Proactive attack detection scheme based on nonlinear transformation and moving target defense*

1st Xue Chen
School of Automation
Shenyang Aerospace University
Shenyang, China
c202205230202@163.com

2nd Hao Liu*
School of Automation
Shenyang Aerospace University
Shenyang, China
lh_hit_1985@163.com

2023 35th Chinese Control and Decision Conference (CCDC) | 979-8-3503-3472-2/23/$31.00 ©2023 IEEE | DOI: 10.1109/CCDC58219.2023.10326625

***Abstract*—This paper proposes a novel active attack defense strategy for cyber-physical systems(CPS) with unknown-but-bounded (UBB) noises. A resourceful and veteran adversary can design a sequence of perfect injection on channel of sensors and remain undetected. To confront such an attacker, a hybrid attack detection approach is designed by combining moving target(MT) and nonlinear transformation to change the original output to increase the uncertainty of the system and the estimation residues under stealthy false-data-injection (FDI) attacks. Furthermore, the attack detection rate can be improved by the hybrid attack detection approach. Finally, an unmanned aerial systems as an example is provided to demonstrate the effectiveness of the proposed method.**

***Index Terms*—Cyber-physical systems, state estimation, moving target defense, false-data-injection (FDI) attacks.**

I. Introduction

CYBER-PHYSICAL systems (CPSs) have attracted great attention and discussion in academia because of their widespread applications in social infrastructure, such as health care systems, smart grids, smart cars, waterway facilities, etc [1]. CPSs are defined as a collaborative system of communication, computing, and control interaction between physical systems and information systems [2]. With the rapid development of CPSs, the new generation of CPSs introduce various vulnerabilities, threats, and malicious attacks [3]. Therefore, the development of new attack detection methods is particularly important for CPSs.

Based on the openness and uncertainty of the network, CPSs as a networked control system are vulnerable to various attacks, especially considering that the knowledgeable and resourceful adversaries can easily utilize system model to launch stealthy attacks, such as Denial-of-Service (DoS) attacks [4, 5], false-data-injection (FDI) attacks [6, 7], replay attacks [8, 9], zero-dynamics attacks [10, 11]. In this paper, it is assumed that the sensor output channel is under a FDI attack.

To counter such an adversary, effective active defensive measures are adopted to detect malicious attacks. For example, in [12], a new physical watermarking scheme is designed to detect replay attacks, which can greatly improve the detection rate. In [13], a sensor switching rule is designed to detect perturbation attacks and replay attacks, which uses watermarking in conjunction with attack mitigation strategies rather than focusing on attack detection. The moving target defense (MTD) [14–16] is another active defense strategy, which introduces stochastic time-varying parameters in the system to increase the uncertainty of the system, so as to limit the attacker to obtain the system model knowledge and simultaneously prevent the attacker from launching covert attacks. The goal in [17] is to detect all sensor attacks by moving target defense and construct a robust estimator that can identify sensors that cause unbounded estimation errors. In [18], a novel MTD strategy is proposed, which randomly changes the availability of sensor data and makes it difficult for adversaries to capture useful information from the system. In [19], moving target defense changes system parameters by utilizing stochastic switching to prevent attackers from reconnoitring the system. Furthermore, the information entropy caused by probabilistic switching is used to optimize the unpredictability of actuators and sensors.

The proactive detection strategies of moving target and physical watermarking mentioned above have made it difficult for attackers to launch covert attacks. However, the impact on system performance is inevitable. Therefore, the trade-off between system performance and attack detection rate needs to be balanced with an optimized control strategy. A switched multiplicative watermarking strategy in [20] is designed to detect replay attacks, and finally the original measurement data is reconstructed by a watermark remover on the controller side. In this way, there will be no loss of system performance. In addition, in [21], the watermarking signal combines with a nonlinear static auxiliary function to act as an unrecognizable moving target. This approach is capable of detecting zero-dynamic, replay, and stealthy attacks.

