (WAMS), which provides the use of information, computing and digital technologies for measuring, transmitting and processing operating parameters when solving control problems. In this regard, the increased vulnerability to cyberattacks of the control system was noted. The control of Smart Grid includes monitoring, forecasting and planning of the system operation based on its Electric Power System state estimation results. Therefore, the goal of the paper is to develop a mathematical instrument to bad data detection under cyberattacks. Particular attention is paid to false data injection attacks which result in distortion of state variables estimates. The result of the research is an algorithm developed for state estimation based on the interior point method and test equation obtained by Crout matrix decomposition. The obtained results showed effectiveness of the algorithm in state estimation.

1 Introduction

The energy power systems (EPS) in the most advanced countries are developing towards the creation and large-scale adoption of Smart Grid which got the name of intelligent energy system in Russia [1]. An attribute of the Smart Grid is cyber-physical intrusion tolerance of the network. Developing the conceptual Smart Grid models and projects, researchers nowadays, pay great attention to the issue of cyber security. In this connection it is necessary to note the elevated vulnerability of EPS information-communication infrastructure. Thus, it becomes essential to upgrade the existing mathematical tools and develop new ones to furnish the EPS control and monitoring under cyberattacks with the data of required quality.

We consider the state estimation tool as a link between physical and information-communication infrastructures of EPS. It acts as a barrier to the corruption of data on current operating conditions of the electric power system in the control problem, including the data corruption caused by cyberattacks on data collection and processing systems of the EPS.

State estimation is a mathematical data processing method which is widely used for calculation of power system state variables on the basis of measurements.

The correct estimate of the system state can only be provided if the measurements do not contain gross errors or bad data.

The reasons for bad data are:

– Random factors related to a failure in the data collection system, personnel errors, etc.
– Cyberattacks on the SCADA system and WAMS and state estimation software.

The most vulnerable facilities in terms of cyberattack consequences for state estimation are the information-communication control subsystems (SCADA and WAMS). Since the input data for the state estimation are represented by the SCADA measurements and PMU data, cyberattacks on the SCADA and WAMS, distort the results of state estimation. If no special measures are taken to identify these distortions and suppress their impact on the state estimation results, serious errors can appear in decisions made by dispatchers using the state estimation results. Therefore, to obtain quality state estimation results, the used measurements should be tested for the presence of bad data.

The Melentiev Energy Systems Institute, SB RAS, has developed a test equation (TE) method for bad data detection and state estimation. The TE method enables us to detect bad data and systematic errors in measurements and identify their variances before state estimation [2].

Therefore, this method was chosen for the analysis of cyber security of SCADA system, WAMS and state estimation.

The paper is concerned with the issue of identification of the malicious cyberattacks in the EPS state estimation. To this end, we consider SCADA and WAMS structures, reveal vulnerable “points”, and analyze potential cyberattacks. Special attention is paid to cyberattacks such as False Data Injection (FDI) Attacks aimed at distorting the state estimation results [3].

In this connection, we propose an algorithm for detection of cyber intrusions. The algorithm is based on test equations obtained by Crout matrix decomposition of the Jacobian matrix. The SCADA data were used to implement the algorithm under simulated cyberattacks. The obtained results showed effectiveness of the

*Corresponding author: gurina@isem.irk.ru

algorithm in state estimation.

2 Cyber security of EPS control system

The research into cyber security of electric power system is focused on two subsystems - information-communication control and controlled (physical) subsystems. The information-communication control subsystem includes the systems of SCADA/EMS, WAMS, WAPS (Wide Area Protection Systems), and WACS (Wide Area Control Systems). The controlled subsystems are represented by the objects of control (electric power plants, substations, transmission and distribution networks).

The SCADA/EMS systems designed to support the actions of the operator in the operational and emergency control of EPS, include: Remote Terminal Units (RTU), installed at substations of EPS and intended to record telesignals about the status of switching equipment and measurements of operating parameters; communications channel; Master Terminal Unit (MTU), which provides Human Machine Interface (HMI) between the Engineering Work Station (EWS) and System.

