Available online at www.sciencedirect.com

ScienceDirect

IFAC PapersOnLine 51-24 (2018) 969–974

Frequency-based detection of replay attacks: application to a multiple tank system [★]

Helem Sánchez [*,**] Damiano Rotondo [*,***] Teresa Escobet [*,****]
Vicenç Puig [*,**,***] Joseba Quevedo [*,**]

[*] *Research Center for Supervision, Safety and Automatic Control (CS2AC) of the Universitat Politècnica de Catalunya (UPC)*
[**] *Automatic Control Department, UPC-ESAII, Rambla de Sant Nebridi, 11, 08222 Terrassa, Spain*
[***] *Institut de Robotica i Informatica Industrial (IRI), UPC-CSIC Carrer de Llorens i Artigas, 4-6, 08028 Barcelona, Spain*
[****] *Department of Mining, Industrial and ICT Engineering, UPC, Av. de les Bases de Manresa, 61-73, 08242 Manresa, Spain*

Abstract: This paper presents a frequency-based method for detecting replay attacks and its application to a multiple tank system. This method introduces a sinusoidal signal with a time-varying frequency (authentication signature) into the closed-loop system and checks whether the output signal is compatible with the signature or not. The effectiveness of the method is illustrated through simulation scenarios using a 9-tank system under different situations.

Keywords: Replay attacks, cyber security, cyber-physical systems, water supply systems.

1. INTRODUCTION

Current societies depend on complex engineering systems with interconnected components working together, which are known as *critical infrastructure systems* (CISs). Among the most important CISs, there are *water supply systems* (WSSs), which are increasing in complexity and dimensions in order to meet the demands of both industry and normal life in growing cities. Due to the interaction between physical elements in the real world and computing elements in the cyber world, WSSs are considered nowadays cyber-physical systems (CPSs) (Park et al., 2012, Kim and Kumar, 2012, Shi et al., 2011), which consist of interconnected subsystems that interact through control, communication, and computation (Wei and Li, 2015).

WSSs are vulnerable to the potential threats brought by natural hazards and terrorism, which may cause temporary disruptions that affect other infrastructures, e.g. the ones devoted to the generation of electrical power (Haimes et al., 1998). In particular, although the application of CPSs to modern WSSs grants superior reliability, autonomy, and efficiency, it creates risks for cyber-physical attacks (Rasekh et al., 2016), which can violate the consumers' privacy, cause intentional damage to the physical water assets (pumps, valves, tanks), decrease the water supply, or impact the water quality (Taormina et al., 2017). A remarkable attack to a water facility happened in 2000 at Maroochy Water Services (Queensland, Australia), affecting the SCADA of a sewage system, which caused the release of almost one million liters of wastewater into waterways and parks (Slay and Miller, 2007). Other relevant incidents are the Pennsylvania Water Company hack in 2006, as well as the Florida's Key Largo Wastewater Treatment District hack, and the computer malfunction blamed for major sewage spill into the Tijuana River in 2012. These events have motivated recent research on cyber security in water systems, see for example Amin et al. (2013b,a), Laszka et al. (2017), Taormina et al. (2017), Ahmed et al. (2017).

Among the most critical cyber-physical attack, there are the *replay attacks*. When an attack of this type is carried out, at first the attacker records the measurements coming from the sensors. Then, in a subsequent phase of the attack, the attacker replaces the real data with the recorded one, causing deterioration of the control system's performance and potentially allowing other types of attacks without being discovered. In the last few years, different approaches have been proposed to detect these attacks, e.g. statistical detection (Mo and Sinopoli, 2009), receding-horizon control (Zhu and Martínez, 2014), data-driven methods (Ma et al., 2017), quantized signals (Kashima and Inoue, 2015) and spectral estimation (Tang et al., 2015).

The main contribution of this paper is to present a method to detect replay attacks using a frequency-based signature and to demonstrate its application to a multiple (nine) tank system, which is described by a complex, highly interconnected and nonlinear model. This method introduces a sinusoidal signal with a time-varying frequency (authentication signal) into the closed-loop system, and checks whether the time profile of the frequency components in the output signals are compatible with the authentication signal or not, by comparing the energies of appropriate signals.

[★] This work has been partially funded by the Spanish State Research Agency (AEI) and the European Regional Development Fund (ERFD) through the projects DEOCS (ref. MINECO DPI2016-76493) and SCAV (ref. MINECO DPI2017-88403-R), by AGAUR of the Generalitat de Catalunya through the Advanced Control Systems (SAC) group grant (2017 SGR 482) and by the Agència de Gestió d'Ajuts Universitaris i de Recerca. This work has been also supported by the Spanish State Research Agency through the Maria de Maeztu Seal of Excellence to IRI (MDM-2016-0656) and the grant Juan de la Cierva - Formacion (FJCI-2016-29019).

10.1016/j.ifacol.2018.09.692

The interest of considering a multiple tank system comes from the fact that this type of system may serve as a first approximation for modeling a portion of a WSS (Georgescu et al., 2010, Ormsbee and Lansey, 1994) and, more generally, as an example of decentralized control system (Johansson, 2000). It is worth mentioning that multiple tank systems are common testbeds for diagnostic monitoring techniques in fault tolerant control (Buciakowski et al., 2017) and cyber security (Gawand et al., 2015, 2017).

The remaining of the paper is structured as follows. Section 2 describes the frequency-based method for detection of replay attacks. Section 3 presents the application of the proposed method to a multiple tank system. The simulation results are presented in Section 4. Finally, the main conclusions are drawn in Section 5.

