2023 62nd IEEE Conference on Decision and Control (CDC)
December 13-15, 2023. Marina Bay Sands, Singapore

Optimal Linear Attack in Cyber-Physical Systems with Periodic Detection

Jia Qi, Chongrong Fang, and Jianping He

Abstract— **Security issues are of significant importance for cyber-physical systems (CPS), where the attack design is a major concern. Most related studies on attack design implicitly consider that the control period and detection period are the same. However, the two periods could be different in practical systems with remote detection such as supervisory control and data acquisition (SCADA) systems, which could lead to new vulnerabilities for attackers. In this paper, we consider the design of innovation-based linear attack strategies for CPS when the control period and detection period are inconsistent. Specifically, we propose an attack framework that consists of attack strategies for detection and non-detection instants under the period discrepancy. On this basis, we design the optimal stealthy innovation-based linear attack strategies for state estimation and LQG control to maximize the estimation error or control cost, respectively. Simulations are given to demonstrate the effectiveness of the proposed attack strategies.**

I. Introduction

Cyber-physical systems (CPS) refer to systems that tightly integrate physical and software components. It enables seamless computation, communication, control, and physical processes, resulting in the development of smarter, more efficient, and safer automated processes. It has attracted much attention in the past decades due to its wide applications such as intelligent power grids [1], transportation system [2], health care [3] and military [4]. However, CPS is vulnerable to malicious attacks, which can result in significant financial losses, utility disruptions, environmental damage, and even harm to human safety. In the past years, lots of malicious attacks on CPS happened such as the BlackEnergy attack on the power grid in Ukraine in 2015 and a ransomware attack on a US fuel pipeline in 2021.

Recently, CPS security issues have attracted great attention from the industry and academia, especially the system vulnerability analysis. Stealthy attack is a kind of malicious attack that can achieve a destructive impact on the control system while remaining stealthy to false data detectors. Lots of work on the stealthy attack has been done in the past few years. For instance, optimal stealthy attack for remote estimation in the sense of generating maximal error covariance has been investigated [5]–[7]. Specifically, Guo *et al.* [5] proposed an innovation-based linear attack to achieve the worst performance degradation while keeping stealthy to χ^2 detector for remote estimation. Li *et al.* [6] consider a group of reliable and unreliable sensors existing in the scenario of remote estimation, and by using the disclosure resources of reliable sensors and disruption resources of unreliable ones, a stealthy attack strategy is proposed. Some works study stealthy attacks for closed-loop systems [8]–[10]. For instance, Shang *et al.* [10] analyzed the impact on closed-loop LQG systems caused by a generic attack model for two different stealthiness constraints.

Despite the tremendously advanced results of the above works, almost all of them consider that the detection period is equal to the control period. However, in practical systems e.g., the supervisory control and data acquisition (SCADA) system, the discrepancy between the detection period and control period could exist due to practical limitations such as the discrepancy of communication rate or bandwidth constraints. A sketch of SCADA is shown in Fig. 1. The remote control center in the SCADA system periodically communicates with the local programmable logic controller (PLC). Usually, the false data detector is placed in the remote control center. The communication period between the control center and PLC is usually longer than that between the PLC and controlled plant [11]. This leads to the possible situation where attacks are out of the monitoring of false data detectors. The period discrepancy incurs new vulnerabilities that malicious attackers can utilize to achieve a more destructive impact while remaining undetected. For designing defense strategies against such attacks, it needs to investigate the system vulnerability under such period discrepancy.

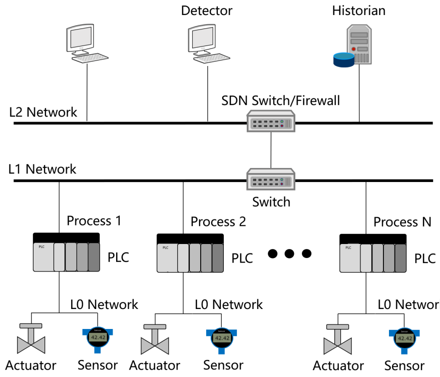


Fig. 1. A sketch of SCADA

To address the problem, we propose optimal innovation-based linear attack strategies for state estimation and LQG control. The main contributions of this paper are summarized

This work was supported by the National Natural Science Foundation of China under Grant 62103266.

J. Qi, C. Fang, and J. He are with the Dept. of Automation, Shanghai Jiao Tong University, the Key Laboratory of System Control and Information Processing, Ministry of Education of China, and Shanghai Engineering Research Center of Intelligent Control and Management, Shanghai, 200240, China. E-mails: {jia.qi, crfang, jphe}@sjtu.edu.cn

979-8-3503-0123-6/23/$31.00 ©2023 IEEE

as follows.

1) We propose an attack framework that consists of two attack modes for detection and non-detection time while considering the discrepancy between the control period and the remote detection period.
2) We present an optimal stealthy innovation-based linear attack strategy for state estimation under periodic detection, which maximizes the state estimation error while remaining stealthy.
3) We characterize the estimation error covariance under stealthy linear attack considering the periodic detection. On this basis, an optimal stealthy innovation-based linear attack strategy that maximizes the LQG controller cost with periodic detection is proposed.

The rest of the paper is organized as follows. Section II formulates the problem. The adversary model is given in section III. Section IV and V present the optimal attack strategy for estimation and LQG control, respectively. Simulation results are given in Section VI. Finally, we conclude our work in Section VII.

