IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received June 1, 2018, accepted July 3, 2018, date of publication August 8, 2018, date of current version September 21, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2856520

Power System Security Under False Data Injection Attacks With Exploitation and Exploration Based on Reinforcement Learning

ZHISHENG WANG, YING CHEN, (Member, IEEE), FENG LIU, (Member, IEEE), YUE XIA, AND XUEMIN ZHANG, (Member, IEEE)

Department of Electrical Engineering, Tsinghua University, Beijing 100084, China

Corresponding author: Ying Chen (chen_ying@tsinghua.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grants 51477081 and 51321005 and in part by the National Key Research and Development Program of China under Grant 2017YFB0903000.

ABSTRACT The false data injection (FDI) attack is a potential threat to the security of smart grids, and therefore, such threats should be assessed carefully. This paper proposes a self-governing FDI attack method with exploitation and exploration mechanisms and then evaluates its threat to power systems. The attack is executed by viruses embedded in monitoring systems in substations. First, the FDI attack is formulated as a partially observable Markov decision process. Then, an improved online reinforcement learning method called nearest sequence memory Q-learning is adopted to make the attack more effective. Finally, propagation, an inherent property of viruses, is described using a propagation-evolution model that serves as the exploration mechanism for the proposed FDI attack. To validate the proposed attack method, cosimulations of daily operations of the IEEE 39-bus system are performed in which both the automatic voltage control system and the proposed FDI attack are modeled. Test results show that the proposed FDI method can cause voltage collapse even if only a few substations are infected.

INDEX TERMS Cybersecurity, false data injection, partially observable Markov decision process, nearest sequence memory Q-learning.

I. INTRODUCTION

A. MOTIVATION

Novel communication technologies and control methods can enable better smart grid regulation; however, they also introduce serious cybersecurity threats [1]. For example, in 2010, the "Stuxnet" worm hit the staff computers of an Iranian nuclear plant and caused irreversible damage [2]. This is an example of a typical cyberattack against a power system. Cyberattacks may also cause cascading failures in a power system, thereby posing a serious threat to national infrastructure [3], [4].

A false data injection (FDI) attack is an important type of cyberattack in which a malicious attacker injects false data into a control and communication system to disrupt its normal operation [5]. Researchers have shown the existence of attack patterns capable of bypassing state estimation (SE) and presented the smallest set of manipulated measurements needed to launch a hidden attack [3]. The threat of FDI attacks cannot be neglected and should be assessed thoroughly.

This study develops a novel self-governing FDI attack model with weak assumptions to make it more realistic. In contrast to existing studies, it is assumed that viruses know nothing about the parameters and topology of a target system. A virus can only obtain local measurements from its host substation. Such attacks are launched by independent viruses enhanced with reinforcement learning. The attacker only has to embed one virus in the target system, following which the virus can propagate through the target system and independently launch decentralized attacks. Studies on FDI attacks enhanced with reinforcement learning may help in identifying vulnerabilities in a power grid and in performing a security assessment.

B. RELATED WORK

Studies on FDI attacks vary depending on the perspective of the attack targets. Sridhar and Govindarasu [6] modeled smart FDI attacks against automatic generation control systems and proved that the system frequency is affected by such attacks.

Sridhar and Manimaran [7] discussed direct and indirect FDI attacks against automatic voltage control (AVC) systems as well as the disturbance of bus voltages through sensitivity analysis. Electricity market operations are also vulnerable to FDI attacks [8], [9].

Most studies usually make some basic assumptions. For example, some studies [5], [10]–[12] assumed that attackers can obtain complete knowledge of the system topology, parameters, and measurements. In this case, constructing an attack pattern becomes an optimization problem. Other studies [6], [7], [13] assumed that FDI attacks are launched based on local topological, electrical parameters, and local measurements and then presented static attacking strategies. Sou *et al.* [14] discussed a dynamic strategy. Yu and Chin [15] proposed a blind FDI attack. Attackers who do not know the system parameters or topology uses principal component analysis to infer critical information from limited measurements. For example, Kim *et al.* [16] inferred information using a subspace method. An effective FDI attack strategy that can be performed using only limited knowledge and information will be more cost effective and less detectable and, consequently, a bigger security threat. From this perspective, this study discusses an FDI attack strategy based on reinforcement learning. In this strategy, an attacker launches attacks through viruses that attack substations independently without centralized control. The viruses do not know the topological or electrical parameters of the substation; therefore, they obtain information from their host substations.

Exploration is a key part of this decentralized attack. Viruses exploit information about the host substation and perform a wide exploration of different substations. They duplicate, propagate, accumulate information, and evolve within the cyber-physical system. Many studies have investigated virus propagation models. Draief *et al.* [17] discussed how the cybersystem topology and probability of spread affect the final propagation state. Other studies [18], [19] modeled virus propagation accurately using a Markov chain and then simplified it. Yang *et al.* [20] discussed a bivirus competing propagation model in cybersystems. In this light, this study presents a simplified propagation-evolution model to describe a multiple substation infection scenario and discusses the possible impact on the power system.

C. CONTRIBUTIONS

This study discusses decentralized FDI attacks with exploitation and exploration mechanisms. It makes the following contributions:

(1) This study proves the feasibility of purely data-driven FDI attacks on power systems. The attacker launches this attack without knowing the topological and electrical parameters of the target system. It is assumed that the attacker can only observe local measurements in a substation. Under these weak assumptions, the FDI attack becomes more realistic but may be less effective. After modeling the FDI attack from a substation as a partially observed Markov decision process (POMDP), we apply online reinforcement learning to find an optimal attack strategy. Then, we develop an effective FDI attack method that is strictly constrained by the above assumptions. Finally, this attack is executed through a malicious virus injected into the substation that aims to disrupt normal power system operations.

(2) This study introduces a virus propagation model to the FDI attack method. This model describes the stochastic process of a virus spreading in a target power system. Although worm spreading has been studied extensively in the field of cybersecurity, ours is the first study of FDI attacks on power systems in consideration of virus-spreading effects. Test results show that after several rounds of spreading, the proposed virus-based FDI attack can create voltage collapses easily.

(3) This study presents a knowledge merging method for accumulating useful experiences from different virus instances. Through remaining dominant utility functions associated with states and actions, a new Q-matrix is generated and referred for late attacks when two viruses meet at the same substation. This method enables virus evolution, thereby speeding-up information accumulation and enhancing the efficacy of proposed FDI attacks.

(4) This study also discusses the weakness of the proposed FDI attack and proposes a defense strategy against it.

II. PROBLEM DESCRIPTION

This study discusses decentralized FDI attacks with exploitation and exploration mechanisms. The target system is an AVC that regulates the active and reactive power outputs of a generator. The AVC system model includes models for power flow (PF), SE and bad data identification (BDI) system, and optimal power flow (OPF). The AVC optimizes a power system's reactive power flow and sends commands to the generators' automatic voltage regulators. Fig. 1 shows the AVC system and an FDI attack against it.

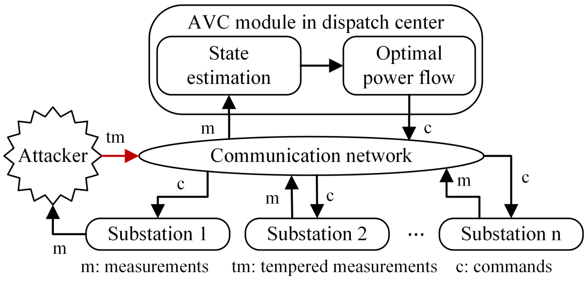


FIGURE 1. AVC system.

A power system's PF is modeled using a set of time-variant nonlinear equations in which the load shape is defined by a load curve. This study considers a distributed slack bus instead of a single slack bus.

The weighted least squares approach is commonly used for modeling SE. It has a simple BDI mechanism that uses residual examination. Detailed models of SE and BDI can be referred elsewhere [21].

