IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 16 July 2024, accepted 1 August 2024, date of publication 5 August 2024, date of current version 10 September 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3438287

RESEARCH ARTICLE

Detection and Cancellation of Multiplicative FDI Attack on Bilateral Encrypted Control System

KATSUMASA KOSHA[1], TETSURO MIYAZAKI[1], (Member, IEEE), KAORU TERANISHI[2], (Member, IEEE), KIMINAO KOGISO[2], (Member, IEEE), AND KENJI KAWASHIMA[1], (Member, IEEE)

[1]Department of Information Physics and Computing, The University of Tokyo, Bunkyo-ku, Tokyo 113-8656, Japan
[2]Department of Mechanical and Intelligent Systems Engineering, The University of Electro-Communications, Chofu-shi, Tokyo 182-8585, Japan

Corresponding author: Tetsuro Miyazaki (Tetsuro_Miyazaki@ipc.i.u-tokyo.ac.jp)

This work was supported by JSPS KAKENHI through Grant-in-Aid for Scientific Research (B) under Grant JP22H01509 and Grant JP23K22779.

ABSTRACT Teleoperation of remote assist robots has recently advanced owing to the development of communication technology. However, anonymous malicious attacks may intercept or falsify the network control system. Therefore it is necessary to improve the security against cyber-attacks. As a countermeasure, an encrypted control method has been applied to prevent intercepting and detect the falsification of control parameters, as well as control signals by performing control operations with encrypted signals and parameters. However, it is difficult to detect False Data Injection (FDI) attacks, which exploit encryption malleability, and countermeasures against such attacks are necessary. Therefore, this study proposes a novel method to detect and cancel the effect of the FDI attack, which aims at the encrypted bilateral control system's malleability. The assumed FDI attack falsifies the plaintext by multiplying a constant factor, which is realized by monitoring the entire system's energy change and estimating the attack parameters. The proposed method can detect and cancel the impact of the FDI attack within finite steps. Moreover, we verified the proposed method for the bilateral control system using ElGamal encryption and experimentally confirmed its effectiveness against the FDI attack.

INDEX TERMS Cyber security, FDI attack, encrypted control, attack detection and cancellation, pneumatically driven system, bilateral control system.

I. INTRODUCTION

Recently, research on remote robot control has increased owing to the development in communication technology. Bilateral control is a method that enables teleoperation [1], [2], [3], and it involves an operator-controlled device known as the leader device and a remotely located device known as the follower device. When both the leader and follower devices are coupled by sending reference values, control that follows each other's position and force is achieved. This facilitates remote control while receiving external force feedback; thus it it critical in tasks requiring precise force perception, such as surgical assistance. Historically, the first mechanical and electrical leader-follower manipulators were developed at the Argonne National Laboratory in the United States [4], [5], [6]. Thereafter, developments shifted toward electrical leader-follower systems, which offer superior work range, installation capabilities, and adjustable force reflection ratios. Alongside this technological shift, control theories have been developed to address challenges inherent in remote control, such as time delays, information loss, and feedback of kinesthetic sensations [3], [7], [8], [9]. Furthermore, control strategies that enhance the stability and accuracy of bilateral teleoperation, especially when considering soft contact with the environment, have been explored [10]. Following extensive developments in control theory, various practical applications of remote robot control have emerged, including spacecraft [11], industrial robots [12], undersea

The associate editor coordinating the review of this manuscript and approving it for publication was Hosam El-Ocla.

teleoperators [13], mine-cleaning devices [14], steer-by-wire systems [15], and remote surgery [16], [17], [18], [19], [20], [21], [22], [23].

In such remote control systems, it is critical to establish security against cyber-attacks [24], [25], [26], [27], [28], [29]. Missing a malicious attack on the control system can result in serious consequences, such as theft of information inside the control device, unauthorized manipulation of the system, or destruction of the controlled object. A prevalent form of attack on control systems is the False Data Injection (FDI), which involves falsifying signals and control commands over the network; thus, enabling unauthorized manipulation or destruction of the control target.

As an effective countermeasure against cyber-attacks on control systems, encrypted control has been proposed in [30] and [31], and it involves encrypting control parameters and signals. In encrypted control, homomorphic encryption [32], [33], [49] is used to compute the controller's output from the encrypted control parameters and encrypted controller's inputs, thereby hiding the parameters and signals of the control system. Moreover, encrypted control is sensitive to data falsification. Owing to encryption properties, decrypting falsified ciphertext generates significant white noise in the decrypted controller's output [34], which can facilitate the detection of several FDI attacks.

However, there is the malleability of encryption schemes used in encrypted control [35], [36], [37], [38]. For example, in ElGamal encryption [49], malleability allows the plaintext to be manipulated by multiplying the second component of the ciphertext by a constant. In conventional bilateral encrypted control systems using wave-variable transformations [39], [40], an FDI attack exploiting the malleability can compromise the stability of the system. It is difficult to detect such FDI attacks because they do not generate the aforementioned white noise, and can inflict serious damage to bilateral encrypted control systems. Therefore, addressing this challenge and developing countermeasures for malleability-based attacks is crucial for enhancing the cybersecurity of bilateral encrypted control systems.

A. OBJECTIVE AND CONTRIBUTIONS

This study aims to propose attack detection and restoration methods for an encrypted bilateral control system. In the assumed force-feedback bilateral control system, the leader receives the follower's force information, while the follower receives the leader's position information. In the attack model, it is assumed that the attacker is knowledgeable about both the control system and the encryption scheme, and the force and position information can be manipulated by multiplying it by an attack parameter. This is feasible owing to the malleability of the encryption scheme. If the attacker chooses a relatively large parameter value, it can destabilize the bilateral control system by increasing the energy of the system, thereby breaking the passivity of the system. This has motivated the development of a passivity observer for each leader and follower to compute the total energy, and to integrate an energy-based detection method into the encrypted bilateral control system.

This study presents a detection and restoration algorithm for the FDI attack by estimating the attack parameters. Considering the information asymmetry — where the attacker has information about the system but the system users lack information about the attacker — detecting and restoring from the attack has become a challenge. Although it is difficult to ascertain the true attack parameters directly, the proposed method employs a passivity observer to estimate these parameters and to detect and cancel the impact of the FDI attack within finite steps. The effectiveness of the proposed attack detection and restoration method for the bilateral encrypted control system are validated both theoretically and experimentally.

The contributions of this study are as follows. Methodically establishing a novel energy-based attack detection and cancellation framework for bilateral control systems, featuring a task-based systematic approach to attack parameter estimation. Theoretically demonstrating that the FDI attack can be detected and restored within finite steps, grounded in the passivity of the bilateral control system. Technically validating the proposed method using a test bed designed for remote surgical robots, contributes to foundational cybersecurity technology in remote surgical operations. Specifically, it involves enhancing the safety and security of surgical operations and advancing the development of cybersecure bilateral control systems.

B. ORGANIZATION OF THIS PAPER

This paper is organized as follows. Section II introduces the force feedback bilateral control system and the concept of encrypted control to describe the attack model for the bilateral encrypted control system. Section III proposes its detection and restoration algorithm, along with a theoretical result as a theorem. Section IV demonstrates the effectiveness of the proposed method though experiments under several attack scenarios. Section V discusses the challenges of the method, which enables the realization of more secure control systems. Finally, Section VI concludes this paper.

Notations: $\mathbb{Z}$ denotes the set of integers, and $\mathbb{Z}_{\geq i}$ denotes the set of integers greater than or equal to $i \in \mathbb{Z}$. The variables in this study include discrete time variables, which are shown with the step $k \in \mathbb{Z}_{\geq 0}$. $(\cdot)^{m\times n}$ indicates a matrix with m rows and n columns. The equation symbol, $\simeq$ is approximately equal. $\dot{(\cdot)}$ indicates the differential value of variable $(\cdot)$.

II. PROBLEM SETUP

This section describes the FDI-attack detection problem for the bilateral encrypted control systems.

A. FORCE-FEEDBACK BILATERAL CONTROL

This section describes the bilateral control system used in this study. There are 3 types of bilateral control: Symmetrical, Force reverse type, and Force reflecting type [41], [42].

Considering situations where the follower's force feedback to the leader device is beneficial, such as a surgical assist robot, this study adopted a force-reflecting type control because it can ideally return the force immediately.

Fig. 1 shows the block diagram of the entire control system. While Shono et.al [39], [40] constructed the bilateral control system with wave variables, in this study, the leader device sends its position, x_l to the follower device, and the follower device sends its force, F_f to the leader device. The control requirements are considered as follows:

$$x_l = x_f, \quad F_l = -F_f,$$

where x_f is the follower position and F_l is the leader force. The first equation is the condition for the motion to follow, while the second is the condition for the action-reaction of the force. Both devices are controlled to satisfy these conditions.

There exists a time delay in the communication between each device [43], [44], [45]. As shown in Fig. 1, if the reference force of the leader is F_{ld}, the reference position of the follower is x_{fd}, and the time delay steps T_1, T_2 between the leader and follower, then the following relationship holds.

$$x_{fd}(k) = x_l(k - T_1), \quad F_{ld}(k) = F_f(k - T_2). \tag{1}$$

The control block diagrams of the leader and follower devices included in Fig. 1 are shown in Fig. 2 and 3, respectively. The leader device performs the force control, which calculates the control voltage, V_l from F_{ld}. In Fig. 2, K_{ap}, K_{ai}, and K_{ad} are the force proportional gain, force integral gain, and force differential gain, respectively. z is z-transformation operator. V_l is calculated as follows:

$$V_l(z) = \left(K_{ap} + \frac{K_{ai}}{1 - z^{-1}}\right)(-F_{ld}(z) - F_l(z)) - K_{ad}(1 - z^{-1})F_l(z), \tag{2}$$

and V_l is input to the actuator. F_h indicates the force on the actuator by the operator. The position, x_l, is sent to the follower device.

The follower device performs the cascade control, that is, the position control, including the force control inside. As shown in Fig. 3, K_{pp}, K_{pi}, and K_{pd} are the position proportional gain, position integral gain, and position differential gain, respectively. The control voltage, V_f is calculated from x_{fd} as follows:

$$V_f(z) = \left(K_{ap} + \frac{K_{ai}}{1 - z^{-1}}\right)(F_{ref}(z) - F_f(z)) - K_{ad}(1 - z^{-1})F_f(z), \tag{3a}$$

$$F_{ref}(z) = \left(K_{pp} + \frac{K_{pi}}{1 - z^{-1}}\right)(x_{fd}(z) - x_f(z)) - K_{pd}(1 - z^{-1})x_f(z), \tag{3b}$$

and V_f is input to the actuator. F_e indicates the external force on the follower device by contact with such an obstacle. The force, F_f, is sent to the leader device. In this study, the leader's motion is a periodic motion with frequency f and amplitude A, which is moved by the human hand. The values of f and A are critical in the proposed method, which is detailed in SECTION V.

