IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received January 13, 2018, accepted February 8, 2018, date of publication February 16, 2018, date of current version March 16, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2806941

Security Analysis of Energy Internet With Robust Control Approaches and Defense Design

HUI GE[1,2], (Member, IEEE), AND ZHENJIANG ZHAO[2]
[1]School of Automation, Nanjing University of Posts and Telecommunications, Nanjing 210046, China
[2]Institute of Shipping and Mechanics Engineering, Taizhou University, Taizhou 225300, China

Corresponding author: Hui Ge (gehuivip0217@163.com)

ABSTRACT Power cyber-physical systems (PCPSs) are becoming increasingly complex and diverse, benefitting from all aspects of societal issues. A PCPS is a next-generation system that is a tight conjunction of computation, communication, control, and power systems. The integration of Internet of Things and PCPS generates a novel energy system, the Internet of Energy, which is the future of the energy system. In the same way, the connection of bulk components and the extensive communications among them has brought many insecurities, uncertainties, and security challenges. Thus, a study of the analysis and synthesis of the reliability and security of PCPS is presented in this paper. By considering the unsafe uncertainties, models are formulated for a general cyber-attack, and a double-loop architecture for security defense is designed. According to this framework, security control scenarios are obtained from the character of each kind of cyber-attack. Finally, a separately excited dc motor with uncertainty is used as an example to demonstrate the problem.

INDEX TERMS Energy Internet, Internet of Things, power cyber-physical system, cyber security, system uncertainty.

I. INTRODUCTION

Cyber-physical systems (CPSs) are an innovated system with integrated computation, communication and control capabilities for the key physical components or infrastructures of heterogeneous engineering systems. A CPS is also ubiquitous in industrial control systems, such as power systems, transportation networks, advanced manufacturing, as well as airplanes and space vehicles [1], [2].

Since the first international workshop held by National Science Foundation (NSF) focusing on CPS in 2006, a new round of scientific and technological competitions taking CPS as the research content is in full swing [3]. CPS, a system with a tight conjunction of physical resources and computational information, is listed as the top priority for information technology in the report of the President's Council of Advisors on Science and Technology (PCAST) in 2007. The U.S. National Science Foundation has provided nearly seven billion dollars for CPS research and education [1]. Furthermore, CPS has been a priority issue for the U. S. to achieve global dominance in the new round of industrial competition [3]. In 2011, the Internet of Energy was first proposed as a political decision to change the energy system in Germany. In 2013, CPS was proposed as a critical technique in Germany's Industry 4.0. At the end of 2014, the plan of ''Made in China 2025'' was established. In Japan, South Korea and the European Union, the government, industry and academia have aroused great research enthusiasm for CPS.

Based on the Internet of Things (IoT), the Internet of Energy is taken as a typical CPS system that is also called a power cyber-physical system (PCPS) [4]. The overall system spans different scales from gigawatt power plants to kilowatt distribution grids. The term ''Internet of Energy'' is proposed for the future of the energy landscape in the context of mature information and communication technologies, which play an critical role in promoting a new energy system [10]. The rapid development of techniques makes the traditional closed system progress to a larger scale, with greater distributed space and more complexity. As the critical infrastructure, the ability to generate and distribute electricity in bulk is highly regulated. It is necessary to note that energy generation and distribution are equally important.

By using advanced sensors, control and software applications, all of the energy production, transmission, consumption and hundreds of millions of pieces of equipment, machines and systems are connected to form the foundation of IoS in the Internet of Energy. Together with intelligent power generation, electricity consumption and power storage equipment will access the network, with the help of the flow of information and the formation of self-talk, the future framework of the Internet of Energy is constructed. Wang *et al.* [6] show that even ambient sensors for assisting living will generate and collect massive amounts of data, which will disturb the effective analysis. A white paper on Internet of Energy [5] reveals that there are 50 billion devices connected in the Internet of Energy and that even a fan within the system will generate up to 15 GB of data. Thus, the overall system will generate incredible amounts of data [9], [11], [15], [19], which is generally the provenance of insecurity for the Internet of Energy.

Moreover, the proper manufacture and distribution of electricity can directly impact our safety [8], [12], such as providing heat in winter or powering irrigation pumps during a drought. The smart grid is a unique facet of industrial networks that brings many new security questions and concerns into the energy industry [21]. It is an update to traditional electrical transmission and distribution systems and can accommodate digital communications for the metered and intelligent delivery of electricity.

Over the past several decades, system and control domain researchers have developed many powerful engineering methods, such as time and frequency domain methodologies, state space analysis, robust control, predictive control and game theory; see [13], [17], [20], and others based on the references therein. Fruitful results have also been achieved in computer science, such as real-time computing techniques, embedded systems architectures and system software, and innovative approaches to ensure system reliability and cyber security [14]. However, the 14 greatest challenges relating environmental, health, and societal issues listed by the U.S. National Academy of Engineering have revealed that CPS research is still in its infancy. In particular, how to satisfy the high reliability and security requirements for heterogeneous physical components is the largest challenge of this research [7], [16].

In recent years, a great deal of incidents referring to CPS security have occurred, such as the Stuxnet virus, which destroyed the Iranian nuclear program; the breakout in Ukraine, which caused thousands of families to lose power; and the weapons-grade RansomWare called ''Wannacry'' and ''Petya'', which disabled nearly 20000 computers in governments, universities and hospitals distributed in the USA, China and Russia.

The integration of the energy landscape of tomorrow must be built on information, communication and technology (ICT), and a control infrastructure. Based on a set of open standards and protocols, all of the components of PCPS join together to form the Internet of Energy. From the theoretical control perspective, stabilization performance is considered first; stable and unstable systems are often difficult to distinguish, but this performance is insufficient to ensure the reliability and security of the system. However, system security is another strict requirement, and heterogeneous information technology (IT) techniques should be adopted. Protecting data and information via a network is the traditional mission of IT security control. However, in the control theory field, how to guarantee the stability of the system from underlying uncertain perturbations or even cyber-attacks is a more important issue. Unlike the traditional information network security, attackers aim to influence the physical device in PCPS by changing the computational information. In [21] and [22], the dramatic differences in security between CPS and general-purpose computing systems have been summarized.

Within PCPS, the critical processes and key instruments are closely related to a hybrid combination of all new technologies [24]. Many works have focused on isolated domains, such as information disclosure [42], denial-of-service (DoS) [29], wormhole attacks [43], stealthy attack models [30], synthesis attacks [31] and control theory, optimization and game theory approaches [27]. Recently, the set-based and event-trigger approaches have been increasingly adopted to analyze cyber-attacks (see [28] and the reference therein). In detail, an explicit characterization of the frequency and duration of DoS attacks is analyzed under the condition that the closed-loop stability is preserved in [37].

The increasing complexity of the PCPS brings in heterogeneous uncertainties, and these uncertainties may mitigate the reliability and security of PCPS. Therefore, uncertainties generally impact system performance. In this paper, the PCPS security issue is investigated from the uncertainty perspective, and many faults and cyber-attack are taken as different uncertainties. A whole model is formulated to describe the PCPS. However, in a practice system, it is difficult to eliminate the error between the models and the practical systems, which are the so-called model errors. In addition, attack processes are often stealthy and seem normal, and the abnormal information is usually unavailable [23], [30].

A. CONTRIBUTIONS AND OUTLINE

The contributions of this study are as follows: (1) Security issues are considered from the fault diagnosis perspective, and the fault diagnosis approach is adopted to formulate models of a cyber-attack. (2) A cyber-attack is assumed to be the uncertainties with the control inputs, and system states and sensor outputs are employed to describe and analyze the attack. (3) Theoretical and technical scenarios are designed to defend against the cyber-attack and enhance the resiliency of the systems.

The remainder of this paper is organized as follows. In section II, four kinds of cyber-attacks are formulated by their physical features. In section III, serval meaningful results are obtained, and corresponding defense algorithms are proposed. In section IV, a separated DC motor is taken as a typical example to demonstrate the process of a cyber-attack. Finally, the conclusions of this paper are given in section V.

B. NOTATIONS

Throughout this paper, $\Delta *$ presents the uncertainties of the parameter $*$. $E(.)$, $D(.)$ and $hash(.)$ are used to describe the encryption function, decryption function and hash function, respectively. For example, for any message x, $h(x)$ is the Hash Value of x, which is also called the message digest. $\lambda_{min}(R) = R_{min}$ represents the smallest eigenvalue of matrix R.

Table 1 summarizes the most frequently used notations throughout the remainder of this paper.

TABLE 1. **Notations.**

Notations	Descriptions
$x(k)$	state of plant in the system
$\mu(k)$	attack vector in the communication channels
$u_c(k)$	control input for the plant
$\Delta u_c(k)$	control uncertainties, which is taken as an attack
$y_p(k)$	sensor measurement or output of the plant
$\Delta y_p(k)$	output caused by uncertain inputs or cyber-attacks
$T^p_{stamp}(k)$	time-stamp of the package from the plant
$U^w_c(k)$	encrypted data of $u_c(k)$
$U^d_c(k)$	detection data by function $hash(U^w_c(k))$
$T^c_{stamp}(k)$	time-stamp of the package from the controller
$Y^w_p(k)$	encrypted sensor measurement (attack vector may be mixed)
$Y^{w-\mu}_p(k)$	encrypted sensor measurement without any attack vector
$Y^d_p(k)$	detection data by function $hash\big(y_p(k)\big)$
$Y^{d-\mu}_p(k)$	detection data by function $hash\big(y_p(k) - \Delta y_p(k)\big)$

II. FORMULATION

A. POWER CYBER-PHYSICAL SYSTEM (PCPS)

A PCPS can be formulated as a framework of networked control systems (NCS). between and among computing and physical entities there exists a communication channel. The framework is shown in Fig.1. The interactions between controller and the physical (plant) within NCSs just reveals the feature of PCPS.

