Electrical Power and Energy Systems 67 (2015) 582–590

Contents lists available at ScienceDirect

Electrical Power and Energy Systems

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

Decentralized control and fair load-shedding compensations to prevent cascading failures in a smart grid

Benyun Shi [a,b], Jiming Liu [b,*]

[a] School of Information Engineering, Nanjing University of Finance and Economics, Nanjing, China
[b] Department of Computer Science, Hong Kong Baptist University, Kowloon Tong, Hong Kong

ARTICLE INFO

Article history:
Received 8 January 2014
Received in revised form 7 December 2014
Accepted 9 December 2014
Available online 27 December 2014

Keywords:
Cascading failures
Decentralized control algorithm
Load-shedding compensations
Embedded feedback mechanism
The proportional fairness criterion

ABSTRACT

Evidence shows that a small number of line contingencies in power systems may cause a large-scale blackout due to the effects of cascading failures. With the development of new technologies and the growing number of heterogeneous participants, a modern/smart grid should be able to self-heal its internal disturbances by continually performing self-assessment to deter, detect, respond to and restore from unpredictable contingencies. Along this line, this research focuses on the problem of how to prevent the occurrence of cascading failures through load shedding by considering heterogeneous shedding costs of grid participants. A fair load-shedding algorithm is proposed to solve the problem in a decentralized manner, where a load-shedding participant need only monitor its own operational status and interact with its neighboring participants. Using an embedded feedback mechanism, the fair load-shedding algorithm can determine a marginal compensation price for each load-shedding participant in real time based on the proportional fairness criterion, without knowing the shedding costs of the participants. Such fairly determined compensations can help motivate loaders/generators to actively participate in the load shedding in the face of internal disturbances. Finally, the properties of the load-shedding algorithm are evaluated by carrying out an experimental study on the standard IEEE 30 bus system. The study will offer new insights into emergency planning and design improvement of self-healing smart grids.

Introduction

Historical data shows that power systems have suffered from series of internal and external disturbances leading to various degrees of blackouts due to the effect of cascading failures [13,25,11]. An unpredictable blackout may severely affect activities reliant on electricity, such as railway and air transportation, water supply and hospital services. For example, during the U.S.-Canada blackout of August 14, 2003, over 400 transmission lines and 531 generating units tripped and approximately 50 million people were affected [31]. According to the modern grid initiative conducted by the National Energy Technology Laboratory of the U.S. Department of Energy, many types of electrical generation (e.g., distributed energy resources), storage options [17], advanced metering infrastructure [12,21], as well as the active participation of consumers (e.g., demand-response programs), will in future be integrated to form a huge network of heterogeneous intelligent participants. This will introduce more rigorous reliability and security requirements due to the increasing interdependency and complexity of electric elements. In this context, this work is dedicated to tackling the problem of how to prevent the occurrence of cascading failures by taking into consideration autonomous behaviors and real-time interactions of heterogeneous participants in a smart grid.

Current reliability policies in power systems focus mainly on secure their normal operation under the most severe single or at most two contingencies, known as the $N-1$ and $N-2$ criteria. Many academic studies have been conducted to impose power balance by solving various contingency-constrained unit commitment problems in a centralized manner [20,30]. The results have shown that, to extend the policy to more tighter criteria (i.e., $N-k$ criteria, where $k>2$) becomes intractable due to the computational burden for the huge number of contingency states $\left(\text{i.e., } \sum_{i=1}^{k}\binom{N}{k}\right)$. Along this line, Street et al. [30] have proposed a computationally efficient framework using robust optimization, which does not depend on the size of the set of credible contingencies. However, in their work, the number of contingencies is required as prior knowledge. Recently, to avoid the computational burden, an inverse problem have been proposed, that is, to identify a small group of line contingencies that can trigger a blackout with a certain level of severity [26,10]. For example, Pinar et al. [26] have

* Corresponding author. Tel.: +852 34117088; fax: +852 34117892.
E-mail address: jiming@comp.hkbu.edu.hk (J. Liu).

http://dx.doi.org/10.1016/j.ijepes.2014.12.041
0142-0615/© 2014 The Authors. Published by Elsevier Ltd.

shown that such a problem can be approximately transferred to be a combinatorial network inhibition problem. Although the centralized optimization approaches adopted by above-mentioned studies are beneficial to analyze and improve system reliability, they are practically infeasible to control power flow in real time and handle multiple and simultaneous contingencies affecting several parts of a power system [27]. Therefore, it would be desirable to control power systems in a decentralized manner [33].

As Amin and Schewe [4] argued, a smart grid that automatically responds to emergencies could reduce the rising number of debilitating blackouts. The last few years have witnessed the development of new technologies, services and concepts to improve grid reliability and security in the face of system disturbances. Specifically, to facilitate the real-time monitoring and control of a grid, a series of advanced communication infrastructure, modern sensors, real-time simulation tools, as well as intelligent protection applications have been introduced. For example, advanced metering infrastructure, which is a means to facilitate two-way communications, will let utilities send real-time pricing signals to consumers and thus encourage consumers to implement direct control of demand-side management [12]. By doing so, a look-ahead simulation tool may send corrective instructions to control devices in less than half a second [4]. All these research and development efforts offer new opportunities for us to systematically design the "immune system" of a smart grid [21], where each individual participant continually performs self-assessments to deter, detect, respond to and restore grid components from unpredictable contingencies, and at the same time optimizing its performance in a decentralized manner [2,16].

Technically speaking, loads and generation in a smart grid must ultimately balance in real time to maintain stable. When line contingencies happen, the power flow will be redistributed to other lines based on the operational conditions of the system. Further, due to the physical capacity constraint of transmission lines, a small number of line contingencies may result in overload or failures in other lines. In this case, it would be necessary to shed some amount of loads and generation such that the power overload on other lines can be avoided. Specifically, this work aims to address two important issues to prevent the occurrence of cascading failures. The first and most important is to quickly shed a *minimum* amount of loads and generation to secure the grid after contingencies happen. This is because both under-shedding and over-shedding may cause unnecessary damages/costs for grid participants. The second is to make appropriate *compensations* for heterogeneous load-shedding participants. In current power systems, the compensations are usually determined by pre-signed contracts [32], which cannot reflect the real-time operational status of the system during the contingencies. As for the increasing number of active participants in a smart grid, it would be desirable to make compensation for individual participants by taking into consideration their heterogeneous shedding costs in real time. By doing so, if the compensations can cover the shedding costs for each individual participant, power loaders and generators will be motivated to participate in the autonomous load shedding process during line contingencies.

To mathematically formulate the above-mentioned two issues, this work first approximate the complex power flows in a smart grid using a dc model, i.e., the linearized active power flow model [14]. Then, the grid is modeled as a directed network, where each node represents a bus and each edge represents a transmission/distribution line. By doing so, the load-shedding problem is formulated as a network optimization problem for minimizing the total shed amount under the power flow constraints. Further, to fairly compensate load-shedding participants, the concept of load-shedding fairness is introduced based on the criterion of proportional fairness [29,23].

