Preprints of the 21st IFAC World Congress (Virtual)
Berlin, Germany, July 12-17, 2020

Security Enhancement of Sampled-Data Systems: Zero Assignment via Generalized Sampler ⋆

Daehan Kim, Kunhee Ryu, and Juhoon Back*

School of Robotics, Kwangwoon University, Seoul, Republic of Korea,
(e-mail: {2018124101, ryuhhh, backhoon}@kw.ac.kr)

Abstract: Remote control systems have advantages in terms of flexibility and efficiency, but at the same time, they are exposed to cyber-attacks. Zero-dynamics attack is one of the most lethal model-based cyber attacks due to its stealthiness. In this paper, a new zero-dynamics attack neutralizing strategy is proposed, which is based on the generalized sampler. By using generalized sampler, the zeros of the discrete-time system can be placed at arbitrary locations, and if all zeros are placed inside the unit circle, the attack signal itself is neutralized. This strategy still works even if all the information is exposed to hackers. Furthermore, the proposed method is insensitive to shifting the intrinsic zeros comparing to the existing zero-assignment based methods. A design procedure of generalized sampler is provided, and theoretical findings are validated through the numerical simulations.

Keywords: Sampled-data systems, Zero-dynamics attack, Security, Zero assignment

1. INTRODUCTION

Thanks to advanced communication technologies developed during the last few decades, it is now possible to control remote dynamic systems, monitor the status of geographically distributed systems such as power systems and smart factories, and make decisions for multi-agent systems, etc. Although the flexibility in system configuration has been greatly increased, these networked control systems are inherently exposed to cyber attacks, as reported in real incidents such as the attack on Ukrainian power plant (Case, 2016). A number of cyber attacks have been modeled and analyzed (see, e.g., Cárdenas et al. (2011); Teixeira et al. (2015); Mo and Sinopoli (2009); Gupta et al. (2010) and references therein), and countermeasures against those attacks have been developed (Mo and Sinopoli, 2009; Teixeira et al., 2012; Back et al., 2017).

Among many cyber attacks, the zero-dynamics attack (ZDA) is one of most dangerous attacks since it is not possible to detect it. It is a model-based attack, and exploits the zero-dynamics of a system (Khalil, 2002) which describes the internal behavior of the system. Suppose that we have a stable system (or it has been stabilized by a controller) and that the zero-dynamics is known to a hacker. The attack signal of ZDA is generated by a dynamic system which is identical to the zero-dynamics of the given system. By stability, the output of the system converges to zero while the internal state corresponding the zero-dynamics converges to that of ZDA's dynamics. Thus, ZDA is fatal to systems which have unstable zero-dynamics, i.e., non-minimum phase systems (Park et al., 2019).

It is noted that most sampled-data systems are vulnerable to ZDA. In fact, in most sampled-data systems, it is common to use zero-order hold (ZOH) to convert the discrete input signal to continuous input and simple sampler to obtain a discrete signal from the continuous output signal. Unfortunately, in this case, the discrete-time system has unstable zero-dynamics if the original (continuous-time) system has a relative degree greater than two and the sampling time is sufficiently small, regardless of the stability of the original zero-dynamics (Yuz and Goodwin, 2014).

Several countermeasures against ZDA have been proposed. Teixeira et al. (2012) characterized geometric properties of the ZDA and provided solutions to reveal the attack by modifying the system structure. Hoehn and Zhang (2016) introduced a (constant or time-varying) modulation matrix in the input channel so that actual input gain matrix is hidden. Both approaches can reveal the presence of attack, but the information on modification or modulation matrix should be hidden. Dual rate control is proposed by Naghnaeian et al. (2015). The idea is to construct a lifted discrete-time system by collecting sufficiently large number of output measurements during a single sampling interval and it is shown that the lifted system has no unstable zeros. Recently, a neutralization strategy employing generalized hold (GH) (Yuz and Goodwin, 2014) has been proposed by Back et al. (2017). This approach exploited the fact that if the hold function of GH is designed properly, the zeros of discrete-time system can be arbitrarily assigned. In fact, a design procedure is provided to make the zeros reside inside the unit circle so that ZDA is not effective anymore.

⋆ This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT) (2019R1A4A1029003) and supported by Korea Electric Power Corporation(Grant number : R18XA06-20).

In this paper, we present a new neutralization strategy against ZDA. We replace the simple sampler (SS) that takes the output at each sampling time by so-called generalized sampler (GS) (Yuz and Goodwin, 2014) which takes several samples during one sampling interval and generates new output that is defined by a weighted average of the samples. It is shown that the zeros of the discrete-time system, with new output, can be placed anywhere desired, which means that ZDA can be neutralized like the case with GH. A design procedure to choose the gains of GS is also presented. It is emphasized that the proposed approach has several advantages over existing strategies; firstly no information needs to be hidden, secondly, the zeros can be arbitrarily assigned, and finally, it is remarkably insensitive to the shift of intrinsic zeros. Numerical simulations on a two mass system are conducted to validate theoretical findings.

