Secure Control for Cyber-Physical Systems under Malicious Attacks

Wu, Chengwei; Yao, Weiran; Pan, Wei; Sun, Guanghui; Liu, Jianxing; Wu, Ligang

DOI
10.1109/TCNS.2021.3094782

Publication date
2021

Document Version
Accepted author manuscript

Published in
IEEE Transactions on Control of Network Systems

Secure Control for Cyber-Physical Systems under Malicious Attacks

Chengwei Wu, Weiran Yao, Wei Pan, Guanghui Sun, Jianxing Liu, and Ligang Wu, *Fellow, IEEE*

***Abstract*— This paper investigates the secure control problem for cyber-physical systems when the malicious data is injected into the cyber realm which is directly connecting to the actuators. Based on moving target defense and reinforcement learning, we propose a novel proactive and reactive defense control scheme. First, the system (A, B) is modeled as a switching system consisting of several controllable pairs $(\mathrm{A}, \mathcal{B}_{\mathrm{l}})$ to facilitate the construction of the moving target defense control scheme. The controllable pairs $(\mathrm{A}, \mathcal{B}_{\mathrm{l}})$ can be altered to update system dynamics under certain unpredictable switching probabilities for each subsystem, which can prevent the adversaries from effective attacks. Second, both attack detection and isolation schemes are designed to accurately locate and exclude the compromised actuators from a switching sequence. Third, a reinforcement learning algorithm based on the zero-sum game theory is proposed to design the defense control scheme when there exist no controllable subsystems to switch. To demonstrate the effectiveness of the defense control scheme, a three-tank system under unknown cyber attacks is illustrated.**

***Index Terms*— Actuator attacks, Moving target defense, Reinforcement learning, Proactive and reactive control, Cyber-physical systems.**

1.. INTRODUCTION

The cyber-physical systems (CPS) have tremendous economic and societal impact and potential [1]. The applications of CPS can range from the military to the civil critical infrastructure, and more. The CPS consist of the cyber realm and the physical layer. The cyber realm is mainly utilized to be in charge of interactions between the cyber world and the physical world, and the physical layer governs the physical dynamics [2]. The quality of cyber realm can affect the performance of the physical process and vice versa. Due to the integration of the cyber layer, there exists a big risk that exogenous cyber attacks can intrude the communication networks/computers, which can vastly degrade the performance or even destroy the CPS. Examples of the attack cases include the elaborately designed Stuxnet [3], attacking the Maroochy water services in Australia [4], etc. Some other attack incidents can refer to [5]. It is extremely urgent to dissuade CPS from being intruded. Effective mechanisms for CPS security need to be designed. Some challenges for securing CPS using the control theory have been outlined in [6]. Researchers from the control field have paid more attention to security problems mainly including attack detection and identification, secure state reconstruction and defense control [7]. In the following, a brief review concerning these three directions is given.

As to attack detection, the χ^2 detector and observer-based detection scheme can be found in existing results. In [8], a countermeasure has been proposed to improve the performance of χ^2 attack detector in the presence of sensor replay attacks. But either the increasing of the control cost or the detection delay can arise. In [9], the undetectable and unidentifiable malicious signals have been characterized based on the system-theoretic and graph-theoretic approaches. The monitors for attacks have been designed. Should the robustness of the proposed schemes be guaranteed, the schemes can perform better in practical applications. In [10], an unknown input observer based attack identification scheme has been proposed. Although the identification scheme is similar to our paper, our paper solves the identification delay problem. Additionally, its control scheme is reactive.

For the remote estimate problem, several important results have appeared in literature [11]–[13]. In [11], a Kalman filtering approach has been designed using intermittent measurements. The critical probability of the measurements that the filter receives has been derived, within which the filtering approach can converge. In [12], the relationship between the number of sensors that allows to be disrupted and the state reconstruction was revealed. If the number of disrupted sensors exceeds half of its total number, states cannot be exactly reconstructed. Such a conclusion have been applied in [14], [15].

For the secure control problem under attacks, it can refer to schemes in [16]–[18]. In [16], novel definitions of attack frequency and attack duration have been proposed to model denial-of-service attacks, based on which a secure control scheme was derived to preserve the input-to-state stability. Using the novel DoS attack model in [16], a secure controller using the sliding mode has been designed in [19], where a defense scheme has been established to guarantee that the attack model can be satisfied. In [17], an adaptive control framework has been presented for systems under false data injection attacks. However, the schemes are designed from the system designer's perspective. An adversary is not considered in the aforementioned results. The game-theoretical approach

This work was supported in part by the National Key R&D Program of China (No. 2019YFB1312001), National Natural Science Foundation of China (62033005, 62022030, 62003114), and the State Grid Heilongjiang Electric Power Company Limited funded project (No. 522417190057).

C. Wu, W. Yao, G. Sun, J. Liu and L. Wu are all with the Department of Control Science and Engineering, Harbin Institute of Technology, Harbin 150001, P.R. China.

W. Pan is with the Department of Cognitive Robotics, Delft University of Technology, Netherlands.

can take both sides into a unified framework to analyze and design CPS [20], [21]. Examples of results concerning game-based CPS design are a remote state estimate scheme in [22], a defense control scheme in [23] and a learning-based defense controller in [24].

However, the current methods of attack detection, state recovering and control mechanisms are reactive which limits the broad applications. An adversary can monitor the system for a long time and acquire the defense scheme. After having knowledge of the defense scheme, attacks can be implemented again. Alternatively, blended attacks are launched on the system, which also makes the defense scheme ineffective. The fundamental reason is that the system structure is fixed. If the system structure is dynamical, this issue can be solved. Moving target defense (MTD), which is often applied to computer networks, provides a novel approach to securing CPS. Such a scheme makes the system to be a dynamical structure, that is, the system structure is regarded a moving target and it can be altered unpredictably. In this way, it is more difficult for adversaries to successfully intrude the CPS. The advantages that adversaries have are also substantially decreased. A few related results have been appeared in literature. In [25], an optimal multi-stage defense scheme was proposed by constructing a time-varying attack surface to create the moving target diagram. In [26], the attack identification and isolation problems for CPS were solved using the MTD scheme. In [27], to create a moving defense diagram, an unpredictable switching sequence was designed to activate a controllable pair $(A, \mathcal{B}_l)$. Using such a switching scheme, the attacks can be effectively mitigated. In [27], the physical process is described as a continuous-time model yet the results can be readily extended to the discrete-time counterpart. Nevertheless, there still exist some limitations. First, the attack isolation scheme is absent, which makes it impossible to determine whether the controllable pair $(A, \mathcal{B}_l)$ to be altered is corrupted or not. Second, potential adversaries cannot attack all actuators simultaneously and there always exist available controllable pairs $(A, \mathcal{B}_l)$ to alter. Once the actuators are simultaneously attacked, or the system designer cannot timely recover attacked actuators, the proposed scheme fails to work.

Either the absence of an attack isolation algorithm or no available controllable pairs $(A, \mathcal{B}_l)$ to alter can make the CPS undergo attacks all the time. The system performance can be steadily deteriorated. To solve the problems, this paper proposes attack detection, isolation and MTD control schemes for CPS in a unified framework. The physical process is described as a linear time-invariant discrete-time model [27]. False data can be injected into the cyber layer to deteriorate the system performance [9]. The initial linear time-invariant discrete-time model is converted into a series of controllable sub-systems. An unpredictable switching sequence, which creates the MTD diagram, is given to activate the sub-system. A controller for each activated dynamics is designed. An observer-based attack detection scheme is designed to detect attacks. An attack isolation algorithm is proposed to accurately locate the attacked actuators by designing a series of parallel unknown input observer. After the attack detector reports an alarm, the attack isolation algorithm is invoked to locate and exclude the attacked actuators from the controllable pairs $(A, \mathcal{B}_l)$. A game-based defense controller is designed to deal with the case, in which there exist no available controllable pairs $(A, \mathcal{B}_l)$ that can be used to make the MTD control diagram work. The main contributions of this paper can be summarized as follows

1) The considered attack case in this paper is general. It includes the attack case in [27], in which actuators in each activated dynamics are attacked simultaneously. When part of used actuators are attacked, the control scheme in [27] cannot maintain the desired performance. For actuators in each activated dynamics, either part or all of them are allowed to be attacked. Corresponding defense schemes are designed.
2) If the system dynamics to be altered include attacked actuators, attacks can continuously deteriorate the system performance even the MTD scheme is used. This paper shows that it is necessary to design an attack isolation algorithm when we design a secure control scheme in a MTD control diagram. Using an attack isolation algorithm, we can exclude the attacked actuators from the sequence to be altered.
3) When no available controllable pairs $(A, \mathcal{B}_l)$ can be used to switch after isolating attacked actuators, the system will be exposed to attacks all the time. The system performance can be deteriorated even destroyed. Once such a case happens, a solution to mitigate attacks is provided in this paper.

The rest of the paper is organized as follows. In Section 2., a switching system representation of the physical process and the problem formulation are given. In Section 3., a MTD control scheme is proposed. In Section 4., attack detection and isolation schemes are provided. In Section 5., a proactive/reactive defense control scheme including the reinforcement leaning algorithm is proposed. In Section 6., simulations of a three-tank system under unknown cyber attacks are illustrated. Finally, we conclude this paper in Section 7..

Notations. The notations used throughout the paper are defined as follows. The superscripts "$\top$" and "-1" respectively denote matrix transposition and matrix inverse; $\mathbb{R}^n$ denotes the n-dimensional Euclidean space; the notation $\mathcal{P} > 0$ means that $\mathcal{P}$ is real symmetric and positive definite; diag$\{\cdot\}$ denotes the matrix with diagonal structure; card(a) means the number of the elements in the vector a and card(a) is the number of the columns of a if a is a matrix. pinv$(\cdot)$ means the pseudo inverse.

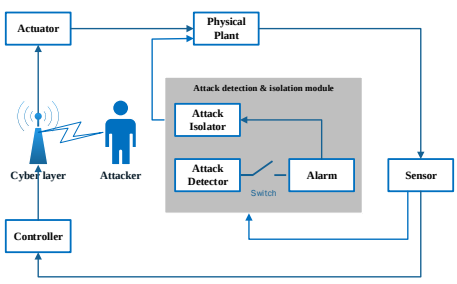

Fig. 1. System blueprint. Here, the "Switch" means that if the residual signal generated by the detector is greater than the predefined threshold, the alarm will be triggered.

2.. System Formulation

Fig. 1 presents the system diagram of interest consisting of the physical plant, sensor, controller, cyber layer, actuator and the attack detection and isolation module. This section gives the physical system and attack models, based on which the physical system under attacks is described. Then, the purpose of this paper is given.

A. Physical system and attack models

The physical system is described as the following linear time-invariant model

$$\begin{aligned} x(k+1) &= Ax(k)+Bu(k), \\ y(k) &= Cx(k), \end{aligned} \tag{1}$$

where $x(k)\in\mathbb{R}^{n_x}$ is the state vector, $u(k)\in\mathbb{R}^{n_u}$ represents the control input. $y(k)\in\mathbb{R}^{n_y}$ denotes the measurement output. A, B and C are compatible matrices. The pair (A,C) is observable.

According to the description of the system (1), we hereby define $u_i(k)$ as the i-th control signal related to the i-th actuator and define $\tilde{B}$ as all possible combinations of the column B_j in the matrix B. j takes values in the set $\{1,\ldots,n_u\}$. $\mathcal{B}_l$ means each specific combination and l takes values in the set $\{1,\ldots,2^{n_u}\}$.

