

Markov Decision Process-based Resilience Enhancement for Distribution Systems: An Approximate Dynamic Programming Approach

Chong Wang, *Member, IEEE*, Ping Ju, *Senior Member, IEEE*, Shunbo Lei, *Member, IEEE*,
Zhaoyu Wang, *Member, IEEE*, Yunhe Hou, *Senior Member, IEEE*

Abstract—Because failures in distribution systems caused by extreme weather events directly result in consumers' outages, this paper proposes a state-based decision-making model with the objective of mitigating loss of load to improve the distribution system resilience throughout the unfolding events. The sequentially uncertain system states, e.g., feeder line on/off states, driven by the unfolding events are modeled as Markov states, and the probabilities from one Markov state to another Markov state throughout the unfolding events are determined by the component failure caused by the unfolding events. A recursive optimization model based on Markov decision processes (MDP) is developed to make state-based actions, i.e., system reconfiguration, at each decision time. To overcome the curse of dimensionality caused by enormous states and actions, an approximate dynamic programming (ADP) approach based on post-decision states and iteration is used to solve the proposed MDP-based model. IEEE 33-bus system and IEEE 123-bus system are used to validate the proposed model.

Index Terms—Approximate dynamic programming, distribution systems, Markov decision processes, resilience enhancement

NOMENCLATURE

Indices and Sets

c	Index of components.
l	Index of lines.
k, k'	Index of terminal buses of line l .
t, τ	Index of time periods.
i, i', j, j'	Index of states.
\mathcal{A}	Set of actions.
\mathcal{B}	Set of buses.
$\tilde{\mathcal{B}}_{i,t}$	Set of non-islanded buses under the state $S_{i,t}$.
$\tilde{\mathcal{B}}$	Set of substation nodes.
$\mathcal{C}_{i,t}^f$	Set of repaired components under state $S_{i,t}$.
\mathcal{F}_t^i	Set of all possible failure components at t .
$\tilde{\mathcal{F}}_t$	Set of actual failure components at t .
$\mathcal{L}_{i,t}$	Set of non-islanded lines under state $S_{i,t}$.
$\mathcal{L}_{i',t}$	Set of dispatchable lines under post-decision state $S_{i',t}^{a_t}$.
$\mathcal{L}_{i,t}^d$	Set of dispatchable lines under state $S_{i,t}$.
$\mathcal{L}_{i,t}^{nd}$	Set of non-dispatchable lines under state $S_{i,t}$.
$N_{k,i,t}$	Set of nodes connected to bus k under state $S_{i,t}$.
$\tilde{\mathcal{R}}_\tau$	Set of repaired components at τ .
\mathcal{S}	Set of Markov states.
\mathcal{S}^{post}	Set of post-decision states of state $S_{i,t}$.
\mathcal{T}	Set of time periods.
Notation for Solution Method	
\mathbf{b}_i	Binary-coded matrix for post-decision states.
\mathbb{E}	Expected value.
m	Number of dispatchable lines.
n	Number of iterations.

N	Maximum number of iterations.
$S_{i,t}^{a_t}$	Post-decision after $S_{i,t}$ with action a_t .
$S_{j',t-1}^{a_{t-1}}$	Post-decision after $S_{j',t-1}$ with action a_{t-1} .
T	Number of decision periods.
$v_t^{a_t}$	Value function of post-decision state.
v_t^n	Value function of state at n^{th} iteration at t .
$\tilde{v}_t^{a_t,n}, \tilde{v}_{t-1}^{a_{t-1},n}$	Approximated value function of post-decision state at n^{th} iteration.
$V_{i'}$	Known values of post-decision states.
x_1, x_2	Binary variables.
$y_{l,i',t}$	Binary variable.
ϵ	A coefficient.
Notation for MDP-based Model	
a_t	Action at t .
C_t	Immediate cost at t .
C_l	Operational cost of line l at t .
$F_{kk',i,t}^p$	Active power flow on line $k - k'$ under $S_{i,t}$ at t .
$F_{kk',i,t}^q$	Reactive power flow on line $k - k'$ under $S_{i,t}$ at t .
$F_{kk'}^s$	Apparent power capacity of line $k - k'$.
$\Delta L_{k,i,t}^p, \Delta L_{k,i,t}^q$	Loss of active/reactive load of bus k under state $S_{i,t}$.
$L_{k,t}^p$	Active load of bus k at t .
$L_{k,t}^q$	Reactive load of bus k at t .
M	Large positive number.
$o_{kk',i,t}$	Binary variable, the value is 1 if bus k' is the parent bus for bus k under state $S_{i,t}$, otherwise 0.
$r_{kk'}/x_{kk'}$	Resistance/reactance of line $k - k'$.
$S_{i,t}, S_{j,t+1}$	Markov state i and j at t and $t + 1$, respectively.
$s_{c,t}$	On-off state of component c at t .
ΔT	Duration of each time period.
T_c^f	Time period from normal state to failure state of component c .
ΔT_c^r	Repair duration for component c .
$U_{k,i,t}$	Squared voltage magnitude of bus k .
$\underline{V}_k, \overline{V}_k$	Low/upper limits of voltage value of bus k .
v_t, v_{t+1}	values functions at t and $t + 1$.
Pr	Transition probability.
$\beta_{l,i,t}, \beta_{c,i,t}$	Binary variables representing on-off states of line l and c , respectively. 1 denotes on state, and 0 denotes off state.
η_t, η_{t-1}	Penalty due to loss of load at t and $t - 1$ (\$/MWh).
ξ_t	Uncertainty of extreme event at t .

I. INTRODUCTION

BECAUSE distribution systems are directly connected to commercial and residential customers with radial topologies, any failures in distribution systems will lead to outages.

Climate change increases the frequency and intensity of severe weather, which is a major cause of severe system failures. For example, weather events caused roughly 679 power outages, each of which affected at least 50,000 customers, between 2003 and 2012. Although transmission system outages did occur, a major portion of outages occurred along distribution systems [1]. The severe consequences have required distribution systems to have resilience against these extreme weather events, and this has been identified by the United States Electric Power Research Institute (EPRI) [2] and the North American Electric Reliability Corporation (NERC) [3].