This work presents a fair load-shedding algorithm to solve the load-shedding problem in a decentralized manner. Based on the algorithm, each agent representing a loader/generator in a grid need only monitor the operational situations in their local areas and interact with their neighboring agents to coordinately determine the amount of loads/generation they should shed during line contingencies. From the perspective of grid security, there are two major steps in the algorithm: the first step aims to prevent the contingencies spreading to other lines by shedding a necessary amount of loads and generation, while the second step aims to recover as many shed loads/generation as possible to reduce the total amount of load shedding. From the perspective of load-shedding fairness, an embedded feedback mechanism is involved in the first step to facilitate communications among agents in terms of load-shedding compensations. Specifically, a focal loader (or generator) determines the total amount of compensations based on compensation requests from its downstream (or upstream) neighbors on the one hand, and adjust its request to its upstream (or downstream) neighbors on the other hand. During this process, the focal loader (or generator) does not need to know the exact cost functions of its downstream/upstream neighbors. By doing so, the marginal compensation prices for per unit of shed loads/generation can be collectively and fairly determined by all load-shedding participants.

The remainder of this paper is organized as follows. In Section 'Problem statement', we briefly introduce the dc power flow model and formally formulate the load-shedding problem by taking into consideration the load-shedding fairness. In Section 'A decentralized fair load-shedding algorithm', we present the decentralized fair load-shedding algorithm associated with an embedded feedback mechanism. In Section 'Experimental results and discussions', we validate and demonstrate the properties of the algorithm by carrying out experimental studies on the standard IEEE 30 bus system. Finally, we conclude this paper and list several issues that are worth to be further pursued in Section 'Conclusion and future work'.

Problem statement

In this section, we formally define the load-shedding problem in the face of line contingencies and present the issue of load-shedding fairness.

Active power flow model

A power system can be modeled as a directed network $G(V,E)$, where each node $v_i \in V$ represents a bus and each edge $e_{kl} \in E$ represents a transmission/distribution line from v_k to v_l. Similar to existing studies [10,26,8], this work focuses mainly on the active power flow in $G(V,E)$, which can be further simplified to be the dc power flow model $P_{kl} = -b_{kl}(\theta_k - \theta_l)$. Here, θ_k represents the phase angle variable at node v_k and b_{kl} represents the susceptance of e_{kl}.

To describe the power flow at the network level, we define an arc-node incidence matrix A of $G(V,E)$ with n nodes and m edges. The ith row represents the ith edge in $G(V,E)$ and the jth column represents the jth node v_j in $G(V,E)$. Specifically, each entry $a_{ij} \in A$ is defined as follows:

$$a_{ij} = \begin{cases} -1, & \text{if the } i\text{th edge is an out-edge from } v_j, \\ 1, & \text{if the } i\text{th edge is an in-edge to } v_j, \\ 0, & \text{otherwise.} \end{cases} \quad (1)$$

Denote D as a $n \times 1$ vector representing the power generation (i.e., $D_i > 0$) or loads (i.e., $D_i < 0$) on each node. Based on the property of flow conservation [19], for each node in $G(V,E)$, the power inflows should be equal to the power outflows. Therefore, the power flow in $G(V, E)$ can be described as $A^TBA\Theta - D = 0$, where B is an $m \times m$ diagonal matrix whose entries correspond to line susceptance (i.e., b_{kl}) and Θ is an $n \times 1$ vector whose entries represent phase angle variables of corresponding buses. Similar to the work of Pinar et al. [26], a binary variable γ_i is adopted to indicate whether the ith edge is in service or not. If the edge is in service, $\gamma_i = 0$. Otherwise, if the edge is out of service, $\gamma_i = 1$. By doing so, the power flow model with line contingencies can be expressed as

$$A^TB\Gamma A\Theta - D = 0, \tag{2}$$

where Γ is an $m \times m$ diagonal matrix, whose i^{th} entry is $1 - \gamma_i$.

The load-shedding optimization problem

In reality, the power flows on each edge e_{kl} is constrained by its transmission capacity. When line contingencies happen, some demand may not be met due to the reduced transmission capability of the power grid. In this case, shedding some amount of loads by cutting off supply to some consumers is necessary to protect the transmission infrastructures and avoid cascading failures [9]. Furthermore, due to the constraint of load-generation balance, the same amount of generation should also be reduced by generators. Denote $\overline{D}$ as the vector of power generation or loads after the load shedding, S and L as the set of generator nodes and load nodes, respectively. Then, the total amount of load shedding can be calculated as $Z = \sum_{v_i \in S}(D_i - \overline{D}_i) = \sum_{v_j \in L}(\overline{D}_j - D_j)$.

The load-shedding problem becomes the following network optimization problem:

$$\min Z = \sum_{i \in S}(D_i - \overline{D}_i), \tag{3}$$

$$s.t.\ A^TB\Gamma A\Theta - \overline{D} = 0, \tag{4}$$

$$P_{kl} \leqslant \mu_{kl}, \forall e_{kl} \in E, \tag{5}$$

$$-\pi/2 \leqslant \Theta \leqslant \pi/2. \tag{6}$$

Here, Eq. (4) represents the power flow constraints under line contingencies. Moreover, the *transmission constraint* of a system consists of both the physical transmission capacity μ of each edge (i.e., Eq. (5)) and the angular constraint of the power flow model (i.e., Eq. (6)). Based on the dc power flow model, the power flow on $e_{kl} \in E$ is constrained by $-b_{kl}\pi/2$, i.e., $P_{kl} \leqslant -b_{kl}\pi/2$. Together with the physical capacity, we have the transmission constraint $P_{kl} \leqslant \tau_{kl}$, where $\tau_{kl} = \min\{\mu_{kl}, -b_{kl}\pi/2\}$. Such a constraint will largely affect the load-shedding results.

A useful property can be given based on the Lemma 1.1 proposed by Bienstock and Verma [10].

Lemma 1. *Suppose a power system G is connected after line contingencies. If the loads and generation are in balance, i.e.,* $\sum_{v_i \in V}\overline{D}_i = 0$, *then the Eqs. (4) and (6) determine a unique power flow on each line.*

According to the lemma, if we could maintain the load-generation balance during the load-shedding process, there must be a solution for the phase angle θ_i at each node $v_i \in V$. Hence, the load-shedding algorithms in this paper focus mainly on the maintenance of the load-generation balance instead of the adjustment of phase angles.

The load-shedding fairness

After a line contingency happens on e_{ij}, j's downstream loaders may need to shed some amount of loads to prevent the cascading failures. While at the same time, i's upstream generators need to reduce the same amount of power generation to maintain the load-generation balance. Here, we mainly introduce the load shedding of j's downstream loaders. Suppose that the system will compensate each loader who participate in the load shedding with respect to the amount of loads it shed. Therefore, the total amount of compensations will depend on the participated loaders L^p. Since in reality loaders may evaluate their shedding costs differently, we use a function $C_i(x_i)$ to represent the cost function of i, where x_i $(= \overline{D}_i - D_i)$ represents the actual shed loads by i. From a systematic viewpoint, it would be reasonable to minimize the following aggregate costs of all participated loaders after a contingency happens:

$$\min \sum_{v_i \in L^p} C_i(x_i), \tag{7}$$

$$s.t. \sum_{v_i \in L^p} x_i \geqslant P_c, \tag{8}$$

$$x_i \geqslant 0. \tag{9}$$

where P_c is the minimal amount of loads that need to be shed after the contingency.