The rest of this paper is organized as follows. In Section 2, we briefly recall the concept of ZDA and several strategies against it. The generalized sampler as a new zero assignment tool is introduced in detail and discuss the novelty in Section 3. In Section 4, we show that the proposed tool can be used as a good countermeasure against ZDA. Section 5 concludes the paper with future research directions.

2. PRELIMINARIES

2.1 Zero-Dynamics Attack on Sampled-Data System

We consider a continuous-time system controlled by a digital controller which is connected through a communication network. At each sampling time, measured information (typically sampled output) is transmitted to the controller and the controller generates a control signal which is delivered to the actuator that applies a control input to the system. Suppose that the network has been compromised so that a malicious attack signal can be added in the control signal. Precisely, the control system under attack is described by

$$\begin{aligned} \dot{x} &= Ax + B(u + a) \\ y &= Cx \end{aligned} \tag{1}$$

where $x \in \mathbb{R}^n$ is the state vector, $u \in \mathbb{R}$ is the control input, $a \in \mathbb{R}$ is the attack signal, and $y \in \mathbb{R}$ is the output. A, B and C are constant matrices with appropriate dimensions. Suppose that the communication between the plant and controller is done at $t = kT_s$, $k = 1, 2, \ldots$, where T_s is the sampling time. In most cases, ZOH and SS are used to interface the system (1) with digital controller. ZOH is used in the input side so that the discrete input signal $u_k := u(kT_s)$ coming from the controller is converted to $u(t) = u_k, kT_s \le t < (k+1)T_s$, and SS converts the continuous-time signal $y(t)$ to a discrete-time signal $y_k := y(kT_s)$ in the output side. It is assumed that the attack signal a_k is injected through the communication network so that the signal transmitted to ZOH becomes $u_k + a_k$.

With ZOH and SS, the system (1) can be rewritten as a discrete-time system given by

$$\begin{aligned} x_{k+1} &= A_{\mathrm{d}} x_k + B_{\mathrm{d}}(u_k + a_k) \\ y_k &= C_{\mathrm{d}} x_k \end{aligned} \tag{2}$$

where $x_k = x(kT_s)$ and

$$A_{\mathrm{d}} = e^{AT_s},\ B_{\mathrm{d}} = \int_0^{T_s} e^{A(T_s-\tau)} B d\tau,\ C_{\mathrm{d}} = C.$$

The attack considered in this paper is constructed using the zero-dynamics of the system (2) which explains the internal behavior of the system (Khalil, 2002). In order to construct it, the system is rewritten in Byrnes-Isidori normal form given by

$$\begin{aligned} \eta_{k+1} &= S_{\mathrm{d}}\eta_k + P_{\mathrm{d}}\bar{C}_{\mathrm{d}}\xi_k \\ \xi_{k+1} &= \bar{A}_{\mathrm{d}}\xi_k + \bar{B}_{\mathrm{d}}(\psi_{\mathrm{d}}^\top \eta_k + \phi_{\mathrm{d}}^\top \xi_k + g_{\mathrm{d}}(u_k + a_k)) \\ y_k &= \bar{C}_{\mathrm{d}}\xi_k. \end{aligned} \tag{3}$$

The dynamics $\eta_{k+1} = S_{\mathrm{d}}\eta_k$ is called the zero-dynamics and eigenvalues of S_{d} correspond to the zeros of the discrete-time system (2). If the system (2) has a relative degree μ, then the dimensions of S_{d}, P_{d}, ψ_{d}, ϕ_{d} and g_{d} are determined so that $\eta_k \in \mathbb{R}^{n-\mu}$, $\xi_k \in \mathbb{R}^{\mu}$, and we have

$$\bar{A}_{\mathrm{d}} = \begin{bmatrix} 0_{\mu-1} & I_{\mu-1} \\ 0 & 0_{\mu-1}^\top \end{bmatrix}, \bar{B}_{\mathrm{d}} = \begin{bmatrix} 0_{\mu-1} \\ 1 \end{bmatrix}, \bar{C}_{\mathrm{d}} = \begin{bmatrix} 1 \\ 0_{\mu-1} \end{bmatrix}^\top.$$

For the system (2), ZDA is constructed as

$$z_{k+1} = S_{\mathrm{d}} z_k, \qquad a_k = -\frac{1}{g_{\mathrm{d}}}\psi_{\mathrm{d}}^\top z_k. \tag{4}$$

It is noted that to construct a ZDA, system parameters should be known.

Suppose the system is stabilized by a static output feedback controller given by $u_k = -L_{\mathrm{d}} y_k$, i.e., the matrix $A_{\mathrm{d}} - B_{\mathrm{d}} L_{\mathrm{d}} C_{\mathrm{d}}$ is Schur. This property results in that, under the attack (4), η_k and ξ_k admit a bound

$$\left\| \begin{bmatrix} \eta_k - z_k \\ \xi_k \end{bmatrix} \right\| \le \lambda_0 \lambda^k \left\| \begin{bmatrix} \eta_0 - z_0 \\ \xi_0 \end{bmatrix} \right\|,\ \lambda_0 > 0, |\lambda| < 1$$

where η_0, ξ_0, and z_0 are initial conditions of corresponding variables, and this relation implies that η_k approaches z_k as k increases. The lethality of ZDA becomes obvious when S_{d} is unstable, i.e., the system (2) is of non-minimum phase. In this case, η_k becomes unbounded whenever z_k is excited by unstable modes of S_{d}, while y_k, that depends solely on ξ_k, converges to zero so that the presence of attack can not be monitored from y_k.