Russian analog of WAMS is the System for Transient Conditions Monitoring. It includes recorders of synchronized phasor measurements (PMU data), phasor data concentrators (PDC), dispatch control at all levels (central (CDC), interregional (IDC) and regional (RDC)), channels for data transfer between the recorders, data concentrators and dispatch control centers of JSC "SO of UES"', and facilities for processing the obtained information. WAMS measurements are synchronized using global navigation satellite systems (GNSS), including GPS and GLONASS. Signals from GNSS are received by time server (TS) intended for the generation of accurate time signals and further synchronization of phasor measurement units. A hierarchical structure of EPS control system in Russia is presented in Fig. 1.

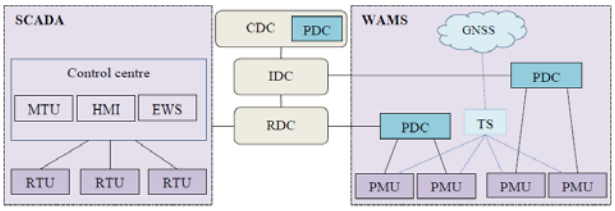

Fig. 1. A hierarchical structure of EPS control system.

The functional components of the control system are a time synchronization subsystem (TSS), a measurement subsystem (MS), a data transfer subsystem (DTS), and a data processing subsystem (DPS) [4]. Possible cyberattacks on EPS control system [5] are presented in Table 1.

Table 1. Possible cyberattacks on control system.

	MS	DTS	DPS	TSS
Reconnaissance attacks	+	+	+	-
False data injection attacks	+	+	+	-
Denial of service attacks	+	+	+	-
Spoofing attacks	-	-	-	+
Replay attacks	-	-	-	+
Jamming	-	-	-	+

We focused on the study of FDI attacks as they are negatively affect the state estimation results.

3 FDI Attacks

False data injection attacks are aimed at breaking integrity, availability and validity of data or operability of the system.

For the case of invalid data that occurred due to cyberattacks, we propose the following model [6] to describe the measurements of parameters of state $\bar{y}$:

$$\bar{y} = y + \xi_y + a\,,$$

where y is a stream of true values of measured parameters; ξ_y is vector of measurement noise that has normal distribution $\xi_y \to (0, \sigma_y^2)$ with zero mean and variance σ_y^2 that characterizes the accuracy of measurements; a is cyberattack.

We described possible FDI attacks (Fig. 2), affecting the accuracy of the state estimation results.

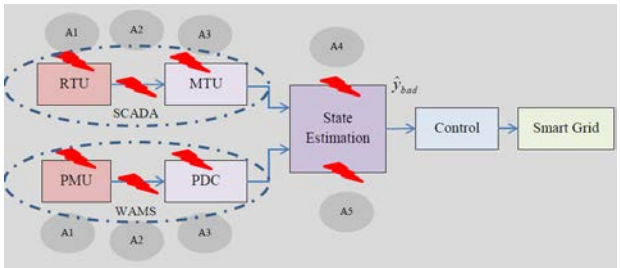

Fig. 2. FDI attacks aimed at distorting the state estimation results.

Mathematical models of measurements with cyberattacks are proposed Table 2.

Table 2. Possible FDI attacks and adequate measurement model.

FDI attacks	Model
$A1, A3$ are malicious data added to the measurements	$\bar{y} = y + \xi_y + a_{1,3}$
$A2$ is noise	$\bar{y} = y + \xi_y + \xi_{a2}$
$A4$ is errors	$\bar{y} = H(x + a_4) + \xi_y$, where H is Jacobian Matrix, $x = \{U, \delta\}$ - state vector
$A5$ is state errors	$\hat{y}_{bad} = \hat{y} + a_5$

