Uncertainty-Inflicted Event-Driven Resilient Recovery for Distribution Systems: A Semi-Markov Decision Process Approach

Chong Wang, *Member, IEEE*, Gengfeng Li, *Member, IEEE*, Can Wan, *Member, IEEE*, Zhaoyu Wang, *Member, IEEE*, Ping Ju, *Senior Member, IEEE*, Shunbo Lei, *Member, IEEE*

***Abstract*—Repair and reconfiguration are vital for power recovery after outages caused by natural disasters in distribution systems, but sequential and uncertainty-inflicted decision points due to uncertain repair periods make power recovery complicated. This paper proposes semi-Markov decision process(SMDP)-based resilient recovery with sequentially event-driven repair and reconfiguration in consideration of uncertainty-inflicted decision-making points. The sequential repair/reconfiguration actions in consideration of uncertain repair periods are considered as uncertainty-inflicted event-driven processes. The sequential repair states with different repair crews are established as semi-Markov states. The whole sequential and uncertain decision-making process is modeled as a semi-Markov decision process-based optimization model, which is an event-driven recursive model. Q-learning is employed to solve the proposed model, and the convergent estimations of Q values for semi-Markov states map the original model into an event-driven deterministic optimization based on the sequential repairs that actually occurred over the time horizon. IEEE 123-bus system is used to validate the proposed model.**

***Index Terms*—repair, resilient recovery, semi-Markov decision process, uncertain decision-making**

NOMENCLATURE

Indices and Sets

e, e'	Index of semi-Markov states indicating different event e and e'.
i	Index of repair crews.
j	Index of damaged components.
l	Index of lines.
k, k'	Index of terminal buses of line l.
g	Index of optimization variables related to semi-Markov states.
$\mathcal{A}$	Set of actions.
$\mathcal{B}$	Set of power buses.
$\tilde{\mathcal{B}}$	Set of substation nodes.
$\mathcal{E}$	Set of semi-Markov states.
$\mathcal{L}$	Set of lines.
$\mathcal{L}'$	Set of dispatchable lines.
$\mathcal{N}_{ke}$	Set of power buses connected to power bus k under semi-Markov state S_e.
Ω_{cr}	Set of repair crews.
Ω_{fc}	Set of damaged components.

Notations for semi-Markov Decision Process-based model

S_e, S'_e	Semi-Markov states.
a_e	Actions for semi-Markov state S_e.
$v_e, v_{e'}$	Value functions for semi-Markov state S_e and $S_{e'}$.
C_e	Immediate cost with semi-Markov state S_e under action a_e.
$P^a_{e'e}$	Probability from semi-Markov state S_e to $S_{e'}$ under action a_e.
$\Delta L_{e'}$	Loss of load with semi-Markov state $S_{e'}$.
$\Delta L^{n+1}_{e'}$	Loss of load with semi-Markov state $S_{e'}$ on the $(n+1)^{th}$ stochastic path.
τ_1, τ_2	Start and end time for one event including the repair and the corresponding transportation.
$Q^*(S_{e'})$	Q value for semi-Markov state $S_{e'}$.
$Q^{*n}(S_{e'})$	Q value for semi-Markov state $S_{e'}$ at the n^{th} iteration.
$Q^{*n}_{e'}$	Estimated Q value for semi-Markov state $S_{e'}$ at the n^{th} iteration.

Parameters

η	Coefficient for loss of load.
c_l, c'_l, c''_l	Coefficient for operational cost, switch-on, and switch-off costs of lines.
$P^{max}_{ke}, Q^{max}_{ke}$	Maximum active load and reactive load at bus k with semi-Markov state S_e.
$r_{kk'}, x_{kk'}$	Resistance/reactance of line $k-k'$.
γ_{le}	Given value representing failure state of line l with semi-Markov state S_e.
M	A large number.
$\overline{S}_{kk'}$	Apparent power capacity of line $k-k'$.
κ_j	A given value representing whether the component j is repaired.
λ	Dynamic learning rate, and it is the mean value of slopes of Q value curves at the n^{th} iteration

This work was supported by the National Natural Science Foundation of China (52277088).

C. Wang, and P. Ju are with the School of Eletrical and Power Engineering, Hohai University, Nanjing 211100, China (e-mail: chongwang@hhu.edu.cn, pju@hhu.edu.cn).

G. Li is with the State Key Laboratory of Electrical Insulation and Power Equipment, Smart Grid Key Laboratory of Shaanxi Province, Department of Electrical Engineering, Xi'an Jiaotong University, Xi'an 710049, China (e-mail: gengfengli@mail.xjtu.edu.cn)

C. Wan is with the College of Electrical Engineering, Zhejiang University, Hangzhou 310027, China (e-mail: canwan@zju.edu.cn).

Z. Wang is with the Department of Electrical and Computer Engineering,Iowa State University, Ames, IA 50011, USA (email: wzy@iastate.edu).

S. Lei is with the School of Science and Engineering, The Chinese University of Hong Kong–Shenzhen, Guangdong 518172, China, and with the Shenzhen Institute of Artificial Intelligence and Robotics for Society, Guangdong 518129, China (email: leishunbo@cuhk.edu.cn).

Variables

$x_{le}, x_{le'}$	Binary variables representing on-off states of line l under semi-Markov state S_e and $S_{e'}$, respectively. 1 denotes on state, and 0 denotes off state.
$x'_{le'}, x''_{le'}$	Binary variables representing switch-on and switch-off actions, respectively.
$y_{ije'}$	Binary variables representing the repair state of the j^{th} damaged component with the i^{th} repair crew.
$o_{kk'e}$	Binary variables, the value is 1 if bus k' is the parent bus of bus k with semi-Markov state S_e.
P_{ke}, Q_{ke}	Active load and reactive load at bus k with semi-Markov state S_e.
$P_{kk'e}, Q_{kk'e}$	Active and reactive power through line $k-k'$ with semi-Markov state S_e.
U_{ke}	Squared voltage magnitude of bus k with semi-Markov state S_e.
$\beta_{ge'}$	Semi-Markov state-related variables including $x_{le'}$, $\forall l$ and $y_{ije'}$, $\forall i, j$.
$\varsigma_{ge'}$	Variables related to $\beta_{ge'}$.

I. Introduction

COMMERCIAL and residential customers are directly connected to radial distribution systems, and any failures due to damaged components caused by natural disasters or cyber attacks will result in massive loss of loads [1]. Many organizations, e.g., the North American Electric Reliability Corporation [2] and the United States Electric Power Research Institute [3], have released reports emphasizing that energy systems should have characteristics of resilience against natural disasters or cyber attacks. As suggested, many preventive and adaptative strategies are investigated to improve power system resilience from the perspectives of network hardening [4], unfolding strategies [5], and cyber security [6]. For example, a two-stage stochastic mixed integer model, in which the first stage aims to obtain the optimal hardening scheme and the second stage includes the potential damages, is investigated in [4] to improve the resilience against extreme weather events. In [7], a state-based decision-making model is proposed to increase system resilience with the objective of minimizing loss of loads during an unfolding sequential event. In [8], a two-step optimization model in consideration of predisaster preparedness and post-disaster resource re-allocation is constructed to improve distribution system resilience.

Although many existing preventive and adaptative strategies are investigated to improve power system resilience [9], damaged components caused by natural disasters are inevitable and in consequence power outages due to damaged components are also unavoidable. To achieve system recovery as quickly as possible, there have been many research studies on system recovery from the perspectives of system reconfiguration [10], microgrid formulation [11], etc. For example, a stochastic optimization model for resource preparation is investigated in [12] to preposition crews and equipment aiming to achieve faster post-disaster deployment of equipment resources and crews to damage locations. An integrated model scheduling mobile battery-carried vehicles and networked microgrids is investigated in [13] to restore multiple outages caused by natural disasters, and the computational complexity of the proposed model is reduced by the auxiliary induce function-based approach. Based on multiple microgrid formulation, [14] proposes a hierarchical outage management scheme to deal with power outages, and [15] investigates a heuristic approach-based microgrid formation to address computational intractability. When forming microgrids after power outages, distributed generators have great impacts on microgrid operation. [16] investigates dynamic characteristics of distributed generators, and the corresponding influences are integrated into the recovery model. During the distribution system recovery, radiality is one of important constraints and the single-commodity flow constraints are presented in [17] to address the limitation of spanning tree constraints. In consideration of the couplings between different systems, the poster recovery on the interdependent power and natural gas system is investigated in [18], and the coordinated recovery model for transmission and distribution systems is established in [19]. A new metric, defined as the number of recovered customers divided by the average outage time of the interrupted customers, is developed in [20] to evaluate the resilience enhancement strategies. With the development of intelligence algorithms, the optimal equipment repair sequence can be searched by means of the heuristic algorithm [21], and machine learning techniques can be used to train an agent with thousands of off-line scenarios to obtain the optimal recovery strategy [22]. In addition, communication systems in the cyber-physical system have also great impacts on power system resilience. A hierarchical two-stage robust optimization is investigated in [23] to mitigate the impacts of false data injection on the integration of electricity and gas systems, increasing the integrated system resilience. In consideration of emergency communication vehicles, an integrated recovery model coordinating physical systems and cyber systems is investigated to restore unserved load as soon as possible [24]. A two-stage sequential disaster recovery model is established in [25] to collaborate cyber-physical optimization.

To achieve a fast recover, repairs on damaged components are critical. In [26], the post-disaster repair process is established as a scheduling problem with a set of newly-defined soft precedence constraints, and a polynomial-time algorithm in consideration of constant performance guarantees is employed to solve the proposed model. In [27], the model of dispatching maintenance/restoration crews is integrated into the disaster recovery model that scheduling the action sequence for controllable switches and controllable load. In [28], a co-optimization model is investigated to schedule repair crews and mobile power sources in transportation networks and distribution systems in consideration of different timescales of distribution system restoration, and one preprocessing approach, which cuts down the number of candidate nodes for connecting mobile power sources and preassigns a minimal set of repair tasks to depots, is used to reduce computational complexity. In [29], a mixed-integer linear programming model is constructed to co-optimize repair crews and resources to restore unserved loads after power outages caused by natural disasters, and an iterative neighborhood search approach is used to solve the proposed model.

