Hindawi
Complexity
Volume 2021, Article ID 6646799, 10 pages
https://doi.org/10.1155/2021/6646799

Research Article

A Brief Overview of Optimal Robust Control Strategies for a Benchmark Power System with Different Cyberphysical Attacks

Bo Hu,[1] **Hao Wang**,[2,3] **Yan Zhao**,[2,3] **Hang Zhou**,[2,3] **Mingkun Jiang**,[2,3] **and Mofan Wei**[2,3]

[1]*State Grid Liaoning Electric Power Supply Co. Ltd.,, Shenyang 110004, China*
[2]*School of Renewable Energy, Shenyang Institute of Engineering, Shenyang 110136, China*
[3]*Key Laboratory of Regional Multi-energy System Integration and Control of Liaoning Province, Shenyang 110136, China*

Correspondence should be addressed to Bo Hu; bohuln@163.com

Received 29 November 2020; Revised 20 January 2021; Accepted 6 February 2021; Published 23 March 2021

Academic Editor: Rui Wang

Security issue against different attacks is the core topic of cyberphysical systems (CPSs). In this paper, optimal control theory, reinforcement learning (RL), and neural networks (NNs) are integrated to provide a brief overview of optimal robust control strategies for a benchmark power system. First, the benchmark power system models with actuator and sensor attacks are considered. Second, we investigate the optimal control issue for the nominal system and review the state-of-the-art RL methods along with the NN implementation. Third, we propose several robust control strategies for different types of cyberphysical attacks via the optimal control design, and stability proofs are derived through Lyapunov theory. Furthermore, the stability analysis with the NN approximation error, which is rarely discussed in the previous works, is studied in this paper. Finally, two different simulation examples demonstrate the effectiveness of our proposed methods.

1. Introduction

With the development of cloud computing, artificial intelligence, and 5th-generation, the power systems regarded as the primary infrastructures in society become typical CPSs [1, 2]. Since there are numerous physical sensors, complex interaction mechanisms, and massive signals in cyberphysical power systems [3], the security of CPSs is inevitably threatened. For example, a large-scale blackout caused by cyberattacks in Ukraine had disrupted the normal lives of many people [4, 5]. Despite there are many advanced control strategies in CPSs, the imperfection of security has not been sufficiently addressed. Meanwhile, the power system composed of distributed energy and multiple loads is multidimensional [6]. It is urgent to further strengthen the security of CPSs.

Generally, the security of CPSs is threatened by attacks from the perception layer, cyberlayer, and decision layer. In particular, the attacks at the perception layer and cyberlayer, known as cyberattacks, severely disrupt the system. In recent years, reliable control strategies against various cyberattacks, such as false data injection attacks, time-delay switch attacks, and denial-of-service attacks, have been presented by many scholars. Denial-of-service attacks, which can jam information transmission channel, are an aggressive threat to CPS security [7–9]. A novel control strategy based on the game theoretic approach was proposed to resist the attacks in discrete systems [7]. Similar to the theory of literature [7], Seo et al. [9] primarily solved jamming attack in the communication between sensor and network, where an adaptive scheduling with energy constraints was presented. Using an evaluation function that quantitatively analyzes the impact of attacks, the optimal attack strategy was investigated under energy constraint in a wireless network, which can maximally destroy the stability of the system [10]. Besides, false data injection attacks, which generally cause state estimation errors, have received widespread attention because these attacks can send inaccurate control signals to the executor [11–14]. Moreover, the critical detection technology for unknown attacks and unpredictable attack areas was proposed in the previous works [9, 15–20]. For instance, considering the characteristics of the network topology and

transmission media, a risk prediction method based on a predictive model was proposed to accurately obtain the characteristics of the physical system, which can judge the fault area in the CPSs [16]. In [20], focusing on undetectable attacks, a dynamic attack detector was proposed.

With the integration of multiple energy sources, the control platform and information transmission are extremely complicated [21]. Thus, other irresistible attacks, sensor, and actuator attacks are a topic of research. For example, in [22], a reliable control with the attack compensator was investigated which can withstand sensor and actuator attacks. In [1], the resilient control strategies were proposed to ensure that the variables converge to the equilibrium point in presence of sensor and actuator attacks.

This paper concentrates on the study of an optimal robust control strategy, where the designed unified control method makes the power system immune to the actuator and sensor attacks. We use optimal control theory, reinforcement learning (RL), and neural networks (NNs) to design the controller under the assumed attacks of multiple characteristics. The main works and contributions can be summarized as follows:

(1) Optimal control theory, RL, and NNs are integrated to address the security issue of a benchmark power system.

(2) A unified way is proposed to deal with the sensor and actuator attacks via the optimal control design.

(3) The stability analysis with the NN approximation error, which is rarely discussed in the previous works, is studied in this paper.

The rest of this paper is arranged as follows.

First, the benchmark power system models with actuator and sensor attacks are formulated. Second, the optimal control issue for the nominal system is investigated, and the state-of-the-art RL methods along with the NN implementations are reviewed. Third, several robust control strategies are proposed for different types of cyberphysical attacks via the optimal control design, and stability proofs are derived through Lyapunov theory. Then, two different simulation examples demonstrate the effectiveness of our proposed methods. Finally, a brief conclusion is given.

2. Problem Statement for Power System

Let us consider the following benchmark power system:

$$\begin{aligned}\Delta\dot{\psi}_f &= -\frac{1}{T_p}\Delta\psi_f + \frac{k_p}{T_p}\Delta\psi_t,\\ \Delta\dot{\psi}_t &= -\frac{1}{T_t}\Delta\psi_t + \frac{1}{T_t}\Delta\psi_g,\\ \Delta\dot{\psi}_g &= -\frac{1}{k_s T_g}\Delta\psi_f - \frac{1}{T_g}\Delta\psi_g + \frac{1}{T_g}u,\end{aligned} \tag{1}$$

where $\Delta\psi_f$, $\Delta\psi_t$, and $\Delta\psi_g$ represent the deviations of frequency, turbine power, and governor position value, respectively; T_t, T_g, and T_p denote the time constants of turbine, governor, and power system, respectively; k_p and k_s represent the gain of power system and the speed regulation coefficient, respectively; and u is the control input.