2. REPLAY ATTACK DETECTION METHOD

2.1 *Attack definition and overview of the detection method*

The replay attack is a type of cyber attack that affects the output of a system, denoted in the following as $y(t)$, and is carried out in two stages:

(1) the attacker gathers the data without disturbing the system, starting from time t_0 until $t_0 + w$, where w is the size of the attack window;
(2) at time t_1, the attacker begins to replay the collected data, such that the real data in the intervals $[t_1 + (N_f - 1)w, t_1 + N_f w], N_f \in \mathbb{N}, N_f \geq 1$, is replaced with the data recorded in stage 1.

Since control systems are not resilient to replay attacks, there is a need to develop methods to detect them. The method presented hereafter introduces a sinusoidal signal with a time-varying frequency (*signature*) into the system and detects if the measured output is compatible with the introduced signature or not. Using a dynamic decoupling technique based on *vector fitting* (VF) (Gustavsen and Semlyen, 1999), it is assured that a signature introduced on a specific input channel will affect only an output. By comparing the energies of band-pass filtered signals, an estimation of the frequency-varying profile in the signature is obtained, which is used by the detector to determine if a replay attack is being carried or not.

2.2 *Signal generation*

Let us consider a linear system described by the following equations

$$\dot{x}(t) = Ax(t) + Bu(t) \tag{1}$$

$$y(t) = Cx(t) \tag{2}$$

which, together with a linear state-feedback law of the type $u(t) = -Kx(t)$, leads to a closed-loop system

$$\dot{x}(t) = (A - BK)x(t) \tag{3}$$

The signature $\varsigma(t)$ is introduced in the input $u(t)$ as an additional signal, such that $u(t) = -Kx(t) + \varsigma(t)$ and

$$x(t) = (A - BK)x(t) + B\varsigma(t) \tag{4}$$

As stated previously, it is desirable to establish a bijection between the available inputs and the available outputs, such that the effect of an element of $\varsigma(t)$ will be observed only on the associated output. Since the closed-loop transfer matrix from $\varsigma(t)$ to $y(t)$, i.e. $G(s) = C(sI - A + BK)^{-1} B$ is coupled, a decoupler $F(s)$ must be introduced such that the series interconnection $G_d(s) = G(s)F(s)$ is *dynamically decoupled* (approximately diagonal) at the frequencies ω_i, $i = 1, \ldots, N$. By requiring that $F(\imath\omega_i) = G(\imath\omega_i)^{-1}$, a set of N constraints that the decoupler should satisfy is obtained. Then, $F(s)$ can be obtained by applying VF (Gustavsen and Semlyen, 1999), a robust numerical method for rational approximation in the frequency domain using poles and residues.

Each element of the input to the decoupler $F(s)$, denoted in the following as $\tilde{\xi}(t)$, is chosen as a frequency-varying sinusoidal signal

$$\tilde{\xi}_l(t) = \tilde{\alpha}_l \cos(\omega_{\sigma_l(t)} t) \quad l = 1, \ldots, n_u \tag{5}$$

where $\tilde{\alpha}_l$ denotes the magnitude, while $\sigma_l(t)$ denotes a piecewise constant signal, which takes integer values between 1 and N, such that at each instant of time $\omega_{\sigma_l(t)}$ equals one of the frequencies ω_i, $i = 1, \ldots, N$, for which $F(s)$ achieves the decoupling. It is assumed that $\sigma_l(t)$ takes a random value between 1 and N at equally-spaced time instants $t_s^{(j)}$, $j \in \mathbb{N}_0$, with $t_s^{(0)} = 0$ and $t_s^{(j+1)} - t_s^{(j)} = T_s$, where T_s is the switching period. The piecewise constant signal $\sigma_l(t)$ is completely known by the detector, whereas the attacker does not have access to this information.

2.3 *Detector logic*

In order to analyze the content of $y(t)$ at the frequencies ω_i, $i = 1, \ldots, N$, used to generate the signature signal $\tilde{\xi}(t)$, the output signal $y(t)$ is introduced into a bank of band-pass filters $H_i(s)$ (Zumbahlen, 2008)

$$H_i(s) = diag\left\{ \frac{\frac{\omega_i}{Q_i} s}{s^2 + \frac{\omega_i}{Q_i} s + \omega_i^2} \right\} \tag{6}$$

where ω_i acts as a peak frequency and Q_i is the selectivity of the filter. In general, to a higher value of Q_i corresponds a narrower frequency response $\|H_i(s)\|$ around ω_i, even though higher values of Q_i will also lead to a slower dynamic response.

Let us denote the output of the band-pass filter with peak frequency ω_i and input $y_l(t)$ as $z_{il}(t)$. Then, the replay attack detection algorithm compares the known signal $\sigma_l(t)$ with $\hat{\sigma}_l(t)$, which is an estimation of $\sigma_l(t)$ based on the signals $z_{il}(t)$, $i = 1, \ldots, N$. In particular, if $\hat{\sigma}_l(t) = \sigma_l(t)$, then the algorithm will state that no replay attack is affecting the output $y_l(t)$, while if $\hat{\sigma}_l(t) \neq \sigma_l(t)$, the algorithm will warn about $y_l(t)$ being corrupted by a replay attack.