To enhance the system resilience, we can take actions in three stages with regard to weather events, i.e., prior to events, during events, and after events [4], [5]. Prior to weather events, the historical data-based models [6]–[8] are used to estimate outages that help the system operators to make preventive actions such as system maintenance [9] and system hardening [10]. System hardening makes physical infrastructural changes to systems so that they are less susceptible to extreme events. For example, a coordinated hardening and distributed generator (DG) allocation strategy has been developed in [10]. Prior to events, preparing enough blackstart generators and emergency generators after potential failures is also a critical measure to improve the system resilience. To this end, a Generic Restoration Milestones (GRMs)-based algorithm is developed to assess blackstart capacities [11], and a procurement plan with a minimal cost while guaranteeing sufficient blackstart capacities is proposed to provide enough blackstart resources at right locations [12]. To effectively isolate possible failures and connect blackstart/emergency generators to systems, a resilience-based model for switch placement in distribution systems is developed prior to events [13]. In addition to physical power systems, hardening communication systems in charge of monitoring/controlling the physical power systems play an important role in enhancing the system resilience [14]. Even though many preventive actions are performed prior to events, it is impossible to avoid outages completely. When outages occur after events, it is necessary to recover outages as quickly as possible to improve the system resilience. A conventional power system restoration includes three stages, i.e., preparation, system restoration and load restoration [15]–[17]. Some algorithms such as expert systems [18] and heuristic approaches [19] are proposed to accelerate load recovery. However, there are unique characteristics associated with outages caused by weather-related events, leading to different restoration strategies such as microgrid-based restoration strategies [20] and decentralized restoration schemes [21].

The above studies mainly focus on strategies prior to events and after events. However, few studies investigate the strategies during the unfolding events. One difficulty in establishing strategies during the unfolding events is to map sequentially varying states caused by the unfolding events to optimal strategies. The commonly used scenario-based stochastic programming [22] and robust stochastic programming are not suitable for mapping sequentially varying states to optimal strategies. To address this difficulty, MDP can be employed to help make state-based decisions on a stochastic environment caused by weather events. Some applications of MDP in power systems have been investigated [23], [24]. For the resilience enhancement, [25] proposes sequentially proactive MDP-based strategies to improve the transmission system resilience, and a linear scalarization method based on the state tree is used to solve the proposed model. However, distribution systems and transmission systems differ in topologies and allowable actions, the developed model and solution in [25] cannot be applied to distribution systems directly. It is necessary to

develop state-based decision-making processes for distribution systems considering their own characteristics.

This paper proposes MDP-based resilience enhancement for distribution systems. The contributions of this paper are three-fold: 1) The sequentially uncertain states, i.e., feeder line on/off states, on the trajectory of an unfolding event are represented as Markov states. Transition probabilities between Markov states are determined by component failure probabilities caused by the unfolding event. 2) A recursive optimization model for each Markov state is constructed to map states to optimal strategies. The allowable action for each state is system reconfiguration. 3) An approximate dynamic programming approach based on post-decision states and value function approximation is employed to solve the proposed model to deal with the curse of dimensionality caused by a mass of states and allowable actions.

The remainder of this paper is organized as follows. Section II shows extreme events' impacts on system states. Section III presents the mathematical formulation, and section IV introduces the solution method. The case studies are demonstrated in Section V, and the work is concluded in Section VI.

II. MODELING INFLUENCES OF EVENTS ON DISTRIBUTION SYSTEMS

This section first introduces Markov states on the trajectories of extreme events, and then presents transition probabilities between different Markov states under allowable actions.

A. Markov states on the event's trajectories

Usually, the impacts of a weather-related event on a distribution system are sequential due to the sequential trajectory, indicating that the components with different locations in the system may be in failure in different time periods. This results in the sequential changes of system states such as on-off states of distribution lines. A system state including on-off states of distribution lines on the trajectory is represented as a Markov state in this paper. Define \mathcal{F}_t as the set of all possible failure components due to the unfolding event at t , and $\tilde{\mathcal{F}}_t$ as the set of the actual failure components at t , and we have $\tilde{\mathcal{F}}_t \subseteq \mathcal{F}_t$. For example, $\mathcal{F}_{t_2} = \{b_{1-2}, b_{1-4}\}$ and $\tilde{\mathcal{F}}_{t_2} = \{b_{1-4}\}$. The failure scenarios of the lines b_{1-2} and b_{1-4} are uncertain before the time period t_2 , and the actual state can only be observed at t_2 and the actual failure on b_{1-4} occurs. Considering the sequential characteristic of the extreme event, the Markov state $S_{i,t}$ at t can be represented as follows.

$$S_{i,t} = \bigcup_{\tau=1}^t (\tilde{\mathcal{F}}_{\tau} - \tilde{\mathcal{R}}_{\tau}) \quad (1)$$

where (1) shows that the Markov state at t is represented as failure components except repaired components from the initial time period to the time period t . $\tilde{\mathcal{R}}_{\tau}$ is the set of components repaired at time τ .

B. Transition probability between Markov states

On the trajectory of the event, the current Markov state at t has a probability of reaching to each future Markov state at $t+1$, and the probability is called as a transition probability, which is determined by component failure rates caused by the event and the allowable system reconfiguration. The transition probability can be expressed as follows.

$$\Pr(S_{j,t+1}|S_{i,t}, a_t, \xi_t) = \prod_{c \in \mathcal{F}_{t+1}} \Pr(s_{c,t+1}|s_{c,t}, a_t, \xi_t) \quad (2)$$

where $\Pr(S_{j,t+1}|S_{i,t}, a_t, \xi_t)$ means the probability from the state $S_{i,t}$ to the state $S_{j,t+1}$ under the action a_t with the uncertainty ξ_t , and $\Pr(s_{c,t+1}|s_{c,t}, a_t, \xi_t)$ represents the probability from the c th component's on-off state $s_{c,t}$ to the on-off state $s_{c,t+1}$ under the action a_t with the uncertainty ξ_t . Three causes can change the on-off states, and they are listed as follows.

- Component failure caused by extreme events: Since the on-off states are uncertain in the next time period due to extreme events, it is necessary to calculate the probability of each scenario. At present, there are many existing studies on components' failure probabilities caused by extreme events [26].
- System reconfiguration: After system reconfiguration, the line state, i.e., on state or off state, is determined. This indicates that the corresponding probability from $s_{k,t}$ to $s_{k,t+1}$ is 0 or 1.
- Repair: Before the failure components are repaired, the state is off. After repaired, the state is on. This shows that the corresponding probability from $s_{k,t}$ to $s_{k,t+1}$ is 0 or 1.

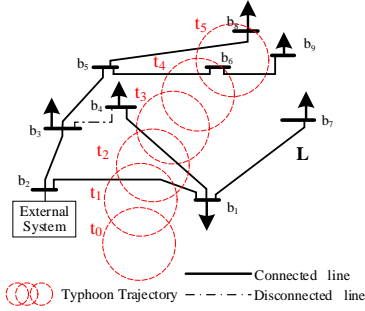


Fig. 1. An example of a distribution system under an unfolding event

III. OPTIMIZATION MODEL BASED ON MARKOV DECISION PROCESSES

This section first introduces a recursive model to map each Markov state to its optimal strategy, and then lists the operational constraints for distribution systems.