Let $x = [\Delta\psi_f, \Delta\psi_t, \Delta\psi_g]^T$. The nominal system (1) can be rewritten as

$$\dot{x} = Ax + Bu, \tag{2}$$

where $A = \begin{bmatrix} -(1/T_p) & (k_p & /T_p) & 0 \\ 0 & -(1/T_t) & (1/T_t) & \\ -(1/k_s T_g) & 0 & -(1/T_g) & \end{bmatrix}$ and $B = \begin{bmatrix} 0 \\ 0 \\ (1/T_g) \end{bmatrix}$.

However, the attacks on the system are generally inevitable, which may affect the control performance. System dynamics (2) suffers from the actuator and sensor attacks, which can be, respectively, described by

$$\dot{x} = Ax + B(\overline{u} + t\Lambda), \tag{3}$$

$$\dot{x} = Ax + B\overline{u} + \Lambda, \tag{4}$$

where $\overline{u}$ is the robust control policy, which will be designed later. Λ denotes the system uncertainties. In this paper, we will consider different types of attacks.

Due to the existence of unknown attacks, it is difficult or even impossible to investigate the systems (3) and (4) directly. Inspired by the idea of classical works [23–26], we convert this robust control issue of the systems (3) and (4) into the optimal control problem of the nominal system (2). The main idea is that, with the system data and models, we can first attain the optimal control policy through ADP algorithms. Subsequently, based on the optimal control form, we can develop different robust control strategies for the systems with various attacks.

3. Optimal Control for the Nominal System

Define the performance index function as

$$J(x(0), u) = \int_0^\infty r(x(\tau), u(\tau))\mathrm{d}\tau, \tag{5}$$

where $r(x, u) = x^T Q x + u^T R u$ with positive definite symmetric matrices Q and R. Given the admissible control policy $u(x)$, the value function is expressed as

$$V(x(t)) = \int_t^\infty r(x(\tau), u(x(\tau)))\mathrm{d}\tau. \tag{6}$$

The optimal value function can be defined as

$$V^*(x(t)) = \min_u\left(\int_t^\infty r(x(\tau), u(x(\tau)))\mathrm{d}\tau\right). \tag{7}$$

According to the stationarity condition [27], the optimal control policy is derived by

$$u^*(x) = -\frac{1}{2}R^{-1}B^T\nabla V^*(x), \tag{8}$$

where $\nabla V^*(x) = \partial V^*(x)/\partial x$ and $V^*(x)$ should satisfy the following Hamilton–Jacobi–Bellman (HJB) equation.

$$0 = r\left(x, u^*(x)\right) + \nabla V^{*T}(x)\left(Ax + Bu^*(x)\right). \tag{9}$$

Thus, the key point to obtain the optimal control policy is to solve the HJB equation.

ADP is a powerful tool to solve the optimal control problems. Traditional ADP methods include two iterative algorithms: policy iteration (PI) and value iteration (VI). Afterwards, two noniterative RL methods are developed.

3.1. Online RL method. The aforementioned iterative ADP methods belong to the offline learning field because the value function and control policies are updated with the iteration index. Quite different from offline algorithms, online RL methods [27, 28] do not involve any iteration processes, and the value function and control policies are updated in real time.

3.2. Event Trigger-Based RL Method. In the online RL methods, the update and delivery of information must be continuous, which causes a waste of communication resources. For this phenomenon, the event trigger-based RL methods [29, 30] are developed. Here, the value function and control policies are updated only once when the system state error reaches the set point, which reduces the communication burden.

By using the aforementioned ADP methods, we can obtain the optimal control form of the nominal system, which will be employed in the following sections.

To implement the proposed algorithms, a critic NN and an actor NN are employed to approximate the iterative value function and control policy:

$$\widehat{V}^{(i)}(x) = \phi_c^T(x) W_c^{(i)}, \tag{10}$$

$$\widehat{u}^{(i)}(x) = \phi_a^T(x) W_a^{(i)}, \tag{11}$$

where $\phi_c(x)$ and $\phi_a(x)$ denote the NN activation functions and $W_c^{(i)}$ and $W_a^{(i)}$ represent the NN weights.

Hence, the optimal value function and control policy have NN representation as

$$V^*(x) = \phi_c^T(x) W_c, \tag{12}$$

$$u^*(x) = \phi_a(x) W_a, \tag{13}$$

where W_c and W_a denote the ideal NN weights.

In previous works, the NN approximation error was rarely discussed. In this paper, we attempt to consider its effect in the stability analysis.

In Figure 1, the sensor attacks, tampering the state values collected by sensors, occur in the sensor and communication network. Meanwhile, the actuator attacks, which generally modify the control instructions in actuator, occur between the decision and physical layer. The changed system state and control command can be eliminated by the robust control strategy which is calculated by RL based on the performance index function. Ultimately, the power system can work at the scheduled operating point under the sensor and actuator attacks.

4. Robust Control Strategies for Actuator Attacks

First, let us consider the system (3) with $\Lambda = \Gamma(x)$, where $\|\Gamma(x)\| \le k_d \|x\|$. The robust controller is designed by

$$\overline{u}(x) = u^*(x), \tag{14}$$

where the parameters for generating $u^*(x)$ will be determined later.