In the real world, the repairs on the damaged components depend on many factors, e.g., transportation conditions, damage conditions, and external environment. Theses factors result in uncertain periods of sequential repairs, and uncertain repair periods cause a consequence that the repair decisions cannot be made at the specified and fixed time. This kind of

decision-making process at uncertain decision time points due to uncertain sequential repairs is different from the decision-making process at the given decision time points, and needs to be investigated. The semi-Markov decision process is an effective framework for the sequential decision-making process at uncertain decision time points. Currently, the semi-Markov decision process has been used in maintenance scheduling [30], reliability evaluation [31], availability analysis [32], etc. In [33], a semi-Markov decision process-based optimization model, estimating conditional reliability under varying conditions, is established to obtain the maintenance scheme. In consideration of random holding time of system states, [34] investigates power-plant reliability analysis by means of semi-Markov processes. With the massive deployment of cyber technologies, the semi-Markov models represent the intrusion process of measurement manipulation [35] and penetration attacks [36] to evaluate power system reliability. The semi-Markov decision process has not been used in system recovery in consideration of the decision-making process at uncertain decision time points. Therefore, a semi-Markov decision process(SMDP)-based resilient recovery model is proposed to investigate the decision-making process at uncertain decision time points due to uncertain sequential repairs. The contributions of the paper are listed as follows. 1) An uncertainty-inflicted event-driven process is used to describe sequential repair/reconfiguration problem in consideration of uncertain decision points due to uncertain repair periods. 2) The uncertainty-inflicted event-driven process is established as a semi-Markov decision process-based model, which is an event-driven recursive model. 3) To address the dimensionality curse of the problem, Q-learning with the dynamic learning rate is used to estimate the expected values of semi-Markov states by means of off-line iterations, and the converged estimations of Q values for semi-Markov states map the original model into an event-driven deterministic optimization. IEEE 123-bus system is used to validate the proposed model, and the results show that the proposed model and the algorithm have a good performance on sequentially event-driven repair/reconfiguration in consideration of uncertainty-inflicted decision-making points.

The remainder of this paper is organized as follows. Section II shows the uncertainty-inflicted event-driven process. Section III presents the semi-Markov decision process-based event-driven optimization, and section IV introduces Q-learning solution and model reformulation. The case studies and the conclusions are demonstrated in Section V and Section VI.

II. Uncertainty-Inflicted Event-Driven Process

A. Description of Uncertainty-Inflicted Event-Driven Process

In the real world, many events occur sequentially one by one after given actions and the duration of each event is uncertain. In this case, the decision time point of each action is also uncertain. This kind of decision-making process at uncertain decision time points due to uncertain sequential events, different from the decision-making process at the given decision time points, is defined as an uncertainty-inflicted event-driven process in our study. Take the repair action in the distribution system as an example. After component failures caused by natural disasters in the distribution system, system repair needs to recover the system to the normal state as quickly as possible. Due to limits of repair resources and system operational constraints, system repair is a sequential process. In addition, component repair depends on transportation conditions, damage conditions, external environment, etc. These make component repair time uncertain, making each repair become an uncertain event. Because repairs on different failure components are implemented sequentially, uncertain repair periods and sequential implementation make the entire recovery become an uncertainty-inflicted event-driven process. Take the recovery process in Fig. 1 as an example. The lines 1-2, 3-4, 6-7, and 9-10 are in failure due to natural disasters, and in consequence power load at the buses 2, 3, 4, 5, 6, 7, 8, and 9 are disconnected. The line 7-9 has a switch for system reconfiguration. To recover the system and reduce the loss of loads, the following recovery processes can be implemented.

- *Process* 1: Dispatch Crew 1 to repair the lines 6-7 (the repair r_1) and 3-4 (the repair r_2) sequentially. Dispatch Crew 2 to repair the line 9-10 (the repair r_3), and meanwhile connect the line 7-9 by means of the switch.
- *Process* 2: Dispatch Crew 1 to repair the line 6-7 (the repair r_1). Dispatch Crew 2 to repair the lines 9-10 (the repair r_3) and 3-4 (the repair r_2) sequentially, and meanwhile connect the line 7-9 by the switch.

For the above two recovery processes, the repair r_2 is implemented after the repair r_1 implemented by the crew 1 or the repair r_3 implemented by the crew 2. Due to uncertain repair periods, it is not sure that which of r_1 and r_3 is first completed due to sequential characteristics. If the repair r_1 is first completed, we could implement *Process* 1. If the repair r_3 is first completed, we could implement *Process* 2. So, it indicates that the implementation of the repair r_2 depends on the previous uncertain event, and we define this as an uncertainty-inflicted event-driven process. This also demonstrates that an optimal strategy needs to be made based on observed repairs that actually occurred. In this study, the semi-Markov decision process is employed to describe and model the uncertainty-inflicted event-driven process.

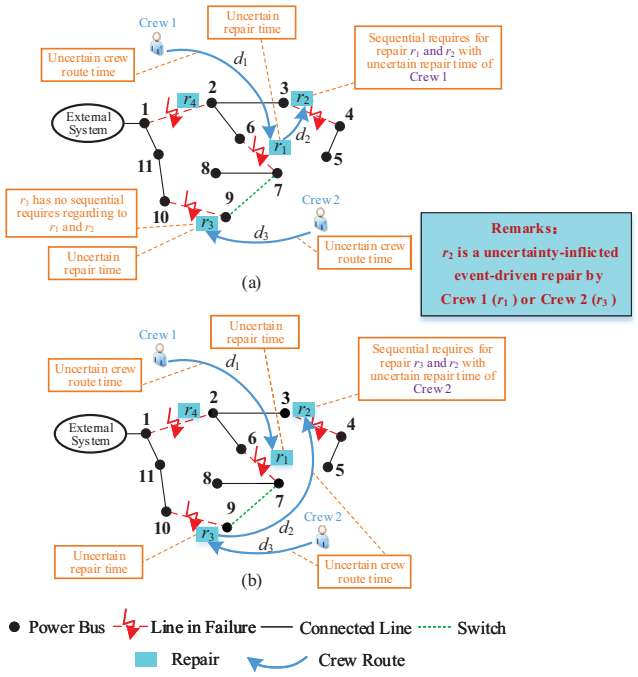


Fig. 1. Uncertainty-inflicted event-driven process.

B. Modeling of Uncertainty-Inflicted Event-Driven Process

Fig. 2 shows the time horizon for the uncertainty-inflicted event-driven process. In Fig. 2 (a), r_1 denotes the time period to repair the line 6-7, and d_1 is the time period to route the crew 1 to the location of the line 6-7. We define the state in this period as a state r_{11}^D, as shown in Fig. 2 (b). During the state r_{11}^D, the crew 1 cannot be re-dispatched. Only when the repair is finished, the crew 1 can be re-dispatched again. Due to uncertain repair periods, the decision points for re-dispatching crews are event-based and uncertain. When there are multiple groups of crews, the scenarios will become more complicated. The decision points of the system are Cartesian products of each decision point of different crews in consideration of their uncertain repair periods and sequential interdependence. For example, decisions should be made in the state $\{r_{11}^E, r_{23}^D, \Pi_t\}$ and $\{r_{12}^D, r_{23}^E, \Pi_t\}$. The superscript '$D$' denotes 'during one scheduled action', indicating no other actions can be implemented by the corresponding crew, and the superscript 'E' denotes 'the scheduled action is completed', indicating new actions can be implemented by the corresponding crew. Π_t represents the system topology at time t. r_{11}^E denotes that the repair implemented by the first crew on the 1^{st} failure component is finished. r_{23}^D denotes the 3^{rd} failure component is being repaired by the second crew. r_{12}^D denotes that the 2^{nd} failure component is being repaired by the first crew . r_{23}^E denotes the repair implemented by the second crew on the 3^{rd} failure component is finished.

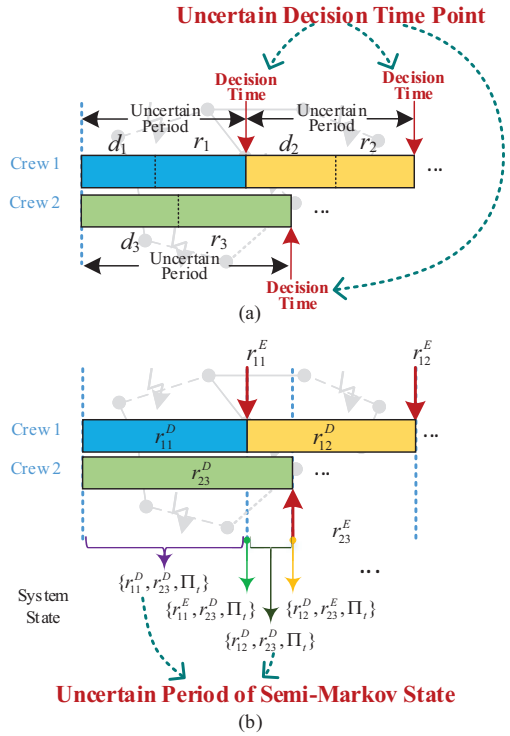


Fig. 2. Semi-Markov states for uncertainty-inflicted event-driven process.

The whole decision process is an event-driven process, and the event has an uncertain time period. There are two difficulties for scheduling repair/system. The first one is that the decision points for dispatching repair crews are uncertain, and the second one is that different repairs have sequential interdependence. In this study, semi-Markov decision process is used to describe the uncertainty-inflicted event-driven process.