A. Markov Decision Processes-based Recursive Model

Transition probabilities in (2) show that the current actions associated with uncertainties caused by extreme events impact the current states and future states. Since different states result in different operational costs, it is necessary to make decisions based on not only current states but also future states impacted by transition probabilities. Take the scenario in Fig. 1 as an example. Since the line b_{1-2} will be impacted by the typhoon at t_1 , the line b_{1-2} may be in failure at t_1 , resulting in the outages of the lines b_{1-4} and b_{1-7} . If we can disconnect the line b_{1-2} and connect the line b_{3-4} before the line b_{1-2} is impacted by the typhoon, we can avoid the outages of the lines b_{1-4} and b_{1-7} at t_1 . For this case, we only consider one time period ahead and there is only one line impacted by the typhoon at t_1 . If all time periods over the unfolding typhoon and numerous components impacted in each time period are considered, we need to develop a model to help make decisions to ensure the minimum operational cost. Considering the sequential time periods and numerous components on the trajectory, we establish a recursive model

with the current cost and the expected future cost listed as follows.

$$v_t(S_{i,t}) = \min_{a_t \in \mathcal{A}} \left(C_t(S_{i,t}, a_t) + \sum_{S_{j,t+1} \in \mathcal{S}} \Pr(S_{j,t+1}|S_{i,t}, a_t, \xi_t) \cdot v_{t+1}(S_{j,t+1}) \right) \quad (3)$$

where $v_t(S_{i,t})$ and $v_{t+1}(S_{j,t+1})$ are the value functions of the states $S_{i,t}$ and $S_{j,t+1}$ at t and $t+1$, respectively. The second term on the right side of (3) shows the expected future cost. $C_t(S_{i,t}, a_t)$ is the current cost caused by the action a_t under the state $S_{i,t}$ at t , and this cost in this study is defined as the sum of the cost of loss of load and the operational cost of controllable lines. It is expressed as follows.

$$C_t(S_{i,t}, a_t) = \sum_{k \in \mathcal{B}} (\eta_t \cdot \Delta L_{b,i,t}^p \cdot \Delta T) + \sum_{l \in \mathcal{L}_{i,t}^d} (\beta_{l,i,t} \cdot C_l) \quad (4)$$

where the first term on the right side of (4) is the cost of loss of load, and the second term is the operational cost of controllable lines.

B. Operational Constraints

For each time period during the unfolding event, the operational constraints, i.e., radial topologies, power balance, power flow, voltage limits, and line capacity, should be satisfied.

1) *Radiality constraint*: Different from transmission systems, distribution systems should operate in radial topologies. When performing system reconfiguration under the state $S_{i,t}$, the spanning tree constraints are used to guarantee the network radiality.

$$o_{kk',i,t} + o_{k'k,i,t} = \beta_{l,i,t} \quad l \in \mathcal{L}_{i,t}, t \in \mathcal{T} \quad (5a)$$

$$\sum_{k' \in \mathcal{N}_{k,i,t}} o_{kk',i,t} = 1 \quad k \in \bar{\mathcal{B}}_{i,t}, t \in \mathcal{T} \quad (5b)$$

$$o_{kk',i,t} = 0 \quad k \in \tilde{\mathcal{B}}, k' \in \mathcal{N}_{k,i,t}, t \in \mathcal{T} \quad (5c)$$

where (5a) and (5b) constrain that the two terminals of a connected line only have one parent bus. In practice, islanded buses, to which power cannot be supplied by the grid, maybe exist due to component failures caused by extreme events, and these islanded buses are not included in the spanning tree constraint. (5c) indicates that the substation bus (i.e., the bus connected to the external system) has no parent buses.

In practice, it is possible that only some lines can be dispatched under the state $S_{i,t}$. In this case, we can add a constraint with regard to non-dispatched lines.

$$\beta_{l,i,t} = 1 \quad l \in \mathcal{L}_{i,t}^{nd}, t \in \mathcal{T} \quad (6)$$

2) *Repair constraint*: When there are components in failure under the state $S_{i,t}$, these failure components and the resulting islanded components cannot participate in system dispatch before they are repaired.

$$\beta_{c,i,t} = 0 \quad c \in \mathcal{C}_{i,t}^f, T_c^f \leq t \leq T_c^f + \Delta T_c^r \quad (7)$$

where (7) means that the state of the failure component c is set to 0 during the repair time periods.

3) *Power flow constraint*: The power flow of each line has relations to bus voltages of the two terminal buses of each line, and can be expressed as follows.

$$U_{k,i,t} - U_{k',i,t} \leq (1 - \beta_{l,i,t}) \cdot M + 2(r_{kk'} \cdot F_{kk',i,t}^p + x_{kk'} \cdot F_{kk',i,t}^q) \quad l \in \mathcal{L}_{i,t}, t \in \mathcal{T} \quad (8a)$$

$$U_{k,i,t} - U_{k',i,t} \geq (\beta_{l,i,t} - 1) \cdot M + 2(r_{kk'} \cdot F_{kk',i,t}^p + x_{kk'} \cdot F_{kk',i,t}^q) \quad l \in \mathcal{L}_{i,t}, t \in \mathcal{T} \quad (8b)$$

where (8a)-(8b) are derived from the DsitFlow model [27]. The quadratic terms in the accurate power flow model are ignored [28]. The big M is a disjunctive parameter. With a sufficiently large M , (8a)-(8b) are redundant when distribution lines are disconnected or outages. Non-islanded buses are included in these constraints.

4) *Power balance constraint*: When reaching to the state $S_{i,t}$ at t , the out-flow/in-flow power of each non-islanded bus in the system should be equal. The constraint can be expressed as follows.

$$\left(L_{k,t}^p - \Delta L_{k,i,t}^p \right) + \sum_{k' \in \mathcal{N}_{k,i,t}} F_{kk',i,t}^p = 0 \quad k \in \bar{\mathcal{B}}_{i,t}, t \in \mathcal{T} \quad (9a)$$

$$\left(L_{k,t}^q - \Delta L_{k,i,t}^q \right) + \sum_{k' \in \mathcal{N}_{k,i,t}} F_{kk',i,t}^q = 0 \quad k \in \bar{\mathcal{B}}_{i,t}, t \in \mathcal{T} \quad (9b)$$

where (9a) and (9b) represent real power balance and reactive power balance, respectively. Only non-islanded buses are included in the constraint. The load connected to the islanded buses in the system is directly considered as loss of load in (4).