Theorem 1. *If the positive definite matrices Q and R are selected appropriately, then system (3) is asymptotically stable under the robust controller (14).*

Proof. Choose the Lyapunov function candidate as follows:

$$V = V^*(x), \tag{15}$$

which, according to (9), implies

$$\begin{aligned}
\dot{V} &= \dot{V}^*(x) \\
&= \nabla V^{*T}(x)\left(Ax + B\left(u^*(x) + \Gamma(x)\right)\right) \\
&= \nabla V^{*T}(x)\left(Ax + Bu^*(x)\right) + \nabla V^{*T}(x) B\Gamma(x) \\
&= -x^T Q x - u^{*T}(x) R u^*(x) + \nabla V^{*T}(x) B\Gamma(x).
\end{aligned} \tag{16}$$

Substituting (8) into (16) yields

$$\begin{aligned}
\dot{V} &= -x^T Q x - \frac{1}{4}\nabla V^{*T}(x) B R^{-1} B^T \nabla V^*(x) + \nabla V^{*T}(x) B\Gamma(x) \\
&\le -x^T Q x + \frac{1}{2}\Gamma^T(x)\Gamma(x) - \frac{1}{4}\nabla V^{*T}(x) B R^{-1} B^T \nabla V^*(x) \\
&\quad + \frac{1}{2}\nabla V^{*T}(x) B B^T \nabla V^*(x) \\
&\le -\left(\lambda_{\min}(Q) - \frac{1}{2}k_d^2\right)\|x\|^2 - \left(\frac{1}{4}\lambda_{\min}\left(R^{-1}\right) - \frac{1}{2}\right)\left\|B^T \nabla V^*(x)\right\|^2,
\end{aligned} \tag{17}$$

where $\lambda_{\min}(\cdot)$ denotes the minimum eigenvalue of a matrix.

To guarantee $\dot{V} \le 0$, one should choose the parameters Q and R to satisfy the following inequalities:

$$\begin{cases}
\lambda_{\min}(Q) \ge \dfrac{1}{2}k_d^2, \\
\lambda_{\min}\left(R^{-1}\right) \ge 2.
\end{cases} \tag{18}$$

The proof is completed. □

Remark 1. By using ADP methods, one can obtain the approximate optimal control policy. However, these ADP methods are finally implemented by NNs or other universal approximators, which will bring approximation errors. In the previous works, NN approximation errors were rarely

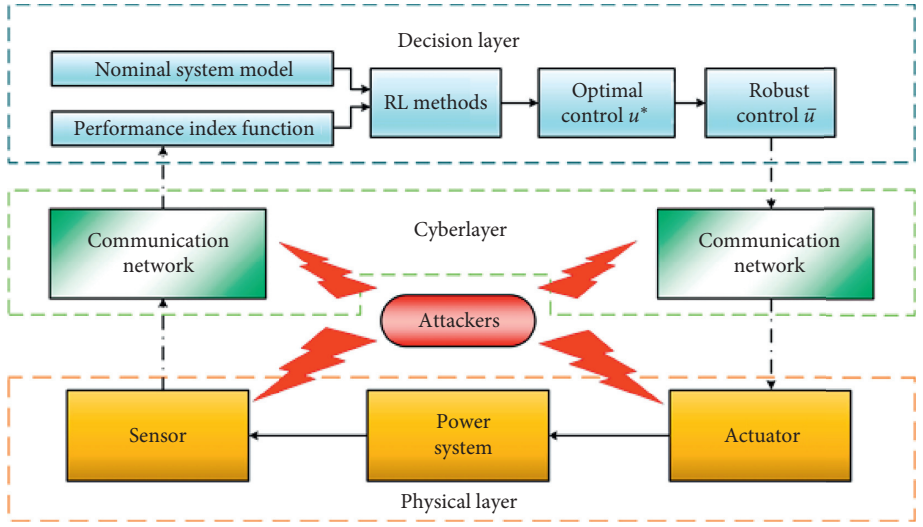

Figure 1: Block diagram of proposed robust control.

discussed. In this paper, we attempt to present the corresponding error analysis.

When NNs finish learning, NN weights will achieve convergence. Based on (11), the NN-based approximate optimal control policy, which is actually applied to the system, is expressed as

$$\hat{u}(x) = \phi_a(x)\hat{W}_a, \tag{19}$$

where $\hat{W}_a$ is the estimation of the ideal NN weight W_a. Let the approximation error be $\tilde{W}_a = \hat{W}_a - W_a$ with $\|\tilde{W}_a\| \le \tilde{W}_{am}$.

By means of (13) and (19), one gets

$$\begin{aligned}\hat{u}(x) &= \phi_a(x)W_a + \phi_a(x)\hat{W}_a - \phi_a(x)W_a \\ &= u^*(x) + \phi_a(x)\tilde{W}_a.\end{aligned} \tag{20}$$

Corollary 1. *If the positive definite matrices Q and R are selected appropriately, then the system (3) is asymptotically stable under the NN-based approximate optimal controller (19) as the NN weight approximation error $\tilde{W}_a$ goes to zero.*

Proof. Utilizing the Lyapunov function candidate (15) yields

$$\begin{aligned}\dot{V} &= \nabla V^{*T}(x)(Ax + B(\hat{u} \\ &= \nabla V^{*T}(x)\left(Ax + B\left(u^*(x) + \Gamma(x)\right)\right) + \nabla V^{*T}(x)B\phi_a(x)\tilde{W}_a.\end{aligned} \tag{21}$$