One important issue for load shedding among heterogeneous loaders is the load-shedding fairness among the participated loaders. To fairly determine the amount of shed loads among loaders in L^p, this work utilizes the concept of *proportional fairness* proposed in [22,23,29], which is defined as follows:

Definition 1. A vector of shed loads $x = \{x_i, v_i \in L^p\}$ is proportionally fair if it is feasible, i.e., $x \geqslant 0$, and if for any other feasible vector x^*, the aggregate of proportional changes satisfies:

$$\sum_{v_i \in L^p} \frac{x_i^* - x_i}{x_i} \leqslant 0. \tag{10}$$

The load-shedding fairness for generators can be defined in a similar way. Specifically, this work aims to solve the load-shedding problem in a decentralized manner, such that (i) the loaders and generators can respond to unpredictable line contingencies to prevent cascading failures; (ii) the load-generation balance can be maintained in real time; and (iii) a minimum amount of loads and generation can be fairly shed by heterogeneous participates.

A decentralized fair load-shedding algorithm

In reality, the advanced communication technologies in a smart grid allows power buses to communicate and interact with each other to maintain grid reliability. Therefore, in this paper, we assume that there is an intelligent agent with certain computational capability located at each node of a given network $G(V, E)$. Each agent updates its profile by communicating with its neighbors, and adaptively determines the amount of shed loads or generation when line contingencies happen.

Agents' profiles

An agent i's profile consists of two components, i.e., $\langle AS_i, AL_i \rangle$ and $\langle MUG_i(\cdot), MUL_i(\cdot) \rangle$, where the first component is used to determine the maximal load shedding that i can conduct and the second component is used to recover as many loads as possible after contingencies happen. To quickly respond to contingencies happened anytime and anywhere in a power grid, the agents should monitor the status of the grid in real time. For a given agent i, we denote AS_i as the *available supply* that can be reduced by i and its upstream neighbors, and AL_i as the *available loads* that can be shed by i and its downstream neighbors. Moreover, we use the notations

$I_i = \{k | e_{ki} \in E\}$ and $O_i = \{k | e_{ik} \in E\}$ as i's upstream and downstream neighbors. Suppose that before the contingencies happen, the grid is stable, i.e., the power flow satisfies transmission constraints on each edge as well as the flow conservation constraint on each node. Denote $\widehat{D}_i$ as the power generation (if $\widehat{D}_i > 0$) and the consumption (if $\widehat{D}_i < 0$) by agent i, respectively. Then, the values of AS_i and AL_i can be calculated as follows: If i has no upstream neighbors, then we have $AS_i = \widehat{D}_i$; if i has no downstream neighbors, then we have $AL_i = \widehat{D}_i$. For other agents (e.g., l), the values of AS_l and AL_l are determined by:

$$AS_l = \widehat{D}_l + \sum_{k \in I_l} P_{kl} \tag{11}$$

$$AL_l = \widehat{D}_l - \sum_{j \in O_l} P_{lj} \tag{12}$$

where P_{lk} and P_{lj} are determined by the dc power flow model. Essentially, the tuple of $\langle AS_i, AL_i \rangle$ represents i's real-time load shedding capability after the contingencies happen.

We further introduce the notion of *residual network* $\widetilde{G}(V, E')$, which represents the incremental transmission capability of the network $G(V, E)$. The node set of $\widetilde{G}(V, E')$ is the same as that of $G(V, E)$. In a residual network, if $e_{kl} \in E$, then $e_{kl} \in E'$ and its capacity corresponds to the unused transmission capacity of e_{kl}, i.e., $\tilde{\tau}_{kl} = \tau_{kl} - P_{kl}$; If $e_{lk} \in E$, then $e_{kl} \in E'$ and its capacity corresponds to the used transmission capacity of e_{lk}, i.e., $\tilde{\tau}_{kl} = P_{lk}$; Otherwise, e_{kl} does not belong to E'. Fig. 1 shows an example of residual network. An *augmenting path* (e.g., $< e_{12}, e_{23}, e_{34}, e_{46} > \in \wp_{16}$ in Fig. 1) from i to j in $\widetilde{G}(V, E')$ is defined as a series of distinct agents connecting i and j, where $\wp_{ij}$ represents the set of augmenting paths from i to j. If there is at least one augmenting path from k to i in $\widetilde{G}(V, E')$, we say that k has access to i (or i is accessible from k).

Denote $\widetilde{D}_i\ (= D_i - \widehat{D}_i)$ as the unused generation (if $\widetilde{D}_i > 0$) or the unsatisfied loads (if $\widetilde{D}_i < 0$) of agent i. Each agent will remember both the *minimal unused generation* (i.e., $MUG_i(g)$) for each generator g that has access to i and the *minimal unsatisfied loads* (i.e., $MUL_i(l)$) for each loader l that has access to i. For each agent i, the values of $MUG_i(g)$ and $MUL_i(l)$ can be updated as follows:

$$MUG_i(g) = \min_{e_{ki} \in \widetilde{I}_i^g} \min\{\tilde{\tau}_{ki}, MUG_k(g)\} \tag{13}$$

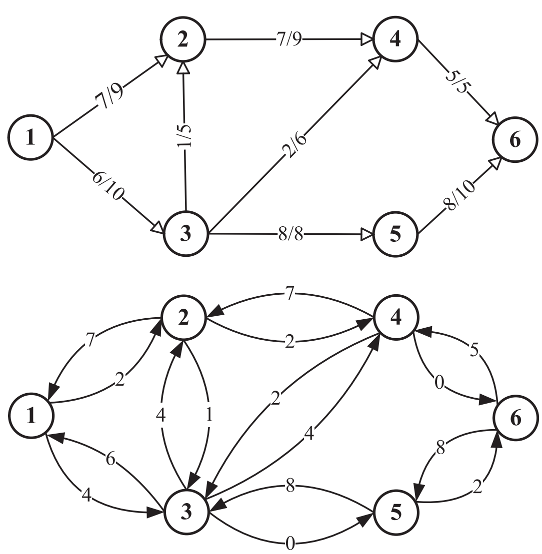

Fig. 1. An illustration of a flow network and its residual network. The upper figure shows the original network with flow/capacity on each line; the lower figure shows its residual network with capacity on each line.

$$MUL_i(l) = \min_{e_{ji} \in \widetilde{I}_i^l} \max\{-\tilde{\tau}_{ji}, MUL_j(l)\} \tag{14}$$

where we have $\widetilde{I}_i^g = \{k | e_{ki} \in \wp_{gi}\}$ and $\widetilde{I}_i^l = \{j | e_{ji} \in \wp_{li}\}$. Moreover, $MUG_g(g) = \widetilde{D}_g$ and $MUL_l(l) = \widetilde{D}_l$. Here $MUG_i(g) > 0$ means that generator g can supply at least $MUG_i(g)$ power generation to i's downstream agents through i, while $MUL_i(l) < 0$ means that l can consume at lease $|MUL_i(l)|$ supply from i's upstream agents through i.