B. PASSIVITY OBSERVER

To monitor the stability of the system, a measure called Passivity Observer (PO) [46], [47], [48] was proposed as follows:

$$PO(k) = \Delta T \Sigma_{\tau=0}^{k}(\dot{x}_l(\tau)F_l(\tau) - \dot{x}_f(\tau)F_f(\tau)), \tag{4}$$

where ΔT is the sampling time. If PO is a positive value, then $PO(k) \geq 0$, and it guarantees that the system is stable. Therefore, this study uses PO as a sufficient stability condition. However, it is difficult to simultaneously share (4) with the leader and follower because merging the communicated values yields at least one communication delay between the leader and follower. In Fig. 1 case, the communication lines have T_1 and T_2 step delays. Therefore, this study computed passivity observers PO_l and PO_f for the leader and follower, respectively, as follows:

$$PO_l(k) = \Delta T \Sigma_{\tau=0}^{k-T_2}(\dot{x}_l(\tau)F_l(\tau) - \dot{x}_l(\tau - T_1)F_l(\tau + T_2)),$$

$$PO_f(k) = \Delta T \Sigma_{\tau=0}^{k-T_1}(\dot{x}_f(\tau + T_1)F_f(\tau - T_2) - \dot{x}_f(\tau)F_f(\tau)).$$

These equations are derived using (1) and assumptions of $x_{fd} \simeq x_f$ and $F_{ld} \simeq F_l$; thus, the control tracking performance is sufficient. Furthermore, this study introduced an algorithm to calculate the amount of change in PO_l and PO_f per one step, denoted as dPO_l and dPO_f, respectively,

$$dPO_l(k) = \dot{x}_l(k - T_2)F_l(k - T_2) - \dot{x}_l(k - T_1 - T_2)F_l(k),$$

$$dPO_f(k) = \dot{x}_f(k)F_f(k - T_1 - T_2) - \dot{x}_f(k - T_1)F_f(k - T_1).$$

C. ELGAMAL ENCRYPTION

This section explains ElGamal encryption used in the encrypted control later. ElGamal encryption [49] is multiplicative homomorphic encryption that conceals the control parameters and signals. ElGamal encryption scheme, which is denoted as $\mathcal{E}^{\times}$, consists of Gen: $p \mapsto (\mathrm{pk}, \mathrm{sk}) = ((\mathbb{G}, q, g, h), s)$, Enc:$(\mathrm{pk}, m) \mapsto c = (c_1, c_2) = (g^r \bmod p, mh^r \bmod p)$, and Dec:$(\mathrm{sk}, c) \mapsto m' = c_1^{-s}c_2 \bmod p$, where $p = 2q + 1$ is a safe prime. g is a generator of a cyclic group $\mathbb{G} = \{g^i \bmod p | i \in \mathbb{Z}_q\}$ such that $g^q \bmod p = 1$: s is a random number in $\mathbb{Z}_q$ generated once by the keygen: r is a random number $\mathbb{Z}_q$ in generated for every encryption instance. : and $h = g^s \bmod p$, where $\mathbb{Z}_q = \{0, 1, 2, \cdots, q - 1\}$. The plaintext and ciphertext spaces, $\mathcal{M}$ and $\mathcal{C}$, are expressed by $\mathcal{M} = \mathbb{G}$ and $\mathcal{C} = \mathbb{G}^2$, respectively. Furthermore, ElGamal encryption has multiplicative homomorphism as follows:

$$\mathrm{Dec}(\mathrm{sk}, \mathrm{Enc}(\mathrm{pk}, m) * \mathrm{Enc}(\mathrm{pk}, m') \bmod p) = mm' \bmod p, \tag{5}$$

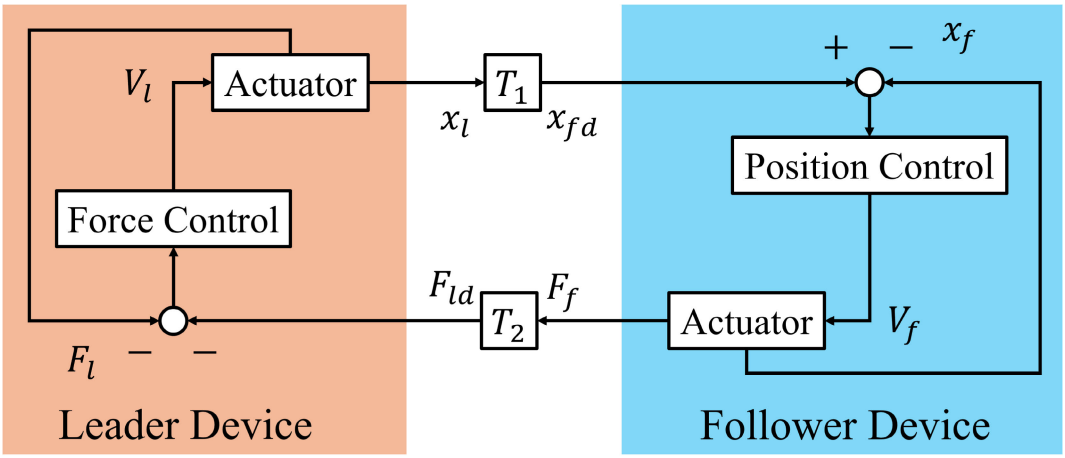


FIGURE 1. Block diagram of the entire control system. The leader device sends the position signal to the follower device, and the follower device sends the force signal to the leader device.

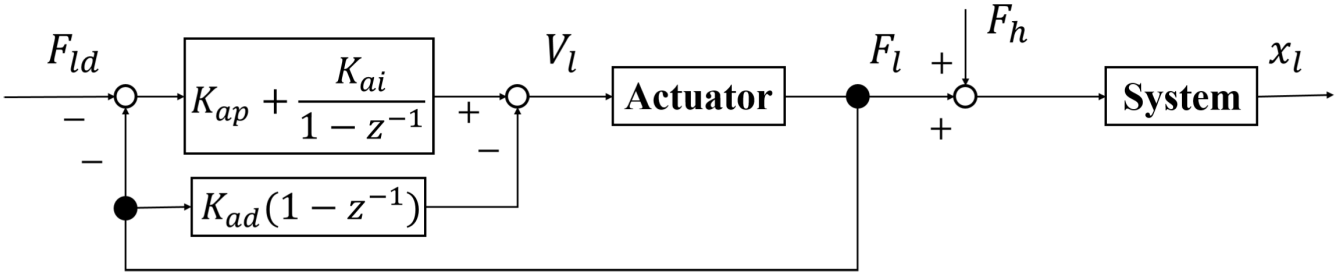


FIGURE 2. Block diagram of the leader device. Calculate the control voltage from the reference signal of the cylinder force.

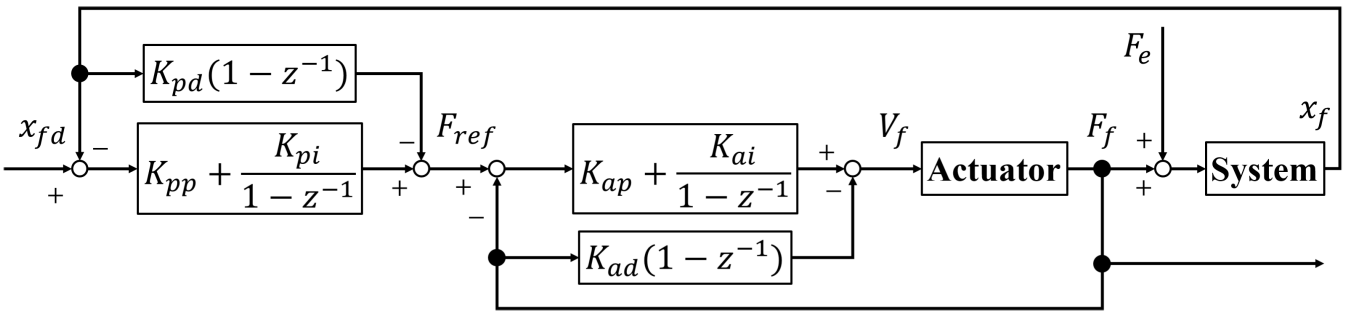


FIGURE 3. Block diagram of the follower device. Calculate the control voltage from the reference signal of the cylinder position.

where $\forall m, m' \in \mathcal{M}$, and $*$ represents the Hadamard product. This property allows the preservation of multiplication over encrypted data.

D. ENCRYPTED CONTROL

In this study, we implemented encryption in a bilateral control system, using ElGamal encryption, as shown in Fig. 4. Enc and Dec^+ denote encryption and decryption, respectively. The superscript E indicates that the variable is encrypted. The leader device sends the encrypted position signal, $x_l^E = \mathrm{Enc}(x_l)$, to the follower device, and the follower device sends the encrypted force signal, $F_f^E = \mathrm{Enc}(F_f)$ to the leader device. F_{ld}^E and x_{ld}^E are the encrypted reference signal with time delay, that is, $F_{ld}^E(k) = F_f^E(k - T_2)$ and $x_{fd}^E(k) = x_l^E(k - T_1)$. In the leader device, the state vector, ξ_l is encrypted to $\xi_l^E = \mathrm{Enc}(\xi_l)$, and updated with F_{ld}^E, which is conducted in Update ξ_l^E block in Fig. 4. Ψ_l^E is calculated with ξ_l^E, which is used to calculate the control voltage, V_l. This is the same on the follower side.

The implementation of the encrypted control allows us to hide the reference value on the communication channel, control the voltage of each device, and calculate the contents of the reference value. The computational domain in plaintext is only the control voltage input to the device and drive. In encrypted control, the control law is expressed as follows: $\psi = f(\Phi, \xi) = \Phi\xi$, where $\Phi = [\Phi_1\ \Phi_2\ \cdots\ \Phi_n] \in \mathbb{R}^{m\times n}$

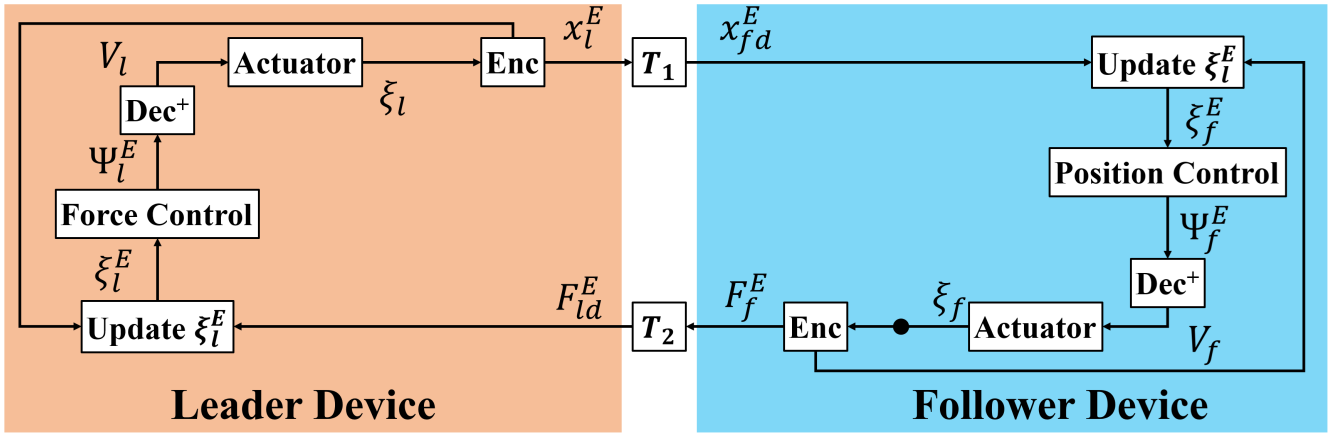


FIGURE 4. Block diagram of encrypted bilateral control system. The superscript E indicates that the variable is encrypted. V_l and V_f are calculated on ciphertext space $\mathcal{C}$.

is the coefficient matrix that arranges the parameters of the controller, and $\xi = [\xi_1\ \xi_2\ \cdots\ \xi_n]^\top \in \mathbb{R}^n$ is the state vector that lays out the control signal. The control computation can be divided into multiplication and addition, such as $f = f^+ \circ f^\times$, where $f^\times$ and f^+ are denoted as follows:

$$f^\times(\Phi, \xi) = [\Phi_1\xi_1\ \Phi_2\xi_2\ \cdots\ \Phi_n\xi_n] = \Psi,$$
$$f^+(\Psi) = \sum_{k=1}^{n} \Psi_k = \psi.$$

The ciphertext corresponding to ψ_{ij} can be calculated directly as $\mathrm{Enc}(\Phi_{ij})\mathrm{Enc}(\xi_j) \bmod p$ owing to multiplicative homomorphism (5). However, ψ cannot be calculated as a ciphertext owing to addition. Therefore, ψ is obtained after decrypting each component of $\mathrm{Enc}(\Psi)$ and adding them together. Defining $\mathrm{Dec}^+ = f^+ \circ \mathrm{Dec}$, it holds that $\psi = \mathrm{Dec}^+(\mathrm{sk}, \mathrm{Enc}(\mathrm{pk}, \Psi))$. To conduct the encrypted control, Φ and ξ, which are real numbers, must be conversed to the components of $\mathcal{M}$. This study employs the mapping used in [39].

Based on the above discussion, we implemented the encrypted control in bilateral control. For (2), by performing inverse z transformation and focusing on each term, the coefficient matrix, Φ_l, and the state vector, ξ_l, are shown as follows:

$$\Phi_l = [-K_{ad}, -K_{ap}, -K_{ai}, -K_{ap}, -K_{ai}], \tag{6a}$$
$$\xi_l(k) = [\dot{F}_l(k), F_l(k), \sum_{\tau=0}^{k} F_l(\tau)\Delta T,$$
$$F_{ld}(k), \sum_{\tau=0}^{k} F_{ld}(\tau)\Delta T]^\top. \tag{6b}$$

By applying the ElGamal encryption to (6a) and (6b), the encrypted coefficient matrix, $\Phi_l^E = \mathrm{Enc}(\Phi_l)$ and the encrypted state vector, $\xi_l^E = \mathrm{Enc}(\xi_l)$ are calculated, where Enc calculates the encrypted value at each component. Therefore, V_l is obtained as follows:

$$V_l = \mathrm{Dec}^+(\Psi_l^E), \quad \Psi_l^E = f^\times(\mathrm{Enc}(\Phi_l), \mathrm{Enc}(\xi_l)), \tag{7}$$

enabling the control operations and state variables of the leader device to be hidden.