It is emphasized that discrete-time systems are vulnerable to ZDA because of 'sampling zeros' appearing from the sampling procedure. In fact, when the continuous-time system has a relative degree greater than two and the sampling time is sufficiently small, it is inevitable that the discrete-time system has unstable zero-dynamics because at least one of the sampling zeros lies outside the unit circle (Yuz and Goodwin, 2014).

2.2 Existing Countermeasures

In order to enhance security against ZDA, several strategies have been developed. Teixeira et al. (2012) analyzed how the system structure affects the stealthiness property of ZDA and suggested to modify input gain matrix, output matrix, and system matrix to reveal the attack. Hoehn and Zhang (2016) introduced a modulation matrix in the input channel so that actual input gain matrix is hidden from hackers. An optimization based design is proposed and time-varying (periodic) modulation matrix is also considered. Although these approaches can reveal ZDA

attack, they have critical drawback that information on modification or modulation matrix should be hidden.

Instead of modifying the internal structure, Naghnaeian et al. (2015) proposed to use dual rate control. The idea is to obtain a sufficiently large number of measurements during a single sampling interval and consider the collection of the measurements as a new output. They proved that the system with new output has no unstable zeros. It is the main advantage that it is not necessary to hide any information from hackers. However, substantially large amount of information should be transmitted.

Recently, a new strategy overcoming the drawbacks mentioned above has been introduced by Back et al. (2017). The key idea is to employ the generalized hold (GH) (Yuz and Goodwin, 2014) so that all the zeros of discrete-time system are placed inside the unit circle. Let us recall briefly. GH involves a function $h_g(t)$ called hold function. It is defined as a piecewise continuous function with $h_g(t) = 0$ for $t < 0$ or $t \geq T_s$. On the time interval $[kT_s, (k+1)T_s)$, the input $u(t)$ is given by $u(t) = u_k h_g(t - kT_s)$. For the sake of implementation, one can use piecewise constant hold function given by

$$h_g(t) = h_i, \ \frac{(i-1)T_s}{N} \leq t < \frac{iT_s}{N}, \ i = 1, \ldots, N \quad (5)$$

where h_i are constant gains and N is the number of subintervals. It is noted that although it has advantages over other strategies, GH may induce undesirable inter-sample behaviors (Kabamba, 1987) which will be illustrated in the subsequent section.

3. ZERO ASSIGNMENT USING GENERALIZED SAMPLER

In this section, we will introduce a new zero-assignment approach using GS. At first, we explain the generalized sampler and then provide a design procedure. Comparison with a recently reported approach employing GH is also presented through a numerical example.

3.1 Generalized Sampler

By GS, we mean a device which generates a discrete-time signal $\breve{y}_k$ from a continuous-time signal $y(t)$ in a way that given measured signals during the time interval $[0, kT_s]$, say $y(t_{k,1}), \ldots, y(t_{k,N_k})$, $t_{k,1} < \cdots < t_{k,N_k}$, the signal $\breve{y}_k$ is defined by $\breve{y}_k = S_k(y(t_{k,1}), \ldots, y(t_{k,N_k}))$ where S_k is a function (or it can be generalized to include some dynamics) and N_k is the number of samples to compute $\breve{y}_k$. The simplest example is SS ($N_k = 1, t_{k,N_k} = kT_s$, $\breve{y}_k = y(kT_s)$).

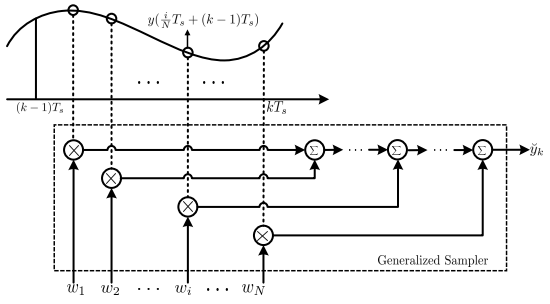


Fig. 1. Concept of generalized sampler.

