IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received November 26, 2020, accepted December 2, 2020, date of publication January 5, 2021, date of current version January 13, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3049228

A Controllable False Data Injection Attack for a Cyber Physical System

JANGHOON YANG, (Member, IEEE)
Department of New Media, Seoul Media Institute of Technology, Seoul 07590, South Korea

e-mail: jhyang@smit.ac.kr

This work was supported by the Basic Science Research Program through the National Research Foundation of Korea funded by the Ministry of Science and ICT under Grant NRF-2017R1A2B4007398.

ABSTRACT With the evolution of the Internet of Things (IoT), various types of devices and massive systems comprising national infrastructures such as a smart grid are connected on a network, which poses various types of security issues in a cyber physical system. In this paper, we propose two false data injection attacks, which are on the forward path and the feedback path of a control system. Both are designed with a controllable parameter which determines the degree of degradation. A defensive method of inversing a linear forward attack through estimating with least square or minimum mean squared method was developed. A conventional Kalman filter was considered as a defensive method for a noise injection attack on the feedback path. The numerical evaluation verifies that the parameters of the proposed attacks control the degree of performance degradation of the control system, and the proposed defenses can effectively defend the proposed attacks.

INDEX TERMS Cyber physical system, false data injection, security, cyber physical system attack, linear quadratic Gaussian control.

I. INTRODUCTION

With the evolution of Internet and mobile technology, everything is connected on a network. While a plant, a controller, and sensors in a conventional control system are usually co-located, separate development of them, and high speed communication over the network accelerate the widespread use of a networked controlled system (NCS) or cyber physical system (CPS) [1]. However, due to the intrinsic nature of connectivity in the network, the control system is vulnerable to a cyber attack. In addition to the attack in a network layer, there can be a physical layer attack which modifies a control input or measurements from sensors. A combined network layer and physical layer attack in the control system is usually called as "CPS attack" since it is usually targeted to degrade the physical performance of the plant through the attack to the controller on the network. The security issue associated with the CPS attack is very important in the sense that the attack can result in a tremendously disastrous effect on national infrastructures such as a power grid, water distribution, a nationwide traffic network, and military operation. The event of Stuxnet which degraded the production performance of the uranium facility covertly [2] arouses the importance of the security of CPS. It is also very important in industrial applications such as smart factory, cloud robotics, and autonomous vehicles [3], since the attacks can result in casualties in human lives and huge economic loss.

To deal with security issues in CPS, a significant amount of research has been done in recent years. CPS attacks may be classified into Denial-of-Service (DoS), Service Degradation (SD) (compromised-key attack and man-in-the-middle(MitM attack), Cyber-physical Intelligence (CPI) (Eavesdropping and system identification attack) [4], [5]. The detectability and identifiability of CPS attack were studied with defining various types of CPS attacks [6]. Depending on the available information, the types of attack and the effect of attack may be different. With the assumption of chi-square failure detection and perfect knowledge of a control system at the attacker, a sufficient and necessary condition for being perfectly attackable was provided [7]. When the attacker does not have full knowledge of the system, it may try to refine it through learning so that it can carry out more sophisticated attack [8]. The fundamental tradeoffs between the system

The associate editor coordinating the review of this manuscript and approving it for publication was Shihong Ding.

performance and security were quantified through modeling a dynamical system with Markov decision process (MDP) which can be seen as partially observable MDP (POMDP) in the perspective of the attacker with limited access to the system [9]. This approach makes it possible for the attacker to find an optimal adversarial policy for the class of stochastic control system. A simple real-time testbed implemented in RTDS and OPNET for the MitM attack showed the potential of providing a realistic testing environment for the CPS attacks [10].

The various methods for the defense to the attacks were also proposed from robust control theory. A sufficient condition for desired security requirement for the FDIA and the derivation of the control gain through some linear matrix inequalities were established with the introduction of the mean square security domain [11]. A receding horizon control of which horizon depended on an attack was proposed to be resilient to the replay attack [12]. A robust sliding mode control was developed for a class of Markovian jump system to deal with FDIA into control signals and peak-bounded external disturbances [13]. To detect the presence of CPS attacks, transformations such as the graph Fourier transform [14] and cross wavelet transform [15] were exploited A control system was transformed into a switched auxiliary system to develop a covert attack detection system using a switched Luenberger observer [16].

Among many different CPS attacks, FDIA can be considered to be one of the most prominent attacks [5]. The FDIA injects a modified control input [13] or modified measurement [17], [18] or both [8] to degrade the performance of a control system. Even though FDIA was considered as a specific deception attack in the sense that the FDIA as the attack to the state estimator [6], MitM, FDIA, a replay attack and a deception attack [11] can be regarded as the attack of the same class. A heuristic MitM attack with system identification attack which changed the scale of the control input was shown to increase the stationary error [4]. The statistical-duplicate attack which deceived the detector with fictitious measurement having the same statistical property as the one without attack was proposed to execute the attack covertly [7]. To improve the state estimation in the presence of FDIA on measurement data, deep learning with a generative adversarial network (GAN) was also introduced [17]. Attack detector and the intermittent message authentication were proposed to improve robust control for CPS operating on the resource-constrained network in the presence of the FDIA at the measurement [18].