For the follower device, we inversely transform z for (3a), (3b) and further split it into the inner product of the coefficient matrix and the state vector as follows:

$$\begin{aligned}\Phi_f &= [-K_{ap}K_{pd}, -(K_{ap}K_{pp} + K_{ai}K_{pd}),\\ &\quad -(K_{ap}K_{pi} + K_{ai}K_{pp}), -K_{ai}K_{pi},\\ &K_{ap}K_{pp}, K_{ap}K_{pi} + K_{ai}K_{pp},\\ &K_{ai}K_{pi}, -K_{ad}, -K_{ap}, -K_{ai}],\end{aligned} \tag{8a}$$
$$\begin{aligned}\xi_f(k) &= [\dot{x}_f(k), x_f(k),\\ &\sum_{\tau=0}^{k} x_f(\tau)\Delta T, \sum_{\upsilon=0}^{k}\sum_{\tau=0}^{\upsilon} x_f(\tau)\Delta T\Delta T,\\ &x_{fd}(k), \sum_{\tau=0}^{k} x_{fd}(\tau)\Delta T, \sum_{\upsilon=0}^{k}\sum_{\tau=0}^{\upsilon} x_{fd}(\tau)\Delta T\Delta T,\\ &\dot{F}_l(k), F_f(k), \sum_{\tau=0}^{k} F_f(\tau)\Delta T]^\top.\end{aligned} \tag{8b}$$

By applying the ElGamal encryption using (8a) and (8b), V_f is obtained as follows:

$$V_f = \mathrm{Dec}^+(\Psi_f^E), \quad \Psi_f^E = f^\times(\mathrm{Enc}(\Phi_f), \mathrm{Enc}(\xi_f)), \tag{9}$$

which also enables the control operations and state variables of the follower device to be hidden.

E. ATTACK MODEL

This study considered the FDI attack under the assumption that the attackers know that the ElGamal encryption scheme is used for the encrypted control system and can access the network of the bilateral control system unauthorized. The attack exploits malleability in the ElGamal encryption, and an attack function $a : \mathcal{C} \times \mathbb{Z}_2 \to \mathcal{C}$ is introduced as follows,

$$a(c(k), \alpha) = (c_1, \alpha c_2 \bmod p), \quad \forall k \geq K, \tag{10}$$

where $\alpha \in \mathbb{Z}_{\geq 2}$ and $K \in \mathbb{Z}_{\geq 0}$ are an attack parameter and a step when the attack starts, respectively. The multiplication

of c_2 by α results in manipulating a corresponding plaintext m, i.e., $\mathrm{Dec}(a(c,\alpha)) = \alpha m$.

The bilateral control system involving the FDI attacks in (10) is shown in Fig. 6. The blocks, $a(x_{fd}^E, \alpha_1)$ and $a(F_{ld}^E, \alpha_2)$, are the functions that falsify the references in ciphertext: $x_{fd}^E = (c_1^{fd}, c_2^{fd})$ and $F_{ld}^E = (c_1^{ld}, c_2^{ld})$. The falsified references, denoted as $x_{fd}^{E'}$ and $F_{ld}^{E'}$, can be described using the attack function a and attack parameters $\alpha_i \in \mathbb{Z}_{\geq 2}, \forall i \in \mathcal{I} := \{1, 2\}$,

$$\begin{aligned} x_{fd}^{E'}(k) &= a(x_{fd}^E, \alpha_1) = (c_1^{fd}, \alpha_1 c_2^{fd} \bmod p), \\ F_{ld}^{E'}(k) &= a(F_{ld}^E, \alpha_2) = (c_1^{ld}, \alpha_2 c_2^{ld} \bmod p). \end{aligned} \tag{11}$$

The attacks correspond to the manipulation of the original plaintexts as follows,

$$\mathrm{Dec}(x_{fd}^{E'}(k)) = \alpha_1 x_{fd}(k), \ \mathrm{Dec}(F_{ld}^{E'}(k)) = \alpha_2 F_{ld}(k). \tag{12}$$

This causes the impairing of tracking performance without the operator noticing. Compared to general FDI attacks on ElGamal-type encrypted control systems, which generate white noise after decryption, the considered FDI attacks do not generate white noise, making their detection challenging. Studies, such as [39] and [40], have demonstrated that missing the detection for a while can destabilize bilateral control systems. Therefore, this study aimed to restore the falsified references to their original values in the event of an attack.

III. PROPOSED METHOD

This section proposes an algorithm for restoring references (12) that have been falsified by the FDI attacks to their original values, thereby canceling the impacts of the attacks on the bilateral control system.

A. ALGORITHM

The proposed algorithm detects the FDI attacks and restores the falsified references, and is summarized in

Algorithm 1 Detect and Cancel Algorithm

Require: $c(k) \in \{F_{ld}^{E'}(k), x_{fd}^{E'}(k)\}$,
$\sigma(k) \in \{0, 1, 2, 3\}, \hat{\alpha}(k), \bar{\gamma}, k \in \mathbb{Z}_{\geq 0}$
$\mathcal{H} = \{\mathcal{H}_l, \mathcal{H}_f\}$,
$\mathcal{H}_l = \{\dot{x}_l(0), \cdots, \dot{x}_l(k), F_l(0), \cdots, F_l(k)\}$,
$\mathcal{H}_f = \{\dot{x}_f(0), \cdots, \dot{x}_f(k), F_f(0), \cdots, F_f(k)\}$
Ensure: $c_c(k), \sigma(k+1), \hat{\alpha}(k+1), \bar{\gamma}$
$\gamma(k) \leftarrow$ (13)
if $\sigma(k) == 0$ **then**
  `CheckGamma`$(c(k), \gamma(k))$
else if $\sigma(k) == 1$ **then**
  `EstimateAttackParam`$(c(k), \bar{\gamma})$
else if $\sigma(k) == 2$ **then**
  `CancelAttack`$(c(k), \gamma(k), \hat{\alpha}(k))$
else if $\sigma(k) == 3$ **then**
  `ModifyAttackParam`$(c(k), \hat{\alpha}(k), \bar{\gamma})$
end if

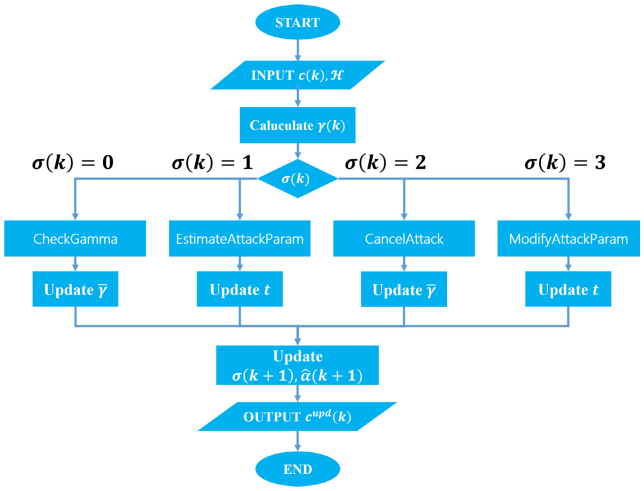


FIGURE 5. Flowchart of the proposed algorithm that selects one of the four functions based on $\sigma(k)$.

Algorithm 1, as well as its flowchart is illustrated in Fig. 5. The algorithm is executed at every step by the receivers, represented by the Detect and Cancel blocks in Fig. 6 on both the leader and follower sides. The inputs for the algorithms on the leader and follower sides are the encrypted reference and history of the data sequence, denoted as $c(k) = F_{ld}^{E'}(k)$ and $\mathcal{H}_l$ for the leader, and $c(k) = x_{fd}^{E'}(k)$ and $\mathcal{H}_f$ for the follower, respectively. The outputs are the restored or modified references, denoted as $c_c(k) = F_{ld}^{Ec}(k)$ for the leader and $c_c = x_{fd}^{Ec}(k)$ for the follower. Additionally, initial values of the estimated attack parameter $\hat{\alpha}_i \in \mathbb{Z}_{\geq 0}, \forall i \in \mathcal{I}$, a function selector $\sigma \in \{0, 1, 2, 3\}$ are set to $\hat{\alpha}_i(0) = 1$ and $\sigma(0) = 0$, respectively, and status variable *st* is set to *Normal*, indicating whether an attack is deemed to be occurring.

The flowchart in Fig. 5 is detailed, with the subscripts for variables indicating the leader and follower omitted because the algorithms are the same. The algorithm receives c and $\mathcal{H}$, updating four memory-type inner variables: $\sigma, \hat{\alpha}, \bar{\gamma} \geq 0$, and $t \in \mathbb{Z}_0$. Based on $\mathcal{H}$, the algorithm calculates γ_l in the leader side and γ_f in the follower side as follows:

$$\gamma_l(k) = \frac{RMS_l(k)}{M_l}, \quad \gamma_f(k) = \frac{RMS_f(k)}{M_f}, \tag{13}$$

where *RMS* represents the amplitude of monitored energy, considering the $k \geq T$,

$$\begin{aligned} RMS_l(k) &= \sqrt{\frac{1}{T}\sum_{\tau=0}^{T-1} (dPO_l(k-\tau))^2}, \\ RMS_f(k) &= \sqrt{\frac{1}{T}\sum_{\tau=0}^{T-1} \left(dPO_f(k-\tau)\right)^2}, \end{aligned} \tag{14}$$

Function 1 `CheckGamma`: Calculate $\gamma(k)$

Input: $c(k)$, $\gamma(k)$
Output: $c_c(k)$, $\sigma(k+1)$, $\bar{\gamma}$

1: **if** $\lfloor \gamma(k) \rceil \geq 2$ **then**
2: $\quad \sigma(k+1) \leftarrow 1$
3: **else**
4: $\quad \sigma(k+1) \leftarrow 0$
5: **end if**
6: $c_c(k) \leftarrow c(k)$
7: $\bar{\gamma} \leftarrow \gamma(k)$

Function 2 `EstimateAttackParam`: Estimate α

Input: $c(k)$, $\bar{\gamma}$
Output: $c_c(k)$, $\hat{\alpha}(k+1)$, $\sigma(k+1)$

1: $c_c(k) \leftarrow c(k)$
2: $c_{can} \leftarrow [a^{-1}(c,2), \cdots, a^{-1}(c, \lfloor \bar{\gamma} \rceil)]^\top$
3: $\mathcal{H}_c \leftarrow \{\mathcal{H}_c, c_{can}\}$
4: $\hat{\alpha}(k+1) \leftarrow 1$
5: **if** $t < T_\alpha$ **then**
6: $\quad \sigma(k+1) \leftarrow 1$
7: $\quad t \leftarrow t+1$
8: **else if** $t == T_\alpha$ **then**
9: $\quad$ **if** $\neg(\delta(\mathrm{Dec}(\mathcal{H}_c)) \geq [\theta, \cdots, \theta]^\top)$ **then**
10: $\quad\quad \sigma(k+1) \leftarrow 2$
11: $\quad\quad \hat{\alpha}(k+1) \leftarrow \max i+1$ s.t. $(\delta(\mathrm{Dec}(\mathcal{H}_c))_i \leq \theta$
12: $\quad\quad st \leftarrow Attacked$
13: $\quad$ **else**
14: $\quad\quad \sigma(k+1) \leftarrow 0$
15: $\quad$ **end if**
16: $\quad t \leftarrow 1$
17: $\quad \mathcal{H}_c \leftarrow \phi$
18: **end if**

and M is defined as:

$$M_l = \frac{1}{N}\sum_{\tau=0}^{N-1} RMS_l(k-\tau),$$
$$M_f = \frac{1}{N}\sum_{\tau=0}^{N-1} RMS_f(k-\tau), \tag{15}$$

where M is the average value of the corresponding *RMS* over steps of window length $N \in \mathbb{Z}_{\geq 0}$ assuming no attacks. M is pre-calculated offline data. The index γ indicates a ratio of *RMS* to M. Under normal conditions, this ratio is approximately one, as $RMS \simeq M$. However, the value tends to increase during the FDI attack. Ideally, if the impacts of the attack are perfectly canceled, the index will again approach one. Subsequently, based on σ, the algorithm selects and executes one of four functions: `CheckGamma`, `EstimateAttackParam`, `CancelAttack`, and `ModifyAttackParam`, detailed in **Functions 1**, **2**, **3**, and **4**, respectively.