To create the MTD scheme, not all actuators are used to stabilize the physical plant. The system dynamics governed by controllable pairs $(A,\mathcal{B}_l)$ will be altered following a specific rule. To distinguish the controllable and uncontrollable pairs, define the following two sets

$$\begin{aligned} \tilde{B}_1 &= \left\{\mathcal{B}_l\in\tilde{B}:\text{rank}\left(\left[\mathcal{B}_l,\ A\mathcal{B}_l,\ldots,A^{n_x-1}\mathcal{B}_l\right]\right)=n_x\right\}, \\ \tilde{B}_2 &= \left\{\mathcal{B}_l\in\tilde{B}:\text{rank}\left(\left[\mathcal{B}_l,\ A\mathcal{B}_l,\ldots,A^{n_x-1}\mathcal{B}_l\right]\right)<n_x\right\}, \end{aligned}$$

where $\tilde{B}_1$ denotes the actuating mode sets which can stabilize the system and $\tilde{B}_2$ denotes the actuating mode sets which cannot stabilize the system.

As shown in Fig. 1, the cyber layer, which is prone to malicious behaviors, is utilized to transmit the control signal to the actuator. It is assumed that the adversary imposes false data injection attacks on the occupied communication channels. Since the system dynamics will be altered to defend attacks, the communication channels will be also changed. For each controllable pair $(A,\mathcal{B}_l)$, corresponding communication channels will be employed to transmit the control signals. Once the adversary successfully intrudes the cyber layer, the physical system (1) is described as

$$x(k+1) = Ax(k)+Bu(k)+B\Gamma(k)u_a(k), \tag{2}$$

where $\Gamma(k)$ is a time-varying diagonal matrix, which is defined to describe which actuators are attacked. $u_a(k)$ is the malicious signal designed by the adversary. If the i-th actuator is attacked, the i-th diagonal element in $\Gamma(k)$ is 1, otherwise it is 0. Define $\mathbb{S}=\text{ind}(\Gamma(k))$ as the attacked actuator index set.

Remark 1: As to $\Gamma(k)$, its diagonal elements consist of 0 and 1. If the i-th actuator is attacked, $\Gamma_{ll}(k)=1$, otherwise $\Gamma_{ll}(k)=0$ $(l=1,2\ldots,n_u)$. Additionally, there doesn't exist any constraints on $\|\Gamma(k)\|_{l_0}$ in this paper. The adversary can attack partial or all actuators, that is, $0\le\|\Gamma(k)\|_{l_0}\le n_u$.

B. Physical system: a switching system representation

The MTD technique is used to design a secure control scheme in this paper. If the system operator detects the attack, it will isolate the attacked actuators and take them offline. To create the MTD diagram, the system operator alters the dynamics by switching to the controllable actuating modes without attacks. If the remaining actuating modes are uncontrollable, a zero-sum game based control scheme will be adopted to defend the attacks and stabilize the system. According to the above description, we can describe the system (1) in the following two cases.

Case 1 Except the current actuating mode, there exist controllable actuating modes without attacks in $\tilde{B}_1$

$$x(k+1) = Ax(k)+\mathcal{B}_l\bar{u}_l(k), \tag{3}$$

where $\bar{u}_l(k)$ consists of control signals without attacks, $l\in\{1,\ldots,\text{card}(\tilde{B}_1)\}$.

Case 2 Except the current actuating mode, the remaining actuating modes without attacks belong to $\tilde{B}_2$

$$x(k+1) = Ax(k)+\mathcal{B}_l\bar{u}_l(k)+\mathcal{B}_l\tilde{u}_{a,l}(k), \tag{4}$$

where $\tilde{u}_{a,l}(k)$ is the malicious signal, $l\in\{1,\ldots,\text{card}(\tilde{B}_1)\}$.

The interest of this paper is to propose a secure control algorithm to preserve the stability under attacks. The algorithm to be proposed includes an attack detection, isolation scheme, proactive defense controller based on the moving target defense diagram and a zero-sum game learning based on the reactive defense control scheme. Using such an algorithm, the attack can be utmostly mitigated and the advantages the adversary has over the defender can be decreased.

3.. MTD Controller Design and Stability Analysis

This section mainly focuses on how to design a controller for each altered dynamics, design the MTD to alter system dynamics, and analyze the stability of the overall system. As to the controller design, a linear quadratic optimal control strategy is derived for each altered dynamics. For the MTD scheme, it is designed as a switching sequence, the elements of which mean each dynamics to be altered. Then, the stability is analyzed by employing the average dwell time approach.

A. Controller and MTD design

When the system dynamics are altered based on the controllable combination set $\tilde{B}_1$, design the following performance index for system (3)

$$V_l\left(x(k)\right) = \min_{\bar{u}_l(i)}\left[\sum_{i=k}^{\infty}x^{\top}(i)\,Q_l\,x(i)+\bar{u}_l^{\top}(i)\,R_l\bar{u}_l(i)\right] \tag{5}$$

where $Q_l\ge 0$, $R_l>0$ and $l\in\{1,\ldots,\text{card}(\tilde{B}_1)\}$.

Based on the results [28], the following lemma is given to design the optimal control signal $\bar{u}_l(k)$.

Lemma 1: [28] For the system (3), the optimal control gain $\bar{u}_l(k)$ can be designed as $\bar{u}_l(k) = K_l x(k)$, where $K_l = -\left(R_l + \mathcal{B}_l^\top P_l \mathcal{B}_l\right)^{-1} \mathcal{B}_l^\top P_l A$ and P_l is the solution of the following Riccati equation

$$P_l = A^\top P_l A + Q_l - A^\top P_l \mathcal{B}_l \left(R_l + \mathcal{B}_l^\top P_l \mathcal{B}_l\right)^{-1} \mathcal{B}_l^\top P_l A.$$

Remark 2: Although the dynamic programming is used to derive linear quadratic optimal controller for each altered dynamics, the controller given in Lemma 1 is only optimal for the current activated dynamics.

In contrast to results reviewed in previous sections, the communication pattern in this paper is dynamical. When the controller is altered, the communication pattern is changed. Thus, the core of MTD is to devise a stochastic switching rule to unpredictably switch the controller. In this way, attacks can be mitigated. Next, a lemma is given to determine the probabilities of activating the controller K_l.

Lemma 2: [27] For the candidate controller K_l, the probability p_l to activate K_l satisfies $p_l = \exp^{-\frac{V_l^*}{\delta} - 1 + \log\left(\exp \sum_{l=1}^{\text{card}(\tilde{B}_1)} \exp^{\frac{V_l^*}{\delta}}\right)}$, where $V_l^* = x^\top(0) P_l x(0)$ with P_l solved from the Riccati equation in Lemma 1 and $\delta > 0$ is a weighting coefficient.

B. Stability analysis

Based on the probability given in Lemma 2, each candidate controller can be activated. The closed-loop system can be regarded as a switched system. To facilitate the stability analysis, each activated system dynamics will run satisfying the least period. Next, the average dwell time definition is introduced to derive the conditions to preserve stability.

Definition 1: [29] For switching signal α_k and any $k_i > k_j > k_0$, define $N_{\alpha_k}(k_j, k_i)$ as the switching numbers of α_k over the interval $[k_j, k_i]$. If for any given $N_0 \geq 0$ and $\tau > 0$, we have $N_{\alpha_k}(k_j, k_i) \leq N_0 + (k_i - k_j)/\tau$, then τ and N_0 are called average dwell time and the chatter bound, respectively.

Based on Definition 1, the following theorem is proposed to show that the stability of the system can be preserved under the specific condition.

Theorem 1: The overall system can be stabilized provided that the average dwell time τ satisfies the inequality $\tau > \text{ceil}\left(-\frac{\ln \mu}{\ln(1-\beta)}\right)$, where

$$\beta = \min_{l \in \{1,2,\ldots,\text{card}(\tilde{B}_1)\}} \frac{\lambda_{\min}(\tilde{Q}_l)}{\lambda_{\max}(P_l)},$$

$$\mu = \max_{l,q \in \{1,2,\ldots,\text{card}(\tilde{B}_1)\}} \frac{\lambda_{\max}(P_l)}{\lambda_{\min}(P_q)}.$$

Proof: The proof is given in Section 8.-A of Appendix. ■

4.. Attack Detection and Isolation

To utmost mitigate the attack, both detection and isolation of attacks are designed to exclude the controllable system dynamics $(A, \mathcal{B}_l)$ under attacks from the MTD sequence. When attacks are detected, the isolation scheme is activated to accurately locate the attacked actuators. Next, we start with designing an attack detection scheme, following which an attack isolation algorithm is provided.

A. Attack detection observer design

First, an attack detection observer is designed to determine whether the actuators are attacked or not. The following attack detection observer is designed

$$\begin{aligned} \hat{x}(k+1) &= A\hat{x}(k) + \mathcal{B}_l \bar{u}_l(k) + \tilde{L}_l\left(y(k) - \hat{y}(k)\right), \\ \hat{y}(k) &= C\hat{x}(k), \end{aligned} \tag{6}$$

where $\hat{x}(k)$ is the estimate of $x(k)$ and $\hat{y}(k)$ is the estimated measurement. $\tilde{L}_l$ denotes the observer gain to be designed. It is designed to ensure $\lim_{k\to\infty}(x(k) - \hat{x}(k)) = 0$ under the attack-free case.

Define $e_x(k) = x(k) - \hat{x}(k)$ as the estimate error. According to the system (3) and the observer (6), the following error system can be obtained

$$e_x(k+1) = \left(A - \tilde{L}_l C\right) e_x(k) + \mathcal{B}_l \tilde{u}_{a,l}(k), \tag{7}$$

The observer gain $\tilde{L}_l$ can be given such that all the eigenvalues of $A - \tilde{L}_l C$ locate in the unit circle, which in turn shows that $\lim_{k\to\infty}(x(k) - \hat{x}(k)) = 0$ holds without attacks.

To determine whether attacks happen or not, define $e_y(k) = Ce_x(k)$ as the residual signal. Then the residual evaluation function R_{resi} is designed as $R_{resi} = \sum_{k=t_1}^{t_2} \|e_y^\top(k) e_y(k)\|_2$.

Based on the designed residual evaluation function R_{resi}, the following attack detection decision can be made

$$\begin{cases} R_{resi} \leq \mathcal{V}, & \text{no attacks}, \\ R_{resi} > \mathcal{V}, & \text{alarm}, \end{cases}$$

where $\mathcal{V}$ is a threshold to be designed.

According to the error system in (7), we know that the system state $x(k)$ can be effectively estimated under attack-free case. Once attacks are implemented, the estimates $\hat{x}(k)$ will deviate the real state $x(k)$. Accordingly, the threshold $\mathcal{V}$ is defined as $\mathcal{V} = \sup_{\tilde{u}_{a,l}=0} e_x(k)$.

B. Attack isolation scheme

Once the detection scheme reports an alarm, it is vital to determine which actuators are attacked and switch to the healthy actuators. In this subsection, an attack isolation scheme is proposed to locate the attacked actuators, which provides access to taking the attacked actuators offline.