In this paper, we consider the case where the measurements used to construct $\breve{y}_k$ are taken from the interval $((k-1)T_s, kT_s]$ and S_k is a functions which computes a weighted average of the measurements. Precisely speaking, let N be the number of measurements to be used and $w_1, \ldots, w_N$ be weights. As illustrated in Fig. 1, we take $y(\frac{1}{N}T_s + (k-1)T_s), y(\frac{2}{N}T_s + (k-1)T_s), \ldots, y(kT_s)$ and compute the weighted average of these signals with weights $w_1, \ldots, w_N$, i.e.,

$$\breve{y}_k = \sum_{i=1}^{N} w_i y(\tfrac{i}{N}T_s + (k-1)T_s). \quad (6)$$

To proceed, we would like to find a discrete-time system whose output is $\breve{y}_k$. Noting that

$$\begin{aligned} &x(\tfrac{i}{N}T_s + (k-1)T_s) \\ &= e^{A\frac{i}{N}T_s}x_{k-1} + \int_0^{\frac{i}{N}T_s} e^{A(\frac{i}{N}T_s - \tau)}B\,d\tau\,u_{k-1}, \end{aligned}$$

we obtain

$$\begin{aligned} x_k &= A_{\mathrm{d}}x_{k-1} + B_{\mathrm{d}}u_{k-1} \\ \breve{y}_k &= \breve{C}_{\mathrm{d}}x_{k-1} + \breve{D}_{\mathrm{d}}u_{k-1} \end{aligned} \quad (7)$$

where

$$\begin{aligned} \breve{C}_{\mathrm{d}} &= \sum_{i=1}^{N} w_i C_d e^{A\frac{i}{N}T_s} \\ \breve{D}_{\mathrm{d}} &= \sum_{i=1}^{N} w_i C_d \int_0^{\frac{i}{N}T_s} e^{A(\frac{i}{N}T_s - \tau)}B d\tau. \end{aligned}$$

From (7), we can compute the discrete-time transfer function from u_k to y_k as

$$G_{\mathrm{d}}(z) = z^{-1}(\breve{C}_{\mathrm{d}}(zI_n - A_{\mathrm{d}})^{-1}B_{\mathrm{d}} + \breve{D}_{\mathrm{d}}). \quad (8)$$

Note that $\breve{C}_{\mathrm{d}}$ and $\breve{D}_{\mathrm{d}}$ contain the sampler weights w_i of GS which are design parameters. Thus, it is expected that by choosing w_i appropriately, the numerator of $G_{\mathrm{d}}(z)$ can be assigned as desired. In fact, this is true under mild assumptions as can be seen in the next subsection.

3.2 Zero Assignment: Design

In this subsection, we show that the zeros of the system (7) (or (8)) can be placed at desired locations by adjusting the gains w_i. Let $z_{d,1}, \cdots, z_{d,n} \in \mathbb{C}$ be the desired zeros and k_d is a gain. Define

$$G^*_{\mathrm{d}}(z) = k_d z^{-1}\frac{(z - z_{d,1})\cdots(z - z_{d,n})}{\det(zI_n - A_{\mathrm{d}})}. \quad (9)$$

Lemma 1. Suppose $(A_{\mathrm{d}}, B_{\mathrm{d}})$ of the system (7) is controllable and $N \geq n+1$. Then, there exist $\breve{C}_{\mathrm{d}} \in \mathbb{R}^{1\times n}$ and $\breve{D}_{\mathrm{d}} \in \mathbb{R}$ such that $G_{\mathrm{d}}(z)$ is identical to $G^*_{\mathrm{d}}(z)$. Furthermore, there exist $w_1, \ldots, w_N$, which realize $\breve{C}_{\mathrm{d}}$ and $\breve{D}_{\mathrm{d}}$ if the matrix M defined below has full column rank.

$$M = \begin{bmatrix} C_d e^{A\frac{1}{N}T_s} & C_d \int_0^{\frac{1}{N}T_s} e^{A(\frac{1}{N}T_s - \tau)}B d\tau \\ C_d e^{A\frac{2}{N}T_s} & C_d \int_0^{\frac{2}{N}T_s} e^{A(\frac{2}{N}T_s - \tau)}B d\tau \\ \vdots & \\ C_d e^{AT_s} & C_d \int_0^{T_s} e^{A(T_s - \tau)}B d\tau \end{bmatrix}$$

$\diamond$

Proof. Firstly, we rewrite (9) in the control canonical form given by

$$\begin{aligned} x_k &= \left[\begin{array}{c|c} 0_{n-1} & I_{n-1} \\ \hline -d_0 \ \cdots & -d_{n-1} \end{array}\right] x_{k-1} + \begin{bmatrix} 0_{n-1} \\ 1 \end{bmatrix} u_{k-1} \\ &:= A_{\rm con} x_{k-1} + B_{\rm con} u_{k-1} \\ y_k &= [c_0 \ \cdots \ c_{n-2} \ c_{n-1}]\, x_{k-1} + k_d u_{k-1} \\ &:= C_{\rm con} x_{k-1} + D_{\rm con} u_{k-1} \end{aligned}$$

where the constants $c_0,\ldots,c_{n-1}$, $d_0,\ldots,d_{n-1}$ and k_d are determined from the relations

$$\begin{aligned} &\det(zI_n - A_{\rm d}) = z^n + d_{n-1}z^{n-1} + \cdots + d_0 \\ &k_d \prod_{i=1}^{n}(z - z_{d,i}) = k_d(z^n + (\tfrac{c_{n-1}}{k_d} + d_{n-1})z^{n-1} + \\ &\qquad \cdots + (\tfrac{c_1}{k_d} + d_1)z + (\tfrac{c_0}{k_d} + d_0)) \end{aligned}$$

so that $G^*_{\rm d}(z) = z^{-1}(C_{\rm con}(zI_n - A_{\rm con})^{-1}B_{\rm con} + D_{\rm con})$.