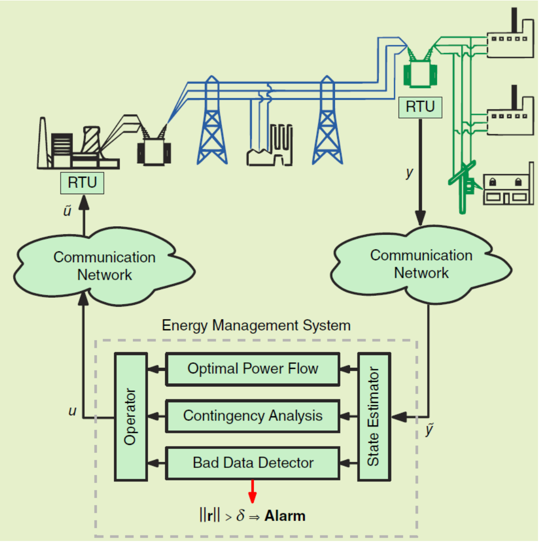


FIGURE 1. **Structure of networked control system for PCPS [26].**

Motivated by [25], the PCPS is formulated as the model of a distributed discrete-time NCS, as shown in Fig. 1, and the model is presented as

$$\begin{cases} X(k+1) = (A + \Delta A)X(k) + (B + \Delta B)U(k) + D\omega(k) \\ Y(k) = (C + \Delta C)X(k) \end{cases} \tag{1}$$

where $A + \Delta A = \sum_{i=1}^{m}(A_i + \Delta A_i)$, $B + \Delta B = \sum_{i=1}^{m}(B_i + \Delta B_i)$ and $C + \Delta C = \sum_{i=1}^{m}(C_i + \Delta C_i)$, $i \in \{1, 2, \cdots, m\}$, m is the number of subsystem.

In another case, systems are formulated as

$$\begin{cases} X(k+1) = A(X(k) + \Delta X(k)) + B(U(k) + \Delta U(k)) + D\omega(k) \\ Y(k) = (C + \Delta C)X(k) \end{cases} \tag{2}$$

with $X(k) + \Delta X(k) = \sum_{i=1}^{m}(x_i(k) + \Delta x_i(k))$, $U(k) + \Delta U(k) = \sum_{i=1}^{m}(u_i(k) + \Delta u_i(k))$ and $Y(k) + \Delta Y(k) = \sum_{i=1}^{m}(y_i(k) + \Delta y_i(k))$ for the i_{th} subsystem, m is the number of subsystem.

Remark 1: From the mathematical model, we note that the parts $\Delta AX(k)$, $\Delta CX(k)$ and $\Delta BU(k)$ in (1) are equal to the parts $A\Delta X(k)$, $C\Delta X(k)$ and $B\Delta U(k)$ in (2). However, the physical features are essentially different.

Considering one of the subsystems, the plant in NCS is formulated as

$$\begin{cases} x(k+1) = (A_p + \Delta A_p)x(k) + B_p\tilde{u}(k) + D_1\omega(k) + Ff(k) \\ y(k) = C_px(k) + D_2\upsilon(k) \end{cases} \tag{3}$$

where $x(k) \in R^{n_p}$, $\tilde{u}(k) \in R$ and $y(k) \in R^{n_y}$ denote the state of the plant, control signal input from controller side and measurement output of the plant, respectively. ΔA_p denotes the uncertainties of the system performance.

With the development of control theory and the widespread application of networks, security threats caused by adversaries in CPS have become an important problem to address. In this paper, uncertainties are considered not only for stabilization analyses but also for the system security performance.

Similar to [33], a LTI feedback controller in NCSs is presented as

$$\begin{cases} z(k+1) = A_cz(k) + B_c\tilde{y}(k) \\ u(k) = C_cz(k) + D_c\tilde{y}_c(k) \end{cases} \tag{4}$$

where $z(k) \in R^{n_z}$, $u(k) \in R^{n_x}$ and $\tilde{y}(k) \in R^{n_y}$ denote the controller state, control signal for the plant and feedback measurement of the plant, respectively. Moreover, $\tilde{y}_c(k) = \tilde{y}(k) + y_{ref}(k)$, where $y_{ref}(k)$ is the reference input, which is part of the controller side input to the controller side output is adjusted against perturbations, faults and cyber-attacks.

Remark 2: The sensor outputs are presented as $\tilde{y}(k) = y(k) + \Delta y(k)$ with $\Delta y(k)$ denoting the uncertainties. However, in this paper, excluding uncertainties, $\Delta y(k)$ is also adopted to describe cyber-attacks. In particular, if a cyber-attack is absent with no system perturbation, from the theoretical perspective, $\Delta y(k) = 0$, then $\tilde{y}(k) = y(k)$.

In controller equations, $\tilde{y}_c(k) \in R^{n_y}$ is the reference information and includes feedback and regulation parameters. The relationship between them can be given as $\tilde{y}_c(k) = y_c(k) + y_{ref}(k)$, with $y_{ref}(k)$ as part of the control input to adjust the controller output against perturbations, faults and cyber-attacks.

Referring to similar methods in previous research [34], [35], an observer-based residual detection structure is given as

$$\begin{cases} s(k+1) = A_e s(k) + B_e u_c(k) + E_e \tilde{y}(k) \\ r(k) = C_e s(k) + D_e u_c(k) + F_e \tilde{y}(k) \end{cases} \tag{5}$$

where $s(k) \in R^{n_s}$ and $r(k) \in R^{n_r}$ are the state of the anomaly detector and the residue between the estimator and virtual exists (normal), respectively.

The combination vector is defined as

$$\begin{aligned} \eta(k) &= \begin{bmatrix} x(k)^T & z(k)^T & s(k)^T \end{bmatrix}^T, \\ \xi(k) &= \begin{bmatrix} \omega(k)^T & \upsilon(k)^T \end{bmatrix}^T, \\ \mu(k) &= \begin{bmatrix} \Delta u_{att}^T(k) & \Delta y_{att}^T(k) \end{bmatrix}^T, \end{aligned}$$

and $\mathcal{R}(k)$ represents the residual error of the detector.

$$\begin{cases} \eta(k+1) = \bar{\mathcal{A}}\eta(k) + \bar{\mathcal{B}}\mu(k) + \bar{\mathcal{E}}\xi(k) + \mathcal{H}f(k) + \mathcal{G}_1 y_{ref}(k) \\ \mathcal{R}(k) = \bar{\mathcal{C}}\eta(k) + \bar{\mathcal{D}}\mu(k) + \bar{\mathcal{F}}\xi(k) + \mathcal{G}_2 y_{ref}(k) \end{cases} \tag{6}$$

where

$$\begin{aligned} \bar{\mathcal{A}} &= \begin{bmatrix} A_p + B_p D_c C_p & B_p C_c & 0 \\ B_c C_p & A_c & 0 \\ B_e D_c C_p + E_e C_p & B_e C_c & A_e \end{bmatrix}, \\ \bar{\mathcal{B}} &= \begin{bmatrix} B_p & B_p D_c \\ 0 & B_c \\ 0 & B_e D_c + E_e \end{bmatrix}, \\ \bar{\mathcal{E}} &= \begin{bmatrix} D_\omega & B_p D_c D_\nu \\ 0 & 0 \\ 0 & B_e D_c D_\nu + E_e D_\nu \end{bmatrix}, \\ \mathcal{H} &= \begin{bmatrix} F \\ 0 \\ 0 \end{bmatrix}, \quad \mathcal{G}_1 = \begin{bmatrix} B_p D_c \\ 0 \\ B_e D_c \end{bmatrix}, \\ \bar{\mathcal{C}} &= \begin{bmatrix} D_e D_c C_p + F_e C_p \\ D_e C_c \\ C_e \end{bmatrix}^T, \quad \bar{\mathcal{D}} = \begin{bmatrix} 0 \; D_e D_c + F_e \end{bmatrix}, \\ \bar{\mathcal{F}} &= \begin{bmatrix} 0 \; D_e D_c D_\nu + F_e D_\nu \end{bmatrix}, \quad \mathcal{G}_2 = D_e D_c. \end{aligned}$$

The reference input $y_{ref}(k)$ is used to adjust the outputs of the controller, which can effectively eliminate the fault or attack.

Remark 3: In (6), $y_{ref}(k)$ is specially designed to adjust the controller output, which is adopted to fight against perturbations $\omega(k)$, $\nu(k)$ and attacks $\Delta u_{att}(k)$, $\Delta y_{att}(k)$. Without losing generality, $y_{ref}(k) = \mathcal{J} x_c(k)$ is employed to evolve the model of (4), and $D_\omega = I_{n_x}$, $D_\nu = I_{n_y}$ are defined.

