Q-Learning-Based Vulnerability Analysis of Smart Grid Against Sequential Topology Attacks

Jun Yan, *Student Member, IEEE*, Haibo He, *Senior Member, IEEE*,
Xiangnan Zhong, and Yufei Tang, *Member, IEEE*

***Abstract*— Recent studies on sequential attack schemes revealed new smart grid vulnerability that can be exploited by attacks on the network topology. Traditional power systems contingency analysis needs to be expanded to handle the complex risk of cyber-physical attacks. To analyze the transmission grid vulnerability under sequential topology attacks, this paper proposes a Q-learning-based approach to identify critical attack sequences with consideration of physical system behaviors. A realistic power flow cascading outage model is used to simulate the system behavior, where attacker can use the Q-learning to improve the damage of sequential topology attack toward system failures with the least attack efforts. Case studies based on three IEEE test systems have demonstrated the learning ability and effectiveness of Q-learning-based vulnerability analysis.**

***Index Terms*— Smart grid security, topology attacks, sequential attacks, Q-learning, line-switching, reinforcement learning, power flow analysis.**

I. Introduction

IN THE last decade, the smart grid has been emerging as the next-generation electrical power infrastructure. It integrates information and communication technologies, intelligent control policies, renewable energies, and complex energy systems. However, the promising economic benefits from the smart grid are accompanied by rising cyber-physical security challenges [1], [2]. On one hand, the conventional physical structure of power grids (as shown in Fig. 1) is lack of critical protection against cyber-penetrations [3], [4]. On the other hand, the cyber-integration exposes unprecedented risk of unauthorized access and attack threats to power grid operation [5]. Understanding the power grid vulnerability beyond normal operational conditions and disturbances has already become a critical task for the smart grid.

Manuscript received March 1, 2016; revised June 24, 2016; accepted August 22, 2016. Date of publication September 8, 2016; date of current version November 3, 2016. This work was supported in part by the National Science Foundation under Grant CNS 1117314 and Grant ECCS 1053717 and in part by the Army Research Office under Grant W911NF-12-1-0378. The associate editor coordinating the review of this manuscript and approving it for publication was Prof. Mauro Conti. *(Corresponding author: Haibo He.)*

J. Yan, H. He, and X. Zhong are with the Department of Electrical, Computer and Biomedical Engineering, The University of Rhode Island, Kingston, RI 02881 USA (e-mail: jyan@ele.uri.edu; he@ele.uri.edu; xzhong@ele.uri.edu).

Y. Tang is with the Department of Computer and Electrical Engineering and Computer Science, Institute for Sensing and Embedded Network Systems Engineering, Florida Atlantic University, Boca Raton, FL 33431 USA (e-mail: tangy@fau.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIFS.2016.2607701

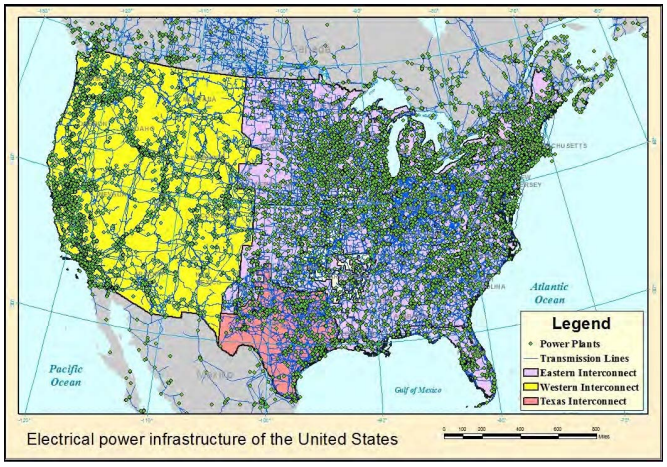


Fig. 1. The electrical power infrastructure in the United States is becoming more interconnected, unpredictable and complicated in the era of smart grid.

Security studies have investigated potential attack schemes and countermeasures. Such attacks exploit different aspects of vulnerability in the smart grid: malicious data attacks [6] inject undetectable false data into the state estimation to mislead grid operations. Line-switching interdiction [7] extends traditional line contingency into interactive attacker-defender scenarios. Cascaded attacks [8] evaluate the risk of cascading outages triggered by attacking a small set of components. Time synchronization attacks [9] target at the critical temporal information of measurement and devices. These attack scheme studies aided traditional power grid contingency and stability analysis by revealing security concerns that are usually outside the scope of normal operating dynamics, random faults, and major disturbances due to extreme natural events.

Smart grid attacks are not limited to a single target when cyber-security are considered: it should be assumed that attackers have the ability to compromise multiple measurements or manipulate control commands from the cyber-space [10]. In addition, despite the standardized $N-1$ security and the forthcoming $N-1-1$ and $N-2$ security standards, it is important for both forensic and security studies to understand the $N-k$ vulnerability of smart grids under potential smart and informed attacks. It remains to date a complex task to thoroughly examine the grid vulnerability under multiple coordinated attacks.

Multi-contingency screenings in power systems have mostly focused on concurrence [11]–[13] instead of sequence.

However, recent studies on the sequential attack have revealed another vulnerability in the smart grid [14]–[16]. Instead of a straightforward concurrent attack scheme, sequential attacks can be launched on critical components consecutively. It is similar to the extension of $N-1-1$ contingency, but the number, target, and timing of attacks could be determined by the attackers to could lead to a maximized damage. A preliminary study [14] has shown that sequential attacks with the same strength can cause comparable damages as the concurrent attacks; nonetheless, sequential attacks require less concurrent resources to coordinate, and the vulnerability of the same targets in such attacks can differ significantly: some line outages that occur back-to-back can lead to much severe system blackouts than when they occur at the same time.

To identify most critical sequences that can lead to large-scale system failures, existing sequential attack and contingency studies rely either on heuristic graph methods [15], exhaustive search [16], or engineering expertise [17]. A more systematic and effective method can be helpful when bulk power systems are considered. False data injection attacks [18] could also lead to topology attacks [19], though the current studies still focus on concurrent schemes and are mostly detection-oriented [20], i.e., to probe concurrent attack schemes that are undetectable by current mechanisms. Instead, this study focuses on impacts of sequential topology attacks with consideration of physical system behaviors when attack has bypassed the detection, as it is equally important to investigate the attack schemes based on its impacts on the physical system to fully understand its threat.

The development of machine learning algorithms provides promising tools to handle complicated security problems with a large, stochastic search space. The patterns underlying in the system dynamics and cascading failures can be adaptively revealed by computational intelligence algorithms. The adaptability, i.e, the ability to self-tune based on previous experience can also aid the vulnerability analysis of a complex system. This paper introduces a novel Q-learning based approach to adaptively identify the more vulnerable attack sequence that can cause critical system failure from sequential topology attacks. In what follows, the term "sequential attack" (SA) exclusively refers to the sequential line-switching interdiction on the power transmission grids.

The major contributions of this paper are as follows:

1) The paper proposed a reinforcement learning based approach for vulnerability analysis of sequential attacks in power transmission grids. The approach evaluates the blackout damage resulting from line-switching interdiction with consideration of overloading-related cascading outages and hidden line failures. It formulates the problem under the reinforcement learning framework and identifies critical sequences in sequential attacks with the Q-learning algorithm;
2) The proposed method, utilizing the Q-learning algorithm and Monte Carlo simulation, effectively identified grid vulnerability under sequential attacks causing complex system outages. Simulation-based case studies showed that critical attack sequences that lead to large blackouts have been identified. Results with different systems and loading levels have shown the effectiveness of the proposed method as it discovers more vulnerable target sequence in sequential attacks;
3) Only topological information has been used to identify the critical attack sequences with the Q-learning approach; this echos the vulnerability observed in [21] as complete information of system dynamics is not required to identify critical sequences in the power grid.