Notations: Denote the sets of natural numbers and real numbers as $\mathbb{N}$ and $\mathbb{R}$, respectively. $\mathbb{N}_+$ represents set of positive integers. $\mathbb{R}^n$ represents n-dimensional Euclidean space. $\{u_k\}_{k\in\mathbb{N}}$ denotes the sequence $\{u_0, u_1, u_2, \ldots\}$. $\mathbb{S}^n_+$ and $\mathbb{S}^n_{++}$ denote the sets of $n \times n$ positive semi-definite and positive definite matrices, respectively. Symmetrical matrices $X \geq 0$ $(X > 0)$ represent $X \in \mathbb{S}^n_+$ $(X \in \mathbb{S}^n_{++})$. $X \geq Y$ if $X - Y \in \mathbb{S}^n_+$. $\mathcal{N}(\mu, \Sigma)$ denotes Gaussian distribution with mean μ and covariance matrix Σ. For matrix X, X^T and $\mathrm{tr}(X)$ represent transposition and trace, respectively. For functions f_1 and f_2 with appropriate domain, $f_1 \circ f_2(\cdot)$ represents function composition $f_1(f_2(\cdot))$.

II. Problem Formulation

The system architecture is shown in Fig. 2, which contains classic components of a closed-loop LQG control system and a remote periodic detector.

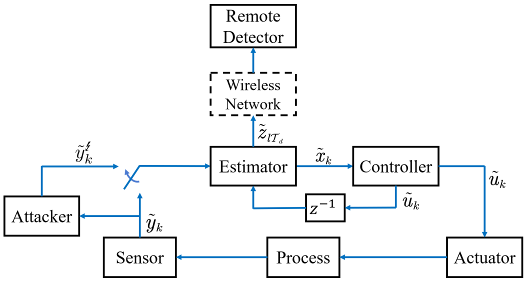


Fig. 2. System Architecture

A. System Model

Consider a discrete-time LTI system

$$x_{k+1} = Ax_k + Bu_k + w_k, \tag{1}$$

$$y_k = Cx_k + v_k. \tag{2}$$

where $x_k \in \mathbb{R}^n$ is the system state, $y_k \in \mathbb{R}^m$ is the system output, $u_k \in \mathbb{R}^q$ is the control input. The terms $w_k \in \mathbb{R}^n$ and $v_k \in \mathbb{R}^m$ represent process and measurement noise which are independent and identically distributed Gaussian with zero mean and covariance $Q \in \mathbb{S}^n_+$ and $R \in \mathbb{S}^m_{++}$, respectively. The initial state x_0 is Gaussian with zero mean and covariance $\Pi_0 \in \mathbb{S}^n_+$. Note that w_k, v_k, and x_0 are independent of each other. A is asymptotically stable. The pair (A, C) is detectable and $(A, \sqrt{Q})$ is stabilizable.

B. Estimator

The Kalman filter is employed as the estimator to estimate the state of the process. It recurs as follows.

$$\hat{x}^-_k = A\hat{x}_{k-1} + Bu_{k-1} \tag{3}$$

$$\hat{x}_k = \hat{x}^-_k + K_k z_k \tag{4}$$

$$P^-_k = AP_{k-1}A^T + Q \tag{5}$$

$$K_k = P^-_k C^T(CP^-_k C^T + R)^{-1} \tag{6}$$

$$P_k = (I - K_k C)P^-_k \tag{7}$$

where $\hat{x}^-_k$ and $\hat{x}_k$ denote the *a priori* and *a posteriori* minimum mean squared error (MMSE) estimates of x_k, respectively. P^-_k and P_k denote the corresponding error covariance, respectively. $z_k = y_k - C\hat{x}^-_k$ is called innovation. The recursion starts from $\hat{x}^-_0 = 0$ and $P^-_0 = \Pi_0$. It is well known that the Kalman filter converges exponentially fast from any initial condition [12]. The steady value of P^-_k, denoted as $\overline{P}$, can be represented by $\overline{P} \triangleq \lim_{k\to\infty} P^-_k$, where $\overline{P}$ is the unique semi-definite solution of $A[X - XC^T(CXC^T + R)^{-1}CX]A^T + Q = X$. To simplify our subsequent analysis, we assume the estimator starts from steady state $\Pi_0 = \overline{P}$, and adopts the fixed gain $K = \overline{P}C^T(C\overline{P}C^T + R)^{-1}$.

C. LQG Controller

The controller aims to obtain sequence $\{u_k\}_{k\in\mathbb{N}}$ to minimize the following control cost.

$$J = \lim_{N\to\infty} \frac{1}{N}\mathbb{E}\left[\sum_{i=0}^{N-1}(x_i^T W x_i + u_i^T U u_i)\right], \tag{8}$$

where $W \in \mathbb{S}^n_+$ and $U \in \mathbb{S}^q_{++}$ are the weight matrices. With dynamic programming [13], the optimal control input is $u_k = -L\hat{x}_k$ where L is determined by

$$L = (B^TSB + U)^{-1}B^TSA \tag{9}$$

$$S = A^TSA + W - A^TSB(B^TSB + U)^{-1}B^TSA. \tag{10}$$

D. Remote Periodic False Data Detector

This paper considers the periodic detector which periodically monitors the process. Denote detection and control period as $\mathcal{T}_d$ and $\mathcal{T}_c$, respectively. Assume $\mathcal{T}_d = h\mathcal{T}_c$, where $h \in \mathbb{N}_+$ is predetermined. W.l.o.g., assume detection instants are at $k = l\mathcal{T}_d$ for all $l \in \mathbb{N}$.