In the OPF model, the objective function is to minimize net loss. The PF and bus voltage constraints are set to ensure that the system runs within the safe zone. The OPF model is expressed as follows:

$$\min_{\boldsymbol{P}_{\mathrm{Gen}}} \boldsymbol{c}_p^T \left(\boldsymbol{P}_{\mathrm{Gen}} - \boldsymbol{P}_{\mathrm{Load}}\right)$$

$$\text{s.t.} \begin{cases} \boldsymbol{P}_{\mathrm{Gen}} - \boldsymbol{P}_{\mathrm{Load}} - \mathrm{P}(\boldsymbol{\theta}, \boldsymbol{V}) = 0 \\ \boldsymbol{Q}_{\mathrm{Gen}} - \boldsymbol{Q}_{\mathrm{Load}} - \mathrm{Q}(\boldsymbol{\theta}, \boldsymbol{V}) = 0 \\ \boldsymbol{P}_{\mathrm{Gmin}} \leq \boldsymbol{P}_{\mathrm{Gen}} \leq \boldsymbol{P}_{\mathrm{Gmax}} \\ \boldsymbol{Q}_{\mathrm{Gmin}} \leq \boldsymbol{Q}_{\mathrm{Gen}} \leq \boldsymbol{Q}_{\mathrm{Gmax}} \\ \Delta\boldsymbol{Q}_{\mathrm{Gmin}} \leq \Delta\boldsymbol{Q}_{\mathrm{Gen}} \leq \Delta\boldsymbol{Q}_{\mathrm{Gmax}} \\ \boldsymbol{V}_{\min} \leq \boldsymbol{V} \leq \boldsymbol{V}_{\max} \\ \boldsymbol{P}_{Br\,\min} \leq \boldsymbol{P}_{Br} \leq \boldsymbol{P}_{Br\,\max} \\ \boldsymbol{Q}_{Br\,\min} \leq \boldsymbol{Q}_{Br} \leq \boldsymbol{Q}_{Br\,\max}. \end{cases} \tag{1}$$

where $\boldsymbol{P}_{\mathrm{Gen}}$ and $\boldsymbol{Q}_{\mathrm{Gen}}$ are respectively the active power and reactive power output of generators, $\boldsymbol{P}_{\mathrm{Gmax}}$ and $\boldsymbol{P}_{\mathrm{Gmin}}$ are respectively the upper and lower limits for the active power output of a generator, $\boldsymbol{Q}_{\mathrm{Gmax}}$ and $\boldsymbol{Q}_{\mathrm{Gmin}}$ are respectively the upper and lower limits for the reactive power output of a generator, $\boldsymbol{P}_{\mathrm{Load}}$ and $\boldsymbol{Q}_{\mathrm{Load}}$ are respectively the active and reactive loads, $\boldsymbol{V}$ and $\boldsymbol{\theta}$ are respectively the magnitude and phase angle of the bus voltage, $\mathrm{P}(\boldsymbol{\theta}, \boldsymbol{V})$ and $\mathrm{Q}(\boldsymbol{\theta}, \boldsymbol{V})$ are respectively the functions for calculating the active and reactive power flows using a complex bus voltage, $\boldsymbol{P}_{Br}$ is the vector of the branch active power flow, $\boldsymbol{P}_{Br\,\min}$ and $\boldsymbol{P}_{Br\,\max}$ are respectively the lower and upper limits for the branch active power flow, $\boldsymbol{Q}_{Br}$ is the vector of the branch reactive power flow, and $\boldsymbol{Q}_{Br\,\min}$ and $\boldsymbol{Q}_{Br\,\max}$ are respectively the lower and upper limits of the branch reactive power flow.

The OPF and SE modules rely on correct and intact data collected from each substation. A malicious attacker embeds viruses into substations and interrupts normal uplink data. This study uses weak assumptions of the information obtained by the attacker to enhance the feasibility and practicality of the proposed FDI attack method. These assumptions are listed below:

- The viruses do not know the topological and electrical parameters of the power system.
- The viruses can only observe local measurements from the host substation and only decide the optimal attack strategy based on local measurements.
- Viruses can spread through the network carrying their previously learned knowledge.

The attack uses both ‘‘exploitation’’ and ‘‘exploration’’ mechanisms. (1) ‘‘Exploration’’: A virus does not know the topological and electrical parameters of the power system. It has to learn these through trial-and-error. In each trial, the virus performs an action and observes the system feedback. Gradually, the virus collects more information about the system, thus making its attacks more threatening. ‘‘Exploitation’’ thus represents the process of obtaining information through reinforcement learning. (2) ‘‘Exploitation’’: Viruses duplicate and spread to other substations through a propagation process, as discussed in Section V. When doing so, they carry previously learned knowledge. Finally, viruses can accumulate knowledge learned from all substations to enhance their attack effects.

This study evaluates the impact of the newly proposed FDI attack on power systems. By understanding the attack mechanism, defense strategies can be developed to enhance power system security.

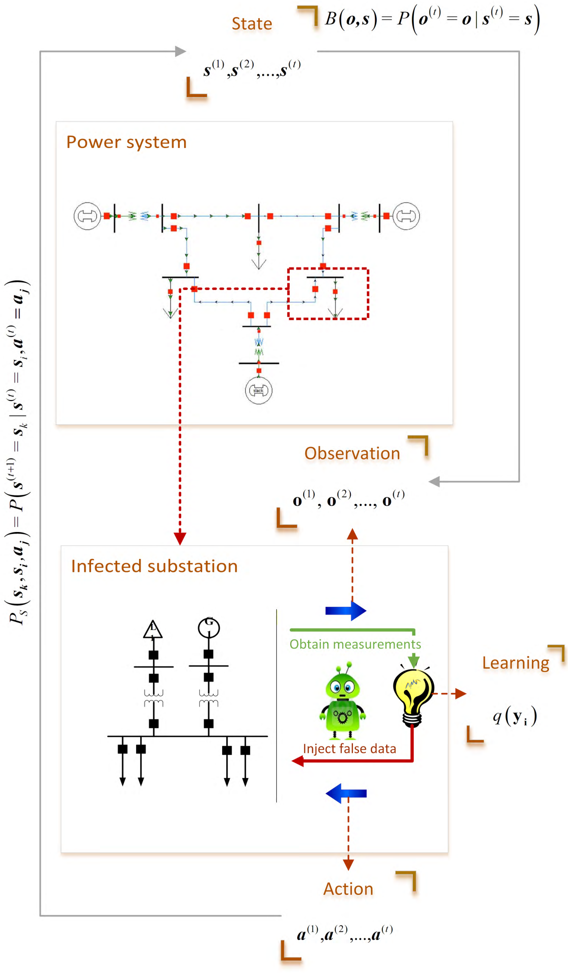
III. FDI ATTACK MODEL

As shown in Fig 2, the POMDP is used to describe an FDI attack launched by viruses. A POMDP consists of the following eight elements: state, observation, action, state-transition probability, state-observation probability, reward, and strategy:

$$\{S, O, A, \mathrm{P}, \mathrm{B}, \mathrm{R}(\boldsymbol{o}, \boldsymbol{a}), \pi(\boldsymbol{o})\}. \tag{2}$$

A. STATE AND OBSERVATION

The complex bus voltage vector is set to be the POMDP state, that is, $\boldsymbol{s} = [\boldsymbol{V}, \boldsymbol{\theta}]$. S denotes the set of all possible $\boldsymbol{s}$.



State
$B(\boldsymbol{o},\boldsymbol{s}) = P\left(\boldsymbol{o}^{(t)} = \boldsymbol{o} \mid \boldsymbol{s}^{(t)} = \boldsymbol{s}\right)$
$\boldsymbol{s}^{(1)}, \boldsymbol{s}^{(2)}, \ldots, \boldsymbol{s}^{(t)}$
Power system
$P_S\left(s_k, s_i, a_j\right) = P\left(s^{(t+1)} = s_k \mid s^{(t)} = s_i, a^{(t)} = a_j\right)$
Observation
$\mathbf{o}^{(1)}, \mathbf{o}^{(2)}, \ldots, \mathbf{o}^{(t)}$
Infected substation
Obtain measurements
Inject false data
Learning
$q(\mathrm{y}_i)$
Action
$\boldsymbol{a}^{(1)}, \boldsymbol{a}^{(2)}, \ldots, \boldsymbol{a}^{(t)}$

FIGURE 2. FDI attacks modeled using POMDP.

A virus can obtain measurements from its host substation through a POMDP observation as expressed below:

$$\boldsymbol{M}_{\text{ob}} = \left[m_1, m_2, \cdots, m_{n_M}\right]^T, \tag{3}$$

where $\boldsymbol{M}_{\text{ob}} \in \mathbb{R}^{n_M}$ are the measurements acquired by the virus from the host substation, n_M is the number of measurements, and m_i represents the ith measurement of the substation. The discrete per unit value of each measurement m_i is defined as follows:

$$s_i = k, \text{ if } m_i \in \left[m_i^0 + (k-1)\Delta m_i,\ m_i^0 + k\Delta m_i\right), \tag{4}$$

where m_i^0 is the lower limit of m_i; Δm_i, the interval between discrete values; k, an integer that satisfies $1 \leq k \leq n_i^{\text{DM}}$; and n_i^{DM}, the number of discrete steps. Hereafter, $\boldsymbol{M}_{\text{ob}}$ is simply denoted as $\boldsymbol{o}$.