For the controllable pair $(A, \mathcal{B}_l)$ under malicious behaviors, the system dynamics are described as

$$x(k+1) = Ax(k) + \mathcal{B}_l \bar{u}_l(k) + \mathcal{B}_l \tilde{u}_{a,l}(k). \tag{8}$$

To remove the effect resulting from the attacks, we treat the signal $\tilde{u}_{a,l}(k)$ as an unknown input. Then, the following unknown input observer is designed

$$\begin{aligned} z_l(k+1) &= \mathcal{E}_l z_r(k) + \mathcal{F}_l \mathcal{B}_l \bar{u}_l(k) + \mathcal{L}_l y(k), \\ \hat{\tilde{x}}_l(k) &= z_l(k) + \mathcal{G}_l y(k), \end{aligned} \tag{9}$$

where $\hat{\tilde{x}}_l(k)$ is the state estimate of the l-th unknown input observer. The matrices $\mathcal{E}_l$, $\mathcal{F}_l$, $\mathcal{L}_l$ and $\mathcal{G}_l$ are given to make $\lim_{k\to\infty}\left(x(k) - \hat{\tilde{x}}_l(k)\right) = 0$ and all eigenvalues of the matrices $\mathcal{E}_l$ should locate in the unit circle.

Define $\bar{e}_l(k) = x(k) - \hat{\bar{x}}_l(k)$ as the estimate error. The matrices $\mathcal{E}_l$, $\mathcal{F}_l$, $\mathcal{L}_l$ and $\mathcal{G}_l$ should satisfy the following equations

$$\begin{aligned}(I - \mathcal{G}_l C) A + \mathcal{E}_l (\mathcal{G}_l C - I) - \mathcal{L}_l C &= 0,\\ (\mathcal{F}_l - (I - \mathcal{G}_l C)) \mathcal{B}_l = 0, \ \ (I - \mathcal{G}_l C) \mathcal{B}_l &= 0.\end{aligned}$$

The following lemma is given to guarantee that the unknown input observer exists

Lemma 3: [30] The unknown input observer in (9) exists provided that the condition $\text{rank}(C\mathcal{B}_l) = \text{rank}(\mathcal{B}_l) = \text{card}(\mathcal{B}_l)$ holds and the pair $(A - \mathcal{G}_l CA, C)$ is detectable.

For the estimate error dynamics $\bar{e}_l(k)$, it is derived as

$$\begin{aligned}\bar{e}_l(k+1) &= (A - \mathcal{G}_l CA - \mathcal{L}_l C + \mathcal{E}_l \mathcal{G}_l C) x(k)\\ &\quad + (\mathcal{B}_l - \mathcal{F}_l \mathcal{B}_l - \mathcal{G}_l C\mathcal{B}_l) \bar{u}_l(k)\\ &\quad + (\mathcal{B}_l - \mathcal{G}_l C\mathcal{B}_l) \tilde{u}_{a,l}(k) - \mathcal{E}_l \hat{\bar{x}}_l(k)\\ &= \mathcal{E}_l \bar{e}_l(k). \end{aligned} \tag{10}$$

As can be seen from (10), the estimate error $\bar{e}_l(k)$ can exponentially converge. After obtaining the estimates $\hat{\bar{x}}_l(k)$, we can use them to exponentially reconstruct the attack signal. Then which actuators are attacked can be identified by observing the reconstructed attack signals.

Combining the dynamics in (8) and the definition of $\bar{e}_l(k)$ yields

$$\begin{aligned}\hat{\bar{x}}_l(k+1) &= A\left(\hat{\bar{x}}_l(k) + \bar{e}_l(k)\right) + \mathcal{B}_l \bar{u}_l(k)\\ &\quad + \mathcal{B}_l \tilde{u}_{a,l}(k) - \bar{e}_l(k+1),\end{aligned}$$

which implies

$$\begin{aligned}\tilde{u}_{a,l}(k) &= \text{pinv}(\mathcal{B}_l)\left(\hat{\bar{x}}_l(k+1) - A\hat{\bar{x}}_l(k)\right)\\ &\quad + \text{pinv}(\mathcal{B}_l)\left(\bar{e}_l(k+1) - A\bar{e}_l(k)\right) - \bar{u}_l(k).\end{aligned}$$

Therefore, we can reconstruct the attack signal as

$$\hat{\tilde{u}}_{a,l}(k) = \text{pinv}(\mathcal{B}_l)\left(\hat{\bar{x}}_l(k+1) - A\hat{\bar{x}}_l(k)\right) - \bar{u}_l(k),$$

where $\hat{\tilde{u}}_{a,l}(k)$ means the estimate of $\tilde{u}_{a,l}(k)$.

For a sufficient large time window, i.e., k is sufficiently large, we can obtain $\tilde{u}_{a,l}(k-1) \to \hat{\tilde{u}}_{a,l}(k)$. Then we can identify which actuators are attacked using the sparsity of the vector $\hat{\tilde{u}}_{a,l}(k)$. Then, $\hat{l}$, the index set of the attacked actuators can be derived as

$$\hat{l} = \{\hat{l} = j | \hat{\tilde{u}}_{a,l,j}(k) \neq 0\}, \tag{11}$$

where $\hat{\tilde{u}}_{a,l,j}(k)$ means the j-th element of $\hat{\tilde{u}}_{a,l}(k)$.

Although we can use (11) to isolate the attacked actuators, the estimate of the attack at time k needs the data at time $k+1$, and $\mathcal{B}_l$ should be column full rank. To remove such restrictions, the following attack isolation scheme with a family of parallel unknown input observers is proposed.

To clearly describe how to design the attack isolation mechanism, define the set $\bar{l}_{1,\eta} \subset \{1, 2, \ldots, \text{card}(\mathcal{B}_l)\}$ and $\bar{l}_{2,\eta}$ as the complementary set of $\bar{l}_{1,\eta}$ in $\{1, 2, \ldots, \text{card}(\mathcal{B}_l)\}$. η takes value from 1 to $2^{\text{card}(\mathcal{B}_l)} - 1$[1].

[1]Since the isolation scheme is activated after the attack detection mechanism reports an alarm, $\bar{l}_{1,\eta} = \emptyset$ is excluded.

First, rewrite the dynamics in (8) as follows

$$\begin{aligned}x(k+1) &= Ax(k) + \mathcal{B}_l \bar{u}_l(k)\\ &\quad + \mathcal{B}_{\bar{l}_{1,\eta}} \tilde{u}_{a,\bar{l}_{1,\eta}}(k) + \mathcal{B}_{\bar{l}_{2,\eta}} \tilde{u}_{a,\bar{l}_{2,\eta}}(k),\end{aligned} \tag{12}$$

where $\mathcal{B}_{\bar{l}_{1,\eta}}$ consists of column vectors in the matrices $\mathcal{B}_l$ indexed by the set $\bar{l}_{1,\eta}$, $\mathcal{B}_{\bar{l}_{2,\eta}}$ mean the matrices which remove column vectors from the matrices $\mathcal{B}_l$ indexed by the set $\bar{l}_{1,\eta}$, $\tilde{u}_{a,\bar{l}_{1,\eta}}(k)$ consists of rows in $\tilde{u}_{a,l}$ indexed by the set $\bar{l}_{1,\eta}$ and $\tilde{u}_{a,\bar{l}_{2,\eta}}(k)$ mean the actuator signals which remove the rows from $\tilde{u}_{a,l}$ indexed by the set $\bar{l}_{1,\eta}$.

Regard the item $\tilde{u}_{a,\bar{l}_{1,\eta}}(k)$ as the unknown input and define $\bar{e}_{l,\bar{l}_{1,\eta}}(k) = x(k) - \hat{\bar{x}}_{l,\bar{l}_{1,\eta}}(k)$ as the estimate error. If the following conditions are satisfied

$$\begin{aligned}\left(I - \mathcal{G}_{l,\bar{l}_{1,\eta}} C\right) A + \mathcal{E}_{l,\bar{l}_{1,\eta}} \left(\mathcal{G}_{l,\bar{l}_{1,\eta}} C - I\right) - \mathcal{L}_{l,\bar{l}_{1,\eta}} C &= 0,\\ \left(\mathcal{F}_{l,\bar{l}_{1,\eta}} - \left(I - \mathcal{G}_{l,\bar{l}_{1,\eta}} C\right)\right) \mathcal{B}_l = 0, \ \ \left(I - \mathcal{G}_{l,\bar{l}_{1,\eta}} C\right) \mathcal{B}_{l,\bar{l}_{1,\eta}} &= 0,\end{aligned}$$

we can construct a series of unknown input observers for (12) as

$$\begin{aligned}z_{l,\bar{l}_{1,\eta}}(k+1) &= \mathcal{E}_{l,\bar{l}_{1,\eta}} z_{l,\bar{l}_{1,\eta}}(k) + \mathcal{F}_{l,\bar{l}_{1,\eta}} \mathcal{B}_l \bar{u}_l(k)\\ &\quad + \mathcal{L}_{l,\bar{l}_{1,\eta}} y(k),\\ \hat{\bar{x}}_{l,\bar{l}_{1,\eta}}(k) &= z_{l,\bar{l}_{1,\eta}}(k) + \mathcal{G}_{l,\bar{l}_{1,\eta}} y(k),\end{aligned} \tag{13}$$

where $\hat{\bar{x}}_{l,\bar{l}_{1,\eta}}(k)$ is the state estimate. The matrices $\mathcal{E}_{l,\bar{l}_{1,\eta}}$, $\mathcal{F}_{l,\bar{l}_{1,\eta}}$, $\mathcal{L}_{l,\bar{l}_{1,\eta}}$ and $\mathcal{G}_{l,\bar{l}_{1,\eta}}$ are designed to ensure $\lim_{k\to\infty}\left(x(k) - \hat{\bar{x}}_{l,\bar{l}_{1,\eta}}(k)\right) = 0$ when all the attacked actuators can be included in the matrices $\mathcal{B}_{l,\bar{l}_{1,\eta}}$.

For such a bank of unknown input observers, the error dynamics can be described as

$$\begin{aligned}\bar{e}_{l,\bar{l}_{1,\eta}}(k+1) &= \mathcal{E}_{l,\bar{l}_{1,\eta}} \bar{e}_{l,\bar{l}_{1,\eta}}(k) + \left(\mathcal{B}_{l,\bar{l}_{2,\eta}} - \mathcal{G}_{l,\bar{l}_{2,\eta}} C\mathcal{B}_{l,\bar{l}_{2,\eta}}\right)\\ &\quad \times \tilde{u}_{a,l,\bar{l}_{2,\eta}}(k).\end{aligned} \tag{14}$$

A series of designed unknown input observers in (13) can guarantee the errors $\bar{e}_{l,\bar{l}_{1,\eta}}(k)$ exponentially converge if all indices of attacked actuators are included in the set $\bar{l}_{1,\eta}$. Using such a conclusion, the attack isolation scheme can be designed.

For the activated model l, there may exist several models $\mathcal{B}_{l,\bar{l}_{1,\eta}}$ which include all attacked actuators. To accurately locate the attacked actuators, define the set

$$\begin{aligned}\tilde{l}_m &= \{\tilde{l}_m = \bar{l}_{1,\eta} | \bar{l}_{1,\eta} : \bar{e}_{l,\bar{l}_{1,\eta}}(k) \leq \bar{e}_{l,\min}(k),\\ &\quad \bar{l}_{1,\eta} \subset \{1, 2, \ldots, \text{card}(\mathcal{B}_l)\}\},\end{aligned}$$

where $\bar{e}_{l,\min}(k)$ is regarded as a threshold, which is calculated without attacks.

Then, $\hat{l}$, the index set of attacked actuators can be obtained as $\hat{l} = \tilde{l}_1 \cap \tilde{l}_2 \cap \ldots \cap \tilde{l}_m$. To show the effectiveness of the designed parallel unknown input observer based isolation scheme, a numerical example is provided.