5) *Line capacity constraint*: The power through each line should be within the limit for the state $S_{i,t}$ with the action a_t . The constraint can be expressed as follows.

$$(F_{kk',i,t}^p)^2 + (F_{kk',i,t}^q)^2 \leq \beta_{l,i,t} \cdot (F_{kk'}^s)^2 \quad l \in \mathcal{L}_{i,t}, t \in \mathcal{T} \quad (10)$$

where (10) is a nonlinear constraint, resulting in computational intractability. To facilitate the model solution, the constraint (10) is relaxed to a group of linear constraints [29], and are rewritten as follows.

$$-\beta_{l,i,t} \cdot F_{kk'}^s \leq F_{kk',i,t}^p \leq \beta_{l,i,t} \cdot F_{kk'}^s \quad l \in \mathcal{L}_{i,t}, t \in \mathcal{T} \quad (11a)$$

$$-\beta_{l,i,t} \cdot F_{kk'}^s \leq F_{kk',i,t}^q \leq \beta_{l,i,t} \cdot F_{kk'}^s \quad l \in \mathcal{L}_{i,t}, t \in \mathcal{T} \quad (11b)$$

$$-\sqrt{2}\beta_{l,i,t} \cdot F_{kk'}^s \leq F_{kk',i,t}^p + F_{kk',i,t}^q \leq \sqrt{2}\beta_{l,i,t} \cdot F_{kk'}^s \quad l \in \mathcal{L}_{i,t}, t \in \mathcal{T} \quad (11c)$$

$$-\sqrt{2}\beta_{l,i,t} \cdot F_{kk'}^s \leq F_{kk',i,t}^p - F_{kk',i,t}^q \leq \sqrt{2}\beta_{l,i,t} \cdot F_{kk'}^s \quad l \in \mathcal{L}_{i,t}, t \in \mathcal{T} \quad (11d)$$

6) *Voltage constraint*: The voltage limits under the state $S_{i,t}$ with the action a_t should be satisfied.

$$\underline{V}_k^2 \leq U_{k,i,t} \leq \bar{V}_k^2 \quad k \in \bar{\mathcal{B}}_{i,t}, t \in \mathcal{T} \quad (12)$$

C. MDP-based Optimization Model

The constructed MDP-based optimization model can be represented as follows.

$$\begin{aligned} \text{Obj.} & \quad (3) \\ \text{s.t.} & \quad (4) - (12) \end{aligned} \quad (13)$$

This is a recursive model for each Markov state which is computationally intractable. The next section will introduce the solution method.

IV. MODEL SOLUTION BASED ON APPROXIMATE DYNAMIC PROGRAMMING

This section first introduces the challenges of solving the proposed model, and then presents the basic idea of ADP in solving the MDP-based model, and finally shows how to solve the proposed model by using the ADP approach.

A. Challenges of model solution

When using the conventional stochastic programming to deal with a sequential decision-making problem with uncertainties, one commonly used approach is to generate some scenarios to represent the uncertainties. Based on these generated scenarios, we optimize a model with an expected objective, and then we can obtain the optimal strategy. The conventional stochastic programming cannot be used to deal with the proposed model because the influences of actions on scenario transitions are not included and the real scenario may not be included in the generated scenarios.

The decision processes used in this paper can be illustrated by using the case in Fig. 2 (a), (b), and (c), which represent the decision making processes at t_1 , t_2 , and t_3 , respectively. At t_1 , the state transition tree under the impacts of actions is constructed, as shown in Fig. 2 (a), and the optimal action can be obtained by optimizing the recursive model (3). After performing the optimal action, the state reaches to a new state at t_2 under uncertainty, as shown in Fig. 2 (b). For this new state, the state transition tree under the impacts of actions needs to be updated again because some state transitions may be invalid. With the new state transition tree, the optimal action for the new state at t_2 can be obtained by optimizing the recursive model (3) again. The optimal action for the new state at t_3 can be obtained with the similar processes. According to the decision processes, constructing the state transition tree under actions in consideration of uncertainties is one critical step to solve the MPD-based model. However, it is a difficult task to construct the state transition tree of the proposed model in consideration of various actions and the resulting complicated state transitions. In addition, a large-scale problem with numerous states is possibly intractable due to ‘‘curse of dimensionality’’. There are three curses of dimensionality: (i) the state space \mathcal{S} may be too huge to calculate the value function $v_t(S_{i,t})$ for each state within acceptable time, (ii) the decision space \mathcal{A} is too large to obtain the optimal action for each state, (iii) the outcome space may be too large to calculate the expectation of future cost.

B. Approximate dynamic programming

Approximate dynamic programming is a modeling framework offering some techniques for dealing with the curses of dimensionality in multi-period, large, and stochastic MDP-based models. There are two critical techniques used by ADP: 1) post-decision states are constructed to deal with the large outcome space, 2) a forward dynamic algorithm based on sample paths is used to solve the recursive model by stepping forward in time, and repeat this procedure for enough iterations.

1) *Post-decision states*: The post-decision state, defined as $S_{i,t}^{a_t}$, is a state immediately after the action a_t but before the arrival of a new state in consideration of uncertainties. To

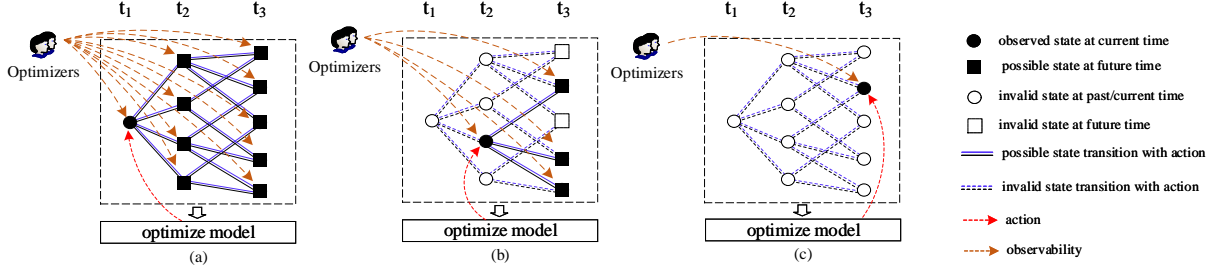


Fig. 2. Markov state-based decision at (a) t_1 , (b) t_2 , and (c) t_3 .

apply the ADP approach, a more generic form of the proposed model in (3) is an expectational form listed as follows.

$$v_t(S_{i,t}) = \min_{a_t \in \mathcal{A}} (C_t(S_{i,t}, a_t) + \mathbb{E}\{v_{t+1}(S_{j,t+1}|S_{i,t}, a_t, \xi_t)\}) \quad (14)$$

where (14) can be rewritten as (15) with the post-decision state $S_{i,t}^{a_t}$.

$$v_t(S_{i,t}) = \min_{a_t \in \mathcal{A}} (C_t(S_{i,t}, a_t) + \mathbb{E}\{v_{t+1}(S_{j,t+1}|S_{i,t}^{a_t}, \xi_t)\}) \quad (15)$$

Define $\mathbb{E}\{v_{t+1}(S_{j,t+1}|S_{i,t}^{a_t}, \xi_t)\}$ by $v_t^{a_t}(S_{i,t}^{a_t})$, we have the following optimality equations.

$$v_t(S_{i,t}) = \min_{a_t \in \mathcal{A}} (C_t(S_{i,t}, a_t) + v_t^{a_t}(S_{i,t}^{a_t})) \quad (16a)$$

$$v_{t-1}^{a_{t-1}}(S_{j',t-1}^{a_{t-1}}) = \mathbb{E}\{v_t(S_{i,t}|S_{j',t-1}^{a_{t-1}}, \xi_{t-1})\} \quad (16b)$$

Substituting (16a) into (16b) results in the optimality equations of the post-decision states as follows.

$$v_{t-1}^{a_{t-1}}(S_{j',t-1}^{a_{t-1}}) = \mathbb{E}\{\min_{a_t \in \mathcal{A}} (C_t(S_{i,t}, a_t) + v_t^{a_t}(S_{i,t}^{a_t}|S_{j',t-1}^{a_{t-1}}, \xi_{t-1}))\} \quad (17)$$

where (17) can be rewritten as the form at t as follows.

$$v_t^{a_t}(S_{i,t}^{a_t}) = \mathbb{E}\{\min_{a_{t+1} \in \mathcal{A}} (C_{t+1}(S_{j,t+1}, a_{t+1}) + v_{t+1}^{a_{t+1}}(S_{j,t+1}^{a_{t+1}}|S_{i,t}^{a_t}, \xi_t))\} \quad (18)$$

where (18) shows the value of the post-decision state.