The way of calculating the signal $\hat{\sigma}_l(t)$ will affect the effectiveness of the algorithm. A possibility to do so is to compare the energies of different $z_{il}(t)$ over the largest period associated with the frequencies ω_i, $i = 1, \ldots, N$, i.e.

$$T_\omega = \max_{i=1,\ldots,N} \frac{2\pi}{\omega_i} \tag{7}$$

and choose $\hat{\sigma}_l(t)$ as the index corresponding to the signal with the biggest energy. However, when a change in the frequency of the signal $\omega_{\sigma_l(t)}$ in (5) occurs, the system will exhibit a transient behavior with respect to the signal $\tilde{\xi}(t)$. For this reason, a proper choice for obtaining the signal $\hat{\sigma}_l(t)$ is to take into account the time needed for the transient to become neglectable, denoted as t_{trans}, and calculate $\hat{\sigma}_l(t)$ as (8) (see top of the next page), where $t_s^* = \lfloor t/T_s \rfloor T_s$ denotes the last switching time.

It is worth noting that, since the band-pass filters $H_i(s)$ determine the frequency content of the output signals, a reasonable

$$\hat{\sigma}_l(t) = \begin{cases} \sigma_l(t) & if\ \sigma_l(t) \neq \sigma_l(t-T_s) \wedge t \in [t_s^*, t_s^* + t_{trans} + T_\omega] \\ \arg \max\limits_{i=1,\dots,N} \int\limits_{t-T_\omega}^{t} |z_{il}(\tau)|^2 d\tau & \text{otherwise} \end{cases} \tag{8}$$

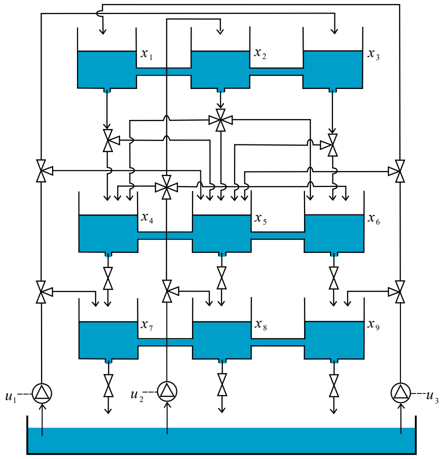

Fig. 1. Schematic diagram of the nine tank process.

estimation of t_{trans} is given by the biggest among the settling times of $H_i(s)$, $i = 1, \dots, N$.

3. APPLICATION TO A MULTIPLE TANK SYSTEM

In this section, the application of the proposed method to a multiple tank system made up by nine tanks, interconnected as shown in Fig. 1, is presented.

3.1 Nonlinear model

The system in Fig. 1 is described by a ninth order state space model, where the states variables x_i, $i = 1, \dots, 9$, are the liquid levels of each tank, while u_1, u_2, u_3 represent the inputs (voltages applied to the pumps). Under the assumption that the water flow at the inlet of each valve is equally distributed among its outlets, the nonlinear state space model is obtained by performing a mass rate balance, as follows

$$\dot{x}_1 = \frac{1}{3}u_3 - \phi_1 - \phi_{12} \tag{9}$$

$$\dot{x}_2 = \frac{1}{4}u_2 - \phi_2 + \phi_{12} - \phi_{23} \tag{10}$$

$$\dot{x}_3 = \frac{1}{3}u_1 - \phi_3 + \phi_{23} \tag{11}$$

$$\dot{x}_4 = \frac{1}{4}u_2 + \frac{1}{2}\phi_1 + \frac{1}{3}\phi_2 - \phi_4 - \phi_{45} \tag{12}$$

$$\dot{x}_5 = \frac{1}{3}u_1 + \frac{1}{3}u_3 + \frac{1}{2}\phi_1 + \frac{1}{3}\phi_2 + \frac{1}{2}\phi_3 - \phi_5 + \phi_{45} - \phi_{56} \tag{13}$$

$$\dot{x}_6 = \frac{1}{4}u_2 + \frac{1}{3}\phi_2 + \frac{1}{2}\phi_3 - \phi_6 + \phi_{56} \tag{14}$$

$$\dot{x}_7 = \frac{1}{3}u_1 + \phi_4 - \phi_7 - \phi_{78} \tag{15}$$

$$\dot{x}_8 = \frac{1}{4}u_2 + \phi_5 - \phi_8 + \phi_{78} - \phi_{89} \tag{16}$$

$$\dot{x}_9 = \frac{1}{3}u_3 + \phi_6 - \phi_9 + \phi_{89} \tag{17}$$

where

$$\phi_i = \alpha_i \sqrt{2gx_i} \tag{18}$$

$$\phi_{ij} = \alpha_{ij} sgn(x_i - x_j)\sqrt{2g|x_i - x_j|} \tag{19}$$

with α_i, α_{ij} denoting the orifice coefficients of each tank or interconnection of tanks, and g denoting the gravity acceleration.