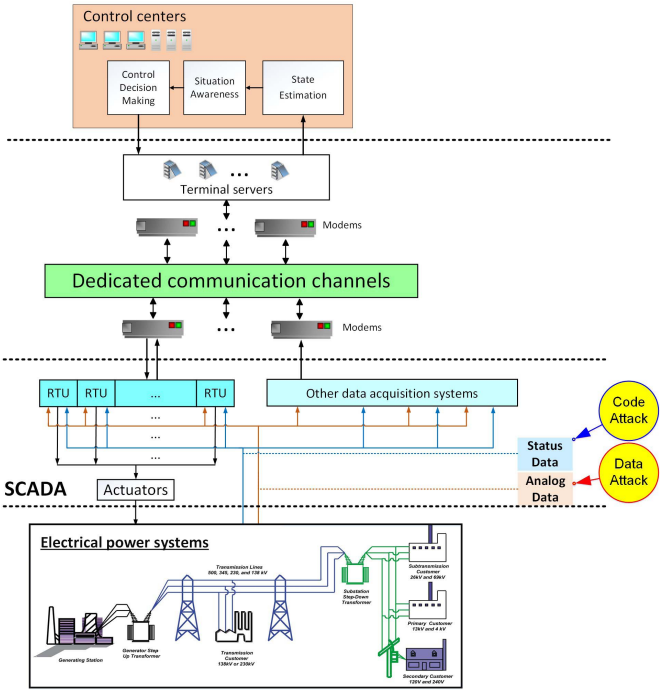


Fig. 2. The smart grid infrastructure and potential attacks on the status/analog data [41].

The rest of the paper is organized as follows: Section II introduces the topology attacks on smart grid, the sequential attacks, the cascading outage vulnerability, and the platform to simulate system responses and outages after line-switching attacks. Section III describes the proposed Q-learning approach to identify critical attack sequences. Section IV demonstrates the simulation results on three benchmark systems of different scales and under different loading. Finally, Section V provides the conclusions and future works.

II. Sequential Topology Attacks in Smart Grids

A. Topology Attacks on Smart Grids

The smart grid integrates two-way communication into system planning and operations [22], where the control centers rely on the supervisory control and data acquisition (SCADA) systems to monitor and operate the grid. However, the entire infrastructure, as illustrated in Fig. 2, has been shown vulnerable to both cybernetic and physical penetrations [4], [23]. Among different threat models, the topology attacks can create large blackouts by manipulating the grid connectivity and creating large disturbances in interconnected power systems [19].

The attack surface consists of both commands and measurements in the smart grids. First, code attacks can manipulate

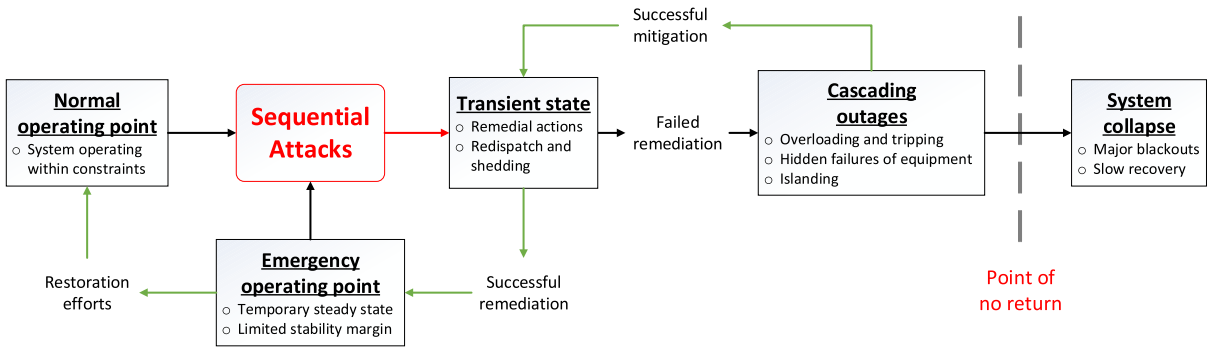


Fig. 3. The intermediate states in cascading blackouts of electrical power grid [29].

or forge control commands [24], [25] to maliciously close a target line from distance [26], [27]. Advanced schemes using tools like the Petri-net [28] can also launch coordinated cyber-physical attacks on transmission components such as transformers and circuit breakers, and maximize their damage when sufficient resources and/or complete information of power systems are available.

Alternatively, the measurements, consisting of status and analog data collected from remote terminal unites (RTUs), are also targets of interest in cyber-attacks. Status data are topological measurements describing the connectivity of the grid, a vital piece of information to determine transmission capability and system reliability. Analog data consist of measurements of system states, including voltage, frequency, active and reactive load, among others, which are vital details of system dynamics. Both data are measured and transmitted to the control room for operational decision making. In this process, topology data attacks can be launched directly by manipulating the status data [30] or indirectly by misleading system operation with false analog data [18], [31]. For either type of attacks, there may also be unforeseen, zero-day type of vulnerabilities to be exploited by smart grid attackers [4].

B. Sequential Topology Attacks

Aforementioned manipulation of control commands or interdiction of controlled electronic devices can pose severe threats to electrical power transmission [32]–[34] as the budgets against interdiction are often limited [25], [35].

In this paper, a sequential attack is defined as a series of coordinated interdiction with malicious topological attacks in a power transmission grid. Specifically, this paper refines an interdiction to the line-switching attack, i.e., the malicious operation that turns an in-service line to out-of-service. Each line-switching directly changes the topology of a transmission grid [36], [37]; in practice, these attacks can be launched from manipulated control commands, false line status data, or physical sabotages. They can also be disguised as irrelevant disturbances or contingencies that are harder to defend [17]. Fig. 3 illustrates the state transitions of power systems in blackouts and the phase where a sequential attack kicks in to trigger a cascading outage, with reference blackout stages described in [29].

Mathematically, a sequential attack scheme can be formulated as a sequence S of ordered and timed 2-tuples [15]:

$$S = \{(a_1, t_1), (a_2, t_2), \ldots, (a_k, t_k)\}, \quad k \leq N \tag{1}$$

where (a_i, t_i) describes the i-th line-switching attack launched on target line a_i at time t_i; k is the number of attacks in the sequence and N is the number of active lines in the power grid. By definition, the time-domain sequence $T = \{t_1, t_2, \ldots, t_k\}$ is non-negative and monotonically non-decreasing for any attack sequence $A = \{a_1, a_2, \ldots, a_k\}$ [16].

The inclusion of selection, ordering and timing in (1) poses a complicated task for vulnerability analysis of sequential attack schemes [16]. The target selection needs to be made from a total of $\binom{N}{k}$ combinations; ordering of the combinations expands the problem to a permutative space $\frac{N!}{(N-k)!}$; in addition, the timing introduces a continuous variable in time domain, where there is little research to provide a background truth or a systematic approach to the best knowledge of the authors. Particularly, attacks with arbitrary timing during transient states with fast system dynamics will lead to complicated and nondeterministic system responses. Therefore, this paper focuses on the handling of the first two aspects in the vulnerability analysis under different sequential attack sequence A; the influence of sequential attack timing T is yet to be analyzed in the future work of this paper.

Consequently, three assumptions are made in this paper to further refine the definition of sequential attacks:

Assumption 1: Attackers can access and manipulate the topological information of power systems. The topological information refers to the connectivity of substations and transmission lines, recorded in the status data and changeable by circuit breaker operations or malicious manipulations as shown in Fig. 2.

Assumption 2: Each line-switching attack $a_i \in A$ is launched during a steady-state of the system, which includes both normal and emergent operating points in Fig. 3. This decomposes a sequential topology attack into a series of k consecutive individual line-switching attacks on a power grid, allowing this study to focus on the identification of attack sequence in vulnerability analysis.

Assumption 3: Without loss of generality, the cost to attack any line is considered equal in this paper.

With the assumptions above, the **attack objective** is to identify a minimal attack sequence that causes a critical system failure through cascading outages. The critical system failure occurs when the number of line outages exceeds a critical threshold N_θ that leads to a system collapse and/or major blackouts, shown as the point of no return in Fig. 3.

C. Vulnerability of Cascading Outages in Power Systems

Cascading outages are among the most severe threats to power grid stability [29], [38]. A malicious attack can trigger cascading outages with various mechanisms, including line overloading [39], frequency and voltage instability [40], grid islanding, among others. Although the initial attack can be few compared to the scale of a power grid, cascading outages often lead to severe and costly power blackouts and economic losses [41]. This vulnerability can thus be targeted and exploited by attacks who aim at triggering cascading outages [42], where the targets can be buses (substations) [21], lines [43], or combined [25].