The semi-Markov decision process (SMDP) can be expressed as a four-tuple $(\mathcal{S}, \mathcal{A}, \mathcal{P}, \mathcal{R})$, in which $\mathcal{S}$ and $\mathcal{A}$ represent the semi-Markov state space and the action space, respectively. $\mathcal{P}$ is the transition probability after actions under a state. $\mathcal{R}$ is the reward after actions. A semi-Markov decision process is an extension of a Markov decision process that includes an agent that makes decisions that affect the evolution of the system over time. For the semi-Markov decision process, one critical difference from the Markov decision process is that the duration of each state is stochastic. Fig. 3 illustrates the semi-Markov decision process. When observing the state S_1', one action A_1 is performed, resulting a reward R_1. After a stochastic duration T_1, the state S_1' reaches the S_2' with the probability P_{12} in consideration of environment uncertainties. Then, this process is repeated for other decision points. The investigated uncertainty-inflicted event-driven process can be mapped into a semi-Markov decision process as follows.

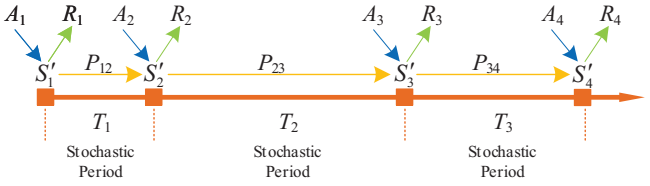


Fig. 3. Semi-Markov decision process.

- $\mathcal{S}$: The semi-Markov states include the system topology and the repair state for each group of crews. Each group of crews is dispatched as a whole, and there may be several groups who can be dispatched simultaneously.
- $\mathcal{A}$: The actions include repair dispatch and system reconfiguration by means of controlling line switches. It is assumed that the implementation time of controlling line switches is ignored, and system reconfiguration will be implemented at the initial time or the time when one failure component is repaired. Repair dispatch can only be implemented when the corresponding crews are unoccupied.
- $\mathcal{P}$: After re-dispatching the crews and implementing system reconfiguration, the next state is deterministic. This indicates that the transition probability from one state to another state with actions is 1.
- $\mathcal{R}$: The reward includes a discrete state-action dependent reward and a time-continuous reward with regard to uncertain repair time periods and future states.

The semi-Markov state S_e for the whole decision process can be expressed as follows.

$$S_e = (s_{1e}, \cdots, s_{ie}, \cdots, s_{Ie}, \Pi_e, \mathcal{L}_e) \quad \forall t \tag{1}$$

where S_e represents the system state, Π_e is the state of the system topology in the system state S_e. $\mathcal{L}_e$ denotes the system load level in the system state S_e. s_{ie} in (1) denotes the state of the ith crew in the system state S_e, and it can be expressed as follows.

$$s_{ie} \in \{r_{i1}^D, r_{i1}^E, \cdots, r_{ij}^D, r_{ij}^E, \cdots\} \; \forall i \in \Omega_{cr}, j \in \Omega_{fc} \tag{2}$$

where r_{ij}^D denotes that the j^{th} failure component is being repaired by the i^{th} crew. r_{ij}^E denotes that the repair implemented by the i^{th} crew on the j^{th} failure component is finished.

It is assumed that the load at each bus is considered as a block and the load at one bus would be connected to the grid when the power supply path is effective. This indicates that one system topology scenario has the corresponding supplied load level when having one given loading factor. Take the scenario in Fig. 2 as an example. If the connected line set is {1-11, 11-10}, the loads at the buses 1, 11, and 10 will be connected to the grid, and the loads at these three buses cannot be partially dispatched. If the connected line set is {1-11, 11-10, 1-2, 2-3, 2-6}, the loads at the buses 1, 10, 11, 2, 3, and 6 will be connected to the grid. However, different loading factors over the time horizon result in different power load even for the same system topology. Therefore, the system load level $\mathcal{L}_e$ in the semi-Markov state S_e is used to represent different loading factors. Because the loading factor is a continuous function with regard to time, the uncertain period of each repair corresponds to a continuous load level, resulting in infinite continuous states, as shown in Fig. 4 (a). The infinite continuous states make the problem difficult to be solved, so discrete states of the load level as shown Fig. 4 (b) are deployed in this study. The time step of discretizing the load level can be set to be 15 minutes or 30 minutes. The value of discretizing the load level can be set as the average value of load in each interval.

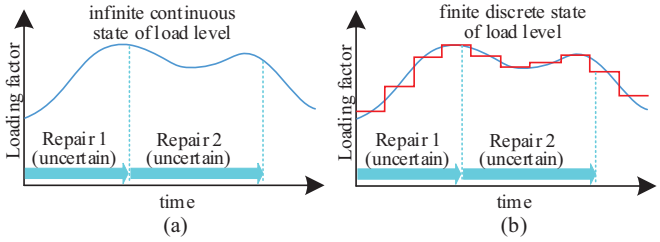


Fig. 4. (a) Infinite continuous state and (b) finite discrete state of load level under uncertain periods of different repairs.

III. Semi-Markov Decision Process-based Event-Driven Optimization

A. Event-Driven Recursive Model for Semi-Markov States

Section II clarifies that the entire decision process of repair and reconfiguration is uncertainty-inflicted and event-driven. Due to the event-driven characteristics, it is expected to construct an event-driven rather than time-driven optimization as follows.

$$v_e(S_e) = \arg\min_{a_e \in \mathcal{A}} \left\{ \begin{array}{l} C_e(S_e, a_e) + \\ \sum\limits_{e'} P^a_{e'e} \cdot \left[\int_{\tau_1}^{\tau_2} (\eta \cdot \Delta L_e) dt + v_{e'}(S_{e'}) \right] \end{array} \right\} \tag{3}$$

where $v_e(S_e)$ and $v_{e'}(S_{e'})$ represent the expected total rewards for the semi-Markov states S_e and $S_{e'}$, and $S_{e'}$ is the triggered state after the semi-Markov state S_e under the action a_e. This indicates the event-driven process. $P^a_{e'e}$ indicates the transition probability from one state to another state under one action. Here, 'one state' represents that the system is being repaired. Once one repair is finished, the next state is deterministic after re-dispatching the crews. This indicates that the transition probability from one state to another state under one action is 1, i.e., the value of $P^a_{e'e}$ is 1. τ_1 and τ_2 are the beginning and ending times of the semi-Markov states $S_{e'}$, and these two times are uncertain over the time horizon. The uncertain times indicate the sequential interdependence and uncertain repair time periods. Therefore, the optimization objective (3) can be rewritten as follows.

$$v_e(S_e) = \arg\min_{a_e \in \mathcal{A}} \left\{ \begin{array}{l} C_e(S_e, a_e) + \\ \left[\int_{\tau_1}^{\tau_2} (\eta \cdot \Delta L_e) dt + v_{e'}(S_{e'}) \right] \end{array} \right\} \tag{4}$$

where the integral term $\int_{\tau 1}^{\tau 2} (\eta \cdot \Delta L_e) dt$ represents the cost of loss of loads. Due to uncertain times, the cost of loss of loads is uncertain, and this leads to difficulties in solving the optimization model. The function $C_e(S_e, a_e)$ represents the immediate cost when implementing the action a_e in the semi-Markov state S_e, and this cost includes the operational cost of dispatchable lines and the cost for line switching. It can be expressed as follows.

$$C_e(S_e, a_e) = \sum_{l \in \mathcal{L}'} (c_l \cdot x_{le'} + c'_l \cdot x'_{le'} + c''_l \cdot x''_{le'}) \tag{5}$$

where the first term is the operational cost of dispatchable lines, the second and the third terms represent the costs of connecting and disconnecting dispatchable lines.

B. Operational Constraints for Semi-Markov States

When the system reaches $S_{e'}$ after the action a_e, the system operational constraints, e.g., radiality for non-islanding buses, power flow, power balance, line capability, and voltage limits, should be satisfied.

1) Radiality constraint: The distribution system needs to operate in a radial topology. Because some components, e.g., distribution lines, are in failure caused by extreme events, some buses may be in island states and these islanded buses cannot be connected to the grid before the fault components are repaired. In consideration of fault components and system reconfiguration, the spanning tree constraints are used to guarantee the network radiality.

$$x_{le} = o_{kk'e} + o_{k'ke} \quad l = (k, k') \in \mathcal{L}, e \in \mathcal{E} \tag{6a}$$

$$\sum_{k' \in \mathcal{N}_{ke}} o_{kk'e} = 1 \quad k \in \mathcal{B}, e \in \mathcal{E} \tag{6b}$$

$$o_{kk'e} = 0 \quad k \in \tilde{\mathcal{B}},\ k' \in \mathcal{N}_{ke},\ e \in \mathcal{E} \tag{6c}$$

$$P_{kk'e} = 0 \quad (k, k') \in \Omega_{fc},\ e \in \mathcal{E} \tag{6d}$$

$$Q_{kk'e} = 0 \quad (k, k') \in \Omega_{fc},\ e \in \mathcal{E} \tag{6e}$$

where (6a) and (6b) show that two terminals of one connected distribution line have one parent bus under the semi-Markov state S_e. In practice, the bus connected to the external system or the substation bus has no parent bus, and this is constrained by (6c). In addition, the directed multicommodity flow-based models of the spanning tree constraints [37] are used as supplementary constraints to ensure system radiality.

Because the loads connected to islanded buses cannot be supplied by power source, (6a)-(6c) cannot include this scenario. We use the constraint (6d) to represent the island scenario. The constraints (6d) and (6e) denote that the power flows through the fault lines are set to be 0.

2) Line switching constraints: From one semi-Markov state to another semi-Markov state, the following constraints regarding line switching need to be satisfied.

$$x'_{le'} \geq x_{le'} - x_{le} \tag{7a}$$
$$x'_{le'} \leq x_{le'} \tag{7b}$$
$$x'_{le'} \leq 1 - x_{le} \tag{7c}$$
$$x''_{le'} \geq x_{le} - x_{le'} \tag{7d}$$
$$x''_{le'} \leq 1 - x_{le'} \tag{7e}$$
$$x''_{le'} \leq x_{le} \tag{7f}$$

where (7a)-(7c) represent the switch-on constraints, and (7d)-(7f) represent the switch-off constraints.