4. State Estimation based on SCADA and WAMS data

The state estimation problem is a search for the calculated values (estimates) of measurable variables of state $\hat{y}$ that are closest to the measured values $\bar{y}$ with respect to some criterion which is most often represented by the sum of squared deviations of estimates from measurements [2]:

$$J(y) = (\bar{y} - \hat{y})^T R^{-1} (\bar{y} - \hat{y}) \quad (1)$$

and satisfies the steady state equations

$$w(y, z) = 0, \quad (2)$$

relating the measured y and unmeasured z state variables. In (1), R_y is the covariance matrix of measurement errors; its diagonal elements are equal to the variances of measurements σ^2, $\bar{y}$ is vector of SCADA and WAMS measurements, including magnitudes U_i and phases δ_i of nodal voltage, generation of active P_{gi} and reactive Q_{gi} power at nodes, and power flows in transformers and lines P_{ij} and Q_{ij}, currents at nodes and in lines I_i and I_{ij}, φ_{ij} is angles between current and voltage:

$$\bar{y} = \{P_i, Q_i, P_{ij}, Q_{ij}, U_i, \delta_i, I_i, I_{ij}, \varphi_{ij}\}.$$

Pseudo-measurements of nodal loads are used in addition to measurements of generation to obtain nodal injections P_i, Q_i.

5. Bad data detection techniques for state estimation under FDI Attacks

5.1 Formation of test equations for EPS state estimation

The test equations are steady-state equations that include only measured variables:

$$w(y) = 0. \quad (3)$$

There are different ways to obtain test equations:

1. Test equations can be obtained by excluding unmeasured variables z from the equations of electric power system steady state. System of equations (2) is nonlinear with respect to y and z, and should be linearized.

2. The vector of state $x = (U_i, \delta_i)$ is introduced to estimate the state of electric power system. This vector includes voltage magnitudes and phases. Explicit x dependences of measured and unmeasured variables are used as equations (2):

$$y = y(x), \quad (4)$$

$$z = z(x). \quad (5)$$

To obtain test equations, the components of state vector $x = (U_i, \delta_i)$ are excluded from (4) [2]. In this paper, we use the test equations obtained in this way.

Write (4) in the form

$$y - y(x) = 0. \quad (6)$$

By linearizing (6) at point $x = x_0$

$$\frac{\partial y}{\partial x}(x - x_0) = y - y(x_0),$$

divide all measurements into base y_b and redundant y_r:

$$\frac{\partial y_b}{\partial x}(x - x_0) = y_b - y_b(x_0),$$

$$\frac{\partial y_r}{\partial x}(x - x_0) = y_r - y_r(x_0).$$

Base measurements y_b are a set of measurements that allows us to uniquely determine all components of state vector x. The number of redundant measurements determines the amount of test equations.

At a linearization point we calculate the Jacobian matrix:

$$H = \frac{\partial y}{\partial x}.$$

Represent matrix H as

$$H = \begin{bmatrix} H_{11} \\ H_{21} \end{bmatrix},$$

where $H_{11} = \frac{\partial y_b}{\partial x}$, $H_{21} = \frac{\partial y_r}{\partial x}$.

By applying Crout matrix decomposition to matrix H, we obtain

$$H = \begin{bmatrix} L_{11} \\ L_{21} \end{bmatrix} U_{11},$$

where L_{11} is a lower triangular matrix whose order is equal to the number of components of vector x, L_{21} is a rectangular matrix whose number of rows is equal to the number of redundant measurements or the number of test equations, U_{11} is an upper triangular matrix.

The matrix of coefficients of the system of test equations is determined by

$$D = L_{21}L_{11}^{-1}. \quad (7)$$

According to (6), the system of test equations will have the form

$$y_r - Dy_b = 0. \quad (8)$$

Thus, (8) in a matrix form is written as follows

$$\begin{bmatrix} E & -D \end{bmatrix} \times \begin{bmatrix} y_r \\ y_b \end{bmatrix} = 0, \quad (9)$$

where $B = \begin{bmatrix} E & -D \end{bmatrix}$, $y = \begin{bmatrix} y_r \\ y_b \end{bmatrix}$, E is a square identity matrix, whose dimension is determined by the number of redundant measurements.

We propose using the Interior Point Method (IPM) based on the TE method to solve the state estimation problem under FDI attacks.