Many of existing FDIAs can be considered as a covert attack which degrades the performance of the control system without being detected as an attack [8]. There can be many different ways to achieve the covert attack. It can be broadly classified into two types, a pretender covert attack, and a stealth covert attack. The pretender covert attack degrades the performance with modified control input while replacing the measurement output with the one predicted from the model without attack. The stealth attack usually tries to degrade the performance gradually or marginally without replacing the measurement. A method of detecting an attack from the error statistics of the output and compensating the controller was developed to deal with the covet attack [19].

As more and more physical systems such as smart city, and smart factory get connected, the importance of CPS security is growing significantly. Thus, the various potential attacks and associated defenses need to be studied further. To this end, CPS attacks in both the forward path and the feedback path at a control system are considered in the perspective of FDIA while existing research focuses on either one not both. Unlike many existing covert attacks which try to degrade the performance of a control system by considering a specific detector or replacing the measurements with ones without attack, we propose controllable covert attacks with a single parameter so that they can slow down the stabilization of the system. In the forward path, we propose a linear forward attack (LFA) which linearly transforms a control input to degrade the performance of a control system with linear quadratic Gaussian (LQG) control marginally while a noise injection attack (NIA) which adds random noise to increase the variance of the measurement noise is proposed in the feedback path. The suboptimal defensive methods are also provided. Estimating the transformation matrix through the least square (LS) method and minimum mean squared error (MMSE) method, the proposed method tries to nullify the LFA by applying the inverse matrix of the LFA. For NIA, Kalman filter is adopted to deal with changes in noise statistics. The numerical simulations verify that the degree of the proposed attacks can be controllable with a parameter, and the proposed defense can effectively nullify the CPS attacks. This paper is constructed as follows. This paper is constructed as follows. In section-II, a system model and problem formulation are given. The proposed LFA and the proposed defensive method to the LFA are presented in section-III. The proposed NIA and the proposed defensive method to the NIA are given in section-IV. In section-V, the numerical evaluations of the proposed attacks and the defensive methods are provided. Some concluding remarks are made in section-VI.

II. SYSTEM MODEL AND PROBLEM FORMULATION

We consider a networked control system in Fig-1, where control input and system output are transmitted through a sensor to a controller (SC) channel, and a controller to actuator (CA) channel respectively. Two types of cyber physical attack are considered. i.e. LFA in a forward channel and NIA in a feedback channel. The LFA can be considered as a special case of MitM attack in [4], [20], [21]. It is assumed that a network delay is zero for simplicity of problem formulation. It is also assumed that the system is controllable and observable. Since we design a controllable attack which can control the degree of attack, a particular detection scheme is not considered. Similarly, since the detection of the CPS attack is out of the scope of this study, a particular detection scheme is not considered for designing a defense scheme.

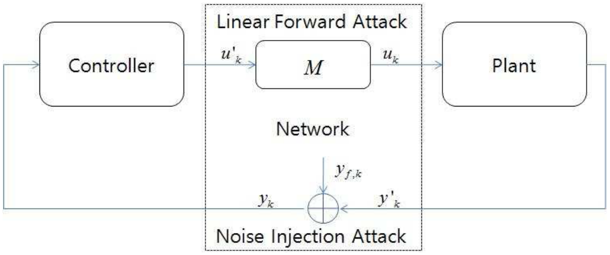


FIGURE 1. **A system model.**

The state equation without cyber physical attack can be expressed as

$$x'_{k+1} = Ax'_k + Bu'_k + v_k \tag{1}$$
$$y'_k = Cx'_k + w_k \tag{2}$$

where $x'_k \in R^n$ is a state at time k, $u'_k \in R^m$ is a control input in the absence of cyber physical attacks, $y'_k \in R^p$ is a system output, and v_k and w_k are process noise and measurement noise which are assumed to be generated from the normal distribution with zero mean and covariance matrices Σ_v and Σ_w, and A, B and C are constant matrices. When LFA and NIA are present, the system will be controlled by a different control input, which results in different states and system outputs. The corresponding state equation under attack is denoted by

$$x_{k+1} = Ax_k + Bu_k + v_k \tag{3}$$
$$y_k = Cx_k + y_{f,k} + w_k \tag{4}$$

where $u_k = Mu'_k$, M is a linear attack matrix, and $y_{f,k}$ is noise for NIA which is assumed to be generated from the normal distribution with zero mean and covariance matrices Σ_{y_f}.

When Kalman filter is exploited for state estimation, the corresponding update equation for the estimated state $\hat{x}_k$ can be expressed as

$$\hat{x}_{k+1} = A\hat{x}_k + Bu'_k + K[y_{k+1} - C(A\hat{x}_k + Bu'_k)] \tag{5}$$

where K is a Kalman gain. It is noted that u'_k is used for the update equation since the control unit does not know that it is transformed to u_k through the LFA. The dynamic equation for the state estimation error can be easily derived as

$$e_{x,k+1} = (A - KCA)e_{x,k} + (B - KCB)e_{u,k} - Ky_{f,k+1} \tag{6}$$

where $e_{x,k} = x_k - \hat{x}_k$ and $e_{u,k} = u_k - u'_k$. (6) explicitly shows that additional error terms can be decomposed into errors resulting from each attack.