Suppose a contingency happens on an edge e_{ij}, the available supply of agent j as well as the available loads of agent i will be directly affected. Accordingly, the AS of j's downstream agents as well as AL of i's upstream agents will also be updated based on Eqs. (11) and (12). However, the AL of j's downstream neighbors and the AS of i's upstream neighbors will not change after the contingency happens. On the other hand, the network $\widetilde{G}(V, E')$ will be updated once the grid reaches a new stable state, i.e., a new load-generation balance is reached. Consequently, starting from each loader and generator and then propagating into the network $\widetilde{G}(V, E')$, agents' MUG and MUL can be updated based on the rules, i.e., Eqs. (13) and (14).

With respect to load shedding, over-shedding of loads or generation may cause unnecessary shedding costs. Meanwhile, under-shedding of loads or generation may fail to prevent the cascading failures, which may cause tremendous damages. Therefore, the load-shedding algorithm in this paper complies with the following two steps:

- *Step 1 (Load shedding)*: The agents coordinate and interact with their neighbors to *fairly* reduce the same amount of loads and generation so as to re-balance the grid.
- *Step 2 (Load recovery)*: The agents gradually recover as much as generation, and at the same time, maintain the load-generation balance.

Load shedding using an embedded feedback mechanism

In reality, loaders or generators may evaluate their shedding costs variously due to different factors such as the operational conditions and the contents of interruptible contracts [6]. Therefore, it would be helpful to appropriately compensate the potential loaders and generators so as to motivate them participate in the load shedding. In the following, we introduce an embedded feedback mechanism to determine the load-shedding compensations in real time based on the proportional fairness criterion, i.e., the optimization problem (7)–(9) in Section 'The load-shedding fairness'.

Decentralized load-shedding fairness

Each agent in the load-shedding algorithm need only communicate with its upstream and downstream neighbors. Suppose a line contingency happens on e_{ij}, an amount of required shed loads P_{ij} and corresponding compensations will be announced by agent j to its downstream neighbors in O_j. Then, each of its downstream neighbors $k \in O_j$ may request a compensation v_k for participating the load shedding. The compensation v_k requested by k consists of two components: (i) the compensation for itself w_k, which is determined by k's cost function $C_k(\cdot)$; and (ii) the compensation w_{O_k}, which is used to compensate the k's downstream neighbors who participate in the load shedding.

We first introduce how to determine the value of $w_k \in O_j$. After receiving all requests from its downstream neighbors, agent j will determine the compensations for per unit amount of shed loads λ_j and allocate shed loads $\phi_k = v_k/\lambda_j = x_k + x_{O_k}$ to k, where $x_k = w_k/\lambda_j$ and $x_{O_k} = w_{O_k}/\lambda_j$. Since the compensation w_{O_k} will be used to compensate the loads x_{O_k} shed by k's downstream neighbors, the rational behavior for each $k \in O_j$ is to shed x_k loads and minimize its own cost by solving the following problem:

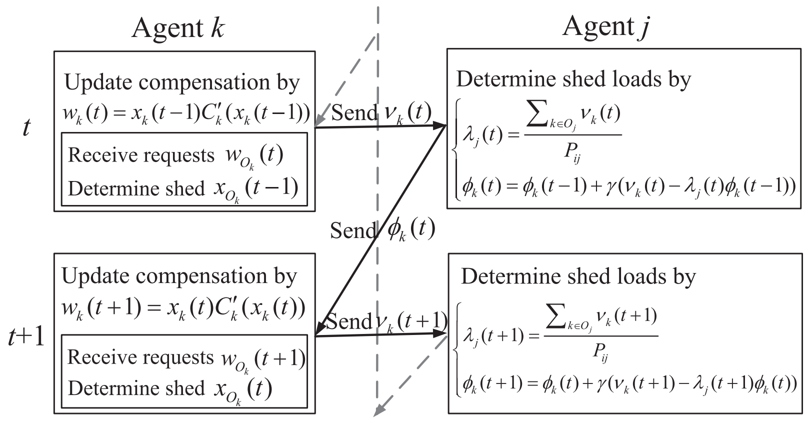

Fig. 2. An illustration of the interaction procedure between agent k and a focal agent j in the embedded feedback mechanism for the load-shedding fairness.

$$\min_{w_k \geqslant 0} C_k(w_k/\lambda_j) - w_k. \tag{15}$$

Given λ_j, the optimal value of w_k should satisfies: $C'_k(w_k/\lambda_j) = \lambda_j$, which is equivalent to

$$w_k = x_k C'_k(x_k). \tag{16}$$

This means that k can optimally respond to the allocated shed loads x_k based on its cost function $C_k(\cdot)$ by adjusting its requested compensation w_k to j.

According to the analysis by Kelly [22], to achieve the proportional fairness for allocating shed loads to j's downstream neighbors, we have the following lemma:

Lemma 2. *If agent j knows the vector $v = \{v_k | k \in O_j\}$ and attempts to determine ϕ_k by solving the following optimization problem:*

$$\min_{\phi_k \geqslant 0} \sum_{k \in O_j} v_k \log \phi_k \tag{17}$$

$$s.t. \sum_{k \in O_j} \phi_k \geqslant P_{ij}, \tag{18}$$

then the optimization problem (7)–(9) can be solved so as to achieve the proportional fairness.

The Lagrangian for the problem (17), (18) is:

$$L(\phi; \lambda_j) = \sum_{k \in O_j} v_k \log \phi_k + \lambda_j \left(P_{ij} - \sum_{k \in O_j} \phi_k \right). \tag{19}$$

After eliminating the terms not dependent on λ_j, the associated dual problem becomes:

$$\min_{\lambda_j \geqslant 0} \sum_{k \in O_j} v_k \log \lambda_j - \lambda_j P_{ij}. \tag{20}$$

Solving the dual problem, the compensations for per unit amount of shed loads can be determined as:

$$\lambda_j = \sum_{k \in O_j} v_k / P_{ij}. \tag{21}$$

By doing so, agent j does not need to know the exact cost functions of its downstream neighbors to determine the compensations for per unit shed loads.