3) Power balance constraint: For each semi-Markov state, the in-flow and out-flow of each bus should be equal.

$$P_{ke} + \sum_{k' \in \mathcal{N}_{ke}} P_{kk'e} = 0 \qquad k \in \mathcal{B}, e \in \mathcal{E} \tag{8a}$$
$$Q_{ke} + \sum_{k' \in \mathcal{N}_{ke}} Q_{kk'e} = 0 \qquad k \in \mathcal{B}, e \in \mathcal{E} \tag{8b}$$
$$0 \leq P_{ke} \leq P_{ke}^{\max} \qquad k \in \mathcal{B}, e \in \mathcal{E} \tag{8c}$$
$$0 \leq Q_{ke} \leq Q_{ke}^{\max} \qquad k \in \mathcal{B}, e \in \mathcal{E} \tag{8d}$$
$$\Delta L_e = \sum_k (P_{ke}^{\max} - P_{ke}) \qquad e \in \mathcal{E} \tag{8e}$$

where (8a) and (8b) represent real power balance and reactive power balance, respectively. Because (6d) and (6e) constrain the power flow through the fault lines as 0, it indicates that the loads at the islanded buses are disconnected. (8c), (8d), and (8e) are the load constraints at bus k with the semi-Markov state S_e.

4) Power flow constraint: For each semi-Markov state, the power flow through each line is a function with regard to terminal bus voltages of each line. Distflow is used to describe this expression as follows.

$$U_{ke} - U_{k'e} \leq (2 - x_{le} - \gamma_{le}) \cdot M + 2(r_{kk'} \cdot P_{kk'e} + x_{kk'} \cdot Q_{kk'e}) \quad l = (k, k') \in \mathcal{L}, e \in \mathcal{E} \tag{9a}$$

$$U_{ke} - U_{k'e} \geq (\gamma_{le} - x_{le} - 2) \cdot M + 2(r_{kk'} \cdot P_{kk'e} + x_{kk'} \cdot Q_{kk'e}) \quad l = (k, k') \in \mathcal{L}, e \in \mathcal{E} \tag{9b}$$

where (9a) and (9b) are the DsitFlow model [38], and the quadratic terms are ignored. γ_{le} denotes the failure state of the line l, and it is a known parameter for the semi-Markov state S_e. $\gamma_{le} = 0$ means that the line l is in failure due to extreme events, and $\gamma_{le} = 1$ means that the line l is in normal state. A sufficiently large M makes (9a)-(9b) redundant when the lines are in failure or disconnected. Associated with (6), (8) and (9), non-islanded buses with the radiality constraint and islanded buses are both included in the model.

5) Line capacity constraint: The power flow through each line should be within the limits in each semi-Markov state.

$$(P_{kk'e})^2 + (Q_{kk'e})^2 \leq x_{le} \cdot \gamma_{le} \cdot (\overline{S}_{kk'})^2 \qquad l = (k, k') \in \mathcal{L}, e \in \mathcal{E} \tag{10}$$

where γ_{le} is a known parameter for the semi-Markov state S_e. Due to the quadratic constraints with regard to the active power and reactive power, (10) is a nonlinear constraint. A relaxed expression is used as follows.

$$-x_{le} \cdot \gamma_{le} \cdot \overline{S}_{kk'} \leq P_{kk'e} \leq x_{le} \cdot \gamma_{le} \cdot \overline{S}_{kk'} \quad \forall l, e \tag{11a}$$
$$-x_{le} \cdot \gamma_{le} \cdot \overline{S}_{kk'} \leq Q_{kk'e} \leq x_{le} \cdot \gamma_{le} \cdot \overline{S}_{kk'} \quad \forall l, e \tag{11b}$$
$$-\sqrt{2} x_{le} \cdot \gamma_{le} \cdot \overline{S}_{kk'} \leq P_{kk'e} + Q_{kk'e} \leq \sqrt{2} x_{le} \cdot \gamma_{le} \cdot \overline{S}_{kk'} \quad \forall l, e \tag{11c}$$
$$-\sqrt{2} x_{le} \cdot \gamma_{le} \cdot \overline{S}_{kk'} \leq P_{kk'e} - Q_{kk',e} \leq \sqrt{2} x_{le} \cdot \gamma_{le} \cdot \overline{S}_{kk'} \quad \forall l, e \tag{11d}$$

where the above constraints are linear inequality constraints because γ_{le} is a known parameter for the semi-Markov state S_e.

6) Voltage constraint: The voltage limits should be satisfied when having the semi-Markov state S_e.

$$\underline{V}_k^2 \leq U_{ke} \leq \overline{V}_k^2 \quad k \in \mathcal{B}, e \in \mathcal{E} \tag{12}$$

where (12) limits the upper and lower voltage bounds of each bus at each semi-Markov state.

IV. Q-learning Solution and Model Reformulation

This section first presents the difficulty in solving the proposed model, and then introduces how to use Q-learning approach to solve the model. Third, model reformulation with regard to semi-Markov states and optimization variables in consideration of Q-learning approach is presented, and finally the solution process is introduced.

A. Difficulty of Model Solution

There are three difficulties in solving the proposed model, and they are listed as follows.

- Recursive optimization model: The optimization model (4) shows that it is a recursive model with regard to different semi-Markov states. The recursive model represents the sequential characteristics of different semi-Markov states under different actions. A large number of semi-Markov states and actions result in curse of dimensionality.
- Uncertainty-inflicted event-driven process: In (4), τ_1 and τ_2 are the beginning and ending times of the semi-Markov states $S_{e'}$. These two times are uncertain, and they are related to sequential repairs and the uncertain time periods of repairs over the time horizon.
- Implicit functions between semi-Markov states and optimization variables: The optimization objective (4) is a function of the semi-Markov states S_e and $S_{e'}$. However, the operational constraints are the functions with regard to optimization variables, e.g., line states and physical variables of power flow. The semi-Markov states cannot be directly mapped to the optimization variables.

The above three challenges make the proposed semi-Markov decision process-based optimization difficult to be solved. In consideration of the recursive characteristics with uncertainties caused by the semi-Markov decision process-based optimization, Bellman's principle of optimality [39] indicates that the future cost can be estimated by a replacement value. Therefore, the Q-learning approach is used to address the first and the second challenges. In addition, a model reformulation approach is employed to address the third challenge.

B. Q-learning Approach

One critical point of the proposed model is to make decision based on the event-driven semi-Markov state. When having the semi-Markov state S_e, the value of the subsequent semi-Markov state $S_{e'}$ and the uncertain repair period $\tau_2 - \tau_1$ make the model difficult to be solved. Q-learning is an approach that uses an optimal estimation $Q^*(S_{e'})$ to replace the uncertain and sequential terms, e.g., we have the following equation.

$$Q^*(S_{e'}) = \int_{\tau_1}^{\tau_2} (\eta \cdot \Delta L_e)dt + v_{e'}(S_{e'}) \tag{13}$$

Based on (13), the optimization objective (4) can be rewritten as follows.

$$v_e(S_e) = \underset{a_e \in A}{\arg\min} \{C_e(S_e, a_e) + Q^*(S_{e'})\} \quad \forall e \tag{14}$$

where $S_{e'}$ is the resulting semi-Markov state from the semi-Markov state S_e under the action a_e. If we know the value of $Q^*(S_{e'})$ for different actions a_e, the optimal strategy for the semi-Markov state S_e can be easily solved subject to with the constraints (6)-(12).

The value of $Q^*(S_{e'})$ cannot be obtained directly due to the uncertainty caused by τ_1 and τ_2. The critical point for the Q-learning approach is to estimate the value of $Q^*(S_{s'})$ via off-line value iteration with stochastic paths by means of monte carlo sampling. To represent the iteration, the superscripts n and $n+1$ are added in the model.

$$v_e^{n+1}(S_e) = \underset{a_e \in A}{\arg\min} \{C_e(S_e, a_e) + Q^{*n}(S_{e'})\} \tag{15}$$

$$Q^{*(n+1)}(S_{e'}) = exp(-\lambda) \cdot \underbrace{Q^{*n}(S_{e'})}_{I} + (1 - exp(-\lambda)) \cdot \underbrace{\left[\int_{\tau_1^{n+1}}^{\tau_2^{n+1}} (\eta \cdot \Delta L_e^{n+1})dt + v_{e''}^n(S_{e''})\right]}_{II} \tag{16}$$

where the optimization objective (15) constrained by (6)-(12) with the known value of $Q^{*n}(S_{e'})$ at the $(n{+}1)^{th}$ iteration. Then, the value of $Q^{*(n+1)}(S_{e'})$ at the $(n{+}1)^{th}$ iteration is updated by means of (16). There are two updated dimensions in (16). The first dimension is the estimation of the corresponding semi-Markov state at the $(n)^{th}$ iteration, i.e., the term I on the right side of (16). The second dimension is the value of resulting semi-Markov state under the action at the $(n{+}1)^{th}$ iteration, i.e., the term II on the right side of (16). The value of $\tau_2^{n+1} - \tau_1^{n+1}$ is the uncertain repair time, which is generated by monte carlo sampling.

The parameter λ is interpreted as the dynamic learning rate, and its value is the mean value of slopes of Q value curves at each iteration. When the slope of the Q value curve tends to be 0, i.e., the value of $exp(-\lambda)$ tends to be 1, the left side of (16) tends to be the term I. This indicates that the Q value is converged. Otherwise, the slope of the Q value curve is not 0. In this case, the estimated values at the previous iteration, i.e., the term I in (16), and the resulting values under uncertainty at the current iteration, i.e., the term II in (16), are employed to updated the estimated values until Q values are converged. With enough off-line iterations, the converged value of $Q^*(S_{e'})$ can be obtained, and online optimization can be implemented based on event-driven semi-Markov states by utilizing (14).