For this system, a conventional LQG control can be easily found through deriving the optimal state estimator and the linear quadratic regulators. To derive an optimal estimator, we first define the following variables

$$\hat{x}_{k|k-1} = E\{x_k|Y_{k-1}\} \tag{7}$$
$$\hat{x}_k = E\{x_k|Y_k\} \tag{8}$$
$$P_{k|k-1} = E\{(x_k - \hat{x}_{k|k-1})(x_k - \hat{x}_{k|k-1})^T|Y_{k-1}\} \tag{9}$$
$$P_k = E\{(x_k - \hat{x}_k)(x_k - \hat{x}_k)^T|Y_k\} \tag{10}$$

where $Y_k = \{y_0, y_1, \cdots, y_k\}$. From the above definition, prediction equations and update equations can be easily derived as follows

$$\hat{x}_{k|k-1} = A\hat{x}_{k-1} + Bu'_{k-1} \tag{11}$$
$$\begin{aligned} P_{k|k-1} = &AP_{k-1|k-1}A^T \\ &+ B(M-I)u'_{k-1}{u'_{k-1}}^T(M-1)^T B^T + \Sigma_v \end{aligned} \tag{12}$$
$$\hat{x}_k = \hat{x}_{k|k-1} + P_{k|k-1}C^T S_k^{-1}(y_k - C\hat{x}_{k|k-1}) \tag{13}$$
$$P_k = P_{k|k-1} - P_{k|k-1}C^T S_k^{-1} CP_{k|k-1} \tag{14}$$
$$K_k = P_{k|k-1}C^T S_k^{-1} \tag{15}$$

where $S_k = (CP_{k|k-1}C^T + \Sigma_w + \Sigma_{y_f})$. It is observed that the LFA affects $P_{k|k-1}$ while the NIA does P_k. (12) shows that since $P_{k|k-1}$ has dependency on u_{k-1}, separation principle does not hold when a perfect defense to LFA cannot be made.

To derive the optimal linear quadratic regulator (LQR), the finite horizon cost function is given as follows.

$$J_{LQR}(T) = \sum_{k=1}^{T}[x_k^T Qx_k + {u'_k}^T Ru'_k] + x_{T+1}^T G_{T+1}x_{T+1} \tag{16}$$

where Q, R, and G_{T+1} are positive definite matrices associated with state cost, input cost, and final cost respectively. The optimal control for this cost and resulting Riccati equation are given as in [8]

$$G_k = Q + A^T G_{k+1}A - A^T G_{k+1}BT_k^{-1}B^T G_{k+1}A \tag{17}$$
$$u'_k = -T_k^{-1}B^T G_{k+1}Ax_k \tag{18}$$

where $T_k = (R + B^T G_{k+1}B)$. Finally, the LQG control is given as

$$u'_k = -T_k^{-1}B^T G_{k+1}A\hat{x}_k \tag{19}$$

III. LINEAR FORWARD ATTACK AND DEFENSE

There can be an infinite number of ways to design physical cyber attacks for a system with LQG control. Since LFA affects a control inputs and NIA affects an observation as their direct effects can be seen in (12) and (14), the linear forward attack is designed to degrade LQR performance while the NIA is designed to degrade the performance of state estimation.

A. SLOW LINEAR FORWARD ATTACK

There can be several different goals to design LFA. One possible direction is to slow down the convergence. To this end, one can determine M such that the system can maintain stability while its convergence speed can be degraded, which we call "Slow LFA" throughout this paper. In this context, the following optimization problem can be posed as.

$$\min \bar{\alpha}$$
$$\begin{bmatrix} -Y & Y(A+BML)^T \\ (A+BML)Y & -Y \end{bmatrix} < 0$$
$$\begin{bmatrix} -\bar{\alpha}I & (A+BML)^T \\ (A+BML) & -I \end{bmatrix} \leq 0,$$

$$\bar{\alpha} > 0,\ Y > 0,\ tr(M) = \xi m \tag{20}$$

where $L = -(R + B^T GB)^{-1} B^T GA$ and $G = \lim_{k\to\infty} G_k$is a control gain of LQR controller in (20). The first LMI is a Lyapunov stability condition which is required for a covert attack. The second LMI is equivalent to $(A + BML)^T (A + BML) < \bar{\alpha} I$ through Schur complement where $\bar{\alpha}$ is equivalent to the maximum eigenvalue of the matrix in the left side of the matrix inequality when the optimal solution exists at the boundary. The equality constraint in (20) was included to control the effect of LFA through varying ξ. As the trace of M increases, the achievable maximum singular value of $(A+BML)$ tends to increase, even though it is not guaranteed. Thus, the proposed method controls the degree of attack through varying ξ properly.

B. A DEFENSE FOR THE LINEAR FORWARD ATTACK

In the presence of LFA, there can be many different ways to deal with this attack. Those can be broadly classified into a linear method and a nonlinear method. However, since it is a linear attack and the considered system is a linear system, a simple linear defense can provide an optimal way to defend the system from LFA.