The innovation sequence z_k is a white Gaussian process with covariance $\mathcal{P} = C\overline{P}C^T + R$. The paper employs χ^2 detector, which uses the Chi-Squared test to determine whether the covariance of a variable exceeds a normal value or not. The detection criteria is given by

$$g_{l\mathcal{T}_d} = z_{j\mathcal{T}_d}^T \mathcal{P}^{-1} z_{j\mathcal{T}_d} \underset{H_1}{\overset{H_0}{\lessgtr}} \delta \tag{11}$$

where δ is the threshold predetermined by a normal false alarm rate. Note that the term $\mathcal{P}^{-1}$ in (11) normalizes the square of the steady innovation. The null hypothesis H_0 and alternative hypothesis H_1 represent that the process operates in a normal state and is compromised by attacks, respectively. $g_{l\mathcal{T}_d}$ follows a χ^2 distribution with m degrees of freedom, and an alarm is triggered if $g_{l\mathcal{T}_d}$ is greater than δ.

E. Motivating Example

Next, we show attackers can utilize the discrepancy between the control period and the detection period to achieve a more destructive impact while remaining undetected.

Consider the system architecture in Fig. 2. Let $n = 1$, $m = 1$, $A = 0.8$, $C = 1.2$, $Q = 1$ and $R = 1$, detection period $\mathcal{T}_d = 10$, and control period $\mathcal{T}_c = 1$. Simulation time interval is $[0, 120]$, and attacks happen during time inteval $[10, 100]$. Consider attack strategy 1: $\tilde{z}_k = -z_k$, which is proved to be the optimal stealthy linear attack for state estimation case [5]. Attack strategy 2: $\tilde{z}_k = -z_k$ when $k = l\mathcal{T}_d$ for all integer $l \in \{i \in \mathbb{N} : 1 \leq i \leq 10\}$, and $\tilde{z}_k = -2z_k$ when $k \neq l\mathcal{T}_d$ for any k in the interval $[10, 100]$. Note that attack strategy 2 is also stealthy. The initial condition is $\hat{x}_0^- = 0$ and $P_0^- = 0$. The simulation result is shown in Fig 3. P_k and $\tilde{P}_k$ represents the estimation error variance under attack strategies 1 and 2, respectively. Fig 3 shows that the attack strategy 2 achieves a more destructive impact than attack strategy 1.

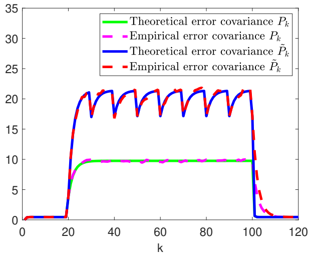


Fig. 3. Estimation error covariance under attack strategy 1 and 2

F. Problem of Interest

Considering the system model, the estimator, the LQG controller, and the periodic detector given in Section II, the main problems we are interested in are as follows.

1) What is the most destructive impact that the attacker can achieve on the system with periodic detection?
2) What are the optimal attack strategies for state estimation and LQG control with periodic detection?

III. Adversary Model

Knowledge: the attacker knows A, B, C, Q, R, K, W, U and the detection instants, and has access to the output signal of the sensor and input signal received by the process.

Capabilities: the attacker can modify the output signal transmitted between the sensor and the estimator.

Assume the system is compromised by the attacks from instant $k_a \geq 0$, and the attacks are continuously added to the system. The attacker makes the following malicious change:

$$\tilde{y}_k \mapsto \tilde{y}_k^{\lightning} \tag{12}$$

where $\tilde{y}_k$ denotes the sensor measurement of the compromised system, and $\tilde{y}_k^{\lightning}$ is the manipulated output of the attacker. At $k = 0$, (12) becomes $y_0 \mapsto \tilde{y}_0^{\lightning}$, where y_0 is the output of the nominal system. With attack (12), the state estimation equations in the estimator become

$$\tilde{x}_k^- = A\tilde{x}_{k-1} + B\tilde{u}_{k-1} \tag{13}$$

$$\tilde{x}_k = \tilde{x}_k^- + K\tilde{z}_k, \tag{14}$$

where $\tilde{x}_k^-$ and $\tilde{x}_k$ denote the *a priori* and *a posteriori* state estimates of the compromised system, respectively. Because $\tilde{x}_0^-$ is not corrupted, we have $\tilde{x}_0^- = \hat{x}_0^-$. $\tilde{z}_k$ is the affected innovation given by

$$\tilde{z}_k = \tilde{y}_k^{\lightning} - C\tilde{x}_k^-. \tag{15}$$

$\tilde{u}_k$ is the control input of the compromised system given by $\tilde{u}_k = -L\tilde{x}_k$. Denote the state of the compromised system at instant k as x_k^*. It recurs as follows.

$$x_{k+1}^* = Ax_k^* + B\tilde{u}_k + w_k, \tag{16}$$

which starts from $x_0^* = x_0$.

