Diseño de un desacoplador dinámico para la detección de ataques de reproducción

Helem S. Sánchez, Marc López Vidal, Damiano Rotondo, Joseba Quevedo
Centro de Investigación en Supervisión, Seguridad y Control Automático (CS2AC)
Rambla de Sant Nebridi 22, 08222 Terrassa, España. e-mail: joseba.quevedo@upc.edu

Resumen

Este trabajo presenta el diseño de un desacoplador dinámico en un rango de frecuencia determinado para la detección de ataques de reproducción. Inicialmente se estudia el comportamiento del desacoplador al variar su orden. Después, se realiza una comparación cuantitativa para la cual se propone una función de coste, que permitirá comparar el valor que toma la función en una frecuencia dada respecto al valor deseado para el funcionamiento correcto del desacoplador. Finalmente, se ilustra este enfoque a través de un ejemplo numérico en el cual los resultados del desacople son satisfactorios y a su vez la efectividad del mismo en la detección de ataques de reproducción.

Palabras clave: ataques cibernéticos, ataques de reproducción, desacopladores, firma frecuencial.

1 Introducción

Los sistemas ciberfísicos integran capacidades de computación, almacenamiento y comunicación para controlar e interactuar con un proceso físico [8]. Es importante mencionar que la integración de componentes cibernéticos y físicos aumenta la eficiencia de los sistemas, pero al mismo tiempo los hace susceptibles a posibles ataques cibernéticos.

Los ataques cibernéticos son genéricos, pueden afectar los procesos físicos a través de la realimentación, y así afectar muchos componentes del sistema de forma coordinada y además podrían rediseñarse para afectar otros CPSs. Los ataques cibernéticos pueden ser maliciosos o involuntarios, y pueden ocurrir en el ciberespacio, en el mundo físico o en ambos. Son capaces de violar la integridad, confiabilidad y disponibilidad de los elementos de un sistema siendo capaces de difundirse en cuestión de segundos, sin ser detectados por la unidad de control.

En este trabajo, nos centraremos en los ataques de reproducción, los cuales son de los más críticos que pueden afectar los sistemas ciberfísicos. Este tipo de ataque se lleva a cabo bajo las siguientes hipótesis:

1. El sistema se encuentra en estado estacionario o periódico cuando el atacante realiza la acción;
2. Suponemos que el atacante tiene control sobre todos los sensores;
3. El lazo de control podría ser interrumpido a causa del ataque.

El ataque de reproducción consta de dos fases:

1. En la primera fase del ataque, que no afecta la dinámica del sistema, se recopilan los datos sin perturbar al sistema, para que luego el adversario pueda utilizarlo para la fase posterior del ataque.
2. En la segunda fase del ataque, el atacante reproduce los datos recopilados, reemplazando los datos reales procedentes de los sensores. De esta manera, el atacante puede llevar a cabo un ataque físico sin ser descubierto y a la vez causar el deterioro del rendimiento del sistema de control.

La detección de los ataques de reproducción se consideró por primera vez en [7], en donde el análisis de las condiciones de detección para este ataque muestra que, asintóticamente, los ataques de repetición son indetectables a menos de interactuar de forma activa con el sistema. Para poder detectar estos ataques, los autores propusieron un esquema de detección a través de marca de agua aditiva, donde un ruido es intencionalmente inyectado en el sistema a través de los actuadores. Sin embargo, en [12] los autores utilizan marcas de agua multiplicativas para evitar que el rendimiento del sistema disminuya y que los actuadores estén cargados con ruido en las entradas. Recientemente, han surgido enfoques alternativos como por ejemplo en [6], donde los autores aplicaron un juego estocástico. Por otro lado, en [14] se investigó la variación de la ley de control de retroceso horizonte, para hacer frente a este tipo de ataque y analizar la degradación del rendimiento del sistema. Otras técnicas propuestas recientemente son: los métodos basados

en datos [5], las señales cuantizadas [4] y la estimación espectral [11].

La aportación de este trabajo es utilizar un método basado en una *firma frecuencial* para los ataques de reproducción en un sistema ciberfísico centrándose en el diseño de uno de sus componentes, el desacoplador dinámico. El caso de estudio consiste en un ejemplo numerico, que permite analizar como cambian los resultados respecto a los diferentes parámetros. Una vez hecho dicho análisis, se llevará acabo la implementación del método intentando de que sea de la forma más eficiente posible.

