IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received May 7, 2018, accepted July 6, 2018, date of publication July 16, 2018, date of current version August 7, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2855153

Mitigating the Risk of Cascading Blackouts: A Data Inference Based Maintenance Method

FENG LIU[1], (Member, IEEE), JINPENG GUO[1], (Student Member, IEEE), XUEMIN ZHANG[1], (Member, IEEE), YUNHE HOU[2], (Senior Member, IEEE), AND SHENGWEI MEI[1], (Fellow, IEEE)

[1]State Key Laboratory of Power Systems, Department of Electrical Engineering, Tsinghua University, Beijing 100084, China
[2]Department of Electrical and Electronic Engineering, The University of Hong Kong, Hong Kong

Corresponding author: Feng Liu (lfeng@tsinghua.edu.cn)

This work was supported by the National Natural Science Foundation of China under Grant 51621065.

ABSTRACT The risk of cascading blackouts (RCB) is of great significance in practice because cascading outages can have catastrophic consequences. As there is a positive relationship between the probability of cascading blackouts and that of component failures, an effective way to mitigate the RCB is to perform maintenance. However, this approach is of limited value when considering extremely complicated cascading outages, such as those in particularly large systems. In this paper, we propose a methodology to efficiently identify the most influential component(s) for mitigating the RCB in a large-power system based on inference from the simulation data. First, we establish a data-based analytic relationship between the adopted maintenance strategies and the estimated RCB. Then, we formulate the component maintenance decision-making problem as a nonlinear 0–1 programming problem. We then quantify the credibility of the estimated RCB and develop an adaptive method to determine the minimum required number of simulations, which is a crucial parameter in the optimization model. Finally, we devise two heuristic algorithms to efficiently identify approximately optimal solutions. The proposed method is then validated by way of numerical experiments based on IEEE standard systems and an actual provincial system.

INDEX TERMS Data inference, maintenance strategy, risk of cascading blackouts.

I. INTRODUCTION

Under certain conditions, system disturbances or component outages in a power system can trigger a sequence of component failures, or cascading outages, that can have serious consequences, even catastrophic blackouts. Although the probability of such catastrophic blackouts is very small, the results of theoretical research and practical experience indicate that the risk of cascading blackouts (RCB) should not be ignored [1]–[4].

Intuitively, component maintenance is seen as an effective means to mitigate RCB since it directly reduces the probability of component failures [5], [6]. In practice, of the large number of components in a power system, a small subset exerts a disproportionately large influence on the RCB [7]–[9]. Therefore, by preferentially maintaining the subset of the most influential components, the RCB can be minimized with limited resources. It should be noted that this is not a new idea, but has been extensively deployed in conventional reliability-centered maintenance or risk-based maintenance methods. In [10], a risk-based resource optimization model for transmission system maintenance is proposed where both the maintenance strategies and corresponding risk reductions serve as input data. From a different perspective, risk can be defined and calculated via scenario enumeration in which components associated with high-risk scenarios are selected for maintenance [11]. In a more rigorous approach, the researchers in [12] leverage 0-1 integer programming to optimize the system risk using limited resources in power systems, where the system risk is defined as the sum of the component risks and is calculated by enumeration. A similar optimization model is presented in [6]. In all of the above methods, maintenance strategies are selected based on evaluating the risk with respect to the maintenance strategies under consideration.

While the aforementioned methods have been shown to work well in both reliability-centered maintenance or risk-based maintenance methods, this is not the case when cascading outages are considered. In fact, few practical methods are available for analyzing cascading outages in large systems due to the complexity of the required computations.

As is well known, the propagation of a cascading outage is a complicated dynamic process involving various random factors, which makes it impossible to analytically calculate the RCB. Hence, statistical methods like the Monte Carlo (MC) method are often employed to indirectly estimate the RCB. However, to realize a credible estimation, the MC method must generate a large number of samples via cascading blackout simulations [13]–[15], which is extremely time consuming [16] for large systems in particular, and maintenance only exacerbates this problem. Since there is no analytic relationship to bridge the estimated RCB with the probability of component failure that vary with maintenance strategies, samples used to estimate the RCB with respect to a particular maintenance strategy cannot be used for another. This implies that whenever the maintenance strategy changes, all blackout samples must be completely regenerated by conducting blackout simulations. Considering the large scale of real power systems and the immense number of possible maintenance strategies, both sample generation and estimation of RCB are extremely time-consuming. This further increases the intractability of the corresponding maintenance optimization problem.

In addition to the MC method, importance sampling [17]–[19] and Splitting [20]–[22] are two other efficient methods to simulate cascading outages. In the importance sampling method, the probability of severe cascading outages is 'amplified' so that more rare cascading blackouts can be captured. As for the Splitting method, the basic idea is to divide the simulation process into sub-levels and then copy samples in the beginning of each sub-level, so that more rare cascading blackouts can be obtained. It is worthy of noting that despite the important sampling method and the Splitting method can remarkably improve the simulation efficiency, they cannot explicitly reveal the relationship between component failure probabilities and the corresponding RCB. Therefore, they face with similar difficulties to the MC method in the maintenance optimization problem.

Guo *et al.* [23] exploit the information pertaining to cascading blackouts that is buried in the data of cascading blackout simulations in their development of a semi-analytic method to characterize the relationship between the unbiased estimation of the RCB and component failure probabilities. With this approach, it is possible to directly estimate the RCB under varying component failure probabilities with no need to regenerate any samples. In other words, this potentially suggests an efficient RCB evaluation approach that can address varying maintenance strategies. In this paper, we extend that approach to address the optimal component maintenance problem while considering the RCB. The major contributions of our work are threefold:

1) An analytic relationship between the estimated RCB and maintenance strategy is established based on inference from the cascading blackout simulation data. On this basis, the optimal component maintenance problem is formulated as a nonlinear 0-1 programming problem. In the model, the evaluation of the RCB with respect to varying maintenance strategies is explicitly expressed based on the initial simulation data.
2) To guarantee the validity of the RCB evaluation result when solving the proposed optimization model, we propose an adaptive method to determine an appropriate sample size via credibility analyses based on inference from the simulation data.
3) The proposed optimization model is a high-dimensional nonlinear 0-1 programming problem. However, for simplicity, we propose two simple heuristic algorithms to search for nearly-optimal solutions with very high efficiency. This guarantees the proposed methodology can be employed in real large power systems.

The rest of this paper is organized as follows. Section II provides an overview of the basic definitions and notations, based on which the optimization problem is formulated. In Section III, based on credibility analyses of the estimated RCB, the sample size, which is the critical parameter, is determined. The two high-efficiency heuristic algorithms and accompanying holistic procedure are presented in Section IV, while several case studies are described in Section V. Finally, our conclusions are presented in Section VI.

II. PROBLEM STATEMENT AND FORMULATION

In this work, we develop a method to mitigate the RCB by maintaining a few key components in large systems. In essence, the influence of maintenance on the component failure probabilities can be estimated based on historical data and experience. Therefore, the critical issue lies in characterizing the relationship between the component failure probabilities and RCB (or more precisely, the estimation of the RCB). As long as the relationship is obtained, it is straightforward to determine suitable maintenance strategies by enumeration or optimization. To this end, an analytic relationship is preferable for optimizing the maintenance strategy. To realize this, we employ the method described in [23] to analytically characterize this relationship through inference from the simulation data, and then to explicitly formulate the optimization problem.

A. DEFINITIONS

We begin by providing basic definitions of cascading outages, the component failure probability function, and the RCB.

