Vulnerability Assessment of Power System to Multi-step Stealthy False Data Injection Attacks

Mohammadmahdi Asghari, Amir Ameli, Mohsen Ghafouri, and Mohammad N. Uddin

***Abstract*—Stealthy false data injection attacks (SFDIAs) targeting state estimation can bypass the bad data detection module, mislead operators with false system states, and potentially result in erroneous decisions and physical damages. While most existing studies focus on single-step SFDIAs, multi-step SFDIAs pose a greater threat due to their forward-looking nature, where each step is strategically planned to amplify the cumulative impact. Therefore, this paper focuses on multi-step SFDIAs and presents a vulnerability assessment framework that leverages a Markov decision process (MDP) and bi-level optimization to quantify the system vulnerability to this type of attack. The MDP models the sequential and strategic nature of these attacks, with states reflecting evolving system conditions influenced by prior actions. At each state, actions derived through bi-level optimization identify attack vectors that maximize line overloads, potentially triggering the tripping of transmission lines. The MDP is solved using *Q*-learning, enabling the calculation of a vulnerability index that assists operators in assessing the impact of multi-step SFDIAs and identifying the attacker's most critical action at each step of multi-step SFDIAs. The effectiveness of the proposed vulnerability assessment framework is validated through simulations on the IEEE 39-bus test system.**

***Index Terms*—Bi-level optimization, vulnerability assessment, Markov decision process, false data injection attack, *Q*-learning.**

I. Introduction

As power systems transition into the digital era, they have become deeply intertwined with advanced communication and control technologies. While this digitization offers enhanced efficiency and management, power systems are exposed to various cyber threats. Among these threats, stealthy false data injection attacks (SFDIAs) targeting state estimation (SE) are particularly important, since they can bypass the bad data detection (BDD) module within SE and present falsified states to the operator in the control center, potentially leading to wrong decisions and actions [1]. Such attacks, in extreme cases, could trigger widespread blackouts [2].

To assess the impact of SFDIAs on the system, bi-level optimization is commonly used [2]-[8]. In this optimization, the upper level models the attacker's objectives and constraints for altering the data, while the lower level models the operator's response to the manipulated data. This framework determines the maximum potential damage an attacker can inflict on the system by manipulating measurements. For example, [3] presents a bi-level optimization that captures the interaction between a strategic attacker aiming to manipulate market outcomes such as locational marginal prices (LMPs), generation dispatch, congestion patterns, and a market operator who responds optimally to mitigate the impact of the attack.

Among the various potential impacts of SFDIAs such as inflicting physical damage [7], [9]-[11], increasing operation costs [4], changing LMPs [6], and modifying network parameters [12], physical damage is particularly critical because it directly compromises the stability and reliability of the system. Physical damage can occur when an attacker misleads the operator into making decisions or taking actions that violate the operation constraints of the grid [7], [9], [10]. For instance, as detailed in [7], manipulating data through an SFDIA can mislead an operator into dispatching generators in a way that causes one or more transmission lines to become significantly overloaded. If these overloads are large enough, protective devices may trip the affected line or lines. References [2] and [11] demonstrate that an SFDIA could overload several lines simultaneously, leading to cascading failures.

In assessing the impacts of SFDIAs on the system, a prevailing assumption in much of the literature is that an attacker executes the SFDIA in a single step [2]-[14]. Based on this premise, SFDIAs can have a considerable physical impact on the system only under specific conditions: when the power system is operating close to its operation constraints [2], or when the SFDIAs are coordinated with a topology attack [15], [16]. If these conditions are not met, the applied attack may not substantially affect system stability or performance. However, this assumption might not always hold, since attackers who have access to a data set can manipulate it over several steps. For instance, an attacker can deploy SFDIAs strategically across multiple steps, with the primary goal of amplifying the cumulative impact.

Manuscript received: December 12, 2024; revised: April 15, 2025; accepted: July 3, 2025. Date of CrossCheck: July 3, 2025. Date of online publication: August 20, 2025.

This work was supported in part by Natural Sciences and Engineering Research Council (NSERC) of Canada (No. RGPIN-2021-04042) and in part by Public Safety Canada (No. PSIMS22168).

M. Asghari (corresponding author), A. Ameli, and M. N. Uddin are with the Eelctrical and Computer Engineering Department, Lakehead University, Thunder Bay, Canada (e-mail: masghar2@lakeheadu.ca; aameli@lakeheadu.ca; muddin@lakeheadu.ca).

M. Ghafouri is with the Concordia Institute for Information Systems Engineering, Concordia University, Montreal, Canada (e-mail: Mohsen.ghafouri@concordia.ca).

DOI: 10.35833/MPCE.2024.001332

Multi-step attacks are commonly proposed in literature for topology attacks, where adversaries directly trip transmission lines [17] - [19]. The existing multi-step topology attacks, however, differ from multi-step SFDIAs in three aspects. ① Operators' decisions in topology attacks are made based on genuine data [17]-[19], while the decision-making process in multi-step SFDIAs is based on manipulated data [2], [4]-[7]. ② The tripping of lines in each step of topology attacks is immediate [17] - [19], whereas an SFDIA is more gradual, which relies on the delayed system response (due to generators' ramp rates) to manipulated data, progressively overloading lines and eventually triggering protection trips [7]. ③ Due to the gradual behavior of multi-step SFDIAs and their distinct threat model, the possibility of restoring previously tripped lines should be considered. To the best of the authors' knowledge, there is only one study in the literature that focuses on multi-step SFDIAs [20]. While this paper addresses multi-step SFDIAs, it does not prioritize maximizing the cumulative impact and does not consider the influence of each step on others. In other words, it develops a sequence of arbitrary independent single-step SFDIAs. Consequently, this method potentially underestimates the severity and effectiveness of multi-step SFDIAs. Additionally, this paper does not address the uncertainties within the system such as the possibility of restoring previously tripped lines.

Based on the above literature review, there is a research gap in methods for assessing power system vulnerability to forward-looking multi-step SFDIAs, in which attackers strategically plan steps to amplify the cumulative impact. These attacks, despite requiring no additional access compared with single-step methods, pose significant challenges to traditional security assessments. To address these research gaps, this paper develops a framework that utilizes the Markov decision process (MDP) to quantify the impact of multi-step SFDIAs on the system and determine the attacker's most critical action at each step of the attack. Each state in this MDP model encapsulates all the information regarding demand, generation, and the admittance matrix of the system, and is the outcome of a preceding action. The states are derived by accounting for the potential restoration of previously tripped lines (due to prior attack steps). To determine each action in the MDP, i.e., a one-step SFDIA that maximizes the power flow of a target line, a bi-level optimization problem is solved. The output of this optimization is an attack vector that potentially leads to overloading and tripping of the target line. The rewards in the proposed MDP are determined by the impact of actions, specifically the amount of load shed, the unused generation capacity, and the proximity to the operation margin. To assess the impact of multi-step SFDIAs on the system and identify the most critical action at each step, the developed MDP is solved using model-free *Q*-learning. In summary, the contributions of this paper are as follows.

1) This paper highlights that focusing defense mechanisms exclusively on the most severe single-step SFDIAs, as proposed in current research, may be inadequate. Instead, the defense mechanism should account for multi-step SFDIAs, in which the attacker employs a forward-looking strategy to meticulously design each step to progressively increase the cumulative impact on subsequent steps.

2) A vulnerability assessment framework is developed that employs an MDP and bi-level optimization to calculate an index quantifying the system vulnerability to multi-step SFDIAs. This framework also identifies the most critical action at each operation state, helping the operator to defend against attacks more effectively.

II. Impacts of Multi-step SFDIAs: A Simple Case Study

The IEEE 39-bus test system shown in Fig. 1 is employed in this paper as the primary platform for assessing the impact of SFDIAs on the system. The parameters of the generators are outlined in Table I, where c_g is the generation cost of generator g; $P_g^{\max}$ and $P_g^{\min}$ are the maximum and minimum possible active power of generator g, respectively; and M_g is the ramp rate of generator g.

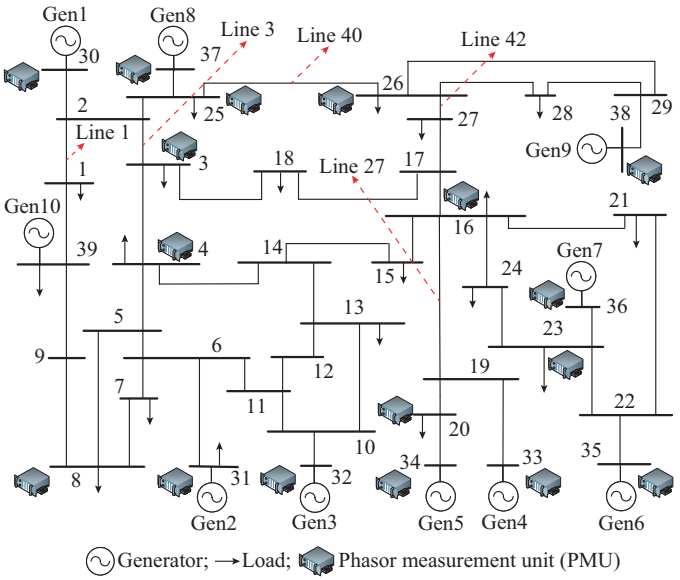

Fig. 1. IEEE 39-bus test system.