A considered strategy is to compensate for the linear attack through the linear feedfoward defense (LFD), M_D, which linearly transforms the control input. After applying LFD, the control input applied to the plant will be

$$u'_k = MM_D u_k \tag{21}$$

Optimal control is given in (21) in the absence of LFA. Thus, the optimal linear defense would be simply given as

$$M_D = M^{-1} \tag{22}$$

It implies that the optimal linear defense requires system identification of the attacker paradoxically. One of the direct and simple methods is to exploit the system equation in (1). However, state information is not available in many cases. Thus, we first introduce a stacked vector form of the system equation with the assumption that there is no process noise and no measurement noise.

$$Y_l(k) = \Gamma_l X_l(k) + H_{l,M} U_l(k) \tag{23}$$

$$\Gamma_l = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{l-1} \end{bmatrix},$$

$$H_{l,M} = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ CBM & 0 & & 0 \\ \vdots & \vdots & \ddots & \vdots \\ CA^{l-2}BM & CA^{l-3}BM & \cdots & 0 \end{bmatrix} \tag{24}$$

where $U_l(k) = \begin{bmatrix} u_k^T & u_{k+1}^T & \cdots & u_{k+l-1}^T \end{bmatrix}^T$, $X_l(k) = \begin{bmatrix} x_k^T & x_{k+1}^T \cdots x_{k+l-1}^T \end{bmatrix}^T$, and $Y_l(k) = \begin{bmatrix} y_k^T & y_{k+1}^T & \cdots & y_{k+l-1}^T \end{bmatrix}^T$. The first term in (23) can be eliminated through multiplying both sides by the matrix of which columns are null space of Γ_l. The resultant equation will be

$$\Gamma_l^{\perp T} Y_l(k) = \Gamma_l^{\perp T} H_l U_l(k) \tag{25}$$

where $\Gamma_l{}^{\perp T} \Gamma_l = 0$. $\Gamma_l^{\perp} \in^{lp\times(lp-n)}$ can be found from left singular vectors of Γ_l of which singular value is 0. The following proposition provides a method of estimating M from the stacked vector form in (23).

Proposition-1: When $\tilde{W}$ has full column rank and $(l-1)p \geq m^2$, M can be calculated as

$$vec(M) = (\tilde{W}^T \tilde{W})^{-1} \tilde{W}^T \Gamma_l^{\perp T} Y_l(k) \tag{26}$$

where $vec(\cdot)$ is the vectorization of the matrix in the bracket.

Proof: (25) can be rewritten through factoring out in the following way.

$$\Gamma_l^{\perp T} Y_l(k) = \Gamma_l^{\perp T} H_{l,I} (I \otimes M) U_l(k) \tag{27}$$

where $\otimes$ represents a Kronecker product operation. (27) can be rearranged as matrix-vector equation through vectorization as [24]

$$(U_l(k)^T \otimes \Gamma_l^{\perp T} H_{l,I}) vec(I \otimes M) = \Gamma_l^{\perp T} Y_l(k) \tag{28}$$

However, due to the special structure in $vec(I \otimes M)$, it is not possible to finding a solution through direct inversion. Let $(U_l(k)^T \otimes \Gamma_l^{\perp T} H_{l,I})$ be denoted by $W \in^{(l-1)p\times(lm)^2}$. W can be represented with its submatrix $W_i \in^{(l-1)p\times m}$ as $W = [\,W_1\ W_2\ \cdots\ W_{l^2 m}]$. The left side of (28) can be rearranged with W as

$$Wvec(I \otimes M) = \sum_{i=1}^{m} \sum_{j=1}^{l} W_{(i-1)l+(j-1)ml+j} [M]_i \tag{29}$$

(28) can be rearranged from using (29) as

$$\tilde{W} vec(M) = \Gamma_l^{\perp T} Y_l(k) \tag{30}$$

where $\tilde{W} = \begin{bmatrix} \tilde{W}_1 & \tilde{W}_2 & \cdots & \tilde{W}_m \end{bmatrix}$, and $\tilde{W}_i = \sum_{j=1}^{l} W_{(i-1)l+(j-1)ml+j}$. As long as $\tilde{W}$ has full column rank and $(l-1)p \geq m^2$, the left inverse of $\tilde{W}$ exists, which proves the proposition.

However, even though (26) provides the perfect recovery of M in the absence of noise, the estimation performance is subject to degrade when there are process noise or measurement noise. In the presence of noise, (30) can be written as

$$\tilde{W} vec(M) + E_l(k) = \Gamma_l^{\perp T} Y_l(k) \tag{31}$$

where $E_l(k)$ is a composite noise due to process noise and measurement noise. It is well known that the optimal estimation with Gaussian noise is MMSE estimation. From the orthogonality principle of the MMSE estimation, the estimation of $vec(M)$ can be easily derived as

$$vec(M)_{MMSE} = \tilde{W}^H (\tilde{W} \tilde{W}^H + \Sigma_E)^{-1} \Gamma_l^{\perp T} Y_l(k) \tag{32}$$

where $\Sigma_E = \Gamma_l^{\perp T} (H_{l,l|B=I} \Sigma_v H_{l,l|B=I}^T + \Sigma_w) \Gamma_l^{\perp}$. It can be observed that the complexity of LS and MMSE estimation are comparable. With the additional information of noise statistics, the MMSE estimation is expected to provide more robust estimation performance than LS estimation.

IV. NOISE INJECTION ATTACK AND DEFENSE

The optimal noise injection without any constraint will be just flooding noise with infinite variance. However, it is impossible physically. Simple large noise inject attack can be easily detected. Thus, noise injection with power constraint will be of interest.