When $v_t^{a_t}(S_{i,t}^{a_t})$ in (16a) is known, it would be easy to solve the optimization model (16a). Based on this idea, ADP is to use the deterministic optimization model (16a) with an initial estimation of $\tilde{v}_t^{a_t}(S_{i,t}^{a_t})$ of $v_t^{a_t}(S_{i,t}^{a_t})$ to make decisions for each state, and then employ the resulting observations to update an estimation $\tilde{v}_t^{a_t}(S_{i,t}^{a_t})$ thereby approximating the expected value in (18).

2) *Forward Dynamic algorithm*: For the forward dynamic algorithm, the recursive model is solved only for one state in each time period, by using the estimation of the post-decision state and performing iterations to update the estimations of the post-decision states on the sample paths. To deal with iterations, we add a superscript n and $n-1$ to the value functions, and (16a) can be expressed as follows.

$$v_t^n(S_{i,t}) = \min_{a_t \in \mathcal{A}} (C_t(S_{i,t}, a_t) + \tilde{v}_t^{a_t, n-1}(S_{i,t}^{a_t})) \quad (19)$$

where the decision that minimizes (19) at n th iteration is shown as follows.

$$a_t = \arg \min_{a_t \in \mathcal{A}} (C_t(S_{i,t}, a_t) + \tilde{v}_t^{a_t, n-1}(S_{i,t}^{a_t})) \quad (20)$$

where $\tilde{v}_t^{a_t, n-1}(S_{i,t}^{a_t})$ can be updated by

$$\tilde{v}_t^{a_t, n-1}(S_{i,t}^{a_t}) = (1 - \epsilon) \cdot \tilde{v}_t^{a_t, n-2}(S_{i,t}^{a_t}) + \epsilon \cdot v_{t+1}^{n-1}(S_{j,t+1}) \quad (21)$$

where the first term on the right side of (21) represent the estimate of the post-decision state $S_{i,t}^{a_t}$ at the $(n-2)$ th iteration, and the second term represent the value of the resulting observations from the post-decision state $S_{i,t}^{a_t}$ at the n th iteration.

C. Reformulation of the proposed model

Based on (19), we just need to solve a deterministic model. In the model, the term $C_t(S_{i,t}, a_t)$ is an explicit objective (4) with regard to variables associated with constraints (5)-(12), however, the term $\tilde{v}_t^{a_t, n}(S_{i,t}^{a_t})$ is just a value with regard to the post-decision state $S_{i,t}^{a_t}$ but has no relations to the variables and actions. In this case, it is not possible to optimize the model (19). Therefore, it is necessary to relate $\tilde{v}_t^{a_t, n}(S_{i,t}^{a_t})$ to the variables and actions.

The Markov state in the study is determined by the on-off states of distributed lines. Since a failure is an observed event and the repair is an activity with continuous time period, whether a failure component is repaired or not at the current period is known. Furthermore, the post-decision states at the current period are defined as states before arrival of uncertainties in the next time period. Therefore, system reconfiguration is the cause of changing the current state $S_{i,t}$ to $S_{i,t}^{a_t}$. It is assumed that there are two reconfigurable lines and the corresponding binaries are x_1 and x_2 . We will have four post-decision states listed in Table I. In this case, the second term $\tilde{v}_t^{a_t, n}(S_{i,t}^{a_t})$ in (19) can be expressed as $(1-x_1)(1-x_2)V_1 + (1-x_1)x_2V_2 + x_1(1-x_2)V_3 + x_1x_2V_4$ which relates the values of post-decision states to decision variables.

TABLE I
MY CAPTION

Post-decision states	x_1	x_2	Estimated Values
$S_{1,t}^{a_t}$	0	0	$\tilde{v}_t^{a_t, n}(S_{1,t}^{a_t}) \triangleq V_1$
$S_{2,t}^{a_t}$	0	1	$\tilde{v}_t^{a_t, n}(S_{2,t}^{a_t}) \triangleq V_2$
$S_{3,t}^{a_t}$	1	0	$\tilde{v}_t^{a_t, n}(S_{3,t}^{a_t}) \triangleq V_3$
$S_{4,t}^{a_t}$	1	1	$\tilde{v}_t^{a_t, n}(S_{4,t}^{a_t}) \triangleq V_4$

Based on this technique, we can rewrite the second term $\tilde{v}_t^{a_t, n}(S_{i,t}^{a_t})$ in (19) in a generic form. For the state $S_{i,t}$, there are m reconfigurable lines, resulting in 2^m post-decision states associated with the corresponding estimates (represented as V_1, V_2, \dots, V_{2^m}) at the n th iteration. The 2^m post-decision states are binary-coded with a $2^m \times m$ matrix \mathbf{b}_i , in which

the $(i', l)^{th}$ entry denotes the on-off state of the line l under the i'^{th} post-decision states. The generic form of the term $\tilde{v}_t^{a_t, n}(S_{i,t}^{a_t})$ at the n^{th} iteration is listed as follows.

$$\sum_{i' \in \mathcal{S}_i^{post}} \left\{ \prod_{l \in \mathcal{L}_{i,t}^d} (1 - \beta_{l,i',t} - \mathbf{b}_i(i', l))(1 - 2\mathbf{b}_i(i', l))V_{i'} \right\} \quad (22)$$

where $\beta_{l,i',t}$ is a binary representing on-off states of the line l under post-decision states. $\mathbf{b}_i(i', l)$ is a known value with 0 or 1, making (22) a sum of multilinear functions. The optimization model (19) in forward dynamic algorithm can be rewritten as follows.

$$\begin{aligned} \min & \quad (4) + (22) \\ \text{s.t.} & \quad (5), (6), (7), (8), (9), (10), (11), (12) \end{aligned} \quad (23)$$