In view of the ideas in [20] and [21], a novel proactive defence scheme is proposed by combining the nonlinear transformation and moving target defense to detect false-data-injection (FDI) attacks, which can not only increase the uncertainty of the system and makes it difficult for adversaries to launch a stealthy attack, but also improves the attack detection rate. Specifics will be discussed in detail later.

The rest of this article is organized as follows. The main components of an proactive defense strategy are formulated in Section II, and the main results in Section III. Numerical examples are provided in Section IV. Finally, Section V

This work is supported by the Project fromthe Liaoning Provincial Science-and Technology Department (2020-KF-11-08).

concludes this article.

II. PROBLEM FORMULATION

In this section, the considered cyber-physical systems(CPS) with unknown-but-bounded (UBB) noises is set up, as shown in Fig.1. It is assumed that measurement channels subject to malicious attacks. Moreover, a estimator is utilized to estimate the system state. More details of the system are given as follows.

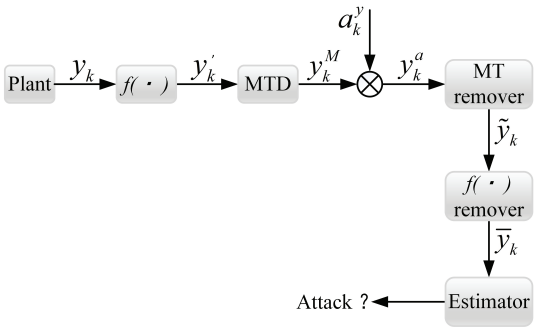

Fig. 1. The proposed proactive attack detection scheme.

A. System Model

In this work, the physical plant in Fig.1 can be represented by the following linear discrete-time system with unknown-but-bounded noises

$$x_{k+1} = Ax_k + \omega_k \tag{1}$$
$$y_k = Cx_k + \upsilon_k \tag{2}$$

where $x_k \in \mathbb{R}^{n_x}$ is the system state, $y_k \in \mathbb{R}^{n_y}$ is the control output, which captures measurements from a set of sensors $S = \{s_1, s_2, ..., s_q\}$, $\omega_k \in \mathbb{R}^{n_\omega}$ and $\upsilon_k \in \mathbb{R}^{n_\upsilon}$ are the process and measurement noises, respectively. *A* and *C* are system matrices with appropriate dimensions. It is assumed that the pair (*A,C*) is observable.

Assumption 1: Suppose that the process noise ω_k and measurement noise υ_k are unknown-but-bounded, i.e.,

$$\|\omega_k\|_2 \leq \delta, \ \|\upsilon_k\|_2 \leq \epsilon \tag{3}$$

where $\|\cdot\|_2$ denotes the 2-Norm operator, $\delta > 0 \in \mathbb{R}^{n_x}, \epsilon > 0 \in \mathbb{R}^{n_y}$ are known vector.

B. Attacks Model

The build of an attack model depends on the attack strategy and resources, i.e., the system model knowledge, the disclosure resources, and the disruptive resources.

Model knowledge: An attacker knows the knowledge of the given system, i.e., *A*, *B*, *C*, *K* and *L*.

Disclosure resources: An attacker can obtain disclosure resources through intercepting transmission channel data.

Disruptive resources: An attacker can inject false data into the transmission channels by utilizing model knowledge.

In this paper, it is assumed that an attacker launches a FDI attack on the output channel, the attacks model a_k^y has the following form

$$a_k^y = g(y_k^M) = g(\Gamma_k y_k') \tag{4}$$

where the function $g(\cdot)$ is the attack strategy designed by attackers. It is clear that a malicious adversary cannot launch a covert attack if he does not know the exact the moving target and the nonlinear transformation strategy.