Function 3 `CancelAttack`: Cancel the Effects of Attack

Input: $c(k)$, $\gamma(k)$, $\hat{\alpha}(k)$
Output: $c_c(k)$, $\hat{\alpha}(k+1)$, $\sigma(k+1)$, $\bar{\gamma}$

1: $c_c(k) \leftarrow a^{-1}(c(k), \hat{\alpha}(k))$
2: $\hat{\alpha}(k+1) \leftarrow \hat{\alpha}(k)$
3: **if** $k \bmod T == 0$ and $\lfloor \gamma(k) \rceil \geq 2$ **then**
4: $\quad \sigma(k+1) \leftarrow 3$
5: **else**
6: $\quad \sigma(k+1) \leftarrow 2$
7: **end if**
8: $\bar{\gamma} \leftarrow \gamma(k)$

Function 4 `ModifyAttackParam`: Modify $\hat{\alpha}$

Input: $c(k)$, $\hat{\alpha}(k)$, $\bar{\gamma}$
Output: $c_c(k)$, $\sigma(k+1)$, $\hat{\alpha}(k+1)$

1: $c_c(k) \leftarrow a^{-1}(c(k), \hat{\alpha}(k))$
2: $\hat{\alpha}(k+1) \leftarrow \hat{\alpha}(k)$
3: $c_{can} \leftarrow a^{-1}(c(k), \lfloor \bar{\gamma} \rceil \hat{\alpha}(k))$
4: $\mathcal{H}_c \leftarrow \{\mathcal{H}_c, c_{can}\}$
5: **if** $t < T_\alpha$ **then**
6: $\quad \sigma(k+1) \leftarrow 3$
7: $\quad t \leftarrow t+1$
8: **else if** $t == T_\alpha$ **then**
9: $\quad$ **if** $\delta(\mathrm{Dec}(\mathcal{H}_c)) \leq \theta$ **then**
10: $\quad\quad \hat{\alpha}(k+1) \leftarrow \lfloor \bar{\gamma} \rceil \hat{\alpha}(k)$
11: $\quad$ **end if**
12: $\quad \sigma(k+1) \leftarrow 2$
13: $\quad t \leftarrow 1$
14: $\quad \mathcal{H}_c \leftarrow \phi$
15: **end if**

1) FUNCTION: `CHECKGAMMA`

This function determines whether $\lfloor \gamma(k) \rceil \geq 2$. It assumes $c(k)$ and $\gamma(k)$ as inputs. If $\lfloor \gamma(k) \rceil \geq 2$, $\sigma(k+1)$ is set to one; Otherwise, $\sigma(k+1)$ is set to zero. In line 6, $c_c(k)$ is assigned the value of $c(k)$. In line 7, $\bar{\gamma}$ is updated to the current $\gamma(k)$, which is used in `EstimateAttackParam`. This function does not update $\hat{\alpha}$, indicating $\hat{\alpha}(k+1) = \hat{\alpha}(k)$.

2) FUNCTION: `ESTIMATEATTACKPARAM`

This function determines whether the FDI attack has occurred by estimating α, requiring T_α steps. Initially, $c_c(k)$ is set to $c(k)$. In line 2, a vector of $\left[2\ 3\ \cdots\ \lfloor \bar{\gamma} \rceil\right]$ is prepared as the candidates of $\hat{\alpha}$, and calculates the component of $c_{can}(k) \in \mathcal{C}^{\lfloor \bar{\gamma} \rceil}$, representing the list of the canceled encrypted references for each candidate for $c(k)$. A cancellation function is defined as follows:

$$a^{-1}(c, \hat{\alpha}) = (c_1, c_2(\hat{\alpha})^{-1} \bmod p), \tag{16}$$

where $(\cdot)^{-1}$ is the modular inverse modulo p. This function serves as the inverse of (10); if $\hat{\alpha} = \alpha$, then (16) returns the ciphertext before falsifying, i.e., $a^{-1}((c_1, \alpha c_2 \bmod p), \hat{\alpha}) = (c_1, \alpha c_2(\hat{\alpha})^{-1} \bmod p) = (c_1, c_2)$, effectively

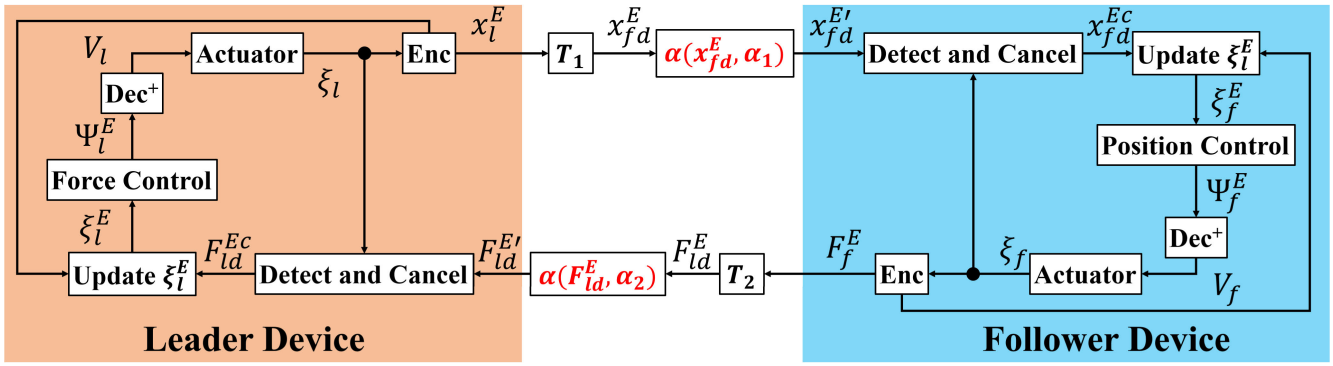


FIGURE 6. Overview of the block diagram in this study. The FDI attack (10) is conducted at $a(\cdot,\cdot)$ block. The proposed algorithm runs at Detect and Cancel block.

canceling the impact of the FDI attack. $c_{can}(k)$ is computed as $[a^{-1}(c(k),2),\cdots,a^{-1}(c(k),\lfloor\bar{\gamma}\rceil)]^\top$. In line 3, $c_{can}(k)$ is stored in $\mathcal{H}_c \in \mathcal{C}^{\lfloor\bar{\gamma}\rceil\times T_\alpha}$, and this storage is repeated over T_α steps (lines 4-7), with t tracking the number of repetitions. Lines 9-15 compute the variance for each component in $\mathcal{H}_c$:

$$\delta(\mathrm{Dec}(\mathcal{H}_c)) = \frac{1}{T_\alpha}\sum_{\tau=0}^{T_\alpha-1}(\mathrm{Dec}(c_{can}(k-\tau))-\mu)^2, \qquad (17)$$

where $\mu = \frac{1}{T_\alpha}\sum_{\tau=0}^{T_\alpha-1}\mathrm{Dec}(c_{can}(k-\tau))$, and Dec denotes the decryption of each component of vector $c_{can}(k)$. If at least one component of δ is less than the threshold θ, the logical expression in line 9 evaluates to true, assigning $\sigma(k+1)$ as two and updating $\hat{\alpha}(k+1)$ to $i+1$, where i is the index of δ. The index shift $i+1$ reflects the mapping of index of $i=1,2,\cdots$ to the candidate $\hat{\alpha}=2,3,\cdots$. The status variable st is set to *Attacked* in line 12. If the condition is false, $\sigma(k+1)$ is reset to zero in line 14. Finally, t and $\mathcal{H}_c$ are reinitialized in lines 16 and 17.

3) FUNCTION: CANCELATTACK

This function computes the encrypted reference before falsification, effectively canceling the effects of the FDI attack. $c(k)$, $\gamma(k)$, and $\hat{\alpha}(k)$ are received as inputs, and then the output $c_c(k)$ is calculated. In line 1, $c_c(k)$ is determined using the cancellation function defined in (16). Line 2 maintains the value of $\hat{\alpha}(k)$ for the next step by setting $\hat{\alpha}(k+1)$. Typically, $\sigma(k+1)$ is set to two; however if $\lfloor\gamma(k)\rceil \geq 2$ holds every T step, then $\sigma(k+1)$ is set to three to trigger an update of $\hat{\alpha}$. This condition is to judge whether the falsified signal has been modified perfectly. This is judged using γ_l and γ_f, which are calculated only using the modified data after T step. Repetition with "$k \bmod T == 0$" is to perform **Function 4** again if the modification fails owing to the energy fluctuations other than the attack. Line 8 updates $\bar{\gamma}$ with the value of $\gamma(k)$ for use in `ModifyAttackParam`.

4) FUNCTION: MODIFYATTACKPARAM

This function is responsible for updating $\hat{\alpha}$. $c(k)$, $\hat{\alpha}(k)$, and $\bar{\gamma}$ are received as inputs. In lines 1 and 2, $c_c(k)$ is recalculated using (16), and $\hat{\alpha}(k+1)$ is directly set to $\hat{\alpha}(k)$. Lines 3 and 4 involve calculating $c_{can}(k)$ and storing in the history $\mathcal{H}_c$. This procedure is repeated T_α times as delineated in lines 5-7. Line 8 calculates $\delta(\mathrm{Dec}(\mathcal{H}_c))$ using (17). In this function, $\delta(\mathrm{Dec}(\mathcal{H}_c))$ is scalar, which is different from that in **Function 2**. Line 9 determines if $\delta(\mathrm{Dec}(\mathcal{H}_c)) \leq \theta$. If true, $\hat{\alpha}(k+1)$ is updated to $\lfloor\bar{\gamma}\rceil\hat{\alpha}(k)$ in line 10; the operation $\lfloor\bar{\gamma}\rceil\hat{\alpha}(k)$ will be detailed in the following section. Finally, lines 12-14 reset $\sigma(k+1)$ to two, and initialize t and $\mathcal{H}_c$.

B. THEORETICAL RESULTS

This section discusses how the proposed algorithm can enable the detection of the attack and restoration of the falsified references in a finite step. **Lemma 1** and **Lemma 2** provide critical support for **Theorem 1** and are discussed as follows.

Lemma 1: Assume that $x_f \simeq x_{fd}$ and $F_l \simeq F_{ld}$. For the FDI attacks with attack parameters, α_i, $\forall i \in \mathcal{I}$, in (II-E), γ_l and γ_f in (13) satisfy the following inequalities, respectively:

$$\gamma_f(k) \simeq \frac{\alpha_1}{\hat{\alpha}_1(k)}, \quad \gamma_l(k) \simeq \frac{\alpha_2}{\hat{\alpha}_2(k)}, \quad \forall k \geq K + K_2,$$

where $K_2 \in \mathbb{Z}_0$ is sufficiently large; $\hat{\alpha}_i$ is updated in **Function 2** and **Function 4**.

Proof: At step K, $\dot{x}_{fd}$ and F_{ld} are multiplied by α_1 and α_2, respectively. The leader and follower algorithms modify the attacked references using $a^{-1}(c',\hat{\alpha})$ to yield $\alpha_2 F_{ld}/\hat{\alpha}_2$ and $\alpha_1 x_{fd}/\hat{\alpha}_1$, respectively. Denoting F_l' and x_f' as the control outputs affected by the FDI attack and modification, it is valid to assume that $F_l' \simeq \alpha_2 F_{ld}/\hat{\alpha}_2$, and $x_f' \simeq \alpha_1 x_{fd}/\hat{\alpha}_1$. Let us denote dPO_l' and dPO_f' as dPO_l and dPO_f under attack, respectively, expressed as follows, $dPO_l'(k) = (\dot{x}_l(k-T_2)F_l'(k-T_2) - \dot{x}_l(k-T_1-T_2)F_l'(k)) \simeq (\alpha_2/\hat{\alpha}_2)(\dot{x}_l(k-T_2)F_l(k-T_2) - \dot{x}_l(k-T_1-T_2)F_l(k)) = \alpha_2 dPO_l(k)/\hat{\alpha}_2$ and $dPO_f'(k) = (\dot{x}_f'(k)F_f(k-T_1-T_2) - \dot{x}_f'(k-T_1)F_f(k-T_1)) \simeq (\alpha_1/\hat{\alpha}_1)(\dot{x}_f(k)F_f(k-T_1-T_2) - \dot{x}_f(k-T_1)F_f(k-T_1)) = \alpha_1 dPO_f(k)/\hat{\alpha}_1$. When steps $K_2 > T$ pass after the attack occurs, (14) can be updated with dPO'. Thus, from the property of the ratios (13) and the assumption that $x_f \simeq x_{fd}$

and $F_l \simeq F_{ld}$, the following approximation is held,

$$\sqrt{\frac{1}{T}\sum_{\tau=0}^{T-1}\left(dPO'_l(k-\tau)\right)^2} \simeq \frac{\alpha_2}{\hat{\alpha}_2}\sqrt{\frac{1}{T}\sum_{\tau=0}^{T-1}(dPO_l(k-\tau))^2}.$$

From the definition of γ, the following approximation is held,

$$\gamma_l(k) \simeq \frac{\alpha_2}{\hat{\alpha}_2}\frac{RMS_l(k)}{M_l} \simeq \alpha_2/\hat{\alpha}_2,$$

and similarly, $\gamma_f(k) \simeq \alpha_1/\hat{\alpha}_1$ holds for $k \geq K + K_2$. ■

Remark 1: The assumption in **Lemma 1** is valid when the leader device operates in a constant cycle. The approximations $x_f \simeq x_{fd}$ and $F_l \simeq F_{ld}$ indicate that the controlled outputs sufficiently track the references for both the leader and follower sides. Because identifying the mathematical conditions that achieve such tracking is challenging, we rely on experimental observations to validate these approximations. In this situation, as shown in Fig. 6, because the attack functions, $a(x^E_{fd}, \alpha_1)$ and $a(F^E_{ld}, \alpha_2)$ are located after the time delay blocks, T_1 and T_2, the effect of the attack appears in F_l and x_f without time delay. However, if the attack occurs before the time delay, the effect appears with the time delay. Therefore, time delay steps, T_1 and T_2, are added to K_2, but the conclusion does not change.