Then, equation (6) is rewritten as follows:

$$\begin{cases} \eta(k+1) = \bar{\mathcal{A}}_c\eta(k) + \bar{\mathcal{B}}_c\mu(k) + \bar{\mathcal{E}}_c\xi(k) + \mathcal{H}f(k) \\ \mathcal{R}(k) = \bar{\mathcal{C}}_c\eta(k) + \bar{\mathcal{D}}_c\mu(k) + \bar{\mathcal{F}}_c\xi(k) \end{cases} \tag{7}$$

where

$$\begin{aligned} \bar{\mathcal{A}}_c &= \begin{bmatrix} A_p + B_p D_c C_p & B_p C_c + B_p D_c \mathcal{J} & 0 \\ B_c C_p & A_c & 0 \\ B_e D_c C + E_e C & B_e C_c + B_e D_c \mathcal{J} & A_e \end{bmatrix}, \\ \bar{\mathcal{B}}_c &= \begin{bmatrix} B_p & B_p D_c \\ 0 & B_c \\ 0 & B_e D_c + E_e \end{bmatrix}, \quad \bar{\mathcal{E}}_c = \begin{bmatrix} I_{n_x} & B_p D_c \\ 0 & 0 \\ 0 & B_e D_c + E_e \end{bmatrix}, \\ \mathcal{H} &= \begin{bmatrix} F \\ 0 \\ 0 \end{bmatrix}, \\ \bar{\mathcal{C}}_c &= \begin{bmatrix} D_e D_c C + F_e C & D_e C_c + D_e D_c \mathcal{J} & C_e \end{bmatrix}, \\ \bar{\mathcal{D}}_c &= \begin{bmatrix} 0 & D_e D_c + F_e \end{bmatrix}, \quad \bar{\mathcal{F}}_c = \begin{bmatrix} 0 \; D_e D_c I_{n_y} + F_e I_{n_y} \end{bmatrix}. \end{aligned}$$

Remark 4: The explicit model (6) in this section is difficult to analyze theoretically. Thus, a simplification step is needed for the model. Motivated by previous research [36], we know that the fault $f(k)$ in the model can be taken as the combination of the actuator $f_a(k)$, plant $f_p(k)$ and sensor $f_s(k)$ faults.

In addition, uncertainties are other important elements for analyzing the fault $f_a(k)$ caused by $u(k)$. Thus, $\Delta u_f(k)$ is taken as a part of the control signal, which is often used to denote the controller-actuator channel attack and actuator faults in a suitable sense.

Assumption 1: Assuming $\Delta u_f(k) = \mathcal{K}\Delta x_f(k)$ for simplicity, we can easily find that $\Delta x_f(k)$ is related to $\Delta u_f(k)$ from state feedback or output feedback control law. To address this issue, we assume $\Delta u_f(k) = \mathcal{K}\Delta x_f(k)$ holds. From the state feedback or output feedback control law, we can easily determine that $\Delta x_f(k)$ is related to $\Delta u_f(k)$.

Assumption 2: Based on previous work, it is reasonable to assume that $\Delta u(k) = \Delta u_{att}(k) + \Delta u_f(k)$ and $\Delta y(k) = \Delta y_{att}(k) + \Delta y_f(k)$. Then, the definition of $\mu(k)$ can be rewritten as

$$\begin{aligned} \tilde{\mu}(k) &= \begin{bmatrix} \Delta u^T(k) & \Delta y^T(k) \end{bmatrix}^T \\ &= \begin{bmatrix} \Delta u_{att}^T(k) + \Delta u_{k_f}^T(k) & \Delta y_{att}^T(k) + \Delta y_{k_f}^T(k) \end{bmatrix}^T \end{aligned}$$

which is equal to the description of nonzero attack ($B_K u_K$, $D_K u_K$) in [31].

Consequently, the system model with faults and cyber-attacks has evolved as

$$\begin{cases} \eta(k+1) = \bar{\mathcal{A}}_z\eta(k) + \bar{\mathcal{B}}_z\tilde{\mu}(k) + \bar{\mathcal{E}}\xi(k) \\ \mathcal{R}(k) = \bar{\mathcal{C}}_z\eta(k) + \bar{\mathcal{D}}_z\tilde{\mu}(k) + \bar{\mathcal{F}}\xi(k) \end{cases} \tag{8}$$

where

$$\begin{aligned} \bar{\mathcal{A}}_z &= \begin{bmatrix} A_p + B_p D_c C_p & B_p C_c + B_p D_c \mathcal{J} & 0 \\ B_c C_p & A_c & 0 \\ B_e D_c C + E_e C & B_e C_c + B_e D_c \mathcal{J} & A_e \end{bmatrix}, \\ \bar{\mathcal{B}}_z &= \begin{bmatrix} B_p + M & B_p D_c \\ 0 & B_c \\ 0 & B_e D_c + E_e \end{bmatrix}, \\ \bar{\mathcal{E}} &= \begin{bmatrix} I_{n_x} & B_p D_c \\ 0 & 0 \\ 0 & B_e D_c + E_e \end{bmatrix}, \end{aligned}$$

$$\bar{\mathcal{C}} = \begin{bmatrix} D_eD_cC + F_eC & D_eC_c + D_eD_c\mathcal{J} & C_e \end{bmatrix},$$
$$\bar{\mathcal{D}} = \begin{bmatrix} 0 & D_eD_c + F_e \end{bmatrix},$$

Remark 5: The coefficient F varies with the function $f(k)$ since $\Delta u_f(k)$, $\Delta x_f(k)$ and $\Delta y_f(k)$ are related with A_p, B_p and C_p, respectively. Thus, the matrix F can be replaced by $\begin{bmatrix} B_p \; A_p \; C_p \end{bmatrix}$.

Remark 6: From the operational process of NCS, we find that that $\Delta x_f(k)$ is a part of $x_p(k)$; thus, $A_p\Delta x_f(k)$ is an inevitable part of $x_p(k)$. In the view of state feedback control law $u(k) = \mathcal{K}x_p(k)$, which is often equal to $\Delta u_f(k) = \mathcal{K}_k\Delta x_f(k)$. Then,we define $A_p = M\mathcal{K}$, where M is the matrix needed to be determined, such that $A_p\Delta x_f(k) = M\mathcal{K}\Delta x_f(k) = M\Delta u_f(k)$.

$$\begin{cases} \eta(k+1) = \bar{\bar{\mathcal{A}}}_z\eta(k) + \bar{\bar{\mathcal{B}}}_z\varpi(k) \\ \mathcal{R}(k) = \bar{\bar{\mathcal{C}}}_z\eta(k) + \bar{\bar{\mathcal{D}}}_z\varpi(k) \end{cases} \tag{9}$$

where $\varpi(k) = \begin{bmatrix} \tilde{\mu}^T(k) \; \xi^T(k) \end{bmatrix}^T$ denotes all coupled perturbations, and $\mathcal{R}(k)$ represents the residue of the synthesis system.

According to [28], [42], and the reference therein, the parameters uncertainties are denoted as

$$\begin{bmatrix} \Delta u_{att}(k) & \Delta u_{k_f}(k) \\ \Delta y_{att}(k) & \Delta y_{k_f}(k) \end{bmatrix} = \begin{bmatrix} H_{u_a}F_{u_a}(k)E_{u_a} & H_{u_f}F_{u_f}(k)E_{u_f} \\ H_{y_a}F_{y_a}(k)E_{y_a} & H_{y_f}F_{y_f}(k)E_{y_f} \end{bmatrix}$$

where $H_{u_a}, H_{u_f}, H_{y_a}, H_{y_f}, E_{u_a}, E_{u_f}, E_{y_a}, E_{y_f}$ are known constant matrices, and $F_{u_a}(k), F_{u_f}(k), F_{y_a}(k)$ and $F_{y_f}(k)$ are unknown matrices with Lebesgue measurable elements satisfied the conditions

$$F_{u_a}(k)F_{u_a}^T(k) \le I, \quad F_{u_f}(k)F_{u_f}^T(k) \le I$$
$$F_{y_a}(k)F_{y_a}^T(k) \le I, \quad F_{y_f}(k)F_{y_f}^T(k) \le I$$

Based on above definition, a mature and widespread lemma of uncertainty is presented as follow.

Lemma 1: For given appropriate dimensions matrix H and E, we have

$$HF(k)E + E^TF^T(k)H^T < 0$$

for all $F(k)$ satisfying $F(k)F^T(k) \le I$ if and only if there exists a positive scalar $\varepsilon > 0$, such that

$$\varepsilon HH^T + \varepsilon^{-1}E^TE < 0$$

For the considered system (9), a threshold is selected ahead, and then the judgement conditions for cyber-attack are adopted as follows:

$$\begin{cases} J_r(k) > J_{th}(k) \Rightarrow \textit{give a alarm} \\ J_r(k) \le J_{th}(k) \Rightarrow \textit{no alarm} \end{cases}$$

where $J_r(k)$ denotes the real-time residual errors for the output measurements, and $J_{th}(k)$ is pre-selected and adjusted to control requirements.

In particular, $J_{th}(k) \le \parallel \mathcal{R} \parallel$ is chosen as the threshold for detecting external faults and threats. In addition, the selection of $\mathcal{R}$ is according to control accuracy.

In control theory, the important definition of stability determination is necessary. Thus, a definition of stability based on system (9) is presented.

Definition 1: The system (9) with $\varpi(k) \equiv 0$ is said to be asymptotically stable (AS) by Lyapunov theory with quadratic Lyapunov function $f(x) = \eta^T(\underline{k})P\eta(\underline{k})$, if there exists a matrix $P > 0$ and the inequality $\bar{\mathcal{A}}_z^TP\bar{\mathcal{A}}_z - P < 0$ holds.