B. ACTION

The virus manipulates measurements sent from substations to the control center. These measurements are expressed as

$$\boldsymbol{M}_{\text{false}} = \left[m_1^f, m_2^f, \cdots, m_{n_M}^f\right]^T$$

with

$$m_i^f = m_i \times r_i, \quad i = 1, 2, \ldots, n_M, \tag{5}$$

where r_i is the error ratio of measurement m_i; it is adjusted according to the POMDP state. We set the error ratio interval as $r_i^{\min} \leq r_i \leq r_i^{\max}$. Then, r_i may be discretized as follows:

$$r_i = r_i^{\min} + a_i \cdot \Delta r_i$$

with

$$a_i = 0, 1, 2, \cdots, \left\lfloor \frac{r_i^{\max} - r_i^{\min}}{\Delta r_i} \right\rfloor, \tag{6}$$

where $\lfloor \cdot \rfloor$ means that fractions are rounded down and a_i is the false data injection action for the ith measurement. Then, the virus action can be represented by vector $\boldsymbol{a}$:

$$\boldsymbol{a} = \left[a_1, a_2, \cdots, a_{n_M}\right]^T. \tag{7}$$

Thus, the number of all possible POMDP actions is

$$n^a = \prod_{i=1}^{n_M} \left\lfloor \frac{r_i^{\max} - r_i^{\min}}{\Delta r_i} \right\rfloor. \tag{8}$$

The set of all possible POMDP actions $\boldsymbol{a}$ is denoted as A.

C. PROBABILITIES OF STATE TRANSITIONS AND STATE OBSERVATIONS

The virus observes measurements of its host substation $\boldsymbol{M}_{\text{ob}}$. Then, it injects certain false data, denoted as action $\boldsymbol{a}$. The power system is disturbed by the injected false data, and the state transitions from $\boldsymbol{s}$ to $\boldsymbol{s}'$ are described by the state transition probability $\mathrm{P}\left(\boldsymbol{s}'|\boldsymbol{s}, \boldsymbol{a}\right)$ in the POMDP model. The state transition probability satisfies the following constraint:

$$\sum_{\boldsymbol{s}' \in S} \mathrm{P}\left(\boldsymbol{s}'|\boldsymbol{s}, \boldsymbol{a}\right) = 1, \forall \boldsymbol{s} \in S, \boldsymbol{a} \in A. \tag{9}$$

However, state $\boldsymbol{s}$ is unobservable for the virus; thus, the state transition probability cannot be acquired.

The state observation probability $\mathrm{B}(\boldsymbol{o}|\boldsymbol{s})$ is defined as the probability of observing $\boldsymbol{o}$ given system state $\boldsymbol{s}$. Similarly, the following constraint exists:

$$\sum_{\boldsymbol{o} \in O} \mathrm{B}\left(\boldsymbol{o}|\boldsymbol{s}\right) = 1, \forall \boldsymbol{s} \in S. \tag{10}$$

D. REWARD

When the bus voltage goes beyond its range, stability problems are likely to occur. Thus, the proposed attack uses bus voltage sag as the immediate reward for the attacker as follows:

$$\mathrm{R}(\boldsymbol{s}, \boldsymbol{a}) = \begin{cases} \mathrm{D}(V_0) - \mathrm{D}(V_t) & \boldsymbol{a} \neq \boldsymbol{a}_0 \\ R_0 & \boldsymbol{a} = \boldsymbol{a}_0, \end{cases} \tag{11}$$

where V_t is the magnitude of the bus voltage of the host substation after false data has been injected; V_0, the normal bus voltage; and $\mathrm{D}(V_0) - \mathrm{D}(V_t)$, the discretized bus voltage sag. The action corresponding to unmanipulated data is denoted as $\boldsymbol{a}_0 = [1, 1, \cdots, 1]^T$. Reward R_0 represents the reward for action $\boldsymbol{a}_0$. $\mathrm{D}(\cdot)$ is the discretization function for bus voltage, and it is defined as follows:

$$\mathrm{D}(V) = k \quad \mathit{if} \begin{cases} V \in \left[V_i^0 + (k-1)\Delta V_i,\ V_i^0 + k\Delta V_i\right) \\ 1 \leq k \leq N_{Vi}^{\text{DS}}, \end{cases} \tag{12}$$

where V_i^0 is the lower limit for the discretized bus voltage; ΔV, the discretization interval; and N_{Vi}^{DS}, the number of discretization levels. Reward R_0 can be expressed as

$$R_0 = \mathrm{D}(V_0) - \mathrm{D}(V_{\text{ex}}), \tag{13}$$

where $\mathrm{D}(V_{\text{ex}})$ is the bus voltage that the attacker hopes to achieve. A larger R_0 value leads to more conservative attacker behavior. For a large R_0 value, viruses tend to choose no data manipulation as the optimal action.

IV. EXPLOITATION USING REINFORCEMENT LEARNING

In the proposed attack, the virus must learn a pragmatic attack strategy. Reinforcement learning is an effective approach for obtaining a suboptimal solution of a POMDP. In this study, nearest sequence memory (NSM) is used to achieve better performance. McCallum [22] proposed NSM Q-learning and showed that it performs well for robot navigation. NSM Q-learning drives the virus to learn online and to generate and adjust its attack strategy. Ordinary Q-learning only considers the current state when updating the Q matrix, whereas NSM Q-learning additionally considers the history states. The NSM Q-learning process can be described as follows:

$$\left\{A, O, r_t, \boldsymbol{Y}, q(\boldsymbol{y}), Q(t, \boldsymbol{a})\right\}, \tag{14}$$

where A and O have the same definitions as those in (2). In addition, r_t is the reward acquired at moment t. The buffer memory $\boldsymbol{Y} = [\boldsymbol{y}_1, \boldsymbol{y}_2, \cdots, \boldsymbol{y}_t]$ records the entire history of

$\boldsymbol{y}_t = (\boldsymbol{o}_t, \boldsymbol{a}_t)$. The depth of this buffer memory must be limited from the viewpoint of complexity. If $t > n_y$, the first $t - n_y$ entries will be deleted, giving

$$\boldsymbol{Y} = \begin{cases} [\boldsymbol{y}_1, \boldsymbol{y}_2, \cdots, \boldsymbol{y}_t] & t \leq n_y \\ [\boldsymbol{y}_{t-n_y}, \boldsymbol{y}_{t-n_y+1}, \cdots, \boldsymbol{y}_t] & t > n_y. \end{cases} \tag{15}$$

NSM Q-learning considers the nearest sequence memory recorded in $\boldsymbol{Y}$. The neighbor function is introduced to describe the "distance" between two memories $\boldsymbol{y}_{i_1}$ and $\boldsymbol{y}_{i_2}$:

$$\mathcal{N}(\boldsymbol{y}_{i_1}, \boldsymbol{y}_{i_2}) = \begin{cases} \mathcal{N}(\boldsymbol{y}_{i_1-1}, \boldsymbol{y}_{i_2-1}) + 1 & \boldsymbol{y}_{i_1} = \boldsymbol{y}_{i_2} \\ 0 & \boldsymbol{y}_{i_1} \neq \boldsymbol{y}_{i_2}. \end{cases} \tag{16}$$

For $\boldsymbol{y}_{i_1}$, the set of its $k_{\mathcal{N}}$ nearest memories is denoted as $\boldsymbol{Y}_{\mathcal{N}}(\boldsymbol{y}_{i_1})$, where $k_{\mathcal{N}}$ is a crucial parameter for the NSM Q-learning algorithm.

Reward $Q(t, \boldsymbol{a})$ is the overall reward for a certain action $\boldsymbol{a}$, and it is calculated as follows:

$$Q(t, \boldsymbol{a}) = \frac{1}{k_{\mathcal{N}}} \sum_{i=1}^{t} q(\boldsymbol{y}_i) \delta_{t,\boldsymbol{a}}(\boldsymbol{y}_i)$$

with

$$\delta_{t,\boldsymbol{a}}(\boldsymbol{y}_i) = \begin{cases} 1 & \boldsymbol{y}_i \in \boldsymbol{Y}_{\mathcal{N}}(\boldsymbol{o}_t, \boldsymbol{a}) \\ 0 & \text{otherwise,} \end{cases} \tag{17}$$

where $q(\boldsymbol{y}_i)$ is the utility function for memory $q(\boldsymbol{y}_i)$. The optimal action is selected according to $Q(t, \boldsymbol{a})$ as

$$\boldsymbol{a}_t = \arg\max_{\boldsymbol{a} \in A} Q(t, \boldsymbol{a}). \tag{18}$$

After the action is executed, the system state transitions from $\boldsymbol{s}_t$ to $\boldsymbol{s}_{t+1}$, and the virus obtains new observation $\boldsymbol{o}_{t+1}$. According to the observed bus voltage, utility function $q(\boldsymbol{y}_t)$ is updated as follows:

$$q(\boldsymbol{y}_t) = \left(1 - \beta\delta(\boldsymbol{y}_t)\right) q(\boldsymbol{y}_t) + \beta\delta(\boldsymbol{y}_t) \left(r_t + \gamma \max_{\boldsymbol{a} \in A} Q(t+1, \boldsymbol{a}) \right). \tag{19}$$

Utility function $q(\boldsymbol{y})$ is equal to the Q function in ordinary Q-learning. Figure 3 shows a simplified diagram of an FDI attack simulation. The simulation consists of three simulation phases: physical system, FDI attack, and control system.