TABLE I
Generator Specifications

Generator	c_g ($/MW)	$P_g^{\max}$ (MW)	$P_g^{\min}$ (MW)	M_g (MW/min)
Gen1	10	900	0	10
Gen2	12	1000	0	10
Gen3	14	900	0	10
Gen4	16	900	0	15
Gen5	18	900	0	20
Gen6	20	900	0	30
Gen7	17	1000	0	25
Gen8	15	900	0	10
Gen9	13	1000	0	10
Gen10	11	1100	0	10

Additionally, the configuration data are sourced from the MATPOWER package [21]. Given that NERC Standard PRC-023-1 recommends setting the tripping threshold for

protective devices of power lines and transformers to be at least 150% of their maximum capacity [22], this paper adopts 150% as the relay pickup threshold [2], [23]. In practice, however, this threshold might be lower. For instance, during the 2003 Northwest Blackout, the relays of a transformer tripped 40 s after its power reached 130% of its capacity [24].

This section considers two scenarios as follows.

1) Scenario 1: the adversary damages the system by injecting false data in multiple steps. In each step, the adversary aims to maximize the attack impact without considering future steps of the attack.

2) Scenario 2: the adversary employs a forward-looking strategy, meticulously designing each step of the attack to progressively increase the cumulative impact on subsequent steps. In this scenario, the adversary aims to maximize the overall impact of the entire attack.

In both scenarios, the attacker intends to trip as many lines as possible to move the system toward a cascading failure. The attacks are executed stealthily, ensuring that ① the attack vector is crafted to bypass the BDD module, ② the manipulation of each measurement remains within 50% of its original value, and ③ only load and line measurements are manipulated, since generator measurements are easily detected if manipulated. Additionally, for conceptual clarity, these scenarios disregard generator ramp rates, allowing lines to become overloaded and trip before previously tripped lines are restored. This simplification is later relaxed in subsequent sections to enhance the practicality of the attack analysis.

At step 1 of scenario 1, the attacker performs a bi-level optimization for each line, as explained in Section III, to find the attack vector that maximizes the power flow of that line through a single-step SFDIA. In Fig. 2, the loading condition of each line is illustrated relative to its capacity after the attack vectors are injected and the operator responds accordingly. As the figure shows, the attacker can only increase the flow of line 3 (between buses 2 and 3) and line 27 (between buses 16 and 19) above the threshold. Given that line 27 is a major line connecting generators 4 and 5 to the rest of the system, its disconnection results in a severe power imbalance. More specifically, after disconnecting line 27, the region containing generators 4 and 5 would experience a surplus of 760 MW, while the remaining system would face a corresponding deficit. If the system lacks sufficient capacity to manage this imbalance, load shedding might be required to stabilize the network. In contrast, the loss of line 3, though impactful, does not result in such a drastic disconnection or severe imbalance. Therefore, in scenario 1, as the attackers seek to inflict maximum damage in each step of the attack, they opt to trip line 27. At step 2 of scenario 1, the attacker runs bi-level optimizations again to find the attack vectors that maximize the power flow of each line after the tripping of line 27. As shown in Fig. 3, following the tripping of line 27, the attacker is unable to trip any other line while maintaining all stealthiness conditions.

At step 1 of scenario 2, the attacker performs a similar bi-level optimization but with a forward-looking objective. On the other hand, if the attackers aim to fulfill the objective outlined in scenario 2, they should opt to trip line 3 instead of line 27. While this action may not cause the maximum immediate damage to the system, it strategically sets the stage for more severe damage in subsequent steps. As shown in Fig. 4, following the tripping of line 3, lines 1, 27, 40, and 42 emerge as potential targets for disconnection through an SFDIA at step 2 of scenario 2. Notably, line 27 appears among the potential lines for disconnection, indicating that scenario 2 not only achieves the objective of scenario 1 but also surpasses it. To maximize the damage within two steps, the attacker may choose to disconnect line 42 after the loss of line 3 at step 1 of scenario 1. This action would force the operator to shed at least 358 MW to ensure safe system operations.

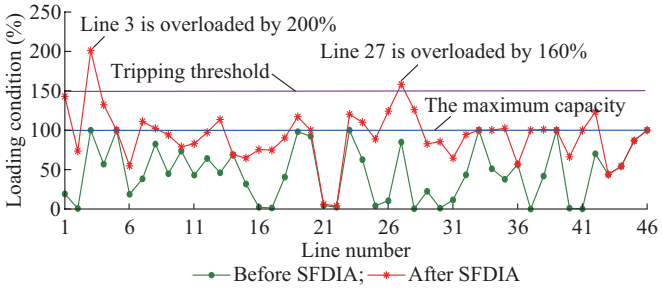

Fig. 2. Loading condition of each line after injecting its associated attack vector at step 1 of scenario 1.

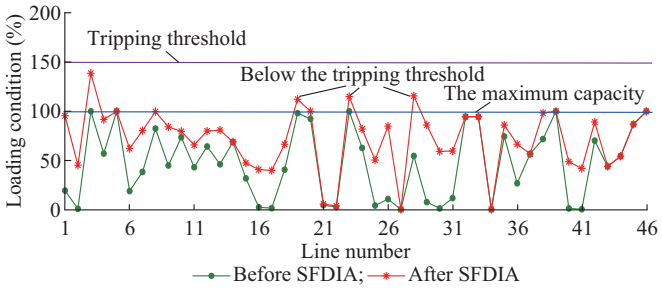

Fig. 3. Loading condition of each line after injecting its associated attack vector at step 2 of scenario 1.

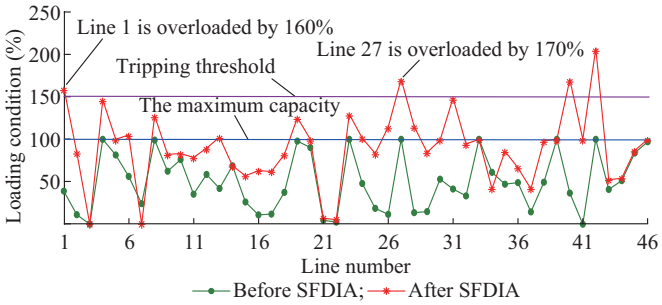

Fig. 4. Loading condition of each line after injecting its associated attack vector at step 2 of scenario 2.

This simplified case study clearly demonstrates that prioritizing defense mechanisms solely based on the most severe single-step SFDIAs, as suggested in the existing literature, may not be adequate. Instead, the defense mechanisms must also account for multi-step SFDIAs, which have the poten-

tial to inflict a more severe impact.

III. Essential Elements of Proposed Vulnerability Assessment Framework: Bi-level Optimization and MDP

This section details the essential tools needed to develop the proposed vulnerability assessment framework.

A. Bi-level Optimization for Overloading Transmission Lines

This subsection presents a bi-level optimization model, as shown in Fig. 5, to identify the attack vector that maximizes the line flow while satisfying stealthiness conditions. The upper level represents the attacker's objective in maximizing the line flow stealthily. The objective of the lower level is to minimize the operation cost, modeling the response of the system operator to the manipulated data and utilizing an optimal power flow (OPF). The upper level of the optimization can be formulated as:

$$\max_{\boldsymbol{Z}_M^r:\boldsymbol{Z}_{M|L}^r,\boldsymbol{Z}_{M|D}^r} \quad PL(r)=\boldsymbol{SF}_r\cdot(\boldsymbol{BG}\cdot\boldsymbol{PG}-\boldsymbol{BL}\cdot\boldsymbol{PD}) \tag{1}$$

where $PL(r)$ is the power flowing through line r; $\boldsymbol{PG}$ is the power generation vector derived from the lower level of the optimization, reflecting the operator's response to the attacked measurements; $\boldsymbol{BG}$ is the bus-generator incidence matrix; $\boldsymbol{BL}$ is the bus-load incidence matrix; $\boldsymbol{PD}$ is the actual demand vector; $\boldsymbol{Z}_M^r$ is the decision variable representing the attack vector injected into the system measurements to maximize the power flow in line r; $\boldsymbol{Z}_{M|L}^r$ and $\boldsymbol{Z}_{M|D}^r$ are the sub-vectors representing the data injected into line and load measurement, respectively; and $\boldsymbol{SF}_r$ represents the line r of the shifting factor matrix $\boldsymbol{SF}$.

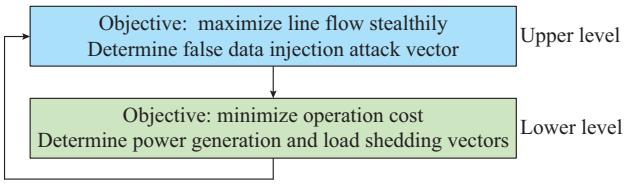

Fig. 5. Bi-level optimization model.

Given that the attacker aims to remain stealthy while altering the data, the upper level is subjected to the following constraints:

$$\left(\boldsymbol{H}\left(\boldsymbol{H}^{\mathrm{T}}\boldsymbol{H}\right)^{-1}\boldsymbol{H}^{\mathrm{T}}-\boldsymbol{I}\right)\boldsymbol{Z}_M^r=\boldsymbol{0} \tag{2}$$

$$\sum_{d\in\mathcal{D}} z_{m|d}^r=0 \tag{3}$$

$$\sigma P_d\le z_{m|d}^r\le\sigma P_d \quad \forall d\in\mathcal{D} \tag{4}$$

$$\eta\le z_{m|l}^r\le\eta \quad \forall l\in\mathcal{L} \tag{5}$$

where $\boldsymbol{H}$ is the linearized Jacobian matrix; d and $\mathcal{D}$ are the index and set of the loads, respectively; l and $\mathcal{L}$ are the index and set of lines, respectively; P_d is the active power of load d; η and σ are the upper bounds for line and load measurement manipulations, respectively; and $z_{m|l}^r$ and $z_{m|d}^r$ are the elements of sub-vectors $\boldsymbol{Z}_{M|L}^r$ and $\boldsymbol{Z}_{M|D}^r$, respectively.