Lemma 2: Denoting the encrypted reference corresponding to $m \in \mathcal{M}$ as $c \in \mathcal{C}$, the encrypted reference falsified by (10) as c', and the modified encrypted reference as $c_{can} = a^{-1}(c', \check{\alpha})$. Subsequently, the decryption of c_{can} is expressed by,

$$\mathrm{Dec}(c_{can}) = \begin{cases} \nu m, & \text{if } \alpha = \check{\alpha}\nu, \\ (\alpha m + lp)/\check{\alpha}, & \text{otherwise,} \end{cases}$$

where $\nu \in \mathbb{Z}_{\geq 1}$ and $l \in \mathbb{Z}$.

Proof: The decryption of c_{can} is $\mathrm{Dec}(c_{can}) = (\check{\alpha})^{-1}\alpha m \bmod p$. When $\check{\alpha}$ is a divisor of α, there exists $\nu \in \mathbb{Z}_{\geq 1}$ such that $\alpha = \check{\alpha}\nu$ holds. In this case, $\mathrm{Dec}(c_{can}) = (\check{\alpha})^{-1}\check{\alpha}\nu m \bmod p = \nu m \bmod p$. Conversely, when $\check{\alpha}$ is not a divisor of α, $\alpha m \bmod \check{\alpha} = i, \forall i \in \{0, 1, \cdots, \check{\alpha} - 1\}$. We can choose $l \in \mathbb{Z}$ such that $lp \bmod \check{\alpha} = \check{\alpha} - i$, which is consistent with the properties of modular arithmetic. This results in $(\alpha m + lp) \bmod \check{\alpha} = i + \check{\alpha} - i = \check{\alpha} = 0$, implying $\alpha m + lp$ is a multiple of $\check{\alpha}$. Using $q_l \in \mathbb{Z}$, $\alpha m + lp$ can be expressed as $\check{\alpha} q_l$, where $q_l = (\alpha m + lp)/\check{\alpha}$. Thus, $\mathrm{Dec}(c_{can})$ is given by, $\mathrm{Dec}(c_{can}) = (\check{\alpha})^{-1}\alpha m \bmod p = (\check{\alpha})^{-1}(\alpha m + lp) \bmod p = (\check{\alpha})^{-1}\check{\alpha} q_l \bmod p = q_l \bmod p = (\alpha m + lp)/\check{\alpha}$. ■

Remark 2: **Lemma 2** indicates that when $\hat{\alpha}$ is not a divisor of α, the time series data of $\mathrm{Dec}(c_{can})$ exhibit significant fluctuations compared to when $\hat{\alpha}$ is a divisor of α. This oscillation behaves similarly to a pseudo random number generator. Therefore, we can appropriately tune θ to ensure that $\hat{\alpha}$, satisfying the condition of line 9 in **Function 2** and line 9 in **Function 4**, is a divisors of α.

Lemma 1 and **Lemma 2** support the following **Theorem 1**.

Theorem 1: Considering the bilateral encrypted controllers (7) and (9), the proposed algorithm is implemented on the leader and follower sides. Assume the attacker initiates the FDI attack (II-E) starting at step $K \in \mathbb{Z}_0$. Subsequently, there exists a finite step K_1 such that $\hat{\alpha}(k) = \alpha, \ \forall k \geq K + K_1$.

Proof: We aim to establish that K_1 can be a finite integer. According to the proposed algorithm and **Lemma 1**, $\gamma(k)$ begins to converge to specific finite values by step K_2, which exceeds T (i.e., $T < K_2$). This results in σ being updated to be one in **Function 1** K_2 steps following the attack's initiation.

In **Function 2**, the algorithm stores $\mathcal{H}_c$ using the cancellation function of a vector $\begin{bmatrix}2 & 3 & \cdots & \lfloor\bar{\gamma}\rceil\end{bmatrix}^\top$ and evaluates the condition at line 9 using (17) and **Lemma 2**. If the current situation meets $r = 1$ as specified in **Lemma 2**, then $\hat{\alpha} = \lfloor\bar{\gamma}\rceil$ holds, indicating that the attack parameter is correctly identified as $\hat{\alpha} = \alpha$. Meanwhile, if $r \neq 1$, then $\hat{\alpha} \neq \alpha$, triggering **Function 2** to execute once within the algorithm; thus necessitating T_α steps to update $\hat{\alpha}$ and the status st in lines 11 and 12, respectively.

When **Function 3** is activated, it continually outputs updated control signals c_c for T steps without updating the estimated attack parameter $\hat{\alpha}$. If $r = 1$ is involved, the actions in line 9 result in a perfect cancellation of the attack's effects because $r = 1$ implies that $\lfloor\bar{\gamma}\rceil = 1$. Therefore, the steps needed to correct the falsified references are finite, specifically, $K_1 = T + T_\alpha + 1$. In instances where $r \neq 1$, it assumes T steps to adjust σ to three, indicating that **Function 4** will be selected in subsequent steps.

Function 4 continuously processes the update of $\hat{\alpha}$ at line 2 or 10 and the attack cancellation at line 1 for T_α steps, after which it reverts to **Function 3** by setting σ to two at line 12.

If the update $\hat{\alpha}(k+1)$ at line 10 indicates $r = 1$ according to **Lemma 2**, then perfect cancellation is realized in **Function 3**; otherwise, the cancellation fails, and scenarios where $r \neq 1$ in **Function 3** are reassessed. Because $r \neq 1$ implies that $\lfloor\bar{\gamma}\rceil \geq 2$ as noted in **Remark 2**, the algorithm opts to update $\hat{\alpha}(k+1)$ such that $\hat{\alpha}(k+1) > \hat{\alpha}(k)$. Consequently, the current $\gamma(k)$ in **Function 3** is reduced compared to $\bar{\gamma}$ used in **Function 4** during the prior step. The iterative execution of **Function 3** and **Function 4** gradually leads $\gamma(k)$ to converge to one as k increases. Therefore, a γ value of one signals the cessation of the attack cancellation, which implies that the restoration of the attached references is achieved within the finite steps: $K_1 = T + T_\alpha + (\kappa - 1)(T + T_\alpha) + 1 = \kappa(T + T_\alpha) + 1$, where $\kappa - 1$ represents the number of updates performed on $\hat{\alpha}$ at line 10 in **Function 4**. ■

Theorem 1 states that the proposed algorithm detects the FDI attack and accurately estimates the true values of attack parameters α_1 and α_2, effectively canceling the impact of the FDI attack. However, as noted in **Remark 1**, the assumptions that $x_f \simeq x_{fd}$ and $F_l \simeq F_{ld}$ are challenging to justify mathematically. In the following section, this study examines the reasonableness of these assumptions and experimentally evaluates the effectiveness of the proposed algorithm.

IV. EXPERIMENTAL DEMONSTRATION

This section experimentally demonstrates that the proposed algorithm allows the detection and cancellation of the impact of the FDI attacks on the constructed bilateral control system.

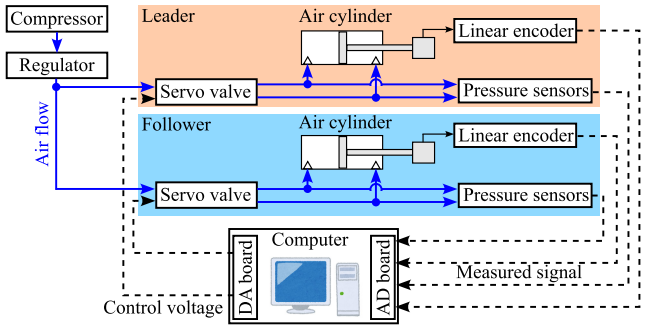


FIGURE 7. Pneumatic circuit and control systems of the experimental setup.

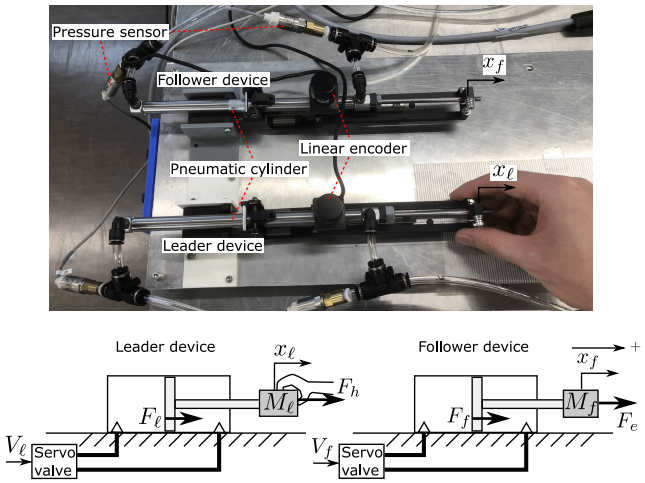


FIGURE 8. Overview and schematic diagrams of the leader and follower devices consisting of masses and pneumatic cylinders.

A. EXPERIMENTAL SETUP

The experimental setup's pneumatic circuit and control systems are shown in Fig. 7, where blue lines indicate airflow and black dotted lines represent electrical signals. The leader and follower devices of the constructed bilateral control system were one-degree-of-freedom, single-rod pneumatic cylinders of identical structure, as shown in Fig. 8. This setup verifies the effectiveness of the proposed algorithm for broader discussions on complex systems. Contrary to the commonly used second-order delay electric drive systems, the pneumatic drive system is a third-order delay system, which introduces additional complexity.

Linear encoders and pressure sensors measure the position and pressure of the leader and follower devices. The control voltages, V_l and V_f, are calculated using the sensor values according to (7) and (9), and then input to the servo valves. The leader device was set up for force control, while the follower device managed position control, which incorporated aspects of force control. The servo valves adjusted the pneumatic cylinder pressures based on the control voltages. An operator periodically maneuvered the leader device with frequency f and amplitude A, while pneumatic pressure actuates the follower device. In this case, as a hand moved the cylinder, F_h changes in balance with cylinder force and friction force. Regarding the external force, an obstacle did not place in forward of the follower device; thus so $F_e = 0$.

TABLE 1. Experimental parameters.

Sampling frequency [Hz]	1000
Supplied Pressure [kPa]	250
Communication delay step (Leader to Follower) T_1	25 (0.025 s)
Communication delay step (Follower to Leader) T_2	25 (0.025 s)
Length of key [bit]	64
Force proportional gain (Leader) K_{ap} [V/N]	1.0
Force integral gain (Leader) K_{ai} [V/(N·s)]	0.2
Force differential gain (Leader) K_{ad} [V·s/N]	0.0
Force proportional gain (Follower) K_{ap} [V/N]	1.0
Force integral gain (Follower) K_{ai} [V/(N·s)]	0.2
Force differential gain (Follower) K_{ad} [V·s/N]	0.0
Position proportional gain (Follower) K_{pp} [N/m]	700.0
Position integral gain (Follower) K_{pi} [N/(m·s)]	40.0
Position differential gain (Follower) K_{pd} [N·s/m]	0.0
Motion frequency f [Hz]	0.25
Motion amplitude A [m]	0.03
Window width T of (14)	10000 (10 s)
Step length N of (15)	20000 (20 s)
Step length T_α of (17)	100 (0.1 s)
Leader-side threshold θ_l [N^2]	25.0
Follower-side threshold θ_f [m^2]	0.0009

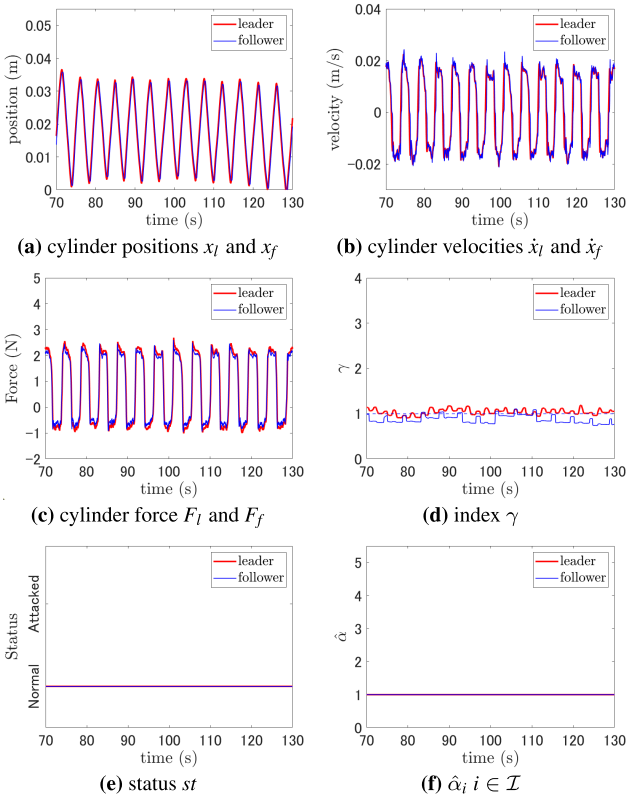
TABLE 2. Type and nature of attacks conducted in the experiment.