2 Detector Lógico

El generador de señales es uno de los componentes del detector de ataques de reproducción basado en la frecuencia (ver la Figura 1 para el diagrama conceptual del sistema). El enfoque propuesto en este trabajo se desarrolló inicialmente en [9], donde se aplicó a un sistema de nueve tanques para la detección de ataques de reproducción.

Para empezar consideraremos un sistema lineal como se muestra a continuación:

$$\dot{x}(t) = Ax(t) + Bu(t) + Dd(t) \tag{1}$$
$$y(t) = Cx(t) + Ev(t) \tag{2}$$

donde x es la variable de estado, u la entrada, y la salida, d una perturbación exógena, v el ruido que afecta a los sensores y A, B, C, D y E son matrices conocidas.

Cuando se utilice dicho enfoque, $u(t)$ está formada por dos señales diferentes:

$$u(t) = u^*(t) + \Delta u(t) \tag{3}$$

donde $u^*(t)$ es la señal de control, elegida como la combinación de una ley de prealimentación $u^*_{ff}(t)$ y realimentación $u^*_{fb}(t)$:

$$u^*(t) = u^*_{ff}(t) + u^*_{fb}(t) \tag{4}$$

mientras que la firma $\Delta u(t)$ debe ser una señal de media cero para que no se introduzca ningún bias en $x(t)$.

La señal $u^*_{fb}(t)$ en (4) es una ley típica de control por realimentación de error lineal. De acuerdo con el principio del modelo interno, si se desea realizar un seguimiento de una trayectoria de referencia $y_{ref}(t)$, es necesario incluir su generador dentro del circuito de control [1]. A continuación, se utilizará una trayectoria de referencia constante, de modo que se debe elegir una estructura de controlador proporcional integral (PI) para el controlador el cual puede describirse de la siguiente forma [2]:

$$u^*_{fb}(t) = K_P\left(y_{ref}(t) - y(t)\right) + K_I x_I(t) \tag{5}$$
$$\dot{x}_I(t) = y_{ref}(t) - y(t) \tag{6}$$

donde K_P y K_I indican la ganancia proporcional e integral, respectivamente. En consecuencia, el sistema (1)-(2) puede ser expresado a través del siguiente sistema aumentado:

$$\begin{aligned}\dot{x}_{aug}(t) = A_{aug}x_{aug}(t) + \begin{bmatrix} BK_P \\ I \end{bmatrix} y_{ref}(t) \\ + B_{aug}\left(u^*_{ff}(t) + \Delta u(t)\right) \\ + \begin{bmatrix} D \\ 0 \end{bmatrix} d(t) + \begin{bmatrix} -BK_PE \\ -E \end{bmatrix} v(t)\end{aligned} \tag{7}$$

$$y(t) = C_{aug}x_{aug}(t) + Ev(t) \tag{8}$$

con $x_{aug}(t) = \begin{bmatrix} x(t)^T & x_I(t)^T \end{bmatrix}^T$ y:

$$A_{aug} = \begin{bmatrix} A - BK_PC & BK_I \\ -C & 0 \end{bmatrix}$$

$$B_{aug} = \begin{bmatrix} B \\ 0 \end{bmatrix} \quad C_{aug} = \begin{bmatrix} C & 0 \end{bmatrix}$$

3 Desacoplamiento Dinámico

La técnica de la firma frecuencial tiene como objetivo detectar un ataque de reproducción introduciendo la señal de autenticación $\Delta u(t)$ en el sistema (1)-(2), y detectando si el resultado medido es compatible con la $\Delta u(t)$ introducida o no. Para hacerlo, es necesario establecer una biyección entre las entradas disponibles y las salidas disponibles, de tal manera que el efecto de un elemento de $\Delta u(t)$, es decir $\Delta u_l(t)$, $l = 1, \ldots, L$, se observará en, y solo en, la salida asociada $y_l(t)$.