For multilinear functions in (22), McCormick proposed a recursive procedure in which additional variables and constraints are added to obtain a formulation of the problem having only bilinear equations, which can be represented by four binary inequations. In (22), the multilinear function with most variables is $\beta_{1,i',t}\beta_{2,i',t}\cdots\beta_{m,i',t}$, which can be represented by additional variables

$$\begin{aligned} y_{2,i',t} &= \beta_{1,i',t}\beta_{2,i',t} \\ y_{3,i',t} &= y_{2,i',t}\beta_{3,i',t} \\ &\dots \\ y_{m,i',t} &= y_{m-1,i',t}\beta_{m,i',t} \end{aligned} \quad (24)$$

and additional constraints

$$\begin{aligned} y_{2,i',t} &\geq \beta_{2,i',t} + \beta_{1,i',t} - 1 \\ y_{2,i',t} &\leq \beta_{1,i',t} \\ y_{l,i',t} &\geq 0 \quad (l = 2, \dots, m) \\ y_{l,i',t} &\leq \beta_{l,i',t} \quad (l = 2, \dots, m) \\ y_{l,i',t} &\geq \beta_{l,i',t} + y_{l-1,i',t} - 1 \quad (l = 3, \dots, m) \\ y_{l,i',t} &\leq y_{l-1,i',t} \quad (l = 3, \dots, m) \end{aligned} \quad (25)$$

where (25) is an exact reformulation of $\beta_{1,i',t}\beta_{2,i',t}\cdots\beta_{m,i',t}$ since $\beta_{l,i',t}, l = 1, 2, \dots, m$ are binary variables. Based on the additional variables and the additional constraints, the optimization model is a mixed integer linear programming, which can be solved by many solvers such as CPLEX and GUROBI.

D. Solution procedure

Based on Section IV.B and Section IV.C, we can solve the proposed MDP-based model by means of the iteration-based ADP approach, and the CPLEX solver is used to solve the MILP model at each iteration. The detailed procedure is listed in **Algorithm 1**.

V. CASE STUDIES

In this section, two test systems are used to verify the proposed model and the algorithm. The first system is the IEEE 33-bus system, and the second system is the IEEE 123-bus system. The cases are tested in MATLAB 2017a using the CPLEX 12.6 solver on computers with 3.1 GHz i5 processors and 8 GB RAMS.

A. IEEE 33-bus system

1) *Data description*: Fig. 3 shows the topology of the IEEE 33-bus system. The typhoon trajectory is also shown in Fig. 3. For the original topology, the lines 8-21, 12-22, 1-18, 9-15, and 25-29 are disconnected to ensure the radial topology. It is assumed that the lines 10-11, 12-13, 25-29, 1-18, 14-15, 12-22, 8-21, and 9-15 are dispatchable and the other lines are non-dispatchable.

Algorithm 1 ADP algorithm

- 1: Step 1. Initialization
- 2: Step 1.1. Set the iteration counter $n = 1$ and the maximum number of iterations N .
- 3: Step 1.2. Set the initial approximation for each state.
- 4: Step 2. Do for $t = 1, \dots, T$
- 5: Step 2.1. Solve (19) and (20) to get $v_t^n(S_{i,t})$ and a_t at the iteration n . To solve (19), (23) and (25) will be used.
- 6: Step 2.2. Update the approximation $\tilde{v}_t^{a_t, n}(S_{i,t}^{a_t})$ for the post-decision $S_{i,t}^{a_t}$ with (21).
- 7: Step 2.3. Obtain the post-decision $S_{i,t}^{a_t}$ from the state $S_{i,t}$ under a_t .
- 8: Step 2.4. According to uncertainties of the extreme event, generate a new state $S_{j,t+1}$ at $t+1$ from the post-decision state $S_{i,t}^{a_t}$ at t .
- 9: Step 3. Increment n . If $n \leq N$ go to Step 1.
- 10: Step 4. Return the value function $\tilde{v}_t^{a_t, N}(S_{i,t}^{a_t})$ and the corresponding action a_t for the state $S_{i,t}$.

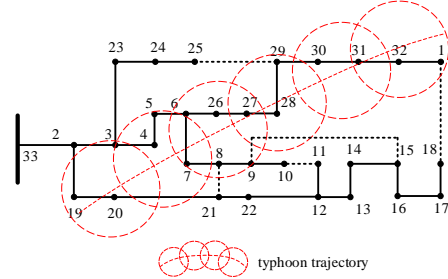


Fig. 3. Topology of IEEE 33-bus system.

2) *Estimated values of post-decision states*: Because post-decision states are introduced to make the proposed recursively state-based model easily to be solved, one important task is to first estimate the values of these post-decision states according to the ADP algorithm. Due to a large number of post-decision states, we only show the estimated values of some post-decision states for the sake of exposition. Fig. 4 shows the estimated values of four post-decision states $S_{1,2}$, $S_{2,2}$, $S_{3,2}$ and $S_{4,2}$, shown in Table II, in the second decision period. 1500 iterations were performed to get the estimated values of post-decision states, and the estimated values of the post-decision states $S_{1,2}$, $S_{2,2}$, $S_{3,2}$ and $S_{4,2}$ converge to 1.29×10^6 \$, 1.37×10^6 \$, 1.45×10^6 \$, and 1.24×10^6 \$, respectively. With the estimated values, the term $v_t^{a_t}(S_{i,t}^{a_t})$ is known when optimizing (16a). In this case, the stochastic problem is transformed into a deterministic problem.

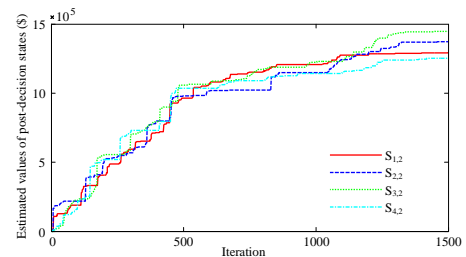


Fig. 4. Iterations for estimated values of post-decision states

TABLE II
POST-DECISION STATES $S_{1,2}$, $S_{2,2}$, $S_{3,2}$ AND $S_{4,2}$

Time Period	Post-decision states	Disconnected lines
2	$S_{1,2}$	10-11, 12-13, 25-29, 1-18, 8-21
2	$S_{2,2}$	10-11, 25-29, 1-18, 14-15, 8-21
2	$S_{3,2}$	10-11, 25-29, 1-18, 12-22, 8-21
2	$S_{4,2}$	25-29, 1-18, 14-15, 12-22, 8-21
1	$S_{1,1}$	10-11, 12-13, 25-29, 1-18, 8-21

Different intensity of severe weather will result in different failure rates, which have great impacts on dispatch strategies. From the perspective of the mathematical model, different failure rates cause different estimated values of each post-decision state. Fig. 5 shows the estimated values of several post-decision states with different failure rates. For example, the estimated values of the post-decision state $S_{1,1}$ in the first period are 1.01×10^6 €, 1.58×10^6 €, and 1.81×10^6 € when the failure rates are 0.02, 0.04, and 0.06, respectively.