Through the result of (17), equation (21) becomes

$$\begin{aligned}\dot{V} &\le -\left(\lambda_{\min}(Q) - \frac{1}{2}k_d^2\right)\|x\|^2 - \left(\frac{1}{4}\lambda_{\min}\left(R^{-1}\right) - \frac{1}{2}\right)\left\|B^T\nabla V^*(x)\right\|^2 + \frac{1}{2}\left\|B^T\nabla V^*(x)\right\|^2 \\ &+ \frac{1}{2}\left\|\phi_a(x)\tilde{W}_a\right\|^2 \le -\left(\lambda_{\min}(Q) - \frac{1}{2}k_d^2 - \frac{1}{2}k_a^2\tilde{W}_{am}^2\right)\|x\|^2 - \left(\frac{1}{4}\lambda_{\min}\left(R^{-1}\right) - 1\right)\left\|B^T\nabla V^*(x)\right\|^2,\end{aligned} \tag{22}$$

where $\|\phi_a(x)\| \le k_a\|x\|$. If $\dot{V} \le 0$, the following condition should be satisfied:

$$\begin{cases} \lambda_{\min}(Q) \ge \frac{1}{2}k_d^2 + \frac{1}{2}k_a^2\widetilde{W}_{am}^2, \\ \lambda_{\min}\left(R^{-1}\right) \ge 4. \end{cases} \tag{23}$$

It can be observed that if the NN weight approximation error $\widetilde{W}_a$ goes to zero or is small enough, condition (23) can be easily realized with the chosen parameters. That is, the NN-based approximate optimal control policy can stabilize system (3). □

5. Robust Control Strategies for Sensor Attacks

In this section, the proposed robust control schemes are modified and extended to deal with sensor attacks [31].

5.1. Extension to Nonlinear Sensor Attacks. Consider the system (4) with nonlinear sensor attacks:

$$\dot{x} = Ax + B\overline{u} + \Gamma(x). \tag{24}$$

The robust controller for (24) is designed the same as (14), i.e., $\overline{u}(x) = u^*(x)$.

Corollary 2. *If the matrix Q is selected appropriately, then the system (24) is asymptotically stable under the robust controller $\overline{u}(x)$.*

Proof. Choose the Lyapunov function candidate as (15). Then, one attains

$$\begin{aligned} \dot{V} &= \nabla V^{*T}(x)\left(Ax + Bu^*(x) + \Gamma(x)\right) \\ &= -x^TQx - u^{*T}(x)Ru^*(x) + \nabla V^{*T}(x)\Gamma(x) \\ &\le -x^TQx - u^{*T}(x)Ru^*(x) + \frac{1}{2}\Gamma^T(x)\Gamma(x) + \frac{1}{2}\nabla V^{*T}(x)\nabla V^*(x) \\ &\le -\left(\lambda_{\min}(Q) - \frac{1}{2}k_d^2 - \frac{1}{2}k_v^2\right)\|x\|^2 - u^{*T}(x)Ru^*(x), \end{aligned} \tag{25}$$

where $\|\nabla V^*(x)\| \le k_v\|x\|$.

To guarantee $\dot{V} \le 0$, select the matrix Q to satisfy the following inequality:

$$\lambda_{\min}(Q) \ge \frac{1}{2}k_d^2 + \frac{1}{2}k_v^2. \tag{26}$$

The proof is completed.

Note that the robust load frequency control problem is a special case of sensor attacks.

Let (24) be rewritten as

$$\dot{x} = Ax + B\overline{u} + \Delta P_d(t), \tag{27}$$

where $\Delta P_d(t)$ denotes the disturbance caused by the load demand change with $\|\Delta P_d(t)\| \le \Delta P_{dm}$. □

Corollary 3. *If the matrix Q is selected appropriately, then the system states of (27) are uniformly ultimately bounded under the robust controller $\overline{u}(x)$.*

According to (15), one gets

$$\begin{aligned} \dot{V} &= \nabla V^{*T}(x)\left(Ax + Bu^*(x) + \Delta P_d(t)\right) \\ &\le -\left(\lambda_{\min}(Q) - \frac{1}{2}k_v^2\right)\|x\|^2 - u^{*T}(x)Ru^*(x) + \frac{1}{2}\Delta P_{dm}^2. \end{aligned} \tag{28}$$

Let $z = \begin{bmatrix} x \\ u^*(x) \end{bmatrix}$, $\Theta = \begin{bmatrix} (\lambda_{\min}(Q) - (1/2)k_v^2)I_x & 0 \\ 0 & R \end{bmatrix}$ with the identity matrix I_x, and $c_m = (1/2)\Delta P_{dm}^2$. If there exists a matrix Q which guarantees Θ to be positive definite, then (28) can be rewritten as

$$\begin{aligned} \dot{V} &\le -\begin{bmatrix} x \\ u^*(x) \end{bmatrix}^T \begin{bmatrix} \left(\lambda_{\min}(Q) - \frac{1}{2}k_v^2\right)I_x & 0 \\ 0 & R \end{bmatrix} \begin{bmatrix} x \\ u^*(x) \end{bmatrix} + \frac{1}{2}\Delta P_{dm}^2 \\ &\le -\lambda_{\min}(\Theta)\|z\|^2 + c_m. \end{aligned} \tag{29}$$

From (29), it can be observed that $\dot{V} \le 0$ if $\|z\| \ge \sqrt{(c_m/\lambda_{\min}(\Theta))}$. That is, the system states are uniformly ultimately bounded according to the Lyapunov extension theorem [27, 32, 33].

5.2. Extension to Constant Sensor Attacks. Consider the system (2) with constant sensor attacks:

$$\dot{x} = Ax + B\overline{u} - \Xi, \tag{30}$$

where $\dot{\Xi} = 0$.

Let $\overline{u} = u^*$ and add an attack compensator $\widehat{\Xi}$ to (30). Then, (30) becomes

$$\dot{x} = Ax + Bu^* + \widehat{\Xi} - \Xi, \tag{31}$$

where $\dot{\widehat{\Xi}} = \beta(-x^T - \nabla V^{*T}(x))^T$.