Constraint (2) ensures that the injected data can bypass the BDD module of SE [25], [26]; (3) preserves stealthy by ensuring load redistribution without altering the total load; and (4) and (5) limit the extent of manipulation in load and line measurements, respectively, preventing large detectable alterations in the data.

On the other hand, the lower level models the response of the operator to the attack vector specified at the upper level. The primary objective at the operator level is to minimize the operation costs of the system via OPF. The objective function of the operator level can be defined as:

$$\min_{\boldsymbol{PG},\boldsymbol{SD}}\left(\sum_{g\in\mathcal{G}} c_g P_g+\sum_{d\in\mathcal{D}} cs_d\cdot S_d\right) \tag{6}$$

where g and $\mathcal{G}$ are the index and set of the generators, respectively; $\boldsymbol{SD}$ is the load shedding vector for all loads; S_d is the active power shed from load d; P_g is the active power of generator g; and cs_d is the load shedding cost for load d.

In (6), the first term computes the operation costs of generators, and the second term is the load shedding cost in the system. This optimization problem should be solved subject to the following constraints:

$$\sum_{g\in\mathcal{G}} BG_{j,g}\cdot P_g=\sum_{i\in\mathcal{B}} Y_{j,i}\theta_i+\sum_{d\in\mathcal{D}} BL_{j,d}\left(P_d-S_d+z_{m|d}^r\right) \quad \forall j\in\mathcal{B} \tag{7}$$

$$-P_l^{\max}\le\boldsymbol{SF}_l\cdot\left[\boldsymbol{BG}\cdot\boldsymbol{PG}-\boldsymbol{BL}\cdot\left(\boldsymbol{PD}+\boldsymbol{Z}_{M|D}^r\right)\right]\le P_l^{\max} \quad \forall l\in\mathcal{L} \tag{8}$$

$$P_g^{\min}\le P_g\le P_g^{\max} \quad \forall g\in\mathcal{G} \tag{9}$$

$$S_d\le P_d \quad \forall d\in\mathcal{D} \tag{10}$$

$$P_g\ge\max\left\{P_g^0-M_g T_h, P_g^{\min}\right\} \quad \forall g\in\mathcal{G} \tag{11}$$

$$P_g\le\min\left\{P_g^0+M_g T_h, P_g^{\max}\right\} \quad \forall g\in\mathcal{G} \tag{12}$$

where $BG_{j,g}$ is the element of bus-generator incidence matrix; θ_i is the voltage angle at but i; i and j are the indexes of buses; $\mathcal{B}$ is the set of buses; $P_l^{\max}$ is the long-term capacity of line l; $\boldsymbol{SF}_l$ represents the row l of $\boldsymbol{SF}$; T_h is the look-ahead time for one period OPF; $Y_{i,j}$ is the admittance in column j and row i; P_g^0 is the initial power output of generator g; and $BL_{j,d}$ is the element in row j and column d of this matrix.

Constraint (7) ensures the power balance at each bus; (8) limits the power flow across lines; (9) and (10) define the limits for generators and load shedding, respectively; and (11) and (12) specify the ramp rate limits of generators.

To solve this bi-level problem, the lower level can be equivalently reformulated using its corresponding Karush-Kuhn-Tucker (KKT) optimality conditions and incorporated into the attacker level. This technique merges the attacker and operator levels into a single-level problem.

B. MDP

A discrete-time MDP is defined by a 5-tuple (S,A,T,R,γ) that frames the decision-making landscape. S is the finite set of states where each state $s_t\in S$ encapsulates a possible sce-

nario or condition of the system at time t. A is the finite set of actions, where $a_t \in A$ is a possible action that can be performed at state s_t. Transition probabilities $T\left(s_{t+1}\middle|s_t,a_t\right)$ dictate the likelihood of moving from state s_t to a new state s_{t+1} following action a_t. The reward function $R\left(s_{t+1}\middle|s_t,a_t\right)$ represents the reward for transitioning from state s_t to s_{t+1} with action a_t. This function quantifies the immediate benefit derived from taking a particular action in a particular state at time t. Lastly, the discount factor γ, which varies between 0 and 1, balances the priority between immediate and future rewards.

The objective of MDP is to identify the optimal policy, π^*, which determines the best action at each state to maximize the expected reward. To solve the MDP, a model-free algorithm such as Q-learning can be utilized. Q-learning, which is a type of reinforcement learning, interacts directly with the environment and continuously updates its strategy based on the feedback received. Q-learning does not require a predefined model of the environment, making it ideal for this problem, where the system is uncertain and the dynamics are complex. By executing actions and observing the resulting rewards and state transitions, Q-learning incrementally refines its action-value function $Q(s,a)$, which predicts the expected utility of taking action a in state s, thereby improving the decision-making process over time. The Q-values are iteratively updated according to the following rule:

$$Q\left(s_t,a_t\right)Q\left(s_t,a_t\right)+\alpha\left(R\left(s_t,a_t,s_{t+1}\right)+\gamma\max_a Q\left(s_{t+1},a_t\right)-Q\left(s_t,a_t\right)\right) \tag{13}$$

where α is the learning rate; and $\max_a Q\left(s_{t+1}\right)$ is the highest Q-value achievable from the new state s_{t+1}. This update is performed iteratively for each state-action pair experienced, converging to the optimal action-value function $Q^*\left(s_t,a_t\right)$ as more updates are applied. Upon convergence, the optimal policy can be derived simply by selecting the action that maximizes the Q-value in each state:

$$\pi^*\left(s_t\right)=\arg\max_{a_t\in A} Q^*\left(s_t,a_t\right) \tag{14}$$

Once the optimal action-value functions are obtained for state s_t, the value function $V\left(s_t\right)$ for this state, which represents the maximum expected reward for being in state s_t and following the optimal policy, can be obtained using (15), indicating that $V\left(s_t\right)$ is the maximum Q-value over all actions a_t at state s_t.

$$V\left(s_t\right)=\max_{a_t\in A} Q^*\left(s_t,a_t\right) \tag{15}$$

IV. Proposed Vulnerability Assessment Framework for Assessing Impacts of Multi-step SFDIAs on Power Grids

This section delves into the structure of the proposed vulnerability assessment framework designed for analyzing the impacts of multi-step SFDIAs. The framework utilizes an MDP, as shown in Fig. 6, to provide a structured method for analyzing sequential decision-making in the context of multi-step SFDIAs. Before detailing the MDP structure, it is important to clarify that the proposed vulnerability assessment framework serves as a decision-support tool for system operators. Its purpose is to highlight critical vulnerabilities and inform defensive strategies in worst-case scenarios. It is assumed that the operator, with full system visibility, uses the framework to identify the most damaging actions an adversary could take at each step. However, if an attacker intends to carry out a multi-step attack, they require the same level of information and access as for a single-step attack (sustained over a longer period). This includes continuous access to PMU measurements (via substation compromise, communication interception, or PDC manipulation), generator parameters (e.g., costs, ramp rates, capacity limits), and transmission system data (e.g., topology, line ratings) throughout the attack sequence.

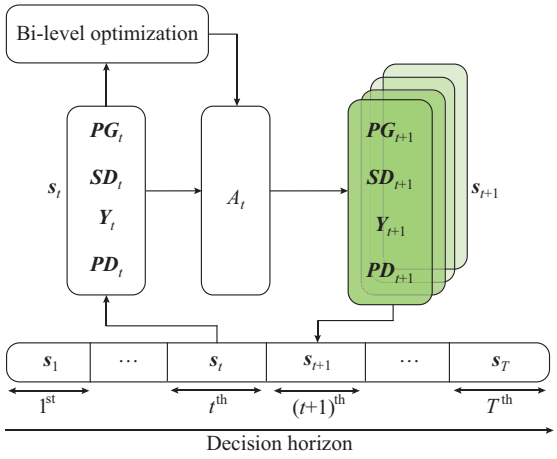

Fig. 6. Proposed MDP.

Within this MDP, each decision or action is an attack vector that overloads a line in the power system, determined by solving a bi-level optimization problem. Since the impacts of SFDIAs depend entirely on the system response to altered data (calculated through OPF), the decision intervals of the proposed MDP align with the intervals at which the operator computes the OPF solutions. Specifically, at each time interval, the attacker selects an action, i.e., an attack vector, by solving a bi-level optimization problem, which yields the corresponding generation dispatch and load shedding vectors. Moreover, using the obtained vectors and the probabilities of line restoration, the system topology is updated accordingly, and the admittance matrix is revised to reflect the new network configuration. Based on this process, the MDP structure, including its states, actions, and other essential elements, is detailed as follows.

A. States

In the MDP model, at any given time step $t\in\{1,2,...,T\}$, the state of a system is represented as:

$$\boldsymbol{s}_t=\left[\boldsymbol{PG}_t \quad \boldsymbol{SD}_t \quad \boldsymbol{PD}_t \quad \boldsymbol{Y}_t\right] \tag{16}$$

$\boldsymbol{Y}_t$ is an important component of $\boldsymbol{s}_t$, as it may change in response to line tripping caused by SFDIAs or the subsequent recovery from previous tripping.

Note that $\boldsymbol{PD}_t$ is independent of the MDP's internal dynamics, as it is assigned based on actual demand at the initial step and on forecasted values at subsequent steps.