Modern power system security studies have brought forth many cascading outage models [38]. In general, the triggering and propagation of wide-spread line outages can be modeled as a stochastic process considering overloading and hidden failures [44]. After the initial attack, lines suffering severe or prolonged overloading are exposed to more risks of outage. Based on these considerations, a dedicated simulation platform has been developed in this paper. The platform is adopted from a dedicated cascading failure simulator (CFS) based on direct current (DC) power flow model [15], [39]. Under the steady-state assumption, the DC model is a linear approximation [39] that can describe the power system cascading outages within a short time period. The DC-CFS assumes that the bus voltages have a uniform magnitude of 1.0 p.u., the transmission lines are lossless, and the voltage angle is small enough so that the angular difference $\Delta\theta$ satisfies $\sin(\Delta\theta) = \Delta\theta$. For the load profile, to prioritize load demand and minimize load shedding, the re-dispatch policy will rely on generation adjustments and only apply load shedding as the last resort. The load demand between re-dispatches is considered constant according to the steady state assumption. The hidden failure in cascading outages [45], which considers random outages of exposed lines, has also been integrated in the simulator. The entire procedure is briefly described below and the pseudo code is provided in the Appendix. Interested readers may refer to [11], [15], and [39] for further information regarding the model.

1) *Topology update*: The simulator updates the grid topology when a line is attacked or tripped;
2) *Power re-dispatching*: With the topology updated, the generation and load in the grid are re-dispatched. The generators are first ramped up if there is a power deficit or ramped down if there is a power surplus. Then, additional generations are tripped in an ascending order to protect the machines if a power surplus persists after the ramping. Last, if a deficit persists after the ramping and tripping of generation, the load demand will be shed in an ascending order to avoid further instability and collapse of the system;
3) *Power flow update*: After the re-dispatch, the simulator recalculates power flow along each transmission line based on the DC power flow model [46];
4) *Overload monitoring*: The simulator monitors the lines in the grid and marks a line as overloaded if the line flow F_l exceeds its thermal rating limit C_l;
5) *Line outage risk*: The simulator checks the overload extent $F_l - C_l$ and duration τ of overloaded lines and computes the accumulated risk $O(l)$ for line l over time [11]:
$$O(l) = \int_t^{t+\tau} (F_l - C_l) d\tau \tag{2}$$
6) *Overload protection*: An overloaded line is tripped if the accumulated overloading risk $O(l)$ exceeds a pre-given critical threshold O_T.
7) *Hidden failure risk*: Trip exposed line with the hidden failure probability. Considering the degree and duration of persisting line overloading, the hidden failure probability is defined as a function of the overloading risk $O(l)$ and critical overloading threshold O_T for exposed lines. In this paper, the hidden failure probability is defined as:
$$p(l) = \begin{cases} O(l)/O_T, & \text{if } O(l) > 0 \\ 0, & \text{otherwise} \end{cases} \tag{3}$$

The above process is repeated until no line is overloaded and no target remains to be attacked. Note that this procedure is a simulation of power system behaviors resulting from attacks. The attacker will only be able to observe and manipulate the line status as discussed in Sect. II-A. Meanwhile, any line outages simulated by the CFS depend on dynamics of the system that are affected by but unknown to the attacker.

III. Vulnerability Analysis of Sequential Attacks With Q-Learning

A. The Q-Learning Algorithm

Q-learning belongs to a category of semi-supervised learning algorithms [47] known as the reinforcement learning (RL). In general, RL seeks an action sequence that produces the maximal cumulative rewards via a trial-and-error manner. A typical framework of reinforcement learning is shown in Fig. 4 [47]. An *agent* takes a sequence of actions at a series of states before it reaches an ultimate goal. The quality of each action is assessed by an evaluative feedback from an *environment*, known as the "reward". By adaptively adjust its actions, the agent has an ultimate objective to learn an optimal policy from the cumulative rewards to maximize the expected total rewards it will receive from the environment.

In general, the expected total rewards Q is computed by a discounted cumulative function of the reward r_t observed upon the action a_t taken at state s_t:
$$Q = \sum_{t=1}^{n} \gamma^{t-1} r_t(s_t, a_t) \tag{4}$$
where γ is a discounted factor. Setting $\gamma = 1$ weights every immediate reward equally in the sequence of actions.

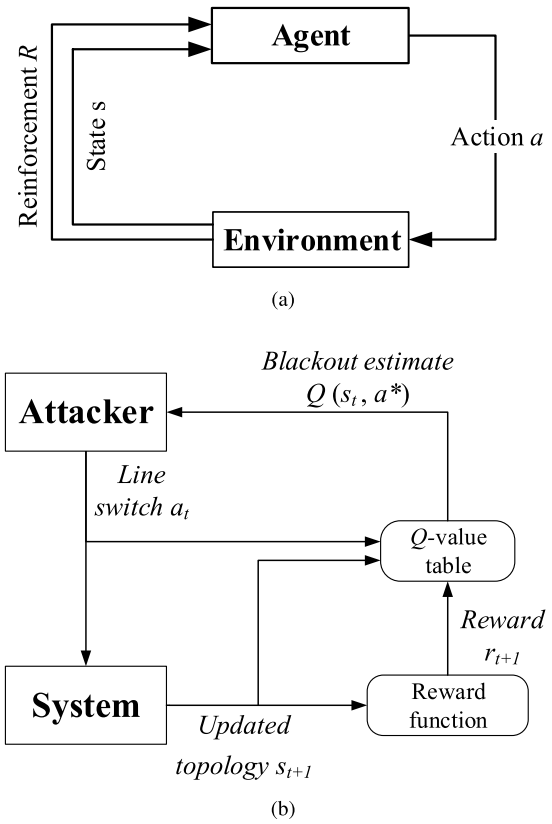


Fig. 4. (a) The general flowchart of typical reinforcement learning process, and (b) the interactions of Q-learning based vulnerability analysis of sequential attacks.

In practice, γ is commonly set slightly smaller than 1.0 to facilitate the convergence of Q value during the learning process [47].

Although the true optimal value of Q^* is usually unknown in practice, it can be approximated by the Q-learning algorithm iteratively. The Q-learning is an off-policy, temporal difference reinforcement learning algorithm that approximates the optimal Q value with Monte Carlo simulation [48]. The general procedure of the algorithm is described below:

In Q-learning, a Q value is assigned to each state-action pair (s, a). Each triplet of Q, s and a creates an entry in a Q-table. Initially, all the Q values are set to zero.

Then, when a state s_t is observed at time t, the agent first searches for a set of available actions A_t. The optimal action a_t^* at s_t is determined by:

$$a_t^* = arg \max_{a_j \in A_t} Q(s_t, a_j) \tag{5}$$

where A_t is the set of available actions at s_t. If multiple a_t^* exists, a random tie-breaker will be chosen as the a_t^*. Note that (5) favors the maximum of total rewards Q instead of an immediate reward r_t to achieve the long-term optimality.

In search for the policy towards the optimal total rewards, random experiments are run repeatedly to update the Q value, during which the quality of the action sequence is improved towards the optimum:

$$Q(s_t, a_t) \leftarrow (1-\alpha) Q(s_t, a_t) + \alpha \{ r_{t+1}(s_t, a_t) + \gamma \max_a Q(s_{t+1}, a) \} \tag{6}$$

where α is the *learning rate* that controls the aggressiveness of learning. Empirically [47], setting $\alpha = 1$ will make the agent extremely aggressive, focusing on the immediate reward r_{t+1} received and the estimated approximate total future rewards $\max_a Q(s_{t+1}, a)$. This can lose the knowledge learned from previous experiments and cause unnecessary oscillations. On the contrary, setting $\alpha = 0$ will make the agent extremely conservative, as it sticks to its initial estimate and learns nothing from its actions. In practice, the value is often chosen as a trade-off between aggressiveness and conservativeness.

The updated Q value for the given state-action pair is saved in the Q-table for future decision-making process. It is possible that the agent makes non-optimal actions at the beginning of training when it tries to learn from the feedback of rewards. Eventually, the algorithm will converge to the optimal action sequence that collects the maximal total rewards [48].