3.2 Linearization around an equilibrium point

By applying a constant input u_e, a steady-state equilibrium point is reached, denoted in the following as x_e. A linearized model can be found by considering deviations Δx of the state variables from the equilibrium point x_e, and performing a first order expansion of the Taylor series. Hence, by taking into account that

$$\frac{\partial \phi_i}{\partial x_i} = \frac{\alpha_i \sqrt{g}}{\sqrt{2x_i}} = \varphi_i(x_i) \tag{20}$$

$$\frac{\partial \phi_{ij}}{\partial x_i} = \frac{\alpha_{ij}\sqrt{g}}{\sqrt{2\,|x_i - x_j|}} = \varphi_{ij}(x_i, x_j) \tag{21}$$

$$\frac{\partial \phi_{ij}}{\partial x_j} = -\frac{\alpha_{ij}\sqrt{g}}{\sqrt{2\,|x_i - x_j|}} = -\varphi_{ij}(x_i, x_j) \tag{22}$$

the following linearized model is obtained

$$\begin{aligned} \dot{\Delta x}_1 = &\frac{1}{3}\Delta u_3 - [\varphi_1(x_{e1}) + \varphi_{12}(x_{e1}, x_{e2})]\Delta x_1 \\ &+ \varphi_{12}(x_{e1}, x_{e2})\Delta x_2 \end{aligned} \tag{23}$$

$$\begin{aligned} \dot{\Delta x}_2 = &\frac{1}{4}\Delta u_2 + \varphi_{12}(x_{e1}, x_{e2})\Delta x_1 + \varphi_{23}(x_{e2}, x_{e3})\Delta x_3 \\ &- [\varphi_2(x_{e2}) + \varphi_{12}(x_{e1}, x_{e2}) + \varphi_{23}(x_{e2}, x_{e3})]\Delta x_2 \end{aligned} \tag{24}$$

$$\begin{aligned} \dot{\Delta x}_3 = &\frac{1}{3}\Delta u_1 + \varphi_{23}(x_{e2}, x_{e3})\Delta x_2 \\ &- [\varphi_3(x_{e3}) + \varphi_{23}(x_{e2}, x_{e3})]\Delta x_3 \end{aligned} \tag{25}$$

$$\begin{aligned} \dot{\Delta x}_4 = &\frac{1}{4}\Delta u_2 + \frac{1}{2}\varphi_1(x_{e1})\Delta x_1 + \frac{1}{3}\varphi_2(x_{e2})\Delta x_2 \\ &- [\varphi_4(x_{e4}) + \varphi_{45}(x_{e4}, x_{e5})]\Delta x_4 + \varphi_{45}(x_{e4}, x_{e5})\Delta x_5 \end{aligned} \tag{26}$$

$$\begin{aligned} \dot{\Delta x}_5 = &\frac{1}{3}\Delta u_1 + \frac{1}{3}\Delta u_3 + \frac{1}{2}\varphi_1(x_{e1})\Delta x_1 + \frac{1}{3}\varphi_2(x_{e2})\Delta x_2 \\ &+ \frac{1}{2}\varphi_3(x_{e3})\Delta x_3 + \varphi_{45}(x_{e4}, x_{e5})\Delta x_4 + \varphi_{56}(x_{e5}, x_{e6})\Delta x_6 \\ &- [\varphi_5(x_{e5}) + \varphi_{45}(x_{e4}, x_{e5}) + \varphi_{56}(x_{e5}, x_{e6})]\Delta x_5 \end{aligned} \tag{27}$$

$$\begin{aligned} \dot{\Delta x}_6 = &\frac{1}{4}\Delta u_2 + \frac{1}{3}\varphi_2(x_{e2})\Delta x_2 + \frac{1}{2}\varphi_3(x_{e3})\Delta x_3 \\ &+ \varphi_{56}(x_{e5}, x_{e6})\Delta x_5 - [\varphi_6(x_{e6}) + \varphi_{56}(x_{e5}, x_{e6})]\Delta x_6 \end{aligned} \tag{28}$$

$$\begin{aligned} \dot{\Delta x}_7 = &\frac{1}{3}\Delta u_1 + \varphi_4(x_{e4})\Delta x_4 - [\varphi_7(x_{e7}) + \varphi_{78}(x_{e7}, x_{e8})]\Delta x_7 \\ &+ \varphi_{78}(x_{e7}, x_{e8})\Delta x_8 \end{aligned} \tag{29}$$

$$\begin{aligned} \dot{\Delta x}_8 = &\frac{1}{4}\Delta u_2 + \varphi_5(x_{e5})\Delta x_5 + \varphi_{78}(x_{e7}, x_{e8})\Delta x_7 \\ &- [\varphi_8(x_{e8}) + \varphi_{78}(x_{e7}, x_{e8}) + \varphi_{89}(x_{e8}, x_{e9})]\Delta x_8 \\ &+ \varphi_{89}(x_{e8}, x_{e9})\Delta x_9 \end{aligned} \tag{30}$$

$$\begin{aligned} \dot{\Delta x}_9 = &\frac{1}{3}\Delta u_3 + \varphi_6(x_{e6})\Delta x_6 - [\varphi_9(x_{e9}) + \varphi_{89}(x_{e8}, x_{e9})]\Delta x_9 \\ &+ \varphi_{89}(x_{e8}, x_{e9})\Delta x_8 \end{aligned} \tag{31}$$