When updating the estimated values of post-decision states by using (21), ϵ is artificially set. Fig. 6 shows the impacts of different values of ϵ on the estimated values of the post-decision state $S_{2,1}$. It is observed that the estimated value are close even when ϵ has different values.

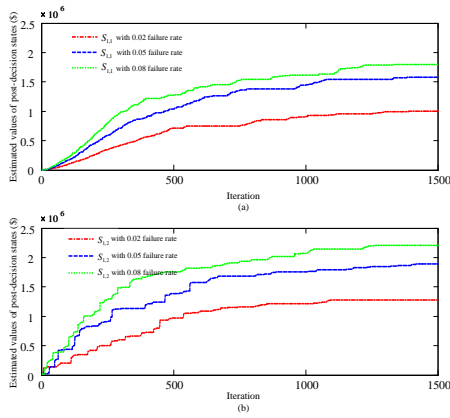


Fig. 5. Iterations for estimated values of post-decision states with different failure rates

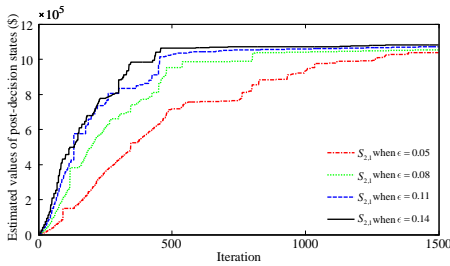


Fig. 6. Iterations for estimated values of post-decision states with different values of ϵ

3) *Dispatch strategies with estimated values of post-decision states:* With the estimated values of each post-decision state based on ADP, the strategy corresponding to one observed real-time state can be obtained by a one-period deterministic optimization problem. Table III and Table IV show the state-based strategies, and the original topology has the disconnected lines 10-11, 8-21, 9-15, 1-18, and 25-29. It is observed that the strategies make that the feeders impacted by the typhoon are downstream. This is reasonable because

downstream feeders cause smaller outages even they are in failure due to the typhoon.

TABLE III
THE FIRST CASE OF STATE-BASED STRATEGY

Time Period	Observed State (Component failure)	Strategy	
		Open lines	Close lines
1	-	19-20	10-11
2	6-7	10-11, 6-26	19-20, 8-21, 1-18
3	-	-	-
4	-	-	-
5	-	-	-
6	-	-	-

TABLE IV
THE SECOND CASE OF STATE-BASED STRATEGY

Time Period	Observed State (Component failure)	Strategy	
		Open lines	Close lines
1	-	19-20	10-11
2	-	6-26	19-20, 1-18
3	-	-	-
4	-	-	-
5	-	-	-
6	-	-	-

B. IEEE 123-bus system

1) *Data description:* Fig. 7 shows the topology of the IEEE 123-bus system and the trajectory of a typhoon. The original topology has the disconnected lines 16-96, 92-120, 115-116, 42-120, 38-43, 39-57, 56-76, 46-65, 51-108, and 71-85. The lines 16-96, 92-120, 56-76, 39-57, 38-43, 42-120, 46-65, 51-108, 71-85, 52-53, 60-57, 60-117, 101-119, 63-64, and 67-117 are dispatchable.

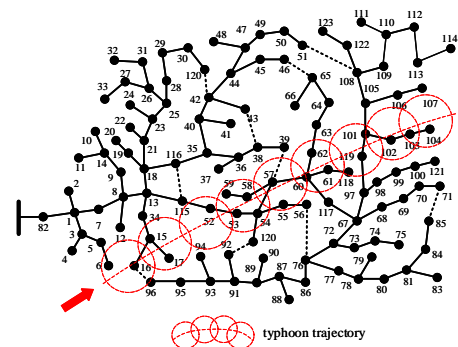


Fig. 7. Topology of IEEE 123-bus system

2) *Simulated results:* Based on the ADP algorithm, the estimated values of post-decision states can be obtained, and then the state-based strategies can be optimized. Table V shows the state-based strategies on the trajectory of the typhoon, and the states on the trajectory are assumed to be generated stochastically based on failure rates caused by the typhoon. Fig. 8 (a) and (b) show the topologies after implementing the state-based strategies in the 3th and 6th periods, respectively. In the 3th period, the line 52-53 is disconnected and the line 46-65 is connected to avoid balck out of downstream feeders if the typhoon fails the line 52-53. In the 6th period, three lines (57-60, 60-117, 101-119) are disconnected and three lines (52-53, 56-76, 51-108) are connected to reduce possible black-out areas. It is observed that the state-based strategies try to make

the feeders on the trajectory locate the terminal of the whole network to reduce potential loss of load.

TABLE V
STATE-BASED STRATEGY FOR IEEE 123-BUS SYSTEM

Time Period	Observed State (Component failure)	Strategy	
		Open lines	Close lines
1	-	-	-
2	15-17	-	-
3	-	52-53	46-65
4	-	-	-
5	58-57	-	-
6	-	57-60, 60-117,	52-53, 56-76
7	-	101-119	51-108
8	102-103	-	-
9	-	-	-

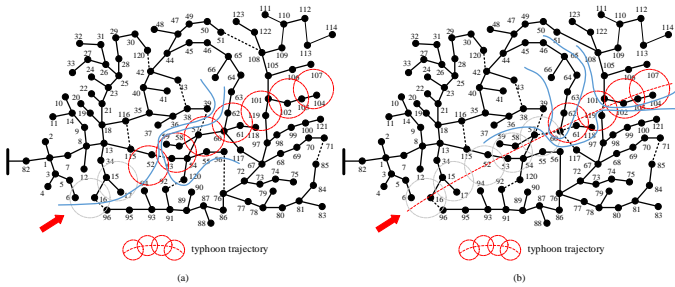


Fig. 8. System Topologies (a) in the 3rd period and (b) in the 6th period.

VI. CONCLUSION

This paper proposed a Markov state-based decision-making model with dispatching system topology to improve the distribution system resilience throughout the unfolding events. The sequentially states of system topologies changed by the unfolding events and actions are modeled as Markov states, and the uncertainties between different Markov states are represented as transition probabilities that are determined by the component failure rates caused by the unfolding events. Based on Markov states, a recursive optimization model based on Markov decision processes, including the current cost and the expected cost in the future, is developed to make state-based actions at each decision time. To deal with ‘curse of dimensionality’ caused by uncertainties, an approximate dynamic programming (ADP) approach with post-decision states and iteration is employed to solve the proposed model. With the estimated values of post-decision states, the stochastic problem with sequential multi-period stochastic optimization problem is transformed into a one-period deterministic problem. Case studies demonstrate that the state-based strategies try to make the feeders on the trajectory locate the terminal of the whole network to reduce potential loss of load, and in consequence to improve system resilience.