A. SLOW NOISE INJECTION ATTACK

When the controller is determined from LQR, noise injection does not have any effect on the controller design. Performance metric associated with the optimal noise injection will be involved with state estimation. The stability condition for the linear state estimator can be posed from the error dynamics as

$$\rho(A - KCA) < 1 \tag{33}$$

where ρ is the spectral radius of the matrix inside the bracket. The corresponding Lyapunov condition can be written as

$$(A - KCA)^T P(A - KCA) - P < 0 \tag{34}$$

where $P > 0$. The stationary Kalman filter in the presence of the noise injection attack can be given as

$$K_s = P_s C^T (CP_s C^T + \Sigma_w + \Sigma_{y_f})^{-1} \tag{35}$$

where P_s is a stationary Riccati matrix associated with Kalman filter. When Σ_{y_f} is much smaller than $CP_sC^T + \Sigma_w$, K_s can be approximated from Talyor series expansion as

$$K \approx K_{s,o}(I - \Sigma_{y_f}(CP_sC^T + \Sigma_w)^{-1}) \tag{36}$$

where $K_{s,0} = P_sC^T(CP_sC^T + \Sigma_w)^{-1}$. Exploiting this approximation, one can determine Σ_{y_f} such that the estimator can maintain stability while its convergence speed can be degraded, which we call 'Slow NIA' throughout this paper. In this context, the following optimization problem can be posed as

$$\begin{aligned} &\min \bar{\beta} \\ &\begin{bmatrix} -Y & Y(\tilde{A} + K_{s,0}\Sigma_{y_f}DCA)^T & I \\ (\tilde{A} + K_{s,0}\Sigma_{y_f}DCA)Y & -Y & 0 \\ I & 0 & \bar{\beta} \end{bmatrix} < 0, \\ &\bar{\beta} > 0,\ Y > 0, \Sigma_{y_f} > 0,\ \Sigma_{y_f} D \leq \gamma I \end{aligned} \tag{37}$$

where $D = (CP_sC^T + \Sigma_w)^{-1}$ and $\tilde{A} = A - K_{s,0}CA$. This problem tries to maximize the largest eigenvalue of the matrix associated with Lyapunov stability condition such that it can slows down the convergence speed implicitly.

B. DEFENSE FOR NOISE INJECTION ATTACK

The NIA simply changes the statistics structure of measurement noise, which affects the performance of the state estimator. Kalman Filter is an optimal filter for estimation in the presence of Gaussian noise. Thus, one can design the Kalman filter simply replacing Σ_w with $\Sigma_w + \Sigma_{y_f}$.

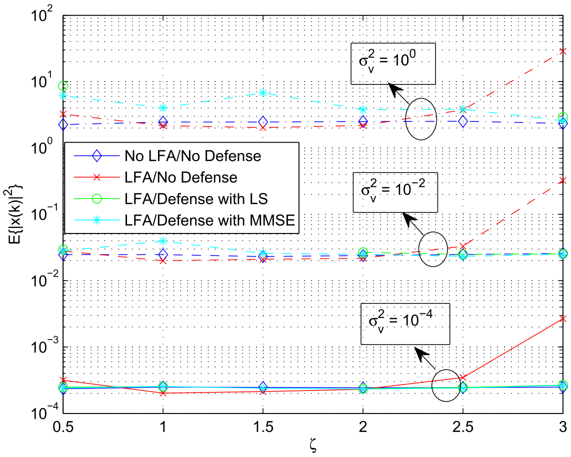


FIGURE 2. The average norm square of state over last 1000 steps for increasing ξ which controls the LFA.

V. NUMERICAL SIMULATIONS

In this section, we assess the performance of the proposed attack and defense through numerical simulations. To this end, we define a hypothetical system as follows.

$$A = \begin{bmatrix} 1 & 0.7098 \\ 0.1 & 0.9653 \end{bmatrix}, \quad B = \begin{bmatrix} 0.0085 & 1 \\ 1.6937 & -1 \end{bmatrix}, \\ C = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \tag{38}$$

This is an unstable system in which the maximum eigenvalue of A is 1.2496. To generate a matrix M for LFA from (20), the VK type iteration [25] was adopted to solve the LMI. An initial state was generated from a standard normal distribution. To evaluate the performance of the estimator for M, the mean squared error was calculated from 500 different realizations. The MMSE estimator for M was designed with an assumption that its mean is I from the consideration of trace constraint. Σ_v and Σ_w are set to be white noise such that $\Sigma_v = \Sigma_w = \sigma_v^2 I = \sigma_w^2 I$.

Figure-2 shows the effect of the proposed LFA for increasing ξ. in terms of $E\{|x(k)|^2\}$ which is calculated from averaging the norm square of states over last 1000 steps for total 5000 steps. When ξ is less than 2, LFA is found to improve the convergence characteristic slightly due to the problem formulation for generating M_n. However $E\{|x(k)|^2\}$ increases proportionally to ξ when ξ is larger than 2. It is also observed that the effect of ξ has the same trend regardless of σ_v^2. However, for this specific system, when ξ is greater than 3.3, it seems that there is no feasible M satisfying the trace constraint. Thus, the attacker needs to find a proper interval to control the effect of LFA. It can be also found that the compensation of LFA with MMSE estimation works efficiently to keep $E\{|x(k)|^2\}$ close to one without attack. However, the compensation of LFA with LS estimation is found to fail often. It makes the state diverge for $(\sigma_v^2, \xi) = (10^{-4}, 1.5), (10^{-2}, 1.0), (10^{-2}, 1.5), (10^0, 1.0), (10^0, 1.5), (10^0, 2.0), (10^{-2}, 2.5)$. It is also found that even