The transfer function $G_{\rm d}(z)$ becomes identical to $G^*_{\rm d}(z)$ if and only if $\breve{C}_{\rm d}A^{k-1}_{\rm d}B_{\rm d} = C_{\rm con}A^{k-1}_{\rm con}B_{\rm con}$, $k = 1,\ldots,N$ (Chen, 1998). Since the pair $(A_{\rm d},B_{\rm d})$ is controllable, this relation is equivalent to that $\breve{C}_{\rm d}\mathcal{C}_{\rm d} = C_{\rm con}\mathcal{C}_{\rm con}$ where $\mathcal{C}_{\rm d}$ is the controllablility matrix of $(A_{\rm d},B_{\rm d})$ and $\mathcal{C}_{\rm con}$ is that of $(A_{\rm con},B_{\rm con})$. Thus, $\breve{C}_{\rm d}$ is given by

$$\breve{C}_{\rm d} = C_{\rm con}\mathcal{C}_{\rm con}\mathcal{C}^{-1}_{\rm d}. \tag{10}$$

It is trivial to see that $\breve{D}_{\rm d} = k_d$. Thus the existence of $\breve{C}_{\rm d}$ and $\breve{D}_{\rm d}$ has been proved.

Regarding the weights w_i, we first rewrite

$$\breve{C}_{\rm d} = \sum_{i=1}^{N} w_i C_d e^{A\frac{i}{N}T_s} = w \begin{bmatrix} C_d e^{A\frac{1}{N}T_s} \\ \vdots \\ C_d e^{AT_s} \end{bmatrix} =: w\, C_{{\rm d},N} \tag{11}$$

where $w = [w_1,\ldots,w_N]$. In addition, we obtain another relation for the weights from $\breve{D}_{\rm d}$, i.e.,

$$k_d = \sum_{i=1}^{N} w_i C_d \int_0^{\frac{i}{N}T_s} e^{A(\frac{i}{N}T_s - \tau)} B d\tau =: w D_{{\rm d},N}. \tag{12}$$

Collecting the conditions (11) and (12) yields

$$w\left[C_{{\rm d},N}, D_{{\rm d},N}\right] = \left[\breve{C}_{\rm d}, \breve{D}_{\rm d}\right] \tag{13}$$

Since $M = \left[C_{{\rm d},N}, D_{{\rm d},N}\right]$, the proof is complete. □

From Lemma 1 we propose a design procedure for GS as follows.

Design Procedure 1

(1) Choose the number of subintervals N.
(2) Choose n desired zeros $z_1, z_2, \ldots, z_n$.
(3) Choose the gain k_d.
(4) Calculate the weights $w_1,\ldots,w_N$ from (13).

Sometimes, it is desirable that $\sum w_i = 1$ to ensure that $\breve{y}_k = y(kT_s)$ when $y(t)$ is constant during $((k-1)T_s, kT_s]$. In this case, solve (13) with $k_d = 1$ and w replaced by w^*, and then set $k_d = 1/(\sum w^*_i)$ and $w = k_d w^*$.

Example 1. We consider a two mass system in Fig. 2 taken from Back et al. (2017). The dynamic equation is given by

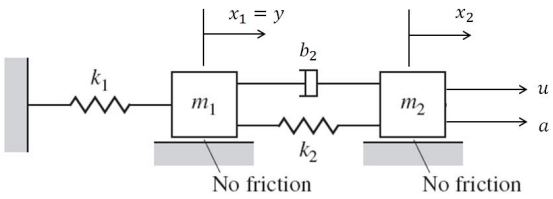


Fig. 2. Two mass system under attack a.

$$\begin{aligned} m_1\ddot{x}_1 &= b_2(\dot{x}_2 - \dot{x}_1) + k_2(x_2 - x_1) - k_1 x_1 \\ m_2\ddot{x}_2 &= u + a - b_2(\dot{x}_2 - \dot{x}_1) - k_2(x_2 - x_1) \\ y &= x_1. \end{aligned}$$

In this example, we assume $a = 0$. With $m_1 = m_2 = 1$kg, $k_1 = k_2 = 1$N/m and $b_2 = 1$Ns/m, the transfer function from u to y becomes $G(s) = (s+1)/(s^4+2s^3+3s^2+s+1)$. Since we have a stable zero at -1, it is a minimum phase system. However, the discrete-time system under ZOH and SS becomes of non-minimum phase for sufficiently small sampling time since the original system has a relative degree 3. In fact, with $T_s = 0.1$s, we have a normal form representation (3) of this system with