Attack scenario	Contents
(i) $\alpha_1 = 1, \alpha_2 = 1$	No attack.
(ii) $\alpha_1 = 1, \alpha_2 = 2$	Doubling the encrypted reference F_{ld} input to the leader.
(iii) $\alpha_1 = 1, \alpha_2 = 4$	Quadrupling the encrypted reference F_{ld} input to the leader.
(iv) $\alpha_1 = 2, \alpha_2 = 1$	Doubling the encrypted reference x_{fd} input to the follower.

Parameter values used in the experiments are listed in TABLE 1. The communication time delays T_1 and T_2 were set to 25. The window width of (14), denoted as T, was set as 10000 to encompass two cycles of the input motion. The step length of (15), N, was set as 20000 to include five cycles within the window. Thresholds for the condition $\neg(\delta(\mathrm{Dec}(\mathcal{H}_c)) \geq [\theta, \cdots, \theta]^\top)$, denoted as θ_l and θ_f, were determined based on the variances of the reference signals. With the input motion in this study, the amplitudes of the leader's force and follower's position reached maximum values of 5.0 N and 0.03 m, respectively. Thus, θ_l was set to 5.0^2 N and θ_f to 0.03^2 m. Other parameters were decided experimentally by trial and error.

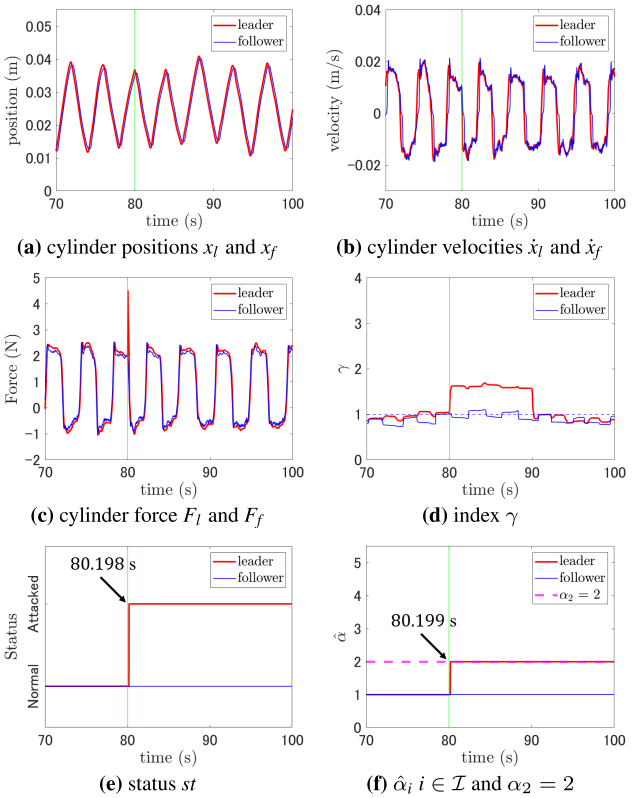
B. ATTACK SCENARIOS

The detailed attack scenarios are outlined in TABLE 2. Our experiments considered the attacks on the encrypted reference of the leader, $F_{ld}^{E'}$, or the encrypted reference of the follower, $x_{fd}^{E'}$. The FDI attacks were executed every step according to (II-E) and commenced at 80 s after the operation begins. The experiments explored the following cases: (i) $\alpha_1 = 1$, $\alpha_2 = 1$: No attack is simulated; (ii) $\alpha_1 = 1$, $\alpha_2 = 2$: The attacker falsifies the leader's reference force, F_{ld} to twice its intended value; (iii) $\alpha_1 = 1$, $\alpha_2 = 4$: The



(a) cylinder positions x_l and x_f (b) cylinder velocities $\dot{x}_l$ and $\dot{x}_f$
(c) cylinder force F_l and F_f (d) index γ
(e) status st (f) $\hat{\alpha}_i\ i \in \mathcal{I}$

FIGURE 9. Result of the attack scenario (i): There is no attack.



(a) cylinder positions x_l and x_f (b) cylinder velocities $\dot{x}_l$ and $\dot{x}_f$
(c) cylinder force F_l and F_f (d) index γ
(e) status st (f) $\hat{\alpha}_i\ i \in \mathcal{I}$ and $\alpha_2 = 2$

FIGURE 10. Results of attack scenario (ii) with $\kappa = 1$.

attacker falsifies the leader's reference force, F_{ld} to four times its intended value; (iv) $\alpha_1 = 2, \alpha_2 = 1$: The attacker falsifies the follower's reference position, x_{fd}, to twice its intended value.

C. EXPERIMENTAL RESULTS

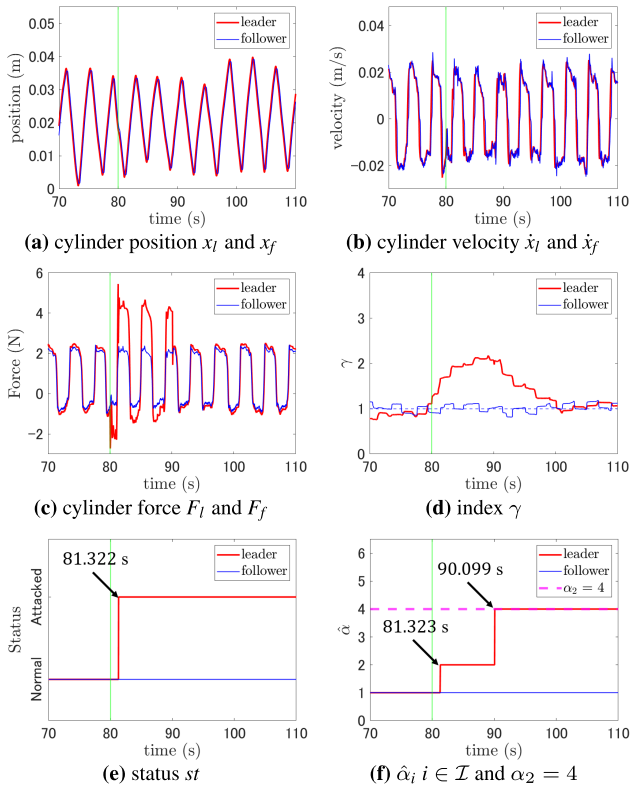
The experimental results of the attack scenarios (i)-(iv) are shown in Figs. 9-12, respectively. Subfigures (a)-(f) display time responses of position, velocity, and force of the cylinder, index, status, and estimation of attack parameters, respectively. In these figures, the red and blue lines represent the leader and follower, respectively. The green line in Figs. 10-12 marks the start of the attack at 80 s. The magenta broken line indicates the true values of attack parameters α_1 and α_2.

According to Fig. 9, the follower and leader devices were controlled to track the leader's position and follower's force, respectively. Fig. 9b displays that both γ_l and γ_f hovered around one, indicating the absence of attacks. Fig. 9e confirms that the algorithm judged the status as *Normal*, and Fig. 9f shows that $\hat{\alpha}_1$ and $\hat{\alpha}_2$ consistently remained at one, resulting in no false positives.

Fig. 10 illustrates that the follower device was controlled to track the position of the leader, while the leader force was manipulated to be twice the value of the force of the follower owing to the attack (12). Fig. 10b indicates that after the attack began, γ_l approached 1.8 and returned to approximately one in 10 s (the window width), validating **Lemma 1**. Fig. 10e shows that after the attack initiation, 0.198 s, the algorithm detected it as the status changed to *Attacked*, and the falsified reference was subsequently modified as demonstrated in Fig. 10f. When the index γ_l returned to approximately one, the algorithm recognized the completion of restoring the falsified reference. The process took approximately 90 s with κ equaling one. Consequently, confirming that the algorithm could detect and cancel the impact of the attack in a finite number of steps.

Fig. 11 depicts the follower device controlled to track the position of the leader, while the force of the leader tracked the falsified reference value of the force of the follower. Fig. 11d shows γ_l rising from 80 s and decreasing back to one around 100 s. Figs. 11e and 11f reveal that 1.322 s post-attack the algorithm detected the anomaly based on the status change to *Attacked*, and then the falsified reference was modified twice to align with the true value $\alpha_2 = 4$. The algorithm recognized the restoration completion at the second modification at 90.099 s, and κ was two. This case verified that **Function 4** operated correctly.

Fig. 12 demonstrates that the leader's force was controlled to track the reference of the follower's force, while the position of the follower tracked the falsified reference value of the position of the leader. Fig. 12d shows that $\lfloor \gamma_l \rceil$ and $\lfloor \gamma_f \rceil$ in **Function 1** exceeded two after the attack, while in the window width, γ_l and γ_f returned approximately to one. The significant numbers in both γ measurements are owing to the impulsive behavior of the follower device's

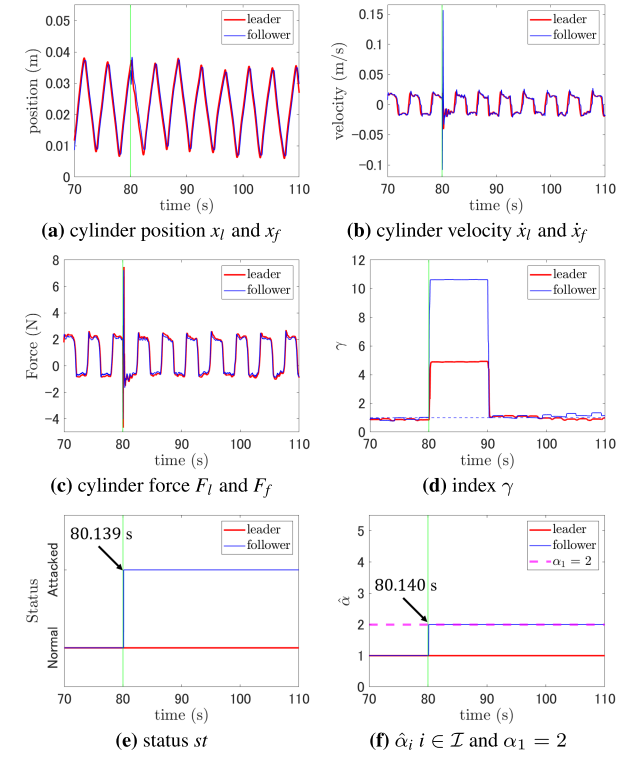


(a) cylinder position x_l and x_f **(b)** cylinder velocity $\dot{x}_l$ and $\dot{x}_f$

(c) cylinder force F_l and F_f **(d)** index γ

(e) status st **(f)** $\hat{\alpha}_i$ $i \in \mathcal{I}$ and $\alpha_2 = 4$

FIGURE 11. Result of attack scenario (iii) with $\kappa = 2$.