Invoking the description in [17], in a broad class of steady-state simulation models, an n-stage cascading outage can be defined as a Markov sequence denoted by

$$Z := \{X_0, X_1, \ldots, X_j, \ldots, X_n, X_j \in \mathcal{X}, \forall j \in \mathbb{N}\}, \tag{1}$$

with respect to probability series $g(Z)$. Here, $\mathbb{N} := \{1, 2, \cdots, n\}$ is the set of cascading stages; j is the stage label; and X_j represents the state variables of the system at stage j, which can be ON/OFF states of components, power injections at each bus, etc. In particular, X_0 is the initial system state of the cascading outage, which is assumed to be deterministic. The state space $\mathcal{X}$ is assumed to be finite. With regard to a specific cascading outage,

$z = \{x_0, \ldots, x_j, \ldots, x_n\}$ and its probability (series) is denoted by $g(z) := g(x_0, \ldots, x_j, \ldots, x_n)$. At stage j, the failure probability of component k is denoted by

$$\varphi_k(x_j) := \mathbf{Pr}(\text{component } k \text{ fails at } x_j\ ), \tag{2}$$

where φ_k is referred to as the *component failure probability function*, which is determined by the inherent characteristics of the component, e.g., the component type, operating condition, etc. In the literature, there are various forms of φ_k [3], [13]–[15], [24]. On the other hand, if a certain component k is maintained, the failure probability function will change accordingly. We denote the failure probability function of component k after maintenance by $\bar{\varphi}_k$. It is worthy of noting that it is difficult to estimate the impact of maintenance on the component failure probability and choose appropriate $\bar{\varphi}_k$ in practice. Because there are many maintenance methods and various kinds of components with different ages, while the accurate data related to the specific component and maintenance method is far from sufficient. However, some existing methods can facilitate dealing with this issue [25]–[27]. In this work, we focus on the influence of component failure probabilities on RCB and the maintenance-based risk management method. Therefore, we simply employ generic forms of φ_k and $\bar{\varphi}_k$, and do not discuss in detail the specific methods to obtain φ_k and $\bar{\varphi}_k$ in this paper.

Note that a cascading outage can involve load shedding. The general definition of the RCB associated with load shedding is as in [17]. Specifically, the load shedding of a cascading outage is denoted Y, which can be considered as a function of the corresponding cascading outage, i.e., $Y = h(Z)$. Then, the RCB with respect to $g(Z)$ and a given load shedding level Y_0 is

$$R_g(Y_0) := \mathbb{E}(Y \cdot \delta_{\{Y \geq Y_0\}}) = \sum_{z \in \mathcal{Z}} g(z)h(z)\delta_{\{h(z)\geq Y_0\}}, \tag{3}$$

where $\mathcal{Z}$ is the set of all possible cascading outages and $\delta_{\{Y \geq Y_0\}}$ is an indicator function of the events $\{Y \geq Y_0\}$, given by

$$\delta_{\{Y \geq Y_0\}} := \begin{cases} 1 & \text{if} \quad Y \geq Y_0; \\ 0 & \text{otherwise.} \end{cases} \tag{4}$$

The RCB defined in (3) is the expectation of load shedding beyond the given level Y_0. When $Y_0 = 0$, this is consistent with the traditional definition of blackout risk. When $Y_0 > 0$, this is consistent with the risk of those events with serious consequences, i.e., those with a load shedding greater than Y_0.

B. ANALYTIC RELATIONSHIP INFERENCE FROM THE SIMULATION DATA

Invoking the Markov property and conditional probability formula, $g(z)$ can be rewritten as

$$g(z) = g(x_n, \cdots, x_1, x_0) = \prod_{j=0}^{n-1} g_{j+1}(x_{j+1}|x_j), \tag{5}$$

where $g_{j+1}(x_{j+1}|x_j)$ represents the transition probability from state x_j to state x_{j+1}. Considering the failure components at stage j, (5) is equivalent to

$$g(z) = \prod_{j=0}^{n-1} \left[\prod_{k \in F_j} \varphi_k(x_j) \cdot \prod_{k \in \bar{F}_j} (1 - \varphi_k(x_j)) \right], \tag{6}$$

where F_j is the component set consisting of the components that are defective at x_{j+1} but operate normally at x_j, while $\bar{F}_j$ consists of the components that function normally at x_{j+1}.

Since maintenance only influences some of the components in the system, we consider the items related to a specific component $k \in K$ in (6) and define

$$\Gamma(\varphi_k, z) := \begin{cases} \displaystyle\prod_{j=0}^{n-1} (1 - \varphi_k(x_j)) & : \text{if} \quad n_k = n \\ \displaystyle\varphi_k(x_{n_k}) \prod_{j=0}^{n_k - 1} (1 - \varphi_k(x_j)) & : \text{otherwise} \end{cases} \tag{7}$$

where K is the set that includes all components; and n_k is the stage at which component k fails. In particular, if component k does not fail in the whole cascading process, let $n_k := n$. Then

$$g(z) = \prod_{k \in K} \Gamma(\varphi_k, z). \tag{8}$$

Substituting (8) into (3) yields

$$R_g(Y_0) = \sum_{z \in \mathcal{Z}} \left(h(z)\delta_{\{h(z)\geq Y_0\}} \cdot \prod_{k \in K} \Gamma(\varphi_k, z) \right). \tag{9}$$

The inherent relationship between $R_g(Y_0)$ and φ_k is indicated in (9). However, since $|\mathcal{Z}|$ and $|K|$ [1] may be very large, it cannot be directly used in optimization. Alternatively, we use an approximation of $R_g(Y_0)$ in terms of a set of samples, which is given by

$$\hat{R}_g(Y_0) = \frac{1}{N} \sum_{i=1}^{N} h(z^i)\delta_{\{h(z^i)\geq Y_0\}}. \tag{10}$$

In (10), N is the number of independent identically distributed (i.i.d.) samples generated using specific blackout simulation models with respect to $g(Z)$ (or more precisely, $\varphi_k, k \in K$); $z^i = \{x_0^i, \ldots, x_j^i, \ldots, x_{n^i}^i\}$ is the i-th sample, where n^i is the number of stages in the i-th sample. In particular, we define $Z_g := \{z^i, i = 1, \cdots, N\}$, which represents the sample set generated with respect to $g(Z)$.

When some of the φ_k values change due to maintenance, $g(Z)$ will be converted into another probability series that is denoted $f(Z)$. With regard to Z_g, Guo *et al.* [23] state that the unbiased estimation of $R_f(Y_0)$ is given by

$$\hat{R}_f(Y_0) = \frac{1}{N} \sum_{i=1}^{N} w(z^i)h(z^i)\delta_{\{h(z^i)\geq Y_0\}}, \tag{11}$$

[1] $|\cdot|$ represents the cardinality of a set.

where $w(z^i)$ is the sample weight of z^i and is defined as

$$w(z^i) := \frac{f(z^i)}{g(z^i)}, \quad (\forall z^i \in \mathcal{Z}). \tag{12}$$

According to (11), it is interesting to note that $\hat{R}_f(Y_0)$ only requires the information of z^i that is generated in the blackout simulations with respect to $g(Z)$. This implies that when $g(Z)$ changes to $f(Z)$, the RCB can be directly obtained using the information of Z_g, rather than by regenerating samples via blackout simulations. Therefore, this approach provides a means to explicitly express the RCB with respect to maintenance strategies, which will greatly facilitate the formulation of optimal maintenance decision-making problems.