B. Actions

The set of available actions A_t directly corresponds to the set of in-service lines $\mathcal{L}_t$, as determined by the $\boldsymbol{Y}_t$ component of $\boldsymbol{s}_t$. Each action $\boldsymbol{a}_t^l \in A_t$ encompasses an attack vector, i.e., $\boldsymbol{Z}_{M|D}^l$, which can overload line l at state $\boldsymbol{s}_t$. The attack vectors can be calculated by solving a bi-level optimization problem, as explained in Section III-A. The mathematical representation of actions is:

$$A_t = \left\{ \boldsymbol{a}_t^l = \boldsymbol{Z}_{M|D}^l \middle| \, l \in \mathcal{L}_t \right\} \tag{17}$$

C. Transition Probabilities

In the developed MDP, the state transition from $\boldsymbol{s}_t$ to $\boldsymbol{s}_{t+1}$ following action $\boldsymbol{a}_t^l$ involves deterministic components such as power generation and load shedding, as well as network configuration, which is treated probabilistically. The load demand in the next time step, $\boldsymbol{PD}_{t+1}$, is often known and available with high accuracy. Additionally, for a given state $\boldsymbol{s}_t$ and action $\boldsymbol{a}_t^l$, $\boldsymbol{PG}_{t+1}$ and $\boldsymbol{SD}_{t+1}$ can be calculated by solving the OPF problem, as detailed in (6)-(12). In this optimization, the initial power generation vector $\boldsymbol{PG}^0$ is set to be $\boldsymbol{PG}_t$, and the attack vector $\boldsymbol{Z}_{M|D}^l$ is the action $\boldsymbol{a}_t^l$.

The transition probability of the admittance matrix depends on the effectiveness of the action $\boldsymbol{a}_t^l$ in overloading line l at state $\boldsymbol{s}_{t+1}$, and the probability of restoring the previously tripped line r at $\boldsymbol{s}_{t+1}$. The transition probability for the admittance matrix can be defined as:

$$T\left(\boldsymbol{Y}_{t+1} \middle| \boldsymbol{Y}_t, \boldsymbol{a}_t^l\right) = \begin{cases} \mathcal{P}_S \mathcal{P}_R\left(t_{lost}\right) & \boldsymbol{Y}_{t+1} = \boldsymbol{Y}_t^{l,r} \\ \mathcal{P}_S\left(1 - \mathcal{P}_R\left(t_{lost}\right)\right) & \boldsymbol{Y}_{t+1} = \boldsymbol{Y}_t^{l} \\ \left(1 - \mathcal{P}_S\right)\mathcal{P}_R\left(t_{lost}\right) & \boldsymbol{Y}_{t+1} = \boldsymbol{Y}_t^{r} \\ \left(1 - \mathcal{P}_S\right)\left(1 - \mathcal{P}_R\left(t_{lost}\right)\right) & \boldsymbol{Y}_{t+1} = \boldsymbol{Y}_t \end{cases} \tag{18}$$

where $\boldsymbol{Y}_t^{l,r}$ is the admittance matrix after tripping line l and recovering line r; $\boldsymbol{Y}_t^l$ is the admittance matrix after tripping line l; $\boldsymbol{Y}_t^r$ is the admittance matrix after recovering line r; $\mathcal{P}_R\left(t_{lost}\right)$ is the probability of restoring a previously tripped line during t_{lost}, which is the period during which the line is out of service; and $\mathcal{P}_S$ is the probability that the action $\boldsymbol{a}_t^l$ increases the load on line l to the point of tripping. This probability is a function of the extent of the overload caused by the action $\boldsymbol{a}_t^l$, and is defined as:

$$\mathcal{P}_S = F\left(\frac{PL_{t+1}(l)}{P_l^{\max}}\right) \tag{19}$$

where F is the probability function that gives the likelihood of line l tripping based on the ratio of its power flow $PL_{t+1}(l)$ to its maximum capacity $P_l^{\max}$. $PL_{t+1}(l)$ can be calculated as:

$$PL_{t+1}(l) = \boldsymbol{SF}_l \cdot \left(\boldsymbol{BG} \cdot \boldsymbol{PG}_{t+1} - \boldsymbol{BL} \cdot \boldsymbol{PD}_{t+1}\right) \tag{20}$$

As discussed in Section II, NERC Standard PRC-023-1 recommends setting the tripping threshold for protective devices of lines and transformers at least 150% of their maximum capacity [22]. Accordingly, this paper adopts 150% as the pickup threshold for relays [2], [23]. Consequently, function F can be defined as a step function, where it equals 1 when the ratio exceeds 1.5 and 0 when it is below 1.5.

D. Reward

The reward function of the MDP is crucial for accurately quantifying the impacts of SFDIAs. Typically, the impact of losing a transmission line is measured by the amount of unserved load or required load shedding. Besides the impact on loads, losing a line can also lead to the loss of generators. Although this event might not directly affect the loads, it can result in significant economic losses, increased pressure on other generators, and a reduced safety margin due to the increased contribution from the remaining generators [27].

Based on the above potential impacts, this paper defines the reward function as:

$$R\left(\boldsymbol{s}_t, \boldsymbol{a}_t, \boldsymbol{s}_{t+1}\right) = RF\left(\boldsymbol{s}_{t+1}\right) - RF\left(\boldsymbol{s}_t\right) \tag{21}$$

$$RF\left(\boldsymbol{s}_t\right) = \beta_1 \sum_{d \in \mathcal{D}} S_{d,t} - \beta_2 \sum_{g \in \mathcal{G}} C_{g,t} + \beta_3 \sum_{d \in \mathcal{D}} \sum_{k \in \Omega^d} \max\left(0, 1 - \frac{\sum_{l \in \Omega^d, l \neq k} P_l^{\max}}{P_{d,t}}\right) P_{d,t} \tag{22}$$

where Ω^d is the set of transmission lines connected to load d; and β_1, β_2, and β_3 are the coefficients determining the relative importance of each term in the reward. In (22), the first term represents the total unserved load, the second term represents the available generation capacity, and the third term quantifies the system operation margin [27].

E. Dynamic Q-learning for Solving MDP

Dynamic Q-learning is employed to solve the developed MDP due to its capability to effectively handle the extensive and uncertain state space inherent in power system operation. This method dynamically adapts by creating Q-table entries only for state-action pairs as they are encountered. This avoids the computational impracticality of initializing a complete Q-table for the vast number of possible states. As a result, the proposed vulnerability assessment framework becomes more scalable and memory-efficient.

Algorithm 1 details the dynamic Q-learning process for solving the proposed MDP. Following initialization, the algorithm iterates over M episodes. In each episode, the initial state $\boldsymbol{s}_1$ of the power system is its current state. Within each episode, for every time step from 1 to T, the algorithm follows a structured sequence of operations. It begins by identifying in-service lines and solving a bi-level optimization problem to determine viable actions A_t. If the state-action pair $\left(\boldsymbol{s}_t, \boldsymbol{a}_t^l\right)$ is not already in the Q-table, it is initialized. An action $\boldsymbol{a}_t^l$ is then selected from A_t using the ϵ-greedy policy [28], which balances exploration and exploitation based on the Q-values. This action is executed to determine the deterministic elements of the next state $\boldsymbol{s}_{t+1}$. Then, the algorithm

calculates the probabilities $\mathcal{P}_S$ and $\mathcal{P}_R(t_{lost})$. Afterwards, the algorithm chooses $\boldsymbol{Y}_{t+1}$ based on $\mathcal{P}_S$ and $\mathcal{P}_R(t_{lost})$ using (18). The reward for the transition from $\boldsymbol{s}_t$ to $\boldsymbol{s}_{t+1}$ using the selected action $\boldsymbol{a}_t^l$ is calculated by (22), and the Q-table is updated using (13).

Algorithm 1: Dynamic Q-learning for solving MDP

1. Initialize a dynamic structure for the Q-table and the system benchmark
2. **for** *episode* = 1 to M **do**
3. Get initial state s_1
4. **for** time step $t = 1$ to T **do**
5. Obtain $\mathcal{L}_t$ and solve bi-level optimization problem (l) to derive A_t
6. If $(s_t, \boldsymbol{a}_t^l)$ is not in Q-table, initialize $Q(s_t, a_t^l)$ for all $a_t^l \in A_t$ to a default value, e.g., 0
7. Select an action a_t^l from A_t
8. Execute a_t^l, which determines the deterministic elements of s_{t+1}
9. Determine $\mathcal{P}_R(t_{\text{lost}})$ based on t_{lost}, and $\mathcal{P}_S$ based on (19)
10. Assign $\boldsymbol{Y}_{t+1}$ by selecting scenarios according to $\mathcal{P}_S$ and $\mathcal{P}_R(t_{\text{lost}})$
11. Calculate the reward $R(s_t, a_t^l, s_{t+1})$ based on (22)
12. Update Q-table dynamically as:

$$Q(s_t, a_t^l) \leftarrow Q(s_t, a_t^l) + \alpha\left(R(s_t, a_t^l, s_{t+1}) + \gamma \max_{a_{t+1}^l \in A_{t+1}} Q(s_{t+1}, a_{t+1}^l) - Q(s_t, a_t^l)\right)$$

13. **end for**
14. **end for**

When the optimal Q-value function $Q^*(\boldsymbol{s}_t, \boldsymbol{a}_t^l)$ is obtained using this algorithm, the operator can determine the next step of the most critical multi-step SFDIA by using (14). Moreover, the vulnerability index, which quantifies the potential impact of the SFDIA at the current state, is derived from (15). This is because the maximum Q-value at a given state $\boldsymbol{s}$ reflects the highest expected cumulative impact that can be inflicted starting from that state. As such, the maximum Q-value serves as a meaningful indicator of the system vulnerability to multi-step SFDIAs.