(a) cylinder position x_l and x_f **(b)** cylinder velocity $\dot{x}_l$ and $\dot{x}_f$

(c) cylinder force F_l and F_f **(d)** index γ

(e) status st **(f)** $\hat{\alpha}_i$ $i \in \mathcal{I}$ and $\alpha_1 = 2$

FIGURE 12. Result of attack scenario (iv) with $\kappa = 1$.

velocity and force resulting from the falsified reference of the follower's position. Here, $\lfloor \gamma_l \rceil$ exceeded two, while the

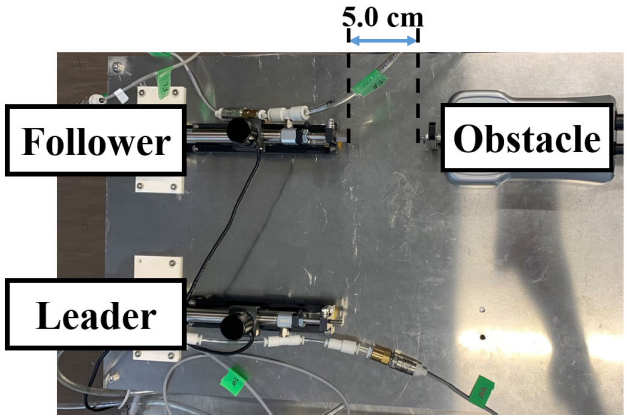


FIGURE 13. Overview of the experimental equipment when there is an obstacle 0.05 m in front of the follower device.

inequality condition in line 9 in **Function 2** are not satisfied. Regarding $\lfloor \gamma_f \rceil$, there existed at least one component that meets the condition in line 9 in **Function 2**. Consequently, the algorithm determined that the follower's reference was attacked while that of the leader was not. Figs. 12e and 12f demonstrate that 0.139 s after the attack, the algorithm detected the anomaly based on the follower's status changing to *Attacked*, and then the falsified reference was modified to reach the true value $\alpha_1 = 2$. This scenario confirmed that **Function 2** functioned correctly.

The experimental results confirmed that the proposed algorithm effectively detected and restored the FDI attack on the encrypted references, as the estimated attack parameters align with the true values within a finite number of steps, as supported by **Theorem 1**.

V. DISCUSSION

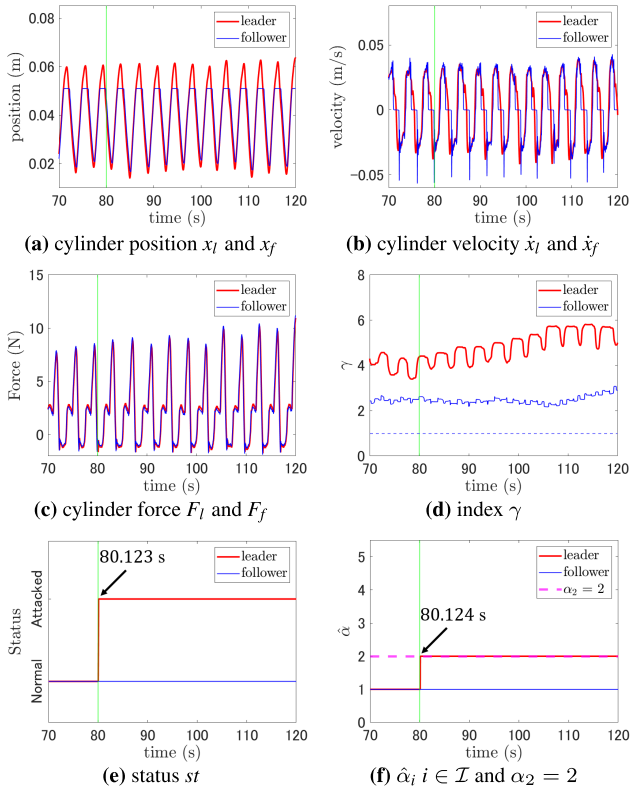
This section discusses the challenges of the method, thereby realizing of more secure control systems.

A. ADDITIONAL VALIDATION UNDER CONTACT WITH OBSTACLE

This section considers another scenario where the follower device physically contacts the environment, indicating the external force F_e is not zero, to experimentally validate the effectiveness of the proposed algorithm. Practically, several situations involve obstacles, such as biological tissues (bones, organs, etc.) in tele-surgical operations. Note that such situations are excluded in **Theorem 1**.

Fig. 13 illustrates the bilateral control system and an obstacle. The distance between the follower cylinder and the obstacle is 0.05 m. The control systems, along with the implemented algorithms and their parameters, are consistent with those used in the experiments, which have been discussed in the previous section. The attack scenario, labeled as (v), involves falsifying the leader's reference to be twice its intended value.

The results of this attack scenario are displayed in Fig. 14, where the format and line meanings are consistent with those



(a) cylinder position x_l and x_f (b) cylinder velocity $\dot{x}_l$ and $\dot{x}_f$

(c) cylinder force F_l and F_f (d) index γ

(e) status st (f) $\hat{\alpha}_i$ $i \in \mathcal{I}$ and $\alpha_2 = 2$

FIGURE 14. Result of attack scenario (v) with $\kappa = 1$.

in Figs. 9-12. From Figs. 14a and 14c, the follower device was controlled to track the position of leader; the position of follower did not exceed 0.05 m; the forces of both the leader and follower became impulsive and larger than in scenarios (i) − (iv), caused by contact with the obstacle. The leader's force was controlled to track the follower's force, indicating successful restoration of the falsified reference. Fig. 14d shows that the γ values for both leader and follower exceeded two owing to the increased energy from the external force. Although $\lfloor \gamma_f \rceil$ exceeded two, the inequalities at line 9 in **Function 2** are not satisfied, implying that the algorithm did not detect an attack on the follower's reference. Meanwhile, for $\lfloor \gamma_l \rceil$, at least one component satisfies the inequality, resulting in the detection of the FDI attack on the leader's reference. Figs. 14e and 14f show that the algorithm detected the attack 0.123 s after its commencement; subsequently, $\hat{\alpha}_2$ adjusted to the true value $\alpha_2 = 2$. Consequently, the experimental result demonstrates the effectiveness of the algorithm in the situation involving contact with an obstacle. The presence of F_e affects the detectability. Compared to the discussion in Section IV where F_h is present but F_e is not, the presence of F_h and F_e increases the energy, making the detection easier.

B. TIME TO ATTACK DETECTION AND CANCELLATION

This section discusses the steps required to restore the attacked reference, and considers appropriate parameter tuning for the proposed algorithm. TABLE 3 lists the times taken to detect the attack and restore the falsified references for each of the attack scenarios (ii) to (v). As shown in TABLE 3, detection time ranges from as fast as 0.123 s to as late as 1.322 s, while restoration time varies from as quick as 0.124 s to as prolonged as 10.099 s. Considering κ, in the attack scenario (ii), (iv), and (v), $\kappa = 1$ while in attack scenario (iii), $\kappa = 2$.

TABLE 3. Time to detect the FDI attack and cancel the effects in the experiments. Cancellation implies to cancel the FDI attack's effect perfectly, that is, $\hat{\alpha} = \alpha$.

Attack scenario	Detection	Restoration
(ii) $\alpha_1 = 1, \alpha_2 = 2$	0.198 s	0.199 s
(iii) $\alpha_1 = 1, \alpha_2 = 4$	1.322 s	10.099 s
(iv) $\alpha_1 = 2, \alpha_2 = 1$	0.139 s	0.140 s
(v) $\alpha_1 = 1, \alpha_2 = 2$ with the obstacle	0.123 s	0.124 s

For example, in tele-surgical operations, the restoration time affects the burden on the patient; thus the proposed algorithm must be designed to minimize this time. As shown in **Theorem 1**, the restoration time K_1 is expressed as $\kappa(T + T_\alpha) + 1$. Therefore, it is important to tune the parameters to shorten the restoration time. K_1 is composed of T and T_α, which can be adjusted by system users, and κ, which cannot be adjusted by them. Regarding the former, future work will consider a parameter-tuning method according to the dynamics of the system. For the latter, although system users cannot directly control κ, devising a method to estimate κ from the impact of the attack will be a challenge in future work.

C. TYPES OF INPUT MOTION

This section examines the types of input motion to which the proposed algorithm can be applied, aiming to expand its applicable domain. This domain is determined by calculating γ. As defined in (14), γ is calculated using $RMS(k)$, which represents online data, and M, which represents offline data. Thus, the domain of the input motion should align closely with that of the offline data. For periodic motion, there exist various types depending on the frequency f. When f is higher than that used in the offline data, the γ values exceed one owing to increased energy, which typically does not result in false positives, as demonstrated in Section IV. Conversely, when f is lower than that of the offline data, the γ values fall below one owing to decreased energy, which may prevent the detection of FDI attacks. Considering the diverse domains applicable to various input operations, it is important to expand the domain in which the proposed algorithm is effective. This expansion will be studied in the future work.

VI. CONCLUSION

In this study, an algorithm to detect and restore FDI attacks on a bilateral encrypted control system was proposed. In particular, the FDI attack model considered here involved falsifying encrypted reference signals using the malleability of encryption schemes, which can significantly impair the performance of control systems. The proposed algorithm

detects and restores such attacks by estimating the attack parameters based on the amount of energy change throughout the system, despite the users lacking information about the attacker's model. Theoretically, the proposed method can detect and restore functionality in finite steps. Moreover, from the experiments, the authors detected and restored the FDI attack, supporting the theoretical results; thus the proposed algorithm can enhance the security of cyber-physical systems.

In future work, the authors will refine the proposed algorithm to further enhance system security. This will include developing parameter-tuning methods to shorten restoration times and expanding the applicable domain of input motion. These improvements are expected to contribute to the realization of more secure cyber-physical systems.

ACKNOWLEDGMENT

(Katsumasa Kosha1 and Tetsuro Miyazaki contributed equally to this work.)

REFERENCES

[1] T. B. Sheridan, "Telerobotics," *Automatica*, vol. 25, no. 4, pp. 487–507, Jul. 1989, doi: 10.1016/0005-1098(89)90093-9.

[2] T. B. Sheridan, "Space teleoperation through time delay: Review and prognosis," *IEEE Trans. Robot. Autom.*, vol. 9, no. 5, pp. 592–606, May 1993, doi: 10.1109/70.258052.

[3] P. F. Hokayem and M. W. Spong, "Bilateral teleoperation: An historical survey," *Automatica*, vol. 42, no. 12, pp. 2035–2057, Dec. 2006, doi: 10.1016/j.automatica.2006.06.027.

[4] R. C. Goertz, "Mechanical master–slave manipulator," *Nucleonics*, vol. 12, no. 11, pp. 45–46, 1954.

[5] R. C. Goertz, "Electronically controlled manipulator," *Nucleonics*, vol. 12, no. 11, pp. 46–47, 1954.

[6] R. C. Goertz, "Manipulator systems developed at ANL," in *Proc. 12th Conf. Remote Syst. Technol.*, Nov. 1964, pp. 117–136.

[7] B. Hannaford, "A design framework for teleoperators with kinesthetic feedback," *IEEE Trans. Robot. Autom.*, vol. 5, no. 4, pp. 426–434, Aug. 1989, doi: 10.1109/70.88057.

[8] D. A. Lawrence, "Stability and transparency in bilateral teleoperation," *IEEE Trans. Robot. Autom.*, vol. 9, no. 5, pp. 624–637, Oct. 1993, doi: 10.1109/70.258054.

[9] Y. Yokokohji and T. Yoshikawa, "Bilateral control of master–slave manipulators for ideal kinesthetic coupling-formulation and experiment," *IEEE Trans. Robot. Autom.*, vol. 10, no. 5, pp. 605–620, Oct. 1994, doi: 10.1109/70.326566.

[10] Y. Deng, Y. Tang, B. Yang, W. Zheng, S. Liu, and C. Liu, "A review of bilateral teleoperation control strategies with soft environment," in *Proc. 6th IEEE Int. Conf. Adv. Robot. Mechatronics (ICARM)*, Jul. 2021, pp. 459–464, doi: 10.1109/ICARM52023.2021.9536056.

[11] D. Adamski and R. Miller, "Unmanned teleoperator spacecraft (UTS) technology," in *Proc. 6th Annu. Meeting Tech. Display*, Oct. 1969, p. 1067, doi: 10.2514/6.1969-1067.

[12] J. R. Parks and D. A. Bell, "Industrial robots and machine intelligence," *Phys. Bull.*, vol. 21, no. 12, pp. 549–553, Dec. 1970, doi: 10.1088/0031-9112/21/12/030.

[13] T. B. Sheridan and W. L. Verplank, "Human and computer control of undersea teleoperators," Massachusetts Inst. Technol., Cambridge, MA, USA, Tech. Rep., 1978. [Online]. Available: https://apps.dtic.mil/sti/citations/ADA057655

[14] A. Kron and G. Schmidt, "Haptic telepresent control technology applied to disposal of explosive ordnances: Principles and experimental results," in *Proc. IEEE Int. Symp. Ind. Electron. (ISIE)*, Nov. 2005, pp. 1505–1510, doi: 10.1109/ISIE.2005.1529155.

[15] J. S. Im, F. Ozawa, T. Matsushita, N. Matsunaga, and S. Kawaji, "Experimental study on steer-by-wire system with bilateral control," in *Proc. IEEE Int. Conf. Mechatronics*, May 2007, pp. 1–6, doi: 10.1109/ICMECH.2007.4280014.

[16] M. Tavakoli, A. Aziminejad, R. V. Patel, and M. Moallem, "Methods and mechanisms for contact feedback in a robot-assisted minimally invasive environment," *Surgical Endoscopy*, vol. 20, no. 10, pp. 1570–1579, Oct. 2006, doi: 10.1007/s00464-005-0582-y.

