Distributed Optimal Load Shedding for Disaster Recovery in Smart Electric Power Grids: A Second-Order Approach

Jia Liu[†] Cathy H. Xia[‡] Ness B. Shroff[†*] Hanif D. Sherali[*]
[†]Dept. of Electrical and Computer Engineering, The Ohio State University
[‡]Dept. of Integrated Systems Engineering, The Ohio State University
[*]Dept. of Industrial and Systems Engineering, Virginia Tech
[†]{liu.1736, xia.52, shroff.11}@osu.edu, [*]hanifs@vt.edu

ABSTRACT

In this paper, we consider the problem of distributed load shedding optimization for disaster recovery in smart grids. We develop distributed second-order interior-point based load shedding algorithms that enjoy a fast quadratic convergence rate. Our main contributions are two-fold: (i) We propose a rooted spanning tree based reformulation that enables our distributed algorithm design; (ii) Based on the spanning tree reformulation, we design distributed computation schemes for our proposed second-order interior-point based load shedding. Collectively, these results serve as an important first step in load shedding and disaster recovery that uses second-order distributed techniques.

1. PROBLEM FORMULATION

Consider a power (sub)network that remains connected upon removal of the failed portion of the system after a disaster strikes. However, we do *not* assume the availability of *any* centralized controller in this sub-network. We represent this subnetwork by a directed graph $\mathcal{G} = \{\mathcal{N}, \mathcal{L}\}$, where $\mathcal{N}$ and $\mathcal{L}$ are the sets of buses and transmission lines, with number of elements $|\mathcal{N}| = N$ and $|\mathcal{L}| = L$, respectively. We use an incidence matrix $\mathbf{A} \in \mathbb{R}^{N\times L}$ to represent the network topology, where the entry $(\mathbf{A})_{nl} = 1$ if edge l is coming out of node n, -1 if edge l is going into node n, and 0 otherwise [1]. We let b_l denote the admittance of line l and let $\mathbf{B} \triangleq \mathrm{Diag}\{b_1,\ldots,b_L\}$ be a diagonal matrix that contains all admittances. Suppose that there are G generator buses in this subnetwork and they are labeled $1,\ldots,G$. The remaining $N-G$ buses are load buses and are labeled $G+1,\ldots,N$. We let $\mathcal{K}$ denote the set of generators, with $|\mathcal{K}| = G$. We use $z_n \geq 0$, $n \notin \mathcal{K}$, to represent the amount of potential power shed at a load bus. At each generator bus $n \in \mathcal{K}$, we let $z_n \geq 0$ denote the power generation to meet the demands. We use $\mathbf{z} \triangleq \left[z_1,\ldots,z_N\right]^\top$ to group all z-variables. Following conventions, we let $p_n < 0$ denote the power demand at a load bus $n \notin \mathcal{K}$. For convenience, we let $\mathbf{p} \triangleq [0,\ldots,0,p_{G+1},\ldots,p_N]^\top$ denote the expanded power demand vector of all buses. We let $f_l \in \mathbb{R}$ represent the power flow on line l, where f_l could be negative in the sense that f_l's direction is against the direction of line l (recall that $\mathcal{G}$ is a directed graph). We use $\mathbf{f} \triangleq [f_1,\ldots,f_L]^\top$ to group all line flow variables. We let $\boldsymbol{\theta} \triangleq [\theta_1,\ldots,\theta_N]^\top$ group the phase angle variables of all buses. Then, the load shedding problem under the DC power flow model can be formulated as follows:

$$\min_{\mathbf{z},\boldsymbol{\theta}} \left\{ \sum_{n\notin\mathcal{K}} C_n(z_n) \,\middle|\, \begin{array}{l} \mathbf{f} = \mathbf{B}\mathbf{A}^\top\boldsymbol{\theta},\ \mathbf{A}\mathbf{f}-\mathbf{z}=\mathbf{p}, \\ |f_l| \leq f_{\max}, \forall l,\ 0 \leq z_n \leq K_n, \forall n. \end{array} \right\}. \quad (1)$$