V. Potential Defensive Mechanism

Since multi-step SFDIAs consist of multiple single-step SFDIAs, existing defense mechanisms designed for single-step attacks can be adapted to multi-step SFDIAs with appropriate modifications. Existing defense mechanisms are examined in the following, highlighting their advantages and disadvantages, and discussing potential modifications for their application in multi-step SFDIAs. Additionally, a novel defense mechanism is proposed for future research to address the limitations of current methods.

A. Physical Layer-based Defense Mechanisms

Physical layer-based defense mechanisms focus on enhancing the system resilience against SFDIAs. For instance, [29]-[31] propose preventive generation dispatch strategies to mitigate the risk of line overloads and potential blackouts caused by SFDIAs. These methods ensure that, even when false data are injected into measurements, the lines are not overloaded. Similarly, [32] introduces a cyber-secured unit commitment method to ensure safe grid operation under SFDIAs. Therefore, since multi-step SFDIAs are composed of multiple single-step attacks, these physical layer-based defense mechanisms can also be employed for them.

Although these defense mechanisms effectively prevent line overloads, they lead to suboptimal economic performance, resulting in increased operation expenses. Consequently, since the proposed vulnerability assessment framework can identify the most critical initial step, it allows the operator to prioritize hardening these specific lines. This strategic focus not only reduces overall defense costs but also maintains system security.

B. Cyber Layer-based Defense Mechanisms

These cyber layer-based defense mechanisms typically focus on the cyber layer and are mainly based on securing certain PMUs from cyber-attacks. These defense mechanisms can be divided into two categories. The research in the first category such as in [33], [34] aims to protect the PMUs to ensure that no stealthy attacks can take place. For example, a bi-level optimization problem is formulated in [34] to identify the minimum number of measurements that need protection to prevent an attacker from executing an SFDIA. Using these methods could also prevent multi-step SFDIAs, since securing PMUs makes the first step of multi-step SFDIAs impossible. The main limitation of these techniques is that they often require a large number of PMUs to be secured, and achieving complete protection may not be feasible in some scenarios.

The research in the second category, as stated in [35] and [36], focuses on preventing high-impact SFDIAs by safeguarding critical PMUs. These methods often employ a tri-level optimization problem to model the interactions between the attacker and the power system operator. For example, a tri-level model is developed in [36] to identify the optimal set of measurements that should be protected to minimize the adverse impact of FDIAs on operation costs. Therefore, the proposed vulnerability assessment framework can assist operators in prioritizing critical PMUs for protection, akin to the methods used in this category. The cyber layer-based defense mechanisms address the main limitation of the first category by reducing the number of PMUs needing protection while maintaining the system resilience. However, the effectiveness of the mechanisms heavily depends on the operation points of the power system, and achieving complete protection for the selected PMUs remains a significant challenge.

C. Proposed Novel Cyber Layer-based Defense Mechanisms

The proposed novel cyber layer-based defense mechanism (hereafter called the proposed defense mechanism) targets the cyber layer, focusing primarily on altering the hierarchical data aggregation structure within the cyber layer of the power system. Changing this aggregation can significantly reduce the risk of SFDIAs. The design of multi-step SFDIAs often involves manipulating a set of PMUs multiple times during an attack. The access to these PMUs can be facilitated by targeting several substation PMUs, the communication links that transfer these data, and the PDCs that aggregate data from multiple PMUs. The current hierarchical data gathering structure in the power system creates multiple highly vulnerable points in the cyber layer, where the access to all critical PMU data can be obtained by attacking these points. By implementing software-defined networking, the hierarchi-

cal data gathering structure can be modified to eliminate these vulnerabilities, thus drastically reducing the likelihood of successfully designing multi-step SFDIAs [37]. Specifically, by using the proposed vulnerability assessment framework, operators can identify the critical PMUs essential for designing such attacks. Then, the operator can modify the hierarchical data aggregation structure in the cyber layer of the power system to increase the complexity for attackers, thereby reducing the risk of this type of attack. The advantage of the proposed defense mechanism is that it does not impose additional costs (unlike physical layer-based defense mechanisms), and it does not rely on securing PMUs, which may not be feasible.

VI. Simulation Results

This section demonstrates the effectiveness of the proposed MDP for assessing the impacts of multi-step SFDIAs on power grids through case studies on the IEEE 39-bus and 57-bus test systems.

A. IEEE 39-bus Test System

This subsection utilizes the IEEE 39-bus test system, which was introduced in Section II, to evaluate the proposed vulnerability assessment framework. Figure 7 illustrates the load profile of the test system. To supply the loads efficiently and reliably, OPF is run every 15 min, enabling the system to adapt to the dynamics within the same time frame.

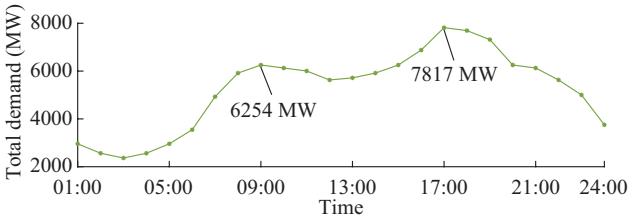

Fig. 7. Load profile of test system.

Next, the probability of restoring tripped lines, $\mathcal{P}_R(t_{\text{lost}})$, is defined. An exponential distribution is widely used to estimate the restoration time of transmission lines affected by physical causes such as weather events, natural disasters, or vegetation interference [38]. In traditional reliability analysis, the restoration rate λ is typically derived from historical outage and repair data, capturing the combined influence of infrastructure conditions, logistics, and workforce availability. In contrast, this paper addresses cyber-induced line trips, where coordinated attacks disrupt protective relays. Unlike physical damage, these events generally do not require extensive on-site repairs. The restoration may involve faster actions such as software reconfiguration, remote relay resets, or operator interventions. While the exponential model remains applicable, the interpretation of λ shifts, reflecting factors such as system monitoring, automation, cyber-response protocols, and the maturity of detection and mitigation mechanisms.

On this basis, the probability of restoring a previously tripped line in this test system is defined as:

$$\mathcal{P}_R(t_{lost}) = 1 - \mathrm{e}^{-\lambda t_{lost}} \tag{23}$$

In fact, the larger the λ, the faster a previously tripped line can be restored.

To develop the proposed vulnerability assessment framework, the bi-level optimization is solved using GAMS software. Additionally, the MDP within this framework is modeled in MATLAB, and its decision interval is set equal to the OPF interval. The bi-level optimization not only provides the attack vector but also yields the corresponding generation dispatch and load shedding resulting from the selected attack vector. Thus, starting from the initial state $\boldsymbol{s}_1$, which is the current operation point of the system, the next state's attack vector and its associated generation dispatch and load shedding are determined by the bi-level optimization. This process is repeated for all possible states (i.e., all possible line overloads) in step one. In addition, based on the probabilities of line tripping and restoration, as well as the initial structure of the power system, the network structure in each new state is determined. Furthermore, the loads in the next state are updated based on forecasted values. Consequently, all elements of the new state are determined, and the bi-level optimization can be re-solved to determine the generation dispatch and load shedding for the subsequent states. The developed MDP is solved in MATLAB using the dynamic Q-learning, which is explained in Section IV-E. The most critical multi-step SFDIA for each step and the potential impact of this attack are found using (14) and (15), respectively, from the optimal Q-action value function $Q^*(\boldsymbol{s}_t, \boldsymbol{a}_t^l)$.

The proposed vulnerability assessment framework is evaluated using ten distinct scenarios, as detailed in Table II.

TABLE II
Scenarios Used to Evaluate Proposed Vulnerability Assessment Framework for IEEE 39-bus Test System

Scenario	β_1	β_2	β_3	γ	λ	α	$V(\boldsymbol{s}_1)$	$\pi^*(\boldsymbol{s}_t)$
1	1.0	0.1	0.1	0.9	0.01	0.10	295.7	$\boldsymbol{a}_1^3, \boldsymbol{a}_2^3, \boldsymbol{a}_3^3, \boldsymbol{a}_4^{42}$
2	1.0	0.1	0.1	0.9	0.03	0.10	219.1	$\boldsymbol{a}_1^3, \boldsymbol{a}_2^3, \boldsymbol{a}_3^3, \boldsymbol{a}_4^{42}$
3	1.0	0.1	0.1	0.5	0.01	0.10	181.7	$\boldsymbol{a}_1^{27}, \boldsymbol{a}_2^3, \ldots$
4	1.0	0.1	0.1	0.5	0.03	0.10	183.8	$\boldsymbol{a}_1^{27}, \boldsymbol{a}_2^3, \ldots$
5	0.5	0.1	0	0.9	0.01	0.10	192.1	$\boldsymbol{a}_1^{27}, \boldsymbol{a}_2^3, \ldots$
6	0.5	0.1	0	0.9	0.03	0.10	201.3	$\boldsymbol{a}_1^{27}, \boldsymbol{a}_2^3, \ldots$
7	1.0	0	0.2	0.5	0.01	0.10	89.1	$\boldsymbol{a}_1^3, \boldsymbol{a}_2^3, \boldsymbol{a}_3^3, \boldsymbol{a}_4^{42}$
8	1.0	0	0.2	0.5	0.03	0.10	65.9	$\boldsymbol{a}_1^3, \boldsymbol{a}_2^3, \boldsymbol{a}_3^3, \boldsymbol{a}_4^{42}$
9	1.0	0.1	0.1	0.9	0.01	0.15	295.7	$\boldsymbol{a}_1^3, \boldsymbol{a}_2^3, \boldsymbol{a}_3^3, \boldsymbol{a}_4^{42}$
10	1.0	0.1	0.1	0.9	0.01	0.30	Divergence	Divergence

These scenarios study the effects of various parameters, e.g., λ, γ, α, and the weighting factors of the reward function, as shown in columns 2-7, on the multi-step SFDIAs. The value of λ varies between 0.01 (slow maintenance) and 0.03 (fast maintenance). The value of γ varies between 0.5 (attacker's preference for immediate reward) and 0.9 (attacker's preference for future reward). The weighting factors also vary between 0 (no weight) and 1 (the maximum weight). The last two columns of this table represent the value func-

tion $V(\boldsymbol{s}_1)$, which quantifies the system vulnerability to multi-step SFDIAs, and the steps of the most critical multi-step SFDIA, i.e., $\pi^*(\boldsymbol{s}_t)$, for each scenario. All scenarios are applied to the test system at 09:00 and continue until two lines are disconnected or the influence of subsequent rewards falls below 5% of their original value due to the discount factor.