[17] H. Tanaka, K. Ohnishi, H. Nishi, T. Kawai, Y. Morikawa, S. Ozawa, and T. Furukawa, "Implementation of bilateral control system based on acceleration control using FPGA for multi-DOF haptic endoscopic surgery robot," *IEEE Trans. Ind. Electron.*, vol. 56, no. 3, pp. 618–627, Mar. 2009, doi: 10.1109/TIE.2008.2005710.

[18] J. Marescaux, J. Leroy, M. Gagner, F. Rubino, D. Mutter, M. Vix, S. E. Butner, and M. K. Smith, "Transatlantic robot-assisted telesurgery," *Nature*, vol. 413, no. 6854, pp. 379–380, Sep. 2001, doi: 10.1038/35096636.

[19] J. Marescaux, J. Leroy, F. Rubino, M. Smith, M. Vix, M. Simone, and D. Mutter, "Transcontinental robot-assisted remote telesurgery: Feasibility and potential applications," *Ann. Surg.*, vol. 235, no. 4, pp. 487–492, Apr. 2002, doi: 10.1097/00000658-200204000-00005.

[20] S. E. Butner and M. Ghodoussi, "Transforming a surgical robot for human telesurgery," *IEEE Trans. Robot. Autom.*, vol. 19, no. 5, pp. 818–824, Oct. 2003, doi: 10.1109/TRA.2003.817214.

[21] P. J. Choi, R. J. Oskouian, and R. S. Tubbs, "Telesurgery: Past, present, and future," *Cureus*, vol. 10, no. 5, pp. e2716–6, May 2018, doi: 10.7759/cureus.2716.

[22] A. Zemmar, A. M. Lozano, and B. J. Nelson, "The rise of robots in surgical environments during COVID-19," *Nature Mach. Intell.*, vol. 2, no. 10, pp. 566–572, Oct. 2020, doi: 10.1038/s42256-020-00238-2.

[23] H. Morohashi, K. Hakamada, T. Kanno, K. Tadano, K. Kawashima, Y. Takahashi, Y. Ebihara, E. Oki, S. Hirano, and M. Mori, "Construction of redundant communications to enhance safety against communication interruptions during robotic remote surgery," *Sci. Rep.*, vol. 13, no. 1, p. 10831, Jul. 2023, doi: 10.1038/s41598-023-37730-9.

[24] A. A. Cardenas, S. Amin, and S. Sastry, "Secure control: Towards survivable cyber-physical systems," in *Proc. 28th Int. Conf. Distrib. Comput. Syst. Workshops*, Jun. 2008, pp. 495–500, doi: 10.1109/icdcs.workshops.2008.40.

[25] M. S. Chong, H. Sandberg, and A. M. H. Teixeira, "A tutorial introduction to security and privacy for cyber-physical systems," in *Proc. 18th Eur. Control Conf. (ECC)*, Jun. 2019, pp. 968–978, doi: 10.23919/ECC.2019.8795652.

[26] W. Knowles, D. Prince, D. Hutchison, J. F. P. Disso, and K. Jones, "A survey of cyber security management in industrial control systems," *Int. J. Crit. Infrastruct. Protection*, vol. 9, pp. 52–80, Jun. 2015, doi: 10.1016/j.ijcip.2015.02.002.

[27] F. Pasqualetti, F. Dörfler, and F. Bullo, "Attack detection and identification in cyber-physical systems," *IEEE Trans. Autom. Control*, vol. 58, no. 11, pp. 2715–2729, Nov. 2013, doi: 10.1109/TAC.2013.2266831.

[28] A. Teixeira, K. C. Sou, H. Sandberg, and K. H. Johansson, "Secure control systems: A quantitative risk management approach," *IEEE Control Syst. Mag.*, vol. 35, no. 1, pp. 24–45, Feb. 2015, doi: 10.1109/MCS.2014.2364709.

[29] A. Teixeira, D. Pérez, H. Sandberg, and K. H. Johansson, "Attack models and scenarios for networked control systems," in *Proc. 1st Int. Conf. High Confidence Networked Syst.*, Apr. 2012, pp. 55–64, doi: 10.1145/2185505.2185515.

[30] K. Kogiso and T. Fujita, "Cyber-security enhancement of networked control systems using homomorphic encryption," in *Proc. 54th IEEE Conf. Decis. Control (CDC)*, Dec. 2015, pp. 6836–6843, doi: 10.1109/CDC.2015.7403296.

[31] M. S. Darup, A. B. Alexandru, D. E. Quevedo, and G. J. Pappas, "Encrypted control for networked systems: An illustrative introduction and current challenges," *IEEE Control Syst. Mag.*, vol. 41, no. 3, pp. 58–78, Jun. 2021, doi: 10.1109/MCS.2021.3062956.

[32] J. H. Cheon, A. Kim, M. Kim, and Y. Song, "Homomorphic encryption for arithmetic of approximate numbers," in *Proc. Int. Conf. Theory Appl. Cryptol. Inf. Secur.*, Dec. 2017, pp. 409–437, doi: 10.1007/978-3-319-70694-8_15.

[33] J. Dyer, M. Dyer, and J. Xu, "Practical homomorphic encryption over the integers for secure computation in the cloud," *Int. J. Inf. Secur.*, vol. 18, no. 5, pp. 549–579, Oct. 2019, doi: 10.1007/s10207-019-00427-0.

[34] K. Kogiso, "Attack detection and prevention for encrypted control systems by application of switching-key management," in *Proc. IEEE Conf. Decis. Control (CDC)*, Dec. 2018, pp. 5032–5037, doi: 10.1109/CDC.2018.8619221.

[35] J. Katz and Y. Lindell, *Introduction to Modern Cryptography*, 3rd ed., Boca Raton, FL, USA: CRC Press, 2020.
[36] K. Teranishi and K. Kogiso, "Control-theoretic approach to malleability cancellation by attacked signal normalization," *IFAC-PapersOnLine*, vol. 52, no. 20, pp. 297–302, 2019, doi: 10.1016/j.ifacol.2019.12.171.
[37] J. Lee, J. Kim, and H. Shim, "Zero-dynamics attack on homomorphically encrypted control system," in *Proc. 20th Int. Conf. Control, Autom. Syst. (ICCAS)*, Oct. 2020, pp. 385–390, doi: 10.23919/ICCAS50221.2020.9268374.
[38] R. Alisic, J. Kim, and H. Sandberg, "Model-free undetectable attacks on linear systems using LWE-based encryption," *IEEE Control Syst. Lett.*, vol. 7, pp. 1249–1254, 2023, doi: 10.1109/LCSYS.2023.3234004.
[39] N. Shono, T. Miyazaki, K. Teranishi, T. Kanno, T. Kawase, K. Kogiso, and K. Kawashima, "Implementation of encrypted control of pneumatic bilateral control system using wave variables," in *Proc. 27th Int. Symp. Artif. Life Robot.*, Jan. 2022, Paper no. OS13-1.
[40] N. Shono, T. Miyazaki, K. Teranishi, K. Kogiso, and K. Kawashima, "A false data injection attack model targeting passivity of encrypted wave variable based bilateral control system," in *Proc. IEEE/SICE Int. Symp. Syst. Integr. (SII)*, Jan. 2023, pp. 1–6, doi: 10.1109/SII55687.2023.10039338.
[41] M. Ota et al., "Field experiment of a telesurgery system using a surgical robot with haptic feedback," *Surg. Today*, vol. 54, no. 4, pp. 375–381, Apr. 2024, doi: 10.1007/s00595-023-02732-7.
[42] X. Deng and D. Tian, "Force feedback bilateral control of multi-DOF cooperative manipulator: Design and realization," in *Proc. IEEE Int. Conf. Mechatronics Autom. (ICMA)*, Aug. 2023, pp. 1056–1061, doi: 10.1109/icma57826.2023.10215857.
[43] W. R. Ferrell, "Delayed force feedback," *Human Factors, J. Human Factors Ergonom. Soc.*, vol. 8, no. 5, pp. 449–455, Oct. 1966, doi: 10.1177/001872086600800509.
[44] J. Vertut, A. Micaelli, P. Marchai, and J. Guittet, "Short transmission delay on a force reflective bilateral manipulator," in *Proc. 4th Symp. Theory Pract. Robot. Manipulation*, 1981, pp. 269–285.
[45] T. Ishii and S. Katsura, "Bilateral control with local force feedback for delay-free teleoperation," in *Proc. 12th IEEE Int. Workshop Adv. Motion Control (AMC)*, Mar. 2012, pp. 1–6, doi: 10.1109/AMC.2012.6197100.
[46] B. Hannaford and J.-H. Ryu, "Time-domain passivity control of haptic interfaces," *IEEE Trans. Robot. Autom.*, vol. 18, no. 1, pp. 1–10, Feb. 1989, doi: 10.1109/70.988969.
[47] J.-H. Ryu, D.-S. Kwon, and B. Hannaford, "Stable teleoperation with time-domain passivity control," *IEEE Trans. Robot. Autom.*, vol. 20, no. 2, pp. 365–373, Apr. 2004, doi: 10.1109/TRA.2004.824689.
[48] J.-H. Ryu, C. Preusche, B. Hannaford, and G. Hirzinger, "Time domain passivity control with reference energy following," *IEEE Trans. Control Syst. Technol.*, vol. 13, no. 5, pp. 737–742, Sep. 2005, doi: 10.1109/TCST.2005.847336.
[49] T. Elgamal, "A public key cryptosystem and a signature scheme based on discrete logarithms," *IEEE Trans. Inf. Theory*, vol. IT-31, no. 4, pp. 469–472, Jul. 1985.

KATSUMASA KOSHA received the B.S. degree from the Department of Mathematical Engineering and Information Physics, The University of Tokyo, Japan, in 2023. He is currently pursuing the master's degree with the Graduate School of Information Science and Technology, The University of Tokyo. His research interests include cyber security of control systems and encrypted control.

TETSURO MIYAZAKI (Member, IEEE) received the Ph.D. degree in engineering from the Department of Mechanical Sciences and Engineering, Tokyo Institute of Technology, in 2014.

From 2014 to 2015, he was a Research Assistant and an Assistant Professor with Yokohama National University, from 2015 to 2017. From 2017 to 2020, he was an Assistant Professor with Tokyo Medical and Dental University. From 2020 to 2021, he was an Assistant Professor and a Lecturer with the Graduate School of Information Science and Technology, The University of Tokyo, since January 2022. His research interests include mechanical engineering, control engineering, power-assistive devices, and medical welfare robotics.

KAORU TERANISHI (Member, IEEE) received the B.S. degree in electronic and mechanical engineering from the National Institute of Technology, Ishikawa College, Ishikawa, Japan, in 2019, and the M.S. and Ph.D. degrees in mechanical and intelligent systems engineering from The University of Electro-Communications, Tokyo, Japan, in 2021 and 2024, respectively.

From October 2019 to September 2020, he was a Visiting Scholar with Georgia Institute of Technology, Atlanta, GA, USA. From April 2021 to March 2024, he was a Research Fellow with Japan Society for the Promotion of Science, Tokyo. He is currently a Research Fellow with The University of Electro-Communications. His research interests include control theory and cryptography for control systems security.

KIMINAO KOGISO (Member, IEEE) received B.E., M.E., and Ph.D. degrees in mechanical engineering from Osaka University, Japan, in 1999, 2001, and 2004, respectively.

In April 2004, he was a Postdoctoral Fellow in the 21st Century COE Program and as an Assistant Professor with the Graduate School of Information Science, Nara Institute of Science and Technology, Nara, Japan, in July 2005. From November 2010 to December 2011, he was a Visiting Scholar with Georgia Institute of Technology, Atlanta, GA, USA. In March 2014, he was promoted to an Associate Professor with the Department of Mechanical and Intelligent Systems Engineering, The University of Electro-Communications, Tokyo, Japan. Since April 2023, he has been a Full Professor with the Department of Mechanical and Intelligent Systems Engineering, The University of Electro-Communications. His research interests include cybersecurity of control systems, constrained control, control of decision-makers, and their applications.

KENJI KAWASHIMA (Member, IEEE) received the Ph.D. degree in engineering from the Department of Control Engineering, Tokyo Institute of Technology, in 1997.

From 1997 to 2000, he was a Research Assistant with Tokyo Metropolitan College of Industrial Technology. Thereafter, he was an Associate Professor with the Precision and Intelligence Laboratory, Tokyo Institute of Technology. From 2013 to 2020, he was a Professor with Tokyo Medical and Dental University. Since April 2020, he has been a Professor with The University of Tokyo, Japan. His research interests include medical robotics, control engineering, fluid measurement, and control.