Definition 2: For a given parameter $\gamma > 0$, the system (9) is AS for any zero initial conditions if the following inequality holds

$$\sum_{k=0}^{\infty}\mathcal{R}^T(k)\mathcal{R}(k) \le \gamma^2\left\{\sum_{k=0}^{\infty}\varpi^2(k)\varpi(k)\right\}$$

B. INFORMATION DISCLOSURE AND PRIVACY PROTECTION

Information disclosure is fundamental to several kinds of cyber-attack. Monitoring, scanning, enumeration as well as destruction, infection, and advance persistent threat (APT) [32] are effective ways for an attacker to capture information. These scenarios also help an attacker find possible bypasses or back doors to the application software.

The information in NCSs is transmitted via the forward channels of the controller-to-actuator and feedback channels of the sensor-to-controller. In time k, adversarial attack scenarios can be modeled as

$$S_{eq}(k) = \begin{bmatrix} \Upsilon_u(k) & 0 \\ 0 & \Upsilon_y(k) \end{bmatrix}\begin{bmatrix} u(k) \\ y(k) \end{bmatrix}, \quad k \in [0, \infty)$$

$\Upsilon_u(k)$ and $\Upsilon_y(k)$ are defined to describe whether the forward channel and feedback channel are secure.

Thus, all the sequences of signal attackers have been obtained and can be described as

$$S(k) = \bigcup_{k=0}^{\infty}\left\{\begin{bmatrix} \Upsilon_u(k) & 0 \\ 0 & \Upsilon_y(k) \end{bmatrix}\begin{bmatrix} u(k) \\ y(k) \end{bmatrix}\right\}$$

In practice, $\{\Upsilon_u(k), \Upsilon_y(k)\} \in \{0, 1\}$ are the general cases, where "1" represents that both the signal transmission and reception processes are successful; otherwise, "0" indicates a failure of transmission and reception.

Furthermore, $\alpha = diag\{\alpha_i, \alpha_j\}$ and $\beta = diag\{\beta_i, \beta_j\}$ can be defined as the probability of transmission data over the forward and feedback transmission channels, respectively. Consequently, the models evolved as follows:

$$S_{eq}(k) = S_{eq}(k-1)\bigcup\left\{\begin{bmatrix} \alpha_i\Upsilon_u & 0 \\ 0 & \beta_i\Upsilon_y \end{bmatrix}\begin{bmatrix} u(k) \\ y(k) \end{bmatrix}\right\}$$

and

$$S_{attack}(k) = S_{attack}(k-1)\bigcup\left\{\begin{bmatrix} \alpha_j\Gamma_{u,a} & 0 \\ 0 & \beta_j\Gamma_{y,a} \end{bmatrix}\begin{bmatrix} u(k) \\ y(k) \end{bmatrix}\right\}$$

where $\{\alpha_i, \alpha_j, \beta_i, \beta_j\} \in [0, 1]$ denotes the probability of a successful attack in the forward and feedback channels, which are often defined as follows:

$$\begin{cases} P\{\Upsilon_u = 1\} = \alpha, & P\{\Upsilon_u = 0\} = 1 - \alpha; \\ P\{\Upsilon_y = 1\} = \beta, & P\{\Upsilon_y = 0\} = 1 - \beta; \end{cases} \tag{10}$$

Generally, resource disclosure is the first step of an attack action, which is also the first indication that the system is under attack. During this process, an adversary may capture the system model information to determine the input control signals, feedback output signals, and even the relationship of the parameters in the system such that an attacker can utilize this information to plan attack actions. Therefore, it is critical to determine whether the attacker can be discovered.

C. DoS ATTACK VERSUS PACKET LOSS

Denial-of-service (DoS) attacks prevent legitimate users from accessing a specific network resource, and the first work on this issue began in the 1980s. The distributed denial-of-service (DDoS) attack incident was first reported [44]. Because CPS is full of distributed information interactions (see [37] and the references therein), it is very important to prevent and effectively defend against DDoS attacks.

For the time interval $[k_0, k_m]$, the transmitted sequence is packaged as $S_{eq}(k) = \{S_{eq}^{k_0}(k), \cdots, S_{eq}^{k_m}(k)\}$, and each $S_{eq}^{k_i}(k)$ contains control input signal $u(k)$ and sensor output signal $y(k)$, which are taken as key data of the NCSs in detailed form as

$$\Gamma^u = \bigcup_{k=1}^{N} u(k) \;\; and \;\; \Gamma^y = \bigcup_{k=1}^{N} y(k)$$

To facilitate the analysis, the following definitions of DoS are given.

Definition 3: For a given closed-loop system, Γ^u and Γ^y are the control and sensor output sequences, respectively. Stacking them as $\Gamma = diag\{\Gamma^u, \Gamma^y\}$, with $\Gamma^u \in R^{n_u}$, $\Gamma^y \in R^{n_y}$, the models for DoS can be constructed as

$$\begin{cases} U_{DoS}^k := \Gamma^u - I_{n_u} \\ Y_{DoS}^k := \Gamma^y - I_{n_y} \end{cases} \tag{11}$$

By checking the values of U_{DoS}^k and Y_{DoS}^k, it can be determined when and which physical signals are suffering a DoS attack.

In $[k_0, k_m]$, with $0 \leq k_0 < k_m$, the DoS presence and absent interval can be generalized as

$$S_{nor} := \bigcup_{i=1}^{\infty} [s_{att}^i + \tau_i, s_{att}^{i+1}) \tag{12}$$

$$S_{att} := \left\{ s_{att}^i \right\} \cup [s_{att}^i, s_{att}^i + \tau_i), \quad i \in [0, m) \tag{13}$$

In addition, two meaningful conclusions can be drawn:

(i) $[k_0, k_m] = S_{att} \bigcup S_{nor}$;

(ii) $S_{att} \bigcap S_{nor} = \phi$.

For better comprehension, Figure 2 is given to show the process of an intermittent DoS attack.

Definition 4: For the i_{th} interval of DoS noted as $[S_{att}^i, S_{att}^i + \tau_i]$, as depicted in Fig.2, τ_i is the i_{th} DoS duration time. If the duration time τ_i can successfully cause the system to lose stability, it is called the effective DoS duration time (EDDT).

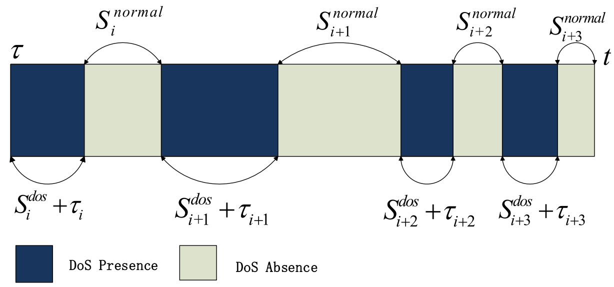


FIGURE 2. Flow diagram for security detection of an attack with an imbedded covert agent.

In a sense, DoS and packet loss exhibit almost the same characteristics for a given length of time. To find out the differences between DoS and traditional packet loss, EDDT with different lengths should be analyzed in details.

The parameters $\mathcal{S}_{PL}$, $\mathcal{S}_{ET}$, and $\mathcal{S}_{ET+PC}$ are defined to denote the maximum allowable upper bound for packet loss, the maximum time interval allowed for event triggering and the maximum effective prediction sequence length under the event-triggering scheme, respectively. Further, $0 < S_{PL} < S_{ET} < S_{ET+PC}$ and S_{ETS+PC} is needed for different control objections.

The maximum EDDT is $\mathcal{I}_i$, defined as $\mathcal{I}_i = [S_i^{DoS}, \tau_i], \forall i \in \{0, 1, \cdots\}$. If $\mathcal{I}_i \in \mathcal{S}_1 = [0, S_{PL})$, traditional methods for packet loss can be adopted to address this kind DoS; otherwise, if $\mathcal{I}_i \in \mathcal{S}_2 = [S_{PL}, S_{ET})$, the well-studied tool of the event-trigger scheme can be employed to mitigate the effect of DoS. If $\mathcal{I}_i \in \mathcal{S}_3 = [S_{ET}, S_{ET+PC})$, both event-trigger predictive control approaches can be applied together to defend against a long-duration DoS to ensure the reliability and security of CPS. Otherwise, human action should be implemented to protect the system and critical fundamentals.

D. STEALTH/COVERT ATTACK

Stealth attacks are more sophisticated than other kinds of attacks. This kind attack aims to inject false data and remain undetectable, of which man-in-the-middle (MITM) attacks are the most typical. Based on the information disclosed, attackers can make full use of the captured system information and then utilize the key-information to attack the system stealthy. Since they have almost full knowledge of the system, they not only enact a successful attack but also hide themselves well. Therefore, this case requires more sophisticated resources and more plant knowledge to find and defend against [30]. Meanwhile, the tradeoffs between utility and delectability should be considered as well.

Considering this type of cyber-attack, NCS in (1) can be rewritten as

$$\begin{cases} y_p(k) = C_p x(k) + D_2 \upsilon(k) \\ \tilde{u}(k) = \mathcal{K} x(k) \end{cases} \tag{14}$$

where $\tilde{u}(k)$ denotes the control input from controller transmitted via the network and $\tilde{u}_c(k) = \mathcal{K} x(k)$ is the state feedback control law. For a linear case, it is represented as $\tilde{u}_c(k) = u_c(k) + \Delta u_{att}(k)$, where $\Delta u_{att}(k)$ is usually taken as a cyber-attack in the controller-to-sensor channel.