C. State Estimator

In this paper, the state observer is described as the following form

$$\hat{x}_{k+1} = A\hat{x}_k + L(\bar{y}_k - C\hat{x}_k)) \tag{5}$$

where $\hat{x}_k \in \mathbb{R}^{n_x}$ is the estimate value of the state, and $L \in \mathbb{R}^{n_x \times n_y}$ is observer gain matrix. Suppose that the initial value of the estimate is $\hat{x}_0 = 0$. Now we define the state estimate error as $e_k = x_k - \hat{x}_k$. By combining (1) and (7), we can obtain the dynamics of estimate error as follows

$$\begin{aligned} e_{k+1} =& x_{k+1} - \hat{x}_{k+1} \\ =& Ax_k + \omega_k - (A\hat{x}_k + L(\bar{y}_k - C\hat{x}_k)) \\ =& Ax_k + \omega_k - A\hat{x}_k - L(y_k + f^{-1}(\Gamma_k^{-1} a_k^y) - C\hat{x}_k) \\ =& (A - LC)e_k + \omega_k - L\upsilon_k - Lf^{-1}(\Gamma_k^{-1} a_k^y) \end{aligned}$$

We all know about that the estimator is stable if and only if the matrix $(A - LC)$ is stable. In this work, the estimator can be stabilized by choosing appropriate estimator gain.

The objectives of this paper can be formulated as follows:
1) How to design the proactive detection mechanism based on nonlinear transformation and moving target such that it can operate well for the system with UBB noises?
2) How to design the proactive detection strategy to improve the attack detection rate?

III. MAIN RESULTS

In this section, a novel proactive detection approach for Cyber-Physical system is modeled to improve attack detection rate, as shown in Fig.1. Note that nonlinear transformation and moving target are adopted, and then removed eventually. The purpose is to eliminate the impact of the nonlinear transformation and moving target on system performance in the absence of attacks. In addition, more importantly, the attack detection issues will be analyzed by using the proactive defense strategy.

A. The Design of Nonlinear Transformation

In order to make it harder for adversaries to obtain system information, it can be seen from Fig.1 that a nonlinear transformation $f(\cdot)$ is placed after the sensor measures to transform the output value y_k into y_k'. As a result, an adversary cannot glean the exact knowledge of output value if he doesn't know specific $f(\cdot)$. The designed nonlinear transformation has the following form

$$y_k' = f(y_k) \tag{6}$$

where the nonlinear transformation $f(y_k)$ is a bijective function. Moreover, $f(y_k)$ can be designed as a linear or nonlinear function. To fight against resourceful and knowledgeable adversaries, a nonlinear function $f(y_k)$ is selected in this paper.

Assumption 2: $f(y_k)$ is a continuous and monotone function and its invertible function $f^{-1}(y_k)$ exists.

Remark 1: It can be seen from Fig.1 that the nonlinear function $f(y_k)$ cannot interfere the normal operation of the system under attack-free. Moreover, if the system is essentially nonlinear, then these nonlinearities can be introduced through system dynamics.

B. The Design of Moving Target

To further increase the uncertainty of the system to prevent attackers from eavesdropping on the transmission channel to launch covert attacks, the moving target is taken into account after the nonlinear transformation to change the transformed output value y'_k. The dynamics of the moving target is given as follows:

$$y_k^M = \Gamma_k y'_k \tag{7}$$

where $\Gamma_k \in \mathbb{R}^{n_y \times n_y}$, and it can be designed as any type except $\Gamma_k = I_{n_y}$. For simplicity, it is defined as the diagonal matrix below

$$\Gamma_k = \begin{bmatrix} \vartheta_k^1 & & & \\ & \vartheta_k^2 & & \\ & & \ddots & \\ & & & \vartheta_k^{n_y} \end{bmatrix}$$

where ϑ_k^i, $i = 1, 2, ..., n_y$, are parameters to be designed.

Remark 2: Unlike traditional moving target defense techniques, the one adopted in this article does not affect system performance in attack-free because it will eventually be removed.