Define the following *a priori* and *a posteriori* estimates:

$$\breve{x}_k^- = \mathbb{E}[x_k^* | \tilde{y}_{0:k-1}, \tilde{u}_{0:k-1}],$$

$$\breve{x}_k = \mathbb{E}[x_k^* | \tilde{y}_{0:k}, \tilde{u}_{0:k-1}].$$

They recur as follows

$$\breve{x}_k^- = A\breve{x}_{k-1} + B\tilde{u}_{k-1} \tag{17}$$

$$\breve{x}_k = \breve{x}_k^- + K(\tilde{y}_k - C\breve{x}_k^-) \tag{18}$$

with $\breve{x}_0^- = \hat{x}_0^-$. $\breve{z}_k = \tilde{y}_k - C\breve{x}_k^-$ is the innovation.

Note that modifying $\tilde{y}_k$ is equivalent to modifying $\breve{z}_k$ from [14], and thus we design attack strategy based on the innovation sequance $\breve{z}_k$ in the subsequent discusstion. The paper considers in this initial study the linear attack strategies where the attacker makes an affine transformation of $\breve{z}_k$:

$$\tilde{z}_k = T_k\breve{z}_k + b_k \tag{19}$$

where $T_k \in \mathbb{R}^{m\times m}$, and $b_k \sim \mathcal{N}(0, \mathcal{L}_k)$ is a Gaussian random variable independent of $\breve{z}_k$.

We consider different attack parameters in detection and non-detection instants:

$$T_k = \overline{T}, \mathcal{L}_k = \overline{\mathcal{L}} \text{ for } k \in \{l\mathcal{T}_d : l \in \mathbb{N}\}, \tag{20}$$

$$T_k = T, \mathcal{L}_k = \mathcal{L} \text{ for } k \in \mathbb{N} \setminus \{l\mathcal{T}_d : l \in \mathbb{N}\}, \tag{21}$$

where $\overline{T}, T \in \mathbb{R}^{m\times m}$, and $\overline{\mathcal{L}}, \mathcal{L}$ are positive definite matrices.

Reasonably, we consider that the attacker needs to satisfy the following energy constraint because its energy is limited.

$$\mathbb{E}[\tilde{z}_k^T \tilde{z}_k] \leq \eta, \tag{22}$$

where η is a predetermined constant representing the upper bound of the energy of the attacker.

IV. Optimal linear attack on state estimation

This section designs the optimal linear attack for state estimation in the sense of generating the maximal trace of estimation error covariance. We first provide the estimation error covariance evolution in the compromised estimator for non-detection instants.

A. Error Covariance Evolution for Non-detection Instants

Denote estimation error covariance of the estimator with attack (19) as $\tilde{P}_k$. The error covariance evolution for non-detection instants is shown in the following lemma.

Lemma 1: Consider the system modeled by (1)–(2), the estimator (3)–(7) and the attack (19)–(21), for non-detection instants $k \in \mathbb{N} \setminus \{l\mathcal{T}_d : l \in \mathbb{N}\}$, $\tilde{P}_k$ follows the recursion

$$\begin{aligned}\tilde{P}_k =& A\tilde{P}_{k-1}A^T + Q + K\tilde{\Sigma}K^T \\ & - KTC\overline{P} - \overline{P}C^T T^T K^T \end{aligned} \tag{23}$$

where $\tilde{\Sigma} = T\Sigma^{-1}T^T + \mathcal{L}$ and $\Sigma = (C\overline{P}C^T + R)^{-1}$.

The proof of the lemma is similar to [14, Lemma 1], and omitted due to space constraints.

B. Optimal Attack Design

We first give the attack feasibility constraint. Since $x_0^* = x_0$ and $\breve{x}_0^- = \hat{x}_0^-$, we have $\mathbb{E}[(x_0^* - \breve{x}_0^-)(x_0^* - \breve{x}_0^-)^T] = \overline{P}$. Thus, $\mathbb{E}[(x_k^* - \breve{x}_k^-)(x_k^* - \breve{x}_k^-)^T] = \overline{P}$ for all $k > 0$. For $k \in \{l\mathcal{T}_d : l \in \mathbb{N}\}$, the covariance of $\tilde{z}_k$ is $\overline{T}\mathcal{P}\overline{T}^T + \overline{\mathcal{L}}$. Thus, for the linear attack to be undetected, the constraint

$$\mathcal{P} = \overline{T}\mathcal{P}\overline{T}^T + \overline{\mathcal{L}} \tag{24}$$

need to be satisfied from (11). For non-detection instants $k \in \mathbb{N} \setminus \{l\mathcal{T}_d : l \in \mathbb{N}\}$, the linear attack is not necessary to meet the stealth requirement. On this basis, we present the optimal attack strategy in the following theorem.

Theorem 1: Consider the system modeled by (1)–(2), the estimator (3)–(7) and the attack strategy (19)–(21) with energy constraint (22), the optimal attack parameters at detection instants are $\overline{T} = -I$, $\overline{\mathcal{L}} = 0$. At non-detection instants, the optimal attack parameters are given by the following optimization problem.

$$\begin{aligned}\min_{T,\tilde{\Sigma}\in\mathbb{R}^{m\times m}} \quad & -\operatorname{tr}(K\tilde{\Sigma}K^T) + 2\operatorname{tr}(KTC\overline{P}) \\ \text{s.t.} \quad & \operatorname{tr}(\tilde{\Sigma}) \leq \eta \\ & \tilde{\Sigma} \geq 0 \\ & \begin{bmatrix} \Sigma & T^T \\ T & \tilde{\Sigma} \end{bmatrix} \geq 0 \end{aligned}$$

Proof: Because the detector only uses the innovation sequence to detect anomalies, the attack parameters for detection and non-detection instants can be designed separately. From the existing work [5], the optimal linear attack at detection instants is $\overline{T} = -I$ and $\overline{\mathcal{L}} = 0$.