Combining the two equations in (13), we obtain the following according to equation (2):

$$\begin{aligned} u_c(k) &= C_c z(k) + D_c \left[y_{p-c}(k) + y_{ref}(k) \right] \\ &= C_c z(k) + D_c \left[y_p(k) + \Delta y_{att}(k) + y_{ref}(k) \right] \end{aligned} \tag{15}$$

Remark 7: On the one hand, for a normal case, $\Delta u_{att}(k) = 0, \Delta y_{att}(k) = 0$ and, thus, $\tilde{u}_c(k) = u_c(k), y_{p-c}(k) = y_p(k)$; then, equations (13) and (15) can be used reduce to NCS (1) and (2). On the other hand, the system security is lost, which means $\mu(k) \neq 0, \Delta y_{att}(k) \neq 0$ and $\tilde{u}_c(k) \neq u_c(k), y_{p-c}(k) \neq y_p(k)$.

Combining $(13) - (15)$ gives, in the nominal case, the closed-loop response of the system as

$$y_p(k) = \Psi \left[C_p \mathcal{K}^{-1} C_c z(k) + C_p \mathcal{K}^{-1} y_{ref}(k) + D_2 \upsilon(k) \right] \tag{16}$$

where $\Psi = \left(I - C_p \mathcal{K}^{-1} D_c \right)^{-1}$.

In the presence of a stealth attack or covert agent, the control signal $u(k)$ is effected as $u(k) + \Delta u_{att}(k)$, and the feedback signal of the sensor-to-controller is $y_p(k) + \Delta y_{att}(k)$.

Since the attacker has been learning and imitating the original system, the model of the attacker can be formulated similarly to the original system as

$$\begin{cases} \Delta y_{att}(k) = \Pi_\mu \Delta u_{att}(k) \\ \Delta u_{att}(k) = \Theta_\mu \Delta y_{att}(k) + \Theta_{ref} \Delta y_{ref}(k) \end{cases} \tag{17}$$

where Π_μ, Θ_u and Θ_{ref} are the matrices that must be determined and adjusted according to the learning errors. This feedback loop is driven by the $\Delta y_{ref}(k)$ input, giving

$$\begin{cases} \Delta u_{att}(k) = \left(I - \Theta_\mu \Pi_\mu \right)^{-1} \Theta_{ref} \Delta y_{ref}(k) \\ \Delta y_{att}(k) = \Pi_\mu \left(I - \Theta_\mu \Pi_\mu \right)^{-1} \Theta_{ref} \Delta y_{ref}(k) \end{cases} \tag{18}$$

According to (14) and (15), the case $\Pi_\mu = C_p \mathcal{K}^{-1} + D_2 \upsilon(k) u_c^{-1}(k)$ is ideal, which indicates that the error between the virtual system and covert agent is zero and that the attacker has learned and mastered the original system.

E. REPLAY ATTACK (RA)

Replay attacks act by recording the history information of the control inputs or sensor measurement outputs and then inject the recorded data into the actuator or controller to disturb the steady state of either the plant or the controller to make wrong decisions.

In the actual process, a replay attack does not work alone because another program or attack behavior is required to complete its own attacks. At first, it attacks by monitoring the network or using other means to steal authentication credentials, usually cookies or an authentication session, and then returns this information to the authentication server.

The replay attack is very strong, although the encryption method can effectively prevent the plain text from being monitored, but it cannot prevent replay attacks because the attack process can be carried out by using the cipher text alone.

Based on the framework presented in Fig. 3, the sequence in the field of information security, the conventional solution is to use a "challenge response", time stamp, serial number and other methods. In the field of control theory, the χ^2 distribution, a physical watermark detection method for detecting replay attacks, has gradually become a mainstream method; see [38]–[40] and the reference therein.

Before analysis for the replay attack, an important definition is given as follow.

Definition 5: CPS has been exposed to information disclosure or infected by a virus program, while the attackers obtain the system information only, they do not act any attack strategy that has a disruptive effect on the system, which is called latency. The attacker obtains all control information and sequence information of the feedback information in this latency period and it is called a complete latency period.

During the interval $[t_k, t_{k+1})$, the attacker records the original information of the system including input and output sequences, then replay them at $t_k + s, s > t_{k+1} - t_k = \mathcal{T}$. In particular, in order to avoid been detected, the points $X(t_k + s) = X(t_k)$ or $Y(t_k + s) = Y(t_k)$ are good choices. Such that, the iteration process of RA is presented as

$$X_a(h_k + s + i) = X(h_k + i), \quad i \in \{0, 1, \cdots, \mathcal{T}\} \tag{19}$$

and

$$Y_a(h_k + s + i) = Y(h_k + i), \quad i \in \{0, 1, \cdots, \mathcal{T}\} \tag{20}$$

for forward and feedback channels. $X_a(\dot{})$ and $Y_a(\dot{})$ denote the state and output signals that have been attacked. The errors between neighboring security states/output is denoted as

$$\Delta_{\mathcal{T}} = |X(t_k + s + T + 1) - X(t_k + s + T)|$$

In order to ensure the stability of the system, $\Delta_{\mathcal{T}}$ is assumed to be under a certain upper bound.

Select a control sequence of length h and output sequence data as contrast sequences, which can be represented as

$$\begin{aligned} u_T(\Delta k) &= \begin{bmatrix} u(k_i) & u(k_i+1) & \cdots & u(k_i+\hbar) \end{bmatrix} \\ y_T(\Delta k) &= \begin{bmatrix} y(k_i) & y(k_i+1) & \cdots & y(k_i+\hbar) \end{bmatrix} \end{aligned}$$

The length value h is the experience value, which can be adjusted. When the same sequence as the contrast sequence data is detected, bidirectional expansion will be carried out based on the two ends of the contrast sequence, respectively. This scenario called "dynamic windows," which is effective for replay attack detection.

III. MAIN RESULTS

Before giving the main results, two significant assumptions are presented that will play important roles in deducing the main results.

Assumption 3: The attacker can learn the system model well. Therefore, he clearly knows the rules between the inputs and outputs of the system plant. For example, the input-output

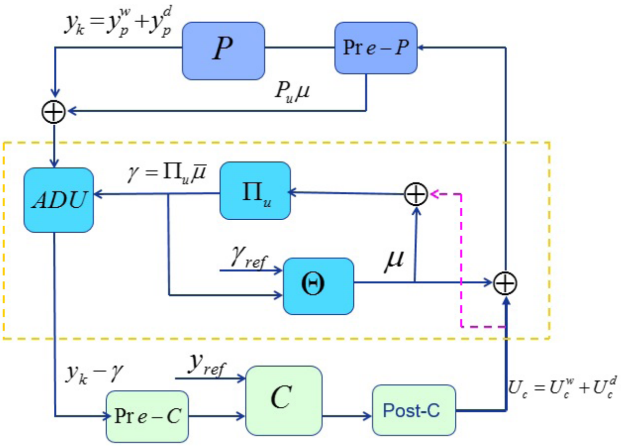


FIGURE 3. A novel defense structure of NCS against deception attack.

set of the plant is (x_1, y_1), and the attack set input-output effect on the plant is (μ, γ). If $\mu = x_1$, then $y_1 = \gamma$.

A. SYSTEM DESIGN OBJECTIVES

The objective of IT security is to protect the information/data in the communication channel rather than the physical resources of the system. This is the inherent difference between networked-based security control and traditional IT security work.

The sequence sent out from the controller is packaged as $\{T_{stamp}(k), U_c^w(k), U_c^d(k)\}$ and includes the time-stamp, the real control and its detection signal by a hash function. $T_{stamp}(k)$ is the time-stamp function at instant k when the signal is transmitted. In this paper, it is defined as

$$T_{stamp}\,(k) = E\,(co\,\{hash_1\,(k)\,, time\,(k)\,, date\,(k)\}) \quad (21)$$

The other parts are obtained by $U_c^w(k) = E(u_c(k))$ and $U_c^d(k) = hash(U_c^w(k))$ to denote the encrypted control signal and the detection signal. Because the hash function is nonreversible in computation, a private shared hash algorithm is insensitively scrutinized to guarantee the transmission data to be unique.

Based on the above works, some significant conclusions can be derived as follows.

Theorem 1: For the packaged transmission sequence $S_{eq}^c(k) = \{T_{stamp}^c(k), U_c^w(k), U_c^d(k)\}$ from controller to actuator, the actuator side receiving a packaged sequence is $\tilde{S}_{eq}^c(k) = \{\tilde{T}_{stamp}^c(k), \tilde{U}_c^w(k), \tilde{U}_c^d(k)\}$. If the controller-to-actuator channel within the system (3)–(4) is secured without any information disclosure or tampering, the following conditions should be satisfied.

(i) $D\left(\tilde{T}_{stamp}^c(k)\right) > D\left(T_{stamp}^c(k)\right)$;

(ii) $hash(\tilde{U}_c^w(k)) = \tilde{U}_c^d(k) = U_c^d(k)$;

Proof 1: If conditions (i) and (ii) are satisfied, the information with $S_{eq}^c(k)$ and $\tilde{S}_{eq}^c(k)$ is identical, which indicates the transmitted data via the forward channel is secure without any modification. The integrity is ensured in this case.

Otherwise, if one or even both conditions are not met, then $S_{eq}^c(k) \neq \tilde{S}_{eq}^c(k)$. Under these circumstances, the integrity of the package is destroyed, and the transmitted package has been intercepted and tampered with.