An embedded feedback mechanism and its implementation

The above analysis shows that there is an embedded feedback mechanism for load-shedding fairness at the system level, where a focal agent (i) determines the marginal compensations for per unit of shed loads based on the requests from its downstream neighbors, and at the same time (ii) update its compensation request to its upstream neighbor based on the announced per unit compensation. An illustration of the interaction procedure between k and j for fair load shedding is shown in Fig. 2. At iteration t, agent $k \in O_j$ updates its compensation request based on the allocated shed load from j at the previous iteration, where $w_k(t) = x_k(t-1)C'_k(x_k(t-1))$. At the same time, k determines the total shed loads $x_{O_k}(t-1)$ to its downstream neighbors and receives a new request $w_{O_k}(t)$ from its downstream neighbors. Then, a new request $v_k(t)$ is send to j. After receiving all requests from its downstream neighbors, j calculates a new compensation for per unit shed loads $\lambda_j(t)$ based on Eq. (21). Given $\lambda_j(t)$, using gradient projection method, the allocated shed load $\phi_k(t)$ can be updated by

$$\phi_k(t) = \phi_k(t-1) + \gamma(v_k(t) - \lambda_j(t)\phi_k(t-1)) \tag{22}$$

where γ represents the step size for updating. Similar to j, k can determine the shed loads for each of its downstream neighbor $l \in O_l$, where the difference is that the minimal shed loads becomes to be $x_{O_k}(t)$. Then, l will update the compensation w_l based on the allocated shed loads x_l. For each agent k, the procedure will stop when the difference $\phi_k(t) - \phi_k(t-1)$ is less than a threshold κ.

Load recovery using the concept of residual network

After the first step is completed, a total amount of P_{ij} load and generation will be fairly shed among a set of load-shedding participants. According to Lemma 1, the system reaches a new stable

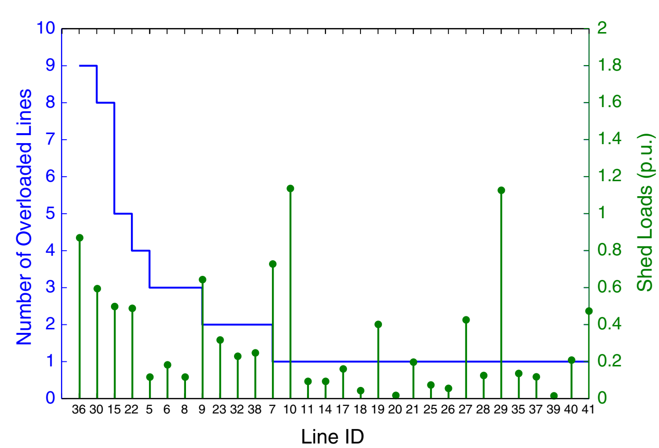

Fig. 3. The relationship between the severity of a line contingency and the amount of shed loads needed to prevent cascading failures with the factor of safety $K = 1.5$ on each line. The stair values show the number of overloaded lines caused by a corresponding line contingency, while the stem values show the amount of shed loads needed to prevent cascading failures.

state with a new phase angle at each node. Since the residual network $\widetilde{G}(V,E')$ is flow dependent, the residual network $\widetilde{G}(V,E')$ as well as the values of $\langle MUG_i(\cdot), MUL_i(\cdot)\rangle$ for each agent i, will also be updated. In the following, we introduce how to recover as many loads and generation as possible after the fist step of the load-shedding algorithm.

For a generator agent g with $\widetilde{D}_g > 0$, denote $MUL_g = \max_l\{|MUL_g(l)|\}$ as the maximum value of minimal unsatisfied loads from loaders that are accessible from g. The loader agent $l_{\max}$ can be identified by g as $l_{\max} = \arg\max_l\{|MUL_g(l)|\}$. It means that $l_{\max}$ can consume at least $MUL_g = |MUL_g(l_{\max})|$ power generation from g. However, g itself can supply $\widetilde{D}_g = MUG_g(g)$ generation. Therefore, an amount of $\min\{MUG_g(g), MUL_g\}$ generation and loads will be recovered simultaneously at generator g and $l_{\max}$. After that, according to Lemma 1, the power flow and phase angles of the system will reach a new stable state. Accordingly, the agents' profiles will also be updated based on the interaction rules in Section 'Agents' profiles'. The above process will not terminate until there is no generator with $(MUG_g(g) > 0)$ and $(MUL_g > 0)$ simultaneously.

Generally, for any generator agent g, if $(MUG_g(g) > 0) \wedge (MUL_g > 0)$, there is at least a positive amount of $\Lambda_g = \min\{MUG_g(g), MUL_g\}$ power generation can be supplied by g to $l_{\max}$. Therefore, at each round, at least a positive amount of shed loads will be recovered. Based on the decentralized algorithms for maximum flow problem in networks proposed by Segall [28], the following lemma shows that after the step 2 of our algorithm, no more shed loads can be recovered before the failed line resumes service again.

Lemma 3. *If there is no generator g whose profile satisfies $(MUG_g(g) > 0) \wedge (MUL_g > 0)$, then no additional loads or generation can be recovered under the transmission constraint of the grid.*

Generally speaking, the load recovery problem is similar to the maximum flow problem in networks [18], which has been extended to other problems such as the maximum flow network interdiction problem [3], and the network inhibition problem related to the vulnerability of power grids [26]. However, the problem and algorithm for power grid are unique in terms of the following two aspects:

- Even though we can identify the loader $l_{\max}$ for a generator g to recover Λ_g generation and loads simultaneously, we cannot exactly know which path(s) the recovered power generation will pass through from g to $l_{\max}$. However, what we are sure is that the power can only flow along the paths through which $l_{\max}$ is accessible from g. The agents' profiles and updating rules in our algorithms are particularly designed based on such an observation.
- Since it is hard to predict the quantity of power flow on each edge before a stable state is reached, we must also prevent the overload on any edges when we recover the same amount of generation and loads for g and $l_{\max}$. The definition of minimal unused generation and minimal unsatisfied loads make sure that all paths from g to $l_{\max}$ can afford Λ_g power flow. By doing so, no edge will overload during the recovery process.

Moreover, for the security reason, the recovery process must be implemented in a sequential order. According to the algorithm, the edge capacity in the residual network will significantly affect the recovery process. In other words, the residual capacity of each edge plays important roles in recovering loads and generation after the load shedding. This provides a new angle of view to make reliability policies in terms of managing the capacity of transmission lines.

Experimental results and discussions

In this section, we demonstrate the properties of the fair load-shedding algorithm by conducting an experimental study on the standard IEEE 30 bus system.

The IEEE 30 bus system and parameter settings

The system consists of 30 buses and 41 transmission lines. There are 6 generators, 6 transmission intermediaries and 18 loaders (see Fig. 3 in the work of Donde et al. [15]. To demonstrate the performance of the fair load-shedding algorithm, we slightly modify the values of active power injections and phase angles in Table 1 such that there is a natural power flow conservation at each bus. For the line reactance, we use the same values as described in the work of Donde et al. [15].