C. Model Reformulation

The Q-learning approach provides a good solving process for the proposed semi-Markov decision process-based event-driven optimization model. However, there is still one challenge in the solving process, i.e., the semi-Markov states are not explicit functions with regard to optimization variables such as line states. The non-explicit functions cause the difficulty in solving the optimization objective (15) constrained by (6)-(12). For a semi-Markov state, repair states and system topology states are related to actions, and the system loading factor is independent on actions. Therefore, additional variables, mapping repair states and system topology states, need to be included in the model.

1) Semi-Markov State Mapping Constraints: For (15) at the $(n{+}1)^{th}$ iteration, $C_e(S_e, a_e)$ can be represented as an explicit function with regard to optimization variables but $Q^{*n}(S_{e'})$ cannot be represented as an explicit function with regard to optimization variables. The semi-Markov state $S_{e'}$ is related to line states and repair states, and in consequence (15) can be expressed as follows.

$$\min \left\{ \begin{array}{l} \sum_l (c_l \cdot x_{le'} + c'_l \cdot x'_{le'} + c''_l \cdot x''_{le'}) + \\ \sum_{e'} \left\{ \prod_g \left((2b_{ge'} - 1)(\beta_{ge'} + b_{ge'} - 1) \right) Q_{e'}^{*n} \right\} \end{array} \right\} \tag{17}$$

where $Q_{e'}^{*n}$ is the estimated Q value of the semi-Markov state $S_{e'}$ at the n^{th} iteration and its value is given. $b_{ge'}$ is the $(g,e')^{th}$ element in the matrix $\mathbf{b}$ that is binary-coded to represent the relations between semi-Markov states and line/repair states, and $b_{ge'}$ is a given value. The term $\prod_g \left((2b_{ge'} - 1)(\beta_{ge'} + b_{ge'} - 1) \right)$ has the multilinear function, and a recursive McCormick envelope is used to transform the multilinear function into a group of linear inequality constraints. For the multilinear function $\beta_{1e'} \cdot \beta_{2e'} \cdots \beta_{ge'}$, we introduce new additional variables listed in (18).

$$\begin{array}{l} \varsigma_{2e'} = \beta_{1e'} \cdot \beta_{2e'} \\ \varsigma_{3e'} = \varsigma_{2e'} \cdot \beta_{3e'} \\ \cdots \\ \varsigma_{ge'} = \varsigma_{(g-1)e'} \cdot \beta_{ge'} \end{array} \tag{18}$$

Based on the new variables $\varsigma_{2e'}$, $\varsigma_{3e'}$, ..., $\varsigma_{ge'}$, the multilinear function $\beta_{1e'} \cdot \beta_{2e'} \cdots \beta_{ge'}$ can be transformed into a group of linear inequality constraints as follows.

$$\begin{array}{l} \varsigma_{2e'} \leq \beta_{1e'} \\ \varsigma_{2e'} \geq \beta_{2e'} + \beta_{1e'} - 1 \\ \varsigma_{ge'} \leq \beta_{1e'} \ (g = 2, \cdots, m) \\ \varsigma_{ge'} \geq 0 \ (g = 2, \cdots, m) \\ \varsigma_{ge'} \leq \varsigma_{(g-1)e'} \ (g = 3, \cdots, m) \\ \varsigma_{ge'} \geq \beta_{ge'} + \varsigma_{(g-1)e'} - 1 \ (g = 3, \cdots, m) \end{array} \tag{19}$$

where the linear inequality constraints (19) based on the recursive McCormick envelope are equivalent to the original constraints (18) because $\varsigma_{2e'}$, $\varsigma_{3e'}$, ..., $\varsigma_{ge'}$, $\beta_{1e'}$, $\beta_{2e'}$, ..., $\beta_{ge'}$ are binary variables.

2) Repair Mapping Constraints: Before each event-driven repair, the repaired components that are not in the repair waiting listing will not be notified to the repair crews, and the repairs in the semi-Markov state $S_{e'}$ have the following constraints.

$$\kappa_j + \sum_i y_{ije'} \leq 1 \quad \forall j \tag{20}$$

$$\sum_j y_{ije'} \leq 1 \quad \forall i \tag{21}$$

where κ_j is a given parameter representing whether the component j is repaired. If the component j is repaired, $\kappa_j = 1$; otherwise, $\kappa_j = 0$. (20) shows that unrepaired components can be scheduled to one repair crew. (21) indicates that one repair crew can only select one unrepaired component at the same time.

D. Q-learning Iteration

Based on (15), (16) and given Q values for semi-Markov states, the optimization model can be solved easily. Q learning utilizes a large amount of off-line iterative optimization and learning to achieve the convergent estimated Q values for semi-Markov states. The learning for Q-learning iteration is listed in **Algorithm 1**.

Algorithm 1 Q-learning Iteration

1: Step 1: Set the iteration counter $n = 1$ and the maximum number of iterations N, respectively.
2: Step 2: Set the initial value of $Q_{e'}^{*n}$ for each semi-Markov state $S_{e'}$ as 0.
3: Step 3: Do until all damaged components are repaired.
4: Step 3.1: Optimize the following model.
5: $Obj.$ (17)
6: $s.t.$ $(6)-(12), (18)-(21)$
7: Step 3.2: With the optimization in Step 3.1, the optimal event-driven strategy (repair and reconfiguration) for the semi-Markov state can be determined. Generate the stochastic repair period $\tau_2 - \tau_1$ on the stochastic paths.
8: Step 3.3: Use (16) to update the Q value of the corresponding semi-Markov state at the $(n+1)^{th}$ iteration.
9: Step 3.4: With the repair periods for all crews repairing components, determine the next event-driven state and repeat Step 3.1 until all damaged components are repaired.
10: Step 4: Set n=n+1. If $n \leq N$ go to Step 3.
11: Step 5: If $n = N$, return the estimated Q values for semi-Markov states.

V. Case Studies

In this section, a revised IEEE 123-bus system is used to verify the proposed model and the algorithm.

A. Data description

Fig. 5 shows the revised IEEE 123-bus system with eight damaged lines. There are thirteen automatic line switches, which can be used for system reconfiguration. There are two groups of repair crews for independent repairs. Each group of repair crew can be arranged to repair damaged lines via different paths. The repair time for one damaged line and the transportation time from the previous location are different. In this study, the transportation time is included in the repair time, and it is assumed that the repair time is satisfied as normal distributions. T in (22) shows the mean values (minutes) of the normal distributions. The element $T_{ij}, i \neq j, i \in \{1, \cdots, 8\}, j \in \{1, \cdots, 8\}$ denotes the total period from the damaged line i to the damaged line j with the addition of the repair time for the damaged line j. The element $T_{ij}, i \in \{9\}, j \in \{1, \cdots, 8\}$ denotes the total period from the depot location to the damaged line j with the addition of the repair time for the damaged line j. The element $T_{ij}, i \in \{1, \cdots, 8\}, j \in \{9\}$ denotes the total period from the damaged line i to the depot location. The variance for each period is 6 minutes. Table I lists the damaged lines, and Table II lists the dispatchable lines.

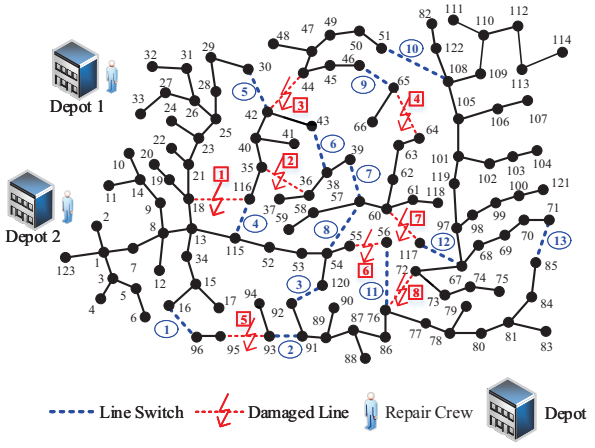


Fig. 5. Test system with damaged lines.

$$T = \begin{bmatrix} - & 50 & 80 & 90 & 100 & 110 & 70 & 80 & 40 \\ 55 & - & 40 & 60 & 80 & 120 & 100 & 90 & 35 \\ 75 & 50 & - & 50 & 70 & 100 & 90 & 80 & 30 \\ 85 & 65 & 55 & - & 80 & 90 & 100 & 100 & 40 \\ 90 & 90 & 75 & 75 & - & 50 & 70 & 70 & 25 \\ 105 & 115 & 105 & 100 & 55 & - & 80 & 90 & 40 \\ 80 & 110 & 85 & 95 & 75 & 85 & - & 100 & 40 \\ 85 & 85 & 90 & 110 & 75 & 95 & 105 & - & 25 \\ 90 & 95 & 55 & 105 & 50 & 100 & 105 & 60 & - \end{bmatrix} \tag{22}$$

TABLE I
Damaged Lines

Label	Lines	Label	Lines
F_1	18-116	F_5	93-95
F_2	42-44	F_6	55-56
F_3	35-36	F_7	60-117
F_4	64-65	F_8	72-76

TABLE II
Dispatchable Lines

No.	Lines	No.	Lines
1	16-96	8	51-108
2	92-120	9	115-116
3	56-76	10	71-85
4	39-57	11	54-57
5	38-43	12	91-93
6	30-42	13	67-117
7	46-65		

B. Estimated values of Q values

For the proposed model of uncertainty-inflicted event-driven repairs and reconfiguration, the Q values for semi-Markov states are important to transform the original recursive model with uncertainty to an event-driven deterministic optimization

model. The Q values for semi-Markov states are estimated by means of off-line learning and optimization. Fig. 6(a), (b), (c), (d), (e), and (f) show Q values of different semi-Markov states with different load levels over the time horizon. It is observed that the Q value for each semi-Markov state is converged to a constant after 15000 iterations. With the converged Q values, i.e., the values of $Q^*(S_{e'})$, the optimization model (12) is transformed into an event-driven deterministic optimization model, which can be optimized easily. In each time period, the load levels are different, and in consequence the converged Q values for semi-Markov states in different load levels are also varying. Fig. 7 shows the converged Q values of the semi-Markov states in different time period. Based on these converged Q values, the system operators can make sequential decisions according to event-driven repairs and reconfiguration. The following section will introduce how to make event-driven decisions in consideration of uncertain events over time horizon.

