Electrical Power and Energy Systems 83 (2016) 241–250

Contents lists available at ScienceDirect

Electrical Power and Energy Systems

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

Distributed algorithms for convexified bad data and topology error detection and identification problems

Yang Weng [a,*], Marija D. Ilić [b], Qiao Li [c], Rohit Negi [b]

[a] Department of Civil and Environmental Engineering, Stanford University, Stanford, CA 94305, USA
[b] Department of Electrical and Computer Engineering, Carnegie Mellon University, Pittsburgh, PA 15213, USA
[c] Qualcomm Research New Jersey, NJ 08807, USA

ARTICLE INFO

Article history:
Received 23 February 2015
Received in revised form 12 February 2016
Accepted 29 March 2016
Available online 22 April 2016

Keywords:
Power systems
Bad data and topology error
Detection and identification
Distributed algorithms
Convexification
Semidefinite programming

ABSTRACT

Large-scale smart grids call for online algorithms that are able to achieve the most accurate estimates. This paper shows how to achieve both the scalability and near globally optimal results for bad data and topology error detection and identification problems, by conducting fully distributed algorithms over convexified problem formulations. The proposed distributed decomposition is realized by (1) reducing a large network into much smaller network "cliques" which do not need extensive information exchange; (2) performing a Lagrangian dual decomposition in each clique and passing messages between cliques; and (3) conducting alternative coordinate descent optimization for robustness. To reduce the relaxation error in the convexification procedure, a nuclear norm penalty is added to approximate original problems. Finally, we propose a new metric to evaluate detection and identification results, which enables a system operator to characterize confidence for further system operations. We show that the proposed algorithms can be realized on IEEE test systems with improved accuracy in a short time.

Introduction

As one of the most significant infrastructures in human society, the electric power grid not only provides electricity as a form of flexible, convenient energy for industrial and individual uses, but also provides it in a clean and relatively easy way to transmit. Despite 130-plus years of development and engineering, electric power systems are still under intense pressure to achieve stability against outages and blackouts as the power grid moves into large scale network with tons of new devices and loads. While having a potential to reduce the impact on the environment, increase fuel diversity, and bring about economic benefits, new grid components also raise tremendous concerns regarding the secure and reliable operation of the backbone EHV/HV power grids.

Speaking more generally, the power system operators of traditional power grids face significant challenges in managing the effects of small scale generations and loads, which include but not limited to renewable energy generators, such as wind/solar generators; responsive small electricity users; and electricity users which can offer storage to utility, such as electric cars. Their topology and data quality need to be estimated to account for their effects on the state of the backbone power grid. Therefore, highly accurate algorithms capable of dealing with increasing number of bad data and constantly varying network topology are needed.

However, as most bad data and topology identification algorithms currently used by the electric power industry is computationally complex, they are typically used for Extra High Voltage (EHV), High Voltage (HV) and, at times, for Medium Voltage (MV) representation of the complex multi-voltage level power grids. This, in turn, makes it very hard to detect and identify bad data and topology error within many new diverse resources and users connected to the Low Voltage (LV) level distribution systems. A multi-layered, distributed implementation of bad data and topology error detection and identification for future electric energy systems is likely to become the preferred approach. This requires a systematic design of distributed algorithms whose performance does not worsen relative to the centralized methods.

In this paper, we first extend the convexification idea in state estimation [1] to bad data and topology error detection (L_2 norm) and identification (L_1 norm) with substation modeling. The goal is to improve accuracy with a near global optimum estimate by embedding states into proper high dimensional space. This is because there

* Corresponding author.
E-mail addresses: yangweng@stanford.edu (Y. Weng), milic@ece.cmu.edu (M.D. Ilić), liqiao02@gmail.com (Q. Li), negi@ece.cmu.edu (R. Negi).

http://dx.doi.org/10.1016/j.ijepes.2016.03.044
0142-0615/

is a strong relationship between SDP formulation for state estimation and the SDP formulation for optimal power flow, e.g., [2,3] show that the rank-one condition is satisfied with most power grid topologies. However, as a centralized implementation of the semidefinite programming (SDP)-based approach is computationally prohibitive [1], we propose to implement a distributed algorithm based on the underlying power system graph for scalability. The objective decomposition is achieved by using the Lagrangian dual decomposition method [4]. To decompose the system-level positive-semidefinite constraint over the state matrix, the notion of induced chordal graphs is used. Although this is similar to the distributed implementation of the SDP-based AC optimal power flow method [5–7], we do not set limits on states, input or output variables due to qualitatively different performance objectives between our problems and AC optimal power flow problem. Then, the proof of equivalence between the centralized algorithm and the distributed algorithm is detailed next. Different than past works, this paper proposes how to reduce approximation error in convex-relaxation by penalizing the objective with a nuclear norm instead of a pure rank-one constraint relaxation.

Therefore, by using message exchange on coupling nodes between neighboring local networks, the original centralized SDP detection and identification problems can be characterized in a distributed manner by performing local SDP computation. Obviously, a direct consequence is computational time reduction if parallel processors are used. A fully distributed implementation can also be achieved by using coordinate descent optimization for robustness [8].