C. FORMULATION OF MAINTENANCE STRATEGY OPTIMIZATION

Binary variable m_k represents the maintenance status of component k. If component k is maintained, $m_k = 1$; otherwise, $m_k = 0$. Vector $M := \{m_k, k \in K\}$ represents the maintenance strategy. For a specific sample z^i, we have

$$f(z^i) = \prod_{k \in K} \Big[m_k \Gamma(\bar{\varphi}_k, z^i) + (1 - m_k)\Gamma(\varphi_k, z^i) \Big]. \tag{13}$$

Substituting (13) into (11) yields

$$\begin{aligned}
&\hat{R}_f(Y_0) \\
&= \frac{1}{N} \sum_{i=1}^{N} \frac{\prod_{k \in K} \left[m_k \Gamma(\bar{\varphi}_k, z^i) + (1-m_k)\Gamma(\varphi_k, z^i) \right]}{\prod_{k \in K} \Gamma(\varphi_k, z^i)} h(z^i)\delta_{\{h(z^i) \geq Y_0\}} \\
&= \frac{1}{N} \sum_{i=1}^{N} \left[\prod_{k \in K^*} \left(1 + m_k \left(\frac{\Gamma(\bar{\varphi}_k, z^i)}{\Gamma(\varphi_k, z^i)} - 1 \right) \right) h(z^i)\delta_{\{h(z^i) \geq Y_0\}} \right],
\end{aligned} \tag{14}$$

where K^* is the set that includes all components available for maintenance. Thus, $K^* \subseteq K$ and $m_k = 0, \forall k \notin K^*$.

The relationship between the estimated RCB and maintenance strategies is detailed in (14), which is an explicit expression that provides an estimate of the RCB. To minimize the RCB with the limited number of components considered to be in maintenance, the following optimization problem can be formulated

$$\begin{aligned}
&\min_{m_k} \frac{1}{N} \sum_{i=1}^{N} \left[\prod_{k \in K^*} \left(1 + m_k \left(\frac{\Gamma(\bar{\varphi}_k, z^i)}{\Gamma(\varphi_k, z^i)} - 1 \right) \right) h(z^i)\delta_{\{h(z^i) \geq Y_0\}} \right] \\
&s.t. \sum_{k \in K^*} m_k \leq M_{max},
\end{aligned} \tag{15}$$

where $m_k, k \in K^*$ are the decision variables, M_{max} is a predefined parameter that stands for the maximum number of components considered in maintenance, and $\Gamma(\varphi, z^i)$, $\Gamma(\bar{\varphi}_k, z^i)$, $h(z^i)$, and $\delta_{\{h(z^i) \geq Y_0\}}$ are variables that rely on the samples. For simplicity, we define an N-dimension vector C and two $|K^*| \times N$ matrices, P and Q, as follows

$$\begin{aligned}
C_i &:= h(z^i)\delta_{\{h(z^i) \geq Y_0\}} \quad \forall i \\
P_{ki} &:= \Gamma(\varphi_k, z^i) \quad \forall i, k \\
Q_{ki} &:= \Gamma(\bar{\varphi}_k, z^i) \quad \forall i, k.
\end{aligned}$$

Remark 1: Note that in practice, cascading outages always occur in power systems that are heavily loaded. Therefore, we assume the initial state is deterministic when formulating cascading outages. Similarly, in practice, the proposed optimization model (15) *is employed in typical heavily loaded system states. On the other hand, it is worth noting that these formulations and the proposed methodology serve as the first step in related analyses. When multiple initial states are considered, the methodology can be separately employed in each initial state, and the final RCB can be calculated by the average of the RCBs in multiple initial states. The proof of this will not be provided here as it is beyond the scope of this paper.*

In large power systems that include thousands of components, (15) is a high-dimensional 0-1 programming problem, the solution for which involves the following two issues.

1) DETERMINING APPROPRIATE SAMPLE SIZE N

In the optimization model described by (15), the sample size N is a crucial parameter because the objective function is an unbiased estimation of the RCB and the estimation error relies on N. To guarantee the credibility of the estimation, a sufficiently large N is required. On the other hand, an N that is too large will drastically increase the computational burden. Unfortunately, it is not trivial to determine an appropriate sample size. In this regard, we analyzed the variance of the estimated RCB based on inferences from the simulation data, which enabled us to develop an adaptive method to determine a suitable sample size that achieves a good trade-off between the estimation accuracy and computational complexity. The details of this method are provided in Sections III and IV-B.

2) REDUCING COMPUTATIONAL COMPLEXITY

Even after an appropriate sample size N has been determined, it is still challenging to solve the optimization problem in (15). The number of decision variables in (15) equals $|K^*|$, and there are $2^{|K^*|} - 1$ product terms of decision variables in the objective function. Therefore, the optimization problem in (15) for a real large system may have thousands of decision variables and an astronomical number of product terms. For this type of high-dimensional 0-1 integer programming problem, enumeration and conventional optimization algorithms, such as branch and bound, are demonstrably ineffective. To address these limitations, we devised two simple yet efficient algorithms to identify nearly-optimal solutions of (15) in a heuristic manner, which is a suitable approach in practice for large systems. The details of these algorithms are provided in Section IV.

III. DETERMINING N BASED ON CREDIBILITY ANALYSES

Due to the inherent uncertainty of the samples in Z_g, the estimated RCB, $\hat{R}_f(Y_0)$, given in (11) always includes a certain amount of estimation error. In general, increasing the sample size reduces the estimation error. However, to determine an appropriate sample size of Z_g, i.e., N, the influence of N on the estimation error must be characterized via a variance analysis based on inference from the simulation data.

A. VARIANCE-BASED CREDIBILITY ANALYSES

We first consider the relative error bound of (11) with respect to a specific Z_g. Specifically, we denote ϵ as the relative error bound of $\hat{R}_f(Y_0)$ with a confidence level β. Then, ϵ should bound the probability of $R_f(Y_0)$ within $I := [\hat{R}_f(Y_0)/(1+\epsilon), \hat{R}_f(Y_0)/(1-\epsilon)]$ with a probability β, i.e.

$$\mathbf{Pr}\left(\frac{|R_f(Y_0)-\hat{R}_f(Y_0)|}{R_f(Y_0)} \leq \epsilon\right) = \beta, \tag{16}$$

where $2\epsilon/(1-\epsilon^2)$ is defined as the relative length of the error bound used to quantify the credibility of the estimated RCB. Note that (16) is equivalent to

$$\mathbf{Pr}\left(-\frac{\epsilon R_f(Y_0)}{\sqrt{D_f(Y_0)}} \leq \frac{R_f(Y_0)-\hat{R}_f(Y_0)}{\sqrt{D_f(Y_0)}} \leq \frac{\epsilon R_f(Y_0)}{\sqrt{D_f(Y_0)}}\right) = \beta, \tag{17}$$

where $D_f(Y_0)$ is the estimation variance of $\hat{R}_f(Y_0)$ given by (11). Invoking the Central Limit Theorem, when N is sufficiently large, we have:

$$\frac{R_f(Y_0)-\hat{R}_f(Y_0)}{\sqrt{D_f(Y_0)}} \sim \mathbf{Norm}(0, 1).$$

In this context, ϵ can be determined by solving (17), yielding

$$\epsilon = \frac{\Phi^{-1}(1/2+\beta/2)\sqrt{D_f(Y_0)}}{R_f(Y_0)}, \tag{18}$$

where Φ is the cumulative density function of the standard normal distribution, and Φ^{-1} is the inverse function of Φ.