Example 1: For the altered dynamics, the system parameters are assumed to be

$$A = \begin{bmatrix} 0.9 & 0.4 & -0.8 \\ 0.2 & -1 & -0.5 \\ -0.9 & -0.9 & -0.2 \end{bmatrix}, \ B = \begin{bmatrix} 0.1 & 0.2 & 0.8 \\ 1.5 & 1 & 1.3 \\ 1 & 0.5 & 0.7 \end{bmatrix},$$
$$C = \text{diag}\{1, 1, 1\}.$$

It is assumed that the adversary implements attacks from $k = 30$ and the target nodes are the actuators 1 and 3. Accordingly, based on the theoretical analysis, only when $\bar{l}_{1,4} = \{1,3\}$ and $\bar{l}_{1,7} = \{1,2,3\}$, can the estimate error $\bar{e}_{l,\bar{l}_{1,\eta}}(k)$ be zero. Fig. 2 depicts the isolation results, with which we can see that the estimate errors can converge to be zero when $\bar{l}_{1,4} = \{1,3\}$ and $\bar{l}_{1,7} = \{1,2,3\}$. That is, the attacked actuator is $\hat{l} = \bar{l}_{1,\eta} \bigcap \bar{l}_{1,\eta} = \{1,3\}$.

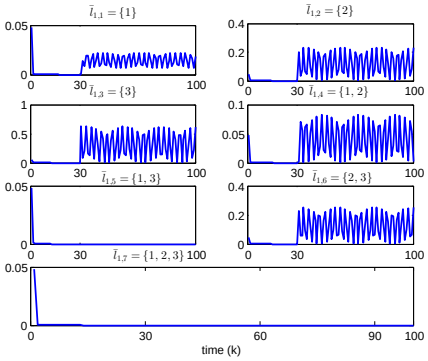

Fig. 2. Attack isolation using the parallel unknown input observers.

5.. SECURE CONTROL ALGORITHM DESIGN

Combining the previous schemes, the proactive reactive defense control algorithm is designed in this section. It is noted that the MTD scheme is feasible with the premise that the remaining sub-models are controllable. Therefore, it is full of importance to find a solution when the premise is not satisfied. To this end, a reactive defense controller is proposed to provide access to mitigating the attack and guaranteeing the stability of the system before giving the secure control algorithm.

A. Reinforcement learning based reactive control scheme

This subsection mainly utilizes the zero-sum game approach and reinforcement learning to design the reactive defense control scheme. Specifically, in such a game, both the adversary and the defender are regarded as two players. The defender's objective is to stabilize the overall system yet the adversary intends to deteriorate the system performance. After deriving the relative closed-form solution of the controller, reinforcement learning is employed to obtain the control gain. In the following, the detailed derivations are provided.

Using the designed attack isolation scheme, the attacked actuators can be located. Thus, the system under actuator attacks can be described as

$$x(k+1) = Ax(k) + \mathcal{B}_l \bar{u}_l(k) + \mathcal{B}_{l,\hat{l}} \tilde{u}_{a,l,\hat{l}}, \tag{15}$$

where $\mathcal{B}_{l,\hat{l}}$ is the attack distribution matrix and $\tilde{u}_{a,l,\hat{l}}$ represents the attack signal.

According to the objectives of both sides, the control problem can be described as a zero-sum game and the value function is given as

$$V_l(x(k)) = \min_{\bar{u}_l(i)} \max_{\tilde{u}_{a,l}(i)} \left[\sum_{i=k}^{\infty} \varphi(i) \right], \tag{16}$$

where $\varphi(i) = x^\top(i) Q_l x(i) + \tilde{u}_l^\top(i) R_{1,l} \tilde{u}_l(i) - \gamma^2 \tilde{u}_{a,l,\hat{l}}^\top(i) R_{2,l} \tilde{u}_{a,l,\hat{l}}(i)$. $\gamma > 0$ can be regarded as an attack rejection index.

Then, the Bellman equation can be written as

$$V_l(x(k)) = \varphi(k) + V_l(x(k+1)). \tag{17}$$

For such a zero-sum game problem, the objective is to find an optimal control scheme and an optimal attack strategy, with which the game can achieve a saddle-point equilibrium and the optimal value function is $V_l^{opt}(x(k)) = x^\top(0) P_l x(0)$. Then, the following lemma is provided to show how to design the optimal schemes for the defender and attacker.

Lemma 4: [31] The system (15) can be stabilized with the following schemes and both players can achieve the saddle-point equilibrium

$$\bar{u}_l(k) = K_l x(k), \tag{18}$$

$$\tilde{u}_{a,l,\hat{l}}(k) = L_{l,\hat{l}} x(k), \tag{19}$$

where

$$\begin{aligned} K_l = & \left(R_{1,l} + \mathcal{B}_l^\top P_l \mathcal{B}_l - \mathcal{B}_l^\top P_l \mathcal{B}_{l,\hat{l}} \left(\mathcal{B}_{l,\hat{l}}^\top P_l \mathcal{B}_{l,\hat{l}} - \gamma^2 R_{2,l} \right)^{-1} \right. \\ & \times \left. \mathcal{B}_{l,\hat{l}}^\top P_l \mathcal{B}_l \right)^{-1} \left(\mathcal{B}_l^\top P_l \mathcal{B}_{l,\hat{l}} \left(\mathcal{B}_{l,\hat{l}}^\top P_l \mathcal{B}_{l,\hat{l}} - \gamma^2 R_{2,l} \right)^{-1} \right. \\ & \times \left. \mathcal{B}_{l,\hat{l}}^\top P_l A - \mathcal{B}_l^\top P_l A \right), \\ L_{l,\hat{l}} = & \left(\mathcal{B}_{l,\hat{l}}^\top P_l \mathcal{B}_{l,\hat{l}} - \gamma^2 R_{2,l} - \mathcal{B}_{l,\hat{l}}^\top P_l \mathcal{B}_l \left(R_{1,l} + \mathcal{B}_l^\top P_l \mathcal{B}_l \right)^{-1} \right. \\ & \times \left. \mathcal{B}_l^\top P_l \mathcal{B}_{l,\hat{l}} \right)^{-1} \left(\mathcal{B}_{l,\hat{l}}^\top P_l \mathcal{B}_l \left(R_{1,l} + \mathcal{B}_l^\top P_l \mathcal{B}_l \right)^{-1} \right. \\ & \times \left. \mathcal{B}_l^\top P_l A - \mathcal{B}_{l,\hat{l}}^\top P_l A \right), \end{aligned}$$

and P_l is the solution of the following Riccati equation

$$\begin{aligned} P_l = & A^\top P_l A + Q_l - \begin{bmatrix} A^\top P_l \mathcal{B}_l & A^\top P_l \mathcal{B}_l \end{bmatrix} \\ & \times \begin{bmatrix} R_{1,l} + \mathcal{B}_l^\top P_l \mathcal{B}_l & \mathcal{B}_l^\top P_l \mathcal{B}_l \\ \mathcal{B}_l^\top P_l \mathcal{B}_l & \mathcal{B}_l^\top P_l \mathcal{B}_l - \gamma^2 R_{2,l} \end{bmatrix}^{-1} \\ & \times \begin{bmatrix} \mathcal{B}_l^\top P_l A \\ \mathcal{B}_l^\top P_l A \end{bmatrix}. \end{aligned} \tag{20}$$

To ensure the existence of the saddle point, the inequalities $\gamma^2 R_{2,l} - \mathcal{B}_l^\top P_l \mathcal{B}_l > 0$ and $R_{1,l} + \mathcal{B}_l^\top P_l \mathcal{B}_l > 0$ should hold.

According to the above discussion, the reinforcement learning based control scheme is designed in the following content. First, rewrite the system in (15) as

$$\begin{aligned} x(k+1) = & \tilde{A}_i x(k) + \mathcal{B}_l \left(\bar{u}_l(k) - K_l^i x(k) \right) \\ & + \mathcal{B}_{l,\hat{l}} \left(\tilde{u}_{a,l,\hat{l}}(k) - L_{l,\hat{l}}^i x(k) \right), \end{aligned} \tag{21}$$

where $\tilde{A}_i = A + \mathcal{B}_l K_l^i + \mathcal{B}_{l,\hat{l}} L_{l,\hat{l}}^i$ and i means each learning step.

In (21), $\bar{u}_l^i(k) = K_l^i x(k)$ and $\tilde{u}_{a,l,\hat{l}}^i(k) = L_{l,\hat{l}}^i x(k)$ are target policies to be learned and updated. $\bar{u}_l(k)$ and $\tilde{u}_{a,l,\hat{l}}(k)$ are behavior policies, which are applied to the system (21) to generate data to learn and update the policies $\bar{u}_l^i(k)$ and $\tilde{u}_{a,l,\hat{l}}^i(k)$. Then, based on the learned policies $\bar{u}_l^i(k)$ and $\tilde{u}_{a,l,\hat{l}}^i(k)$, the Bellman equation in (17) can be described as

$$V_l^{i+1}(x(k)) - V_l^{i+1}(x(k+1)) = \varphi_i(k), \tag{22}$$

where $\varphi_i(k) = x^\top(k)\,Q_l x(k) + \tilde{u}_l^{i^\top}(k)\,R_{1,l}\tilde{u}_l^i(k) - \gamma^2\tilde{u}_{a,l,\hat{l}}^{i^\top}(k)\,R_{2,l}\tilde{u}_{a,l,\hat{l}}(k)$.

Using Taylor's theorem to expand $V_l(x(k))$ at the point $x(k+1)$ yields

$$\begin{aligned} V_l(x(k)) &= V_l(x(k+1)) + 2x^\top(k+1)P_l(x(k)-x(k+1)) \\ &\quad + (x(k)-x(k+1))^\top P_l(x(k)-x(k+1)), \end{aligned}$$

which implies

$$\begin{aligned} & V_l^{i+1}(x(k)) - V_l^{i+1}(x(k+1)) \\ =\ & 2x^\top(k+1)P_l^{i+1}(x(k)-x(k+1)) \\ & + (x(k)-x(k+1))^\top P_l^{i+1}(x(k)-x(k+1)) \\ =\ & -x^\top(k)\tilde{A}_i^\top P_l^{i+1}\tilde{A}_i x(k) + x^\top(k)P_l^{i+1}x(k) \\ & - \left(\bar{u}_l(k)-K_l^i x(k)\right)^\top \mathcal{B}_l^\top P_l^{i+1}x(k+1) \\ & - \left(\bar{u}_l(k)-K_l^i x(k)\right)^\top \mathcal{B}_l^\top P_l^{i+1}\tilde{A}_i x(k) \\ & - \left(\tilde{u}_{a,l,\hat{l}}(k)-L_{l,\hat{l}}^i x(k)\right)^\top \mathcal{B}_{l,\hat{l}}^\top P_l^{i+1}x(k+1) \\ & - \left(\tilde{u}_{a,l,\hat{l}}(k)-L_{l,\hat{l}}^i x(k)\right)^\top \mathcal{B}_{l,\hat{l}}^\top P_l^{i+1}\tilde{A}_i x(k). \end{aligned} \tag{23}$$