Next, we analyze the optimal attack parameters for those instants of non-detection, i.e., $k \in \mathbb{N} \setminus \{l\mathcal{T}_d : l \in \mathbb{N}\}$. From the covariance evolution (23), the trace of $\tilde{P}_k$ is

$$\operatorname{tr}(\tilde{P}_k) = \operatorname{tr}(A\tilde{P}_{k-1}A^T + Q) + \operatorname{tr}(K\tilde{\Sigma}K^T) - 2\operatorname{tr}(KTC\overline{P}).$$

Thus, maximizing $\operatorname{tr}(\tilde{P}_k)$ is equivalent to minimizing the term $-\operatorname{tr}(K\tilde{\Sigma}K^T) + 2\operatorname{tr}(KTC\overline{P})$.

The energy constraint (22) can be transformed as follows.

$$\mathbb{E}[(\tilde{z}_k)^T\tilde{z}_k] = \mathbb{E}[(T\breve{z}_k + b_k)^T(T\breve{z}_k + b_k)] = \operatorname{tr}(\tilde{\Sigma}) \leq \eta.$$

It is clear that $\tilde{\Sigma} = T\Sigma^{-1}T^T + \mathcal{L} \geq 0$. Since $\mathcal{L} = \tilde{\Sigma} - T\Sigma^{-1}T^T \geq 0$, the constraint can be described by linear matrix inequality using Schur complement $\begin{bmatrix} \Sigma & T^T \\ T & \tilde{\Sigma} \end{bmatrix} \geq 0$.

After the optimization problem has been solved, the $\mathcal{L}$ is given by $\mathcal{L} = \tilde{\Sigma} - T\Sigma^{-1}T^T$. This ends the proof. ■

Remark 1: The optimization problem in Theorem 1 is an SDP problem, which can be solved by CVX toolbox [15].

V. Optimal linear attack on LQG controller

This section analyzes the maximal impact of the linear attack on the LQG controller. The objective of the attacker is to maximize the infinite-horizon LQG controller cost:

$$\tilde{J} = \lim_{N\to\infty} \frac{1}{N}\mathbb{E}\left[\sum_{i=0}^{N-1}\left((x_i^*)^T W x_i^* + \tilde{u}_i^T U \tilde{u}_i\right)\right]$$

A. Performance Index with Attack

Compared with the evolution of estimation error covariance at non-detection instants, the stealthiness constraint (24) needs to be satisfied at detection instants. From (23), (24) and $K = \overline{P}C^T\Sigma$, the evolution of estimation error covariance at detection instants is given as

$$\tilde{P}_k = A\tilde{P}_{k-1}A^T + Q + \overline{P}C^T(\Sigma - \overline{T}^T\Sigma - \Sigma\overline{T})C\overline{P} \tag{25}$$

for $k \in \{l\mathcal{T}_d : l \in \mathbb{N}\}$.

For notational brevity, denote the evolution of state estimation error covariance (25) and (23) as

$$\tilde{P}_k = f(\tilde{P}_{k-1}, \overline{T}, \overline{\mathcal{L}}) \text{ for } k \in \{l\mathcal{T}_d : l \in \mathbb{N}\}, \tag{26}$$

$$\tilde{P}_k = g(\tilde{P}_{k-1}, T, \mathcal{L}) \text{ for } k \in \mathbb{N} \setminus \{l\mathcal{T}_d : l \in \mathbb{N}\}, \tag{27}$$

respectively, where f and g are proper mapping functions. Denote $g^{k+1}(\cdot, T, \mathcal{L}) = g[g^k(\cdot, T, \mathcal{L}), T, \mathcal{L}]$ for any integer $k \in \mathbb{N}_+$. According to (26)–(27), $\tilde{P}_k$ is affected by $\tilde{P}_{k-1}$ and attack parameters at both detection and non-detection instants. Thus, $\tilde{P}_k$ for all $k \in \mathbb{N}$ can be obtained by the combination of (26) and (27).

We next transform the LQG controller cost $\tilde{J}$, by Lemmas 2–4, into a form that is convenient to be analyzed.

Lemma 2: The LQG controller cost under the linear attack (19)–(21) is given as follows.

$$\tilde{J} = \operatorname{tr}(SQ) + \lim_{N\to\infty}\frac{1}{N}\operatorname{tr}\left[\sum_{k=0}^{N-1}\left(W + A^TSA - S\right)\tilde{P}_k\right] \tag{28}$$

Proof: According to [10, Lemma 1], for the infinite-horizon LQG controller cost, we have

$$\tilde{J} = \operatorname{tr}(SQ) + \lim_{N\to\infty}\frac{1}{N}\operatorname{tr}\left[\sum_{k=0}^{N-1}A^TSBL\tilde{P}_k\right],$$

and (28) holds due to (9)–(10). ■

Lemma 3: Consider the system modeled by (1)–(2), the estimator (3)–(7) and the attack (19)–(21), the estimation error covariance evolves periodically as k goes to infinity:

$$\lim_{k\to\infty} \tilde{P}_k = \lim_{k\to\infty} \tilde{P}_{k+\mathcal{T}_d}. \tag{29}$$

Proof: In the proof, without loss of generality, we consider $k_a = 0$, and the other case that $k_a > 0$ can be analyzed similarly.