In particular, by considering the parameters $\alpha_1 = \alpha_2 = \alpha_3 = 0.20$, $\alpha_4 = \alpha_5 = \alpha_6 = 0.35$ and $\alpha_7 = \alpha_8 = \alpha_9 = 0.50$, and the input $u_e = [2, 3.5, 3]^T$, which corresponds to the equilibrium point $x_e = [1.08, 0.94, 0.74, 1.40, 2.32, 1.30, 1.42, 1.84, 1.67]^T$, equations (23)-(31) can be expressed in a form akin to (1)

$$\dot{\Delta x}(t) = A\Delta x(t) + B\Delta u(t) \tag{32}$$

with

$$A = \begin{bmatrix} -0.73 & 0.30 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0.30 & -1.01 & 0.25 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.25 & -0.76 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0.21 & 0.15 & 0 & -0.77 & 0.12 & 0 & 0 & 0 & 0 \\ 0.21 & 0.15 & 0.26 & 0.12 & -0.73 & 0.11 & 0 & 0 & 0 \\ 0 & 0.15 & 0.26 & 0 & 0.11 & -0.79 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.66 & 0 & 0 & -1.10 & 0.17 & 0 \\ 0 & 0 & 0 & 0 & 0.51 & 0 & 0.17 & -1.26 & 0.27 \\ 0 & 0 & 0 & 0 & 0 & 0.68 & 0 & 0.27 & -1.13 \end{bmatrix}$$

$$B = \begin{bmatrix} 0 & 0 & 1/3 & 0 & 1/3 & 0 & 1/3 & 0 & 0 \\ 0 & 1/4 & 0 & 1/4 & 0 & 1/4 & 0 & 1/4 & 0 \\ 1/3 & 0 & 0 & 0 & 1/3 & 0 & 0 & 0 & 1/3 \end{bmatrix}^T$$

which, together with a linear error-feedback law of the type $\Delta u(t) = -K\Delta x(t)$, with controller gain designed by pole assignment as

$$K = \begin{bmatrix} -6.48 & -3.62 & 0.79 & 0.01 & 6.47 & 4.67 & -1.62 & -0.55 & 0.14 \\ -4.56 & -3.96 & -3.79 & 2.23 & 0.88 & 5.66 & 4.34 & 4.90 & 4.56 \\ -0.27 & -4.94 & -8.39 & 12.24 & 9.14 & -4.16 & -0.90 & -2.28 & -2.78 \end{bmatrix}$$

leads to a closed-loop system

$$\dot{\Delta x}(t) = (A - BK)\Delta x(t) \tag{33}$$

In the following, we will assume that the tank levels x_7, x_8, x_9 are monitored by a supervision station, which can be hacked through a replay attack. For this reason, the equation (2) is characterized by the output matrix

$$C = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

or, equivalently, by defining $\Delta y(t) = y(t) - y_e = y(t) - Cx_e$

$$\Delta y(t) = C\Delta x(t) \tag{34}$$

3.3 Application of the method

Let us consider a replay attack detector, as described in the previous section, with $N = 2$, and $\omega_2 = 2\omega_1$. In order to choose appropriately the parameters $\tilde{\alpha}_1$, $\tilde{\alpha}_2$, $\tilde{\alpha}_3$ and ω_1, let us note that, by design of the decoupler, the following holds

$$\begin{bmatrix} \bar{\alpha}_1 \\ \bar{\alpha}_2 \\ \bar{\alpha}_3 \end{bmatrix} = \left[\max_{i=1,2} \left| [C(\imath\omega_i - A + BK]^{-1} \right| \right] \begin{bmatrix} \tilde{\alpha}_1 \\ \tilde{\alpha}_2 \\ \tilde{\alpha}_3 \end{bmatrix} \tag{35}$$

where $\bar{\alpha}_1$, $\bar{\alpha}_2$, $\bar{\alpha}_3$ denote the maximum magnitudes of $\varsigma_1(t)$, $\varsigma_2(t)$, $\varsigma_3(t)$ and max is understood as an element-wise maximum. By requiring that $\bar{\alpha}_l \leq \kappa_l u_{e,l}$, $l = 1, \ldots, n_u$, where $\kappa_l \ll 1$ and $u_{e,l}$ is the l-th element of u_e, in order to make the effect of the signature signal on $u(t)$ much smaller than u_e, such that the attacker does not realize about its presence, then a set of feasible frequencies can be calculated, and ω_1 can be chosen as the maximum value among these frequencies. Note that the higher the values of κ_l, $l = 1, \ldots, n_u$, are chosen, the smaller becomes the set of feasible frequencies among which ω_1 can be chosen, which provides a rule of thumb for choosing the magnitude of the elements of the signature signal $\varsigma(t)$ (it is desirable that these elements are chosen as big as possible, in order for the effect of $\varsigma(t)$ on $\Delta y(t)$ to overcome the effect of exogenous disturbances and measurement noise).

According to this reasoning, by using $\kappa_1 = \kappa_2 = \kappa_3 = 1/10$, the following values have been found: $\tilde{\alpha}_1 = 0.012$, $\tilde{\alpha}_2 = 0.016$, $\tilde{\alpha}_3 = 0.021$, $\omega_1 = 1.6\,rad/s$, which corresponds to $T_\omega = 4s$. Then, by requiring an attenuation of $-20dB$ at frequency ω_2 for the first band-pass filter, and at frequency ω_1 for the second band-pass filter, the selectivity parameters are calculated as $Q_1 = Q_2 = 2\sqrt{11}$. Following Section 2.2, the specification of dynamic decoupling for the frequencies ω_1 and $\omega_2 = 2\omega_1$ is satisfied if $F(s)$ is chosen such that

$$F(\imath\omega_1) = \begin{bmatrix} 9.33 + 1.47\imath & -1.77 + 3.03\imath & -0.22 - 0.61\imath \\ -1.06 + 2.26\imath & 14.5 - 0.10\imath & -2.14 + 3.60\imath \\ -0.47 - 0.06\imath & -1.81 + 3.38\imath & 9.86 + 0.82\imath \end{bmatrix}$$

$$F(\imath\omega_2) = \begin{bmatrix} 9.05 + 7.90\imath & -1.90 + 1.45\imath & 0.04 - 0.27\imath \\ -1.23 + 1.07\imath & 14.04 + 9.48\imath & -2.18 + 1.73\imath \\ -0.23 - 0.01\imath & -1.92 + 1.61\imath & 9.57 + 7.60\imath \end{bmatrix}$$