V. EXPLORATION THROUGH PROPAGATION

The proposed FDI attack is launched through viruses. The viruses exploit knowledge about the system by using reinforcement learning. By contrast, existing studies mostly focused on centralized attacks in which attackers collect information from various substations and decide whether to inject false data into particular substations. When a virus directly launches an attack, it can only obtain and manipulate local measurements. In both biology and cybersecurity, propagation and mutation are important features of viruses. Viruses perform exploitation to collect information about the host substation as well as vast exploration through different substations. They duplicate, propagate, and evolve in the cyber-physical system. We develop a simplified propagation-evolution model to describe these features.

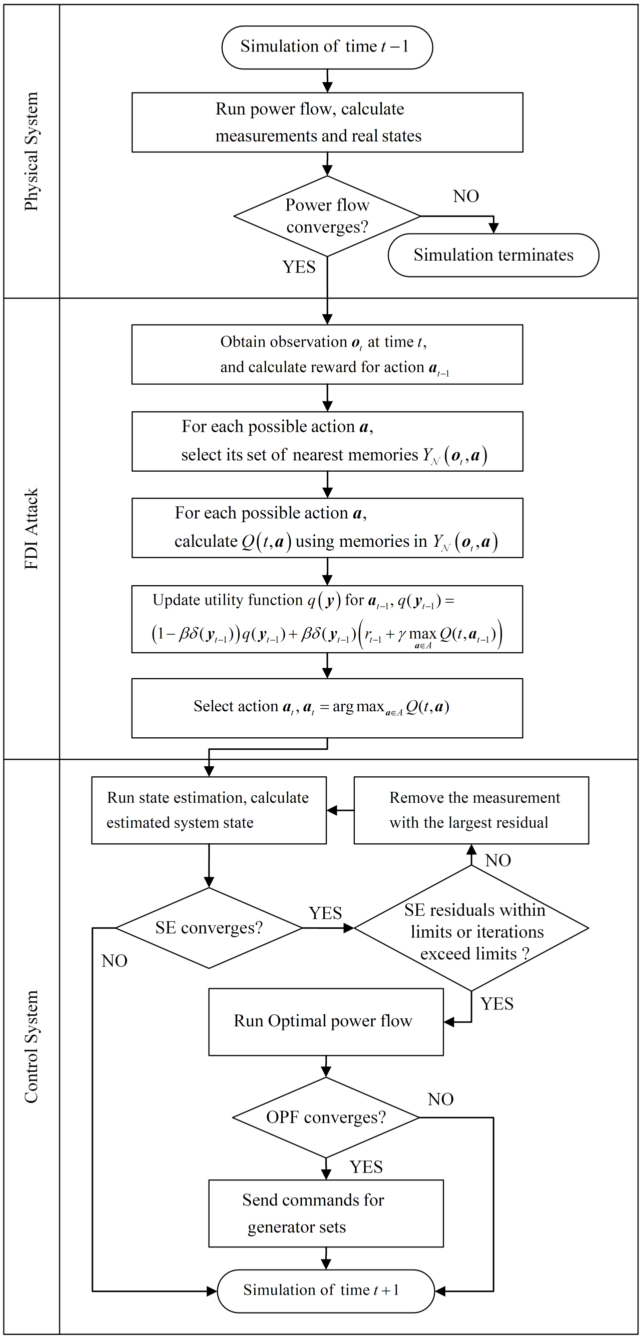

Physical System
FDI Attack
Control System
Simulation of time $t-1$
Run power flow, calculate measurements and real states
Power flow converges?
NO
YES
Simulation terminates
Obtain observation $\boldsymbol{o}_t$ at time t, and calculate reward for action $\boldsymbol{a}_{t-1}$
For each possible action $\boldsymbol{a}$, select its set of nearest memories $Y_{\mathcal{N}}(\boldsymbol{o}_t, \boldsymbol{a})$
For each possible action $\boldsymbol{a}$, calculate $Q(t, \boldsymbol{a})$ using memories in $Y_{\mathcal{N}}(\boldsymbol{o}_t, \boldsymbol{a})$
Update utility function $q(\boldsymbol{y})$ for $\boldsymbol{a}_{t-1}$, $q(\boldsymbol{y}_{t-1}) = (1-\beta\delta(\boldsymbol{y}_{t-1}))q(\boldsymbol{y}_{t-1}) + \beta\delta(\boldsymbol{y}_{t-1})\left(r_{t-1} + \gamma \max_{\boldsymbol{a}\in A} Q(t, \boldsymbol{a}_{t-1})\right)$
Select action $\boldsymbol{a}_t$, $\boldsymbol{a}_t = \arg\max_{\boldsymbol{a}\in A} Q(t, \boldsymbol{a})$
Run state estimation, calculate estimated system state
Remove the measurement with the largest residual
NO
SE converges?
YES
NO
SE residuals within limits or iterations exceed limits ?
YES
Run Optimal power flow
OPF converges?
NO
YES
Send commands for generator sets
Simulation of time $t+1$

FIGURE 3. Flow diagram of FDI attack simulation.

A. PROPAGATION

Propagation is an inherent feature of viruses. A power system is a large-scale and sophisticated system. It would be difficult to seriously damage such a system with a single virus injecting false data into a single substation. Therefore, a pragmatic virus-launched FDI attack on a large system requires an exploration mechanism for propagating through and learning from the system. In this study, the virus has no knowledge before infection and learns everything through trial-and-error. It duplicates and explores other substations.

The propagation model used in this study is a simplified version of the model proposed in [18]. Figure 4 shows a simple illustration of this model that describes the propagation probabilistically. As mentioned in Section I, in the initial state, the attacker knows nothing about the topological structure of the target power system. Thus, the initially infected substation can be considered a random selection. We denote the probability that substation i is the first infected one as $P_{\text{init}}(i)$. Thus, the following constraint exists:

$$\sum_{i=1}^{n_{\text{bus}}} P_{\text{init}}(i) = 1. \tag{20}$$

Assume that the propagation process for each virus is independent. At every time step, each virus has a probability of propagating to a certain substation. Let $P_{t,i}^{\text{infect}}$ be the probability that substation i has been infected at time t. The infection probability is denoted as $P_{t,i}(j)$. Here, $P_{t,i}(j), i \neq j$ represents the probability of a virus spreading from substation i to substation j. Furthermore, $P_{t,i}(i)$ means that substation i is infected at time t and remains infected at the next time; in other words, the virus is not eliminated at time t. Similarly, the probability that the virus is removed is defined as $1 - P_{t,i}(i)$. Then, the probability of substation j being infected at time $t + 1$ can be described as

$$P_{t+1,j}^{\text{infect}} = 1 - \prod_{i \in \mathcal{L}} \left(1 - P_{t,i}(j)\right), \tag{21}$$

where $\mathcal{L}$ in (21) denotes the set of all infected substations. In the simulation, the propagation is a random process described by $P_{\text{init}}(i)$ and $P_{t,i}(j)$.

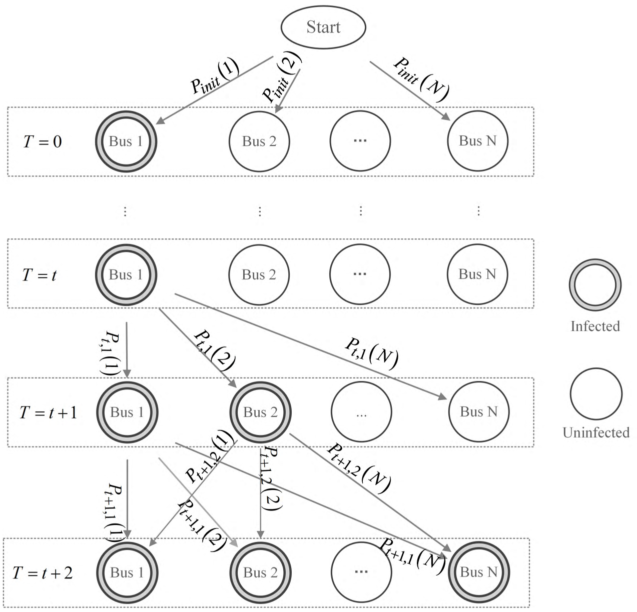


FIGURE 4. Virus propagation model.