According to (18), a small enough $D_f(Y_0)$ is required to ensure the credibility of the blackout risk estimation. However, as the distribution of $\hat{R}_f(Y_0)$ is unknown, $D_f(Y_0)$ cannot be calculated accurately in practice. To address this, we propose a proposition for estimating $D_f(Y_0)$ based on a specific Z_g.

Proposition 1: Given Z_g, $g(Z)$, and $f(Z)$, an unbiased estimation of $D_f(Y_0)$ is given by

$$\hat{D}_f(Y_0) = \frac{1}{(N-1)N}\sum_{i=1}^{N}\left[w(z^i)h(z^i)\delta_{\{h(z^i)\geq Y_0\}} - \hat{R}_f(Y_0)\right]^2. \tag{19}$$

Proof: We first define random variable $L_i(Y_0)$ as $L_i(Y_0) := \frac{f(z^i)}{g(z^i)}h(z^i)\delta_{\{h(z^i)\geq Y_0\}}$, $i = 1, \cdots, N$. Since z^i is i.i.d., $L_i(Y_0)$ is also i.i.d. Therefore, the variance of random variables $L_i(Y_0)$ are the same. Specifically, we denote the variance of L_i by $d_f(Y_0)$. We then have

$$D_f(Y_0) = \mathbb{D}\left(\frac{1}{N}\sum_{i=1}^{N}L_i(Y_0)\right) = \frac{1}{N}\mathbb{D}(L_i(Y_0)) = \frac{1}{N}d_f(Y_0), \tag{20}$$

where $\mathbb{D}$ is an operator of the variance calculation. The second equation holds by noting that $L_i(Y_0)$, $i = 1, \cdots, N$ are i.i.d.

According to the definition of the sample variance in [28], the unbiased estimation of $d_f(Y_0)$ based on Z_g is given by

$$\hat{d}_f(Y_0) = \frac{1}{N-1}\sum_{i=1}^{N}\left[w(z^i)h(z^i)\delta_{\{h(z^i)\geq Y_0\}} - \hat{R}_f(Y_0)\right]^2. \tag{21}$$

Therefore, by substituting (21) into (20), the unbiased estimation of $D_f(Y_0)$ can be determined using (19). ■

Combining (11), (18), and (19), ϵ can be estimated by

$$\hat{\epsilon} = \frac{\Phi^{-1}(1/2+\beta/2)\sqrt{\hat{D}_f(Y_0)}}{\hat{R}_f(Y_0)}. \tag{22}$$

On the other hand, according to (18) and (20), we have

$$\epsilon \propto \frac{\sqrt{d_f(Y_0)}}{\sqrt{N}R_f(Y_0)}. \tag{23}$$

Since $d_f(Y_0)$ is deterministic for specific Z_g and $f(Z)$, when the error bound of the estimated RCB in (11) does not satisfy the predefined requirement, the sample size N should be enlarged and Z_g should be updated, as explained next.

B. DETERMINING THE SAMPLE SIZE N

In this subsection, the previous credibility analyses are employed to determine an appropriate N based on inference from the simulation data in Z_g. In essence, if the estimation error bound is determined (according to the requirement of practical utilization), i.e.,$\bar{\epsilon} = 10\%$, then (22) gives

$$\bar{D}_f(Y_0) = \left(\frac{\bar{\epsilon}\hat{R}_f(Y_0)}{\Phi^{-1}(1/2+\beta/2)}\right)^2. \tag{24}$$

In (24), $\bar{D}_f(Y_0)$ represents the maximal estimation variance required so that the relative error is within a given error bound $\bar{\epsilon}$ with a confidence level of β.

To ensure that $D_f(Y_0)$ is less than $\bar{D}_f(Y_0)$, the required sample size $\bar{N}$ can be obtained using (20) and (21) as

$$\bar{N} = \frac{\hat{d}_f(Y_0)}{\bar{D}_f(Y_0)} = \frac{\hat{d}_f(Y_0)}{\hat{R}_f^2(Y_0)}\left(\frac{\Phi^{-1}(1/2+\beta/2)}{\bar{\epsilon}}\right)^2. \tag{25}$$

According to (25), when $N \leq \bar{N}$, the sample size is not large enough to guarantee the credibility of the blackout risk estimation and more samples are required to reduce the estimation variance. This provides a systematic approach to the optimization problem (15) when determining an appropriate sample size. The corresponding algorithm can be easily implemented in practice, as described in Section IV-B.

It is noteworthy that both $\hat{d}_f(Y_0)$ and $\hat{R}_f(Y_0)$ in (25) are calculated based only on the sample set Z_g. Hence, $\bar{N}$ is essentially inferred from the simulation data in Z_g and there is no need to regenerate any additional samples as long as the sample size condition $N > \bar{N}$ holds.

IV. SOLUTION ALGORITHMS

A. HEURISTIC ALGORITHMS

As mentioned previously, the optimization model (15) is a high-dimensional nonlinear 0-1 programming problem that is considered to be NP-hard. Many nearly-optimal solutions are able to mitigate the RCB in practice, as will be demonstrated in the Case Studies. We now devise two heuristic algorithms to identify nearly-optimal solutions to (15) on a fixed Z_g.

Since one significant obstacle is the large number of decision variables, a simplistic approach is to reduce the number of components considered for maintenance, i.e., $|K^*|$. With this in mind, we employ a sensitivity-based approach in the design of the following algorithm.

Algorithm 1

- ***Step 1**: Sensitivity analysis*. Construct $|K^*|$ scenarios in which the φ_k of a single component $k \in K$ changes to $\bar{\varphi}_k$. Then, estimate the RCB using (11).
- ***Step 2**: Scenario reduction*. Choose the M_k components with larger reductions in the RCB to make up the K_m. Here, M_k is a predefined parameter based on M_{max} and the experience. In particular, $M_{max} < M_k = |K_m| < |K^*|$.
- ***Step 3**: Solving the optimization problem*. Substitute K^* in (15) with K_m and solve (15) by enumeration.

Since (11) only requires a few algebraic calculations, it is efficient to estimate the RCB in each scenario. Suppose the average calculation time of a single scenario is t_s, then the total calculation time is $(|K^*| + C(M_k, M_{max}))t_s$.Here,$C(M_k, M_{max})$ is the number of $M_{max}-$combinations from M_k elements. As the maximal number of components that are considered for maintenance M_{max} is often small in practice, M_k can be chosen to be a moderate number. Then, the number of scenarios and the computation time by enumeration will be acceptable.

Nevertheless, as M_k and M_{max} grow larger, such as in an actual large power system, the number of scenarios $(|K^*|+C(M_k, M_{max}))$ and computation time increase dramatically. To address this, we propose the following additional algorithm.

Algorithm 2

- ***Step 1**: Initialization*. Let $K_m = \emptyset$. Then its complementary set, which is denoted by $\bar{K}_m$, equals K^*.
- ***Step 2**: Estimation of RCB*. Construct $|\bar{K}_m|$ scenarios. In each scenario, for all components in K_m and a single component in $\bar{K}_m$, change their φ_k to $\bar{\varphi}_k$ and estimate the RCB with (11).
- ***Step 3**: Iteration*. Choose the scenario with the lowest RCB. Move the corresponding component from $\bar{K}_m$ to K_m. If $|K_m| = M_{max}$, the procedure ends; otherwise, go back to Step 2.