1) Exploitation vs. Exploration: It is notable that Q-learning could be sensitive to deteriorate initialization and local optima problems. This can cost more learning time in practice. Therefore, *exploration* is commonly used in reinforcement learning. This paper utilizes the optimistic initial guess and the ϵ-greedy method for the exploration purpose.

The optimistic initial guess overcomes deteriorate initialization of Q-learning. It initializes the Q values of all valid actions of any state s_0 encountered for the first time to be a positive constant, e.g., $+1$, so that the agent is first encouraged to explore different actions and later adjust the Q value estimates towards the actual long term rewards.

The ϵ-greedy method is used to address the local optima problem during the learning process. Specifically, when the agent queries an action a_t from the Q table for the current state s_t, the ϵ-greedy method forces the agent to take a non-optimal valid action, i.e., any action other than a^* in Eqn. (5), with a small probability ϵ. Numerically, this means that the probability that the agent takes the optimal action is given by:

$$P(a_t = a_t^* \mid s_t) = 1 - \epsilon \tag{7}$$

where a_t^* is the optimal action at s_t according to (5). A proper choice of ϵ can balance the trade-off between exploitation and exploration so that the algorithm converges to the optimal policy in an effective and efficient manner. To avoid excessive exploration after the agent has learned much from the trials, the exploration parameter ϵ can starts from a larger initial value ϵ_0 then linearly decreases with a certain step distance $\delta\epsilon$ to a near-zero value ϵ_f, after which it remains constant. This allows the agent to sufficiently explore the searching space and then fine-tunes the Q value in a timely manner during the learning process.

B. Vulnerability Analysis With Q-Learning

The paradigm of Q-learning applies well to the vulnerability analysis of sequential attacks in the smart grid, as shown

TABLE I
PSEUDO CODE OF Q-LEARNING BASED VULNERABILITY ANALYSIS

Initialization: Initialize the Q-table and the benchmark system
for $current\ number\ of\ trials \leq maximal\ trials$ **do**
  Reset: $N_o = 0$, $s_0 = \mathbf{1}$;

  while $N_o \leq N_\theta$ **do**
    1. Acquire attack candidates: Obtain all valid line targets A_t from the current steady state s_t;
    2. Initiate an attack: Choose a line l from A_t and set its status $s_t(l) = 0$. Set $a_t = l$;
    3. Simulate cascading outages: With the attack updated in s_t, run the CFS until a new post-attack steady-state s_{t+1};
    4. Obtain evaluative feedback: Obtain N_o from s_{t+1} and generate the reward r_{t+1} according to (9);
    5. Learning from trial: Update the value of $Q(s_t, a_t)$ according to (6).
  end while
end for

in Fig. 4. First, an attacker, who seeks to identify the more vulnerable components in the power grid in sequential attacks, is considered as the agent in Q-learning. Next, the electrical power grid can be viewed as an independently operating environment that responds to the malicious actions of attackers. The action of the attacker is the malicious line-switching and the states can be defined by parameters of the system. Finally, the goal of attack is to produce a critical system failure where a fatal fraction of lines are out of service. The learning objective of the attacker is to find the optimal policy that reaches this goal with the least number of lines attacked. The pseudo code of the vulnerability analysis is shown in Table I, and the design of the state s_t, the action a_t, and the reward r_t are given as follows.

According to Assumption 1, a state is exclusively defined with the system topology as only the topological information is assumed available to a potential attacker. According to Assumption 2, a *state* also exclusively refers to the steady-state prior to sequential attacks. The initial state is $s = \mathbf{1}$; the intermediate states s_t are post-attack states after any cascading outages triggered by previous attacks. A complete system failure occurs when $s = \mathbf{0}$ and a critical failure occurs when the number of non-zero elements in s is dangerously low for a system. It is notable that a number of transitional states could exist between two consecutive attacks if there is a series of cascading outages. To evaluate the quality of sequential attack action, these transitional states are not considered as individual states during the Q-learning but only as intermediate transitions between s_t and s_{t+1}.

First, the topological system state s_t is defined as a vector of line status $s_t = \{s_t(1), s_t(2), \ldots, s_t(N)\}$, where:

$$s_t(l) = \begin{cases} 0, & \text{if line } l \text{ is in-service at time } t \\ 1, & \text{if line } l \text{ is out-of-service at time } t \end{cases} \tag{8}$$

Second, an action is defined as the line-switching attack as in (1). An attack a_i on line l switches the status of l from in-service to out-of-service. The corresponding value in s_t is set to zero.

TABLE II
BENCHMARK SYSTEM INFORMATION

Benchmark	N_{bus}	N_{line}	Load (MW)	Capacity (MW)
IEEE 5-bus	5	6	1,000.0	1,530.0
IEEE RTS-79	24	38	2,850.0	3,405.0
IEEE 300-bus	300	411	23,525.8	32,678.4

Finally, the evaluative feedback, or the reward r, is a critical variable in reinforcement learning. Given the aforementioned attack objective to create a critical system failure, this paper proposes the following reward function:

$$r_{t+1}(s_t, a_t) = \begin{cases} +1, & \text{if } N_o \geq N_\theta \text{ and } k < N_\theta \\ -1, & \text{if } N_o \geq N_\theta \text{ and } k \geq N_\theta \\ 0, & \text{otherwise} \end{cases} \tag{9}$$

where N_o is the total number of lines outages used to define the *blackout size*. N_θ is the critical threshold at the point-of-no-return (the attack objective), and k is the number of attacks launched sequentially. A sequential attack scheme is successful if it achieves the objective with an amplification effect ($r = +1$), i.e., the entire scheme causes N_θ or more line outages with less than N_θ actions. Otherwise, it is either unsuccessful if it takes $k = N_\theta$ attacks to achieve the objective blackout size ($r = -1$) or neutral if the number of attacks and line outages are both still under N_θ ($r = 0$). In the last case, the sequential attack will be continued until the objective blackout size is reached.

IV. SIMULATIONS

A. Simulation Setup

The performance of the proposed Q-learning based vulnerability analysis is tested on three benchmarks: a small-size IEEE 5-bus test system [49], a mid-sized IEEE 24-bus reliability test system (RTS-79) [50], and a large-scale IEEE 300-bus system [51]. Parameters of the three test systems are provided in Table II and the one-line diagrams for the 5-bus and RTS-79 are shown in Fig. 5.

Different systems will have different critical levels when cascading outages lead to a system collapse. We defined the **blackout size** as the combined number of line outages caused by the direct line-switching attack and the cascading failures triggered by the direct attacks. For the smaller IEEE 5-bus benchmark, we consider the attack objective to be a complete system failure, i.e., $\theta = 100\%$ and $N_\theta = 6$; for the mid-sized IEEE RTS-79 system and the large-scale IEEE 300-bus system, we consider more realistic attack objective of a critical system failure, where after a fatal number of lines are taken out the system will lose its functionality. Subsequently, we consider $\theta = 20\%$ ($N_\theta = 8$) for the RTS-79 system and $\theta = 2.5\%$ ($N_\theta = 11$) for the 300-bus system.