Theorem 2. *If the positive definite matrices Q and R are selected appropriately, then the system (31) is asymptotically stable under the optimal controller and the attack compensator.*

Proof. Construct a Lyapunov function candidate as follows:

$$V = \frac{1}{2}x^Tx + V^*(x) + \frac{1}{2\beta}\widetilde{\Xi}^T\widetilde{\Xi}, \tag{32}$$

where $\widetilde{\Xi} = \Xi - \widehat{\Xi}$. Then, one has

$$\begin{aligned}\dot{V} &= x^T\left(Ax + Bu^* - \widetilde{\Xi}\right) + \nabla V^{*T}(x)\left(Ax + Bu^* - \widetilde{\Xi}\right) - \frac{1}{\beta}\dot{\widetilde{\Xi}}^T\widetilde{\Xi} \\ &= x^TAx + x^TBu^* - \left(x^T + \nabla V^{*T}(x)\right)\widetilde{\Xi} + \nabla V^{*T}(x)\left(Ax + Bu^*\right) - \frac{1}{\beta}\beta\left(-x^T - \nabla V^{*T}(x)\right)\widetilde{\Xi}.\end{aligned} \tag{33}$$

After some mathematical derivation, equation (33) becomes

$$\begin{aligned}\dot{V} &= x^TAx + x^TBu^* - x^TQx - u^{*T}Ru^* \\ &\le A_m\|x\|^2 + \frac{1}{2}\|x\|^2 + \frac{1}{2}B_m^2\|u^*\|^2 - \lambda_{\min}(Q)\|x\|^2 - \lambda_{\min}(R)\|u^*\|^2 \\ &= -\left(\lambda_{\min}(Q) - A_m - \frac{1}{2}\right)\|x\|^2 - \left(\lambda_{\min}(R) - \frac{1}{2}B_m^2\right)\|u^*\|^2,\end{aligned} \tag{34}$$

where $\|A\| \le A_m$ and $\|B\| \le B_m$.

To ensure $\dot{V} \le 0$, one should set the parameters Q and R to satisfy the following inequalities:

$$\begin{cases}\lambda_{\min}(Q) \ge A_m + \dfrac{1}{2}, \\ \lambda_{\min}(R) \ge \dfrac{1}{2}B_m^2.\end{cases} \tag{35}$$

This completes the proof.

When NNs finish learning, the approximate optimal value function can be acquired:

$$\widehat{V}(x) = \phi_c^T(x)\widehat{W}_c, \tag{36}$$

where $\widehat{W}_c$ is the estimation of the ideal NN weight W_c. Let the approximation error be $\widetilde{W}_c = \widehat{W}_c - W_c$ with $\|\widetilde{W}_c\| \le \widetilde{W}_{cm}$.

Based on (31) and (36), the NN-based robust control scheme should be designed by

$$\begin{aligned}\dot{x} &= Ax + B\widehat{u} + \widehat{\Xi} - \Xi \\ &= Ax + Bu^*(x) + B\phi_a(x)\widetilde{W}_a + \widehat{\Xi} - \Xi,\end{aligned} \tag{37}$$

where

$$\begin{aligned}\dot{\widehat{\Xi}} &= \beta\left(-x^T - \nabla\widehat{V}^T(x)\right)^T \\ &= \beta\left(-x^T - \nabla V^{*T}(x) - \widetilde{W}_c^T\nabla\phi_c(x)\right)^T.\end{aligned} \tag{38}$$

□

Corollary 4. *If the positive definite matrices Q and R are selected appropriately, then the NN-based robust control scheme can stabilize the system (37) as the NN weight approximation errors $\widetilde{W}_a$ and $\widetilde{W}_c$ go to zero.*

Proof. Employing the Lyapunov function candidate (32) yields

$$\begin{aligned}\dot{V} &= x^T\left(Ax + B\left(u^* + \phi_a(x)\widetilde{W}_a\right) - \widetilde{\Xi}\right) + \nabla V^{*T}(x)\left(Ax + B\left(u^* + \phi_a(x)\widetilde{W}_a\right) - \widetilde{\Xi}\right) \\ &\quad - \frac{1}{\beta}\beta\left(-x^T - \nabla V^{*T}(x) - \widetilde{W}_c^T\nabla\phi_c(x)\right)\widetilde{\Xi} \\ &= x^TAx + x^TBu^* + \nabla V^{*T}(x)\left(Ax + Bu^*\right) + \left(x + \nabla V^*(x)\right)^TB\phi_a(x)\widetilde{W}_a + \widetilde{W}_c^T\nabla\phi_c(x)\widetilde{\Xi} \\ &\le -\left(\lambda_{\min}(Q) - A_m - \frac{1}{2}\right)\|x\|^2 - \left(\lambda_{\min}(R) - \frac{1}{2}B_m^2\right)\|u^*\|^2 \\ &\quad + \left(x + \nabla V^*(x)\right)^TB\phi_a(x)\widetilde{W}_a + \widetilde{W}_c^T\nabla\phi_c(x)\widetilde{\Xi}.\end{aligned} \tag{39}$$

From (39), in the limit as the NN weight approximation errors go to zero, i.e., $\widetilde{W}_a \longrightarrow 0$ and $\widetilde{W}_c \longrightarrow 0$, one can easily set the parameters Q and R to guarantee $\dot{V} \le 0$. If the NN approximation errors are not small enough or NNs fail to approximate optimal values, the result may not be asymptotically stable and the robust control scheme may be invalid. Therefore, the design of ADP learning algorithm is the key point. □

Table 1: Values of system parameters.