REFERENCES

- [1] Executive Office of the President, USA, “Economic benefits of increasing electric grid resilience to weather outages,” Aug. 2013. [Online]. Available: https://www.energy.gov/sites/prod/files/2013/08/f2/20130808%20Resiliency%20Report_FINAL.pdf
- [2] Electric Power Research Institute, “Enhancing distribution resiliency: Opportunities for applying innovative technologies.” [Online]. Available: <http://www2.epri.com/abstracts/Pages/ProductAbstract.aspx?ProductId=00000000001026889>

- [3] North American Electric Reliability Corporation, “Severe impact resilience: Considerations and recommendations.” [Online]. Available: http://www.nerc.com/comm/OC/SIRTF%20Related%20Files%20DL/SIRTF_Final_May_9_2012-Board_Accepted.pdf
- [4] M. Panteli and P. Mancarella, “The grid: Stronger, bigger, smarter?: Presenting a conceptual framework of power system resilience,” *IEEE Power and Energy Magazine*, vol. 13, no. 3, pp. 58–66, May 2015.
- [5] National Academy of Sciences USA, *National Research Council: Disaster Resilience: A National Imperative*, Washington, DC, USA: National Academies Press, 2012.
- [6] H. Liu, R. A. Davidson, J. R. Stedinger, D. V. Rosowsky, and D. V. Rosowsky, “Negative binomial regression of electric power outages in hurricanes,” *J. Infrastructure Syst.*, vol. 11, no. 4, pp. 258–267, Dec. 2005.
- [7] S. D. Guikema, R. Nateghi, S. M. Quiring, A. Staid, A. C. Reilly, and M. Gao, “Predicting hurricane power outages to support storm response planning,” *IEEE Access*, vol. 2, pp. 1364–1373, Nov. 2014.
- [8] R. Nateghi, S. Guikema, and S. Quiring, “Forecasting hurricane-induced power outage durations,” *Natural Hazards*, vol. 74, no. 3, p. 17951811, Dec. 2014.
- [9] C. Wang, Y. Hou, Z. Qin, C. Peng, and H. Zhou, “Dynamic coordinated condition-based maintenance for multiple components with external conditions,” *IEEE Trans. Power Del.*, vol. 30, no. 5, pp. 2362–2370, Oct. 2015.
- [10] W. Yuan, J. Wang, F. Qiu, C. Chen, C. Kang, and B. Zeng, “Robust optimization-based resilient distribution network planning against natural disasters,” *IEEE Trans. Power Syst.*, vol. 7, no. 6, pp. 2817–2826, Nov. 2016.
- [11] W. Sun, C. C. Liu, and S. Liu, “Black start capability assessment in power system restoration,” in *IEEE Power and Energy Society General Meeting*, Jul. 2011, pp. 1–7.
- [12] F. Qiu, J. Wang, C. Chen, and J. Tong, “Optimal black start resource allocation,” *IEEE Trans. Power Syst.*, vol. 31, no. 3, pp. 2493–2494, May 2016.
- [13] M. Zare-Bahramabadi, A. Abbaspour, M. Fotuhi-Firuzabad, and M. Moeini-Aghaie, “Resilience-based framework for switch placement problem in power distribution systems,” *IET Generation, Transmission Distribution*, vol. 12, no. 5, pp. 1223–1230, 2018.
- [14] S. Zhang and V. Vittal, “Wide-area control resiliency using redundant communication paths,” *IEEE Trans. Power Syst.*, vol. 29, no. 5, pp. 2189–2199, Sep. 2014.
- [15] M. M. Adibi and L. H. Fink, “Overcoming restoration challenges associated with major power system disturbances - restoration from cascading failures,” *IEEE Power and Energy Magazine*, vol. 4, no. 5, pp. 68–77, Sept. 2006.
- [16] Y. Hou, C. C. Liu, K. Sun, P. Zhang, S. Liu, and D. Mizumura, “Computation of milestones for decision support during system restoration,” *IEEE Trans. Power Syst.*, vol. 26, no. 3, pp. 1399–1409, Aug. 2011.
- [17] W. Sun, C. C. Liu, and L. Zhang, “Optimal generator start-up strategy for bulk power system restoration,” *IEEE Trans. Power Syst.*, vol. 26, no. 3, pp. 1357–1366, Aug. 2011.
- [18] C. C. Liu, S. J. Lee, and S. S. Venkata, “An expert system operational aid for restoration and loss reduction of distribution systems,” *IEEE Trans. Power Syst.*, vol. 3, no. 2, pp. 619–626, May 1988.
- [19] S. Toune, H. Fudo, T. Genji, Y. Fukuyama, and Y. Nakanishi, “Comparative study of modern heuristic algorithms to service restoration in distribution systems,” *IEEE Trans. Power Del.*, vol. 17, no. 1, pp. 173–181, Jan. 2002.
- [20] A. Castillo, “Microgrid provision of blackstart in disaster recovery for power system restoration,” in *IEEE International Conference on Smart Grid Communications*, Oct. 2013, pp. 534–539.
- [21] F. Ren, M. Zhang, D. Soetanto, and X. Su, “Conceptual design of a multi-agent system for interconnected power systems restoration,” *IEEE Trans. Power Syst.*, vol. 27, no. 2, pp. 732–740, May 2012.
- [22] J. R. Birge and F. Louveaux, *Introduction to Stochastic Programming*, Springer New York, 2011.
- [23] E. Byon and Y. Ding, “Season-dependent condition-based maintenance for a wind turbine using a partially observed markov decision process,” *IEEE Trans. Power Syst.*, vol. 25, no. 4, pp. 1823–1834, Nov. 2010.
- [24] S. K. Abeygunawardane, P. Jirutitijaroen, and H. Xu, “Adaptive maintenance policies for aging devices using a markov decision process,” *IEEE Trans. Power Syst.*, vol. 28, no. 3, pp. 3194–3203, Aug. 2013.
- [25] C. Wang, Y. Hou, F. Qiu, S. Lei, and K. Liu, “Resilience enhancement with sequentially proactive operation strategies,” *IEEE Trans. Power Syst.*, vol. 32, no. 4, pp. 2847–2857, Jul. 2017.
- [26] M. Ouyang and L. Dueas-Osorio, “Multi-dimensional hurricane resilience assessment of electric power systems,” *Struct. Safety*, vol. 48, p. 1524, May 2014.
- [27] M. E. Baran and F. F. Wu, “Network reconfiguration in distribution systems for loss reduction and load balancing,” *IEEE Trans. Power Del.*, vol. 4, no. 2, pp. 1401–1407, Apr. 1989.
- [28] J. A. Taylor and F. S. Hover, “Convex models of distribution system reconfiguration,” *IEEE Trans. Power Syst.*, vol. 27, no. 3, pp. 1407–1413, Aug. 2012.
- [29] X. Chen, W. Wu, and B. Zhang, “Robust restoration method for active distribution networks,” *IEEE Trans. Power Syst.*, vol. 31, no. 5, pp. 4005–4015, Sept. 2016.