By applying the VFIT3 routine [1] , which is an implementation of fast relaxed VF (Gustavsen and Semlyen, 1999, Gustavsen, 2006, Deschrijver et al., 2008), the decoupler is calculated as

$$F(s) = \begin{bmatrix} \frac{252s + 1022}{s + 112.7} & \frac{-1.52s - 5.30}{s + 0.2737} & \frac{-0.24s + 0.84}{s + 0.7079} \\ \frac{-s - 3.742}{s + 0.1487} & \frac{15s + 0.7703}{s + 0.4881} & \frac{-2.12s - 5.77}{s + 0.0244} \\ \frac{-0.4s - 0.15}{s + 0.5644} & \frac{-1.79s - 5.48}{s + 0.0633} & \frac{10s - 3.77}{s + 0.2245} \end{bmatrix}$$

which achieves the above specification, as demonstrated in Fig. 2, where a comparison between the magnitude Bode plot of the non-decoupled (blue line) and the decoupled one (red line) is depicted. It can be seen that, at the frequencies ω_1 and ω_2, the series interconnection $G(s)F(s)$ approximates an identity matrix, such that a good decoupling is achieved. Then, following the discussion in Section 2.3, t_{trans} is calculated as $t_{trans} = 32.7s$ and, by choosing $T_s = 4t_{trans}$, $T_s = 130.8s$ is obtained.

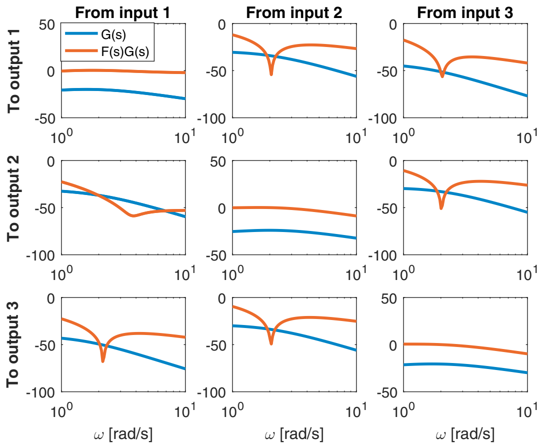

Fig. 2. Decoupling (Bode plot).

4. SIMULATION RESULTS

In order to assess the effectiveness of the proposed strategy, three different simulation scenarios are considered.

4.1 Scenario 1

In the first scenario, the system is working without replay attacks being performed. In Fig. 3, the outputs of the band-pass filters $z_{il}(t)$, $i = 1, 2$, $l = 1, 2, 3$, are plotted with the

[1] https://www.sintef.no/projectweb/vectfit

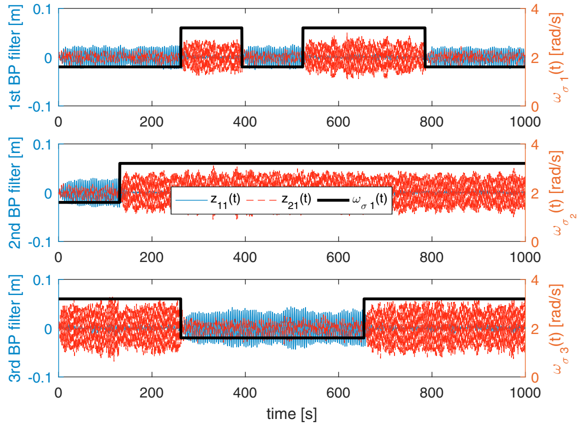

Fig. 3. Scenario 1. Outputs of the band-pass filters $z_{il}(t)$ and varying frequency $\omega_\sigma(t)$.

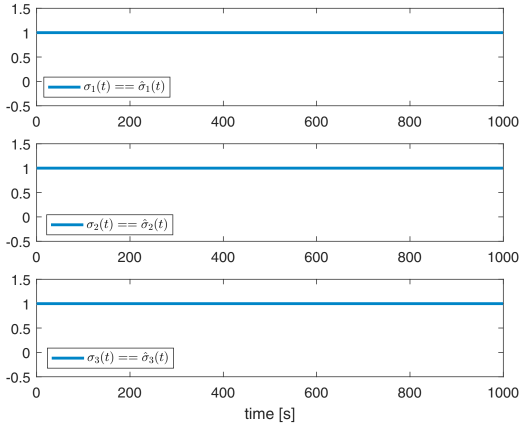

Fig. 4. Scenario 1. Result of the replay attack detection test.

signals $\omega_{\sigma_1(t)}$, $\omega_{\sigma_2(t)}$, $\omega_{\sigma_3(t)}$, which determine the time-varying frequency profile of the signal (5). It can be seen that $z_{1l}(t)$ is the signal with the strongest energy when $\omega_{\sigma_l} = \omega_1 = 1.6\,rad/s$, while $z_{2l}(t)$ is the strongest signal when $\omega_{\sigma_l} = \omega_2 = 3.2\,rad/s$. Using (8), $\hat{\sigma}_1(t)$, $\hat{\sigma}_2(t)$ and $\hat{\sigma}_3(t)$ are calculated, and by comparing them with $\sigma_1(t)$, $\sigma_2(t)$ and $\sigma_3(t)$, respectively, the information about the absence of replay attacks is obtained, as shown in Fig. 4.