In scenario 1, the maintenance speed is low, and the attacker prioritizes the accumulated impact of the attack over its immediate effects. The operator places the highest importance on minimizing the unserved load, while giving low priority to both available generation capacity and the stability margin. To evaluate this scenario, the MDP model is developed, with a concise representation illustrated in Fig. 8. The initial state of this model, denoted as $\boldsymbol{s}_1$, comprises the admittance matrix $\boldsymbol{Y}_1$, the generation power vector $\boldsymbol{PG}_1$, the load shedding vector $\boldsymbol{SD}_1$, and the demand vector $\boldsymbol{PD}_1$. Given that the demand is a state variable independent of the actions, it is not shown in Fig. 8. The operator solves the OPF at 09:00 based on this initial state. The set of available actions for the attacker at $\boldsymbol{s}_1$, denoted as A_1, comprises 46 actions corresponding to the number of lines in the system. Each $\boldsymbol{a}_1^l \in A_1$ can be utilized by the attacker to influence the OPF solution and line l in the system. Considering the ramp rates of generators, only $\boldsymbol{a}_1^3$ and $\boldsymbol{a}_1^{27}$ can trigger a line outage. Specifically, action $\boldsymbol{a}_1^{27}$ can mislead the OPF into a dispatch that causes the flow on line 27 to reach 160% of its capacity. The reward of this action based on 22 is 180. Although $\boldsymbol{a}_1^{27}$ could trip line 27 immediately, no subsequent actions would have an immediate or future impact on the system following this event. On the other hand, action $\boldsymbol{a}_1^3$ can increase the flow on line 3 to 121% of its capacity in the next interval, as shown on the right side of Fig. 8. Although the immediate impact of this action in the first interval is negligible due to the ramp rate limitation of generators, its subsequent actions, i.e., $\boldsymbol{a}_2^3$ and $\boldsymbol{a}_3^3$, can increase the flow on line 3 to 142% and 163% of its capacity, respectively. Therefore, after three intervals, the attacker can overload and disconnect line 3.

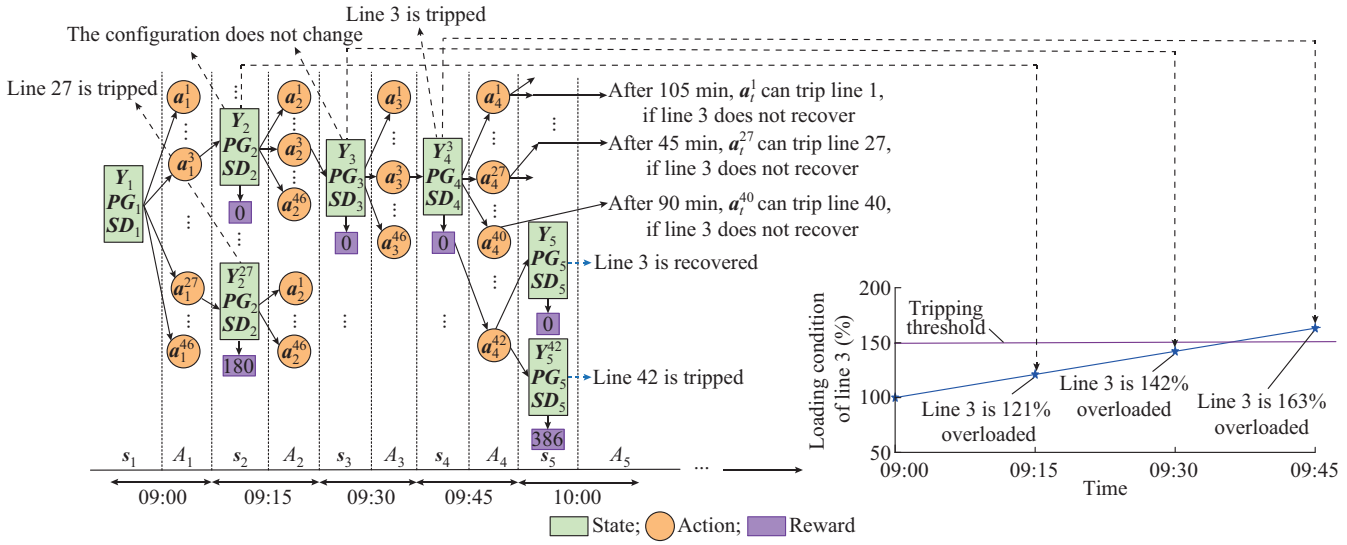

Fig. 8. Developed MDP for scenario 1.

Although the loss of line 3 does not have an immediate negative impact on the system (i.e., the reward remains zero), the new configuration of the system paves the way for subsequent actions to have more significant impacts. In the fourth interval, only actions $\boldsymbol{a}_4^1$, $\boldsymbol{a}_4^{27}$, $\boldsymbol{a}_4^{40}$, and $\boldsymbol{a}_4^{42}$ can overload and eventually trip lines 1, 27, 40, and 42 in the system. However, lines 1, 27, and 40 can be tripped after at least three intervals if and only if line 3 is not restored by then. The rewards for tripping these lines are 7.51, 180, and 28.7, respectively. However, if line 3 is recovered in the meantime, these lines cannot be tripped further. Additionally, action $\boldsymbol{a}_4^{42}$ can increase the flow of line 42 beyond 150% of its capacity in the next interval if line 3 remains disconnected, which is very likely since its restoration probability after 15 min is almost 14%. The reward for tripping line 42 while line 3 is disconnected is 386. However, if line 3 is restored, this reward becomes zero. Therefore, the best courses of actions in this scenario are $\boldsymbol{a}_1^3$, $\boldsymbol{a}_2^3$, $\boldsymbol{a}_3^3$, and $\boldsymbol{a}_4^{42}$, which result in $V(\boldsymbol{s}_1)=295.7$. This number can be used as a vulnerability index for the operator to quantify the potential impacts of multi-step SFDIAs at the current state of the system.

Scenario 2 maintains consistency with scenario 1 across all parameters, except for λ. In this scenario, λ is increased from 0.01 to 0.03 to explore the effect of this parameter on the vulnerability index for multi-step SFDIAs. As Table II shows, this modification reduces the vulnerability index (representing the potential attack impact) from 295.7 to 219.1 at 09:00. This decrease in the vulnerability index highlights the critical importance of quickly restoring systems and components after malicious activities. To further investigate the impact of this parameter, the vulnerability index of the test system under scenarios 1 and 2 has been calculated throughout the day, as shown in Fig. 9. As this figure demonstrates, for both scenarios, the vulnerability index is zero from 01:00 to 07:00, since the system loading is light and no line can be overloaded. Even if a line is overloaded and tripped, it does

not impact the system. Conversely, during peak time such as from 16: 00 to 19: 00, the system is highly vulnerable to multi-step SFDIAs, since overloading and tripping lines are easier during these hours, and the impact of losing a line is greater. This figure also reveals that the line recovery rate significantly affects the power system vulnerability, particularly during peak hours. For instance, at 17: 00, increasing the line recovery rate can reduce the system vulnerability index from 825 to 590.

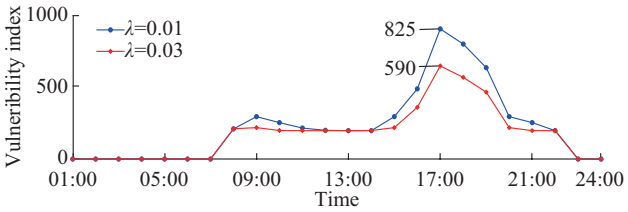

Fig. 9. Daily vulnerability system to multi-step SFDIAs for IEEE 39-bus test system.