The Monte Carlo simulations are used to address the randomness from the stochastic hidden line failures [45] in the proposed vulnerability analysis. As a common practice in the validation of power systems security analysis, repeated simulations on typical operating points help

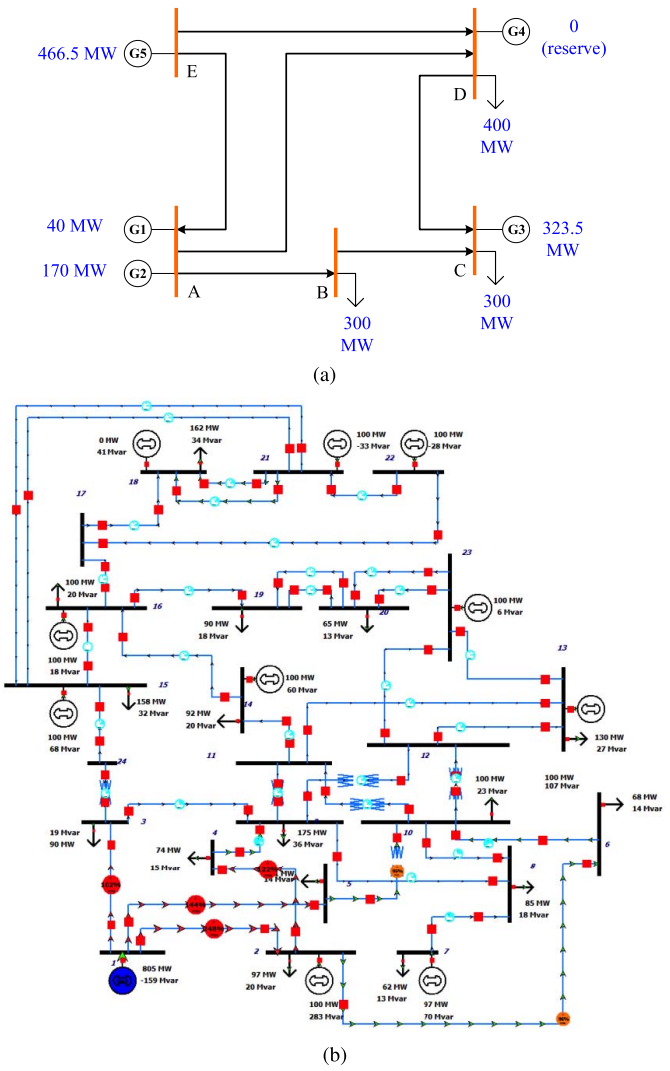


Fig. 5. The IEEE (a) 5-bus test system and (b) RTS-79 test system (second diagram from [50]).

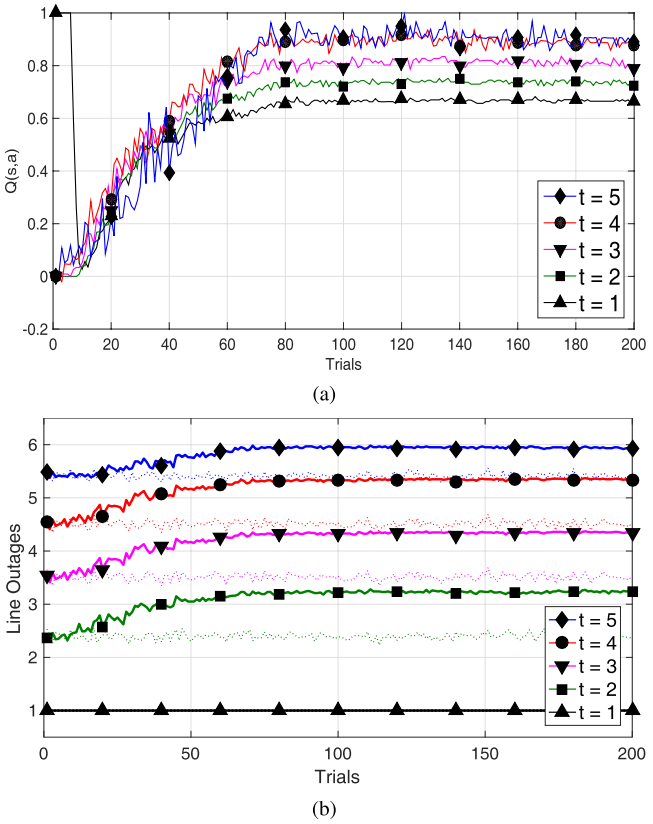


Fig. 6. Results from the IEEE 5-bus system: (a) the $Q(s, a)$ values of actions taken in each attack; (b) the number of line outages after each attack.

verify the effectiveness of the method. Specifically, a set of 100 independent simulations is first performed on each of the three benchmark systems. Every experiment consists of up to 1,000 trials, during which the Q-learning agent seeks attack target sequences to reach the critical blackout size N_θ as the attack objective. At the beginning of the first trial in each experiment, the initial Q values are set to $+1.0$ and the same for any new state-action pairs onward, which is an "optimistic" estimate to encourage exploration during the early stages. In each subsequent trials, the initial state is reset to the attack-free state with no line outages, while the Q values learned from previous trials are retained.

We start each experiment with a relatively large exploration probability of $\epsilon_0 = 0.3$ and decrease it to a small final value of $\epsilon_f = 0.005$, with a step-down by $\Delta\epsilon = -0.005$ after each trial. Meanwhile, following the common practice of Q-learning, we choose $\alpha = 0.1$ for a less aggressive learning process and $\gamma = 0.9$ to slightly weight more on the recent reinforcements [47].

B. Attack Performance Evaluation Under Typical Operating Points

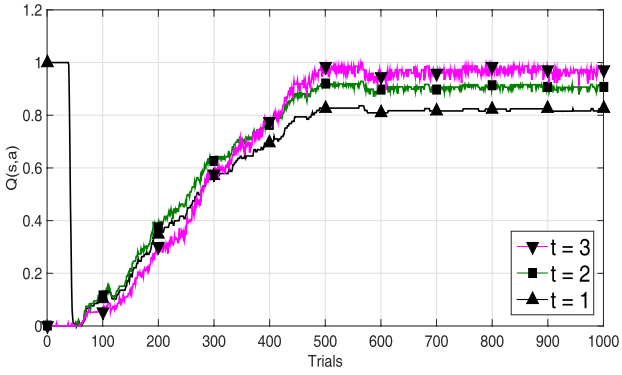
We first evaluates the performance on the IEEE 5-bus test system, for which the performance saturated after about 200 trials. Fig. 6 shows Q values of the action taken in each attack and the number of line outages afterwards. In Fig. 6, we first observed that the initialized Q value for $t = 1$ was quickly decreased from $+1.0$ to near-zero after 10 trials, showing that these early trials were mostly unsuccessful. Because of this cold start effect, these early trials from 100 experiments were also close to random search. This sharp decrease suggested that random or exhaustive search would be ineffective in identifying the vulnerable attack sequence. Meanwhile, after the initial trails, the Q-learning started to explore successful attack sequences and update the Q values quickly, reflecting the expected total reward learned from its trials. After about 100 trials, the Q values were stably increased to different ranges of expected total rewards for different rounds of attack in the sequence.

In addition to the estimated Q values, the number of line outages (the blackout size) is also shown in Fig. 6. As a comparison, the number of line outages caused by random attacks have also been plotted as unmarked dash-lines with the corresponding colors for different values of t. Except for the first attack ($t = 1$), the average blackout size after each attack in the sequence was increased during the Q-learning process, as shown in Table III. The improvement was most significant for the second attack ($t = 2$), with an increase

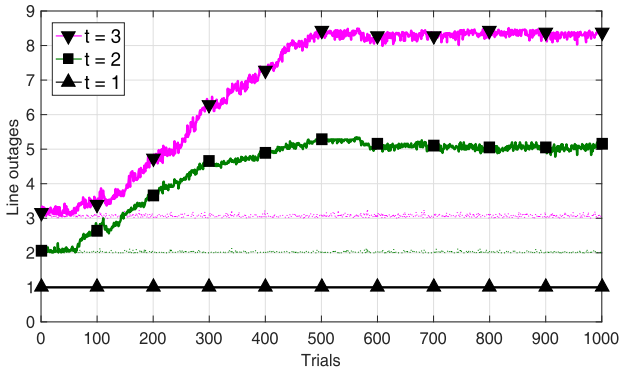
TABLE III

NUMBER OF LINE OUTAGES FROM SEQUENTIAL ATTACKS ON THE IEEE 5-BUS SYSTEM INCREASED BY Q-LEARNING

Order of attack	Initial	Eventual	Best of random attack
$t = 2$	2.37	**3.23**	2.62
$t = 3$	3.54	**4.34**	3.70
$t = 4$	4.54	**5.34**	4.77
$t = 5$	5.48	**5.94**	5.73



(a)

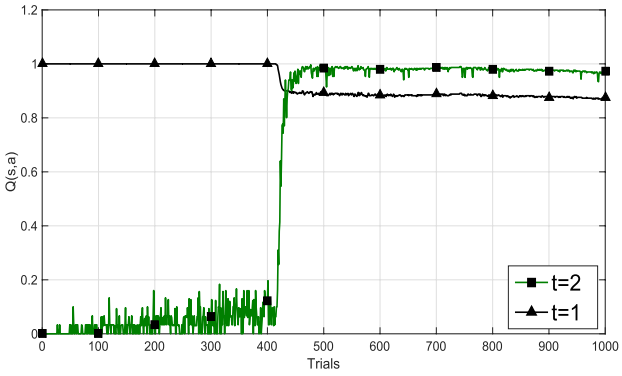


(b)