In comparison with existing methods, the novelty of the proposed approach lies in (1) past WLS-based bad data and topology error detection and identification have local optimum issues. In this paper, we not only show how to convexify such problems but also show how to conduct them in a distributed way. (2) We show that the distributed computation can be generalized. This means that, in addition to L_2 norm, L_1 norm can also be used for error identification. (3) We derive lower bounds to characterize the performance of the arbitrary detection and identification methods. (4) [9] shows results about distributed state estimation based on boundary division, but our results are about distributed bad data and topology detection and identification based on the network structure decomposition, e.g., chordal graphs decomposition. [9] relaxes the rank-one constraint, but we add a nuclear norm penalty to approximate such a constraint.

If implemented, such algorithms are likely to form the basis for "smart" grids by enabling even many small system users to participate in enhancing system operation in predictable ways. For example, the topology (i.e. circuit breaker status) and data quality (i.e. power injection measurements) of smart meters would not have to be estimated by the operator of the backbone system. Instead, topology error and bad data of small users get estimated in a distributed manner by message-passing with neighboring system users which have smart meters. The aggregated information is then communicated in a bottom-up way to the backbone system operator. Highly accurate distributed detection and identification as well as greatly reduced computational time are illustrated through simulations.

This paper is organized as follows: In Section "Review", we review detection and identification methods for bad data and topology error; then a convex-relaxation-based approach is proposed to resolve the local optimum issues; in Section "Distributed implementation for bad data and topology error detection and identification", we propose a distributed SDP-based algorithm; in Section "Simulation results", we describe the simulation results for IEEE test systems; finally, we conclude this paper in Section "Conclusions"; an Appendix on using power flow measurements in the distributed algorithm is also included.

Review

Bad data and topology error detection

As both bad data and topology error can have a dramatic influence on measurements, they usually cause large residuals in weighted least square minimization (1). Therefore, the resulting objective value is used in a chi-squared test to detect them.

$$\min_{\boldsymbol{v}} J_2(\boldsymbol{v}) = \sum_{i=1}^{m}\left|\frac{z_i - h_i(\boldsymbol{v})}{\sigma_i}\right|^2, \tag{1}$$

where vector $\boldsymbol{v} = (|v_1|e^{j\delta_1}, |v_2|e^{j\delta_2}, \ldots, |v_n|e^{j\delta_n})^T$ represents the system states to be estimated for error detection. $h_i(\cdot)$ relates the unknown state variable $\boldsymbol{v}$ to the ith noiseless measurement. z_i is the ith telemetered measurement, such as power flow and voltage magnitude. σ_i is the standard deviation of Gaussian noise u_i in z_i, where noises in different measurements are assumed to be independent. Finally, m represents measurement number.

Next, one looks into a chi-squared distribution table to find a threshold, i.e., $J_{(m-n),95\%}$, corresponding to a confidence probability 95% and $m-n$ degrees of freedom. If $J_2(\boldsymbol{v}) > J_{(m-n),95\%}$, a bad data or topology error is declared [10,11]. The bus associated with the largest measurement residual is called a suspicious bus. Deciding whether a test failure indicates a sensor error or a topology error is out of the scope for this paper. Interested readers are referred to the Section 8.6.1 of [12,10,11].

Bad data and topology error identification

For bad data identification, one approach is to assume that the measurement associated with the largest residual in (1) to be bad data. So one iterates a chi-squared test with a measurement removal until the test is passed. Another approach is to use Weighted Least Absolute Value (WLAV), or L_1 norm minimization in (2) to find bad data (sparse noises) [13] altogether.

$$\min_{\boldsymbol{v}} J_1(\boldsymbol{v}) = \sum_{i=1}^{m}\left|\frac{z_i - h_i(\boldsymbol{v})}{\sigma_i}\right|. \tag{2}$$

For topology identification, one may adjust topology connection associated with large residuals found in (1), but extending suspicious bus into a sub-station model may be a better idea. This approach starts by changing measurement model $z_i = h_i(\boldsymbol{v}) + u_i$ into (3) to explicitly account substation circuit breaker status (Chapter 8 of [12]).

$$z_i = h_i(\boldsymbol{v}) + M_i\boldsymbol{f} + u_i, \tag{3}$$

where the state $\boldsymbol{v}$ is extended by adding the line power flow vector $\boldsymbol{f}$ through circuit breakers in the suspicious substation. In Eq. (3), the term $M_i\boldsymbol{f}$ represents the effect of these flows in each measurement z_i; M_i is an incident matrix defining the interconnection of suspected circuit breakers [12].

To decide the digital status of a circuit breaker, one needs to estimate $\boldsymbol{f}$. One can choose $p = 2$ (for detection) or $p = 1$ (for identification) in (4) for this purpose, where the parameter p ($p \geqslant 0$) is chosen to achieve the desired performance. The goal is to obtain a joint extended state $(\hat{\boldsymbol{v}}, \hat{\boldsymbol{f}})$ that best fits the measurement set $\boldsymbol{z}$ according to the measurement model in (3).

$$\min_{\boldsymbol{v},\boldsymbol{f}} J_p(\boldsymbol{v}, \boldsymbol{f}) = \sum_{i=1}^{m}\left|\frac{z_i - h_i(\boldsymbol{v}) - M_i\boldsymbol{f}}{\sigma_i}\right|^p. \tag{4}$$

Subsequently, $\hat{\boldsymbol{f}}$ is used to estimate digital status of bus breakers for topology identification.