Based on the equation in (22), we can obtain

$$Q - P_l^{i+1} + K_l^{i^\top}R_{1,l}K_l^i - \gamma^2 L_{l,\hat{l}}^{i^\top}R_{2,l}L_{l,\hat{l}}^i + \tilde{A}_i^\top P_l^{i+1}\tilde{A}_i = 0,$$

combining which and (23) yields

$$\begin{aligned} & V_l^{i+1}(x(k)) - V_l^{i+1}(x(k+1)) \\ =\ & x^\top(k)Qx(k) + x^\top(k)K_l^{i^\top}R_{1,l}K_l^i x(k) \\ & -\gamma^2 x^\top(k)L_l^{i^\top}R_{2,l}L_l^i x(k) \\ & - \left(\bar{u}_l(k)-K_l^i x(k)\right)^\top \mathcal{B}_l^\top P_l^{i+1}x(k+1) \\ & - \left(\bar{u}_l(k)-K_l^i x(k)\right)^\top \mathcal{B}_l^\top P_l^{i+1}\tilde{A}_i x(k) \\ & - \left(\tilde{u}_{a,l,\hat{l}}(k)-L_{l,\hat{l}}^i x(k)\right)^\top \mathcal{B}_{l,\hat{l}}^\top P_l^{i+1}x(k+1) \\ & - \left(\tilde{u}_{a,l,\hat{l}}(k)-L_{l,\hat{l}}^i x(k)\right)^\top \mathcal{B}_{l,\hat{l}}^\top P_l^{i+1}\tilde{A}_i x(k). \end{aligned} \tag{24}$$

Performing some mathematical operations to (24) yields

$$\begin{aligned} & x^\top(k)P_l^{i+1}x(k) - x^\top(k+1)P_l^{i+1}x(k+1) \\ & +2\left(\bar{u}_l(k)-K_l^i x(k)\right)^\top \mathcal{B}_l^\top P_l^{i+1}Ax(k) \\ & +2\left(\tilde{u}_{a,l,\hat{l}}(k)-L_{l,\hat{l}}^i x(k)\right)^\top \mathcal{B}_{l,\hat{l}}^\top P_l^{i+1}Ax(k) \\ & + \left(\bar{u}_l(k)-K_l^i x(k)\right)^\top \mathcal{B}_l^\top P_l^{i+1}\mathcal{B}_l \\ & \times \left(\bar{u}_l(k)+K_l^i x(k)\right) + \left(\tilde{u}_{a,l,\hat{l}}(k)-L_{l,\hat{l}}^i x(k)\right)^\top \\ & \times \mathcal{B}_{l,\hat{l}}^\top P_l^{i+1}\mathcal{B}_{l,\hat{l}}\left(\tilde{u}_{a,l,\hat{l}}(k)+L_{l,\hat{l}}^i x(k)\right) \\ & + \left(\bar{u}_l(k)-K_l^i x(k)\right)^\top \mathcal{B}_l^\top P_l^{i+1}\mathcal{B}_{l,\hat{l}} \\ & \times \left(\tilde{u}_{a,l,\hat{l}}(k)+L_{l,\hat{l}}^i x(k)\right) \\ & + \left(\tilde{u}_{a,l,\hat{l}}(k)-L_{l,\hat{l}}^i x(k)\right)^\top \mathcal{B}_{l,\hat{l}}^\top P_l^{i+1}\mathcal{B}_l \\ & \times \left(\bar{u}_l(k)+K_l^i x(k)\right) \\ =\ & x^\top(k)Qx(k) + x^\top(k)K_l^{i^\top}R_{1,l}K_l^i x(k) \\ & -\gamma^2 x^\top(k)L_{l,\hat{l}}^{i^\top}R_{2,l}L_{l,\hat{l}}^i x(k). \end{aligned} \tag{25}$$

Using the Kronecker product, (25) is rewritten as

$$\begin{aligned} & \left(x^\top(k)\otimes x^\top(k) - x^\top(k+1)\otimes x^\top(k+1)\right)\mathrm{vec}(P_l^{i+1}) \\ & +2\left(\left(\bar{u}_l(k)-K_l^i x(k)\right)^\top \otimes x^\top(k)\right)\mathrm{vec}(\mathcal{B}_l^\top P_l^{i+1}A) \\ & +2\left(\left(\tilde{u}_{a,l,\hat{l}}(k)-L_{l,\hat{l}}^i x(k)\right)^\top \otimes x^\top(k)\right)\mathrm{vec}(\mathcal{B}_{l,\hat{l}}^\top P_l^{i+1}A) \\ & + \left(\left(\bar{u}_l(k)-K_l^i x(k)\right)^\top \otimes \left(\bar{u}_l(k)+K_l^i x(k)\right)^\top\right) \\ & \times \mathrm{vec}(\mathcal{B}_l^\top P_l^{i+1}\mathcal{B}_l) \\ & + \left(\left(\tilde{u}_{a,l,\hat{l}}(k)-L_{l,\hat{l}}^i x(k)\right)^\top \otimes \left(\tilde{u}_{a,l,\hat{l}}(k)+L_{l,\hat{l}}^i x(k)\right)^\top\right) \\ & \times \mathrm{vec}(\mathcal{B}_{l,\hat{l}}^\top P_l^{i+1}\mathcal{B}_{l,\hat{l}}) \\ & + \left(\left(\bar{u}_l(k)-K_l^i x(k)\right)^\top \otimes \left(\tilde{u}_{a,l,\hat{l}}(k)+L_{l,\hat{l}}^i x(k)\right)^\top\right) \\ & \times \mathrm{vec}(\mathcal{B}_l^\top P_l^{i+1}\mathcal{B}_{l,\hat{l}}) \\ & + \left(\left(\tilde{u}_{a,l,\hat{l}}(k)-L_{l,\hat{l}}^i x(k)\right)^\top \otimes \left(\bar{u}_l(k)+K_l^i x(k)\right)^\top\right) \\ & \times \mathrm{vec}(\mathcal{B}_{l,\hat{l}}^\top P_l^{i+1}\mathcal{B}_l) \\ =\ & x^\top(k)Qx(k) + x^\top(k)K_l^{i^\top}R_{1,l}K_l^i x(k) \\ & -\gamma^2 x^\top(k)L_{l,\hat{l}}^{i^\top}R_{2,l}L_{l,\hat{l}}^i x(k). \end{aligned} \tag{26}$$

According to (26), we can simultaneously obtain the pair $(P_l^{i+1}, K_l^{i+1}, L_{l,\hat{l}}^{i+1})$ by using the least square approach. Obviously, there exist ζ ($\zeta = n_x^2 + \mathrm{card}(l)^2 + \mathrm{card}(\hat{l})^2 + n_x\mathrm{card}(l) + n_x\mathrm{card}(\hat{l}) + 2\mathrm{card}(\hat{l})\mathrm{card}(l)$) unknown elements to be solved in (26). Accordingly, we at least need to collect ζ data to solve (26). To utilize the least square approach, define the following variables

$$\begin{aligned} \mathcal{W}_{l,o}^i &= \left[\ \mathcal{W}_{1,l,o}^i \quad \mathcal{W}_{2,l,o}^i \quad \mathcal{W}_{3,l,o}^i \quad \mathcal{W}_{4,l,o}^i \quad \mathcal{W}_{5,l,o}^i \right. \\ &\quad \left. \mathcal{W}_{6,l,o}^i \quad \mathcal{W}_{7,l,o}^i\ \right], \\ \mathcal{W}_{1,l,o}^i &= x^\top(k+o)\otimes x^\top(k+o) \\ &\quad -x^\top(k+o+1)\otimes x^\top(k+o+1), \\ \mathcal{W}_{2,l,o}^i &= 2\left(\bar{u}_l(k+o)-K_l^i x(k+o)\right)^\top \otimes x^\top(k+o), \\ \mathcal{W}_{3,l,o}^i &= 2\left(\tilde{u}_{a,l,\hat{l}}(k+o)-L_{l,\hat{l}}^i x(k+o)\right)^\top \otimes x^\top(k+o), \\ \mathcal{W}_{4,l,o}^i &= \left(\bar{u}_l(k+o)-K_l^i x(k+o)\right)^\top \\ &\quad \otimes \left(\bar{u}_l(k+o)+K_l^i x(k+o)\right)^\top, \\ \mathcal{W}_{5,l,o}^i &= \left(\tilde{u}_{a,l,\hat{l}}(k+o)-L_{l,\hat{l}}^i x(k+o)\right)^\top \\ &\quad \otimes \left(\tilde{u}_{a,l,\hat{l}}(k+o)+L_{l,\hat{l}}^i x(k+o)\right)^\top, \\ \mathcal{W}_{6,l,o}^i &= \left(\bar{u}_l(k+o)-K_l^i x(k+o)\right)^\top \\ &\quad \otimes \left(\tilde{u}_{a,l,\hat{l}}(k+o)+L_{l,\hat{l}}^i x(k+o)\right)^\top, \\ \mathcal{W}_{7,l,o}^i &= \left(\tilde{u}_{a,l,\hat{l}}(k+o)-L_{l,\hat{l}}^i x(k+o)\right)^\top \\ &\quad \otimes \left(\bar{u}_l(k+o)+K_l^i x(k+o)\right)^\top, \\ \Psi_l^{i+1} &= \left[\ \Psi_{1,l}^{i+1^T} \quad \Psi_{2,l}^{i+1^T} \quad \Psi_{3,l}^{i+1^T} \quad \Psi_{4,l}^{i+1^T} \quad \Psi_{5,l}^{i+1^T} \right. \\ &\quad \left. \Psi_{6,l}^{i+1^T} \quad \Psi_{7,l}^{i+1^T}\ \right]^\top, \end{aligned}$$

$$
\begin{aligned}
\Psi_{1,l}^{i+1} &= \mathrm{vec}(P_l^{i+1}),\ \Psi_{2,l}^{i+1} = \mathrm{vec}(\mathcal{B}_l^\top P_l^{i+1} A),\\
\Psi_{3,l}^{i+1} &= \mathrm{vec}(\mathcal{B}_{l,\hat{l}}^\top P_l^{i+1} A),\ \Psi_{4,l}^{i+1} = \mathrm{vec}(\mathcal{B}_l^\top P_l^{i+1} \mathcal{B}_l),\\
\Psi_{5,l}^{i+1} &= \mathrm{vec}(\mathcal{B}_{l,\hat{l}}^\top P_l^{i+1} \mathcal{B}_{l,\hat{l}}),\ \Psi_{6,l}^{i+1} = \mathrm{vec}(\mathcal{B}_l^\top P_l^{i+1} \mathcal{B}_{l,\hat{l}}),\\
\Psi_{7,l}^{i+1} &= \mathrm{vec}(\mathcal{B}_{l,\hat{l}}^\top P_l^{i+1} \mathcal{B}_l),\\
\Phi_{l,o}^{i} &= x^\top(k+o) Q x(k+o) + x^\top(k+o) K_l^{i\top} R_{1,l} K_l^i\\
&\quad \times x(k+o) - \gamma^2 x^\top(k+o) L_{l,\hat{l}}^{i\top} R_{2,l} L_{l,\hat{l}}^i x(k+o),
\end{aligned}
$$

where $o = 0, 1, \ldots, \zeta$.

After collecting ζ data, (26) can be solved as

$$
\Psi_l^{i+1} = \left(\tilde{\mathcal{W}}_{l,o}^{i\,T} \tilde{\mathcal{W}}_{l,o}^i\right)^{-1} \tilde{\mathcal{W}}_{l,o}^{i\top} \tilde{\Phi}_{l,o}^i, \tag{27}
$$

where $\tilde{\mathcal{W}}_{l,o}^i = \begin{bmatrix} \mathcal{W}_{l,0}^{i\top} & \mathcal{W}_{l,1}^{i\top} & \ldots & \mathcal{W}_{l,\zeta-1}^{i\top} \end{bmatrix}^\top$, $\tilde{\Phi}_{l,o}^i = \begin{bmatrix} \Phi_{l,0}^i & \Phi_{l,1}^i & \ldots & \Phi_{l,\zeta-1}^i \end{bmatrix}^\top$.