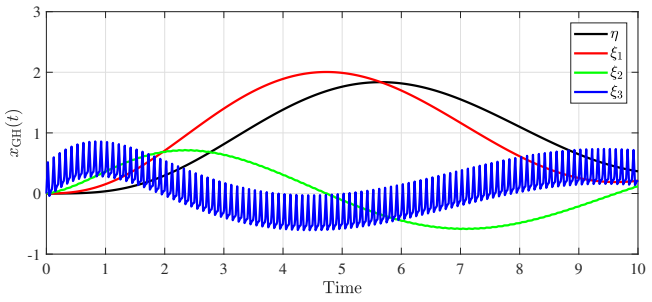
$$\begin{aligned} &S_{\rm d} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0.86 & 2.58 & -2.99 \end{bmatrix}, P_{\rm d} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \psi_{\rm d} = \begin{bmatrix} 5.02 \\ 20.04 \\ -28.28 \end{bmatrix}, \\ &\phi_{\rm d} = 6.78, g_{\rm d} = 1.62 \times 10^{-4}, \mu = 1. \end{aligned} \tag{14}$$

Since the eigenvalues of $S_{\rm d}$ are -3.63, -0.26, 0.90, the discrete-time system is now of non-minimum phase.

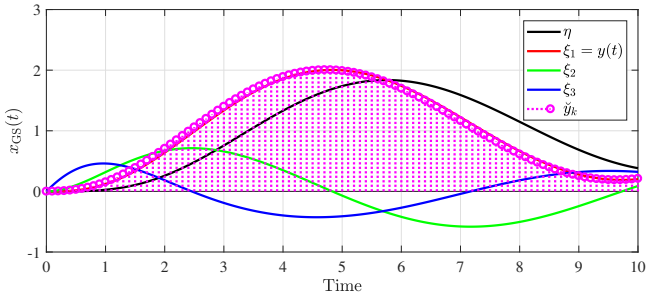
As mentioned earlier, zero assignment can be done by using GH instead of ZOH (Back et al., 2017). We would like to illustrate that the zero assignment method developed in this paper has advantages over GH-based approach. To see this, we follow the procedure proposed by Back et al. (2017) with $N = 4$ to place the zeros at $z_{d,1} = e^{-T_s}$ and $z_{d,2} = z_{d,3} = 0$, which results in $h = [20.89, -21.97, 3.14, 1.94]^\top$. Similarly, following **Design Procedure 1** with $N = 5$, we assign the discrete-time zeros at $z_{d,1} = e^{-T_s}$, $z_{d,2} = z_{d,3} = z_{d,4} = 0$. The weights are computed as $w = [-6.75, 17.49, -1.41, -27.64, 19.30]$.

Fig. 3 shows step responses of both systems. For the case with GH, shown in Fig. 3a, it is observed that the blue line (ξ_3: the third component of $\xi(t)$) fluctuates severely between sampling instants. It is noted that the state variables η and ξ mentioned in the figures are those of continuous-time system represented in normal form (hence, we have three component of ξ and one of η). This phenomenon is typically observed when GH is employed because even if the (discrete-time) input applied to the system is constant, the actual input $u(t)$ generated by GH depends on the pattern which is not constant. On the contrary, for the system using GS, the inter-sample behavior is significantly improved as shown in Fig. 3b. It is noted that if GS is used instead of SS, the output signal $\breve{y}_k$ might be different from the output of the continuous-time system in the transient.

Now, we would like to investigate the sensitivity to the shift of intrinsic zeros. Suppose the desired zero is shifted by 0.002% from the intrinsic zero e^{-T_s} corresponding to the zero at -1 of the continuous-time system. The gains for GH is obtained as $h = [-2.56, 47.76, -66.01, 24.81]^\top$

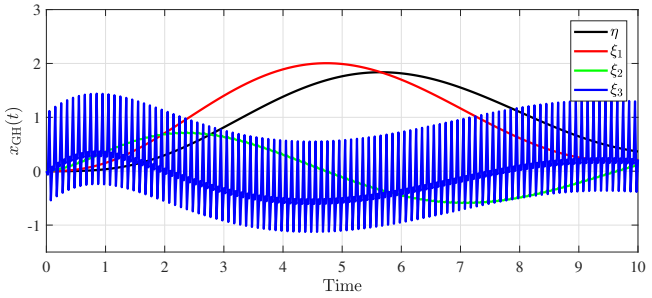


(a) Continuous-time state trajectory under GH.

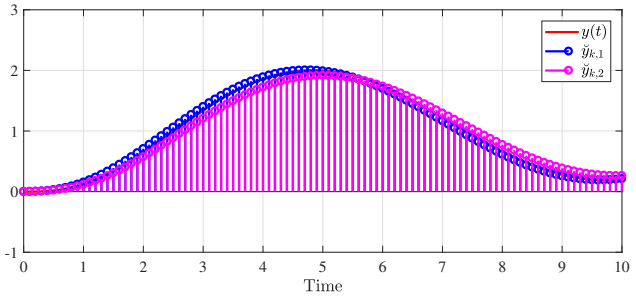


(b) Continuous-time state trajectory and discrete-time output from GS.

Fig. 3. Comparison of system behavior: GH versus GS.



(a) Continuous-time state trajectory under GH, 0.002% shift of intrinsic zero.



(b) Continuous-time output and discrete-time output from GS, $\breve{y}_{k,1}$: 0.002% shift of intrinsic zero, $\breve{y}_{k,2}$: 5% shift of intrinsic zero.