In (1), $C_n(z_n)$ is a convex and twice-differentiable function that evaluates the cost incurred by shedding z_n units of load (a standard cost model in power systems analysis [2]); $\mathbf{f} = \mathbf{B}\mathbf{A}^\top\boldsymbol{\theta}$ corresponds to the DC power flow model (assuming fixed bus voltages and thus $\boldsymbol{\theta}$ can fully describe power constraints [2]); $\mathbf{A}\mathbf{f}-\mathbf{z}=\mathbf{p}$ represents the power flow conservation law (i.e., incoming plus generated power equals outgoing and consumed power at all buses); and $|f_l| \leq f_{\max}$ corresponds to line flow security constraints. In (1), the constraint $0 \leq z_n \leq K_n$ models the fact that we cannot shed more load than the actual demand and cannot produce more power than the limit of each generator bus, where $K_n \triangleq -p_n$ for $n \notin \mathcal{K}$ and $K_n \triangleq B_n^{\max}$ for $n \in \mathcal{K}$, and where $B_n^{\max} > 0$ denotes the capacity of generator n.

2. A DISTRIBUTED INTERIOR-POINT SECOND-ORDER ALGORITHM

In this paper, we solve Problem (1) by developing an *interior-point second-order distributed* approach. The fundamental rationale behind our approach is that the operating point of the system enforced by the interior-point framework is guaranteed to be feasible throughout *all* iterations of the algorithm. In contrast, most existing first-order gradient based schemes in the literature (see, e.g., [3] and references therein) do not have such feasibility guarantee and would produce iterates that constantly violate the feasibility constraints. This is particular problematic for ensuring the grid's security if the algorithm needs to be terminated in a predefined finite number iterations. In what follows, we outline the key design steps and refer readers to our online technical report [4] for further details.

Step 1) Rooted Spanning Tree-Based Reformulation: Suppose that we have a *rooted spanning tree* (RST) rooted at a generator bus. Without loss of generality, we let this generator bus be labeled "0". We attach an artificial root arc l_0 to the root with arbitrary admittance b_0 and an artificial line

SIGMETRICS'14, June 16–20, 2014, Austin, Texas, USA.
ACM 978-1-4503-2789-3/14/06.
http://dx.doi.org/10.1145/2591971.2592036 .

flow f_0.Let $\mathcal{T} = \{\mathcal{N}, \mathcal{L}_0'\}$ denote this RST, where the link set $\mathcal{L}_0' \triangleq \mathcal{L}' \cup \{l_0\}$ with $\mathcal{L}' \subseteq \mathcal{L}$ and $|\mathcal{L}'| = N - 1$. Let $\widetilde{\mathbf{A}}$ be the incidence matrix of $\mathcal{T}$. Let vector $\widetilde{\mathbf{f}} \triangleq [f_l : l \in \mathcal{L}_0']^\top \in \mathbb{R}^N$ group all the line flows in $\mathcal{T}$. Let $\widetilde{\mathbf{B}} \triangleq \mathrm{Diag}\{b_l : l \in \mathcal{L}_0'\} \in \mathbb{R}^{N \times N}$ be the diagonal line admittance matrix with respect to $\mathcal{T}$. Then, it can be shown that Problem (1) can be equivalently reformulated based on the RST as follows [4]:

$$\min_{\mathbf{z},\boldsymbol{\theta}} \left\{ \sum_{n \notin \mathcal{K}} C_n(z_n) \middle| \begin{array}{l} (\widetilde{\mathbf{A}}\widetilde{\mathbf{B}}\widetilde{\mathbf{A}}^\top)\boldsymbol{\theta} - \mathbf{z} = \mathbf{p},\ \mathbf{1}^\top(\mathbf{p}+\mathbf{z}) = 0, \\ |\theta_{\mathrm{Tx}(l)} - \theta_{\mathrm{Rx}(l)}| \le \beta_l,\ 0 \le z_n \le K_n. \end{array} \right\}, \quad (2)$$