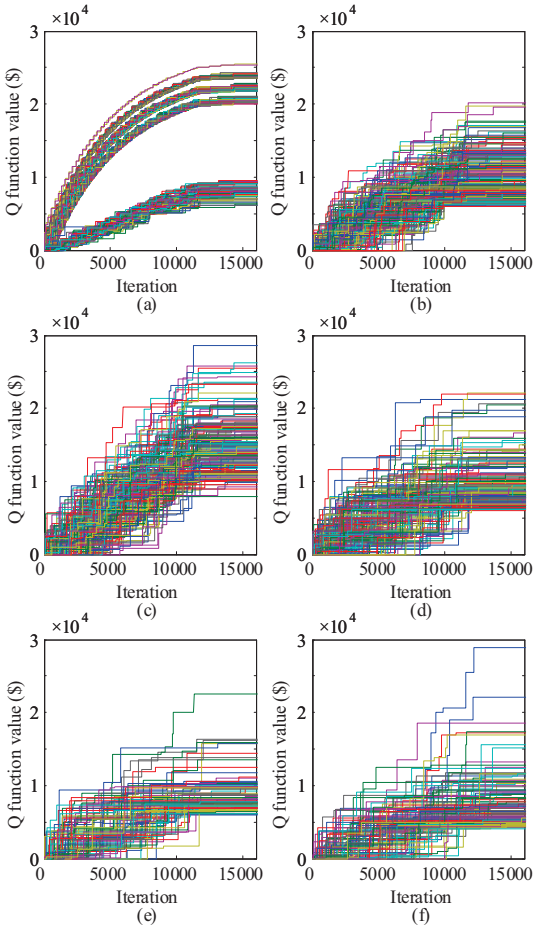


Fig. 6. Iterations for semi-Markov states at different load levels, and (a)-(f) correspond to load level 1 to load level 6.

C. Uncertainty-inflicted event-driven repairs and reconfiguration

Termination of one repair is considered as the start of another event-driven repair/reconfiguration. Because the period of repair is uncertain, the decision point is uncertain and is dependent on the termination of previous repairs. It is necessary to translate the findings of this study into practical

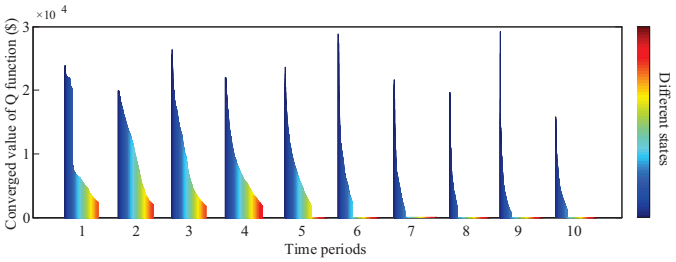


Fig. 7. Convergent estimated Q values for semi-Markov states at different load levels.

applications. Fig. 8 shows three scenarios to illustrate the uncertainty-inflicted event-driven decision making process.

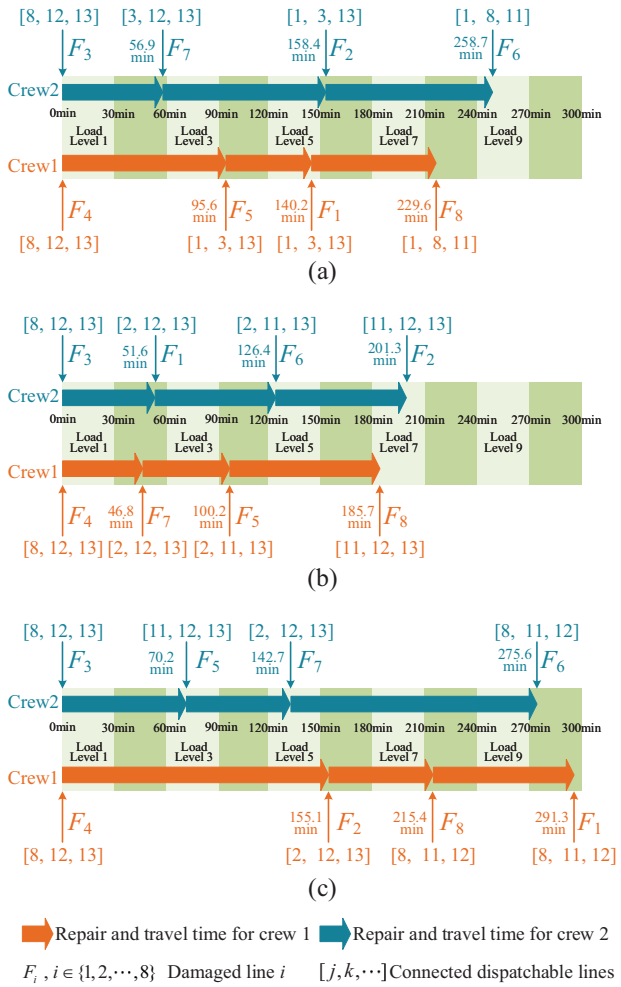


Fig. 8. Uncertainty-inflicted event-driven repairs and reconfiguration for case I (a), case II (b), and case III (c).

At the beginning, crew 1 and crew 2 are located at depots, and this can be considered as the same event when making decision from the depots. In consequence, the repair strategy (i.e., crew 1 needs to repair F_4 and crew 2 needs to repair F_3) and the reconfiguration strategy (i.e., the 8th, 12th, and 13th dispatchable lines are connected) are the same for the three cases. After making decision at the initial time, uncertain

This article has been accepted for publication in IEEE Transactions on Power Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TPWRS.2024.3386851

periods of repairs will lead to different 'events'. Different events cause different decision-making actions.

For the case I in Fig. 8 (a), the repair on F_3 implemented by crew 2 is finished at 56.9 mins, but the repair on F_4 implemented by crew 1 is not finished at this time. At this time point, the event-driven strategy is that the crew 2 starts to repair F_7 and the 3rd, 12th, and 13th dispatchable lines need to be connected. Then, the repair on F_4 implemented by crew 1 is finished at 95.6 mins and the repair on F_7 implemented by crew 2 is not finished at this time, the event-driven strategy is that the crew 1 starts to repair F_5 and the 1st, 3rd, and 13th dispatchable lines need to be connected. All damaged lines are repaired by means of similar event-driven strategies dependent on uncertain time periods of repairs implemented by the crews in the real world.

For the case II in Fig. 8 (b), the repair on F_4 implemented by crew 1 is finished at 46.8 mins, but the repair on F_3 implemented by crew 2 is not finished at this time. At this time point, the event-driven strategy is that the crew 1 starts to repair F_7 and the 2rd, 12th, and 13th dispatchable lines need to be connected. Then, the repair on F_3 implemented by crew 2 is finished at 51.6 mins and the repair on F_7 implemented by crew 1 is not finished at this time, the event-driven strategy is that the crew 2 starts to repair F_1 and the 2st, 12rd, and 13th dispatchable lines need to be connected. All damaged lines are repaired by means of similar event-driven strategies dependent on uncertain time periods of repairs.

For the case III in Fig. 8 (c), the repair on F_3 implemented by crew 2 is finished at 70.2 mins, but the repair on F_4 implemented by crew 1 is not finished at this time. At this time point, the event-driven strategy is that the crew 2 starts to repair F_5 and the 11rd, 12th, and 13th dispatchable lines need to be connected. Then, the repair on F_5 implemented by crew 2 is finished at 142.7 mins and the repair on F_4 implemented by crew 1 is still not finished at this time, the event-driven strategy is that the crew 2 starts to repair F_7 and the 2st, 12rd, and 13th dispatchable lines need to be connected. All damaged lines are repaired by means of similar event-driven strategies dependent on uncertain time periods of repairs.

Based on the above three cases, the repair decisions are made based on the 'actual event'. In the real world, 'actual event' depends on the uncertain periods of the repairs. Take the case II as an example. The 'actual event' can be interpreted as that the repair on F_4 implemented by crew 1 is finished at 46.8 mins and the repair on F_3 implemented by crew 2 is not finished at 46.8 mins. For this uncertainty-inflicted 'actual event', the event-driven strategy is that the crew 1 starts to repair F_7 and the 2rd, 12th, and 13th dispatchable lines are connected. Therefore, the strategies in practice are event-driven with full consideration of the uncertain decision points due to uncertain events.

D. Impacts of Q-learning parameter λ on converged Q Values

When updating the estimated Q values, the parameter λ in the (16) has an impact on the convergence. Fig. 9 shows the curve of the Q values of four semi-Markov states with different λ. When fixing the value of $exp(-\lambda)$ at 0.95 and 0.85, the converged Q values approximately equal. With a smaller fixed value of $exp(-\lambda)$, e.g., 0.75, 0.65, and 0.55, the converged Q values have larger deviations but have faster convergence rates. The red lines represent the proposed dynamic λ for the four semi-Markov states. It is observed that the proposed approach

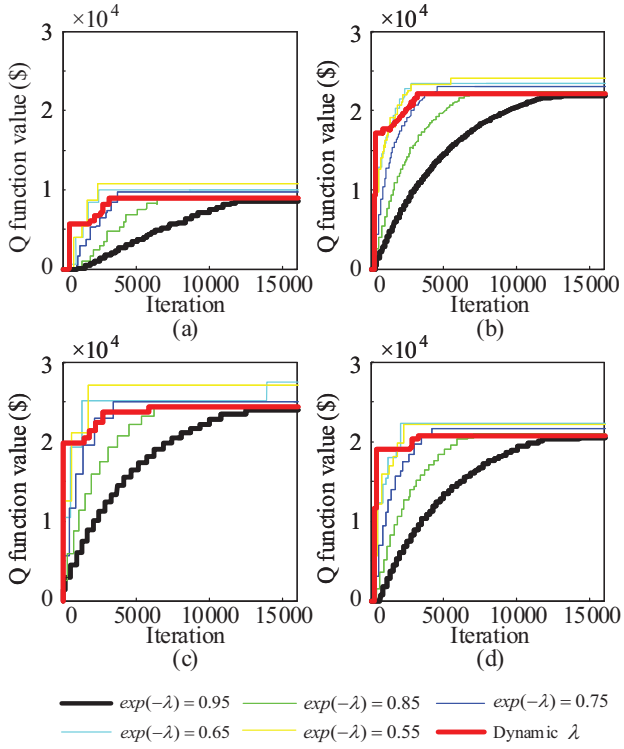


Fig. 9. Iterations for semi-Markov states with different fixed learning rates and dynamic learning rates.