Fig. 4. Comparison of system behavior under intrinsic zero shift: GH versus GS.

which is quite different from the case with $z_{d,1} = e^{-Ts}$. It is remarkable that, on the contrary, the weights for GS is obtained approximately the same as the case with $z_d = [e^{-T_s}, 0, 0, 0]$. If we shift the intrinsic zero by 5%, we have $w = [0.89, -3.59, 5.43, -3.64, 0.91] \times 10^6$, which illustrates that the weights are less sensitive to the shift of intrinsic zeros. The effect of intrinsic zero shift is shown in Fig. 4. As can be seen in Fig. 4a, we have larger fluctuations between sampling instants and it comes from larger gains. On the other hand, in the cases under GS, the system trajectories remain the same because the same inputs are applied to the system. It is noted that even if 5% shift is considered, the output of GS, $\breve{y}_k$ remains near the continuous-time output signal. From these simulation results, we can observe that GS-based zero assignment is more insensitive to the intrinsic zero shift.

4. NEUTRALIZATION OF ZDA: TWO MASS SYSTEM

As described in the previous section, GS allows us to place zeros of the discrete-time system at desired locations. This section explains how to use GS to enhance security against ZDA.

Consider the situation where a hacker launches ZDA on a network-controlled system in Fig. 2 where the physical (continuous-time) system's dynamics is known to the hacker. Suppose that using GS, the discrete-time system's zeros are all assigned inside the unit circle ($z_d = [e^{-T_s},\ 0,\ 0,\ 0]$), and the control input received over the network u_k is zero. In this situation, two cases are considered depending on the hacker's choice of ZDA.

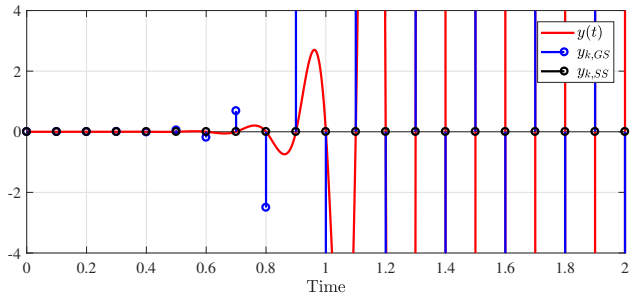
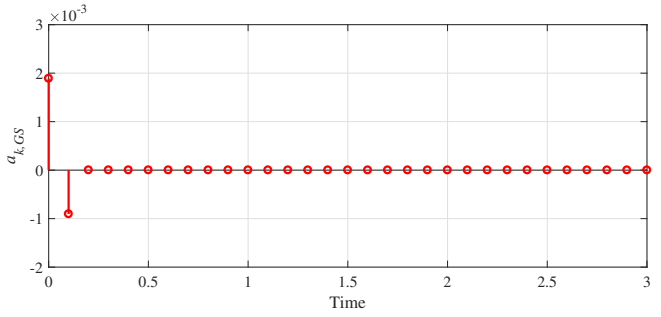


Fig. 5. Continuous-time state and the sampled output for *Case 1*.

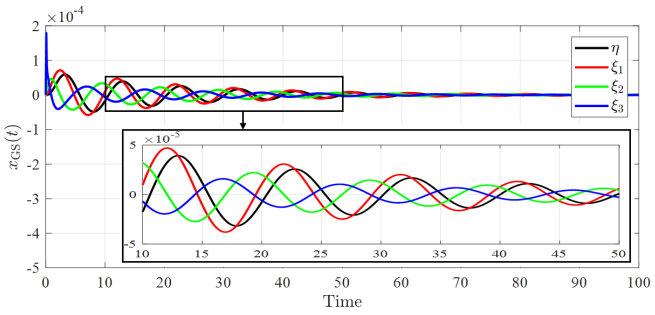
Case 1: Suppose a hacker assumes that the network-controlled system is implemented using ZOH and SS. Then, the hacker injects the attack signal generated by the dynamics (4). Fig. 5 shows the response of the system. In the figure, the red line is the output of the physical system, and the blue circles and black circles are the output signal generated by GS and SS, respectively. The signal from SS remains zero, while the signal from GS is diverging along with the continuous-time state. One can observe that at $t = 0.8$s, we have $y_{k,\mathrm{GS}}$ sufficiently away from the nominal value, and thus conclude that an attack is present. This illustrates that ZDA can be detected by using GS.