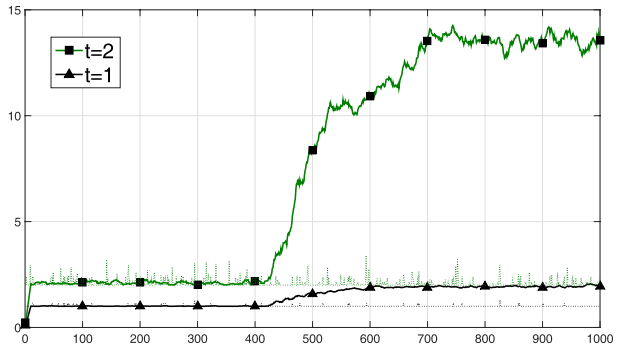
Fig. 7. Results from the IEEE RTS-79 system: (a) the $Q(s, a)$ values of the chosen action for each attack in the sequence and (b) the number of line outages after each attack.

of 0.86 additional line outages. The improvement from Q-learning for the fifth attack ($t = 5$) was 0.46 additional line outages, and the eventual blackout size on average (5.94) was close to the complete system failure of all six lines. Note that the system is $N - 1$ secured so that there is only one line outage after the first attack, but the corresponding Q value is about 0.66, indicating the expected total line outages in the future. On the contrary, the random attacks did not utilize any knowledge from previous trials and thus failed to identify more vulnerable sequences over time. These results on the 5-bus system have exhibited the effectiveness of the Q-learning based vulnerability analysis, learning from trials to reach the objective blackout size with purely topological information.

Fig. 7 shows the performance on the IEEE RTS-79 system. Similar to the 5-bus system, we also observed similar changes of Q value for the first to the third attack in the sequence in Fig. 7, with respective expected total rewards from the



(a)



(b)

Fig. 8. Results from the IEEE 300-bus system: (a) the $Q(s, a)$ values of the chosen action for each attack in the sequence and (b) the number of line outages after each attack.

attack. Meanwhile, with the Q-learning, as shown in Fig. 7, the blackout size was improved from 2.06 to 5.15 after the second attack ($t = 2$) and from 3.15 to 8.39 after the third attack ($t = 3$), respectively. Although the RTS-79 system is $N - 1$ secured, with the attack objective set as $N_\theta = 8$, it only took three sequential attacks to cause a critical 8-line blackout after 500 trials with Q-learning. In contrast, random attacks (unmarked dash-lines) did not show any improvement over time. The largest line outages achieved by random attacks was trivial compared to the Q-learning approach. With the same level of solely topological information, the proposed Q-learning based scheme was more effective in finding critical attack sequence on the RTS-79 system.

With the two case studies above, the proposed approach is further validated on a large-scale benchmark, the IEEE 300-bus system. The results are shown in Fig. 8. Similarly, the Q values converged through the learning process (Fig. 8) and the reinforcement improved the blackout sizes from the second attacks towards the attack objective, an 11-line outage, after 598 trials (Fig. 8) on average. Some oscillations persisted after 700 trials, as the final blackout size differed slightly due to cascading outages and hidden failures when the objective was achieved. The results validated that the proposed approach is scalable to bulk power systems.

The attack cost can also be evaluated by the average number of attacks required to achieve the objective blackout size, as shown in Fig. 9. The dashed lines indicate the respective objective blackout sizes, which are the unsuccessful (worst)

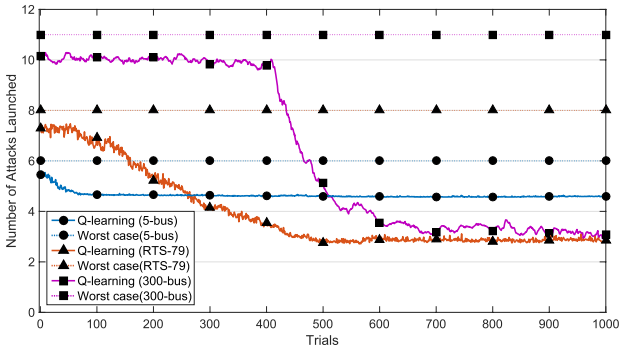


Fig. 9. Number of attacks taken to achieve the objective blackout size (dashed line) for the IEEE 5-bus, RTS-79, and 300-bus systems. The numbers reduced by the Q-learning exhibited the effectiveness of Q-learning in identifying more vulnerable attack sequences.

case for the attackers as they would need to attack as many lines as possible to achieve the attack objective. According to Fig. 9, the initial number of attacks required to reach the objective blackout sizes are 5.48, 7.29, and 10.14 for the 5-bus, the RTS-79, and the 300-bus systems, respectively. These numbers were reduced to 4.67 for the 5-bus system after 100 trials, 2.82 for the RTS-79 system after 500 trials, and 3.16 for the 300-bus system after 700 trials, respectively. Afterwards, their values remained in respective stable ranges, and the eventual numbers of sequential attacks launched to reach the attack objectives were 4.60, 2.85, and 3.09, respectively, for the three systems. From the simulations, the Q-learning based scheme also effectively reduced the number of attacks through the learning process.

C. Attack Performance With Varying Operating Points

As the load of power systems fluctuates with time, the robustness of the proposed approach shall be tested under different operating points (OP). Assuming a benchmark system's total default load is 100%, we solved the optimal power flow to obtain an OP with a peak load at 120% and another with a reduced load at 80%. Simulations were repeated on the three benchmark systems and the eventual blackout size when the attack objectives were achieved are shown in Table IV with different loading settings.

From the simulation, the proposed Q-learning based approach has successfully identified critical sequential topology attacks with respective objectives on all three benchmarks with different loading levels. In general, lowering loading levels reduced the system stress and make them more resilient to cascading outages and sequential topology attacks. In contrast, the peak loads turn $N-1$ secured systems (under default loading) more stressed and vulnerable. The blackout sizes resulting from the proposed Q-learning based vulnerability analysis are consistent with the statement above. As load increases, sequential attacks caused more line outages with the same number of attacks and attack objective were achieved faster.

V. Conclusions and Future Work

This paper has presented a novel Q-learning based vulnerability analysis of electrical power grid in the sequential topological attacks. By monitoring topology change in the system, the Q-learning based sequential scheme was able to find out vulnerable sequences that led to critical blackouts in the system. Not only did the scheme increase the number of line outages through the learning process, but it also reduced the number of attacks launched by excluding unsuccessful attack sequences that did not take advantage of the cascading outage vulnerability. Simulation results on the IEEE 5-bus, RTS-79, and IEEE 300-bus systems have demonstrated the learning ability and the effectiveness of the proposed approach.

From the perspective of a grid defender/operator, the Q-learning based vulnerability analysis can identify critical components in a potential sequential attack scheme. It also gives a warning sign that topological status information of the system could be utilized to conceive disastrous attack schemes. These insights will be helpful in improving situation awareness of the smart grid against cyber-attacks.

The future work will develop detection and mitigation strategies against sequential attacks. While the proposed approach utilizes reinforcement learning to screen potential vulnerable sequences of topological line-switching, it is not limited to the DC power flow model or the hidden failure model and adaptions can be made to consider other factors such as voltage and frequency for vulnerability analysis. Meanwhile, it is valuable to extend the manipulatable information to more sources and parameters to identify if they may result in more catastrophic attack impacts on the critical power infrastructure. Finally, the online learning ability can also be added to refine the learned vulnerabilities in real-world applications.