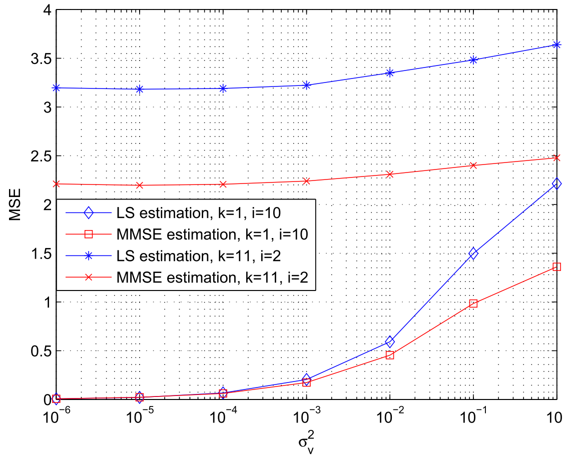


FIGURE 3. **The MSE performance of the proposed LS and MMSE estimator for the LFA attack with varying noise power.**

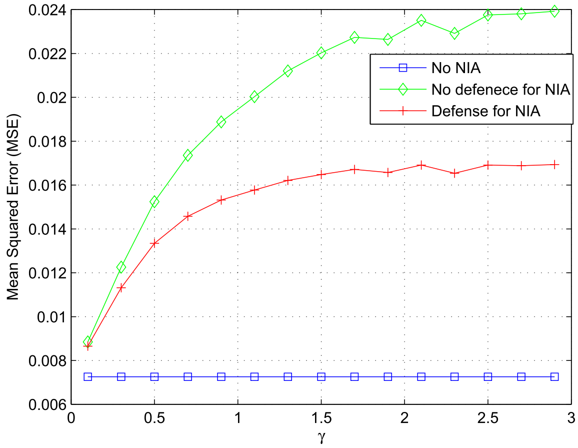


FIGURE 4. **The average norm square of state over last 1000 steps for increasing γ which controls the NIA.**

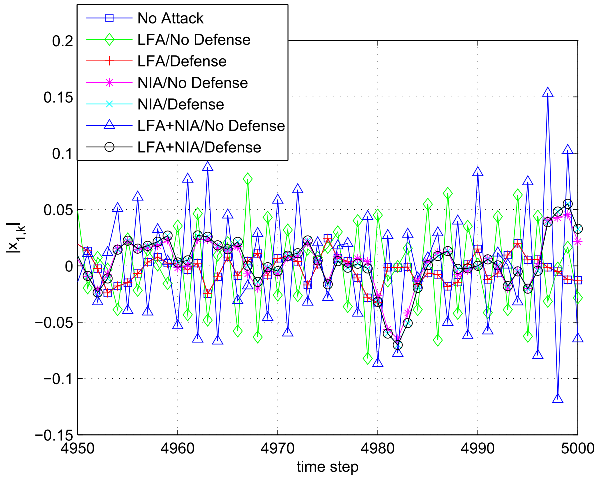


FIGURE 5. **The norm of the first state at the convergence with and without the proposed attack and the proposed defense.**

when LFA does not degrade the convergence performance, the compensation with the proposed estimators may degrade the performance marginally due to the estimation error. However, when the LFA degrades the performance significantly, the proposed compensation with MMSE estimator provides the performance comparable to one without LFA.

To evaluate the performance of the LS and MMSE estimator which are used for the compensation of the LFA, Figure 3 shows the performance of the proposed defense for LFA in terms of MSE. ξ was also set to be 3. It can be observed that LS and MMSE estimators provide the almost the same performance when σ_v^2 is very small and $k = 1$, which corresponds to the case of high signal to noise ratio (SNR). However, as σ_v^2 increases, MMSE estimator provides better performance as expected. When $k = 11$ and $i = 2$ the performance of both estimators degrades significantly. The performance degradation is mainly due to $i = 2$ rather than $k = 11$. When $i = 2$ the number of observed variables is simply smaller than the number of unknown variables. This result shows that i needs to be set properly such that it is proportional to m^2.

Figure-4 shows how the NIA operates with a controlling parameter γ where $\Sigma_v = \Sigma_w = 0.01I$. States were generated over 5000 steps of which last 1000 steps were used to take the average of norm square of states. It is observed that the average norm square of the state increases with increasing γ while the effect of the proposed NIA is small with small γ. However, it does not increase significantly as γ is larger than some value. Even though increasing γ increases the feasible region imposed by γ, the stability condition may limit the feasible region of the space for the covariance matrix of NIA. This implies that NIA can be effectively controlled with γ even though significant attack may be limited due to the stability condition and the problem formulation. The proposed defense which just considers the additional measurement noise covariance shows better performance than one without defense. It is observed that the proposed defense effectively deals with NIA though its average norm square of the state is larger than one without NIA. In table-1, the effect of NIA and the efficiency of the proposed defense were compared for different disturbance and measurement noise configurations. It is assumed that disturbance noise and the measurement noise are white noise. The average norm square of the state increases by 100 times or so when both noise power increases by 100 times. The proposed defense is observed to reduce the average norm square of the state by 20∼40% depending on the noise configurations, which may also depend on γ. When the measurement noise power and disturbance noise power are asymmetric, the measurement power seems to have a greater effect on the average norm square of the state. The constraint associated with γ can be rewritten as $\Sigma_{y_f} \leq \gamma(CP_sC^T + \Sigma_w)$. It shows that for a fixed γ, feasible region of Σ_{y_f} is restricted by Σ_w, which explains the results in table-1.

