Electric Power Systems Research 164 (2018) 212–219

Contents lists available at ScienceDirect

Electric Power Systems Research

journal homepage: www.elsevier.com/locate/epsr

The statistical interval estimation of the mean and the hypothesis testing of population proportions for transformer tap position estimation

M. Rezaei[a,*], S.S. Mortazavi[a], M. Razaz[a], M.S. Ghazizadeh[b]

[a] Department of Electrical Engineering, Shahid Chamran University, Ahvaz, Iran
[b] Department of Electrical Engineering, Shahid Beheshti University, Tehran, Iran

ARTICLE INFO

Keywords:
Energy management system
Hypothesis testing
Power system estimation
Transformer tap position
Parameter estimation

ABSTRACT

Frequent changes in the modern electrical power systems necessitate network state estimating for energy management system. Efficient network state estimation depends on deriving a correct model of the power system. To best model, the precise values of system parameters should be known. Therefore, many studies have paid attention to estimate unknown network parameters. Among them, transformer tap position (TTP) is one of the most important parameters ignored in the power system. This paper estimates TTP based on the statistical interval estimation of the mean and the hypothesis testing of population proportions. We exploit statistical rules jointly with weighted least squares (WLS) method to efficiently obtain unknown TTPs. Because of the non-recursive form, this method does not suffer the calculation divergence problem and is also simpler than recursive methods such as Bayesian approach. The proposed approach is also compared with Seidel type Recursive Bayesian Approach (SRBA) in the IEEE 14-bus system to show its advantages. Furthermore, it is applied on a real power system.

1. Introduction

One of the most important parts in the energy management system (EMS) is state estimation (SE) [1]. Frequent changes in the modern electrical power systems reveal the importance of SE [2]. The error in telemetered measurements and parameters of the system model is the biggest problem in the estimation [3–6]. Ref. [7] has studied the effect of the parameter error in SE. Transformer tap position (TTP) is a fundamental parameter in the system model and its quadratic form in power equations shows its crucial impact on SE [8].

Thus, different methods have been proposed for power system parameters estimation. In Ref. [9], some of the most important methods and their problems have been mentioned. Some papers have used the state vector augmentation approach [10,11]. In Ref. [10], a numerical algorithm has been utilized for observability analysis of the model and the transformer's tap has been supposed to be a state variable. The need for lots of measurements and the numerical instability are the problems of these algorithms [9]. Another group of papers has used the residual sensitivity analysis approach [12–14]. In Ref. [13], an algorithm has been supposed for the TTP estimation of generators that gets into trouble in relation to the residual contaminations.

Other groups of studies are based on the dynamic signal modeling approach [15,16]. This method needs the dynamic modeling of the system in switching transient state that is usually unavailable in TTP estimation studies [9]. Another major group is based on phasor measurement units (PMU) [17–19]. This method needs a lot of measurement units in all of the systems that limits its application in practice [17,20]. Kalman filter is a powerful method in estimation [21–23], but it needs the state space model of the system, and it is difficult to make the state space model of large and complicated power systems.

To improve the power system state estimation, studies on recursive methods became more significant [3,24]. Some papers used the Recursive Bayesian Approach (RBA), which is an efficient method in the estimation process, in various fields [25,26]. References that used RBA and its Seidel type (SRBA) in power system parameter estimation, have shown their better performance [9,25,27]. Ref. [27] has used the Bayesian approach for identification of network configuration changes in the distribution system. In this method, different models of the system are stored in a model bank based on the power circuit breakers status and changes in system injections due to connection of distributed generators (DG) and major loads. The model bank includes all of the feasible mathematical models for the system.

Firstly, the system estimated states are calculated by weighted least squares (WLS) iterative method for all different models of the system. Then, estimated measurements are gained from these estimated states and differences between estimated measurements and system

* Corresponding author.
E-mail address: m-rezaei@phdstu.scu.ac.ir (M. Rezaei).

https://doi.org/10.1016/j.epsr.2018.08.007
Received 14 May 2018; Received in revised form 23 July 2018; Accepted 10 August 2018

measurements make the residual vectors. Finally, by using the residuals in the recursive Bayesian formula, probabilities of all models are calculated and the correct tap is identified.

The problem of this method is the recursive form of RBA. The speed of the convergence and more important than that, the feasibility of the convergence are the significant challenges in the recursive methods. Refs. [4,25] improved the proposed method of Ref. [27] by using the Seidel type of the Recursive Bayesian Approach. Results of this paper show the better performance of this method. Ref. [9] uses SRBA for the TTP estimation. This paper utilizes some methods to improve the estimation operation such as the largest normalized residual (LNR) test for identifying bad data and preserving its effect on calculations convergence, identifying suspected transformers with incorrect tap ratios, fast calculating residuals of SE by prediction–correction method, proposing a criterion to reduce iterations of the recursive approach and effecting position set (EPS) reduction.

Identifying and separating the measurements to form the residual vectors is the first problem in Bayesian method. This will be discussed in the next parts of this study. All the above-mentioned methods are solutions to prevent SRBA divergence and reduce the calculations speed because of its iterative form, especially regarding real time applications. Actually, like other methods which apply residuals, SRBA is very sensitive to the residual vectors. All the above-mentioned methods have a series of calculations some of which need iterative solution that increase the calculations load and make the estimation slow [9]. Furthermore, calculations convergence is not guaranteed by using these methods. These problems are more severe in complicated networks with parallel and three-winding transformers having common nodes.