where $\beta_l \triangleq f_l^{\max}/b_l$, $\mathbf{1} \in \mathbb{R}^N$ is an all-one vector, and $\mathrm{Tx}(l)$ and $\mathrm{Rx}(l)$ represent the transmitting and receiving ends of line l, respectively. Note that the generation-load balance constraint $\mathbf{1}^\top(\mathbf{p}+\mathbf{z}) = 0$ ensures that $f_0 \equiv 0$, implying that Problem (1) is equivalent to Problem (2) [4].

Step 2) An Interior-Point Algorithmic Framework Coupled with Newton's Method: Utilizing an interior-point approach [5], we can rewrite Problem (2) as follows:

$$\begin{array}{rl} \text{Minimize} & g_\mu(\mathbf{y}) \\ \text{subject to} & \mathbf{M}\mathbf{y} = \mathbf{d}, \end{array} \quad (3)$$

where $\mathbf{y} \triangleq \left[z_1 \cdots z_N, \theta_1 \cdots \theta_N\right]^\top$ groups all variables; $g_\mu(\mathbf{y}) \triangleq \mu \sum_{n \notin \mathcal{K}} C_n(z_n) - \sum_{n=1}^N (\log(z_n) + \log(K_n - z_n)) - \sum_{l=1}^L (\log(\beta_l - \theta_{\mathrm{Tx}(l)} + \theta_{\mathrm{Rx}(l)}) + \log(\beta_l + \theta_{\mathrm{Tx}(l)} - \theta_{\mathrm{Rx}(l)}))$ is an augmented objective function, where we apply a logarithmic barrier function to all inequality constraints in Problem (2) and then accommodate them into the objective function; and matrix $\mathbf{M}$ and vector $\mathbf{d}$ groups all network topology and demand information, respectively (see [4] for details). In (3), $\mu > 0$ is the barrier parameter such that as μ gets large, the solution of Problem (3) will approach the original solution of Problem (2). The equality-constrained formulation in (3) allows us to employ Newton's method to solve the Karash-Kuhn-Tucker (KKT) system of (3). Starting from some initial feasible solution $\mathbf{y}^0$, the (centralized) Newton's method is the following iterative scheme: $\mathbf{y}^{k+1} = \mathbf{y}^k + s^k \Delta\mathbf{y}^k$, where $s^k > 0$ is the step-size in the kth iteration and $\Delta\mathbf{y}^k$ denotes the primal Newton direction, which can be computed as:

$$\Delta\mathbf{y}^k = -\mathbf{H}_k^{-1}(\nabla g_\mu(\mathbf{y}^k) + \mathbf{M}^T\mathbf{w}^k), \quad (4)$$

$$\mathbf{w}^k = (\mathbf{M}\mathbf{H}_k^{-1}\mathbf{M}^T)^{-1}(-\mathbf{M}\mathbf{H}_k^{-1}\nabla g_\mu(\mathbf{y}^k)). \quad (5)$$

where $\nabla g_\mu(\mathbf{y}^k)$ and $\mathbf{H}_k$ denote the gradient vector and Hessian matrix of $g_\mu(\cdot)$ evaluated at $\mathbf{y}^k$, respectively; $\mathbf{w}^k$ contains dual variables associated with $\mathbf{M}\mathbf{y} = \mathbf{d}$ in the kth iteration. Note that the scheme in (4)–(5) is a *second-order* algorithmic framework since it exploits not only the first-order gradient information $\nabla g_\mu(\mathbf{y}^k)$, but also the second-order Hessian information $\mathbf{H}_k$. Hence, it enjoys a powerful quadratic rate of convergence performance [6]. That is, given a desired control precision ϵ, the distributed algorithmic framework outlined in (4)–(5) converges in at most $\log_2 \log_2(\Phi/\epsilon)$ iterations, where Φ is some constant that depends on the network topology and the objective function.