Parameters	T_g	T_t	T_p	k_s	k_p
Values	1	1	1	0.2	2

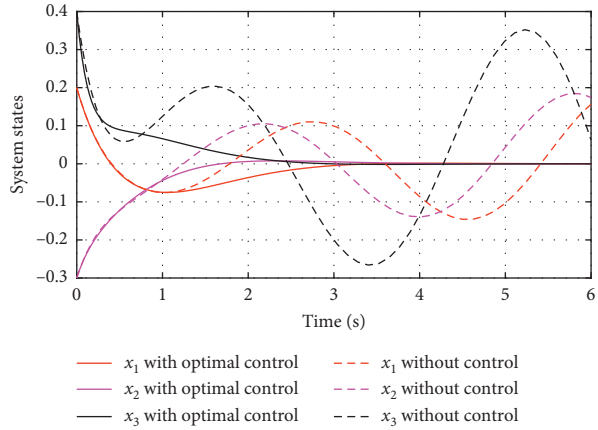

Figure 2: System states under actuator attacks.

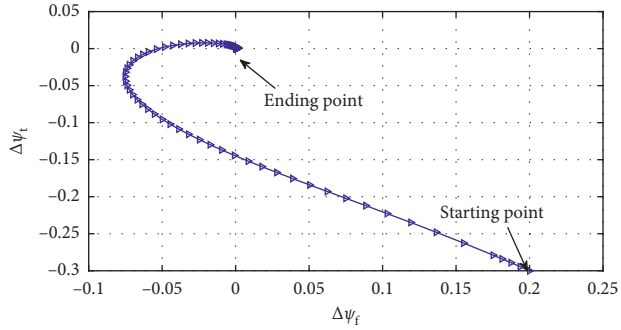

Figure 3: 2D plot of $\Delta\psi_f$ and $\Delta\psi_t$.

6. Simulation Example

In this section, to verify the proposed robust control strategy, two simulation examples of power systems are presented for two different types of attacks, respectively.

6.1. Design against Actuator Attacks. In this case, the actuator attack affecting the controller is considered in power system. The values of system parameters for this simulation are given in Table 1. Then, we can obtain system matrices $A = \begin{bmatrix} -1 & 2 & 0 \\ 0 & -1 & 1 \\ -5 & 0 & -1 \end{bmatrix}$ and $B = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$.

Let the initial system state values be $x(0) = [0.2; -0.3; 0.4]$. When we insert the actuator attacks $\Lambda = \Gamma(x)$ into the system, the states of power systems become unstable, which is shown in Figure 2.

In this case, the parameters are selected as $Q = 10I_3$ and $R = 0.2$, respectively. One can attain the optimal control $u^*(x)$ via Matlab command CARE or other RL methods. By using the robust controller (14), the system states under actuator attacks can be stabilized within 6 seconds in Figure 2. The 2D plot of state convergence trajectory is given in Figure 3, which indicates the nice performance of our control design.

Table 2: Values of system parameters.

Parameters	T_g	T_t	T_p	k_s	k_p
Values	0.5	0.5	0.5	1	1

6.2. Design against Constant Sensor Attacks. In this case, the designed controller is proved by numerical simulation results that it can effectively resist the sensor attacks. The values of system parameters for this simulation are given in Table 2.

Then, we can obtain system matrices $A = \begin{bmatrix} -2 & 2 & 0 \\ 0 & -2 & 2 \\ -2 & 0 & -2 \end{bmatrix}$ and $B = \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix}$. The controller parameters are selected as $Q = 8I_3$ and $R = 5$. Let the initial system state values be

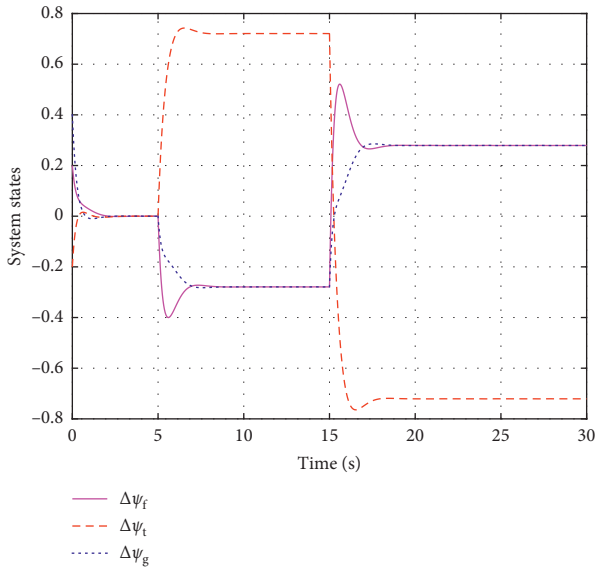

Figure 4: System states without the attack compensator.

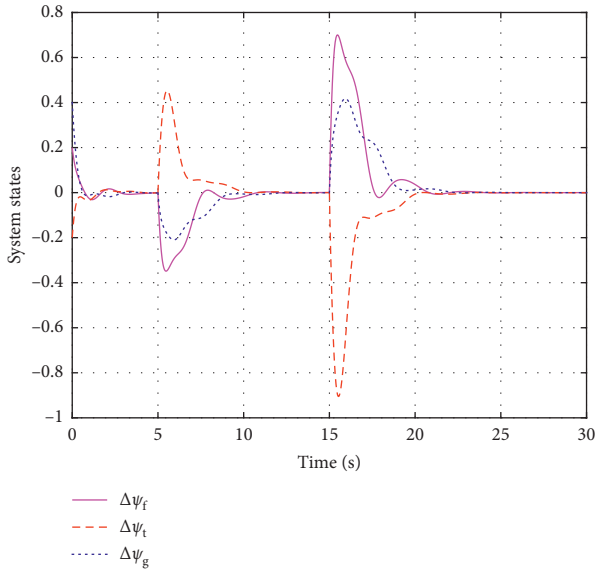

Figure 5: System states with the attack compensator.