The aforementioned assumptions indicate that $U_c^d(k) = \tilde{U}_c^d(k)$, which is ensured by the practical technique's hidden coding. Then, we can judge whether the information of channel transmission is attacked by condition 2.

Similar to the detection approach in Theorem 1, a similar result is obtained for security detection of the sensor-to-controller channel.

Theorem 2: For the packaged sequence $\Gamma_p(k) = \{T_{stamp}^p(k), Y_p^w(k), Y_p^d(k)\}$ from the sensor to controller, the controller side receiving an encrypted sequence is $\tilde{\Gamma}_p(k) = \{\tilde{T}_{stamp}^p(k), \tilde{Y}_p^w(k), \tilde{Y}_p^d(k)\}$. If the sensor-to-controller channel within the system (3)–(4) is secure without any information disclosure or tampering, the following conditions should be satisfied.

(i) $D\left(\tilde{T}_{stamp}^p(k)\right) > D\left(T_{stamp}^p(k)\right)$;

(ii) $hash(\tilde{Y}_p^w(k)) = \tilde{Y}_p^d(k) = Y_p^d(k)$;

Referring to the proof of Theorem 1, the proof of this theorem can be obtained similarly. The detection approach in Theorem 2 is based on a hidden condition that the controller-to-actuator channel is secure.

Remark 8: By applying the two judgement conditions of Theorem 1, we find the attack vector $\mu(k)$ and then extract and separate it as the uncertainties $\Delta u_c(k) = \mu(k)$. According to $u_c(k)$ and $u_c(k) + \Delta u_c(k)$, the post-plant unit can obtain the sensor output $\Delta y_p(k)$ and $y_p(k)$. In order to detect covert agent, we define $y_p^{w-\mu}(k) = y_p(k) - \Delta y_p(k)$. From the hash function, the detection vectors $Y_p^{d-\mu}(k)$ and $Y_p^d(k)$ are derived.

Based on above remark, a further result can be deduced.

Theorem 3: The package sequence $\Gamma_p(k) = \{T_{stamp}^p(k), Y_p^w(k), Y_p^{d-\mu}(k), Y_p^d(k)\}$ will transmitted from the sensor side to the controller side via the network, and the controller side receiving the package is $\tilde{\Gamma}_p(k) = \{\tilde{T}_{stamp}^p(k), \tilde{Y}_p^w(k), \tilde{Y}_p^{d-\mu}(k), \tilde{Y}_p^d(k)\}$. Using Theorems 1 and 2, it can be deduced that the sensor-to-controller channel within the system Σ composed by (3) - (5) is secure with no covert agent, if the following conditions are satisfied:

(i) $D\left(T_{stamp}^p\right) \leq D\left(\tilde{T}_{stamp}^p\right)$;

(ii) $hash\left(\tilde{Y}_p^w(k)\right) = \tilde{Y}_p^d(k) = Y_p^d(k)$;

(iii) $hash\left(\tilde{Y}_p^w(k)\right) \neq \tilde{Y}_p^{d-\mu}(k)$;

Furthermore, for the transmission sequence, $\Gamma_p(k) = \left\{T_{stamp}^p(k), Y_p^{w-\mu}(k), Y_p^{d-\mu}(k), Y_p^d(k)\right\}$ and the received sequence $\bar{\Gamma}_p(k) = \left\{\bar{T}_{stamp}^p(k), \bar{Y}_p^{w-\mu}(k), \bar{Y}_p^{d-\mu}(k), \bar{Y}_p^d(k)\right\}$, we can determine that there exists a covert agent between the channels of plant and controller if the following conditions are satisfied:

(I) $D\left(T_{stamp}^p\right) \leq D\left(\bar{T}_{stamp}^p\right)$;

(II) $hash\left(\bar{Y}_p^{w-\mu}(k)\right) = \bar{Y}_p^{d-\mu}(k)$;

(III) $hash\left(\bar{Y}_p^{w-\mu}(k)\right) \neq \bar{Y}_p^d(k)$;

Proof 2: If the sensor-to-controller channel within the system is secure, the condition (i) $\tilde{D}\left(T_{stamp}^p(k)\right) \geq D\left(T_{stamp}^p(k)\right)$ can be derived directly because the condition (i) is not sufficient for judgment. However, information disclosure and tampering do not occur, so $\Gamma_p(k) = \tilde{\Gamma}_p(k)$ holds. This is also the case for Theorem 2. In addition, referring to the definition of parts within $\Gamma_p(k)$ and $\tilde{\Gamma}_p(k)$, before data transmission, $hash\left(Y_p^w(k)\right) = Y_p^d(k)$ and $hash\left(E\left(y_p(k) - \mu(k)\right)\right) = Y_p^{d-\mu}(k)$ are determined. Hence, the above definitions are used to judge the security of the system and whether a covert agent existed.

Otherwise, if there exists an undetectable covert agent (CA) attacker, the injected uncertainties of the control input are detected and separated based on the approach of Theorem 1. From the package schedule, we know $hash\left(D(y(k))\right) = hash(Y_p^w(k)) = Y_p^d(k)$ and $hash\left(D(y(k) - \Delta y(k))\right) = Y_p^{d-\mu}(k)$. Since the CA attacker has full knowledge of the plant, together with the ability of listening the communication channels between sensor and controller, thus it can remove the effects. When the parts of the controller side received the sequence $\bar{\Gamma}_p(k)$, conditions (I), (II), and (III) alert the users to the existence of the CA attacker.

Remark 9: First, because the hash function has one-way features, the attacker cannot parse the original message from the hash values. Second, if the hash function is determined, the hash value will be same for the same information. Finally, the industrial system is always functionally periodic, and the sampled data follow certain operating rules, which ensure that the historical database is healthy and trustable.

In order to describe the design philosophy of defense scenarios clearly, a flow chart is given in Fig.4.

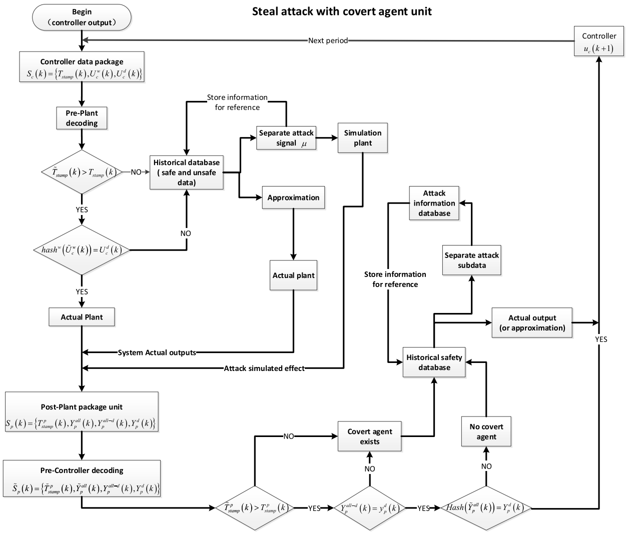

Steal attack with covert agent unit
Begin (controller output)
Next period
Controller $u_c(k+1)$
Controller data package $S_c(k) = \{T_{stamp}(k), U_c^w(k), U_c^d(k)\}$
Pre-Plant decoding
$\tilde{T}_{stamp}(k) > T_{stamp}(k)$
NO
YES
$hash^w(\tilde{U}_c^w(k)) = U_c^d(k)$
NO
YES
Historical database (safe and unsafe data)
Store information for reference
Separate attack signal μ
Simulation plant
Approximation
Actual plant
Actual Plant
System Actual outputs
Attack simulated effect
Attack information database
Separate attack subdata
Store information for reference
Actual output (or approximation)
YES
Historical safety database
Post-Plant package unit $S_p(k) = \{T_{stamp}^p(k), Y_p^{all}(k), Y_p^{all-d}(k), Y_p^d(k)\}$
Pre-Controller decoding $\tilde{S}_p(k) = \{\tilde{T}_{stamp}^p(k), \tilde{Y}_p^{all}(k), Y_p^{all-d}(k), Y_p^d(k)\}$
$\tilde{T}_{stamp}^p(k) > T_{stamp}^p(k)$
NO
Covert agent exists
YES
$Y_p^{all-d}(k) = y_p^d(k)$
NO
YES
$Hash(\tilde{Y}_p^{all}(k)) = Y_p^d(k)$
NO
No covert agent

FIGURE 4. Flow-chart of defense scenario.

IV. EXAMPLES

A. EXAMPLE 1

In this section, a well-known example of DC motor speed regulation presented to illustrate the usefulness of the proposed approach. The system diagram can be expressed in Fig. 5.

Due to the previous adjustable speed performance in a wide range, DC motors have been widely used in industrial systems. Although these motors have nonlinear characters in performance, the magnetization curve is considered linear in practice. The armature voltage u_a is generally formulated as

$$u_a = R_a i_a + L_a \frac{di_a}{dt} + K_m \phi \omega \tag{22}$$

where ϕ is the pole flux, which has a hysteretic nonlinearity. In the virtual case, it is usually operated in the linear region for simplicity as $\phi = L_f i_f$.

The field voltage u_f is formulated by the equation

$$u_f = R_f i_f + \frac{d\phi}{dt} \tag{23}$$

The electric torque generated by the motor is calculated with the following equation:

$$T_e = K_m \phi i_a \tag{24}$$

This is a static torque. Because most DC motors run under load mode, the dynamic torque should be considered, and it is formulated as

$$T_e = J\frac{d\omega}{dt} + B_m \omega + T_L \tag{25}$$

where T_L is the mechanical load torque, which is often taken as the unmeasured load disturbance torque.