Case 2: Suppose a hacker has full information about the discrete-time system, including the presence of GS and the exact location of assigned zeros. In this case, the hacker can take a more sophisticated ZDA with the full information on the system, but the situation is still unfavorable to the hacker. In fact, the hacker will build a ZDA given by

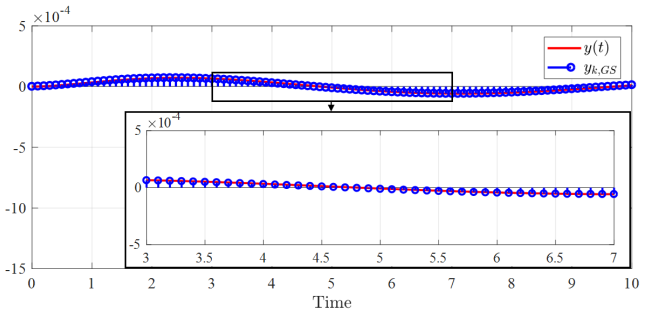
$$\begin{aligned} z_{k+1,\mathrm{GS}} &= S_{\mathrm{d,GS}} z_{k,\mathrm{GS}} \\ a_{k,\mathrm{GS}} &= -\frac{1}{g_{\mathrm{d,GS}}} \psi_{\mathrm{d,GS}}^{\top} z_{k,\mathrm{GS}}. \end{aligned} \tag{15}$$



(a) Attack signal generated by (15).



(b) Dynamic response to ZDA under *Case 2*.



(c) Discrete-time output to ZDA under *Case 2*.

Fig. 6. Attack and response under *Case 2*.

This attack signal, however, has no significant effect on the system since $S_{\mathrm{d,GS}}$ is Schur stable so that $a_{k,\mathrm{GS}}$ converges to zero exponentially, as shown in Fig. 6a. When the attack $a_{k,\mathrm{GS}}$ is injected into the system, as expected, the effect of ZDA on the internal state diminishes to zero as time goes by, as shown in Fig. 6b. The output signal also converges to zero as shown in Fig. 6c,

From two cases considered above, we can conclude that if GS is used and properly designed, ZDA is neutralized in the sense that if the hacker does not have sufficient information on the system, the ZDA is detectable and if the hacker has full information, then the attack will not be effective anymore.

5. CONCLUSION

In this paper, a new countermeasure against the zero dynamics attack has been proposed. It employs the generalized sampler, which takes a weighted average of inter-samples, instead of simple sampler that is frequently used in practice. Although this approach shares the same idea of zero assignment with the generalized hold based approach, the proposed strategy seems to be more effective since the unfavorable inter-sample behavior can be avoided. Compared to other approaches, the proposed idea does not need to hide information on system as well as the generalized sampler, which is an additional benefit. As future research topics, we plan to investigate the robust zero assignment problem using generalized sampler and apply the theory into real systems.

REFERENCES

Back, J., Kim, J., Lee, C., Park, G., and Shim, H. (2017). Enhancement of security against zero dynamics attack via generalized hold. In *2017 IEEE 56th Annual Conference on Decision and Control (CDC)*, 1350–1355. IEEE.

Cárdenas, A.A., Amin, S., Lin, Z.S., Huang, Y.L., Huang, C.Y., and Sastry, S. (2011). Attacks against process control systems: Risk assessment, detection, and response. In *Proceedings of the 6th ACM Symposium on Information, Computer and Communications Security*, ASIACCS '11, 355–366. ACM, New York, NY, USA.

Case, D.U. (2016). Analysis of the cyber attack on the Ukrainian power grid. *Electricity Information Sharing and Analysis Center (E-ISAC)*.

Chen, C.T. (1998). *Linear system theory and design*. Oxford University Press, Inc.

Gupta, A., Langbort, C., and Başar, T. (2010). Optimal control in the presence of an intelligent jammer with limited actions. In *49th IEEE Conference on Decision and Control (CDC)*, 1096–1101. doi: 10.1109/CDC.2010.5717544.

Hoehn, A. and Zhang, P. (2016). Detection of covert attacks and zero dynamics attacks in cyber-physical systems. In *2016 American Control Conference (ACC)*, 302–307. IEEE.

Kabamba, P. (1987). Control of linear systems using generalized sampled-data hold functions. *IEEE Transactions on Automatic Control*, 32(9), 772–783.

Khalil, H.K. (2002). *Nonlinear Systems*, Third Ed. Prentice-Hall, Upper Saddle River, NJ.

Mo, Y. and Sinopoli, B. (2009). Secure control against replay attacks. In *2009 47th annual Allerton conference on communication, control, and computing (Allerton)*, 911–918. IEEE.

Naghnaeian, M., Hirzallah, N., and Voulgaris, P.G. (2015). Dual rate control for security in cyber-physical systems. In *2015 54th IEEE Conference on Decision and Control (CDC)*, 1415–1420. doi:10.1109/CDC.2015.7402409.

Park, G., Lee, C., Shim, H., Eun, Y., and Johansson, K.H. (2019). Stealthy adversaries against uncertain cyber-physical systems: Threat of robust zero-dynamics attack. *IEEE Transactions on Automatic Control*, 64(12), 4907–4919.

Teixeira, A., Shames, I., Sandberg, H., and Johansson, K.H. (2012). Revealing stealthy attacks in control systems. In *2012 50th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, 1806–1813. IEEE.

Teixeira, A., Shames, I., Sandberg, H., and Johansson, K.H. (2015). A secure control framework for resource-limited adversaries. *Automatica*, 51, 135 – 148.

Yuz, J.I. and Goodwin, G.C. (2014). *Sampled-data models for linear and nonlinear systems*. Springer.