Then, the control scheme and the attack strategy using the zero-sum game approach can be designed as

$$
\begin{aligned}
K_l^{i+1} &= \left(R_{1,l} + \Psi_{4,l}^{i+1} - \Psi_{6,l}^{i+1}\left(\Psi_{5,l}^{i+1} - \gamma^2 R_{2,l}\right)^{-1} \Psi_{7,l}^{i+1}\right)^{-1}\\
&\quad \times \left(\Psi_{6,l}^{i+1}\left(\Psi_{5,l}^{i+1} - \gamma^2 R_{2,l}\right)^{-1} \Psi_{3,l}^{i+1} - \Psi_{2,l}^{i+1}\right), \qquad (28)\\
L_{l,\hat{l}}^{i+1} &= \left(\Psi_{5,l}^{i+1} - \gamma^2 R_{2,l} - \Psi_{7,l}^{i+1}\left(R_{1,l} + \Psi_{4,l}^{i+1}\right)^{-1} \Psi_{6,l}^{i+1}\right)^{-1}\\
&\quad \times \left(\Psi_{7,l}^{i+1}\left(R_{1,l} + \Psi_{4,l}^{i+1}\right)^{-1} \Psi_{2,l}^{i+1} - \Psi_{3,l}^{i+1}\right). \qquad (29)
\end{aligned}
$$

According to the least square approach in (27), the control scheme (28) and the attack policy (29) can be learned and updated. The detailed process is provided in Algorithm 1.

Algorithm 1 Learning based control scheme

1: Set the initial learning step $i = 0$ and the learning error ϵ, which is a small positive scalar
2: Give a control gain K_l and $\bar{u}_l(k) = K_l(k) + e(k)$ with the probing noise $e(k) \neq 0$
3: Collect ζ data and solve (27) to obtain Ψ_l^{i+1}
4: Update K_l^{i+1} and $L_{l,\hat{l}}^{i+1}$ using (28) and (29)
5: **if** $\|K_l^{i+1} - K_l^i\| < \epsilon$ & $\|L_{l,\hat{l}}^{i+1} - L_{l,\hat{l}}^i\| < \epsilon$ **then**
6: Output K_l^{i+1} and $L_{l,\hat{l}}^{i+1}$
7: **else**
8: $i = i + 1$
9: Return to Step 3
10: **end if**

For the convergence of Algorithm 1, it is concluded in Theorem 2, the proof of which is omitted for want of space, see [32].

Theorem 2: The gains K_l^{i+1} in (28) and $L_{l,\hat{l}}^{i+1}$ in (29) can converge to K_l in (18) and $L_{l,\hat{l}}$ in (19), respectively. Moreover, the system in (15) can be stabilized.

B. Proactive and reactive defense control algorithm design

Combining the above MTD proactive control scheme and the zero-sum game based reactive controller, the proactive and reactive defense control algorithm is proposed in Algorithm 2.

Algorithm 2 Proactive and reactive defense control algorithm

1: Set $k = 0$ as the initial time and give the initial state value $x(k)$
2: Find the combination set $\tilde{B}_1$, that is, all possible controllable pairs $(A, \mathcal{B}_l)$
3: **for** $l = 1 : \mathrm{card}(\tilde{B}_1)$ **do**
4: For each controllable pair $(A, \mathcal{B}_l)$, solve the solution P_l of the Riccati equation in Lemma 1
5: Compute the controller gains K_l in Lemma 1
6: Compute the value function $V_l^* = x^\top(0) P_l x(0)$
7: Set the weighting coefficient δ and solve the probability p_l using Lemma 2
8: **end for**
9: Active the pair $(A, \mathcal{B}_l)$ in accordance with $l = \arg \min_{l=1,2\ldots,\mathrm{card}(\tilde{B}_1)} x^\top(0) P_l x(0)$
10: $\sigma = 0$
11: **while** $k + \sigma < k + \tau$ **do**
12: Run the system in (3)
13: Run the attack detector in (6)
14: $\sigma = \sigma + 1$
15: **end while**
16: **if** $R_{resi} \leq \mathcal{V}$ **then**
17: Return to Step 9
18: **else**
19: Active the designed attack isolator and find the index set of attacked actuators $\hat{l}$
20: Define $\tilde{\tilde{B}}_1$ as a subset of $\tilde{B}_1$ and all pairs $(A, \mathcal{B}_l)$ in $\hat{\tilde{B}}_1$ relate to $\hat{l}$
21: **if** $\tilde{B}_1 \cap \hat{\tilde{B}}_1 \neq \tilde{B}_1$ **then**
22: Alter the system dynamics based on $l = \arg \min_{l=1,2\ldots,\mathrm{card}(\tilde{B}_1)-\mathrm{card}(\hat{\tilde{B}}_1)} x^\top(0) P_l x(0)$ with P_l being solved using $\tilde{B}_1 \cap \hat{\tilde{B}}_1$
23: Return to Step 10
24: **else**
25: Active the reactive defense control scheme using Algorithm 1
26: Run the system (4) at least τ time constants, and then return to Step 9
27: **end if**
28: **end if**

Remark 3: From Algorithm 2, we can find that more available dynamics in $\tilde{B}_1$ (i.e., more actuators are equipped) can result in a higher performance of the MTD scheme. Accordingly, some redundant actuators can be equipped with the physical system. Another effective approach to improving the performance of the MTD scheme is to timely recover the attacked channels after they are isolated.

Next, a theorem is provided to show that Algorithm 2 can stabilize the physical system (1) under attacks.

Theorem 3: If the moving target scheme is designed in Lemma 2, the active defense controller is designed using Lemma 1 and the reactive defense controller is given in Algorithm 1, Algorithm 2 can guarantee that the physical process (1) under attacks is stable.

Proof. The proof can be readily completed by using the results in Theorems 1 and 2, thus omitted here. ■

6.. Simulation Results

This section provides simulation results to show the effectiveness and advantages of the proposed secure control scheme. A three-tank system used in [33] is regarded as the target physical system in this example. As described in [33], from left to right, three tanks in the system are respectively labeled as tank 1, tank 3 and tank 2 and the corresponding levels of tanks 1, 2, 3 are defined as h_1, h_2 and h_3. When the system is stabilized, $h_1 > h_3 > h_2$ holds. In this example, the objective is to maintain the levels of the three tanks under malicious behaviors. The modeling process and meaning of the system parameters can refer to [33]. With the sampling period being 1 s, the system matrices in (1) are given as

$$A = \begin{bmatrix} 0.9889 & 0.0001 & 0.0110 \\ 0.0001 & 0.9774 & 0.0119 \\ 0.0110 & 0.0119 & 0.9770 \end{bmatrix},$$
$$B = \begin{bmatrix} 64.5993 & 0.0015 \\ 0.0015 & 64.2236 \\ 0.3604 & 0.3910 \end{bmatrix}, \ C = \text{diag}\{1, 1, 1\}.$$

By direct calculation, the controllable set $\tilde{B}_1$ is obtained as $\tilde{B}_1 = \{\mathcal{B}_1, \mathcal{B}_2, \mathcal{B}_3\}$, where

$$\mathcal{B}_1 = \begin{bmatrix} 64.5993 \\ 0.0015 \\ 0.3604 \end{bmatrix}, \ \mathcal{B}_2 = \begin{bmatrix} 0.0015 \\ 64.2236 \\ 0.3910 \end{bmatrix}, \ \mathcal{B}_3 = B.$$

Based on Lemma 1, the corresponding controller gains are solved as $K_1 = \begin{bmatrix} -0.0153 & -0.0023 & -0.0047 \end{bmatrix}$, $K_2 = \begin{bmatrix} -0.0054 & -0.0153 & -0.0061 \end{bmatrix}$, $K_3 = \begin{bmatrix} -0.0153 & 0.0000 & -0.0034 \\ 0.0000 & -0.0152 & -0.0037 \end{bmatrix}$.

Setting $x(0) = [0.3182\ 0.1517\ 0.2314]^\top$ as the initial condition, the switching probabilities to activate each dynamics can be calculated as $p_1 = 0.3843$, $p_2 = 0.1470$, $p_3 = 0.4687$, based on which we can see that the system operator preferably chooses K_3 as the controller. The reason is that K_3 is the optimal control scheme for the system. Fig. 3 depicts the switching signal under the probabilities p_l, which is also consistent with the previous theoretical analysis. First of all, the simulation is conducted without attacks. Fig. 4 shows the levels of tanks 1, 2 and 3 without attacks using the optimal control gain K_3. The levels of tanks 1, 2 and 3 without attacks using the MTD scheme are given in Fig. 5. Fig. 6 provides the comparisons of control cost between the optimal control gain K_3 and the MTD control scheme, which also indicates that it is unavoidable to sacrifice the cost when the MTD control scheme is implemented. It is noted that we can tune the weighting coefficient δ in Lemma 2 to change the switching probabilities, and then change the control cost. Next, we will provide two cases to show the effectiveness and advantages of the proposed secure control scheme.

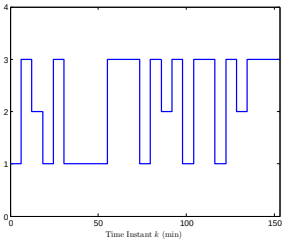

Fig. 3. The moving target switching signal under the probabilities p_l.

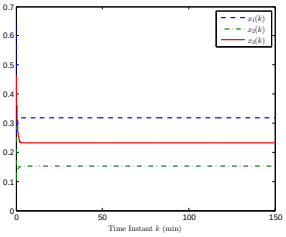

Fig. 4. The levels of tank 1, tank 2 and tank 3 without attacks using the optimal control gain K_3.

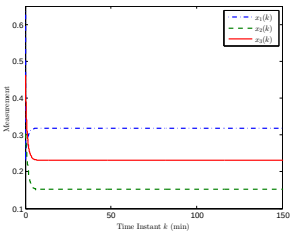

Fig. 5. The levels of tank 1, tank 2 and tank 3 without attacks using the MTD control scheme.

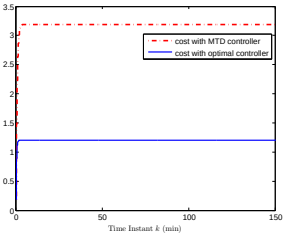

Fig. 6. Control costs respectively under the optimal control gain K_3 and MTD control scheme.

Case 1. Both actuators are attacked at once

The main purpose of this case is to show the importance of designing the learning based reactive secure control scheme. After having access to intruding the cyber layer, the adversary hijacks both actuators. In this case, the MTD control scheme fails to work because no available dynamics can be altered. To show the simulation results, we assume that the adversary intrudes the cyber layer when the 9-th switching happens and the attack signal is defined as $0.01[|\cos(x_1(k))|\ |\cos(x_2(k))|]^\top$. Along with the running of the designed attack detector, the detector can report an alarm timely. Fig. 7 shows that the residual signal is greater than the predefined threshold at the 55-th min. Then, the isolation scheme is activated. The estimates of attack signals are provided in Fig. 8, with which we can determine that both actuators are successfully intruded. There exist no available dynamics without attacks that can be altered. The MTD control scheme fails to work. Fig. 9 gives the responses of the system states, which is still in the framework of MTD control scheme. As can be seen from Fig. 9, the performance of the three-tank system is deteriorated under malicious behaviors. To cope with such a scenario, the zero-sum game based reactive control scheme is proposed in this paper. According to Algorithm 2, Algorithm 1 is invoked. Fig. 10 shows the error evolutions of $\|K_l^i - K_l^*\|$ and $\|L_l^i - L_l^*\|$. Using the reactive secure control scheme, Fig. 11 provides the levels of tanks 1, 2 and 3. It can be found that the reactive secure control scheme can recover the system performance.