In Algorithm II, the components that must be maintained are determined successively. In each round, the component

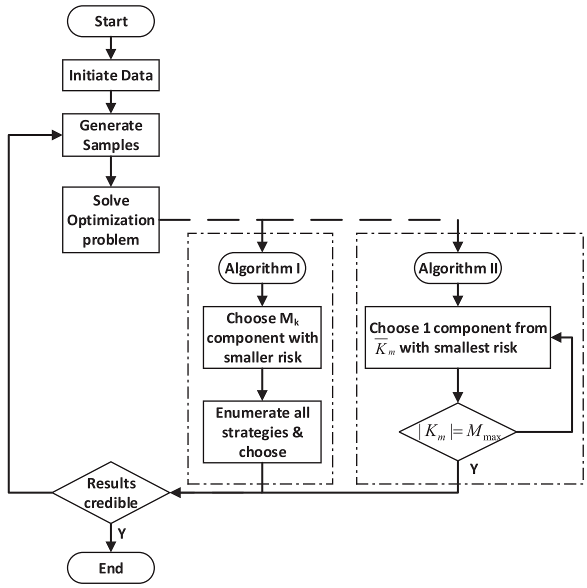


FIGURE 1. Flowchart of the methodology.

that can most effectively reduce the RCB is selected. Therefore, Algorithm II actually belongs to a family of so called greedy algorithms. The total number of scenarios is $(2|K^*| - M_{max} + 1)M_{max}/2$. Since the number of scenarios increases linearly with $|K^*|$, the optimization process remains efficient even in a large system.

B. PROCEDURE

The methodology (see Fig. 1) can be summarized as follows.

Procedure

- **Step 1: Initialization**. Initialize the system. In particular, determine φ_k and $\bar{\varphi}_k$ for each component in the system. The initial sample size is $N = N_0$.
- **Step 2: Sample generation.** Generate N samples based on the specific blackout model and $\varphi_k, k \in K$. The sample set is Z_g. Based on Z_g, φ_k, and $\bar{\varphi}_k$, calculate matrices C, P, and Q.
- **Step 3: Maintenance optimization strategy.** Solve the optimization problem (15) using Algorithm I or II.
- **Step 4: Credibility evaluation.** For the optimal maintenance strategy, calculate the necessary sample size $\bar{N}$ with (25). If $\bar{N} \geq N$, generate another $\bar{N} - N$ samples and add them into Z_g. Then, update C, P, Q and go back to Step 3; Otherwise, directly choose the optimal one according to the results in Step 3.

The proposed procedure incorporates an adaptive selection of the sample size based on a credibility evaluation, as shown in Step 4. In this way, the sample size can be minimized while the estimation error is within a predefined limit with respect to a given confidence level. It is noteworthy that to benefit from the inference from blackout simulation data, all samples are generated with respect to $g(Z)$ only and not for different $f(Z)$. This results in a significant reduction in the sampling time and simplifies the implementation.

V. CASE STUDIES

A. SETTINGS

In this section, numerical experiments are carried out on a IEEE 57-bus system, IEEE 300-bus system, and an actual provincial power system in China with a simplified ORNL-PSerc-Alaska (OPA) model that omits slow dynamics [13]. In the simulations, a traditional MC method is employed to consider the random failures of both the transmission lines and power transformers. The failure probability functions of the transmission lines and power transformers are the same as those in [14] and [24], respectively. In all cases in this work, typical parameters are applied, and for simplicity, only the maintenance of power transformers is considered.

The simulation model with specific component failure probability functions mentioned above is only employed to demonstrate the proposed method, and it is expected that more realistic models and settings will be adopted in actual situations.

B. IEEE 57-BUS SYSTEM

In this simulation, we employed a small system that included 53 transmission lines and 17 power transformers to demonstrate the difficulties encountered in mitigating the RCB via component maintenance and to highlight some salient features of the proposed method.

1) RCB ESTIMATION

Here, we demonstrate the unbiasedness of (11) and the inherent estimation error caused by the randomness of the samples. First, we constructed a series of Z_g with increasing sample sizes and randomly changed the φ_k of four components. Then, we calculated the RCB with $Y_0 = 200$ using (11). Next, we estimated ϵ with respect to $\beta = 90\%$ using (22) and calculated the corresponding error bounds. For comparison, we directly regenerated the samples with respect to the new $\bar{\varphi}_k$ and estimated the RCB using (9). The results are presented in Fig. 2.

Fig. 2 shows that the estimation errors between the calculations and direct sampling become smaller as the sample size increases until they are almost the same when the sample size is large enough. This indicates the effectiveness of the RCB estimation given by (11). This can be illustrated more rigorously by noting the relative length of the error bounds (see Fig. 3) also becomes smaller as the sample size increases.

At the same time, it is worth noting that Fig. 3 also indicates that some amount of estimation error always exists. More importantly, when the sample size is small, the estimation error can be very large. Intuitively, optimized maintenance strategies based on a small Z_g may have a large deviation in their performance of mitigating the RCB. Therefore, considering the calculation efficiency and possible estimation error, it is important that the sample size of Z_g, i.e., N, should be selected appropriately.

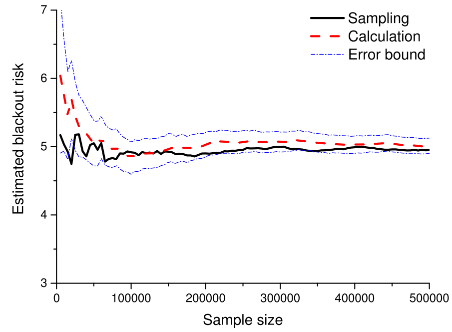


FIGURE 2. RCB estimation with sampling and calculation.

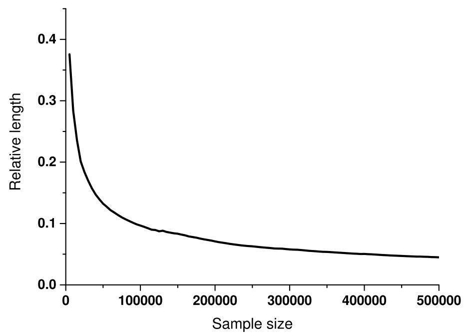


FIGURE 3. Relative length of the error bounds.

TABLE 1. RCB after component maintenance.

Component index	$\hat{R}_f(0)$	Reduction ratio(%)	Component index	$\hat{R}_f(200)$	Reduction ratio(%)
7	6.55	2.8	7	4.86	5.4
6	6.64	1.5	6	5.00	2.7
2	6.66	1.1	3	5.02	2.2
3	6.67	1.0	2	5.02	2.2
5	6.67	1.0	15	5.03	2.1
...	...	...	...	...	...
Mean	6.68	0.9	Mean	5.02	2.2

2) INFLUENCE OF MAINTENANCE ON THE RCB

We constructed a Z_g that included 100,000 samples, based on which we considered the simultaneous maintenance of components. Since there were only 17 transformers in this small system, we directly enumerated all the possible strategies and compared the respective risk reduction performance. We first considered the maintenance of a single component. The RCBs with respect to $Y_0 = 0$ and $Y_0 = 200$ after maintenance are given in Table 1. In particular, the original RCBs were $\hat{R}_g(0) = 6.74$ and $\hat{R}_g(200) = 5.14$.

The results in Table 1 intuitively indicate that component maintenance is an effective way to mitigate the RCB. Moreover, a few critical components in the system have much greater influence on the RCB than the others. For example,

TABLE 2. Optimal maintenance strategies ($Y_0 = 0$).