$x(0) = [0.2; -0.2; 0.4]$ and constant sensor attack be $\Xi = \begin{cases} [0,0,0]^T & 0s \le t \le 5s \\ [2,-2,1]^T & 5s < t \le 15s \\ [-2,2,-1]^T & 15s < t \le 30s \end{cases}$.

First, we present the simulation result without the attack compensator in Figure 4, which implies the system states affected by constant attacks converge to unexpected values. Then, we employ the attack compensator-based robust control scheme, and simulation results are obtained in Figure 5, which shows the proposed scheme can quickly stabilize the system after the attacks occur. Figure 6 displays the dynamics of the attack compensator. Compared with the given attack, the compensator can estimate the attack value in a short time, which indicates the compensator can successfully get rid of the impact caused by the constant sensor attacks.

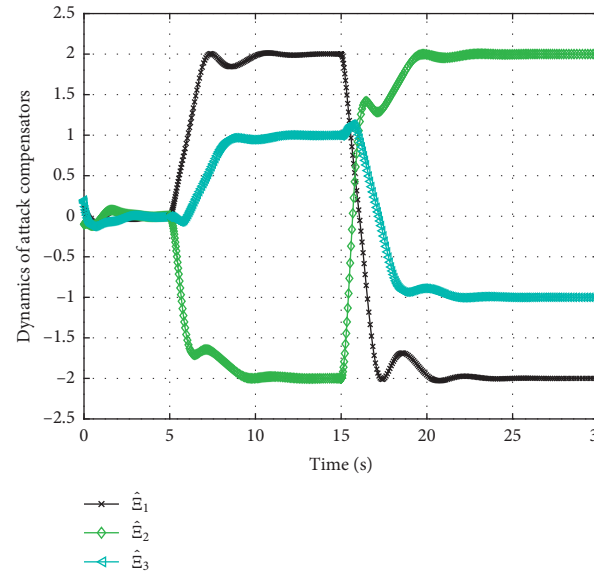

Figure 6: Dynamics of the attack compensator.

7. Conclusions

This paper has integrated optimal control theory, RL, and NNs to address the robust control issues of a benchmark power system. The optimal control theory for nominal systems and state-of-the-art RL methods along with the NN implementations have been reviewed. Multiple types of attacks in power systems, such as actuator attacks, nonlinear sensor attacks, and constant sensor attacks, are discussed. Then, several robust control schemes have been designed for different types of attacks, respectively. The control parameters have been derived through the Lyapunov stability theory. Furthermore, the stability analysis with the NN approximation error, which is rarely discussed in the previous works, has been presented in this paper. Simulation results have demonstrated the effectiveness of our proposed schemes.

Data Availability

Data are available upon request to the corresponding author.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work was supported by the Science and Technology Foundation of SGCC (SGLNDK00DWJS1900036), the Liaoning Revitalization Talents Program (XLYC1907138), the Doctoral Scientific Research Foundation of Liaoning Province (2020-BS-181), the Natural Science Foundation of Liaoning Province (2019-MS-239), the Key R&D Program of

Liaoning Province (2020JH2/10300101), the Technology Innovation Talent Fund of Shenyang (RC190360), and the Science and Technology Project of State Grid Liaoning Electric Power Company Limited (SGLNSY00HLJS2002775).

References

[1] X. Jin, W. M. Haddad, and T. Yucelen, "An adaptive control architecture for mitigating sensor and actuator attacks in cyber-physical systems," *IEEE Transactions on Automatic Control*, vol. 62, no. 11, pp. 6058–6064, 2017.

[2] G. Volkan, P. Steffen, G. Tony, and V. Frank, "A survey on concepts, applications, and challenges in cyber-physical systems," *KSII Transactions on Internet and Information Systems*, vol. 8, no. 12, pp. 4242–4268, 2014.

[3] R. Wang, Q. Sun, D. Ma, and X. Hu, "Line impedance cooperative stability region identification method for grid-tied inverters under weak grids," *IEEE Transactions on Smart Grid*, vol. 11, no. 4, pp. 2856–2866, 2020.

[4] G. Liang, S. R. Weller, J. Zhao, F. Luo, and Z. Y. Dong, "The 2015 Ukraine blackout: implications for false data injection attacks," *IEEE Transactions on Power Systems*, vol. 32, no. 4, pp. 3317-3318, 2017.

[5] G. Liang, S. R. Weller, F. Luo, J. Zhao, and Z. Y. Dong, "Distributed blockchain-based data protection framework for modern power systems against cyber attacks," *IEEE Transactions on Smart Grid*, vol. 10, no. 3, pp. 3162–3173, 2019.

[6] R. Wang, Q. Sun, D. Ma, D. Qin, Y. Gui, and P. Wang, "Line inductance stability operation domain assessment for weak grids with multiple constant power loads," *IEEE Transactions on Energy Conversion*, 2020.

[7] Y. Yuan, H. Yuan, L. Guo, H. Yang, and S. Sun, "Resilient control of networked control system under DoS attacks: a unified game approach," *IEEE Transactions on Industrial Informatics*, vol. 12, no. 5, pp. 1786–1794, 2016.

[8] Y. Yuan, F. Sun, and H. Liu, "Resilient control of cyber-physical systems against intelligent attacker: a hierarchal stackelberg game approach," *International Journal of Systems Science*, vol. 47, no. 9, pp. 2067–2077, 2015.

[9] D. Seo, H. Lee, A. Perrig, and APFS, "APFS: adaptive probabilistic filter scheduling against distributed denial-of-service attacks," *Computers & Security*, vol. 39, pp. 366–385, 2013.