5.2 IPM based on TE method

Minimize objective function (1)
Subject to (9) and

$$y_{\min} \le y \le y_{\max}, \quad (10)$$

i.e. the estimated values of some state variables obtained during the state estimation must be within certain technological limits. Thus, generation of active and reactive power at nodes should be within the limits determined by the power generation schedule; for power flows in transformers and lines, the limits determined by line transfer capability can be assigned; at load nodes it is necessary to provide correct direction (sign) of the nodal injection, etc.

The Interior Point Algorithm based on the TE method consists of two stages:

Stage 1. Calculation of initial parameters meeting the feasibility conditions;

Stage 2. Optimization in the feasibility region consists in iterative calculation:

$$y^{(k+1)} = y^{(k)} + \lambda^{(k)}\Delta y^{(k)}, \quad (11)$$

where $\Delta y^{(k)}$ is direction of improving the solution in iteration k, $\lambda^{(k)}$ is the step value in this direction.

In stage 1, the vector $\Delta y^{(k)}$ is the solution to the auxiliary problem

$$F^{(k)} = \frac{1}{2}\sum_{j=1}^{m}\frac{(\Delta y_j^{(k)})^2}{g_j^{(k)}} \to \min, \quad (12)$$

$$\begin{bmatrix} E & -D \end{bmatrix} \times \begin{bmatrix} \Delta y_r \\ \Delta y_b \end{bmatrix} = r^{(k)}, \quad (13)$$

where $r^{(k)}$ is a residual vector in the k-th iteration. Square weighted coefficients $g_j^{(k)}$ are defined as

$g_j^{(k)} = (\min\{y_{\max j} - y_j^{(k)}, y_j^{(k)} - y_{\min j}\})^2, j = 1,...,m$.

Denote the diagonal matrices

$$G = diag(g^{(k)}).$$

The problem is solved by Lagrange multiplier method. Proceeding from the optimality conditions, we express

$$\Delta y = G^{(k)}B^T u, \quad (14)$$

where u is Lagrange multipliers vector.

Substituting (13) into (12), we obtain a system of linear equations with respect to u,

$$BG^{(k)}B^T u = r^{(k)}. \quad (15)$$

We find vector u, and use (14) to determine the direction of improving the solution $\Delta y^{(k)}$.

In stage 2, we solve the problem

$$F^{(k)} = \frac{1}{2}\sum_{j=1}^{m}\frac{(\Delta y_j^{(k)})^2}{g_j^{(k)}} + J(y^{(k)} + \Delta y) \to \min$$

$$\begin{bmatrix} E & -D \end{bmatrix} \times \begin{bmatrix} \Delta y_r \\ \Delta y_b \end{bmatrix} = 0, \quad (16)$$

where

$$(y^{(k)} + \Delta y) = J(y^{(k)}) + \frac{1}{2}\sum_{j=1}^{n}\sigma_j^{-2}(\Delta y_j)^2 + c^{(k)}\Delta y,$$

$$c^{(k)} = \nabla J(y).$$

Using the Lagrange multiplier method, we obtain

$$\Delta y = ((G^{(k)})^{-1} + R^{-1})B^T u^{(k)} + c^{(k)}. \quad (17)$$

The step value is determined by the rule

$$\lambda^{(k)} = \min\{1, \tilde{\lambda}^{(k)}\},$$

where

$$\tilde{\lambda}^{(k)} = \gamma \max\{\lambda : y_{\min} \le y^{(k)} + \lambda\Delta y^{(k)} \le y_{\max}\}, \gamma \in (0,1).$$

The iterative transition is carried out according to the rules (11).

The stopping criterion is the satisfaction of the condition:

$$\xi = \sqrt{2F^{(k)}} < \varepsilon,$$

i.e., it is assumed that the optimal solution is obtained [7].

6 Case study

To demonstrate the proposed approach, we considered the IEEE 14-bus test system (Fig. 3). Nodal voltage U_i $(i = \overline{1,14})$, injection nodes P_i, and power flows in lines P_{i-j} were given as measurements.