Component index	RCB	Reduction ratio(%)	Sum of individual reduction ratio(%)
(7,6,2,5)	6.505	3.5	6.4
(7,6,5,15)	6.506	3.5	6.2
(7,6,2,15)	6.507	3.4	6.4
(7,6,3,5)	6.509	3.4	6.3
(7,6,5,14)	6.511	3.4	6.2
...	...	...	...
Mean	6.643	1.4	3.6

TABLE 3. Optimal maintenance strategies ($Y_0 = 200$).

Component index	RCB	Reduction ratio(%)	Sum of individual reduction ratio(%)
(7,6,5,15)	4.819	6.2	12.3
(7,6,2,5)	4.822	6.2	12.3
(7,6,3,15)	4.823	6.1	12.4
(7,6,3,5)	4.823	6.1	12.3
(7,6,2,15)	4.823	6.1	12.4
...	...	...	...
Mean	4.991	2.9	3.6

the maintenance of No. 7 transformer can result in a 2.8% reduction in the RCB with respect to $Y_0 = 0$, while the average reduction is only 0.9%. This result confirms the efficacy of optimizing the maintenance strategy in mitigating the RCB.

Next, we considered maintaining four transformers simultaneously and then estimated the RCBs with all possible maintenance strategies. The strategies with the largest reductions in the RCBs with respect to the two Y_0 are presented in Tables 2 and 3, respectively. The results illustrate the complicated relationship between maintenance strategies (or more precisely, component failure probabilities) and the RCB. Note that the degree of risk reduction associated with the four components are much smaller than the sum of the four individual components. Moreover, the optimal components to be maintained (see Table 2 and Table 3) are not simply combinations of components with smaller RCBs (see Table 1). This indicates that the influence of component failure probabilities on the RCB is essentially nonlinear, which implies that one cannot quantify the influence of maintaining multiple components solely based on the influence of the individual components.

Moreover, Tables 2 and 3 show that there are many nearly-optimal maintenance strategies whose risk reduction ratios are very close to the optimal one, even though the resulting components to be maintained are not identical. Considering the possible estimation error of the blackout risk estimation, it is reasonable to also deploy these nearly-optimal maintenance strategies. In other words, there is no need to rigorously solve the optimization problem (15), particularly when finding a solution is very time consuming. With this in mind, we employed two heuristic algorithms to efficiently find (nearly-)optimal solutions in this work.

TABLE 4. Optimal maintenance strategies with different algorithms ($Y_0 = 200$).

Method	Strategy	Risk reduction (%)	Number of Scenarios
Alg. I ($M_k = 8$)	(7,6,2,3)	2.2	70
Alg. I ($M_k = 12$)	(7,6,2,3)	2.2	495
Alg. II	(7,6,2,3)	2.2	62
Enumeration	(7,6,2,3)	2.2	2380

TABLE 5. Optimal maintenance strategies with different algorithms ($Y_0 = 100$).

Method	Strategy	Risk reduction (%)	Number of Scenarios
Alg. I ($M_k = 8$)	(6,2,5,14)	1.0	70
Alg. I ($M_k = 12$)	(6,2,3,7)	1.4	495
Alg. II	(6,2,5,14)	1.0	62
Enumeration	(6,2,3,7)	1.4	2380

3) HEURISTIC ALGORITHMS

To compare the two heuristic algorithms, we selected different parameters, $|Z_g|$, Y_0, M_{max}, and M_k, and compared the corresponding optimal maintenance strategies. Here, we present two typical cases. Specifically, $|Z_g| = 35000$ and $M_{max} = 4$. We employed Algorithm I with $M_k = 8$, 12 and Algorithm II to determine the optimal maintenance strategies with respect to $R_f(100)$ and $R_f(200)$, respectively, and the results are shown in Tables 4 and 5.

As shown in Table 4, both algorithms produced the same maintenance strategy and achieved the global optimum,[2] while in Table 5, Algorithm I with $M_k = 8$ and Algorithm II provided sub-optimal solutions. The reason is due to the inherent nonlinearity between the component failure probabilities and the RCB, as well as the inevitable uncertainty in the samples. However, it is worth noting that, based on our experience, Algorithm I with a moderate M_k and Algorithm II produce effective maintenance strategies in most cases, and are therefore suitable for large systems.

C. IEEE 300-BUS SYSTEM

In this case, the complete proposed method is tested using data from the IEEE 300-bus system with 304 lines and 107 transformers. Specifically, we set $\epsilon = 10\%$, $\beta = 95\%$, $M_{max} = 8$, $M_k = 30$, $N_0 = 5000$, $Y_0 = 100$.

1) SOLVING PROCESS

First, the solving process with Algorithm I is presented in Table 6.

As shown in Table 6, we began with a Z_g that included 5,000 samples, with which we obtained an optimal maintenance strategy. However, the relative error requirement ϵ could not be satisfied. Therefore, 8, 0000 samples were added

[2]Note that since $|Z_g|$ is different in Tables 3 and 4, the largest reductions in the RCB obtained by enumeration are different, which indicates that determining the sample size based on credibility analyses has practical significance.

TABLE 6. Solving process with Algorithm I.

Step	Sample size	Strategy	RCB	$\hat{\epsilon}(\%)$
1	5000	(6,17,46,53,68,75,88,106)	2.272	41.9
2	85000	(10,25,29,66,68,97,100,106)	2.007	10.3
3	95000	(10,25,29,66,68,77,100,106)	2.011	9.7

TABLE 7. Solving process with Algorithm II.

Step	Sample size	Strategy	RCB	$\hat{\epsilon}(\%)$
1	5000	(46,106,68,53,88,75,6,17)	2.272	41.9
2	85000	(106,25,68,66,100,29,10,97)	2.007	10.3
3	95000	(106,25,68,29,66,100,77,10)	2.011	9.7

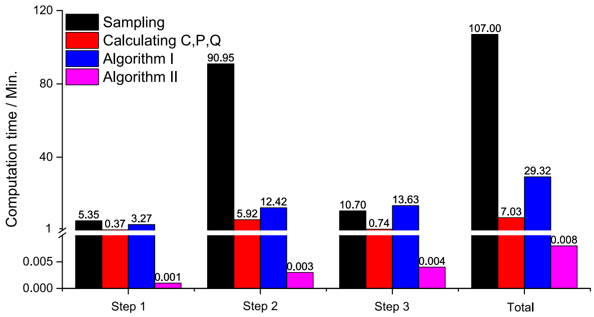


FIGURE 4. Computation time with two algorithms.

into Z_g according to (25) in step 2. Then, the optimal maintenance strategy and $\hat{\epsilon}$ were recalculated. This process was repeated in step 3 to determine the final optimal maintenance strategy. The level of reduction in the RCB with respect to $Y_0 = 100$ was 21.5%.

The solution process with Algorithm II is summarized in Table 7. In this case, the optimal strategies in each step are actually the same as the ones of Algorithm I. Note that the components of each strategy in Table 7 are ordered according to the decision process. Even though the results provided by the two algorithms are similar in steps 2 and 3, the decision processes are different. This observation again demonstrates the complicated relationship between the component failure probabilities and the RCB.