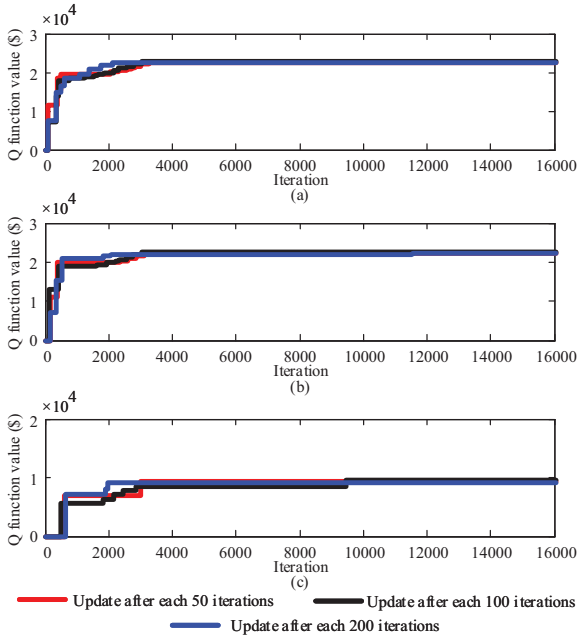


Fig. 10. Iterations for semi-Markov states with dynamic learning rates in consideration of different updated data.

with the dynamic λ has the fast convergence rate and smaller deviations.

In the simulation, the dynamic λ is defined as the mean value of slopes of Q value curves. To simplify the calculation, it is assumed that the dynamic λ is updated after a given iteration and the updated value is the mean value of slopes of

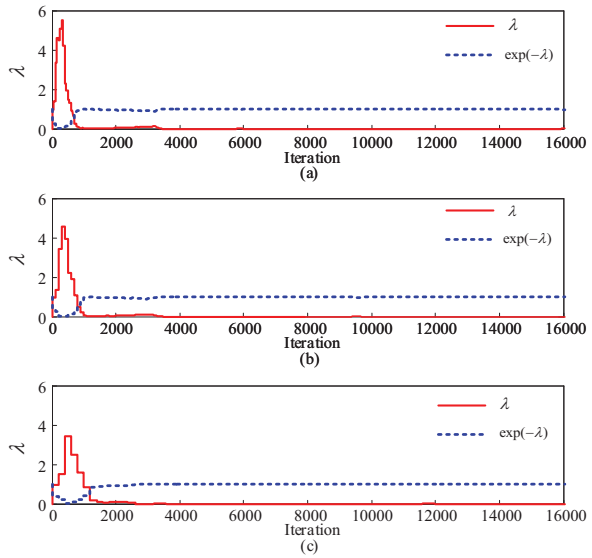


Fig. 11. Dynamic learning rates in consideration of different updated data.

Q value curves at the current iteration. Fig. 10 shows different scenarios for the dynamic λ. The red line, the black line and the blue line represent that the dynamic λ is updated after each 50, 100, and 200 iterations, respectively. The results show that three scenarios for the dynamic λ have similar convergent characteristics. Fig. 11 shows the values of the dynamic λ and $exp(-\lambda)$. It is observed that the curve profiles for three scenarios are similar, and this indicates the proposed dynamic update for the parameter λ can achieve better convergence with acceptable iterations.

E. Discussion

1) Model accuracy: When modeling the operational constraints, the constraints (11a)-(11d) are used to approximate the constraints (10). The key idea of the constraints (11a)-(11d) is the circular constraint linearization method, i.e., using the square constraints to approximate the circular constraint [40]. In the current study, two square constraints are used to approximate the circular constraint. In practice, more square constraints can be used to improve the accuracy of the model but also increase the solution time. Because this study focuses on the uncertainty-inflicted event-driven process, more accurate approximations of the constraints (10) are not investigated. In addition, the current approximation for the constraints (10) is used in some existing studies, also indicating the acceptable accuracy.

2) Computation time: The whole process includes offline Q learning and online decision-making. In our case study, the average calculation time of offline Q learning takes 110 hours. Although Q learning is a time-consuming task, it is acceptable due to offline calculation. In addition, parallel calculation can be used to accelerate the computational speed. Based on the estimated values from the offline Q learning, the online decision-making problem can be transformed into an event-driven deterministic optimization model, i.e., a one-step deterministic optimization model listed in (14). The calculation time for this one-step deterministic optimization model in the case study is 5 seconds on average, and this is acceptable for uncertainty-inflicted event-driven repair/reconfiguration action.

3) Application Scope: As described in the paper, the proposed model can be directly used for radial distribution systems. We would like to clarify that the concept and the framework of the uncertainty-inflicted event-driven process can be used for different power systems, e.g., non-radial distribution systems and transmission systems. The difference between different systems on the detailed model is the operational constraints, i.e., (6)-(12). In addition, more reinforcement learning approaches, e.g., deep deterministic policy gradient (DDPG), advantage actor critic (A2C), asynchronous advantage actor-critic (A3C), and soft actor-critic (SAC), can be used to solve the proposed model. With deep reinforcement learning approaches, the action for one semi-Markov state and the corresponding reward after the action can be generated via two deep neural networks. Solving the proposed model with deep reinforcement learning approaches is a promising point in the future.

VI. Conclusion

In consideration of sequential and uncertain decision point caused by uncertain repair time, this paper investigated semi-Markov decision process-based resilient recovery with uncertainty-inflicted event-driven repairs. A uncertainty-inflicted event-driven process is employed to represent the sequential repair/reconfiguration in consideration of uncertain repair time. Sequential repair implemented by different repair crews are modeled as semi-Markov states, with which the whole process is established as a semi-Markov decision process-based optimization model, which is an event-driven recursive model. Q-learning is employed to solve the proposed model, and the dynamic learning rates based on the slopes Q value curves for semi-Markov states are used to improve the convergence.

Based on the simulations, we have the following conclusions. (1) The semi-Markov decision process-based recursive optimization model can well represent the uncertainty-inflicted and sequential event-driven repairs. (2) The Q-learning approach with the dynamic learning rates based on the slopes Q value curves for semi-Markov states has faster convergence. (3) With the convergent estimations for semi-Markov states, the original event-driven recursive model with uncertainty can be transformed into an event-driven deterministic optimization model, i.e., a one-step deterministic optimization model.

References

[1] C. Wang, P. Ju, F. Wu, X. Pan, and Z. Wang, "A systematic review on power system resilience from the perspective of generation, network, and load," *Renewable and Sustainable Energy Reviews*, vol. 167, p. 112567, 2022.

[2] 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

[3] Electric Power Research Institute, "Enhancing distribution resiliency: Opportunities for applying innovative technologies." [Online]. Available: http://www2.epri.com/abstracts/Pages/ProductAbstract.aspx?ProductId=000000000001026889

[4] S. Ma, L. Su, Z. Wang, F. Qiu, and G. Guo, "Resilience enhancement of distribution grids against extreme weather events," *IEEE Trans. Power Syst.*, vol. 33, no. 5, pp. 4842–4853, 2018.

[5] 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, 2017.

[6] P. Zhao, Z. Cao, D. D. Zeng, C. Gu, Z. Wang, Y. Xiang, M. Qadrdan, X. Chen, X. Yan, and S. Li, "Cyber-resilient multi-energy management for complex systems," *IEEE Trans. Ind. Informat.*, vol. 18, no. 3, pp. 2144–2159, 2022.

[7] C. Wang, P. Ju, S. Lei, Z. Wang, F. Wu, and Y. Hou, "Markov decision process-based resilience enhancement for distribution systems: An approximate dynamic programming approach," *IEEE Trans. Smart Grid*, vol. 11, no. 3, pp. 2498–2510, 2020.

[8] Z. Yang, A. Martí, Y. Chen, and J. R. Martí, "Optimal resource allocation to enhance power grid resilience against hurricanes," *IEEE Trans. Power Syst.*, vol. 38, no. 3, pp. 2621–2629, 2023.

[9] X. Yang, Z. Zhou, Y. Zhang, J. Liu, J. Wen, Q. Wu, and S. Cheng, "Resilience-oriented co-deployment of remote- controlled switches and soft open points in distribution networks," *IEEE Trans. Power Syst.*, vol. 38, no. 2, pp. 1350–1365, 2023.

[10] S. Sun, G. Li, C. Chen, Y. Bian, and Z. Bie, "A novel formulation of radiality constraints for resilient reconfiguration of distribution systems," *IEEE Trans. Smart Grid*, vol. 14, no. 2, pp. 1337–1340, 2023.

[11] Y. Huang, G. Li, C. Chen, Y. Bian, T. Qian, and Z. Bie, "Resilient distribution networks by microgrid formation using deep reinforcement learning," *IEEE Trans. Smart Grid*, vol. 13, no. 6, pp. 4918–4930, 2022.

[12] A. Arif, Z. Wang, C. Chen, and B. Chen, "A stochastic multi-commodity logistic model for disaster preparation in distribution systems," *IEEE Trans. Smart Grid*, vol. 11, no. 1, pp. 565–576, 2020.

[13] T. Ding, Z. Wang, W. Jia, B. Chen, C. Chen, and M. Shahidehpour, "Multiperiod distribution system restoration with routing repair crews, mobile electric vehicles, and soft-open-point networked microgrids," *IEEE Trans. Smart Grid*, vol. 11, no. 6, pp. 4795–4808, 2020.