Since $\tilde{P}_{i+n\mathcal{T}_d} = g^i(\tilde{P}_{n\mathcal{T}_d}, T, \mathcal{L})$ for any integer $i \in (0, \mathcal{T}_d)$ and any $n \in \mathbb{N}$, without loss of generality, we only analyze the convergence of sequence $\{\tilde{P}(n)\}_{n\in\mathbb{N}} \triangleq \{\tilde{P}_{n\mathcal{T}_d} : n \in \mathbb{N}\}$. And we have

$$\begin{aligned}&\tilde{P}(0) = \tilde{P}_0,\\ &\tilde{P}(n) = f \circ g^{\mathcal{T}_d-1}[\tilde{P}(n-1), T, \mathcal{L}]\end{aligned}$$

for $n \in \mathbb{N}_+$, where $\tilde{P}_0 = \overline{P} + \overline{P}C^T(\Sigma - \overline{T}^T\Sigma - \Sigma\overline{T})C\overline{P}$ according to (25). From (23) and (25), we have

$$\tilde{P}(n) = A^{\mathcal{T}_d}\tilde{P}(n-1)(A^T)^{\mathcal{T}_d} + H,$$

where $H = \sum_{i=1}^{\mathcal{T}_d-1} A^i(Q + K\tilde{\Sigma}K^T - KTC\overline{P} - \overline{P}C^TT^TK^T)(A^T)^i + Q + \overline{P}C^T(\Sigma - \overline{T}\Sigma - \Sigma\overline{T})C\overline{P}$. Note that H and $\tilde{P}_0$ are constant matrices for any predetermined attack parameters.

We next prove the sequence $\{\tilde{P}(n)\}_{n\in\mathbb{N}}$ converges. Denote $\tilde{P}(n)$ with n going to infinity as $\tilde{P}(\infty)$.

$$\begin{aligned}\tilde{P}(\infty) &= \lim_{n\to\infty} \tilde{P}(n)\\ &= \lim_{n\to\infty} A^{n\mathcal{T}_d}\tilde{P}_0(A^T)^{n\mathcal{T}_d} + \sum_{i=0}^{\infty} A^{i\mathcal{T}_d}H(A^T)^{i\mathcal{T}_d}\\ &= \sum_{i=0}^{\infty} A^{i\mathcal{T}_d}H(A^T)^{i\mathcal{T}_d}\end{aligned}$$

where we use the fact that A is asymptotically stable in the last equality. Then, according to [12, Lemma 2.1], $\tilde{P}(\infty)$ exists and is finite. Multiply $A^{\mathcal{T}_d}$ and $(A^T)^{\mathcal{T}_d}$ on the left and right side of $\tilde{P}(\infty)$, we have

$$A^{\mathcal{T}_d}\tilde{P}(\infty)(A^T)^{\mathcal{T}_d} = \tilde{P}(\infty) - H. \tag{30}$$

Based on this we will prove the uniqueness of $\tilde{P}(\infty)$. Assume $\tilde{P}'(\infty)$ is another solution of equation (30).

$$\tilde{P}'(\infty) = A^{\mathcal{T}_d}\tilde{P}'(\infty)(A^T)^{\mathcal{T}_d} + H.$$

Subtract $\tilde{P}'(\infty)$ from $\tilde{P}(\infty)$, yields

$$\tilde{P}(\infty) - \tilde{P}'(\infty) = A^{\mathcal{T}_d}[\tilde{P}(\infty) - \tilde{P}'(\infty)](A^T)^{\mathcal{T}_d}$$

from which for all k

$$\begin{aligned}&A^{(k-1)\mathcal{T}_d}[\tilde{P}(\infty) - \tilde{P}'(\infty)](A^T)^{(k-1)\mathcal{T}_d}\\ &\qquad = A^{k\mathcal{T}_d}[\tilde{P}(\infty) - \tilde{P}'(\infty)](A^T)^{k\mathcal{T}_d}.\end{aligned}$$

Adding such relations,

$$\tilde{P}(\infty) - \tilde{P}'(\infty) = A^{k\mathcal{T}_d}[\tilde{P}(\infty) - \tilde{P}'(\infty)](A^T)^{k\mathcal{T}_d}.$$

It follows that $\tilde{P}(\infty) = \tilde{P}'(\infty)$ as k goes to infinity. The uniqueness of $\tilde{P}(\infty)$ is proved. Thus the convergence of $\{\tilde{P}(n)\}_{n\in\mathbb{N}}$ is proved, and this guarantees that (29) holds, which ends the proof. ■

We next transform the LQG controller cost with the attack in infinite-horizon into that in a detection period $\mathcal{T}_d$. From Lemma 3 and the boundedness of the finite iterations of (23), $\left\{\sum_{j=0}^{\mathcal{T}_d-1} \tilde{P}_{k+j}\right\}_{k\in\mathbb{N}}$ is convergent. Furthermore, define $\overline{S} = W + A^TSA - S$, and from (28), we have

$$\begin{aligned}\tilde{J} &= \mathrm{tr}(SQ) + \lim_{v\to\infty}\frac{1}{v\mathcal{T}_d}\,\mathrm{tr}\left(\sum_{i=0}^{v}\overline{S}\sum_{j=0}^{\mathcal{T}_d-1}\tilde{P}_{i\mathcal{T}_d+j}\right)\\ &= \mathrm{tr}(SQ) + \frac{1}{\mathcal{T}_d}\lim_{v\to\infty}\mathrm{tr}\left(\overline{S}\sum_{j=0}^{\mathcal{T}_d-1}\tilde{P}_{v\mathcal{T}_d+j}\right)\\ &= \mathrm{tr}(SQ) + \frac{1}{\mathcal{T}_d}\,\mathrm{tr}\left\{\overline{S}\left[\tilde{P}(\infty) + \sum_{i=1}^{\mathcal{T}_d-1} g^i\left(\tilde{P}(\infty), T, \mathcal{L}\right)\right]\right\}\end{aligned} \tag{31}$$

where we use Stolz–Cesàro theorem in the second equality.