Sin embargo, la matriz de lazo cerrado de $\Delta u(t)$ a $y(t)$, es decir, $G(s) = C_{aug}(sI - A_{aug})^{-1}B_{aug}$ suele estar *acoplada*, ya que cada entrada individual influye en todas las salidas, lo que dificulta el establecimiento de dicha biyección. El manejo de estos acoplamientos (términos no diagonales en $G(s)$) es un problema bastante conocido para el cual existen resultados disponibles en la literatura, como por ejemplo [10, 13]. Para lograr este objetivo, un desacoplador $F(s)$ podría introducirse en el ciclo de modo que la interconexión en serie de $F(s)$ y $G(s)$ sea *dinámicamente desacoplada*, es decir, la matriz de transferencia $G_d(s) = G(s)F(s)$ sea diagonal y el sistema aumentado se pueda percibir como consistente de subsistemas independientes. Para evitar una ley de desacople compleja y altamente sensible, se considerará el *desacoplamiento dinámico en un rango de frecuencia determinado.*

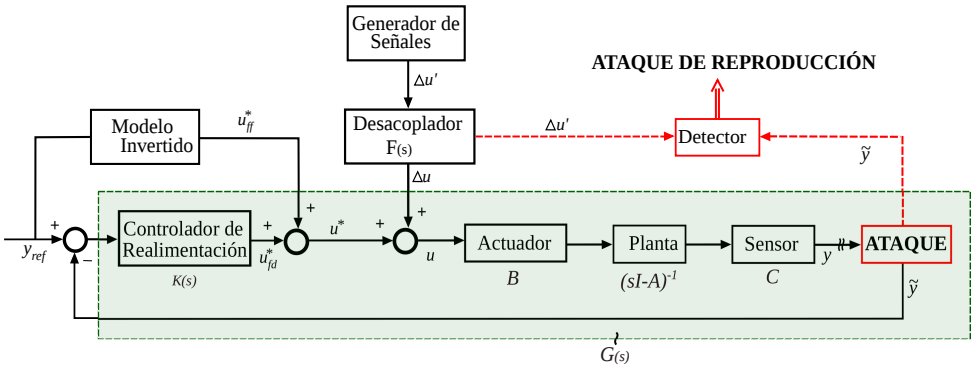

Figura 1: Diagrama del sistema.

La solución desarrollada en [9] se basa en el *vector fitting* (VF) [3], el cual es un método numérico robusto para la aproximación racional en el dominio de frecuencia utilizando polos y residuos. Por ejemplo, dado el sistema (1)-(2), se desea diseñar el desacoplador:

$$\dot{x}_d(t) = A_d x_d(t) + B_d \Delta u'(t) \tag{9}$$
$$\Delta u(t) = C_d x_d(t) + D_d \Delta u'(t) \tag{10}$$

de tal forma que $\forall i = 1, \ldots, N$, $G_d(\iota\omega_i)$ calculada usando $F(s) = C_d(sI - A_d)^{-1}B_d + D_d$ aproxima a una matriz de identidad. Para lograr este objetivo, se necesita $F(\iota\omega_i) = G(\iota\omega_i)^{-1}$, lo que proporciona un conjunto de restricciones N que el desacoplador (9)-(10) debe cumplir [1].

4 Generación de la firma frecuencial

La idea del enfoque de la firma frecuencial es introducir en el sistema (1)-(2) y el desacoplador (9)-(10) señales sinusoidales cuyas frecuencias varían en el tiempo. La posibilidad más simple es considerar señales de la forma:

$$\Delta u'_l(t) = \alpha_l \cos\left(\omega_{\sigma_l(t)} t\right) \qquad l = 1, \ldots, L \tag{11}$$

donde α_l indica la magnitud, mientras que $\sigma_l(t)$ indica una señal constante a trozos, que toma valores enteros entre 1 y N, de tal manera que en cada instante de tiempo $\omega_{\sigma_l(t)}$ sea igual a una de las ω_i, $i = 1, \ldots, N$, para las cuales se logra el desacoplamiento mediante el desacoplador (9)-(10), como se explicó en la sección anterior. Se supone que la señal $\sigma_l(t)$ cambia su valor a un valor aleatorio entre 1 y N, que podría ser el mismo que el valor anterior, en instantes de tiempo igualmente espaciados $t_s^{(j)}$, $j \in \mathbb{N}_0$, con $t_s^{(0)} = 0$ y $t_s^{(j+1)} - t_s^{(j)} = T_s$, donde T_s es el periodo de cambio. Seguidamente, indicaremos el valor tomado por $\omega_{\sigma_l(t)}$ en el intervalo $[t_s^{(j)}, t_s^{(j+1)}]$ como ω_{jl}. Vale la pena señalar que la señal constante a trozos $\sigma_l(t)$ es completamente conocida por el detector, mientras que el atacante no tiene acceso a esta información.

[1] Nótese que $G(\iota\omega_i)$ es un número complejo, no una función de transferencia y, en general $F(s) \neq G(s)^{-1}$. Además, en casos donde $G(s)$ tiene ceros con partes reales positivas, una $F(s)$ estable que satisface $F(\iota\omega_i) = G(\iota\omega_i)^{-1}$, $i = 1, \ldots, N$, puede ser calculada.