Figure-5 shows the magnitude of the first state over the last 50 steps after generating the states for 5000 steps. For the simplicity of the plot, the second state is omitted since its characteristics are pretty much same as those of the first

TABLE 1. Effect of NIA and the proposed defense for the different configurations of process noise covariance matrix and measurement noise covariance matrix.

γ	σ_v	σ_w	No NIA	No Defense for NIA	Optimal Defense for NIA
1	10^{-4}	10^{-4}	7.26×10^{-5}	2.04×10^{-4}	1.62×10^{-4}
1	10^{-2}	10^{-4}	9.98×10^{-5}	1.10×10^{-2}	7.52×10^{-3}
1	10^{-4}	10^{-2}	4.90×10^{-3}	7.27×10^{-3}	4.62×10^{-3}
1	10^{-2}	10^{-2}	7.26×10^{-3}	1.96×10^{-2}	1.57×10^{-2}

state. Σ_v, Σ_w, γ and ξ are set to be $10^{-4}I$, $10^{-4}I$, 1, and 3 respectively. The figure shows that the proposed defense effectively makes a defense over the proposed attack. Even though NIA seems to degrade the performance more significantly than LFA in this figure, the effect of NIA and LFA may vary depending on disturbance noise covariance, measurement noise covariance, and parameterizations of the attack and the defense. One interesting result is that the performance with LFA, NIA, and no defense is better than one with NIA, and no defense. It can be expected from the result shown in the figure-2 in which LFA improves the convergence slightly for $\xi < 2$

VI. CONCLUSION

In this paper, we provided a method of LFA and NIA with controllable parameters such that they can degrade the performance of a control system marginally with a controllable degree of amount rather than destabilizing a system. The defensive methods were also proposed to deal with the proposed attack. The numerical simulations verified that the proposed attack could control the degree of degradation with parameterizations, and the proposed defensive methods could effectively defend the proposed attack.

There are several remaining problems which need to be addressed in future research. The proposed LFA and NIA were developed separately. If they were designed jointly under a single performance measure, it is expected that there can be more flexibility in attackers. Similarly, a joint defense method is expected to provide more robust performance. In this research, disturbance and system uncertainties were not considered. However, they need to be considered further to provide a more robust method to keep the sanity of the system away from the CPS attack in a real control system. The Takagi-Sugeno (T-S) fuzzy model to approximate the control systems with uncertainties such as parameter perturbation [26] and packet dropout [27] was exploited to estimate the state of the system. As such, a fuzzy model may be exploited to develop a more robust defense to noise injection attacks with system uncertainties. A sliding mode control [28] may be also exploited to make the control system more robust to the uncertainties incurred by the false data injection. While CPS attack and defense were considered, the detection of the attack was not addressed. False data detection can be considered to be in line with the fault data detection. Some design methodologies for fault diagnosis such as total measurable fault information residual (ToMFIR) based approach [29] or a robust observer-based detection method [30] can be extended to the detection of CPS attack.

REFERENCES

[1] E. A. Lee, "Cyber physical systems: Design challenges," in *Proc. Int. Symp. Object Oriented Real-Time Distrib. Comput.*, Orlando, FL, USA, 2008, pp. 363–369.

[2] T. M. Chen and S. Abu-Nimeh, "Lessons from Stuxnet," *Computer*, vol. 44, no. 4, pp. 91–93, Apr. 2011.

[3] J. Petit and S. E. Shladover, "Potential cyberattacks on automated vehicles," *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 2, pp. 546–556, Apr. 2015.

[4] A. O. de Sá, L. F. R. da Costa Carmo, and R. C. S. Machado, "Covert attacks in cyber-physical control systems," *IEEE Trans. Ind. Informat.*, vol. 13, no. 4, pp. 1641–1651, Aug. 2017.

[5] E. Ke Wang, Y. Ye, X. Xu, S. M. Yiu, L. C. K. Hui, and K. P. Chow, "Security issues and challenges for cyber physical system," in *Proc. IEEE/ACM Int. Conf. Green Comput. Commun. Int. Conf. Cyber, Phys. Social Comput.*, Hangzhou, China, Dec. 2010, pp. 733–738.

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

[7] Y. Mo and B. Sinopoli, "False data injection attacks in control systems," in *Proc. 1st Workshop Secure Control Syst. (CPS Week)*, Stockholm, Sweden, Apr. 2010, pp. 1–6.

[8] M. J. Khojasteh, A. Khina, M. Franceschetti, and T. Javidi, "Learning-based attacks in cyber-physical systems," *IEEE Trans. Control Netw. Syst.*, early access, Sep. 30, 2020, doi: 10.1109/TCNS.2020.3028035.

[9] P. Venkitasubramaniam, J. Yao, and P. Pradhan, "Information-theoretic security in stochastic control systems," *Proc. IEEE*, vol. 103, no. 10, pp. 1914–1931, Oct. 2015.

[10] B. Chen, K. L. Butler-Purry, A. Goulart, and D. Kundur, "Implementing a real-time cyber-physical system test bed in RTDS and OPNET," in *Proc. North Amer. Power Symp. (NAPS)*, Pullman, WA, USA, Sep. 2014, pp. 1–6.