B. Optimal Attack Design

Define $G(k) = g^k(\tilde{P}(\infty), T, \mathcal{L})$ and $G'(k) = g^k(\tilde{P}(\infty), T', \mathcal{L}')$, where $T' \in \mathbb{R}^{m\times m}$ and $\mathcal{L}' \in \mathbb{R}^{m\times m}$ are predetermined matrices. Besides, $\mathcal{L}'$ is positive semi-definite. Define $\Delta G(k) = G(k) - G'(k)$ for $k \in \mathbb{N}$.

Lemma 4: If the initial condition $\Delta G(1) > 0$, then $\Delta G(k) > 0$ for all $k \in \mathbb{N}_+$.

Next, we provide the optimal linear attack strategy for LQG control by the following theorem.

Theorem 2: Consider the system modeled by (1)–(2), the estimator (3)–(7) and the attack strategy (19)–(21) with energy constraint (22), the optimal attack parameters at detection instants are $\overline{T} = -I$, $\overline{\mathcal{L}} = 0$, At non-detection instants, the optimal attack parameters are given by the following optimization problem.

$$\begin{aligned}\min_{T,\tilde{\Sigma}\in\mathbb{R}^{m\times m}}\quad & \mathrm{tr}[-\overline{S}(K\tilde{\Sigma}K^T - KTC\overline{P} - \overline{P}C^TT^TK^T)]\\ \text{s.t.}\quad & \mathrm{tr}(\tilde{\Sigma}) \le \eta\\ & \tilde{\Sigma} \ge 0\\ & \begin{bmatrix}\Sigma & T^T\\ T & \tilde{\Sigma}\end{bmatrix} \ge 0\end{aligned}$$

Due to the space limitations, the proofs of Lemma 4 and Theorem 2 are given in the technical report [16].

Remark 2: The optimization problem in Theorem 2 is an SDP problem, which can be solved by CVX toolbox [15].

VI. Simulation results

In this section, we present simulation examples to validate the result of this paper. Consider a system with parameters

$$A = \begin{bmatrix}0.7 & 0.2\\ 0.05 & 0.64\end{bmatrix}, B = \begin{bmatrix}1\\ 1\end{bmatrix}, C = \begin{bmatrix}0.5 & -0.8\\ 0 & 0.7\end{bmatrix},$$

$$Q = \begin{bmatrix}0.5 & 0\\ 0 & 0.7\end{bmatrix}, R = \begin{bmatrix}1 & 0\\ 0 & 0.8\end{bmatrix}, W = \begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}, U = 1.$$

Let detection period $\mathcal{T}_d = 10$, control period $\mathcal{T}_c = 1$, and energy threshold $\eta = 4$.

A. Estimation under Attack

Consider the optimal linear attack to the estimation. The optimal attack parameters given by Theorem 1 are

$$T = \begin{bmatrix} -0.6115 & 0.4665 \\ 0.4665 & -1.2552 \end{bmatrix}, \mathcal{L} = 0, \overline{T} = -I, \overline{\mathcal{L}} = 0. \quad (32)$$

The simulation results of the estimation error covariance under different attacks are shown in Fig. 4(a). The simulation iterates from $P_0^- = 0$ and $\hat{x}_0^- = 0$. During time $[20, 60]$ and $[80, 120]$, the system is attacked. We compare our results with the optimal linear attack for remote estimation [5] where attack parameters are $T_k = -I$ and $\mathcal{L}_k = 0$ for all $k \in \mathbb{N}$, and the normal estimation error covariance of the Kalman filter. Despite the advanced result given by [5], it can not achieve optimality with periodic detection.

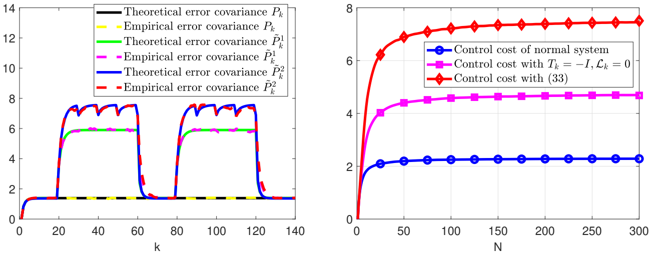
We use P_k, $\tilde{P}_k^1$ and $\tilde{P}_k^2$ to represent the estimation error covariance of the normal Kalman filter, under the linear attack strategy given in [5] and under attack parameters given in (32), respectively. The results for empirical error covariance in Fig. 4(a) are averaged over 10000 Monte Carlo simulations. When the attacker begins compromising the system at instant 20 and 80, $\tilde{P}_k^1$ and $\tilde{P}_k^2$ increase, and $\tilde{P}_k^2$ is larger than $\tilde{P}_k^1$, which shows that the attack strategy proposed in this paper achieves maximal error covariance.