[14] H. Farzin, M. Fotuhi-Firuzabad, and M. Moeini-Aghtaie, "Enhancing power system resilience through hierarchical outage management in multi-microgrids," *IEEE Trans. Smart Grid*, vol. 7, no. 6, pp. 2869–2879, 2016.

[15] K. S. A. Sedzro, X. Shi, A. J. Lamadrid, and L. F. Zuluaga, "A heuristic approach to the post-disturbance and stochastic pre-disturbance microgrid formation problem," *IEEE Trans. Smart Grid*, vol. 10, no. 5, pp. 5574–5586, 2019.

[16] Y. Xu, C.-C. Liu, K. P. Schneider, F. K. Tuffner, and D. T. Ton, "Microgrids for service restoration to critical load in a resilient distribution system," *IEEE Trans. Smart Grid*, vol. 9, no. 1, pp. 426–437, 2018.

[17] Y. Wang, Y. Xu, J. Li, J. He, and X. Wang, "On the radiality constraints for distribution system restoration and reconfiguration problems," *IEEE Trans. Power Syst.*, vol. 35, no. 4, pp. 3294–3296, 2020.

[18] Y. Lin, B. Chen, J. Wang, and Z. Bie, "A combined repair crew dispatch problem for resilient electric and natural gas system considering reconfiguration and dg islanding," *IEEE Trans. Power Syst.*, vol. 34, no. 4, pp. 2755–2767, 2019.

[19] R. Roofegari nejad, W. Sun, and A. Golshani, "Distributed restoration for integrated transmission and distribution systems with ders," *IEEE Trans. Power Syst.*, vol. 34, no. 6, pp. 4964–4973, 2019.

[20] A. Arjomandi-Nezhad, M. Fotuhi-Firuzabad, M. Moeini-Aghtaie, A. Safdarian, P. Dehghanian, and F. Wang, "Modeling and optimizing recovery strategies for power distribution system resilience," *IEEE Systems Journal*, vol. 15, no. 4, pp. 4725–4734, 2021.

[21] H. Liang and Q. Xie, "Resilience-based sequential recovery planning for substations subjected to earthquakes," *IEEE Transactions on Power Delivery*, vol. 38, no. 1, pp. 353–362, 2023.

[22] Z. Yang and J. R. Martí, "Real-time resilience optimization combining an ai agent with online hard optimization," *IEEE Trans. Power Syst.*, vol. 37, no. 1, pp. 508–517, 2022.

[23] P. Zhao, C. Gu, and D. Huo, "Coordinated risk mitigation strategy for integrated energy systems under cyber-attacks," *IEEE Trans. Power Syst.*, vol. 35, no. 5, pp. 4014–4025, 2020.

[24] H. Zhang, C. Chen, S. Lei, and Z. Bie, "Resilient distribution system restoration with communication recovery by drone small cells," *IEEE Trans. Smart Grid*, vol. 14, no. 2, pp. 1325–1328, 2023.

[25] X. Sun, J. Chen, H. Zhao, W. Zhang, and Y. Zhang, "Sequential disaster recovery strategy for resilient distribution network based on cyber–physical collaborative optimization," *IEEE Trans. Smart Grid*, vol. 14, no. 2, pp. 1173–1187, 2023.

[26] Y. Tan, F. Qiu, A. K. Das, D. S. Kirschen, P. Arabshahi, and J. Wang, "Scheduling post-disaster repairs in electricity distribution networks," *IEEE Trans. Power Syst.*, vol. 34, no. 4, pp. 2611–2621, 2019.

[27] G. Zhang, F. Zhang, X. Zhang, K. Meng, and Z. Y. Dong, "Sequential disaster recovery model for distribution systems with co-optimization of maintenance and restoration crew dispatch," *IEEE Trans. Smart Grid*, vol. 11, no. 6, pp. 4700–4713, 2020.

[28] S. Lei, C. Chen, Y. Li, and Y. Hou, "Resilient disaster recovery logistics of distribution systems: Co-optimize service restoration with repair crew and mobile power source dispatch," *IEEE Trans. Smart Grid*, vol. 10, no. 6, pp. 6187–6202, 2019.

[29] A. Arif, Z. Wang, C. Chen, and J. Wang, "Repair and resource scheduling in unbalanced distribution systems using neighborhood search," *IEEE Trans. Smart Grid*, vol. 11, no. 1, pp. 673–685, 2020.

[30] T. Van Acker and D. Van Hertem, "Stochastic process for the availability assessment of single-feeder industrial energy system sections," *IEEE Trans. Rel.*, vol. 67, no. 4, pp. 1459–1467, 2018.

[31] H. Dui, S. Si, M. J. Zuo, and S. Sun, "Semi-markov process-based integrated importance measure for multi-state systems," *IEEE Trans. Rel.*, vol. 64, no. 2, pp. 754–765, 2015.

[32] T. Uemura, T. Dohi, and N. Kaio, "Availability analysis of an intrusion tolerant distributed server system with preventive maintenance," *IEEE Trans. Rel.*, vol. 59, no. 1, pp. 18–29, 2010.

[33] C. Duan and P. Chen, "Adaptive maintenance scheme for degrading devices with dynamic conditions and random failures," *IEEE Trans. Ind. Informat.*, vol. 19, no. 3, pp. 2508–2519, 2023.

[34] M. Perman, A. Senegacnik, and M. Tuma, "Semi-markov models with an application to power-plant reliability analysis," *IEEE Trans. Rel.*, vol. 46, no. 4, pp. 526–532, 1997.

[35] Y. Xiang, Z. Ding, Y. Zhang, and L. Wang, "Power system reliability evaluation considering load redistribution attacks," *IEEE Trans. Smart Grid*, vol. 8, no. 2, pp. 889–901, 2017.

[36] Y. Zhang, L. Wang, and Y. Xiang, "Power system reliability analysis with intrusion tolerance in scada systems," *IEEE Trans. Smart Grid*, vol. 7, no. 2, pp. 669–683, 2016.

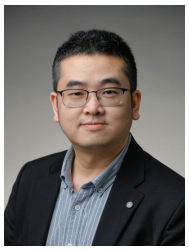
[37] S. Lei, C. Chen, Y. Song, and Y. Hou, "Radiality constraints for resilient reconfiguration of distribution systems: Formulation and application to microgrid formation," *IEEE Trans. Smart Grid*, vol. 11, no. 5, pp. 3944–3956, 2020.

[38] 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.

[39] R. Bellman, *Dynamic Programming*, Princeton, NJ, USA: Princeton University Press; 1957.

[40] 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, 2016.

Chong Wang received the B.E and M.S degrees in electrical engineering from Hohai University, Nanjing, China, in 2009 and 2012, and the Ph.D. degree in electrical engineering from The University of Hong Kong, Hong Kong, in 2016. He was a postdoctoral researcher at The University of Hong Kong in 2016, and was a postdoctoral researcher at Iowa State University from 2017 to 2018. He is currently a Professor with the College of Energy and Electrical Engineering, Hohai University, Nanjing, China. His research interests include resilient power system, power system uncertainty analysis and control.

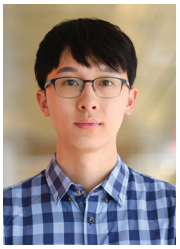


Gengfeng Li received the B.S. and Ph.D. degrees in electrical engineering from Xi'an Jiaotong University (XJTU), Xi'an, China, in 2008 and 2014, respectively. He is currently a Professor with the School of Electrical Engineering, Xi'an Jiaotong University. His research interests include resilient power systems, reliability evaluation and enhancement of distribution systems, demand side management, and artificial intelligence applications in power systems.

Can Wan received the B.Eng. degree from Zhejiang University, Hangzhou, China, in 2008 and the Ph.D. degree from The Hong Kong Polytechnic University in 2015. He is currently a Professor with the College of Electrical Engineering, Zhejiang University. He was a Postdoctoral Fellow with the Department of Electrical Engineering, Tsinghua University, Beijing, China, and held a Research positions with the Technical University of Denmark, Kongens Lyngby, Denmark, The Hong Kong Polytechnic University, and City University of Hong Kong, Hong Kong. His research interests include forecasting, power system uncertainty analysis and control, renewable energy integration, and machine learning.

Zhaoyu Wang is the Harpole-Pentair Assistant Professor with Iowa State University. He received the B.S. and M.S. degrees in electrical engineering from Shanghai Jiaotong University in 2009 and 2012, respectively, and the M.S. and Ph.D. degrees in electrical and computer engineering from Georgia Institute of Technology in 2012 and 2015, respectively. He was a Research Aid at Argonne National Laboratory in 2013 and an Electrical Engineer Intern at Corning Inc. in 2014. His research interests include power distribution systems, microgrids, renewable integration, power system resilience, and power system modeling.

Ping Ju received the B.S. and M.S. degrees from Southeast University, Nanjing, China, in 1982 and 1985, respectively, and the Ph.D. degree from Zhejiang University, Hangzhou, China, all in electrical engineering. He is currently a Professor of electrical engineering with Hohai University, Nanjing, China and Zhejiang University, Hangzhou, China. From 1994 to 1995, he was an Alexander-von Humboldt Fellow with the University of Dortmund, Germany. His research interests include modeling and control of power system with integration of renewable generation.



Shunbo Lei received the B.E. degree in electrical engineering from Huazhong University of Science and Technology, Wuhan, China, in 2013, and the Ph.D. degree in electrical engineering from The University of Hong Kong, Hong Kong SAR, in 2017. He was a visiting scholar at Argonne National Laboratory, Lemont, IL USA, from 2015 to 2017, a postdoctoral researcher with The University of Hong Kong from 2017 to 2019, and a research fellow with the University of Michigan, Ann Arbor, MI USA, from 2019 to 2021. He is currently an assistant professor with the School of Science and Engineering, The Chinese University of Hong Kong – Shenzhen, China. His research interests include power systems, resilience, grid-interactive efficient buildings, optimization and learning.