The parameter notations in above equations are given in Table 2 together with example values.

TABLE 2. Parameter notations of the DC motor system.

Parameters	Physical concept	value
P	Rated power	3.73 kw
ω_{ref}	Rated speed	183.26 Rad/s
$i_{a(norm)}$	Rated armature current	16.74 A
$i_{f(norm)}$	Rated field current	4 A
T_e	Rated torque	18 Nm
u_f	Field voltage (maximum)	240 V
L_a	Armature inductance	0.01 H
L_f	Field winding inductance	60 H
J_m	Motor inertia	$0.208kgm^2$
B_m	Motor damping	$0.011kgm^2$
K_m	Motor torque constant	$0.3Nm/A^2$
R_a	Armature resistance	0.21Ω
R_f	Field resistance	146.7Ω

From (22)-(25), we have

$$\begin{bmatrix} \frac{di_a}{dt} \\ \frac{di_f}{dt} \\ \frac{d\omega}{dt} \end{bmatrix} = \begin{bmatrix} -\frac{R_a}{L_a} i_a - K_m i_f \omega \\ -\frac{R_f i_f}{L_f} \\ \frac{1}{J}\left(-B\omega + K_m i_a i_f - T_L\right) \end{bmatrix} + \begin{bmatrix} \frac{1}{L_a} u_a \\ 0 \\ 0 \end{bmatrix} \tag{26}$$

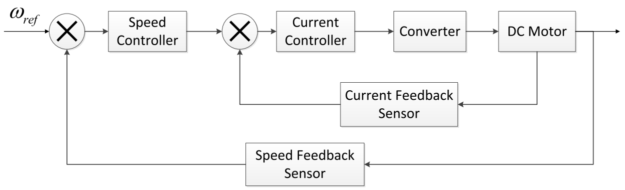


FIGURE 5. **Conventional structure of DC motor control system.**

Similar to [41], we define the system state as $x(t) = \left[\, i_a(t)\ \omega(t)\ x_3(t) \,\right]^T$, $u(t) = U_a(t)$, $x_3(t) = \int(\omega - \omega_{ref})dt$.

According to above definitions, the DC dynamic model can be given as

$$\dot{x}(t) = (A + \Delta A)x(t) + B\left(u(t) + \Delta u(t)\right) + \tilde{\omega}(t) \qquad (27)$$

where

$$A = \begin{bmatrix} -\dfrac{R_a}{L_a} & -\dfrac{K_m u_f}{L_a R_f} & 0 \\ \dfrac{K_m u_f}{J R_f} & -\dfrac{B}{J} & 0 \\ 0 & -1 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} \dfrac{1}{L_a} \\ 0 \\ 0 \end{bmatrix},$$

$$\tilde{\omega}(t) = \begin{bmatrix} 0 \\ 0 \\ \omega_{ref} \end{bmatrix}, \quad u(t) = U_a(t).$$

The obvious relationship of uncertainties and the normal parameters is $R_a = R_{a_{normal}} + \Delta R_a$, $R_f = R_{f_{normal}} + \Delta R_f$, and $T_L = T_{L_{normal}} + \Delta T_L$. Furthermore, the control input is presented as $u_a(t) = u_{a_{normal}}(t) + \Delta u_a(t)$, which is usually neglected. However, because of its important influence to the system, it has been reconsidered in many recent publications.

The uncertainties are often caused by the armature and field resistance as well as the load torque. Hence, ΔR_a, ΔR_f and ΔT_L are used to denote the errors from the nominal values. From recent studies, we find that the uncertainties $\Delta u_a(t)$ are generally caused by cyber-attacks.

Based on above description, we have

$$A = \begin{bmatrix} -\dfrac{R_{a_{normal}}}{L_a} & -\dfrac{K_m U_f}{L_a R_{f_{normal}}} & 0 \\ \dfrac{K_m U_f}{J R_{f_{normal}}} & -\dfrac{B}{J} & 0 \\ 0 & -1 & 0 \end{bmatrix},$$

$$\Delta A = \begin{bmatrix} -\dfrac{\Delta R_a}{L_a} & -\dfrac{K_m U_f}{L_a \Delta R_f} & 0 \\ \dfrac{K_m U_f}{J \Delta R_f} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$
$$= \begin{bmatrix} \Delta A_{11} & \Delta A_{12} & 0 \\ \Delta A_{21} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

where the parameter uncertainties are often decomposed as $\Delta A = HF(t)E$ and H and E are known real constant matrices with appropriate dimensions. Meanwhile, $F(t)$ is an unknown matrix, which satisfies $F(t)F^T(t) \leq I$.

In this study, the static state feedback controller is designed as

$$u(t) = -Kx(t) = -\left[\, k_1 \quad k_2 \quad k_3 \,\right] x(t) \qquad (28)$$

According to the robust controller design theory ([41] and the references therein), the above controller can be obtained using the [42, Lemma 2.5] to obtain a symmetric definite matrix P. The system parameters are same as those of [41] such that $K = \left[\, 0.37265\ 1.1029\ {-8.0814} \,\right]$ can be borrowed for simulation works.

The parameter uncertainties are varying with ±10% of the normal uncertainties; thus, we have

$$\Delta A = \begin{bmatrix} -12 & -133.3 & 0 \\ 64.1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

or

$$\Delta A = \begin{bmatrix} 12 & 133.3 & 0 \\ -164.1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Applying the uncertainties as an effect of the attack, we have

Fig. 6 shows that the uncertainties caused a decline of the motor speed; the blue line represents the normal case, and

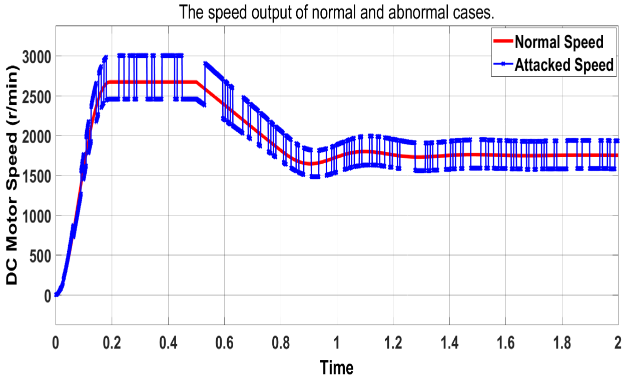


FIGURE 6. **Speed of the DC motor of the normal case and with uncertainties.**

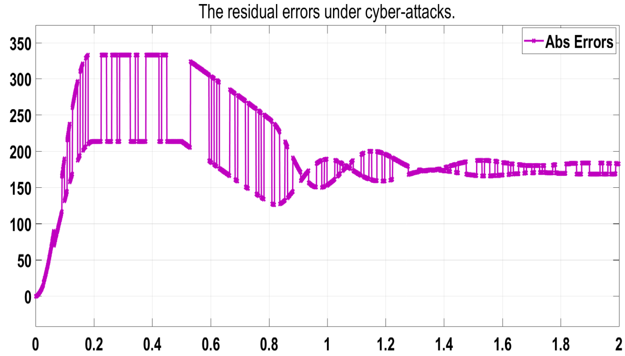


FIGURE 7. **The speed errors under cyber-attack.**

the red line denotes the abnormal case. In the sophisticated method of [30], the speed decline cannot be detected easily. How to guarantee the detectable of the varying of the parameters, especially for critical parameters, requires further study. Fig. 7 shows the speed errors between the normal and abnormal cases of the DC motor. After the unstable initial process, the errors caused by an uncertain attack are limited in the span of $0-200r/min$ and greater than 10% of the rated speed.

V. CONCLUSION

Serial main cyber-attack have been analyzed from the uncertainty perspective. From the analysis, we realized that modeling errors, actuator faults, sensor faults, and even the typical cyber-attack – stealth attacks, covert attacks and denial-of-service – can be formulated in a united form. This is also the main contribution of this paper. To eliminate these attacks, a double closed-loop structure is designed for attack defense. This is another contribution of our work.

REFERENCES

[1] Presidents Council of Advisors on Science and Technology. *Leadership Under Challenge: Information Technology R&D in a Competitive World.* Aug. 2007. [Online]. Available: http://www.nitrd.gov/Pcast/reports/PCAST-NIT-FINAL.pdf

[2] R. Rajkumar, I. Lee, L. Sha, and J. Stankovic, "Cyber-physical systems: The next computing revolution," in *Proc. Design Autom. Conf.*, 2010, pp. 731–736.

[3] Y. Rechtman, "Guide for conducting risk assessments: Information security," *CPA J.*, vol. 83, no. 3, 2013.

[4] K. Wang, Y. Wang, Y. Sun, S. Guo, and J. Wu, "Green industrial Internet of Things architecture: An energy-efficient perspective," *IEEE Commun. Mag.*, vol. 54, no. 12, pp. 48–54, Dec. 2016.

[5] Y. Cao, "The road of China's energy Internet white paper," *China Elect. Equipment Ind.*, no. 7, pp. 38–44, 2015.

[6] K. Wang, Y. Shao, L. Shu, G. Han, and C. Zhu, "LDPA: A local data processing architecture in ambient assisted living communications," *IEEE Commun. Mag.*, vol. 53, no. 1, pp. 56–63, Jan. 2015.

[7] C. Neuman, "Challenges in security for cyber-physical systems," in *Proc. DHS Workshop Future Directions Cyber-Phys. Syst. Secur.*, 2009, pp. 22–24.