B. EVOLUTION

The proposed attack uses a simple evolution model of the virus considering previously learned knowledge. When a virus duplicates itself and spreads to another substation, the copies inherit the previously learned knowledge; however, they are not the same viruses anymore. They continue to exploit this knowledge in different substations, and therefore, the knowledge they carry differs. When exploring different substations, some viruses will accumulate knowledge and launch more dangerous attacks. Knowledge inheritance is possible because of the same structure of knowledge in NSM Q-learning and the homogeneity of the power system. In other words, similar types of measurements are made in different substations. The values of certain types of measurement in different substations, for example, bus voltage magnitude, are similar in per-unit values. Because of these properties, data alignment is relatively easy in NSM Q-learning. Furthermore, knowledge learned in one substation is probably valid in another substation because of these properties.

If two viruses meet in a substation, for example, substation j is already infected when another virus spreads from substation i to j, the evolution model needs to be considered. This situation may not occur if the infection probability $P_{t,i}(j)$ is small. However, it is likely to occur if the infection probability is large. In this situation, the different "knowledge" carried by viruses is merged.

The merging of two viruses refers to their knowledge gained from previous attacks being merged to formulate an updated attack strategy. Then, using this updated attack strategy may make FDI attacks on a newly intruded substation more efficient and effective. The merging of two utility functions is a point-to-point merging. It depends on the property of the function, which is discretized and is expressed as a matrix. Merging the "knowledge" of two neural networks is much more difficult. In (22), $\boldsymbol{y} = (\boldsymbol{o}, \boldsymbol{a})$ represents taking action $\boldsymbol{a}$ under observation $\boldsymbol{o}$. Merging is performed separately for each $\boldsymbol{y}_i$, through which the "experience" of two viruses under similar conditions is merged. For two utility functions $q_1(\boldsymbol{y})$ and $q_2(\boldsymbol{y})$ that represent two viruses, the merged utility function $q^*(\boldsymbol{y})$ is defined as follows:

$$\forall \boldsymbol{y} \quad q^*(\boldsymbol{y}) = \begin{cases} 0 & q_1(\boldsymbol{y}) = 0 \text{ and } q_2(\boldsymbol{y}) = 0 \\ \max(q_1(\boldsymbol{y}), q_2(\boldsymbol{y})) & q_1(\boldsymbol{y}) \neq 0 \text{ or } q_2(\boldsymbol{y}) \neq 0 \\ \frac{S_{q1}}{S_q} q_1(\boldsymbol{y}) + \frac{S_{q2}}{S_q} q_2(\boldsymbol{y}) & q_1(\boldsymbol{y}) \neq 0 \text{ and } q_2(\boldsymbol{y}) \neq 0, \end{cases} \tag{22}$$

where

$$\begin{cases} S_{q_1} = \sum_{\forall \boldsymbol{y}} q_1(\boldsymbol{y}) \\ S_{q_2} = \sum_{\forall \boldsymbol{y}} q_2(\boldsymbol{y}) \\ S_q = S_{q_1} + S_{q_2}. \end{cases} \tag{23}$$

Note that $\forall \boldsymbol{y}$ means $\boldsymbol{y} = (\boldsymbol{o}, \boldsymbol{a}), \forall \boldsymbol{o} \in O, \forall \boldsymbol{a} \in A$, rather than $\forall \boldsymbol{y} \in \boldsymbol{Y}$.

The merging process is guided by the simple idea that useful knowledge should always be retained. There are three different conditions: neither has the experience (also called memory) $\boldsymbol{y}$, one of the two has the experience $\boldsymbol{y}$, and both have the experience $\boldsymbol{y}$.

(1) When $q_1(\boldsymbol{y})$ and $q_2(\boldsymbol{y})$ both equal zero, none of the viruses have experienced such a situation, and thus, the combined utility function equals zero. (2) When $q_1(\boldsymbol{y})$ or $q_2(\boldsymbol{y})$ is nonzero, one of the viruses has memory $\boldsymbol{y}$, and thus, the combined utility selects the nonzero element. (3) When $q_1(\boldsymbol{y})$ and $q_2(\boldsymbol{y})$ are all nonzero, both viruses have "knowledge" of such circumstances, and thus, the merged utility function is a weighted combination of two values. The weights are chosen as S_{q_1}/S_q and S_{q_2}/S_q; they represent how "rich" the knowledge is.

VI. CASE STUDY

A cyber-physical cosimulation is used to study the influence of FDI attacks launched by viruses. The cosimulation platform was implemented by combining different off-the-shelf power system analysis toolboxes and was developed in MATLAB. The analysis functions from the MATPOWER and PSAT toolboxes are used. Details of the design of the cosimulation platform can be referred elsewhere [23]. An IEEE 39-bus system [24] is selected as the test system with fully functioning SE and OPF modules. In the test case, each bus in the IEEE 39-bus system is considered a substation. Appendix A shows details of the configurations of the test system, SE, OPF, POMDP, and NSM Q-learning. Appendix B shows the topology of the IEEE 39-bus system.

The load curve $\lambda_D(t)$ shown in Fig. 5 describes the load over a day. During the simulation, the active and reactive loads change according to the load curve $\lambda_D(t)$ and the load ratio λ_L:

$$\begin{aligned} \boldsymbol{P}_{\text{Load}} &= \boldsymbol{P}_{\text{Load0}} \times \lambda_D(t) \times \lambda_L \\ \boldsymbol{Q}_{\text{Load}} &= \boldsymbol{Q}_{\text{Load0}} \times \lambda_D(t) \times \lambda_L, \end{aligned} \tag{24}$$

where $\boldsymbol{P}_{\text{Load0}}$ and $\boldsymbol{Q}_{\text{Load0}}$ respectively represent the basic active and reactive load (vector) of the IEEE 39-bus system.

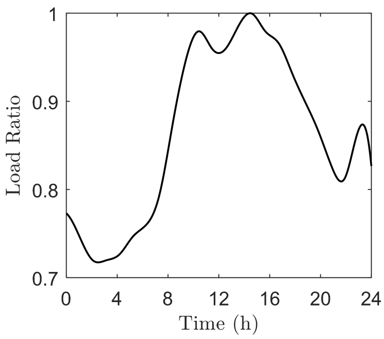


FIGURE 5. Daily load curve of test IEEE 39-bus system.

This section consists of three subsections that explain different aspects of the proposed FDI attack. Section VI-A illustrates a typical single substation attack scenario. The process of misleading SE and OPF is shown to show how the attack strategy works. Section VI-B illustrates a typical propagation attack to provide an intuitive understanding of how a propagation attack could undermine the stability of the power system. Section VI-C presents general test cases consisting of scenarios in which different buses are attacked under different load levels. In this part, we argue the following points: (1) the proposed FDI attack method can pose a severe threat to a power system regulated by an AVC and (2) if the virus can propagate and evolve, such an attack can become even more dangerous.

A. SINGLE SUBSTATION ATTACK

This section discusses an attack on a single substation. In this scenario, a virus infects a substation (bus #38) and injects manipulated data into the power system. Through NSM Q-learning, the virus gradually learns to undermine the system's stability. The load ratio is set to $\lambda_L = 2.4$.

Fig. 6 shows the minimum bus voltage $\min_{\text{bus}} V_{\text{bus}}$ and active power loss. The attack causes system collapse. Immediately before collapse ($t = 9.3$ h), the virus injects manipulated data and the SE is misled, as shown in Table 1, where $\boldsymbol{P}_L^{\text{True}}$ and $\boldsymbol{Q}_L^{\text{True}}$ represent the true active and reactive power loads, respectively, and $\boldsymbol{P}_L^{\text{SE}}$ and $\boldsymbol{Q}_L^{\text{SE}}$ are the output data of the SE. Owing to this misestimation, the OPF converges to false results and sends incorrect commands to the generators, as shown in Table 2. Finally, the commands sent to generators buses #37 and #38 lead to the collapse of the target power system.