Step 3) RST-Based Decentralization: Lastly, the decentralization of the scheme in (4)–(5) boils down to the distributed computations of $\mathbf{H}_k^{-1}$ and $(\mathbf{M}\mathbf{H}_k^{-1}\mathbf{M}^\top)^{-1}$, which are challenging due to their *non-separable* structure. Fortunately, by exploiting the nice *upper triangular* structure of the RST incidence matrix $\widetilde{\mathbf{A}}$, we are able to develop a *double Sherman-Morrison-Woodbury* (dSWM) distributed scheme

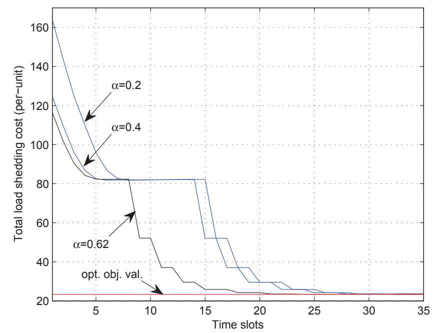

Figure 1: The convergence process of the proposed distributed second-order load shedding algorithm with different choices of initial scaling factors.

for $\mathbf{H}_k^{-1}$ and $(\mathbf{M}\mathbf{H}_k^{-1}\mathbf{M}^\top)^{-1}$, which leads to the following result on distributed computations [4]:

THEOREM 1. *The second-order scheme in (4)–(5) can be decentralized by using the RST structure: the Newton direction of load shed variables z_n can be distributedly computed at bus n with local information; the Newton direction of phase angle variables θ_n can be distributedly computed following any link-ordering on the RST in exactly N steps.*

To identify a starting point, we let $\widehat{\boldsymbol{\theta}}$ be the phase angles before load shedding. Our proposed initialization is to shrink all $\widehat{\theta}$-variables by the same scaling factor $0 < \alpha < 1$, i.e., $\mathbf{f} = \mathbf{B}\mathbf{A}^T(\alpha\widehat{\boldsymbol{\theta}})$. Clearly, there always exists a small enough α such that $\mathbf{f}$ will not exceed any line capacity (i.e., in the interior of the new feasible domain).

To illustrate the convergence speed of our algorithm, we use the IEEE 30-bus benchmark system [7] as an example. For this 30-bus benchmark system, the convergence processes of our proposed distributed second-order load shedding algorithm with different choices of initial scaling factors α are illustrated in Figure 1. We can see that, in all cases, our distributed second-order method converges in less than 35 iterations, which demonstrates the powerful second-order convergence speed.

3. REFERENCES

[1] M. S. Bazaraa, J. J. Jarvis, and H. D. Sherali, *Linear Programming and Network Flows*, 4th ed. New York: John Wiley & Sons Inc., 2010.

[2] J. J. Grainger and W. D. Stevenson, *Power System Analysis*. New York: McGraw-Hill Inc., 1994.

[3] D. Bienstock, "Optimal control of cascading power grid failures," in *Proc. IEEE CDC*, Orlando, FL, Dec. 12-15, 2011, pp. 2166–2173.

[4] J. Liu, C. H. Xia, N. B. Shroff, and H. D. Sherali, "Distributed optimal load shedding for disaster recovery in smart electric power grids: A second-order approach," *Technical Report, Dept. of ECE, The Ohio State University*, 2014. [Online]. Available: http://www2.ece.ohio-state.edu/~liu/publications/

[5] Y. Nesterov and A. Nemirovskii, *Interior-Point Polynomial Algorithms in Convex Programming*, 3rd ed. Philadelphia, PA: SIAM, 2001.

[6] M. S. Bazaraa, H. D. Sherali, and C. M. Shetty, *Nonlinear Programming: Theory and Algorithms*, 3rd ed. New York, NY: John Wiley & Sons Inc., 2006.

[7] "Power systems test case archive." [Online]. Available: http://www.ee.washington.edu/research/pstca/