C. Abnormal Detector

By utilizing the residual detection system, it can be effectively determined whether the current system has been damaged by an attack. The residual of estimation error is defined as follows

$$r_k^a = \bar{y}_k - C\hat{x}_k \tag{8}$$

Then the detector can be defined as below

$$flag = \begin{cases} 0, & \text{if } \|r_k^a\| \le \bar{r} \\ 1, & \text{if } \|r_k^a\| > \bar{r} \end{cases} \tag{9}$$

where $\bar{r} = [\bar{r}_1, \bar{r}_2, ..., \bar{r}_q]^\mathsf{T}$ is the threshold of the detector. $flag = 1$ means that the system is under attack, and alarm will be triggered; otherwise, the system is attack-free.

Then, we can summarize the proactive detection strategy-operations as the following pseudo process.

Algorithm 1: The Proactive Detection Strategy-Operations

1. Input: x_0, y_k, C
2. For $k = 1 : N$
   Calculate: y'_k, y_k^M according to (6) and (7), respectively
   $\bar{y}_k$ can be obtained by the $f(\cdot)$ and MT remover
   $\hat{x}_k$ according to (5)
   $r_k^a = \bar{y}_k - C\hat{x}_k$
   If $\|r_k^a\| \le \bar{r}$
   $flag = 0$(Attack-free)
   else
   $flag = 1$(Attacked)
   End If
   End For
3. Output: $flag$

D. The System Performance in Absence of Attacks

The impact of the proactive detection approach that combining the nonlinear transformation and moving target defense on system performance will be analyzed in this section.

Theorem 1: Consider the proactive detection scheme of cyber-physical systems given in Fig.1, the resulting system cannot be interfered by the proposed framework when the system is attack-free.

proof. From Fig.1, one can derive that

$$\begin{aligned} \bar{y}_k = f^{-1}(\tilde{y}_k) =& f^{-1}(\Gamma_k^{-1}(y_k^M + a_k^y)) \\ =& f^{-1}(\Gamma_k^{-1}(\Gamma_k y'_k + a_k^y)) \\ =& f^{-1}(\Gamma_k^{-1}(\Gamma_k f(y_k) + a_k^y)) \end{aligned}$$

If the system is attack-free, i.e., $a_k^y \equiv 0$, then, it is obvious that $\bar{y}_k \equiv y_k$, which completes the proof.

E. Detection of FDI Attacks

It is assumed that an adversary launches a FDI attack on the output channel, the attack model is shown in (6) .

Theorem 2: The attack detection rate can be improved by proactive defense consisting of moving target and nonlinear transformation in the following two cases: (1) $f(x)$ is a monotonic increasing function and choose a larger ϑ_k^i, $i = 1, 2, ..., n_y$. (2) $f(x)$ is a monotonic decreasing function and choose a smaller ϑ_k^i, $i = 1, 2, ..., n_y$.

proof. From (8) and $e_k = x_k - \hat{x}_k$, we can calculate

$$\begin{aligned} r_k^a =& \bar{y}_k - C\hat{x}_k \\ =& y_k + f^{-1}(\Gamma_k^{-1} a_k^y) - C\hat{x}_k \\ =& Ce_k + \upsilon_k + f^{-1}(\Gamma_k^{-1} a_k^y) \end{aligned} \tag{10}$$

Then, one can derive

$$\begin{aligned} \|r_k^a\| &\le \|C\| \|e_k\| + \|\upsilon_k\| + \left\|f^{-1}(\Gamma_k^{-1} a_k^y)\right\| \\ &\le \|C\| \|e_k\| + \epsilon + \left\|f^{-1}(\Gamma_k^{-1} a_k^y)\right\| \end{aligned} \tag{11}$$

In order to increase the estimation residues (8) under FDI attacks, i.e., the value of $\left\|f^{-1}(\Gamma_k^{-1} a_k^y)\right\|$ should be made as large as possible. Case (1): If $f(x)$ is a monotonic increasing function, then $f^{-1}(x)$ is monotonically decreasing. To

increase $\left\|f^{-1}(\Gamma_k^{-1}a_k^y)\right\|$, should make Γ_k^{-1} smaller, i.e., a larger ϑ_k^i should be chosen. Consequently, the attack detection rate will be improved on account of $\|r_k^a\| \gg \left\|r_k^{r,smaller}\right\|$, where the $r_k^{r,smaller}$ is the value when a smaller ϑ_k^r is selected. Similarly, the proof of case (2) is the same logic as case (1). Therefore, the detailed proof of case (2) is omitted here.