2) COMPUTATION TIME

We conducted all of the tests on a computer with an Intel Xeon E5-2670 processor running at 2.6 GHz and 64 GB of memory. The computation times of the two algorithms are presented in Fig. 4. Note that the computation time for sampling and calculating C, P, Q in each step depends on the number of additional samples, while the optimization time depends on the number of total samples. From the figure, it can be seen that the computation time for sampling was much larger than that for the other times. In our cases, 107 min were required to complete sampling. The requirement for repeated sampling is why traditional methods are extremely inefficient at directly estimating the RCB for all maintenance strategies. In contrast, our methodology enables a very efficient estimation of the RCB under varying maintenance strategies without requiring the regeneration of any samples. In addition, the optimization time of Algorithm II is smaller than that of Algorithm I as it involves fewer maintenance scenarios. However, as Algorithm II may result in sub-optimal solutions in some cases, e.g., Table 4, it is preferable for large-scale systems with many candidate components considered for maintenance.

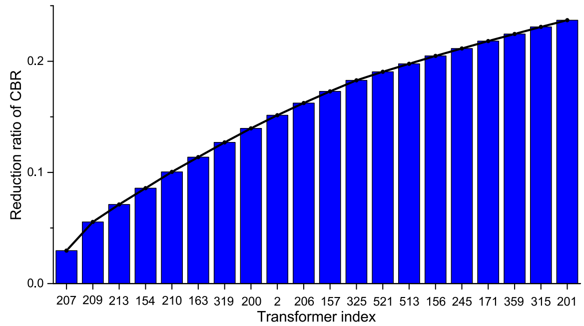


FIGURE 5. Optimal maintenance strategy in the provincial system.

D. AN ACTUAL PROVINCIAL SYSTEM

In this case, an actual provincial power system in China is employed to show the advantages of the proposed method. In particular, this system includes 1,122 buses, 1,230 transmission lines, and 562 transformers. Considering the large scale of the system, the chosen parameters were $Y_0 = 1500$ and $M_{max} = 20$. Under these conditions, even if a 'small' M_k is chosen in Algorithm I, the number of possible maintenance scenarios is astronomical, e.g., if $M_k = 40$, the number of scenarios is greater than 100 billion. Therefore, we only employed Algorithm II when determining the maintenance strategy.

Algorithm II realizes the final optimal maintenance strategy shown in Fig. 5. Specifically, the transformers on the horizontal axis are ordered by the decision process in Algorithm II, and the reduction ratio of the RCB after each transformer is maintained is shown on the vertical axis. Since the most effective transformer is maintained in each step of Algorithm II, the increase in the reduction ratio of the RCB becomes slower along with the decision process. The final reduction ratio is 23.1%. The computational time required for the entire procedure was 3,134 min, and Fig. 6 shows the specific proportions of different parts. From the figure, it can be seen that the sampling required almost all of the time in this large system. In contrast, Algorithm II is extremely efficient. This not only verifies the applicability of the proposed algorithm in actual large systems, but also demonstrates that high-efficiency sampling methods [17], [20] can be employed to further improve the scalability and efficiency of the methodology.

E. SAMPLING TIME

In this case, we further compare the computational time taken for sampling when using the traditional MC method. In three

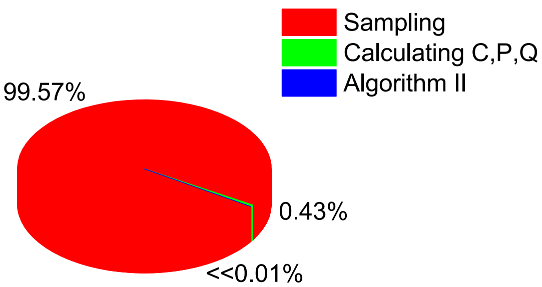


FIGURE 6. Computation time in the provincial system.

TABLE 8. Average time to generate one sample.

System	Bus number	Line number	Transformer number	time (s)
IEEE 57-bus	57	53	17	0.03
IEEE 300-bus	300	304	107	0.06
Actual provincial	1122	1230	562	8.96

test systems, the average times to generate *one* sample are listed in Table 8.

According to Table 8, as the scale of the system grows up, the average time to generate one sample increases obviously. The main reason can be attributed to two aspects:

1) As the system scale increases, the calculation of individual power flows turns to be much more time-consuming.
2) As the number of components in the system increases, the average path length of individual cascading outages becomes longer. That further leads to higher intensity of computation.

Additionally, noting that both the state space of cascading outages and the number of possible maintenance strategies are fairly large, the sample size required for cascading outage simulations considering impacts of maintenance is inevitably huge. It indicates that, in a maintenance optimization problem, the total compuational burden for generating samples are very heavy, particularly when a large scale system is considered. In such a circumstance, traditional methods relying on conducting additional simulations appear to be intractable.

VI. CONCLUSION WITH REMARKS

In this paper, we propose an efficient methodology to optimize component maintenance strategies for effectively mitigating the RCB. In the proposed method, an analytic relationship between the estimated RCB and maintenance strategies is developed through inference from simulation data. The results of theoretical analyses and numerical experiments confirm that:

1) A small set of components in power systems exert great influence on the propagation of cascading outages, and reducing their failure probabilities through component maintenance can significantly mitigate the RCB.
2) Variance-based analyses are employed to further elucidate the credibility of the estimated RCB while considering different maintenance strategies.
3) The proposed heuristic algorithms are found to efficiently optimize the component maintenance strategies in large power systems.

The proposed method also can be applied in other fileds. For example, it can be employed to identify critical buses or branches in complex networks. Moreover, note that as the scale of the system increases, the sampling process will take more time. We intend to incorporate some recently developed high-efficiency sampling methods, such as Sequential Importance Sampling [17] and Splitting [20], to further improve the performance of this method in future work. Another area of ongoing work is to include a more practical formulation of maintenance strategy optimization and the corresponding solving algorithms.

REFERENCES

[1] B. Liscouski and W. Elliot, "Final report on the August 14, 2003 blackout in the United States and canada: Causes and recommendations," U.S. Dept. Energy, 2004, vol. 40, p. 4.
[2] G. A. Maas, M. Bial, and J. Fijalkowski, "Final report; system disturbance on November 2006," Union Coordination Transmiss. Electr. Eur., 2007.
[3] B. A. Carreras, D. E. Newman, I. Dobson, and A. B. Poole, "Evidence for self-organized criticality in a time series of electric power system blackouts," *IEEE Trans. Circuits Syst. I, Reg. Papers*, vol. 51, no. 9, pp. 1733–1740, Sep. 2004.
[4] M. Vaiman *et al.*, "Risk assessment of cascading outages: Methodologies and challenges," *IEEE Trans. Power Syst.*, vol. 27, no. 2, pp. 631–641, May 2012.
[5] A. Abiri-Jahromi, M. Fotuhi-Firuzabad, and M. Parvania, "Optimized midterm preventive maintenance outage scheduling of thermal generating units," *IEEE Trans. Power Syst.*, vol. 27, no. 3, pp. 1354–1365, Aug. 2012.
[6] A. D. Janjic and D. S. Popovic, "Selective maintenance schedule of distribution networks based on risk management approach," *IEEE Trans. Power Syst.*, vol. 22, no. 2, pp. 597–604, May 2007.
[7] M. J. Eppstein and P. D. H. Hines, "A 'random chemistry' algorithm for identifying collections of multiple contingencies that initiate cascading failure," *IEEE Trans. Power Syst.*, vol. 27, no. 3, pp. 1698–1705, Aug. 2012.
[8] P. D. H. Hines, I. Dobson, and P. Rezaei, "Cascading power outages propagate locally in an influence graph that is not the actual grid topology," *IEEE Trans. Power Syst.*, vol. 32, no. 2, pp. 958–967, Mar. 2016.
[9] X. Zhang, F. Liu, and R. Yao, "Identification of key transmission lines in power grid using modified K-core decomposition," in *Proc. 3rd Int. Conf. IEEE Electr. Power Energy Convers. Syst. (EPECS)*, Oct. 2013, pp. 1–6.
[10] Y. Jiang, J. D. McCalley, and T. V. Voorhis, "Risk-based resource optimization for transmission system maintenance," *IEEE Trans. Power Syst.*, vol. 21, no. 3, pp. 1191–1200, Aug. 2006.
[11] S. Natti and M. Kezunovic, "A risk-based decision approach for maintenance scheduling strategies for transmission system equipment," in *Proc. 10th Int. Conf. IEEE Probabilistic Methods Appl. Power Syst. (PMAPS)*, May 2008, pp. 1–6.
[12] S. R. K. Yeddanapudi, Y. Li, J. D. McCalley, A. A. Chowdhury, and W. T. Jewell, "Risk-based allocation of distribution system maintenance resources," *IEEE Trans. Power Syst.*, vol. 23, no. 2, pp. 287–295, May 2008.
[13] B. A. Carreras, V. E. Lynch, I. Dobson, and D. E. Newman, "Critical points and transitions in an electric power transmission model for cascading failure blackouts," *Chaos*, vol. 12, no. 4, pp. 985–994, Dec. 2002.
[14] J. Chen, J. S. Thorp, and I. Dobson, "Cascading dynamics and mitigation assessment in power system disturbances via a hidden failure model," *Int. J. Elect. Power Energy Syst.*, vol. 27, no. 4, pp. 318–326, May 2005.
[15] R. Yao, S. Huang, K. Sun, F. Liu, X. Zhang, and S. Mei, "A multi-timescale quasi-dynamic model for simulation of cascading outages," *IEEE Trans. Power Syst.*, vol. 31, no. 4, pp. 3189–3201, Jul. 2016.