With the simulation results in Case 1, we can conclude that Algorithm 2 is still effective when all actuator signals are compromised. Since the results in [27] impose an assumption

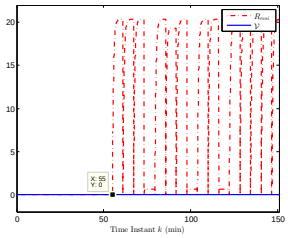

Fig. 7. The attack alarm.

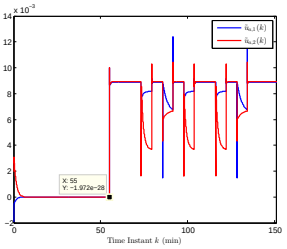

Fig. 8. The estimates of attacks.

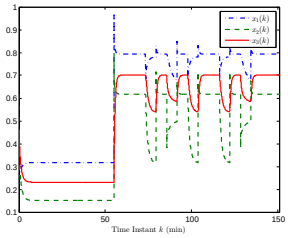

Fig. 9. Levels of tanks 1, 2, 3 when all actuator signals are compromised using the scheme in [27].

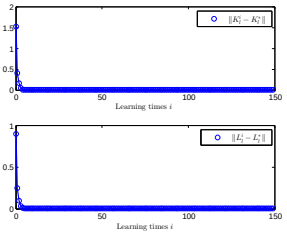

Fig. 10. Error evolutions of $\|\boldsymbol{K}_l^i - \boldsymbol{K}_l^*\|$ and $\|\boldsymbol{L}_l^i - \boldsymbol{L}_l^*\|$.

on the number of attacked actuators, its scheme cannot be applied here, demonstrating the advantages of the Algorithm 2. Next, a general case is considered.

Case 2. Only one actuator is compromised

This case mainly focuses on showing the necessity of designing the attack isolation scheme. It is assumed that only one actuator in the three-tank system is intruded. According to the moving target switching signal in Fig. 3, the first actuator is chosen as an example to run Algorithm 2. First, the simulation without using the isolation scheme is conducted. Figs. 12-14 provide the simulation results, where Fig. 12 shows that the attack detection scheme works well and an alarm can be reported timely. Without the isolation scheme, the system operator cannot exactly locate the attacked actuator. Even the current attacked dynamics are excluded from the switching sequence, the remaining dynamics can still include compromised actuators. Therefore, once the available dynamics which include the compromised actuators are altered, the MTD scheme cannot work well. Fig. 13 gives the evolution of the moving target switching signal. Obviously, after receiving the attack alarm, target 1 is excluded from the sequence. But target 1 is a part of target 3. Accordingly, as can be seen

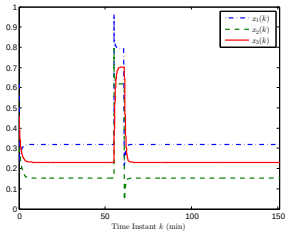

Fig. 11. Levels of tanks 1, 2, 3 when all actuator signals are compromised using Algorithm 2.

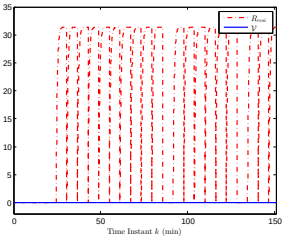

Fig. 12. The attack alarm in Case 2.

from Fig. 14, the performance of the three-tank system is not recovered. Also, when target 2 is activated, the system performance tends to be recovered. Due to the limited dwell time, the desired performance is not obtained.

Next, invoking the attack isolation scheme, Fig. 15 depicts the estimates of attacks, with which we can exactly determine that the first actuator is attacked. Using the isolation scheme, targets 1 and 3 are excluded from the switching sequence. Only target 2 can be available though it sacrifices more optimality. Fig. 16 presents the levels of tanks 1, 2 and 3. It can be found that the levels can be recovered and maintained and that the final levels are the same as those in Figs. 4 and 5. Additionally, since the target 2 is the only one that the system operator can be altered, the attack detector will not report an alarm and the estimates of attack signals are zero. Figs. 17 and 18 provide the corresponding simulation results, which validates the previous analysis. Based on the simulation results and sufficient discussions, we can conclude that the proposed secure control scheme is effective.

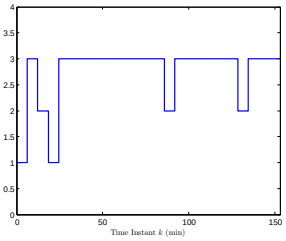

Fig. 13. Moving target switching signal in Case 2.

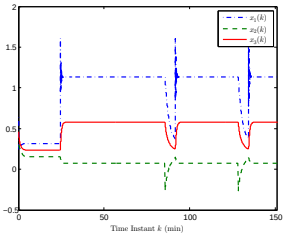

Fig. 14. Levels of tanks 1, 2 and 3 in Case 2.

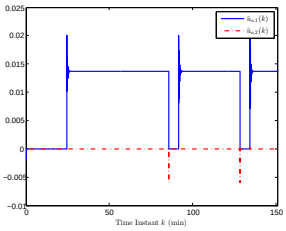

Fig. 15. Estimates of attack signals in Case 2.

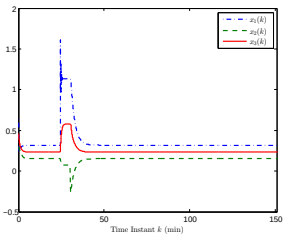

Fig. 16. Levels of tanks 1, 2 and 3 using Algorithm 2.

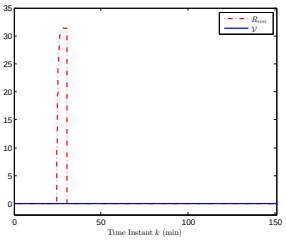

Fig. 17. The attack alarm using Algorithm 2.

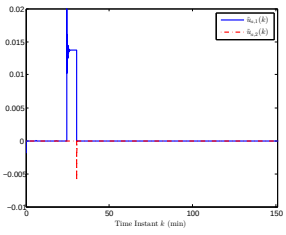

Fig. 18. Estimates of attack signals using Algorithm 2.

7.. Conclusion

The problem of secure defense control for CPS under actuator false data injection attacks has been studied in this

paper. By introducing a well-designed switching sequence, the system dynamics have been altered by unpredictably switching controllable pair $(A, \mathcal{B}_l)$. For each activated model, the linear quadratic optimal approach has been adopted to derive the corresponding controller, and the stability has been proven based on the definition of the average dwell time. By designing an attack detection observer, an attack detector has been proposed to report an attack alarm. Using the unknown input observer, an attack isolator has been derived to accurately determine which actuators were corrupted by false data. With the zero-sum game theory and reinforce learning technique, a learning-based reactive defense control scheme has been proposed to solve the problem existing in the scenario that no controllable pairs $(A, \mathcal{B}_l)$ can be chosen for the MTD design. Integrating all the above designs, a secure control algorithm has been proposed for the CPS. Finally, the proposed algorithm has been applied to the system, and the simulation results show the effectiveness and the advantages. In the future, we will investigate how to integrate the event-triggered scheme [34] into the secure algorithm of this paper for multi-agent systems [35].

REFERENCES

[1] E. A. Lee, "Cyber physical systems: Design challenges," in *11th International Symposium on Object and Component-Oriented Real-Time Distributed Computing*, pp. 363–369, IEEE, 2008.

[2] R. Rajkumar, I. Lee, L. Sha, and J. Stankovic, "Cyber-physical systems: the next computing revolution," in *Design Automation Conference*, pp. 731–736, IEEE, 2010.

[3] J. P. Farwell and R. Rohozinski, "Stuxnet and the future of cyber war," *Survival*, vol. 53, no. 1, pp. 23–40, 2011.

[4] J. Slay and M. Miller, "Lessons learned from the maroochy water breach," in *International Conference on Critical Infrastructure Protection*, pp. 73–82, Springer, 2007.

[5] A. A. Cárdenas, S. Amin, and S. Sastry, "Research challenges for the security of control systems," in *HotSec*, 2008.

[6] A. A. Cárdenas, S. Amin, and S. Sastry, "Secure control: Towards survivable cyber-physical systems," in *The 28th International Conference on Distributed Computing Systems Workshops*, pp. 495–500, IEEE, 2008.

[7] F. Pasqualetti, F. Dörfler, and F. Bullo, "Control-theoretic methods for cyberphysical security: Geometric principles for optimal cross-layer resilient control systems," *IEEE Control Syst. Mag.*, vol. 35, no. 1, pp. 110–127, 2015.

[8] Y. Mo and B. Sinopoli, "Secure control against replay attacks," in *47th Annual Allerton Conference on Communication, Control, and Computing*, pp. 911–918, IEEE, 2009.

[9] 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, 2013.

[10] T. Yang, C. Murguia, M. Kuijper, and D. Nesic, "An unknown input multi-observer approach for estimation and control under adversarial attacks," *IEEE Trans. Control Netw. Syst.*, vol. 8, no. 1, pp. 475–486, 2021.

[11] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, M. I. Jordan, and S. S. Sastry, "Kalman filtering with intermittent observations," *IEEE Trans. Autom. Control*, vol. 49, no. 9, pp. 1453–1464, 2004.

[12] H. Fawzi, P. Tabuada, and S. Diggavi, "Secure estimation and control for cyber-physical systems under adversarial attacks," *IEEE Trans. Autom. Control*, vol. 59, no. 6, pp. 1454–1467, 2014.

[13] L. Hu, Z. Wang, Q.-L. Han, and X. Liu, "State estimation under false data injection attacks: Security analysis and system protection," *Automatica*, pp. 176–183, 2018.

[14] Y. Shoukry and P. Tabuada, "Event-triggered state observers for sparse sensor noise/attacks," *IEEE Trans. Autom. Control*, vol. 61, no. 8, pp. 2079–2091, 2015.

[15] C. Wu, Z. Hu, J. Liu, and L. Wu, "Secure estimation for cyber-physical systems via sliding mode," *IEEE Trans. Cybern.*, vol. 48, no. 12, pp. 3420–3431, 2018.

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

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

[18] C. Wu, W. Pan, G. Sun, J. Liu, and L. Wu, "Learning tracking control for cyber-physical systems," *IEEE Internet Things J., DOI: 10.1109/JIOT.2021.3056633*, 2021.

[19] C. Wu, L. Wu, J. Liu, and Z.-P. Jiang, "Active defense-based resilient sliding mode control under denial-of-service attacks," *IEEE Trans. Inf. Forensics Security*, vol. 15, pp. 237–249, 2020.

[20] 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. Mag.*, vol. 35, no. 1, pp. 46–65, 2015.

[21] C. Wu, X. Li, W. Pan, J. Liu, and L. Wu, "Zero-sum game based optimal secure control under actuator attacks," *IEEE Trans. Autom. Control, DOI: 10.1109/TAC.2020.3029342*, 2020.