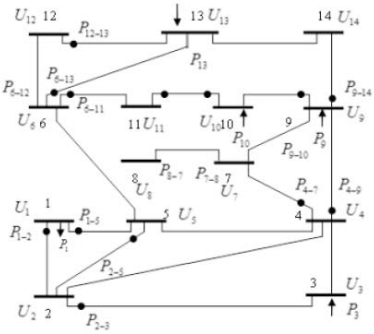

Fig. 3. IEEE 14-bus test system.

FDI attacks were modeled by adding gross errors to the measurement values $U_1, P_1, P_{1-5}, P_{1-2}$ in a range of $3\sigma - 20\sigma$. Errors were set to meet the active power balance at node 1.

The state estimation problem was solved in three ways:

1. Test equations method, implemented in software "Otsenka" [2].
2. IPM [8] for constraints (3) and inequality constraint (10).
3. IPM based on the TE method.

For case 3, initial data are the state vector $x = (U_i, \delta_i)$, measurement values, and measurement variances σ_y^2. Test equations in the form (5) are obtained by decomposing the Jacobi matrix H by the Crout matrix decomposition. The basic measurements include U_i $(i = \overline{1,14})$, P_1, P_3, P_{4-5}, P_{2-5}, P_{13}, P_9, P_{8-7}, P_{10}, P_{9-10}, P_{6-11}, P_{6-12}, P_{6-13}, P_{9-14}. The measurements P_{1-2}, P_{1-5}, P_{2-3}, P_{2-4}, P_{3-4}, P_{4-7}, P_{4-9}, P_{7-8}, P_{12-13} are redundant.

Table 3 presents the results of a comparative analysis of the values of the objective function (1) for three cases.

As evidenced by the analysis of the obtained results (Table 3), the traditional approach does not detect errors. In the second case, the IPM does not detect gross errors and distorts all measurement estimates. In the third case, we can determine at which node a cyberattack occurred by the value and the components of objective function (1).

Table 3. The results of a comparative analysis of the values of the objective function.

	Software "Otsenka"	IPM	IPM (Crout matrix decomposition)
Objective function	39,27	3484,15	382,26
Constraints	$w(y) = 0$	$w(y) = 0$	$y_r - Dy_b = 0$
Inequality constraint	$342 \le U_i \le 418$ $(i = \overline{1,14})$	$342 \le U_i \le 418$ $(i = \overline{1,14})$ $0 \le P_1 \le 280$	$342 \le U_i \le 418$ $(i = \overline{1,14})$ $0 \le P_1 \le 280$

To eliminate the consequences of these attacks, we propose duplicating the data with WAMS measurements in real time to obtain redundant measurements [9].

7 Conclusions

1. Smart grids based on the sophisticated computer and communications equipment are characterized by elevated vulnerability to different types of cyberattacks.
2. The proposed approach to state estimation based on the Crout matrix decomposition allows detecting bad data under FDI attacks on SCADA system and WAMS.
3. To improve the performance of the methods for the verification of data used in the EPS state estimation, it is necessary to increase the redundancy of SCADA measurements, supplement the SCADA measurements with the PMU measurements obtained from WAMS, combine various bad data detection methods (a priori, a posteriori, robust), and use the criterion of maximum probability of bad data detection when placing PMUs.

References

1. N. Voropai, Energy Policy, **2**, 9 (2010).
2. A. Gamm, I. Kolosok, *Bad data detection in measurements in electric power systems* (2000).
3. Yao Liu, Peng Ning, Michael K. Reiter, *CCS'09 Proc.*, 21-32 (2009).
4. Y. Ivanov, A. Cherepov, D. Dubinin, Energy of Unified Grid, **3 (26)**, 62-70 (2016).
5. I. Kolosok, L. Gurina, Intelligent Systems Research, **158**, 94-99 (2018).
6. I. Kolosok, L. Gurina, Information and mathematical technology in science and control, **1(5)**, 19-29 (2017).
7. I. Dikin, *The interior points method in linear and nonlinear programming* (2010).
8. L. Gurina, V. Zorkaltsev, I. Kolosok, E. Korkina, I. Mokry, *EPS state estimation: algorithms and examples of linearized problems* (2016).
9. I. Kolosok, L. Gurina, Lecture Notes in Computer Science, **8985**, 172-177 (2016).