5 Filtros paso banda

La respuesta del sistema aumentado formado por (7)-(8) será la suma de la respuesta natural, que puede ser despreciada debido a la suposición de que el sistema esté en estado estacionario, y las respuestas forzadas debido a las entradas que actúan en él, o sea, $y_{ref}(t)$, $u^*_{ff}(t)$, $\Delta u'(t)$, $d(t)$ y $v(t)$. Con el objetivo de analizar solo el contenido de $y(t)$ en las frecuencias ω_i, $i = 1, \ldots, N$, usadas para generar la señal de firma $\Delta u'(t)$, el sistema aumentado se conecta en cascada con un banco de filtros $H_i(s)$. En particular, cada $H_i(s)$ es una matriz de transferencia diagonal $n_y \times n_y$, con cada elemento en la diagonal elegido como un filtro paso banda de segundo orden, es decir [15]:

$$H_i(s) = diag\left\{ \frac{\frac{\omega_i}{Q_i} s}{s^2 + \frac{\omega_i}{Q_i} s + \omega_i^2} \right\} \tag{12}$$

donde ω_i es la frecuencia a la que el filtro alcanza su punto máximo y Q_i es la selectividad del filtro.

Filtrando la diferencia entre $y_{ref,l}(t)$ y $y_l(t)$ para que el filtro paso banda (12) extraiga solo la información relevante para la detección del ataque de reproducción, se obtiene la señal $z_{il}(t)$, que es utilizada por el algoritmo de detección.

6 Algoritmo de detección

El algoritmo de detección para el ataque de reproducción se basa en la comparación de la señal

constante a trozos $\sigma_l(t)$ con $\hat{\sigma}_l(t)$, la cual es una reconstrucción basada en las señales $z_{il}(t)$ obtenidas como salidas de los filtros paso banda (12). Por lo general, siempre que $\hat{\sigma}_l(t) = \sigma_l(t)$, $l = 1, \ldots, L$, el algoritmo proporcionará la información de que no hay ataque de reproducción en la salida $y_l(t)$. Por otro lado, si $\hat{\sigma}_l(t) \neq \sigma_l(t)$, el algoritmo advertirá que la salida $y_l(t)$ ha sido afectada por un ataque de reproducción.

Es importante señalar que la efectividad del algoritmo depende de cómo se calcula la señal $\hat{\sigma}_l(t)$. Una elección simple sería comparar las energías de los diferentes $z_{il}(t)$ durante el período más grande asociado con las frecuencias ω_i, $i = 1, \ldots, N$, es decir durante los intervalos $[t - T_\omega, t]$, con:

$$T_\omega = \max_{i=1,\ldots,N} \frac{2\pi}{\omega_i} \tag{13}$$

y determinar $\hat{\sigma}_l(t)$ como el índice correspondiente a la señal con la mayor energía, tal que:

$$\hat{\sigma}_l(t) = \arg \max_{i=1,\ldots,N} \int_{t-T_\omega}^{t} |z_{il}(\tau)|^2 d\tau \tag{14}$$

Sin embargo, cuando ocurre un cambio en la frecuencia de la señal $\omega_{\sigma_l}(t)$ en (11), el sistema experimentará un comportamiento transitorio con respecto a la señal $\Delta u'(t)$, lo cual afectará la coincidencia entre $\sigma_l(t)$ y $\hat{\sigma}_l(t)$. En estos casos, la mejor opción es tener en cuenta el tiempo necesario para que dicho transitorio se vuelva despreciable, indicado a continuación como t_{trans}, y calcular $\hat{\sigma}_l(t)$ como (15) (ver página siguiente), donde $t_s^* = \lfloor t/T_s \rfloor T_s$ indica el último tiempo de conmutación.

Vale la pena señalar que el cálculo analítico de t_{trans}, aunque sea posible, no es una tarea fácil, ya que el sistema global compuesto por desacoplador, planta, controlador y el filtro paso banda es un sistema de alto orden. Sin embargo, dado que los filtros paso banda $H_i(s)$ determinan el contenido de frecuencia de las señales de salida, una estimación razonable de t_{trans} es dada por el mayor entre los tiempos de establecimiento de $H_i(s)$, $i = 1, \ldots, N$.