One of the most important above-mentioned solutions is the LNR test which will be explained in the next sections briefly and has been explained in detail in Refs. [9] and [28]. Considering normalized residual vectors, The LNR test identifies bad data and replaces them with their pseudo values. Convergence of SRBA by using LNR is not assured, and it is dependent on the type and form of the bad data [28].

The innovation of this study is to identify TTP by using the statistical interval estimation of the mean and the hypothesis testing [6,29,30] of population proportions. The method's main features are summarized as follows.

- We use local network model rather than the entire network model for TTP estimation. Hence, getting the system more complex and larger does not affect the method's performance significantly.
- We use statistical interval estimation of the mean to get the range of the noisy measurement means of the power system. It is a simple method with low computational complexity.
- We also define a correctness index in order to compare all tap positions. The estimated measurements are produced from the estimated states obtained by WLS method. Through this index, we check whether each estimated measurement is in its calculated mean range. All correct responses make the correctness index for each tap. The sequence and the type of measurements is not an important issue in this step.
- We utilize hypothesis testing of population proportions to compare the correctness indexes for identifying the correct tap position. It is a test like statistical interval estimation of the mean. The proposed method is simpler than recursive methods with less calculation steps, so it shows better performance in real time environments.
- The proposed method does not need prior information about the previous transformers' tap positions. It does not also suffer from the problem of the calculation's convergence because of the non-recursive form.
- The simulation results of the new method are compared with SRBA and it is also implemented on a real power system.

In the next section of the present study, RBA and SRBA with their problems are explained briefly. Then the statistical interval estimation of mean and the hypothesis testing of population proportions are presented. The next part of the study is about applying the statistical rule in the proposed method. The significant advantages of SRBA over the others are the better calculation speed and lower convergence limitations [9]. Holding the strengths of SRBA, the presented method in this study shows more calculation speed and high performance even in the presence of gross error as demonstrated in the simulation results. The rest of this paper compares the method with the SRBA in the IEEE 14-bus system. Finally, the presented method is applied on a part of real power system in Iran. The power system simulations are performed by the DigSILENT software and the mathematical calculations are coded in MATLAB software. The results of the codes show the effectiveness of the proposed method in large number of tap positions and complicated states, such as parallel and three-winding transformers, even without any knowledge of the previous values of TTPs.

2. Recursive Bayesian Approach for transformer tap position estimation

In this section, identifying TTP by using RBA and SRBA is explained. These methods and their mathematical relations have been discussed in Refs. [9,25,27]. The relationship between the state variables and measurements is:

$$z = h(x) + \tau \tag{1}$$

where $z = [z_1 z_2 ... z_m]^T$ is the measurement vector, $x = [x_1 x_2 ... x_n]^T$ is the state vector, $h(x) = [h_1(x) h_1(x) ... h_m(x)]^T$ is the vector of nonlinear and known functions, $\tau \sim N(0, R)$ is zero mean Gaussian noise with covariance matrix $= diag\{\sigma_{z1}^2, \sigma_{z2}^2, ..., \sigma_{zm}^2\}$, and σ_{zi}^2 is the variance of ith measurement. The state vector can be calculated by use of the Newton method with iterated formula as:

$$x^{k+1} = x^k + [G(x^k)]^{-1} H^T(x^k) R^{-1}(z - h(x^k)) \tag{2}$$

k is the iteration number, x^k is the calculated states in ith iteration, $G(x^k) = H^T(x^k) R^{-1} H(x^k)$ is the gain matrix, and $H(x) = \frac{\partial h(x)}{\partial x}$ is the Jacobian matrix. Defining $\hat{x} = lim\, x^k$, the residual vector of estimated measurements is derived by:

$$e = z - h(\hat{x}) \tag{3}$$

To mitigate the effect of bad data in Bayesian method, they should be identified by the LNR test. Then they are replaced with virtual data. Firstly, the normalized error vector is calculated:

$$e_i^N = \frac{|e_i|}{\sqrt{R_{ii} S_{ii}}} \tag{4}$$

where $S = I - H(H^T R^{-1} H)^{-1} H^T R^{-1}$ is the sensitivity matrix, S_{ii} is the ith diagonal element of it, and R_{ii} is the ith diagonal element of R.

If the largest normalized residual is larger than a threshold, it is known as a bad data and should be replaced by its pseudo data [28]. Because of the error of measurements, they are produced in 100 snapshots, and the above-mentioned operation along with LNR test should be performed for all tap models [9,27]. The first problem is the deceleration in operation because of the matrix S calculation and bad data removal. The second problem is inappropriate response of the LNR test in multiple bad data situations [28].

The z vector includes three categories of measurements [9]. The first group is measurements obtained from power system measurement devices, called real measurements, which demonstrate minor errors. They are produced by adding a zero-mean normally distributed random noise to the values obtained from the load flow equations. This noise has 1% error within the real value. The second category is pseudo-measurements with 10% error value. There are no measuring devices in the system for this group and they are calculated by observers (if the system is observable). The third category, called virtual measurements, refers to zero injection buses with a very low error value (10^{-6}).

Due to minor errors, only real measurements are used in the residual