There are four parameters that should be specified in this paper:

- *Line capacity*: In practice, the values of line capacity are hard to obtain and are usually affected by many factors. Based on the concept of N-resilient grids [7], we consider over-provisioning of lines capacity by a constant fraction (the *factor of safety K*) of the initial power flow, i.e., $\mu_{kl} = K \cdot P_{kl}$. The real power grid is usually assumed to be at least N-resilient with $K = 1.2$. In this paper, we will consider the scenarios with $K = 1.2, 1.3, 1.4$ and 1.5, respectively.
- *Shedding cost functions*: We assume that agents shedding cost functions have the form $C_k(x_k) = C_0 + A_k \log x_k$, where C_0 represents a constant cost and A_k is the available loads/generation to be shed by k with respect to its corresponding neighbor.
- *Updating step size*: We set the update size $\gamma = 0.2, 0.3$ and 0.4 to demonstrate the convergence speed of the fair load-shedding algorithm.

Table 1
IEEE 30 Bus system data (buses).

Bus #	P inject (net) *p.u.*	Ang., θ rad	Bus #	P inject (net) *p.u.*	Ang., θ rad	Bus #	P inject (net) *p.u.*	Ang., θ rad
1	0.1750	0.0387	11	0	0.0315	21	−0.7671	0.1073
2	1.0366	0.0390	12	−0.4073	0.2047	22	1.4782	0.1351
3	−0.1100	0.0045	13	2.1664	0.5080	23	1.2807	0.3332
4	−0.2815	0.0017	14	−0.3536	0.1463	24	−0.4166	0.2049
5	0	−0.0469	15	−0.3582	0.1721	25	0	0.3135
6	0	−0.0298	16	−0.2046	0.1041	26	−0.1727	0.2479
7	−1.2878	−0.0985	17	−0.5022	0.0474	27	2.1604	0.4189
8	−1.8064	−0.0842	18	−0.1722	0.0469	28	0	0.0053
9	0	0.0315	19	−0.5584	−0.0047	29	−0.1036	0.2985
10	−0.1638	0.0637	20	−0.1105	0.0066	30	−0.5211	0.2161

- *Stop criterion*: The threshold κ is set to a very small value, i.e., $\kappa = 1 \times 10^{-6}$.

Results and observations

We first investigate the relationship between the severity of a line contingency and the amount of shed loads needed to prevent cascading failures. When a line contingency happens, if no load-shedding algorithm is implemented, the contingency may cause the overload of other lines and trigger cascading failures. This work uses the number of overloaded lines caused by a line contingency to represent its severity. Fig. 3 shows the experimental results for each line contingency, where the factor of safety K is set to 1.5. The stair values show the number of overloaded lines caused by a corresponding line contingency, while the stem values show the amount of shed loads needed to prevent cascading failures. Intuitively, the less severe a line contingency is, the less loads it may need to shed. However, it is not the case. It can be observed that although lines 10 and 29 can only cause one line to be overload, the amount of loads needed to shed is more than those that may cause large number of overloaded lines (e.g., line 36). Actually, the shed loads depend on the position of the line in the system as well as the operational conditions of the system in the face of contingencies, which emphasizes the necessity of the real-time decentralized load-shedding algorithms. Moreover, some contingencies may cause only small number of overloaded lines at the earlier stage and suddenly result in a large number of overloaded lines. For example, the contingency on line 32 can cause the overload of lines 20 and 30. However, the failure of line 30 may further result in at least 8 overloaded lines. Such phenomenon can be verified by real-world blackouts that small failures trigger the failure of critical lines and further cause a large-scale blackout [5].

We then focus on demonstrating the properties of our algorithms for different scenarios of line contingencies. Donde et al. [15] have identified three transmission lines (i.e., lines 28, 29 and 36) as critical ones in the IEEE 30 bus system, whose removal

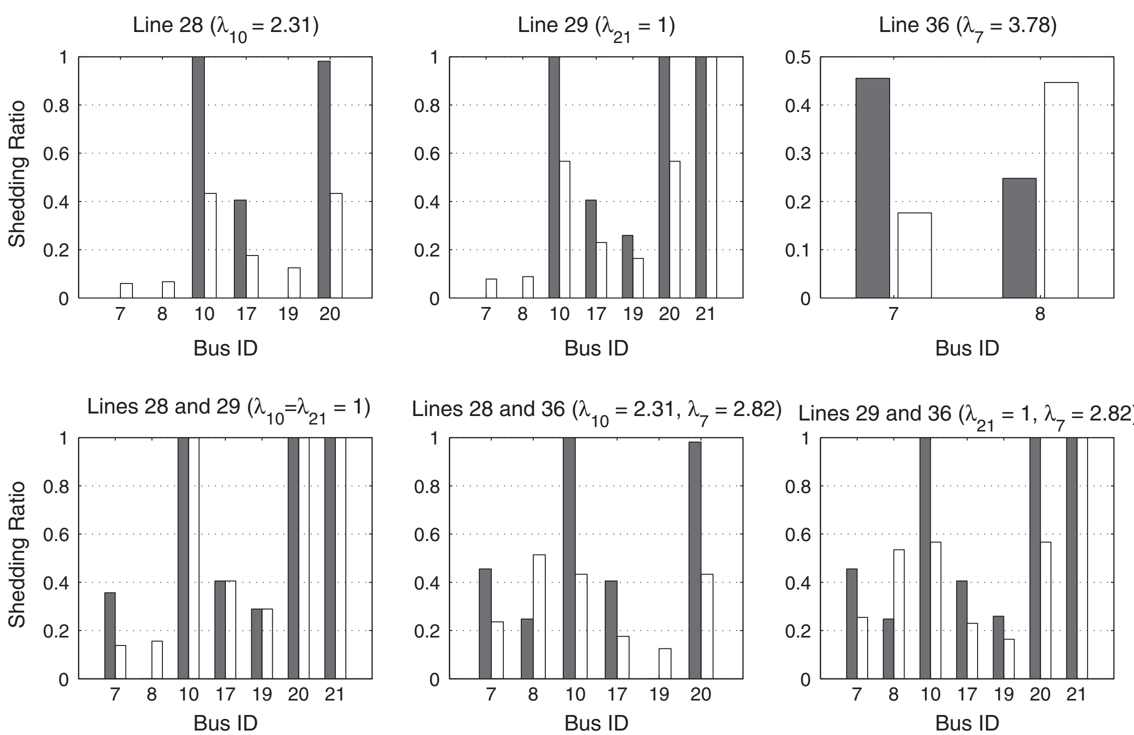

Fig. 4. An illustration of load-shedding results after the first step of load-shedding algorithms for six scenarios of line contingencies. The black bars show the results of an algorithm without considering load-shedding fairness, while the white bars show the results of the fair load-shedding algorithm.