[22] Y. Li, L. Shi, P. Cheng, J. Chen, and D. E. Quevedo, "Jamming attacks on remote state estimation in cyber-physical systems: A game-theoretic approach," *IEEE Trans. Autom. Control*, vol. 60, no. 10, pp. 2831–2836, 2015.

[23] F. Miao, Q. Zhu, M. Pajic, and G. J. Pappas, "A hybrid stochastic game for secure control of cyber-physical systems," *Automatica*, vol. 93, pp. 55–63, 2018.

[24] Y. Zhou, K. G. Vamvoudakis, W. M. Haddad, and Z.-P. Jiang, "A secure control learning framework for cyber-physical systems under sensor and actuator attacks," *IEEE Trans. Cybern., DOI: 10.1109/TCYB.2020.3006871*, 2020.

[25] Q. Zhu and T. Başar, "Game-theoretic approach to feedback-driven multi-stage moving target defense," in *International Conference on Decision And Game Theory for Security*, pp. 246–263, Springer, 2013.

[26] P. Griffioen, S. Weerakkody, and B. Sinopoli, "A moving target defense for securing cyber-physical systems," *IEEE Trans. Autom. Control*, vol. 66, no. 5, pp. 2016–2031, 2021.

[27] A. Kanellopoulos and K. G. Vamvoudakis, "A moving target defense control framework for cyber-physical systems," *IEEE Trans. Autom. Control*, vol. 65, no. 3, pp. 1029–1043, 2020.

[28] F. L. Lewis, D. Vrabie, and V. L. Syrmos, *Optimal Control*. John Wiley & Sons, 2012.

[29] J. Liu, L. Wu, C. Wu, W. Luo, and L. G. Franquelo, "Event-triggering dissipative control of switched stochastic systems via sliding mode," *Automatica*, vol. 103, pp. 261–273, 2019.

[30] S. X. Ding, *Model-Based Fault Diagnosis Techniques: Design Schemes, Algorithms, and Tools*. Springer Science & Business Media, 2008.

[31] A. Al-Tamimi, F. L. Lewis, and M. Abu-Khalaf, "Model-free Q-learning designs for linear discrete-time zero-sum games with application to H-infinity control," *Automatica*, vol. 43, no. 3, pp. 473–481, 2007.

[32] B. Kiumarsi, F. L. Lewis, and Z.-P. Jiang, "H_∞ control of linear discrete-time systems: Off-policy reinforcement learning," *Automatica*, vol. 78, pp. 144–152, 2017.

[33] X. He, Z. Wang, Y. Liu, L. Qin, and D. Zhou, "Fault-tolerant control for an internet-based three-tank system: Accommodation to sensor bias faults," *IEEE Trans. Ind. Electron.*, vol. 64, no. 3, pp. 2266–2275, 2016.

[34] H. Ma, H. Li, R. Lu, and T. Huang, "Adaptive event-triggered control for a class of nonlinear systems with periodic disturbances," *Sci. China In. Sci*, vol. 63, no. 5, pp. 1–15, 2020.

[35] M. Lv, W. Yu, J. Cao, and S. Baldi, "Consensus in high-power multiagent systems with mixed unknown control directions via hybrid nussbaum-based control," *IEEE Trans Cybern., DOI: 0.1109/TCYB.2020.3028171*, 2020.

8.. APPENDIX

A. Proof for Theorem 1

Proof: For each activated model, define the Lyapunov function as $\mathcal{V}_l(k) = x^\top(k) P_l x(k)$.

Then, we can obtain

$$\begin{aligned}\Delta\mathcal{V}_l(k) &= x^\top(k)\left(A+BK_l\right)^\top P_l \left(A+BK_l\right) x(k) \\ &= -x^\top(k)\tilde{Q}_l x(k),\end{aligned}$$

where $\tilde{Q}_l = Q_l + K_l^\top R_l K_l$.

Considering the fact that $Q_l \geq 0$ and $R_l > 0$, we can obtain $\Delta\mathcal{V}_l(k) \leq 0$. Furthermore, the inequality $\Delta\mathcal{V}_l(k) \leq -\lambda_{\min}(\tilde{Q}_l)\|x(k)\|^2$ holds. Using the inequality $\mathcal{V}_l(k) \leq \lambda_{\max}(P_l)\|x(k)\|^2$ yields $\|x(k)\|^2 \geq \frac{1}{\lambda_{\max}(P_l)}\mathcal{V}_l(k)$, which further implies $\Delta\mathcal{V}_l(k) \leq -\frac{\lambda_{\min}(\tilde{Q}_l)}{\lambda_{\max}(P_l)}\mathcal{V}_l(k) \leq -\beta\mathcal{V}_l(k)$.

Then, we can obtain

$$\mathcal{V}_l(k) \leq (1-\beta)\,\mathcal{V}_l(k-1). \tag{30}$$

For the activated l-th sub-model over the interval $[k_i,\ k_{i+1})$, the inequality $\mathcal{V}_l(k) \leq (1-\beta)^{(k-k_i)}\,\mathcal{V}_l(k_i)$ holds.

According to $\mathcal{V}_l(k) \leq \lambda_{\max}(P_l)\|x(k)\|^2$ and $\mathcal{V}_q(k) \geq \lambda_{\min}(P_q)\|x(k)\|^2$, the following inequality holds

$$\mathcal{V}_l(k) \leq \frac{\lambda_{\max}(P_l)}{\lambda_{\min}(P_q)}\mathcal{V}_q(k) \leq \mu\mathcal{V}_q(k). \tag{31}$$

To facilitate describing the proof, define $\alpha(k)$ as the switching signal and $\alpha(k)$ is equivalent to subscript l. Over the overall time window, combining the inequalities in (30) and (31) yields

$$\begin{aligned}\mathcal{V}_{\alpha(k)}(k) &\leq (1-\beta)^{(k-k_i)}\,\mu\mathcal{V}_{\alpha(k_{i-1})}(k_{i-1})\\ &\leq \cdots\\ &\leq (1-\beta)^{(k-k_0)}\,\mu^{\frac{k-k_0}{\tau}}\mathcal{V}_{\alpha(k_0)}(k_0)\\ &= \left((1-\beta)\mu^{\frac{1}{\tau}}\right)^{(k-k_0)}\mathcal{V}_{\alpha(k_0)}(k_0),\end{aligned} \tag{32}$$

Notice that $\mathcal{V}_{\alpha(k)}(k) \geq \underline{\lambda}(P_{\alpha(k)})\|x(k)\|^2$ and $\mathcal{V}_{\alpha(k_0)}(k_0) \leq \bar{\lambda}(P_{\alpha(k_0)})\|x(k_0)\|^2$. Then, (32) can be rewritten as

$$\|x(k)\|^2 \leq \frac{\bar{\lambda}(P_{\alpha(k_0)})}{\underline{\lambda}(P_{\alpha(k)})}\left((1-\beta)\mu^{\frac{1}{\tau}}\right)^{(k-k_0)}\|x(k_0)\|^2.$$

Accordingly, if $0 < (1-\beta)\mu^{\frac{1}{\tau}} < 1$, that is, $\tau > \text{ceil}\left(-\frac{\ln\mu}{\ln(1-\beta)}\right)$ holds, the overall system can be exponentially stabilized. The proof is completed. ■

Chengwei Wu received the B.S. degree in management from the Arts and Science College, Bohai University, Jinzhou, China, in 2013, the M.S. degree from Bohai University, in 2016, and the Ph.D. degree from Harbin Institute of Technology, China, 2021. From July 2015 to December 2015, he was a Research Assistant in the Department of Mechanical Engineering, The Hong Kong Polytechnic University. From 2019 to 2021, he was a joint-PhD student at Department of Cognitive Robotics, Delft University of Technology, Netherlands. He is currently an Assistant Professor with the Harbin Institute of Technology, Harbin, China. His research interests include sliding mode control and networked control systems.

Weiran Yao received the Bachelor's (with honors) degree, the Master's degree, and the Doctor's degree in aeronautical and astronautical science and technology from the School of Astronautics, Harbin Institute of Technology (HIT), Harbin, China in 2013, 2015, and 2020, respectively. From 2017 to 2018, he was a visiting PhD student at the Department of Mechanical and Industrial Engineering, University of Toronto (UofT), Toronto, Canada.

He is currently an Assistant Professor with the School of Astronautics, HIT. His research interests include unmanned vehicles, multi-robot mission planning, and multi-agent control systems.

Wei Pan received the Ph.D. degree in Bioengineering from Imperial College London in 2016. He is currently an Assistant Professor at Department of Cognitive Robotics, Delft University of Technology. Until May 2018, he was a Project Leader at DJI, Shenzhen, China, responsible for machine learning research for DJI drones and AI accelerator. He is the recipient of Dorothy Hodgkin's Postgraduate Awards, Microsoft Research Ph.D. Scholarship and Chinese Government Award for Outstanding Students Abroad, Shenzhen Peacock Plan Award. He is an active reviewer and committee member for many international journals and conferences. His research interests include machine learning and control theory with applications in robotics.

Guanghui Sun received the B.S. degree in automation and the M.S. and Ph.D. degrees in control science and engineering from Harbin Institute of Technology, Harbin, China, in 2005, 2007, and 2010, respectively. He is currently a Professor in the Department of Control Science and Engineering, Harbin Institute of Technology. His research interests include fractional-order systems, networked control systems, and sliding mode control.

Jianxing Liu received the B.S. degree in mechanical engineering in 2008, the M.E. degree in control science and engineering in 2010, both from Harbin Institute of Technology, Harbin, China and the Ph.D. degree in Automation from the Technical University of Belfort-Montbeliard (UTBM), France, in 2014. Since 2014, he joined Harbin Institute of Technology, Harbin, China. His current research interests include nonlinear control algorithms, sliding mode control, and their applications in industrial electronics systems and renewable energy systems.

Ligang Wu (M'10-SM'12-F'19) received the B.S. degree in Automation from Harbin University of Science and Technology, China in 2001; the M.E. degree in Navigation Guidance and Control from Harbin Institute of Technology, China in 2003; the Ph.D. degree in Control Theory and Control Engineering from Harbin Institute of Technology, China in 2006. From January 2006 to April 2007, he was a Research Associate in the Department of Mechanical Engineering, The University of Hong Kong, Hong Kong. From September 2007 to June 2008, he was a Senior Research Associate in the Department of Mathematics, City University of Hong Kong, Hong Kong. From December 2012 to December 2013, he was a Research Associate in the Department of Electrical and Electronic Engineering, Imperial College London, London, UK. In 2008, he joined the Harbin Institute of Technology, China, as an Associate Professor, and was then promoted to a Full Professor in 2012. Prof. Wu was the winner of the National Science Fund for Distinguished Young Scholars in 2015, and received China Young Five Four Medal in 2016. He was named as the Distinguished Professor of Chang Jiang Scholar in 2017, and was named as the Highly Cited Researcher in 2015-2019.

Prof. Wu currently serves as an Associate Editor for a number of journals, including IEEE TRANSACTIONS ON AUTOMATIC CONTROL, IEEE/ASME TRANSACTIONS ON MECHATRONICS, IEEE TRANSACTIONS ON INDUSTRIAL ELECTRONICS, *Information Sciences, Signal Processing,* and *IET Control Theory and Applications.* He is an Associate Editor for the Conference Editorial Board, IEEE Control Systems Society. He is also a Fellow of IEEE. Prof. Wu has published 7 research monographs and more than 170 research papers in international referred journals. His current research interests include switched systems, stochastic systems, computational and intelligent systems, sliding mode control, and advanced control techniques for power electronic systems.