4.2 Scenario 2

In the second scenario, it is assumed that an attacker records the measurements of the first output in the first 200s and then replays the recorded data periodically starting from $t = 200s$. In this case, the signals $z_{i1}(t)$ do not follow anymore the profile of $\omega_{\sigma_1(t)}$, as shown in Fig. 5. By detecting a mismatch between $\sigma_1(t)$ and $\hat{\sigma}_1(t)$ (see Fig. 6), a replay attack affecting the first output channel is detected at time 204.1s.

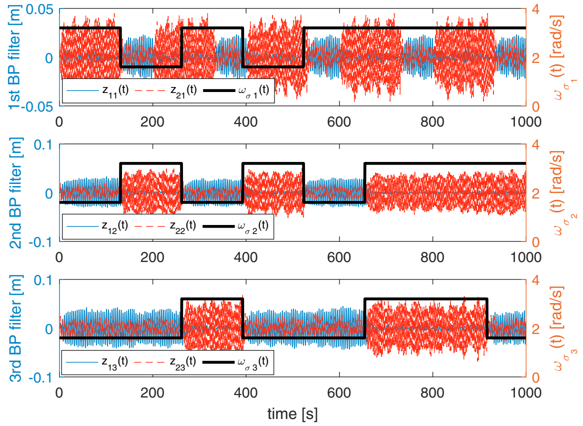

Fig. 5. Scenario 2. Outputs of the band-pass filters $z_{il}(t)$ and varying frequency $\omega_\sigma(t)$.

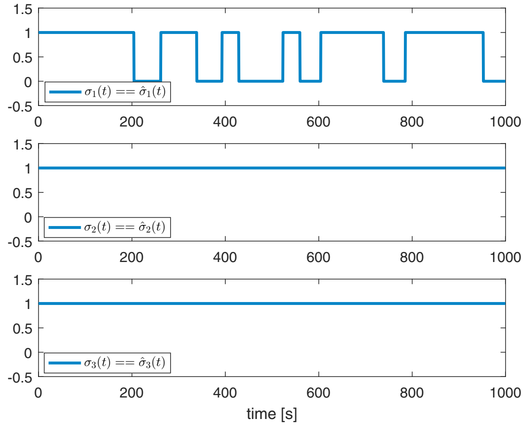

Fig. 6. Scenario 2. Result of the replay attack detection test.

4.3 Scenario 3

In the third scenario, both $y_2(t)$ and $y_3(t)$ are corrupted by the replay attack starting from $t = 200s$. The resulting outputs of the band-pass filters $z_{il}(t)$ are plotted in Fig. 7, together with the varying frequencies $\omega_{\sigma_l(t)}$. The mismatches $\hat{\sigma}_2(t) \neq \sigma_2(t)$ and $\hat{\sigma}_3(t) \neq \sigma_3(t)$ allow detecting the replay attacks affecting the second and the third output channel at $t = 205.6s$ and $t = 428.6s$, respectively. Note that the algorithm needs a longer time to detect the replay attack in the third output channel since until the time $t = 392.5s$, the profile of the randomly generated signal $\sigma_3(t)$ matches with the one in the recorded data ($\sigma_3(t) = 1$).

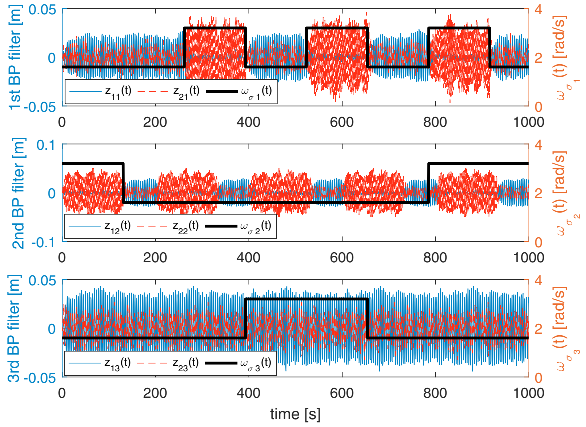

Fig. 7. Scenario 3. Outputs of the band-pass filters $z_{il}(t)$ and varying frequency $\omega_\sigma(t)$.

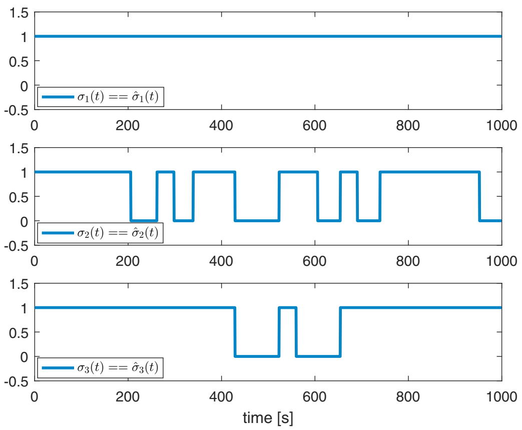

Fig. 8. Scenario 3. Result of the replay attack detection test.