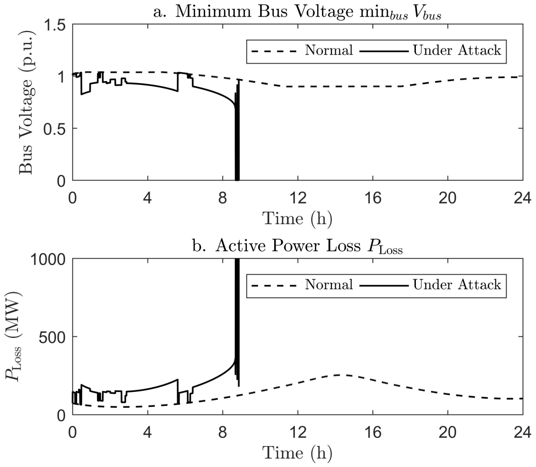


FIGURE 6. Single substation attack: (a) minimum bus voltage during attack and (b) active power loss.

TABLE 1. Real load and misled SE output at $t = 9.3$ h (per-unit value).

Load No.	Bus No.	P_L^{True}	P_L^{SE}	Q_L^{True}	Q_L^{SE}
17	29	6.201	-13.70	0.588	11.02

In contrast to many studies that assume that the attacker uses only one action to control the system, this study assumes

TABLE 2. Correct and misled OPF commands at $t = 9.3$ h (per-unit value).

Generator No.	Bus No.	P_G^{True}	P_G^{SE}	Q_G^{True}	Q_G^{SE}
8	37	9.517	6.277	3.339	2.735
9	38	13.19	0.000	3.647	-7.762

that the virus can infect a substation. Subsequently, it can continually perform malicious actions. Initially, the virus possesses no knowledge about the target system and no communication is feasible. It then learns how to bypass the SE and damage the grid. Through trial-and-error, the virus finally gains enough knowledge and successfully performs a malicious action at $t = 9.3$ h. As illustrated above, after enough information has been accumulated, the virus learns an effective action under that state to bypass the SE, mislead the OPF, and ultimately collapse the whole system.

The numerical results for the bus voltage and active power loss near the collapse point are unrealistic. These indicate that the system is no longer stable. In our simulation, the power system dynamics are modeled by continuously changing the power flow following load variations and reactive power regulations. The power flow is generally represented by a set of nonlinear equations that is solved by the Newton-Raphson method. When Newton-Raphson iterations do not converge, the maximum loading level causing voltage collapse can be identified. Near this collapse point, the numerical results of bus voltages may drop to zero, and the corresponding Jacobian matrix might be singular. The numerical values can only indicate that the system is no longer stable.

B. PROPAGATION ATTACK

In a propagation attack, the virus spreads in the test system according to the propagation-evolution model described in Section IV. The propagation parameters are set as

$$\begin{cases} P_{\text{init}}(i) = 1/39 & \forall i \\ P_{T,i}(j) = 0.05\% & \forall i \neq j,\ \forall T \\ P_{T,i}(i) = 1 & \forall T. \end{cases} \tag{25}$$

These parameters describe a spreading process with a very low infection probability and zero cure probability [18].

Initially, a virus infects a substation (bus #34) and then spreads to buses #34, #22, #38, #17, #14, and #5. Note that the single substation attack on bus #34 does not cause system collapse according to the simulation result shown in Fig. 6. However, when propagation is enabled, viruses collapse the system, as shown in Fig. 7. Tracing back to the last converged OPF period (6.4 h), Tables 3 and 4 show the misled estimations and commands. As illustrated, additional commands are misled in this case. After 6.4 h, the SE never converges again; thus, the system loses centralized control. This alone will not cause a big disturbance. However, after 6.4 h, the system experiences a sharp increase in load (see Fig. 5) that ultimately overruns the test system.

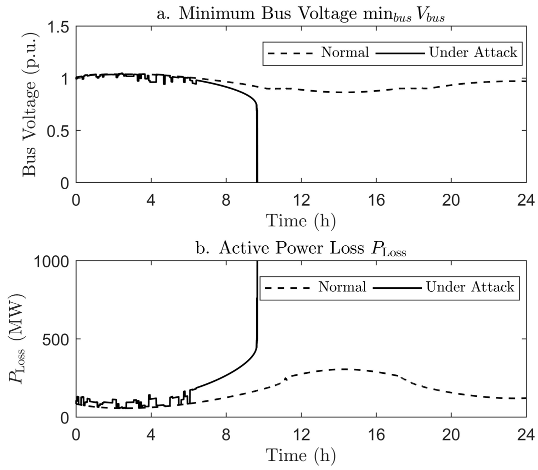


FIGURE 7. Propagation attack: (a) minimum bus voltage during attack and (b) active power loss.

TABLE 3. Real load and misled SE output at $t = 6.4$ h (per-unit value).

Load No.	Bus No.	P_L^{True}	P_L^{SE}	Q_L^{True}	Q_L^{SE}
9	20	12.381	11.152	1.875	-6.288
17	29	5.162	30.216	0.490	-10.54

TABLE 4. Correct and misled OPF commands at $t = 6.4$ h (per-unit value).

Generator No.	Bus No.	P_G^{True}	P_G^{SE}	Q_G^{True}	Q_G^{SE}
1	30	6.800	9.100	1.504	4.322
2	31	14.453	16.760	7.009	10.385
3	32	5.760	9.318	4.844	7.870
4	33	6.526	9.586	1.738	1.428
5	34	7.200	9.614	2.488	-0.498

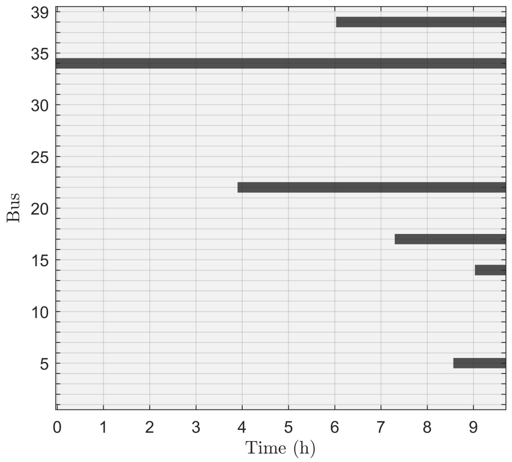


FIGURE 8. Sequence diagram of virus propagation. Black bars indicate that the bus is infected. Buses #34, #22, #38, #17, #14, and #5 are successively infected.

Fig. 8 shows the sequence of infection spread. A virus first infects bus #34. It learns to destabilize the system through trial-and-error. This procedure is the same as that described in Section VI-A until the virus successfully spreads to bus #22 4 h later. By using the proposed exploration mechanism, the virus takes the previously learned Q matrix with it to the new host substation. However, no communication is allowed among these viruses. Thus, the two viruses on buses #34 and #22 begin an uncoordinated attack using the knowledge they have learned. The spreading procedure continues. As is well-known in the field of biology, infection spread becomes faster as more buses are infected. The inheritance of knowledge helps the uncoordinated attack. These viruses cannot communicate, and therefore, forming a certain pattern that contaminates sufficient measurements to bypass the SE is very difficult. Propagation attacks mainly damage the system by forcing the SE offline. Then, the load fluctuation finally disables the system because there is no centralized control.

C. GENERAL CASE

Fig. 9 shows a general case in which 195 propagation attack scenarios (five different load ratios and 39 initial infection buses) were tested. Then, a general case of a single substation attack was simulated as a control experiment, as shown in Fig. 10. In these two figures, black bars indicate that voltage collapse has occurred in this scenario. Here, voltage collapse is defined as the bus voltage dropping below its operational limit of 0.8 p.u..

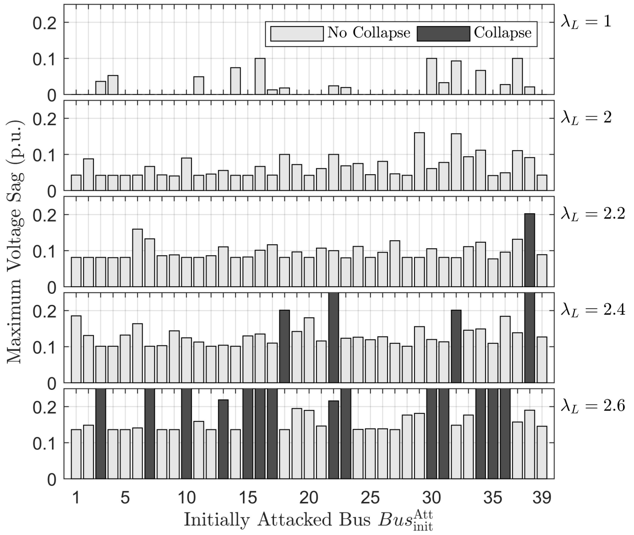


FIGURE 9. Propagation attacks in 195 scenarios. Maximum voltage sag is defined as $\max_{t,\text{bus}} V_0 - V_{\text{bus}}(t)$, $V_0 = [1, 1, \ldots, 1]$. The 19 black bars indicate that in these 19 scenarios, the bus voltage drops below the operational limit of 0.8 p.u..