TABLE IV

Influence of Loading on the Eventual Blackout Sizes for the Q-Learning Based Sequential Topology Attacks

IEEE 5-bus	Default Load	Peak Load	Reduced Load
$t=1$	1.00	1.99	1.00
$t=2$	3.39	4.02	3.29
$t=3$	4.44	5.02	4.36
$t=4$	5.44	5.98	5.38
$t=5$	5.98	**6.00**	5.62
$t=6$	**6.00**	-	**6.00**
IEEE RTS-79	**Default Load**	**Peak Load**	**Reduced Load**
$t=1$	1.00	1.81	1.00
$t=2$	5.15	5.85	2.03
$t=3$	**8.38**	**8.40**	3.45
$t=4$	-	-	4.35
$t=5$	-	-	5.10
$t=6$	-	-	6.05
$t=7$	-	-	7.11
$t=8$	-	-	**8.19**
IEEE 300-bus	**Default Load**	**Peak Load**	**Reduced Load**
$t=1$	1.84	1.92	1.02
$t=2$	**13.91**	**14.72**	3.08
$t=3$	-	-	5.59
$t=4$	-	-	9.73
$t=5$	-	-	**13.10**

Note: A '-' indicates that the attack objective has already been achieved.

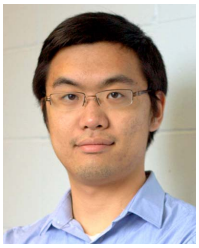
APPENDIX
THE CASCADING FAILURE SIMULATOR

Initialization: For a new attack a_i on line l, if a_i is valid, set $s^l \leftarrow 0$, $NewFailure = True$;
while $NewFailure = True$ **do**
1. Topology update: Remove $\forall l \in \{l|s^l = 0\}$ from the grid;
2. Re-dispatch: Recalculate power generation P_G and load P_D;
3. Power flow update: Re-calculate F_L according to the DC model;
4. Overload monitoring: If $F_l > C_l$, mark line l as overloaded;
5. Line outage risk: Calculate accumulated failure risk $O(l')$;
6. Overcurrent protection: If $O(l) \geq O_T$, then $s_l \leftarrow 0$ and $NewFailure = True$;
7. Hidden failure risk: Calculate hidden failure probability $p(l)$. If a hidden failure is triggered on line l, then $s_l \leftarrow 0$ and $NewFailure = True$; otherwise, if there is no overloading ($F_l < C_l$ for any remaining line l), $NewFailure = False$.
end while
Evaluation: Report blackout statistics.

REFERENCES

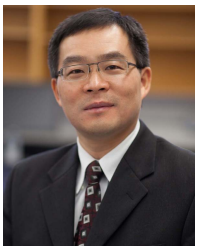
[1] Y. Mo *et al.*, "Cyber-physical security of a smart grid infrastructure," *Proc. IEEE*, vol. 100, no. 1, pp. 195–209, Jan. 2012.
[2] S. Sridhar, A. Hahn, and M. Govindarasu, "Cyber-physical system security for the electric power grid," *Proc. IEEE*, vol. 100, no. 1, pp. 210–224, Jan. 2012.
[3] (2009). *Public Notice Letter: Critical Cyber Asset Identification*. [Online]. Available: http://online.wsj.com/public/resources/documents/CIP-002-Identification-Letter-040609.pdf
[4] M. Govindarasu, A. Hann, and P. Sauer, "Cyber-physical systems security for smart grid," in *Proc. PSERC*, Feb. 2012, pp. 1–29.
[5] P.-Y. Chen, S.-M. Cheng, and K.-C. Chen, "Smart attacks in smart grid communication networks," *IEEE Commun. Mag.*, vol. 50, no. 8, pp. 24–29, Aug. 2012.
[6] O. Kosut, L. Jia, R. J. Thomas, and L. Tong, "Malicious data attacks on the smart grid," *IEEE Trans. Smart Grid*, vol. 2, no. 4, pp. 645–658, Dec. 2011.
[7] A. Delgadillo, J. Arroyo, and N. Alguacil, "Analysis of electric grid interdiction with line switching," *IEEE Trans. Power Syst.*, vol. 25, no. 2, pp. 633–641, May 2010.
[8] J.-W. Wang and L.-L. Rong, "Cascade-based attack vulnerability on the us power grid," *Safety Sci.*, vol. 47, no. 10, pp. 1332–1336, 2009.
[9] Z. Zhang, S. Gong, A. D. Dimitrovski, and H. Li, "Time synchronization attack in smart grid: Impact and analysis," *IEEE Trans. Smart Grid*, vol. 4, no. 1, pp. 87–98, Mar. 2013.
[10] X. Li, X. Liang, R. Lu, H. Zhu, X. Lin, and X. Shen, "Securing smart grid: Cyber attacks, countermeasures, and challenges," *IEEE Commun. Mag.*, vol. 50, no. 8, pp. 38–45, Aug. 2012.
[11] M. J. Eppstein and P. D. H. Hines, "A 'random chemistry' algorithm for identifying collections of multiple contingencies that initiate cascading failure," *IEEE Trans. Power Syst.*, vol. 27, no. 3, pp. 1698–1705, Aug. 2012.
[12] Q. Chen and J. D. McCalley, "Identifying high risk N-k contingencies for online security assessment," *IEEE Trans. Power Syst.*, vol. 20, no. 2, pp. 823–834, May 2005.
[13] V. Donde, V. López, B. Lesieutre, A. Pinar, C. Yang, and J. Meza, "Severe multiple contingency screening in electric power systems," *IEEE Trans. Power Syst.*, vol. 23, no. 2, pp. 406–417, May 2008.
[14] Y. Zhu, J. Yan, Y. Tang, Y. Sun, and H. He, "The sequential attack against power grid networks," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Jun. 2014, pp. 616–621.
[15] Y. Zhu, J. Yan, Y. Tang, Y. L. Sun, and H. He, "Resilience analysis of power grids under the sequential attack," *IEEE Trans. Inf. Forensics Security*, vol. 9, no. 12, pp. 2340–2354, Dec. 2014.
[16] J. Yan, Y. Tang, Y. Zhu, H. He, and Y. L. Sun, "Smart grid vulnerability under cascade-based sequential line-switching attacks," in *Proc. IEEE Global Commun. Conf., Sel. Areas Commun., Smart Grid Commun. (GC)*, San Diego, CA, USA, Dec. 2015, pp. 1–7.
[17] N. Fan, R. Chen, and J.-P. Watson, "N-1-1 contingency-constrained optimal power flow by interdiction methods," in *Proc. IEEE Power Energy Soc. General Meeting*, Jul. 2012, pp. 1–6.
[18] Y. Liu, P. Ning, and M. K. Reiter, "False data injection attacks against state estimation in electric power grids," *ACM Trans. Inf. Syst. Secur.*, vol. 14, no. 1, p. 13, 2011.
[19] J. Kim and L. Tong, "On topology attack of a smart grid: Undetectable attacks and countermeasures," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 7, pp. 1294–1305, Jul. 2013.
[20] G. Liang, J. Zhao, F. Luo, S. Weller, and Z. Y. Dong, "A review of false data injection attacks against modern power systems," *IEEE Trans. Smart Grid*, to be published.
[21] J. Yan, Y. Zhu, H. He, and Y. Sun, "Multi-contingency cascading analysis of smart grid based on self-organizing map," *IEEE Trans. Inf. Forensics Security*, vol. 8, no. 4, pp. 646–656, Apr. 2013.
[22] F. Li *et al.*, "Smart transmission grid: Vision and framework," *IEEE Trans. Smart Grid*, vol. 1, no. 2, pp. 168–177, Sep. 2010.
[23] M. Amin, "Energy infrastructure defense systems," *Proc. IEEE*, vol. 93, no. 5, pp. 861–875, May 2005.
[24] J. Yan, Y. Zhu, H. He, and Y. Sun, "Revealing temporal features of attacks against smart grid," in *Proc. IEEE PES Innov. Smart Grid Technol. (ISGT)*, Feb. 2013, pp. 1–6.
[25] Y. Zhu, J. Yan, Y. Tang, Y. L. Sun, and H. He, "Joint substation-transmission line vulnerability assessment against the smart grid," *IEEE Trans. Inf. Forensics Security*, vol. 10, no. 5, pp. 1010–1024, May 2015.
[26] *Analysis of the Cyber Attack on the Ukrainian Power Grid*, accessed Mar. 25. [Online]. Available: http://www.nerc.com/pa/CI/ESISAC/Documents/E-ISAC_SANS_Ukraine_DUC_18Mar2016.pdf
[27] *Cyber-Attack Against Ukrainian Critical Infrastructure*, accessed Mar. 22. [Online]. Available: https://ics-cert.us-cert.gov/alerts/IR-ALERT-H-16-056-01
[28] T. M. Chen, J. C. Sanchez-Aarnoutse, and J. Buford, "Petri net modeling of cyber-physical attacks on smart grid," *IEEE Trans. Smart Grid*, vol. 2, no. 4, pp. 741–749, Dec. 2011.
[29] P. Pourbeik, P. Kundur, and C. Taylor, "The anatomy of a power grid blackout—Root causes and dynamics of recent major blackouts," *IEEE Power Energy Mag.*, vol. 4, no. 5, pp. 22–29, Sep. 2006.
[30] J. Yan, Y. Tang, B. Tang, H. He, and Y. L. Sun, "Power grid resilience against false data injection attacks," in *Proc. IEEE Power Energy Soc. General Meeting*, Jul. 2016, pp. 1–5.
[31] M. Ozay, I. Esnaola, F. T. Y. Vural, S. R. Kulkarni, and H. V. Poor, "Machine learning methods for attack detection in the smart grid," *IEEE Trans. Neural Netw. Learn. Syst.*, vo. 27, no. 8, pp. 1773–1786, Aug. 2016.
[32] N. Romero, N. Xu, L. Nozick, I. Dobson, and D. Jones, "Investment planning for electric power systems under terrorist threat," *IEEE Trans. Power Syst.*, vol. 27, no. 1, pp. 108–116, Feb. 2012.
[33] J. Salmeron, K. Wood, and R. Baldick, "Worst-case interdiction analysis of large-scale electric power grids," *IEEE Trans. Power Syst.*, vol. 24, no. 1, pp. 96–104, Jan. 2009.
[34] L. Zhao and B. Zeng, "Vulnerability analysis of power grids with line switching," *IEEE Trans. Power Syst.*, vol. 28, no. 3, pp. 2727–2736, Aug. 2013.
[35] X. Liu, K. Ren, Y. Yuan, Z. Li, and Q. Wang, "Optimal budget deployment strategy against power grid interdiction," in *Proc. IEEE INFOCOM*, Apr. 2013, pp. 1160–1168.
[36] S. Liu, S. Mashayekh, D. Kundur, T. Zourntos, and K. Butler-Purry, "A framework for modeling cyber-physical switching attacks in smart grid," *IEEE Trans. Emerg. Topics Comput.*, vol. 1, no. 2, pp. 273–285, Dec. 2013.
[37] S. Liu, X. Feng, D. Kundur, T. Zourntos, and K. Butler-Purry, "A class of cyber-physical switching attacks for power system disruption," in *Proc. 7th Annu. Workshop Cyber Secur. Inf. Intell. Res. (CSIIRW)*, 2011, pp. 16:1–16:1.
[38] M. Vaiman *et al.*, "Risk assessment of cascading outages: Methodologies and challenges," *IEEE Trans. Power Syst.*, vol. 27, no. 2, pp. 631–641, May 2012.
[39] J. Yan, Y. Tang, H. He, and Y. Sun, "Cascading failure analysis with DC power flow model and transient stability analysis," *IEEE Trans. Power Syst.*, vol. 30, no. 1, pp. 285–297, Jan. 2015.
[40] C. Luo, J. Yang, Y. Sun, J. Yan, H. He, and M. Liu, "A cascading failure simulation model considering frequency dynamics and power flow distribution," in *Proc. North Amer. Power Symp. (NAPS)*, Oct. 2015, pp. 1–6.
[41] "Final report on the August 14, 2003 blackout in the United States and Canada: Causes and recommendations," U.S.-Canada Power Syst. Outage Task Force, Tech. Rep., Apr. 2004.
[42] R. Albert, H. Jeong, and A. Barabási, "Error and attack tolerance of complex networks," *Nature*, vol. 406, no. 6794, pp. 378–382, 2000.