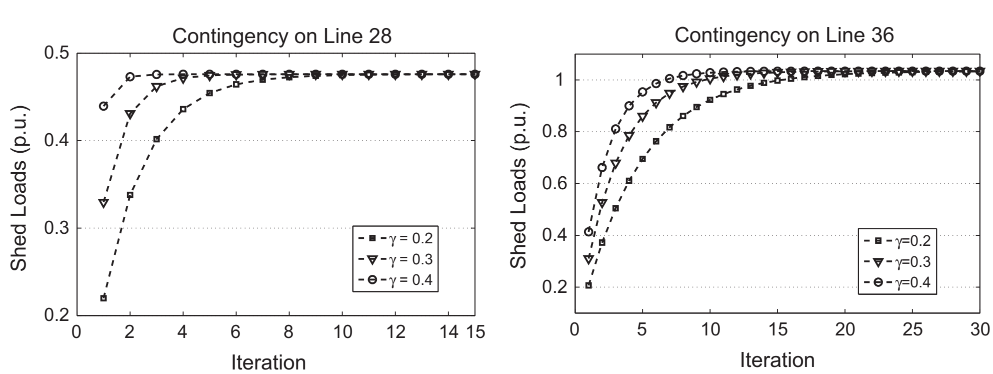

Fig. 5. The convergence speed of our fair load-shedding algorithm with respect to updating step size $\gamma = 0.2, 0.3$ and 0.4. The left figure shows the load shedding for the failure of line 28; the right figure shows the load shedding for the failure of line 36.

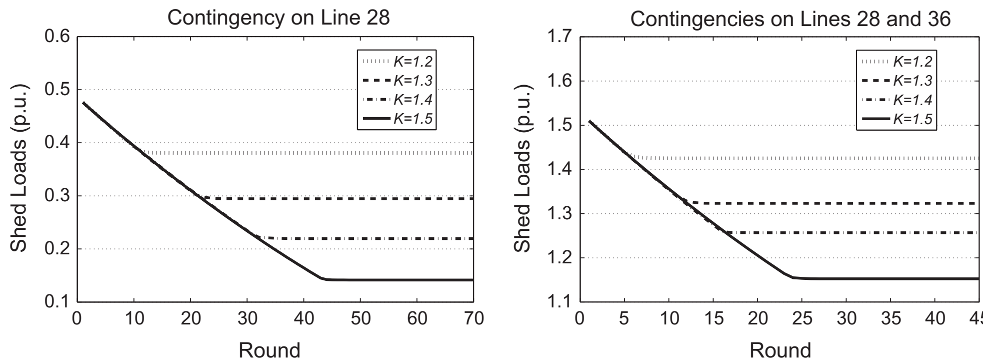

Fig. 6. An illustration of load recovery processes with respect to different factor of safety *K*. The left figure shows the load recovery for the contingency on line 28; the right figure shows the load recovery for the contingencies on lines 28 and 36.

Table 2
Load Shedding results for selected line contingencies with the IEEE 30 bus system.

Line ID	Load-shedding algorithms	Shed loads at each bus (*p.u.*)							Total shedding
		Bus 7	Bus 8	Bus 10	Bus 17	Bus 19	Bus 20	Bus 21	(*p.u.*)
28	Fair	0.0186	0.0635	0.0125	0.0298	0.0114	0.0057	0	0.1415
29	Fair	0.0852	0.1440	0.0772	0.0999	0.0758	0.0470	0.7515	1.2805
36	Fair	0.1447	0.7247	0	0	0	0	0	0.8694
28, 29	Fair	0.1464	0.2136	0.1323	0.1723	0.0506	0.0790	0.7357	1.5297
28, 36	Fair	0.2387	0.8637	0.0091	0.0230	0.0081	0.0099	0	1.1526
29, 36	Fair	0.3029	0.9418	0.0680	0.0907	0.0667	0.0378	0.7423	2.2501

may cause severe contingencies. Accordingly, this work implements the proposed algorithms for the same line failures and their combinations. Fig. 4 shows the load-shedding results and marginal compensation prices for some agents after the first step of the fair load-shedding algorithm for six scenarios of line contingencies, where the shedding ratio for each loader is calculated by the fraction between actual shed loads and its original loads. The black bars show the results of an algorithm without considering load-shedding fairness, while the white bars show the results of the fair load-shedding algorithm. It can be observed that there are more participants affording the required shed loads in the fair load-shedding algorithm (e.g., contingencies on Line 28, Line 29, Lines 28 and 29, and Lines 28 and 36). More importantly, the fair load-shedding algorithm can determine the marginal compensation price λ for each agent based on real-time power flows to achieve load-shedding fairness. For different focal agent in our embedded mechanism, there will be different marginal compensation prices. Based on the shedding cost functions adopted in our experiments, the more loads are required to shed, the lower the marginal compensation price will be. (Note that in reality, different loaders may have different shedding cost functions.) For example, for the line contingency on line 29 requires to shed all available loads of bus 21, the marginal compensation price λ_{21} is equal to 1. While the compensation for other agent (e.g., λ_{10}) will be larger than 1. In this case, to achieve the criterion of proportional fairness, more compensations may be needed. Moreover, for different contingency scenarios, agents may have different marginal compensation prices. For example, in Fig. 4, the marginal compensation prices for bus 7 are different for scenarios Line 36 and Lines 28 and 36.

With respect to the convergence speed of the fair load-shedding algorithm, we carry out simulations for different step sizes, i.e., $\gamma = 0.2, 0.3$ and 0.4. Fig. 5 demonstrates the updating results for contingencies on Line 28 and Line 36, respectively. It can be observed that the step size can significantly affect the convergence speed of the algorithm. The larger the step size is, the faster the algorithm converges. The results also show that even with different step size, the algorithm can finally converge to the same value. That is, the required loads can be shed successfully. However, we cannot immoderately use large step size to improve the convergence speed because our simulations also reveal that if the step size is too large, the algorithm will not converge any more. In reality, the step size can be determined based on expert experiences. We further study how line capacity affects the recovery results (i.e., step 2) of our load-shedding algorithms. Fig. 6 demonstrates the recovery processes for two contingency scenarios (i.e., Line 28, and Lines 28 and 36) with respect to different factor of safety $K = 1.2, 1.3, 1.4$ and 1.5. The results show that after the contingencies happen, a large mount of loads are shed at the first round based on the first step in our algorithm. Subsequently, some load will be recovered gradually based on the second step in our algorithm. However, due to capacity limitation, not all loads can be recovered. The larger the line capacity is, the more loads will be recovered, the less loss the line contingencies will cause. Table 2 shows the final results of the fair load-shedding algorithm for six contingency scenarios with the factor of safety $K = 1.5$. The first column shows the contingencies happened on different line(s); the second column shows our load-shedding algorithms; the third column shows the amount of loads shed at each loader buses; and the final column shows the total amount of shed loads calculated by corresponding algorithms.

Conclusion and future work

In this paper, we have concentrated on the problem of how to prevent the occurrence of cascading failures in a smart grid. Different from the practical under-voltage load shedding and under-frequency load shedding, we have paid special attention to the problem of how to shed the loads and generation before cascading failures happen. Specifically, the power flow dynamics is approximated by the dc power flow model. Then, the load-shedding problem is formulated as a constrained network optimization problem by taking into consideration the load-shedding fairness. Along this line, we have presented a decentralized fair load-shedding algorithm with an embedded feedback mechanism to (i) determine the load-shedding compensations in real time based on the proportional fairness criterion, and (ii) minimize the total amount of shed loads and generation after line contingencies happen. Experimental studies on the IEEE 30 bus system have been carried out to exhibit the properties and effectiveness of the algorithm.