[10] H. Zhang, P. Cheng P, L. Shi, and J. Chen, "Optimal DoS attack scheduling in wireless networked control system," *IEEE Transactions on Control Systems Technology*, vol. 24, no. 2, pp. 843–852, 2016.

[11] A. Chattopadhyay and U. Mitra, "Security against false data-injection attack in cyber-physical systems," *IEEE Transactions on Control of Network Systems*, vol. 7, no. 2, pp. 1015–1027, 2020.

[12] Y. Guan and X. Ge, "Distributed attack detection and secure estimation of networked cyber-physical systems against false data injection attacks and jamming attacks," *IEEE Transactions on Signal and Information Processing Over Networks*, vol. 4, no. 1, pp. 48–59, 2018.

[13] M. A. Rahman and H. Mohsenian-Rad, "False data injection attacks with incomplete information against smart power grids," in *Proceedings of the 2012 IEEE Global Communications Conference*, pp. 3153–3158, Anaheim, CA, USA, December 2012.

[14] A. Anwar, A. N. Mahmood, and Z. Tari, "Identification of vulnerable node clusters against false data injection attack in an AMI based smart grid," *Information Systems*, vol. 53, pp. 201–212, 2015.

[15] J. Park, R. Ivanov, J. Weimer, M. Pajic, and I. Lee, "Sensor attack detection in the presence of transient faults," in *Proceedings of the 2015 ACM/IEEE 6th International Conference on Cyber-Physical Systems*, pp. 1–10, Seattle, WA, USA, April 2015.

[16] P. M. Lima, L. K. Carvalho, and M. V. Moreira, "Detectable and undetectable network attack security of cyber-physical systems," *IFAC-PapersOnLine*, vol. 51, no. 7, pp. 179–185, 2018.

[17] H. Zhang, P. Cheng, L. Shi, and J. Chen, "Optimal denial-of-service attack scheduling with energy constraint," *IEEE Transactions on Automatic Control*, vol. 60, no. 11, pp. 3023–3028, 2015.

[18] L. Che, X. Liu, Z. Li, and Y. Wen, "False data injection attacks induced sequential outages in power systems," *IEEE Transactions on Power Systems*, vol. 34, no. 2, pp. 1513–1523, 2019.

[19] L. K. Carvalho, Y.-C. Wu, R. Kwong, and S. Lafortune, "Detection and mitigation of classes of attacks in supervisory control systems," *Automatica*, vol. 97, pp. 121–133, 2018.

[20] K. Pan, A. Teixeira, M. Cvetkovic, and P. Palensky, "Cyber risk analysis of combined data attacks against power system state estimation," *IEEE Transactions on Smart Grid*, vol. 10, no. 3, pp. 3044–3056, 2019.

[21] R. Wang, Q. Sun, D. Ma, and Z. Liu, "The small-signal stability analysis of the droop-controlled converter in electromagnetic timescale," *IEEE Transactions on Sustainable Energy*, vol. 10, no. 3, pp. 1459–1469, 2019.

[22] X. Huang and J. Dong, "Reliable control policy of cyber-physical systems against a class of frequency-constrained sensor and actuator attacks," *IEEE Transactions on Cybernetics*, vol. 48, no. 12, pp. 3432–3439, 2018.

[23] D. Wang, C. Li, D. Liu, and C. Mu, "Data-based robust optimal control of continuous-time affine nonlinear systems with matched uncertainties," *Information Sciences*, vol. 366, pp. 121–133, 2016.

[24] D. Wang, D. Liu, H. Li, and H. Ma, "Neural-network-based robust optimal control design for a class of uncertain nonlinear systems via adaptive dynamic programming," *Information Sciences*, vol. 282, pp. 167–179, 2014.

[25] X. Yang, D. Liu, B. Luo, and C. Li, "Data-based robust adaptive control for a class of unknown nonlinear constrained-input systems via integral reinforcement learning," *Information Sciences*, vol. 369, pp. 731–747, 2016.

[26] B. Zhao, D. Liu, and Y. Li, "Observer based adaptive dynamic programming for fault tolerant control of a class of nonlinear systems," *Information Sciences*, vol. 384, pp. 21–33, 2017.

[27] K. G. Vamvoudakis and F. L. Lewis, "Online actor-critic algorithm to solve the continuous-time infinite horizon optimal control problem," *Automatica*, vol. 46, no. 5, pp. 878–888, 2010.

[28] D. Wang, H. He, C. Mu, and D. Liu, "Intelligent critic control with disturbance attenuation for affine dynamics including an application to a microgrid system," *IEEE Transactions on Industrial Electronics*, vol. 64, no. 6, pp. 4935–4944, 2017.

[29] D. Wang, H. He, X. Zhong, and D. Liu, "Event-driven nonlinear discounted optimal regulation involving a power system application," *IEEE Transactions on Industrial Electronics*, vol. 64, no. 10, pp. 8177–8186, 2017.

[30] X. Zhong and H. He, "An event-triggered ADP control approach for continuous-time system with unknown internal states," *IEEE Transactions on Cybernetics*, vol. 47, no. 3, pp. 683–694, 2017.

[31] X. Yang and H. He, "Self-learning robust optimal control for continuous-time nonlinear systems with mismatched disturbances," *Neural Networks*, vol. 99, pp. 19–30, 2018.

[32] Q. Wei, R. Song, and P. Yan, "Data-driven zero-sum neuro-optimal control for a class of continuous-time unknown nonlinear systems with disturbance using ADP," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 27, no. 2, pp. 444–458, 2016.

[33] D. Zhao, Q. Zhang, D. Wang, and Y. Zhu, "Experience replay for optimal control of nonzero-sum game systems with unknown dynamics," *IEEE Transactions on Cybernetics*, vol. 46, no. 3, pp. 854–865, 2016.