In scenarios 3 and 4, where the discount factor is lower than that in scenarios 1 and 2, the optimal action is to start with $\boldsymbol{a}_1^{27}$, which has a more immediate impact and yields more immediate rewards compared with other actions. However, after tripping line 27, no other line can be tripped unless the attacker waits for line 27 to be restored. During this waiting period, the attacker can increase the flow on line 3, enabling its immediate tripping once line 27 is restored. Following this, tripping line 42 becomes the most critical next action. In fact, these scenarios eventually converge to scenarios 1 and 2, but with lower vulnerability indices due to the influence of the discount factor. Scenarios 5-8 are developed to study the influence of these coefficients. Comparing scenarios 1 with 5, 2 with 6, 3 with 7, and 4 with 8, as detailed in Table II, demonstrates that the coefficients assigned by the operator to the reward function can significantly alter both the vulnerability index and the sequence of the most critical multi-step SFDIA actions. For example, comparing scenarios 1 with 5 demonstrates that if the operator reduces the coefficient of unserved load, β_1, the sequence of the most critical multi-step SFDIA actions changes from $\boldsymbol{a}_1^3$, $\boldsymbol{a}_2^3$, $\boldsymbol{a}_3^3$, $\boldsymbol{a}_4^{42}$ to $\boldsymbol{a}_1^{27}$, $\boldsymbol{a}_2^3$.

Additionally, scenarios 9 and 10 are used to evaluate the impact of α and assess the sensitivity of the proposed vulnerability assessment framework to this parameter. Comparing scenario 1 with scenario 9 shows that increasing α from 0.1 to 0.15 achieves the same results with faster convergence. However, this improvement comes with a higher risk of instability. Further increasing α to 0.3, as in scenario 10, leads to divergence of the algorithm.

The main takeaways of the simulation are summarized as follows. First, the results show that multi-step SFDIAs, where the adversary uses a forward-looking strategy and carefully plans each step to amplify the impact of future actions, can cause significantly more damage than single-step attacks, which focus only on immediate effects without considering long-term consequences. Second, higher cumulative damage does not necessarily result from multiple individually severe actions. In some cases, tripping a line that causes minimal immediate damage can set the stage for more critical failures later, ultimately leading to the most severe overall impact. Third, lines that connect generators are often among the most critical components in the grid and should receive the highest level of protection against cyber-attacks. Finally, reducing the restoration time of lines can significantly decrease the system vulnerability to multi-step SFDIAs.

B. IEEE 57-bus Test System

This subsection uses the IEEE 57-bus test system to evaluate the proposed vulnerability assessment framework. The configuration data for this test system are sourced from the MATPOWER package [21]. The load profile used for the IEEE 39-bus test system is also applied to this test system. Moreover, the probability of restoration defined in (23) is used for this test system as well.

Similar to the previous section, ten different scenarios are defined for this test system, as detailed in Table III. The vulnerability index and best courses of action for these scenarios are also summarized in Table III. For instance, for scenario 3, the vulnerability index is 47.3, and the best course of action is $\boldsymbol{a}_1^3$, $\boldsymbol{a}_2^3$, $\boldsymbol{a}_3^8$, $\boldsymbol{a}_4^8$. Moreover, the vulnerability index of the test system under scenarios 1 and 2, as detailed in Table III, has been calculated throughout the day, as shown in Fig. 10. As this figure shows, the power system is highly vulnerable to multi-step SFDIAs, and quickly recovering the tripped lines can reduce the system vulnerability.

TABLE III
SCENARIOS USED TO EVALUATE PROPOSED VULNERABILITY ASSESSMENT FRAMEWORK FOR IEEE 57-BUS TEST SYSTEM

Scenario	β_1	β_2	β_3	γ	λ	α	$V(\boldsymbol{s}_1)$	$\pi^*(\boldsymbol{s}_t)$
1	1.0	0.1	0.1	0.9	0.01	0.10	80.2	$\boldsymbol{a}_1^3$, $\boldsymbol{a}_2^3$, $\boldsymbol{a}_3^8$, $\boldsymbol{a}_4^8$
2	1.0	0.1	0.1	0.9	0.03	0.10	50.8	$\boldsymbol{a}_1^3$, $\boldsymbol{a}_2^3$, $\boldsymbol{a}_3^8$, $\boldsymbol{a}_4^8$
3	1.0	0.1	0.1	0.5	0.01	0.10	47.3	$\boldsymbol{a}_1^3$, $\boldsymbol{a}_2^3$, $\boldsymbol{a}_3^8$, $\boldsymbol{a}_4^8$
4	1.0	0.1	0.1	0.5	0.03	0.10	46.8	$\boldsymbol{a}_1^3$, $\boldsymbol{a}_2^3$, $\boldsymbol{a}_3^8$, $\boldsymbol{a}_4^8$
5	0.5	0.1	0	0.9	0.01	0.10	9.3	$\boldsymbol{a}_1^{22}$, $\boldsymbol{a}_2^{22}$, $\boldsymbol{a}_3^3$, ...
6	0.5	0.1	0	0.9	0.03	0.10	9.1	$\boldsymbol{a}_1^{22}$, $\boldsymbol{a}_2^{22}$, $\boldsymbol{a}_3^3$, ...
7	1.0	0	0.2	0.5	0.01	0.10	150.4	$\boldsymbol{a}_1^3$, $\boldsymbol{a}_2^3$, $\boldsymbol{a}_3^8$, $\boldsymbol{a}_4^8$
8	1.0	0	0.2	0.5	0.03	0.10	110.3	$\boldsymbol{a}_1^3$, $\boldsymbol{a}_2^3$, $\boldsymbol{a}_3^8$, $\boldsymbol{a}_4^8$
9	1.0	0.1	0.1	0.9	0.01	0.15	80.2	$\boldsymbol{a}_1^3$, $\boldsymbol{a}_2^3$, $\boldsymbol{a}_3^8$, $\boldsymbol{a}_4^8$
10	1.0	0.1	0.1	0.9	0.01	0.30	Divergence	Divergence

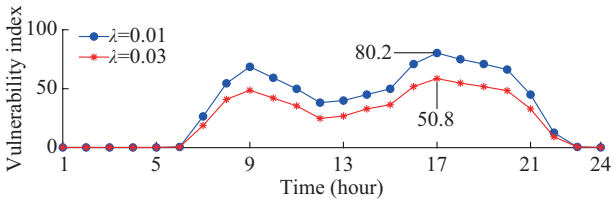

Fig. 10. Daily vulnerability system to multi-step SFDIAs for IEEE 57-bus test system.

VII. CONCLUSION

This paper first presents that prioritizing defense mechanisms solely based on the most severe single-step SFDIAs, as suggested by existing literature, may be insufficient. Instead, the defense mechanisms must also consider multi-step

SFDIAs, which can cause more severe impacts. Subsequently, the proposed vulnerability assessment framework is developed employing MDP and bi-level optimization to calculate an index quantifying the system vulnerability to multi-step SFDIAs. The vulnerability index produced by the proposed vulnerability assessment framework can serve as a quantitative metric for utilities to weigh the cost of defense investments such as securing PMUs, enhancing monitoring, or reconfiguring network topology, against the projected consequences of multi-step SFDIAs. Moreover, the proposed vulnerability assessment framework also identifies the most critical action at each step of the attack, aiding the operator in defending against attacks more effectively. Simulation results on the IEEE 39-bus test system demonstrate the effectiveness of the proposed vulnerability assessment framework in pinpointing the most critical multi-step SFDIA and the attack vector associated with each step, underscore the importance of the vulnerability index across different scenarios, and highlight the critical role of line recovery rates in mitigating the impacts of multi-step SFDIAs.

A promising future research direction involves incorporating advanced machine learning techniques to enhance the proposed vulnerability assessment framework. Specifically, the bi-level optimization can be potentially improved by replacing the lower-level optimization with a physics-informed neural network (PINN), serving as a high-fidelity surrogate for OPF. By leveraging such models, the bi-level structure can be transformed into a reinforcement learning framework, where deep Q-learning can be employed to jointly determine the optimal attack vector along with the corresponding generation dispatch and load shedding decisions. This method offers the potential to significantly reduce the computational burden, thereby enabling the efficient solution of the MDP on large-scale power systems within practical time constraints.

REFERENCES

[1] P. Verma and C. Chakraborty, "Load redistribution attacks against smart grids – models, impacts, and defense: a review," *IEEE Transactions on Industrial Informatics*, vol. 20, no. 8, pp. 10192-10208, Aug. 2024.

[2] L. Che, X. Liu, Z. Li *et al*., "False data injection attacks induced sequential outages in power systems," *IEEE Transactions on Power Systems*, vol. 34, no. 2, pp. 1513-1523, Mar. 2019.

[3] Q. Zhang and F. Li, "Cyber-vulnerability analysis for real-time power market operation," *IEEE Transactions on Smart Grid*, vol. 12, no. 4, pp. 3527-3537, Jul. 2021.

[4] Y. Yuan, Z. Li, and K. Ren, "Modeling load redistribution attacks in power systems," *IEEE Transactions on Smart Grid*, vol. 2, no. 2, pp. 382-390, Jun. 2011.

[5] R. Liu, X. Wang, B. Zeng *et al*., "Modeling load redistribution attacks in integrated electricity-gas systems," *IEEE Transactions on Smart Grid*, vol. 15, no. 4, pp. 4115-4127, Jul. 2024.

[6] L. Jia, J. Kim, R. J. Thomas *et al*., "Impact of data quality on real-time locational marginal price," *IEEE Transactions on Power Systems*, vol. 29, no. 2, pp. 627-636, Mar. 2014.

[7] J. Liang, L. Sankar, and O. Kosut, "Vulnerability analysis and consequences of false data injection attack on power system state estimation," *IEEE Transactions on Power Systems*, vol. 31, no. 5, pp. 3864-3872, Sept. 2016.

[8] M. Asghari, A. Ameli, M. Ghafouri *et al*., "On the economic vulnerability analysis of power grids to false data injection attacks against wide area measurement systems," in *Proceedings of 2022 IEEE 1st Industrial Electronics Society Annual On-line Conference*, Kharagpur, India, Jun. 2022, pp. 1-6.