IV. Numerical Examples

Consider utilizing an unmanned aircraft system (UAS) in [22] to demonstrate the effectiveness of the main results. The dynamics of UAS can be given by

$$x_{k+1} = Ax_k + Bu_k + \omega_k$$
$$y_k = Cx_k + \upsilon_k$$

with

$$A = \begin{bmatrix} 1 & 0 & (1-0.5\Delta t)\Delta t & 0 \\ 0 & 1 & 0 & (1-0.5\gamma\Delta t)\Delta t \\ 0 & 0 & 1-\gamma\Delta t & \\ 0 & 0 & 0 & 1-\gamma\Delta t \end{bmatrix}$$

$$B = \begin{bmatrix} 0.5\Delta t^2 & 0 \\ 0 & 0.5\Delta t^2 \\ \Delta t & 0 \\ 0 & \Delta t \end{bmatrix}, C = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

where Δt=0.01 and γ=0.25 are the sampling time and damping parameters, respectively. $x_k = [P_x\ P_y\ V_x\ V_y]^T$ are the four states of the unmanned aerial system. The initial state is set to be $x_0 = [10\ \ -20\ 30\ \ -10]^T$. By solving pole placement, the controller gain K and the state observer gain L can be solved

$$K = \begin{bmatrix} 40.0400 & 0 & 29.5498 & 0 \\ 0 & 20.2002 & 0 & 68.7490 \end{bmatrix}$$

$$L = \begin{bmatrix} 0.2000 & 0 & 0.0499 & 0 \\ 0 & 0.2000 & 0 & 0.0499 \\ 0 & 0 & 0.4975 & 0 \\ 0 & 0 & 0 & 0.0975 \end{bmatrix}$$

According to Theorem 1, nonlinear transformation and moving target will not affect the system performance in the absence of attacks. The simulation results are shown in Figs.2-4. The trajectory of P_X and P_y is depicted in Fig.2. The state estimation error e_k is given in Fig.3. It can be seen from Fig.3 that the estimation error e_k can converge to zero quickly.

Now, it is assumed that FDI attacks occur during the time interval [20,30]. The simulation results are shown in Figs.5-6. It can be seen from Figs.5-6 that the trajectory of the system state and state estimation error changes significantly under attacks. In addition, Fig.7 is the detection results, which illustrates the attack can be detected by the proposed proactive defense strategy. It should be emphasized that $flag = 1$ means that the system is attacked, and $flag = 0$ means no attack. Moreover, the impact of the attack on the estimation residue

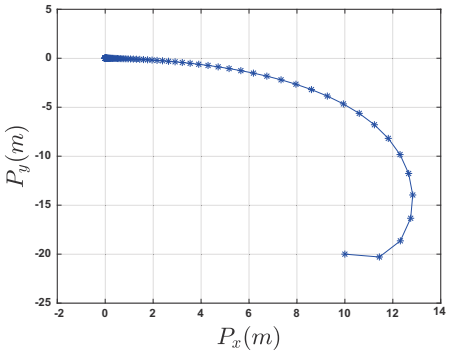

Fig. 2. The trajectory of P_x and P_y in the attack-free.

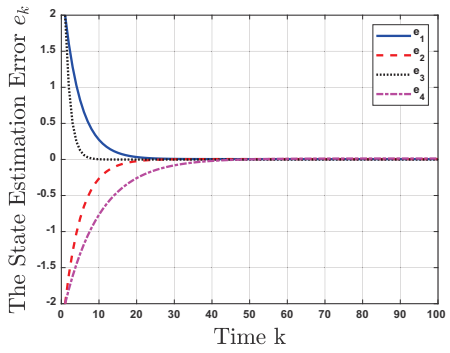

Fig. 3. The estimation error e_k in the attack-free.