The decentralized algorithm proposed in this paper may help gain more insights into the prevention of cascading failures in a smart grid. However, there are still some issues that are worth to be considered in the future:

- Although a focal agent in the embedded feedback mechanism do not need to know the exact shedding costs of each load-shedding participant, one important issue is how to motivate the participants to announce compensation requests based on their real shedding cost functions.
- We have proposed one way to determine the load-shedding compensations based on the proportional fairness criterion, however, there may be other realistic ways to determine compensations in real time such as the criterion to reduce the total load-shedding costs.
- Since it is difficult to obtain the reactance values of transmission lines in a real-world power system, in this paper, we have only simulated our algorithms in the IEEE 30 bus system. In the future, we would apply use our algorithms to analyze and evaluate the reliability of real-world power systems, such as the North American power grid [1,24], by approximately determining their system parameters.

Acknowledgements

The authors would like to acknowledge the funding support from the Hong Kong Research Grants Council under Grant HKBU211212, the National Natural Science Foundation of China under Grant NSFC81402760, and the National Center for International Joint Research on E-Business Information Processing under Grant 2013B01035 for the research work being presented in this article.

References

[1] Albert R, Albert I, Nakarado GL. Structural vulnerability of the North American power grid. Phys Rev E 2004;69:025103.
[2] Alrifai MT, Hassan MF, Zribi M. Decentralized load frequency controller for a multi-area interconnected power system. Int J Electr Power Energy Syst 2011;33:198–209.
[3] Altner DS, Özlem Ergun, Uhan NA. The maximum flow network interdiction problem: valid inequalities, integrality gaps, and approximability. Oper Res Lett 2010;38:33–8.
[4] Amin M, Schewe P. Preventing blackouts. Sci Am Mag 2007;296:60–7.
[5] Andersson G, Donalek P, Farmer R, Hatziargyriou N, Kamwa I, Kundur P, et al. Causes of the 2003 major grid blackouts in North America and Europe, and recommended means to improve system dynamic performance. IEEE Trans Power Syst 2005;20:1922–8.
[6] Baldick R, Kolos S, Tompaidis S. Interruptible electricity contracts from an electricity retailer's point of view: valuation and optimal interruption. Oper Res 2006;54:627–42.
[7] Bernstein A, Bienstock D, Hay D, Uzunoglu M, Zussman G. Power grid vulnerability to geographically correlated failures – analysis and control implications. In: Technical report 2011-05-06. Columbia University, Electrical Engineering; 2011
[8] Bienstock D. Adaptive online control of cascading blackouts. In: Proceedings of the IEEE power and energy society general meeting, 2011a. p. 1–8.
[9] Bienstock D. Optimal control of cascading power grid failures. in: Proceedings of the fiftieth IEEE conference on decision and control and European control conference, 2011b. p. 2166–73.
[10] Bienstock D, Verma A. The n-k problem in power grids: new models, formulations, and numerical experiments. SIAM J Optim 2010;20:2352–80.
[11] Bompard E, Huang T, Wu Y, Cremenescu M. Classification and trend analysis of threats origins to the security of power systems. Int J Electr Power Energy Syst 2013;50:50–64.
[12] Bouhafs F, Michael M, Madjid M. Links to the future: communication requirements and challenges in the smart grid. IEEE Power Energy Mag 2012;10:24–32.
[13] Carreras B, Newman D, Dobson I, Poole A. Evidence for self-organized criticality in a time series of electric power system blackouts. IEEE Trans Circ Syst I 2004;51:1733–40.
[14] Chen J, Thorp J, Dobson I. Cascading dynamics and mitigation assessment in power system disturbances via a hidden failure model. Int J Electr Power Energy Syst 2005;6274:318–26.
[15] Donde V, López V, Lesieutre B, Pinar A, Yang C, Meza J. Identification of severe multiple contingencies in electric power systems. IEEE Trans Power Syst 2008;23:406–17.
[16] Dou CX, Yang J, Li X, Gui T, Bi Y. Decentralized coordinated control for large power system based on transient stability assessment. Int J Electr Power Energy Syst 2013;46:153–62.
[17] Dunn B, Kamath H, Tarascon J. Electrical energy storage for the grid: a battery of choices. Science 2011;334:928–35.
[18] Ford L, Fulkerson D. Flows in networks. Princeton: Princeton University Press; 1962.
[19] Grainger JJ, Stevenson WD. Power system analysis. New York: McGraw-Hill; 1994.
[20] Hedman K, Ferris W, O'Neill R, Fisher E, Oren S. Co-optimization of generation unit commitment and transmission switching with n-1 reliability. IEEE Trans Power Syst 2010;25:1052–63.
[21] Horowitz S, Phadke A. Boosting immunity to blackouts. IEEE Power Energy Mag 2003;1:47–53.
[22] Kelly FP. Charging and rate control for elastic traffic. Eur Trans Telecommun 1997;8:33–7.
[23] Kelly FP, Maulloo AK, Tan DKH. Rate control for communication networks: shadow prices, proportional fairness and stability. J Oper Res Soc 1998;49:237–52.
[24] Kinney R, Crucitti P, Albert R, Latora V. Modeling cascading failures in the North American power grid. Eur Phys J B – Condens Matter Complex Syst 2005;64:101–7.
[25] Liang C, Wu Z. Performance and reliability of electrical power grids under cascading failures. Int J Electr Power Energy Syst 2011;33:1410–9.
[26] Pinar A, Meza J, Donde V, Lesieutre B. Optimization strategies for the vulnerability analysis of the electric power grid. SIAM J Optim 2010;20:1786–810.
[27] Ray S, Venayagamoorthy GK, Chaudhuri B, Majumder R. Comparison of adaptive critic-based and classical wide-area controllers for power systems. IEEE Trans Syst, Man, Cybern, Part B: Cybern 2008;38:1002–7.
[28] Segall A. Decentralized maximum-flow protocols. Networks 1982;12:213–30.
[29] Shakkottai S, Srikant R. Network optimization and control. Found Trends Network 2007;2:271–379.
[30] Street A, Oliveira F, Arroyo J. Contingency-constrained unit commitment with n-k security criterion: a robust optimization approach. IEEE Trans Power Syst 2011;26:1581–90.
[31] U.S.-Canada Power System Outage Task Force. Final report on the Augest 14, 2003 blackout in the United States and Canada: causes and recommendations. Technical report; 2004.
[32] Zareipour H, Canizares C, Bhattacharya K. An overview of the operation of Ontario's electricity market. In: Proceedings of IEEE power & energy society general meeting, 2005. p. 2463–70.
[33] Zribi M, Al-Rashed M, Alrifai M. Adaptive decentralized load frequency control of multi-area power systems. Int J Electr Power Energy Syst 2005;27:575–83.