[16] I. Dobson, B. A. Carreras, and D. E. Newman, "How many occurrences of rare blackout events are needed to estimate event probability?" *IEEE Trans. Power Syst.*, vol. 28, no. 3, pp. 3509–3510, Aug. 2013.

[17] J. Guo, F. Liu, and J. Wang, "Toward efficient cascading outage simulation and probability analysis in power systems," *IEEE Trans. Power Syst.*, vol. 33, no. 3, pp. 2370–2382, May 2017.

[18] J. S. Thorp, A. G. Phadke, S. H. Horowitz, and S. Tamronglak, "Anatomy of power system disturbances: Importance sampling," *Int. J. Electr. Power*, vol. 20, no. 2, pp. 147–152, 1998.

[19] P. Henneaux and P.-E. Labeau, "Improving monte carlo simulation efficiency of level-I blackout probabilistic risk assessment," in *Proc. Int. Conf. IEEE Probabilistic Methods Appl. Power Syst. (PMAPS)*, Jul. 2014, pp. 1–6.

[20] J. Kim, J. A. Bucklew, and I. Dobson, "Splitting method for speedy simulation of cascading blackouts," *IEEE Trans. Power Syst.*, vol. 28, no. 3, pp. 3010–3017, Aug. 2013.

[21] S. P. Wang, A. Chen, C. W. Liu, C. H. Chen, J. Shortle, and J. Y. Wu, "Efficient splitting simulation for blackout analysis," *IEEE Trans. Power Syst.*, vol. 30, no. 4, pp. 1775–1783, Jul. 2015.

[22] S.-P. Wang, A. Chen, C.-W. Liu, C.-H. Chen, and J. Shortle, "Rare-event splitting simulation for analysis of power system blackouts," in *Proc. IEEE Power Energy Soc. General Meeting*, Jul. 2011, pp. 1–7.

[23] J. Guo, F. Liu, and J. Wang. (Nov. 2017). "Quantifying the influence of component failure probability on cascading blackout risk." [Online]. Available: https://arxiv.org/abs/1711.02580

[24] S. K. E. Awadallah, J. V. Milanovic, P. N. Jarman, and Z. Wang, "Probabilistic indicators for assessing age- and loading-based criticality of transformers to cascading failure events," *IEEE Trans. Power Syst.*, vol. 29, no. 5, pp. 2558–2566, Sep. 2014.

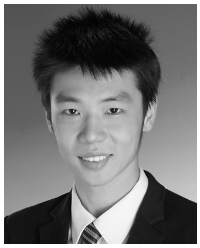
[25] D. Zhang, W. Li, and X. Xiong, "Overhead line preventive maintenance strategy based on condition monitoring and system reliability assessment," *IEEE Trans. Power Syst.*, vol. 29, no. 4, pp. 1839–1846, Jul. 2014.

[26] L. Zhou, Y. Wang, and Y. Li, "A maintenance decision optimization method based on life cycle cost of converter transformer," in *Proc. IEEE Elect. Insulation Conf. (EIC)*, Jun. 2016, pp. 288–291.

[27] D. Hughes, G. Dennis, and J. Walker, "Condition based risk management(CBRM)-enabling asset condition information to becentral to corporate decision making," in *Proc. 1st World Congr. Eng. Asset Manage.*, Gold Coast, QLD, Australia, 2006, pp. 1212–1217.

[28] G. Casella and R. L. Berger, *Statistical Inference*, vol. 2. Pacific Grove, CA, USA: Duxbury, 2002.

FENG LIU (M'12) received the B.Sc. and Ph.D. degrees in electrical engineering from Tsinghua University, Beijing, China, in 1999 and 2004, respectively. From 2015 to 2016, he was a Visiting Associate with the California Institute of Technology, CA, USA. He is currently an Associate Professor with Tsinghua University. He has authored/co-authored over 100 peer-reviewed technical papers and two books, and holds over 20 issued/pending patents. His research interests include power system stability analysis, optimal control and robust dispatch, game theory and learning theory, and their applications to smart grids. He was a Guest Editor of the IEEE TRANSACTIONS ON ENERGY CONVERSION.



JINPENG GUO (S'17) received the B.Sc. degree in electrical engineering from Tsinghua University, Beijing, China, in 2013, where he is currently pursuing the Ph.D. degree in electrical engineering. His research interests include the modeling and simulation of blackouts, risk assessment, and management.

XUEMIN ZHANG (M'06) received the B.S. and Ph.D. degrees in electrical engineering from Tsinghua University, Beijing, China, in 2001 and 2006, respectively. She is currently an Associate Professor with the Department of Electrical Engineering, Tsinghua University. Her research interests include power system analysis and control.

YUNHE HOU (M'08–SM'15) received the B.E. and Ph.D. degrees from the Huazhong University of Science and Technology, Wuhan, China, in 1999 and 2005, respectively. He was a Post-Doctoral Research Fellow at Tsinghua University, Beijing, China, from 2005 to 2007. He was a Visiting Scholar with Iowa State University, Ames, IA, USA, and a Research Scientist with University College Dublin, Dublin, Ireland, from 2008 to 2009. He was also a Visiting Scientist with the Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, Cambridge, MA, USA, in 2010. He is currently with The University of Hong Kong as a Research Assistant Professor. His research interests include electricity market theory, power system resource planning, and self-healing smart grid construction.

SHENGWEI MEI (SM'06–F'15) received the B.Sc. degree in mathematics from Xinjiang University, Ürümqi, China, the M.Sc. degree in operations research from Tsinghua University, Beijing, China, and the Ph.D. degree in automatic control from the Chinese Academy of Sciences, Beijing, in 1984, 1989, and 1996, respectively. He is currently a Professor with Tsinghua University. His research interests include power system analysis and control, game theory, and its application in power systems.