B. LQG Control under Attack

Consider the optimal linear attack to LQG control. The optimal attack parameters given by Theorem 2 are

$$T = \begin{bmatrix} -0.0003 & 0.0212 \\ 0.0212 & -1.7828 \end{bmatrix}, \mathcal{L} = 0, \overline{T} = -I, \overline{\mathcal{L}} = 0. \quad (33)$$

The simulation time interval is $[0, 300]$. The system is compromised from the initial instant 0, and the simulation result is shown in Fig. 4(b). The longitudinal coordinate represents $J(N) = \frac{1}{N}\left[\sum_{i=0}^{N-1}(x_i^T W x_i + u_i^T U u_i)\right]$. As N goes to infinity, $J(N)$ converges to the infinite-horizon LQG controller cost. The results in Fig. 4(b) are averaged over 10000 Monte Carlo simulations. The blue circle-mark line, the pink square-mark line, and the red diamond-mark line represent $J(N)$ of the normal system, under linear attack with $T_k = -I$, $\mathcal{L}_k = 0$ for all $k \in \mathbb{N}$, and under linear attack with parameters in (33), respectively. We can see from Fig. 4(b) that the red diamond-mark line is larger than other lines, which shows the optimal attack strategy to LQG control given in this paper leads to maximal LQG controller cost.



(a) Estimation error covariance

(b) LQG controller cost

Fig. 4. Simulation results

VII. Conclusion

This paper considers the design of optimal stealthy innovation-based linear attack to state estimation and LQG control in CPS where a remote periodic detector is employed to detect anomalies. For state estimation, we illustrate the attack strategy under instants of detection and non-detection which can be designed separately. Then, we give an optimal stealthy linear attack strategy by solving a convex optimization problem. For LQG control, we first rigorously prove that the LQG cost can be analyzed within a detection period as time goes to infinity. Further, we prove that the attack strategy at detection instants is the same as that for state estimation. At non-detection instants, we formulate an optimization problem to obtain the attack parameters. Simulations are provided to show the effectiveness of the proposed optimal attack strategies. Future directions include designing defense strategies for the proposed attacks.

References

[1] Y. Liu, P. Ning, and M. K. Reiter, "False data injection attacks against state estimation in electric power grids," *ACM Transactions on Information and System Security (TISSEC)*, vol. 14, no. 1, pp. 1–33, 2011.

[2] G. Xiong, F. Zhu, X. Liu, X. Dong, W. Huang, S. Chen, and K. Zhao, "Cyber-physical-social system in intelligent transportation," *IEEE/CAA Journal of Automatica Sinica*, vol. 2, no. 3, pp. 320–333, 2015.

[3] Y. Zhang, M. Qiu, C.-W. Tsai, M. M. Hassan, and A. Alamri, "Health-cps: Healthcare cyber-physical system assisted by cloud and big data," *IEEE Systems Journal*, vol. 11, no. 1, pp. 88–95, 2017.

[4] M. Vierhauser, J. Cleland-Huang, S. Bayley, T. Krismayer, R. Rabiser, and P. Grünbacher, "Monitoring cps at runtime - a case study in the uav domain," in *2018 44th Euromicro Conference on Software Engineering and Advanced Applications (SEAA)*, 2018, pp. 73–80.

[5] Z. Guo, D. Shi, K. H. Johansson, and L. Shi, "Optimal linear cyber-attack on remote state estimation," *IEEE Transactions on Control of Network Systems*, vol. 4, no. 1, pp. 4–13, 2017.

[6] Y. Li, Y. Yang, Z. Zhao, J. Zhou, and D. E. Quevedo, "Deception attacks on remote estimation with disclosure and disruption resources," *IEEE Transactions on Automatic Control*, 2022.

[7] J. Zhou, J. Shang, and T. Chen, "Optimal deception attacks on remote state estimators equipped with interval anomaly detectors," *Automatica*, vol. 148, p. 110723, 2023.

[8] R. Zhang and P. Venkitasubramaniam, "Stealthy control signal attacks in linear quadratic gaussian control systems: Detectability reward tradeoff," *IEEE Transactions on Information Forensics and Security*, vol. 12, no. 7, pp. 1555–1570, 2017.

[9] C. Fang, Y. Qi, J. Chen, R. Tan, and W. X. Zheng, "Stealthy actuator signal attacks in stochastic control systems: Performance and limitations," *IEEE Transactions on Automatic Control*, vol. 65, no. 9, pp. 3927–3934, 2019.

[10] J. Shang, D. Cheng, J. Zhou, and T. Chen, "Asymmetric vulnerability of measurement and control channels in closed-loop systems," *IEEE Transactions on Control of Network Systems*, vol. 9, no. 4, pp. 1804–1815, 2022.

[11] R. Krutz, *Securing SCADA Systems*. New York, NY, USA: Wiley Publishing, Inc., 2005.

[12] B. D. Anderson and J. B. Moore, *Optimal filtering*. Courier Corporation, 2012.

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

[14] Z. Guo, D. Shi, K. H. Johansson, and L. Shi, "Worst-case stealthy innovation-based linear attack on remote state estimation," *Automatica*, vol. 89, pp. 117–124, 2018.

[15] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 2.1," http://cvxr.com/cvx, Mar. 2014.

[16] J. Qi, C. Fang, and J. He, "Optimal linear attack in cyber-physical systems with periodic detection," [Online].Available:https://iwin-fins.com/wp-content/uploads/2023/09/CDC2023qi.pdf, 2023.