Under different load ratios, FDI attacks disrupt normal operations to varying degrees. The maximum voltage sag (defined as $\max_{t,\text{bus}} V_0 - V_{\text{bus}}(t)$, $V_0 = [1, 1, \ldots, 1]$) indicates the severity of sabotage. When a system operates under heavy load, the stability margin decreases and the proposed FDI attack can cause a larger voltage sag or even voltage collapse. A comparison of Fig. 10 with Fig. 9 shows that when propagation is enabled, viruses usually cause larger voltage sag. Voltage collapse occurs in 19 scenarios; this is much higher than the number of single substation attacks. This indicates that a propagation attack poses a greater threat to the power system.

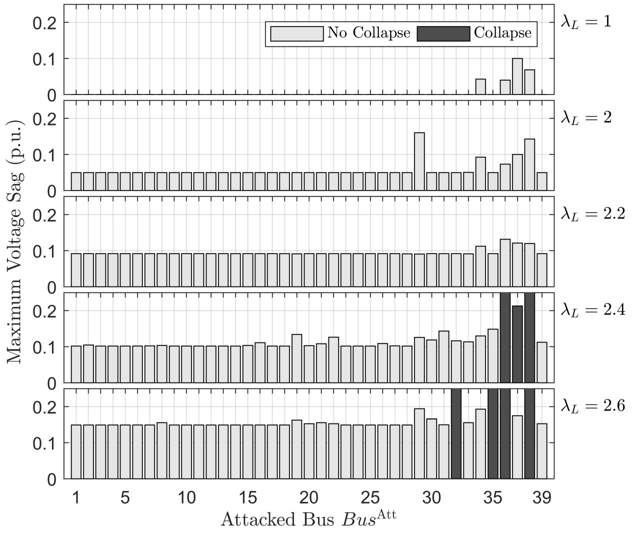


FIGURE 10. Single substation attacks in 195 scenarios. Maximum voltage sag is defined as $\max_{t,\text{bus}} V_0 - V_{\text{bus}}(t)$, $V_0 = [1, 1, \ldots, 1]$. The seven black bars indicate that in these seven scenarios, bus voltage drops below the operational limit of 0.8 p.u..

Figs. 9 and 10 show that a single attack may pose a threat to the power system and that propagation may deteriorate this situation. When several viruses are involved in the attack, more measurements are interpolated, and this increases the information entropy. Chaotic measurements cause the SE to deviate from the real state. The strategy of a virus can also evolve in different directions: (1) in a single attack, some viruses choose to cheat the SE, that is, make the SE converge to an incorrect state, and (2) in a propagation attack, after several trial-and-error iterations, the SE fails to converge in almost all cases. The FDI attacks then actually become denial-of-service attacks. This tendency is a direct result of the basic assumption: once a virus compromises a substation, it cannot communicate because of the firewall. When viruses cannot exchange information, they cannot collaborate to cheat the SE. Reinforcement learning then guides the virus in another direction, that is, toward paralyzing the SE.

VII. DISCUSSION: MITIGATION OF PROPOSED ATTACK

The attacks described in this study can be mitigated. Two mitigation methods are presented for the weaknesses of the attacks.

(1) This type of FDI attack requires a long learning period. The injected virus initially does not possess any pragmatic strategy. It needs to learn through trial-and-error. In the above test cases, the learning process repeats approximately 10^3 times. In each round, the viruses observe local measurements and inject manipulated data. If we can spot and interrupt the learning process before the virus adopts a pragmatic strategy, such an attack will probably fail.

TABLE 5. Configuration of test cases.

Symbol	Quantity	Value			
		Single Substation Attack	Propagation Attack	Single Substation Attack (General Case)	Propagation Attack (General Case)
λ_L	Load ratio	2.4	2.6	$\{1, 2, 2.2, 2.4, 2.6\}$	$\{1, 2, 2.2, 2.4, 2.6\}$
Bus^{Att}	Infected Bus No.	38	-	$\{1, 2, \cdots, 39\}$	-
$Bus^{\text{Att}}_{\text{init}}$	Initial infected Bus No.	-	34	-	$\{1, 2, \cdots, 39\}$
$P_{\text{init}}(i)$	Initial infected bus probability	-	$1/39, \forall i$	-	$1/39, \forall i$
$P_{T,i}(j)$	Independent propagation probability	-	$0.05\%, \forall i \neq j, \forall T$	-	$0.05\%, \forall i \neq j, \forall T$
$P_{T,i}(i)$	Probability of virus not eliminated	-	$1, \forall T$	-	$1, \forall T$

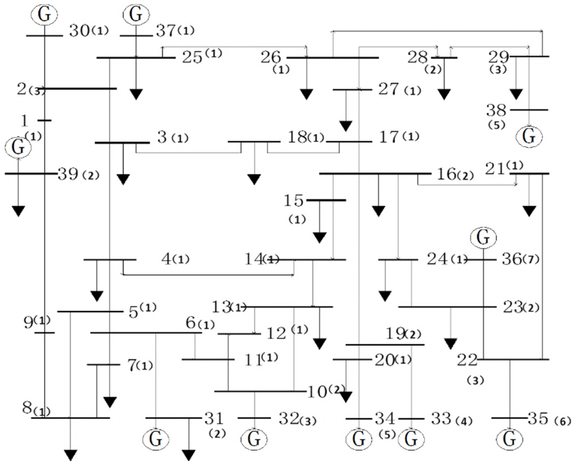


FIGURE 11. Topology of IEEE 39-bus system.

(2) Attacks on different substations are not equally damaging. The detailed patterns of these vulnerable buses still need further investigation. As a simple example, buses with both generators and loads are significantly more vulnerable. Because the set of critical buses is relatively small, we can improve their security to reduce the probability of infection. When considering the propagation process, there may also be certain combinations of buses that can cause severe damage when infected together. Thus, if we can prevent the virus from infecting these combinations of buses, the viruses will have limited effect. To achieve this, we can increase the heterogeneity of the cyber-system; specifically, we can use equipment from different manufacturers in power system planning to prevent a critical combinations of infections.

VIII. CONCLUSION

This study investigated a type of FDI attack on a power system with exploitation and exploration mechanisms. This type of attack only requires local measurements. An attack strategy based on NSM Q-learning was proposed and validated. Then, a simplified virus exploration model was introduced.

This study proposed a new type of decentralized FDI attack. The virus can learn from limited information and propagate carrying its previously learned knowledge. In other words, no human intervention is needed once the virus is embedded. These properties can significantly lower the barriers to launching an FDI attack. Test cases illustrate that this attack method can pose a severe threat to power systems regulated by an AVC. The test results reveal a new cyber-physical threat to power systems.

The proposed attacks can be mitigated. First, if we can spot and interrupt the learning process before the virus adopts a pragmatic strategy, the attack will probably fail. Second, attacks on different substations are not equally damaging. Because the set of critical buses is relatively small, we can improve their security to reduce the probability of infection.

The results of this study may help in recognizing vulnerabilities and enhancing security. We will further study the

TABLE 6. Configuration of test System.

Symbol	Value	Quantity
$V_{\min}$	0.9 p.u.	Lower limit of bus voltage
$V_{\max}$	1.1 p.u.	Upper limit of bus voltage
$\Delta \boldsymbol{Q}_{\text{Gmin}}$	$-0.1\boldsymbol{Q}_{\text{Gen}}$	Lower limit of gradient of $\boldsymbol{Q}_{\text{Gen}}$
$\Delta \boldsymbol{Q}_{\text{Gmax}}$	$0.1\boldsymbol{Q}_{\text{Gen}}$	Upper limit of gradient of $\boldsymbol{Q}_{\text{Gen}}$
$\boldsymbol{P}_{\text{Gmin}}$	0 p.u.	lower limit of $\boldsymbol{P}_{\text{Gen}}$
$\boldsymbol{P}_{\text{Gmax}}$	$\lambda_L \boldsymbol{P}_{\text{Gmax0}}$	Upper limit of $\boldsymbol{P}_{\text{Gen}}$, adjusted according to λ_L in different scenarios
$\boldsymbol{P}_L(t)$	$\lambda_L \boldsymbol{P}_{L0}(t)$	Load, adjusted according to λ_L in different scenarios
$\boldsymbol{Q}_L(t)$	$\lambda_L \boldsymbol{Q}_{L0}(t)$	Load, adjusted according to λ_L in different scenarios
$\boldsymbol{Q}_{\text{Gmin}}$	$-\infty$	Lower limit of $\boldsymbol{Q}_{\text{Gen}}$
$\boldsymbol{Q}_{\text{Gmax}}$	$+\infty$	Upper limit of $\boldsymbol{Q}_{\text{Gen}}$
Δt_{pf}	10 s	Simulation step for PF
Δt_{ctrl}	60 s	Simulation step for SE and OPF

TABLE 7. Configuration of POMDP.