[11] D. Ding, Z. Wang, G. Wei, and F. E. Alsaadi, "Event-based security control for discrete-time stochastic systems," *IET Control Theory Appl.*, vol. 10, no. 15, pp. 1808–1815, Oct. 2016.

[12] M. Zhu and S. Martinez, "On the performance analysis of resilient networked control systems under replay attacks," *IEEE Trans. Autom. Control*, vol. 59, no. 3, pp. 804–808, Mar. 2014.

[13] Z. Cao, Y. Niu, and J. Song, "Finite-time sliding-mode control of Markovian jump cyber-physical systems against randomly occurring injection attacks," *IEEE Trans. Autom. Control*, vol. 65, no. 3, pp. 1264–1271, Mar. 2020.

[14] E. Drayer and T. Routtenberg, "Detection of false data injection attacks in smart grids based on graph signal processing," *IEEE Syst. J.*, vol. 14, no. 2, pp. 1886–1896, Jun. 2020.

[15] B. Chen, H. Li, and B. Zhou, "Real-time identification of false data injection attacks: A novel dynamic-static parallel state estimation based mechanism," *IEEE Access*, vol. 7, pp. 95812–95824, 2019.

[16] C. Schellenberger and P. Zhang, "Detection of covert attacks on cyber-physical systems by extending the system dynamics with an auxiliary system," in *Proc. IEEE 56th Annu. Conf. Decis. Control (CDC)*, Melbourne, VIC, Australia, Dec. 2017, pp. 1374–1379.

[17] Y. Li, Y. Wang, and S. Hu, "Online generative adversary network based measurement recovery in false data injection attacks: A cyber-physical approach," *IEEE Trans. Ind. Informat.*, vol. 16, no. 3, pp. 2031–2043, Mar. 2020.

[18] V. Lesi, I. Jovanov, and M. Pajic, "Integrating security in resource-constrained cyber-physical systems," *ACM Trans. Cyber-Phys. Syst.*, vol. 4, no. 3, pp. 1–27, May 2020.

[19] F. Farivar, M. S. Haghighi, S. Barchinezhad, and A. Jolfaei, "Detection and compensation of covert service-degrading intrusions in cyber physical systems through intelligent adaptive control," in *Proc. IEEE Int. Conf. Ind. Technol. (ICIT)*, Melbourne, VIC, Australia, Feb. 2019, pp. 1143–1148.

[20] R. Smith, "A decoupled feedback structure for covertly appropriating networked control systems," in *Proc. 18th IFAC World Congr.*, Milano, Italy, 2011, pp. 90–95.

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

[22] M. G. Safonov, "Data-driven robust control design: Unfalsified control," in *Proc. 21st Century Mil. Appl.*, Nueilly-sur-Seine Cedex, France, 2006, pp. 4–18.

[23] J.-C. Bourin, "Matrix versions of some classical inequalities," *Linear Algebra Appl.*, vol. 416, nos. 2–3, pp. 890–907, Jul. 2006.

[24] K. Schcke, "On the Kronecker product," Univ. Waterloo, Waterloo, ON, Canada, Tech. Rep., 2013.

[25] L. El Ghaoui and V. Balakrishnan, "Synthesis of fixed-structure controllers via numerical optimization," in *Proc. 33rd IEEE Conf. Decis. Control*, Buena Vista, FL, USA, 1994, pp. 2678–2683.

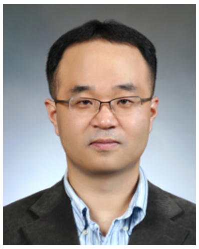
[26] X.-H. Chang, J. Xiong, and J. H. Park, "Estimation for a class of parameter-controlled tunnel diode circuits," *IEEE Trans. Syst., Man, Cybern. Syst.*, vol. 50, no. 11, pp. 4697–4707, Nov. 2020.

[27] X.-H. Chang, Q. Liu, Y.-M. Wang, and J. Xiong, "Fuzzy peak-to-peak filtering for networked nonlinear systems with multipath data packet dropouts," *IEEE Trans. Fuzzy Syst.*, vol. 27, no. 3, pp. 436–446, Mar. 2019.

[28] S. Ding and S. Li, "Second-order sliding mode controller design subject to mismatched term," *Automatica*, vol. 77, pp. 388–392, Mar. 2017.

[29] Y. Wu, B. Jiang, N. Lu, H. Yang, and Y. Zhou, "Multiple incipient sensor faults diagnosis with application to high-speed railway traction devices," *ISA Trans.*, vol. 67, pp. 183–192, Mar. 2017.

[30] Y. Wu, B. Jiang, and Y. Wang, "Incipient winding fault detection and diagnosis for squirrel-cage induction motors equipped on CRH trains," *ISA Trans.*, vol. 99, pp. 488–495, Apr. 2020.



JANGHOON YANG (Member, IEEE) received the Ph.D. degree in electrical engineering from the University of Southern California, Los Angeles, CA, USA, in 2001. From 2001 to 2006, he was with the Communication Research and Development Center, Samsung Electronics. From 2006 to 2010, he was a Research Assistant Professor with the Department of Electrical and Electronic Engineering, Yonsei University. Since 2010, he has been with the Seoul Media Institute of Technology, Seoul, South Korea, where he is currently an Associate Professor with the Department of New Media. He has published numerous papers in the area of multi-antenna transmission, signal processing, and control. His research interests include wireless systems and networks, artificial intelligence, control theory, neuroscience, affective computing, and intervention for special education.