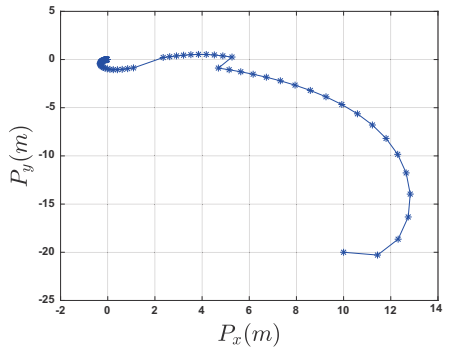

Fig. 4. The trajectory of P_x and P_x under FDI attacks.

still exists in the short time after the attack ends, in other words, the residue value is still greater than the detection threshold, i.e., $flag = 1$ during the time interval [30,35].

V. Conclusion

This paper studied the proactive detection problem in the cyber-physical systems. We proved that the attack detection rate can be improved by combining nonlinear transformation and moving target. Simulation and comparison results shown the effectiveness of the proposed strategy to detect attacks. Moreover, the proposed strategy does not affect system performance when the system is attack free. In our future work,

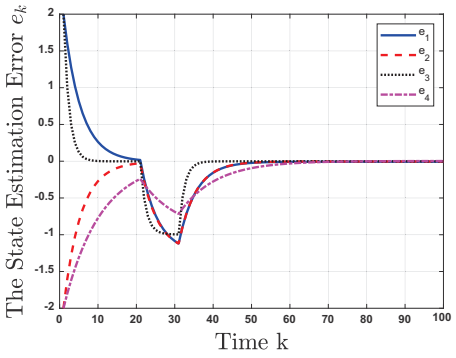

Fig. 5. The estimation error e_k under FDI attacks.

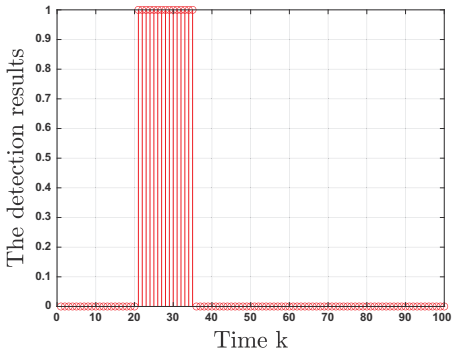

Fig. 6. The attack detection results.

the proactive defense approaches for cyber-physical systems will be further investigated.

REFERENCES

[1] K. D. Kim, and P. R. Kumar, "CyberPhysical Systems: A Perspective at the Centennial," *Proceedings of the IEEE*, vol. 100, pp. 1287-1308, 2012.

[2] E. A. Lee, "Cyber physical systems: Design challenges," *11th IEEE International Symposium on Object and Component-Oriented Real-Time Distributed Computing (ISORC)*. Orlando, FL, USA, 2008, pp. 363-369.

[3] A. Humayed, J. Lin, F. Li, and B. Luo, "Cyber-physical systems securitya survey," *IEEE Internet of Things Journal*, vol. 4, no. 6, pp. 1802-1831, 2017.

[4] M. Wakaiki, A. Cetinkaya, and H. Ishii, "Stabilization of networked control systems under DoS attacks and output quantization," *IEEE Transactions on Automatic Control* , vol. 65, no. 8, pp. 3560-3575, 2020.

[5] G. K. Befekadu, V. Gupta, and P. J. Antsaklis, "Risk-sensitive control under markov modulated denial-of-service (DoS) attack strategies," *IEEE Transactions on Automatic Control* , vol. 60, no. 12, pp. 3299-3304, 2015.

[6] Z. H. Pang, G. P. Liu, D. Zhou, F. Hou, and D. Sun, "Two-channel false data injection attacks against output tracking control of networked systems," *IEEE Transactions on Industrial Electronics* , vol. 63, no. 5, pp. 3242-3251, 2016.