5. CONCLUSIONS

This work has presented an innovative method for detecting replay attacks based on adding a frequency-based authentication signature and its application to a complex and nonlinear system, i.e. a multiple tank system, made up by nine highly interconnected tanks. Three simulation scenarios have illustrated the main characteristics of the method, which is capable of not triggering false alarms while being able to identify successfully which output channels have been corrupted by replay attacks. Future work will aim at extending the method to descriptor systems, in order to apply it to a more realistic model of water supply network.

REFERENCES

C. M. Ahmed, V. R. Palleti, and A. P. Mathur. Wadi: a water distribution testbed for research in the design of secure cyber physical systems, 2017.

S. Amin, X. Litrico, S. Sastry, and A. M. Bayen. Cyber security of water scada systems - part i: Analysis and experimentation of stealthy deception attacks. *IEEE Transactions on Control Systems Technology*, 21(5):1963–1970, 2013a.

S. Amin, X. Litrico, S. Sastry, and A. M. Bayen. Cyber security of water scada systems - part ii: Attack detection using enhanced hydrodynamic models. *IEEE Transactions on Control Systems Technology*, 21(5):1679–1693, 2013b.

M. Buciakowski, M. Witczak, V. Puig, D. Rotondo, F. Nejjari, and J. Korbicz. A bounded-error approach to simultaneous state and actuator fault estimation for a class of nonlinear systems. *Journal of Process Control*, 52:14–25, 2017.

D. Deschrijver, M. Mrozowski, T. Dhaene, and D. De Zutter. Macromodeling of multiport systems using a fast implementation of the vector fitting method. *IEEE Microwave and Wireless Components Letters*, 18(6):383–385, 2008.

H. L. Gawand, A. K. Bhattacharjee, and K. Roy. Online monitoring of a cyber physical system against control aware cyber attacks. *Procedia Computer Science*, 70:238–244, 2015.

H. L. Gawand, A. K. Bhattacharjee, and K. Roy. Securing a cyber physical system in nuclear power plants using least square approximation and computational geometric approach. *Nuclear Engineering and Technology*, 49(3):484–494, 2017.

S.-C. Georgescu, R. Popa, and A.-M. Georgescu. Pumping stations scheduling for a water supply system with multiple tanks. *University Politehnica of Bucharest Scientific Bulletin, Series D Mechanical Engineering*, 72(3):129–140, 2010.

B. Gustavsen. Improving the pole relocating properties of vector fitting. *IEEE Transactions on Power Delivery*, 21(3): 1587–1592, 2006.

B. Gustavsen and A. Semlyen. Rational approximation of frequency domain responses by vector fitting. *IEEE Transactions on Power Delivery*, 14(3):1052–1061, 1999.

Y. Y. Haimes, N. C. Matalas, J. H. Lambert, B. A. Jackson, and J. F. R. Fellows. Reducing vulnerability of water supply systems to attack. *Journal of Infrastructure Systems*, 4(4): 164–177, 1998.

K. H. Johansson. The quadruple-tank process: A multivariable laboratory process with an adjustable zero. *IEEE Transactions on control systems technology*, 8(3):456–465, 2000.

K. Kashima and D. Inoue. Replay attack detection in control systems with quantized signals, 2015.

K.-D. Kim and P. R. Kumar. Cyber–physical systems: A perspective at the centennial. *Proc. of the IEEE*, 100(Special Centennial Issue):1287–1308, 2012.

A. Laszka, W. Abbas, Y. Vorobeychik, and X. Koutsoukos. Synergic security for smart water networks: redundancy, diversity, and hardening, 2017.

M. Ma, P. Zhou, D. Du, C. Peng, M. Fei, and H. M. AlBuflasa. Detecting replay attacks in power systems: A data-driven approach. In *Advanced Computational Methods in Energy, Power, Electric Vehicles, and Their Integration*, pages 450–457. Springer, 2017.

Y. Mo and B. Sinopoli. Secure control against replay attacks, 2009.

L. E. Ormsbee and K. E. Lansey. Optimal control of water supply pumping systems. *Journal of Water Resources Planning and Management*, 120(2):237–252, 1994.

K.-J. Park, R. Zheng, and X. Liu. Cyber-physical systems: Milestones and research challenges. *Computer Communications*, 36(1):1–7, 2012.

A. Rasekh, A. Hassanzadeh, S. Mulchandani, S. Modi, and M. K. Banks. Smart water networks and cyber security, 2016.

J. Shi, J. Wan, H. Yan, and H. Suo. A survey of cyber-physical systems, 2011.

J. Slay and M. Miller. Lessons learned from the maroochy water breach. *Critical infrastructure protection*, pages 73–82, 2007.

B. Tang, L. D. Alvergue, and G. Gu. Secure networked control systems against replay attacks without injecting authentication noise, 2015.

R. Taormina, S. Galelli, N. O. Tippenhauer, E. Salomons, and A. Ostfeld. Characterizing cyber-physical attacks on water distribution systems. *Journal of Water Resources Planning and Management*, 143(5):04017009, 2017.

Y. Wei and S. Li. Water supply networks as cyber-physical systems and controllability analysis. *IEEE/CAA Journal of Automatica Sinica*, 2(3):313–319, 2015.

M. Zhu and S. Martínez. On the performance analysis of resilient networked control systems under replay attacks. *IEEE Transactions on Automatic Control*, 59(3):804–808, 2014.

H. Zumbahlen. *Linear circuit design handbook*. Elsevier Newnes Press, 2008.