7 Ejemplo Numérico

El método de detección se ha aplicado a un sistema multivariable descrito por las ecuaciones (1)-(2), del modelo en espacio de estado:

$$A = \begin{bmatrix} -3.15 & 4.06 & -3.73 & 4.13 \\ 1.32 & -4.02 & -2.21 & 0.47 \\ 4.58 & 4.65 & -3.42 & 4.71 \\ 4.57 & -0.15 & 3.00 & -3.58 \end{bmatrix}$$

$$B = \begin{bmatrix} -0.78 & 4.16 & 3.15 \\ 2.92 & 4.59 & 4.06 \\ 1.56 & -4.64 & 4.13 \\ 3.49 & 4.34 & -3.73 \end{bmatrix}$$

$$C = \begin{bmatrix} 1.79 & 2.58 & 2.43 & -1.08 \\ 1.55 & -3.29 & 2.06 & -4.68 \\ -2.23 & -4.54 & -4.03 & 3.23 \end{bmatrix}$$

$$D = 10^{-3} \begin{bmatrix} 1.95 & -1.83 & 4.50 & -4.66 \end{bmatrix}^T$$

$$E = 10^{-3} \begin{bmatrix} -0.61 & -1.18 & 2.66 \\ 2.95 & -3.13 & -0.10 \\ -0.54 & 1.46 & 2.09 \end{bmatrix}$$

el cual sigue la referencia constante $y_{ref}(t) = \begin{bmatrix} 99.2 & -175.7 & -104.8 \end{bmatrix}^T$ que corresponde a la acción de pre-alimentación $u_{ff}^*(t) = \begin{bmatrix} 281.12 & -177.06 & -64.53 \end{bmatrix}^T$

Para seguir la referencia constante de $y_{ref}(t)$ con error nulo a régimen, se ha implementado un controlador proporcional integral (PI) diseñado por asignación de polos:

$$K_I = \begin{bmatrix} 1.72 & -0.20 & -2.30 \\ 0.94 & 0.16 & -0.59 \\ -1.12 & 2.13 & -1.49 \end{bmatrix}$$

$$K_P = \begin{bmatrix} 2.73 & 0.40 & 0.04 \\ 0.96 & 0.12 & 0.02 \\ -3.05 & 1.27 & -1.28 \end{bmatrix}$$

8 Diseño del Desacoplador

Para calcular las matrices del desacoplador, se ha usado la rutina VFIT3[2], que proporciona una herramienta relativamente sencilla para aplicar el método del *VF*. Una vez se haya calculado el desacoplador usando la rutina VFIT3 hay que comprobar su funcionamiento. Para eso, se grafican los diagramas de Bode de cada entrada y salida, de manera que se obtendrá una cuadrícula donde cada fila representa una salida y cada columna una entrada. Se considera que el desacoplador ha logrado un buen funcionamiento si cada salida queda afectada únicamente por la entrada correspondiente. De manera que cualquier relación entre entradas y salidas que no sea ésta, debe estar totalmente desacoplada. Para que esto ocurra, los gráficos de la diagonal (posiciones 1-1, 2-2 y 3-3) no deben tener atenuación ($0dB$) en los puntos de frecuencia ω_1 y $\omega_2 = 2\omega_1$ y los gráficos correspondientes a las relaciones que no sean las de la diagonal, en las frecuencias ω_1 y $\omega_2 = 2\omega_1$, deberían tener una atenuación superior al valor considerado aceptable, elegido como $20dB$.

Al verificar el funcionamiento del desacoplador con el ejemplo numérico, se ha observado que un

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

$$\hat{\sigma}_l(t) = \begin{cases} \sigma_l(t) & si\ \sigma_l(t) \neq \sigma_l(t - T_s)\ \wedge\ t \in [t_s^*, t_s^* + t_{trans} + T_\omega] \\ \arg \max_{i=1,\ldots,N} \int_{t-T_\omega}^{t} |z_{il}(\tau)|^2 d\tau & \text{de otro modo} \end{cases} \quad (15)$$

parámetro de diseño muy importante es el orden del desacoplador M. Según el orden del desacoplador, para el funcionamiento correcto del VFIT3, habrá que definir las restricciones con las cuales el desacoplador deberá cumplir en un número de frecuencias igual a $M+1$. Por eso, para un desacoplador de orden uno, habrá que definir las restricciones en dos frecuencias, para el de orden dos en tres frecuencias, y así sucesivamente. Seguidamente, se analiza el comportamiento de los desacopladores al variar su orden y el número de frecuencias.