[43] J. Yan, H. He, and Y. Sun, "Integrated security analysis on cascading failure in complex networks," *IEEE Trans. Inf. Forensics Security*, vol. 9, no. 3, pp. 451–463, Mar. 2014.
[44] I. Dobson, B. A. Carreras, V. E. Lynch, and D. E. Newman, "Complex systems analysis of series of blackouts: Cascading failure, critical points, and self-organization," *Chaos, Interdiscipl. J. Nonlinear Sci.*, vol. 17, no. 2, p. 026103, 2007.
[45] J. Chen, J. S. Thorp, and I. Dobson, "Cascading dynamics and mitigation assessment in power system disturbances via a hidden failure model," *Int. J. Elect. Power Energy Syst.*, vol. 27, no. 4, pp. 318–326, May 2005.
[46] B. Stott, J. Jardim, and O. Alsac, "DC power flow revisited," *IEEE Trans. Power Syst.*, vol. 24, no. 3, pp. 1290–1300, Aug. 2009.
[47] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*, vol. 1. Cambridge, MA, USA: MIT Press, 1998.
[48] C. J. C. H. Watkins, "Learning from delayed rewards," Ph.D. dissertation, King's College, Cambridge, U.K., 1989.
[49] F. Li and R. Bo, "Small test systems for power system economic studies," in *Proc. IEEE Power Energy Soc. General Meeting*, Jul. 2010, pp. 1–4.
[50] *IEEE 24-Bus System*, accessed on Jan. 4, 2016. [Online]. Available: http://publish.illinois.edu/smartergrid/ieee-24-bus-system/
[51] R. D. Zimmerman, C. E. Murillo-Sánchez, and R. J. Thomas, "MATPOWER: Steady-state operations, planning, and analysis tools for power systems research and education," *IEEE Trans. Power Syst.*, vol. 26, no. 1, pp. 12–19, Feb. 2011.



Jun Yan (S'13) received the B.Eng. degree in information and communication engineering from Zhejiang University, Hangzhou, China, in 2011, and the M.S. degree in electrical engineering from the University of Rhode Island, Kingston, RI, USA, in 2013, where he is currently pursuing the Ph.D. degree with the Department of Electrical, Computer and Biomedical Engineering.

His current research interest includes smart grid security and vulnerability, cyber-physical systems, reinforcement learning, deep learning, and computational intelligence.



Haibo He (SM'11) received the B.S. and M.S. degrees in electrical engineering from the Huazhong University of Science and Technology, China, in 1999 and 2002, respectively, and the Ph.D. degree in electrical engineering from Ohio University in 2006. He was an Assistant Professor with the Department of Electrical and Computer Engineering, Stevens Institute of Technology, from 2006 to 2009. He is currently the Robert Haas Endowed Chair Professor with the Department of Electrical, Computer, and Biomedical Engineering, University of Rhode Island.

His research interests include computational intelligence, machine learning and data mining, and various applications. He has authored one sole-author research book (Wiley), edited one book (Wiley-IEEE), and six conference proceedings (Springer). He has authored and coauthored over 200 peer-reviewed journal and conference papers. He severed as the General Chair of the IEEE Symposium Series on Computational Intelligence (SSCI 2014). He was a recipient of the National Science Foundation CAREER Award in 2011, the Providence Business News Rising Star Innovator Award in 2011, the IEEE International Conference on Communications Best Paper Award in 2014, and the IEEE Computational Intelligence Society Outstanding Early Career Award in 2014. He is currently the Editor-in-Chief of the IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS.

Xiangnan Zhong received the B.Eng. degree in automation and the M.S. degree in control theory and control engineering from Northeastern University, Shenyang, China, in 2010 and 2012, respectively. She is currently pursuing the Ph.D. degree with the Department of Electrical, Computer, and Biomedical Engineering, University of Rhode Island, RI, USA.

Her research interests include adaptive dynamic programming, reinforcement learning, neural network, and optimal control.

Yufei Tang (M'16) received the B.Eng. and M.Eng. degrees from Hohai University, Nanjing, China, in 2008 and 2011, respectively, and the Ph.D. degree in electrical engineering from the Department of Electrical, Computer, and Biomedical Engineering, The University of Rhode Island, Kingston, RI, USA, in 2016.

He is currently an Assistant Professor with the Department of Computer & Electrical Engineering and Computer Science, and a Faculty Fellow with the Institute for Sensing and Embedded Network Systems Engineering, Florida Atlantic University, Boca Raton, FL, USA. His research interests include power systems stability and control, smart grid, computational intelligence, and cyber-physical systems.