Symbol	Value	Quantity
$(r_P^{\max}, r_P^{\min})$	$(0, 4)$	Error ratio interval for active power
$(r_Q^{\max}, r_Q^{\min})$	$(-2, 2)$	Error ratio interval for reactive power
Observations: $\boldsymbol{M}_{\text{ob}}$		Bus voltage: V Power flow of each branch: P_{br}, Q_{br} Load: P_{br}, Q_{br}
Manipulated: $\boldsymbol{M}_{\text{false}}$		Power flow of each branch: P_{br}, Q_{br} Load: P_L, Q_L

TABLE 8. Configuration of NSM Q-learning.

Symbol	Value	Quantity
γ	0.2	Learning rate
$k_{\mathcal{N}}$	3	Number of nearest memory
n_y	3	Depth of sequence memory
R_0	5	Fixed reward for keeping silent

propagation path of such attacks while proposing detection and defense strategies.

APPENDIX A TOPOLOGY OF IEEE 39-BUS SYSTEM

Fig. 11 shows the topology of the IEEE 39-bus system. There is no modification in the original IEEE 39-bus circuit except for the distributed slack bus mentioned in section II.

APPENDIX B CONFIGURATIONS

Table 5 shows the detailed configurations of the test cases. Tables 6, 7, and 8 show the detailed configurations of the models.

REFERENCES

[1] 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.

[2] Y. Katz and W. Jpost, "Stuxnet virus set back Iran's nuclear program by 2 years," Jerusalem Post, Jerusalem, Israel, Tech. Rep., 2010.

[3] G. Hug and J. A. Giampapa, "Vulnerability assessment of AC state estimation with respect to false data injection cyber-attacks," *IEEE Trans. Smart Grid*, vol. 3, no. 3, pp. 1362–1370, Sep. 2012.

[4] A. Hahn, A. Ashok, S. Sridhar, and M. Govindarasu, "Cyber-physical security testbeds: Architecture, application, and evaluation for smart grid," *IEEE Trans. Smart Grid*, vol. 4, no. 2, pp. 847–855, Jun. 2013.

[5] Y. Liu, P. Ning, and M. K. Reiter, "False data injection attacks against state estimation in electric power grids," in *Proc. 16th ACM Conf. Comput. Commun. Secur.*, 2009, pp. 21–32.

[6] S. Sridhar and M. Govindarasu, "Model-based attack detection and mitigation for automatic generation control," *IEEE Trans. Smart Grid*, vol. 5, no. 2, pp. 580–591, Mar. 2014.

[7] S. Sridhar and G. Manimaran, "Data integrity attack and its impacts on voltage control loop in power grid," in *Proc. IEEE Power Energy Soc. Gen. Meeting*, Jul. 2011, pp. 1–6.

[8] L. Xie, Y. Mo, and B. Sinopoli, "Integrity data attacks in power market operations," *IEEE Trans. Smart Grid*, vol. 2, no. 4, pp. 659–666, Dec. 2011.

[9] L. Xie, Y. Mo, and B. Sinopoli, "False data injection attacks in electricity markets," in *Proc. 1st IEEE Int. Conf. Smart Grid Commun.*, Oct. 2010, pp. 226–231.

[10] D. Wang, X. Guan, T. Liu, Y. Gu, Y. Sun, and Y. Liu, "A survey on bad data injection attack in smart grid," in *Proc. IEEE PES Asia–Pacific Power Power Energy Eng. Conf.*, Dec. 2013, pp. 1–6.

[11] A. Giani, E. Bitar, M. Garcia, M. McQueen, P. Khargonekar, and K. Poolla, "Smart grid data integrity attacks," *IEEE Trans. Smart Grid*, vol. 4, no. 3, pp. 1244–1253, Sep. 2013.

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

[13] L. Liu, M. Esmalifalak, Q. Ding, V. A. Emesih, and Z. Han, "Detecting false data injection attacks on power grid by sparse optimization," *IEEE Trans. Smart Grid*, vol. 5, no. 2, pp. 612–621, Mar. 2014.

[14] K. C. Sou, H. Sandberg, and K. H. Johansson, "On the exact solution to a smart grid cyber-security analysis problem," *IEEE Trans. Smart Grid*, vol. 4, no. 2, pp. 856–865, Jun. 2013.

[15] Z. H. Yu and W. L. Chin, "Blind false data injection attack using PCA approximation method in smart grid," *IEEE Trans. Smart Grid*, vol. 6, no. 3, pp. 1219–1226, May 2015.

[16] J. Kim, L. Tong, and R. J. Thomas, "Subspace methods for data attack on state estimation: A data driven approach," *IEEE Trans. Signal Process.*, vol. 63, no. 5, pp. 1102–1114, Mar. 2015.

[17] M. Draief, A. Ganesh, and L. Massoulié, "Thresholds for virus spread on networks," *Ann. Appl. Probab.*, vol. 18, no. 2, pp. 359–378, 2008.

[18] X. Wang, W. Ni, K. Zheng, and R. P. Liu, "Virus propagation modeling and convergence analysis in large-scale networks," *IEEE Trans. Inf. Forensics Security*, vol. 11, no. 10, pp. 2241–2254, Nov. 2016.

[19] P. V. Mieghem, J. Omic, and R. Kooij, "Virus spread in networks," *IEEE/ACM Trans. Netw.*, vol. 17, no. 1, pp. 1–14, Feb. 2009.

[20] L.-X. Yang, X. Yang, and Y. Y. Tang, "A bi-virus competing spreading model with generic infection rates," *IEEE Trans. Netw. Sci. Eng.*, vol. 5, no. 1, pp. 2–13, Jan./Mar. 2018.

[21] J. Kim, L. Tong, and R. J. Thomas, "Data framing attack on state estimation," *IEEE J. Sel. Areas Commun.*, vol. 32, no. 7, pp. 1460–1470, Jul. 2014.

[22] R. A. McCallum, "Hidden state and reinforcement learning with instance-based state identification," *IEEE Trans. Syst., Man, Cybern. B. Cybern.*, vol. 26, no. 3, pp. 464–473, Jun. 1996.

[23] Y. Chen and X. Sun, "Cyber security assessment of wide area controlled power system based on co-simulations," in *Proc. Int. Conf. Power Syst. Technol.*, Oct. 2014, pp. 1986–1991.

[24] A. Pai, *Energy Function Analysis for Power System Stability*. Norwell, MA, USA: Kluwer, 1989.

ZHISHENG WANG received the B.E. degree in electrical engineering from Tsinghua University, Beijing, China, in 2017.

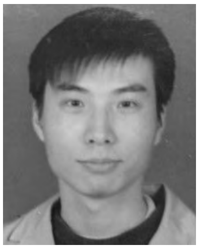
Mr. Wang is currently pursuing the master's degree with the Department of Electrical Engineering and Applied Electronic Technology, Tsinghua University. His research interests include cyber-physical system modeling and cybersecurity of a smart grid.

YING CHEN (M'07) received the B.E. and Ph.D. degrees in electrical engineering from Tsinghua University, Beijing, China, in 2001 and 2006, respectively.

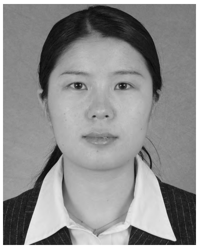
He is currently an Associate Professor with the Department of Electrical Engineering and Applied Electronic Technology, Tsinghua University. His research interests include parallel and distributed computing, electromagnetic transient simulation, cyber-physical system modeling, and cybersecurity of smart grid.

FENG LIU received the B.Sc. and Ph.D. degrees in electrical engineering from Tsinghua University, Beijing, China, in 1999 and 2004, respectively.

From 2015 to 2016, he was a Visiting Associate with the California Institute of Technology, CA, USA. He is currently an Associate Professor with Tsinghua University. He has authored or co-authored over 100 peer-reviewed technical papers and two books. He holds over 20 issued/pending patents. His research interests include power system stability analysis, optimal control and robust dispatch, game theory and learning theory, and their applications to smart grids. He was a Guest Editor of the IEEE Transactions on Energy Conversion.



YUE XIA received the B.S. and M.S. degrees in electrical engineering from China Agricultural University, Beijing, China, in 2009 and 2011, respectively, and the Ph.D. degree in electrical engineering from the Technische Universität Berlin, Germany, in 2016. He is currently pursuing the Ph.D. degree with Tsinghua University, Beijing. His research interests include modeling of power system transients, computational methods, and integration of wind power.



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