[7] C. Liu, H. Liang, and T. Chen, "Network parameter coordinated false data injection attacks against power system AC state estimation," *IEEE Transactions on Smart Grid* , vol. 12, no. 2, pp. 1626-1639, 2021.

[8] M. Zhu, and S. Martinez, "On the performance analysis of resilient networked control systems under replay attacks," *IEEE Transactions on Automatic Control*, vol. 59, no. 3, pp. 804-808, 2014.

[9] Y. Su, J. Wu, C. Long, and S, Li, "Event-triggered control for networked control systems under replay attacks," *2018 Chinese Automation Congress (CAC)*, Xi'an, China, Nov. 2018. pp. 2636-2641.

[10] G. Park, C. Lee, H. Shim, Y. Eun, and K. H. Johansson, "Stealthy adversaries against uncertain cyber-physical systems: threat of robust zero-dynamics attack," *IEEE Transactions on Automatic Control*, vol. 64, no. 12, pp. 4907-4919, 2019.

[11] A. Baniamerian, K. Khorasani, and N. Meskin, "A special class of zero dynamics cyber-attacks for SISO time-delay systems," *2021 60th IEEE Conference on Decision and Control (CDC)*, Austin, TX, USA, Dec. 2021. pp. 4182-4187.

[12] S. Weerakkody, O. Ozel, and B. Sinopoli, "A bernoulli-gaussian physical watermark for detecting integrity attacks in control systems," *2017 55th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, Monticello, IL, USA, Oct. 2017, pp. 966-973.

[13] P. Hespanhol, M. Porter, R. Vasudevan, and A. Aswani, "Sensor switching control under attacks detectable by finite sample dynamic watermarking tests," *IEEE Transactions on Automatic Control*, vol. 66, no. 10, pp. 4560-4574, 2021.

[14] J. Tian, R. Tan, X. Guan, Z. Xu, and T. Liu, "Moving target defense approach to detecting stuxnet-like attacks," *IEEE Transactions on Smart Grid*, vol. 11, no. 1, pp. 291-300, 2020.

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

[16] S. Lakshminarayana, E. V. Belmega, and H. V. Poor, "Moving-target defense against cyber-physical attacks in power grids via game theory," *IEEE Transactions on Smart Grid* , vol. 12, no. 6, pp. 5244-5257, 2021.

[17] S. Weerakkody, and B. Sinopoli, "A moving target approach for identifying malicious sensors in control systems," *2016 54th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, Monticello, IL, USA, Sept. 2016, pp. 1149-1156.

[18] J. Giraldo, A. Cardenas, and R.G. Sanfelice, "A moving target defense to detect stealthy attacks in cyber-physical systems," *American Control Conference (ACC)*, Philadelphia, PA, USA, Jul. 2019. pp. 391-396.

[19] A. Kanellopoulos, and K.G. Vamvoudakis, "A moving

target defense control framework for cyber-physical systems," *IEEE Transactions on Automatic Control*, vol. 65, no. 3, pp. 1029-1043, 2020.
[20] R. M. G. Ferrari, and A. M. H. Teixeira, "A switching multiplicative watermarking scheme for detection of stealthy cyber-attacks," *IEEE Transactions on Automatic Control*, vol. 66, no. 6, pp. 2558-2573, 2021.
[21] M. Ghaderi, K. Gheitasi, and W. Lucia, "A blended active detection strategy for false data injection attacks in cyber-physical systems," *IEEE Transactions on Control of Network Systems*, vol. 8, no. 1, pp. 168-176, 2021.
[22] S.S. Hassan, W. Kang, and C. Seon, "Unmanned aerial vehicle waypoint guidance with energy efficient path planning in smart factory," *2019 20th Asia-Pacific Network Operations and Management Symposium (APNOMS)*, Matsue, Japan, Sept. 2019. pp. 1-4.