[8] K. Wang *et al.*, "A survey on energy Internet: Architecture, approach, and emerging technologies," *IEEE Syst. J.*, to be published, doi: 10.1109/JSYST.2016.2639820.

[9] K. Wang, H. Li, Y. Feng, and G. Tian, "Big data analytics for system stability evaluation strategy in the energy Internet," *IEEE Trans. Ind. Inform.*, vol. 13, no. 4, pp. 1969–1978, Aug. 2017.

[10] H.-J. Appelrath, O. Terzidis, and C. Weinhardt, "Internet of energy," *Bus. Inf. Syst. Eng.*, vol. 4, pp. 1–2, Feb. 2012.

[11] H. Jiang, K. Wang, Y. Wang, M. Gao, and Y. Zhang, "Energy big data: A survey," *IEEE Access*, vol. 4, pp. 3844–3861, 2016.

[12] K. Wang, X. Hu, H. Li, P. Li, D. Zeng, and S. Guo, "A survey on energy Internet communications for sustainability," *IEEE Trans. Sustain. Comput.*, vol. 2, no. 3, pp. 231–254, May 2017.

[13] K. Wang, M. Du, S. Maharjan, and Y. Sun, "Strategic honeypot game model for distributed denial of service attacks in the smart grid," *IEEE Trans. Smart Grid*, vol. 8, no. 5, pp. 2474–2482, Feb. 2017.

[14] R. Baheti and H. Gill, "Cyber-physical systems," *Impact Control Technol.*, vol. 12, pp. 161–166, Mar. 2011.

[15] K. Wang *et al.*, "Wireless big data computing in smart grid," *IEEE Wireless Commun.*, vol. 24, no. 2, pp. 58–64, Apr. 2017.

[16] A. A. Cárdenas, S. Amin, B. Sinopoli, A. Perrig, and S. Sastry, "Challenges for securing cyber physical systems," in *Proc. Workshop Cyber-Phys. Syst. Secur.*, 2006, pp. 1–7.

[17] K. Wang, Z. Ouyang, R. Krishnan, L. Shu, and L. He, "A game theory-based energy management system using price elasticity for smart grids," *IEEE Trans. Ind. Informat.*, vol. 11, no. 6, pp. 1607–1616, Dec. 2015.

[18] K. Wang *et al.*, "Distributed energy management for vehicle-to-grid networks," *IEEE Netw.*, vol. 31, no. 2, pp. 22–28, Mar./Apr. 2017.

[19] K. Wang, C. Xu, Y. Zhang, S. Guo, and A. Zomaya, "Robust big data analytics for electricity price forecasting in the smart grid," *IEEE Trans. Big Data*, to be published.

[20] K. Wang, M. Du, D. Yang, C. Zhu, J. Shen, and Y. Zhang, "Game-theory-based active defense for intrusion detection in cyber-physical embedded systems," *ACM Trans. Embedded Comput. Syst.*, vol. 16, no. 1, 2016, Art. no. 18.

[21] K. Wang, M. Du, Y. Sun, A. Vinel, and Y. Zhang, "Attack detection and distributed forensics in machine-to-machine networks," *IEEE Netw.*, vol. 30, no. 6, pp. 49–55, Nov./Dec. 2016.

[22] A. Cardenas, S. Amin, and S. Sastry, "Secure control: Towards survivable cyber-physical systems," in *Proc. 28th Int. Conf. Distrib. Comput. Syst. Workshops (ICDCS)*, Jun. 2008, pp. 495–500.

[23] H. Sandberg, S. Amin, and K. Johansson, "Cyberphysical security in networked control systems: An introduction to the issue," *IEEE Control Syst.*, vol. 35, no. 1, pp. 20–23, Feb. 2015.

[24] J. P. How, "Cyberphysical security in networked control systems [about this issue]," *IEEE Control Syst.*, vol. 35, no. 1, pp. 8–12, Feb. 2015.

[25] P. Kundur, *Power System Stability and Control*. New York, NY, USA: McGraw-Hill Professional, 1994.

[26] A. Teixeira, K. C. Sou, H. Sandberg, and K. H. Johansson, "Secure control systems: A quantitative risk management approach," *IEEE Control Syst.*, vol. 35, no. 1, pp. 24–45, Feb. 2015.

[27] Q. Zhu and T. Basar, "Game-theoretic methods for robustness, security, and resilience of cyberphysical control systems: Games-in-games principle for optimal cross-layer resilient control systems," *IEEE Control Syst.*, vol. 35, no. 1, pp. 46–65, Feb. 2015.

[28] D. Yue, E. Tian, and Q.-L. Han, "A delay system method for designing event-triggered controllers of networked control systems," *IEEE Trans. Autom. Control*, vol. 58, no. 2, pp. 475–481, Feb. 2013.

[29] A. Householder, A. Manion, L. Pesante, and G. Weaver, "Managing the threat of denial-of-service attacks," *Cert Coordination Center*, vol. 33, no. 4, pp. 99–110, 2001.

[30] R. S. Smith, "Covert misappropriation of networked control systems: Presenting a feedback structure," *IEEE Control Syst.*, vol. 35, no. 1, pp. 82–92, Feb. 2015.

[31] F. Pasqualetti, F. Dörfler, and F. Bullo, "Attack detection and identification in cyber-physical systems," *IEEE Trans. Autom. Control*, vol. 58, no. 11, pp. 2715–2729, Nov. 2013.

[32] E. D. Knapp and J. T. Langill, *Industrial Network Security: Securing Critical Infrastructure Networks for Smart Grid, SCADA, and Other Industrial Control Systems*. Boston, MA, USA: Syngress, 2011.

[33] K. Zhou, J. C. Doyle, and K. Glover, *Robust and Optimal Control* (Prentice Hall Information and System Sciences Series). Upper Saddle River, NJ, USA: Prentice-Hall; 1996, p. 538.

[34] S. Ding, *Model-Based Fault Diagnosis Techniques: Design Schemes, Algorithms, and Tools*. Springer, 2008.

[35] I. Hwang, S. Kim, Y. Kim, and C. E. Seah, "A survey of fault detection, isolation, and reconfiguration methods," *IEEE Trans. Control Syst. Technol.*, vol. 18, no. 3, pp. 636–653, May 2010.

[36] S. Simani, C. Fantuzzi, and R. J. Patton, *Model-Based Fault Diagnosis in Dynamic Systems Using Identification Techniques*. London, U.K.: Springer, 2003.

[37] C. D. Persis and P. Tesi, "Input-to-state stabilizing control under denial-of-service," *IEEE Trans. Autom. Control*, vol. 60, no. 11, pp. 2930–2944, Nov. 2015.

[38] Y. Mo and B. Sinopoli, "Secure control against replay attacks," in *Proc. 47th Annu. Allerton Conf. Commun., Control, Comput.*, Sep./Oct. 2009, pp. 911–918.

[39] Y. Mo, R. Chabukswar, and B. Sinopoli, "Detecting integrity attacks on SCADA systems," *IEEE Trans. Control Syst. Technol.*, vol. 22, no. 4, pp. 1396–1407, Jul. 2014.

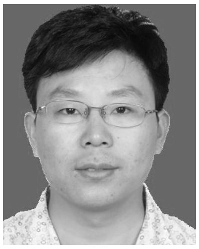
[40] Y. Mo, S. Weerakkody, and B. Sinopoli, "Physical authentication of control systems: Designing watermarked control inputs to detect counterfeit sensor outputs," *IEEE Control Syst.*, vol. 35, no. 1, pp. 93–109, Feb. 2015.

[41] J. Zhou, Y. Wang, and R. Zhou, "Global speed control of separately excited DC motor," in *Proc. Power Eng. Soc. Winter Meeting*, Jan./Feb. 2001, pp. 1425–1430.

[42] A. Teixeira, I. Shames, H. Sandberg, and K. H. Johansson, "A secure control framework for resource-limited adversaries," *Automatica*, vol. 51, pp. 135–148, Jan. 2015.

[43] P. Lee, A. Clark, L. Bushnell, and R. Poovendran, "A passivity framework for modeling and mitigating wormhole attacks on networked control systems," *IEEE Trans. Autom. Control*, vol. 59, no. 12, pp. 3224–3237, Dec. 2013.

[44] P. J. Criscuolo, "Distributed denial of service Trin00, tribe flood network, tribe flood network 2000, and stacheldraht CIAC-2319," Dept. Energy Comput. Incident Advisory Capability, Lawrence Livermore Nat. Lab., Livermore, CA, USA, Tech. Rep. UCRL-ID-136939, Feb. 2000.



HUI GE (M'17) received the B.S. degree in mechanotronics engineering and automation and the M.S. degree in theory and new technology of electrical engineering from Nanjing Normal University, Nanjing, China, in 2006 and 2009, respectively. He is currently pursuing the Ph.D. degree in information acquisition and control with the Nanjing University of Posts and Telecommunications, Nanjing. He is also a Lecturer with Taizhou University, Taizhou, China. His current research interests include the analysis and synthesis of networked control systems, fault diagnosis, and CPS security control.

ZHENJIANG ZHAO received the B.S. degree from the Luoyang Institute of Science and Technology, Luoyang, China, in 1986. He is currently a Professor and the Dean of the School of Shipping and Mechatronics Engineering, Taizhou University, Taizhou, China. His current research interests include the analysis and synthesis of networked control systems, multiagent systems, optimal control of power systems, and Internet of Things.