[9] Z. Chu, J. Zhang, O. Kosut *et al*., "$N-1$ reliability makes it difficult for false data injection attacks to cause physical consequences," *IEEE Transactions on Power Systems*, vol. 36, no. 5, pp. 3897-3906, Sept. 2021.

[10] J. Zhang, Z. Chu, L. Sankar *et al*., "Can attackers with limited information exploit historical data to mount successful false data injection attacks on power systems?" *IEEE Transactions on Power Systems*, vol. 33, no. 5, pp. 4775-4786, Sept. 2018.

[11] D. T. Peng, J. Dong, and Q. Peng, "Overloaded branch chains induced by false data injection attack in smart grid," *IEEE Signal Processing Letters*, vol. 27, pp. 426-430, Aug. 2020.

[12] C. Liu, W. He, R. Deng *et al*., "False-data-injection-enabled network parameter modifications in power systems: attack and detection," *IEEE Transactions on Industrial Informatics*, vol. 19, no. 1, pp. 177-188, Jan. 2023.

[13] X. Liu, Z. Li, Z. Shuai *et al*., "Cyber attacks against the economic operation of power systems: a fast solution," *IEEE Transactions on Smart Grid*, vol. 8, no. 2, pp. 1023-1025, Mar. 2017.

[14] Y. Yuan, Z. Li, and K. Ren, "Quantitative analysis of load redistribution attacks in power systems," *IEEE Transactions on Parallel and Distributed Systems*, vol. 23, no. 9, pp. 1731-1738, Sept. 2012.

[15] H. M. Chung, W. T. Li, C. Yuen *et al*., "Local cyber-physical attack for masking line outage and topology attack in smart grid," *IEEE Transactions on Smart Grid*, vol. 10, no. 4, pp. 4577-4588, Jul. 2019.

[16] Y. Xiang, L. Wang, and N. Liu, "Coordinated attacks on electric power systems in a cyber-physical environment," *Electric Power Systems Research*, vol. 149, pp. 156-168, Aug. 2017.

[17] Y. Zhu, J. Yan, Y. Tang *et al*., "Resilience analysis of power grids under the sequential attack," *IEEE Transactions on Information Forensics and Security*, vol. 9, no. 12, pp. 2340-2354, Dec. 2014.

[18] J. Yan, Y. Tang, Y. Zhu *et al*., "Smart grid vulnerability under cascade-based sequential line-switching attacks," in *Proceedings of 2015 IEEE Global Communications Conference*, San Diego, USA, Sept. 2015, pp. 1-7.

[19] J. Yan, H. He, X. Zhong *et al*., "Q-learning-based vulnerability analysis of smart grid against sequential topology attacks," *IEEE Transactions on Information Forensics and Security*, vol. 12, no. 1, pp. 200-210, Jan. 2017.

[20] Y. Liu, S. Gao, J. Shi *et al*., "Sequential-mining-based vulnerable branches identification for the transmission network under continuous load redistribution attacks," *IEEE Transactions on Smart Grid*, vol. 11, no. 6, pp. 5151-5160, Nov. 2020.

[21] R. D. Zimmerman, C. E. Murillo-Sánchez, and D. Gan, "Matpower: a MATLAB power system simulation package," Power Systems Engineering Research Center, Ithaca, Tech. Memo, 1997.

[22] North American Electric Reliability Corporation (NERC). (2024, May). Standard transmission relay load ability. [Online]. Available: https://www.nerc.com/pa/Stand/Reliability%20Standards/PRC-023-1.pdf

[23] L. Che, X. Liu, and Z. Li, "Fast screening of high-risk lines under false data injection attacks," *IEEE Transactions on Smart Grid*, vol. 10, no. 4, pp. 4003-4014, Jul. 2019.

[24] Blackout Final. (2003, Dec.). Final report on the August 14, 2003 blackout in the United States and Canada: causes and recommendations. [Online]. Available: https://energy. gov/sites/prod/files/oeprod/DocumentsandMedia/BlackoutFinal-Web.pdf

[25] Y. Liu, P. Ning, and M. K. Reiter, "False data injection attacks against state estimation in electric power grids," *ACM Transactions on Information and System Security*, vol. 14, no. 1, pp. 1-33, May 2011.

[26] R. Deng and H. Liang, "False data injection attacks with limited susceptance information and new countermeasures in smart grid," *IEEE Transactions on Industrial Informatics*, vol. 15, no. 3, pp. 1619-1628, Mar. 2019.

[27] P. Akaber, B. Moussa, M. Ghafouri *et al*., "CASeS: concurrent contingency analysis-based security metric deployment for the smart grid," *IEEE Transactions on Smart Grid*, vol. 11, no. 3, pp. 2676-2687, May 2020.

[28] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*. Cambridge: MIT Press, 2018.

[29] L. Che, X. Liu, and Z. Li, "Mitigating false data attacks induced overloads using a corrective dispatch scheme," *IEEE Transactions on Smart Grid*, vol. 10, no. 3, pp. 3081-3091, May 2019.

[30] M. Du, X. Liu, Z. Li *et al*., "Robust mitigation strategy against dummy data attacks in power systems," *IEEE Transactions on Smart Grid*, vol. 14, no. 4, pp. 3102-3113, Jul. 2023.

[31] K. Wu, J. Li, B. Zhang *et al*., "Preventive dispatch strategy against FDIA induced overloads in power systems with high wind penetration," *IEEE Access*, vol. 8, pp. 210452-210461, Jan. 2020.

[32] H. Shayan and T. Amraee, "Network constrained unit commitment under cyber attacks driven overloads," *IEEE Transactions on Smart Grid*, vol. 10, no. 6, pp. 6449-6460, Nov. 2019.
[33] K. C. Sou, "Protection placement for power system state estimation measurement data integrity," *IEEE Transactions on Control of Network Systems*, vol. 7, no. 2, pp. 638-647, Jun. 2020.
[34] X. Liu, Z. Li, and Z. Li, "Optimal protection strategy against false data injection attacks in power systems," *IEEE Transactions on Smart Grid*, vol. 8, no. 4, pp. 1802-1810, Jul. 2017.
[35] M. Tian, M. Cui, Z. Dong *et al*., "Multilevel programming-based coordinated cyber physical attacks and countermeasures in smart grid," *IEEE Access*, vol. 7, pp. 9836-9847, Jan. 2019
[36] J. Fu, W. Zhang, B. Hu *et al*., "A tri-level defense model against load redistribution attacks," in *Proceedings of 2019 IEEE Sustainable Power and Energy Conference*, Beijing, China, Oct. 2019, pp. 1606-1611.
[37] M. Asghari, A. Ameli, M. Ghafouri *et al*., "Optimal data aggregation reconfiguration scheme to mitigate stealthy false data injection attacks in energy management systems," *IEEE Transactions on Smart G*rid, vol. 16, no. 4, pp. 3269-3281, Jul. 2025.
[38] S. Kancherla and I. Dobson, "Heavy-tailed transmission line restoration times observed in utility data," *IEEE Transactions on Power Systems*, vol. 33, no. 1, pp. 1145-1147, Jan. 2018.

Mohammadmahdi Asghari received the B.Sc. and M.Sc. degrees in electrical engineering from the Amirkabir University of Technology (AUT), Tehran, Iran, in 2017 and 2019, respectively. He is currently pursuing the Ph.D. degree in electrical engineering with Lakehead University, Thunder Bay, Canada. His research interests include cybersecurity of wide area monitoring, protection, and control systems, and machine learning application in modern power systems.

Amir Ameli received the B.Sc. degree in electrical engineering from Iran University of Science and Technology, Tehran, Iran, in 2011, the M.Sc. degree in electrical engineering from the Sharif University of Technology, Tehran, Iran, in 2013, and the Ph.D. degree in electrical engineering from the University of Waterloo, Waterloo, Canada, in 2019. He was a Postdoctoral Fellow with the Electrical and Computer Engineering Department, University of Waterloo, from August 2019 to July 2020. Currently, he is an Assistant Professor with the Electrical Engineering Department, Lakehead University, Thunder Bay, Canada. He is a registered Professional Engineer in Canada. His current research interests include power system cybersecurity and protection.

Mohsen Ghafouri received the B.Sc. and master's degrees in electrical engineering from the Sharif University of Technology, Tehran, Iran, in 2009 and 2011, respectively, and the Ph.D. degree in electrical engineering from Polytechnique Montréal, Montreal, Canada, in 2018. He was a Researcher with Iranian Power System Research Institute, Sharif University, Tehran, Iran, from 2011 to 2014. In 2018, he was a Researcher with CYME International, Eaton Power System Solutions, Montreal, Canada. In August 2018, he joined as the Horizon Postdoctoral Fellow of the Security Research Group, Concordia Institute for Information Systems Engineering (CIISE), Montreal, Canada, where he is currently an Associate Professor. His research interests include cybersecurity of smart grids, power system modeling, microgrid, wind energy, and control of industrial processes.

Mohammad N. Uddin received the B.Sc. and M.Sc. degrees in electrical and electronic engineering from the Bangladesh University of Engineering and Technology, Dhaka, Bangladesh, in 1993 and 1996, respectively, and the Ph.D. degree in electrical engineering from the Memorial University of Newfoundland, St. John's, Canada, in 2000. He is currently a Professor with the Department of Electrical Engineering, Lakehead University, Thunder Bay, Canada. His research interests include power electronics, electric motor drive, and application of neural networks.