Para obtener una comparación cuantitativa entre los distintos desacopladores obtenidos, se ha propuesto una función de coste J, expresada de la siguiente manera:

$$J(\omega^*) = \sum_{i=1}^{3} (F_{ii}(\omega^*) - F_{ii}^*)^2 + \sum_{i=1}^{3} \sum_{\substack{j=1 \\ i \neq j}}^{3} \left(\max\left(F_{ij}(\omega^*), -20dB\right) - F_{ij}^*\right) \quad (16)$$

donde ω^* es la frecuencia que se esté analizando, F es el valor que toma la función en la frecuencia correspondiente y F^* es el valor deseado que debería tomar la función para el correcto funcionamiento del desacoplador. Los subíndices i y j corresponden respectivamente a la entrada y la salida que se esté analizando. De manera que si los subíncides son 2 y 3 y se está calculando para una $\omega^* = 0.6$ s^{-1}, esto corresponde al valor de la función en 0.6 s^{-1} del sub-gráfico de la fila 2 columna 3. Esta función de coste suma diferentes términos: por un lado, en el primer término de la ecuación (16) acumula los valores de los gráficos de la diagonal, es decir los que representan el desacople entre la salida y la entrada del mismo canal y se calculan con el error al cuadrado entre el valor real y el deseado, o sea $0dB$. Para el resto de gráficos (los que no están en la diagonal) se calcula el error al cuadrado del máximo entre el valor deseado $-20dB$ y el valor real y se resta con el valor deseado de -20 dB, lo cual corresponde al segundo término de la ecuación (16). Para las dos frecuencias donde se quiere obtener el desacople (ω_1 y ω_2), se obtienen dos valores diferentes de la función de coste. La función de coste total que evalúa la bondad del desacoplador se obtiene posteriormente como la media de esos valores. Eso quiere decir, que para los casos de desacopladores de orden superior a uno, se calcula el valor de J solo para los valores de frecuencia ω_1 y ω_2, ya que las otras frecuencias utilizadas han sido introducido solo para satisfacer los vínculos necesarios por el V3FIT.

En primer lugar se ha analizado el comportamiento del desacoplador de orden uno ($\omega_1 = 0.6$ s^{-1} y $\omega_2 = 1.2$ s^{-1}). En la Figura 2 se muestra el comportamiento del desacoplador. Como se puede ver, el desacople no es del todo efectivo, ya que para las frecuencias seleccionadas no atenúa como debería.

Seguidamente, se ha propuesto un desacoplador de segundo orden, de manera que en vez de dos frecuencias para la definición de las restricciones, se han utilizado tres (ω_a, ω_b y ω_c). Esto hace que aumenten los estados del sistema total y por lo tanto se requiere mayor potencia para efectuar las simulaciones, pero también permite obtener una mayor precisión y eficiencia en cuanto al desacople. Para ello, se ha vuelto a utilizar como frecuencia base $\omega_a = \omega_1 = 0.6$ s^{-1}, como segunda una intermedia entre ω_a y ω_c ($\omega_b = 0.8$ s^{-1}) y como tercera el doble de la primera ($\omega_c = \omega_2 = 1.2$ s^{-1}). Se puede analizar el resultado en la Figura 3, que enseña que, no obstante la mejora obtenida respecto al desacoplador de orden uno, el de orden dos aún no permite obtener un desacople satisfactorio.

Por esa razón, como última modificación en cuanto al orden, se ha intentado con un desacoplador de cuatro frecuencias, es decir, de tercer orden. Este cambio aumenta mucho la precisión al par que la potencia de cálculo requerida. La selección de frecuencias para este caso es como frecuencia base $\omega_a = \omega_1 = 0.6\,s^{-1}$, como segunda ($\omega_b$) y tercera ($\omega_c$) los puntos medios entre la primera y la última, $0.8\,s^{-1}$ y $0.9\,s^{-1}$ respectivamente, y por último $\omega_d = \omega_2 = 1.2$ s^{-1}. Con esta modificación se ha obtenido una notable mejora tal como se puede ver en el gráfico 4.

Esta conclusión queda confirmada por los valores de la función de coste, tal como se puede ver en la Tabla 1, donde se presentan los resultados de la comparación entre los distintos ordenes de desacopladores utilizando la función de coste J previamente definida.

Finalmente, se ha decidido que el mejor desacoplador es el de tercer orden, ya que, pese al requerimiento de potencia de cálculo que necesita, los resultados obtenidos son altamente satisfactorios en cuanto al desacople y por tanto en la de-

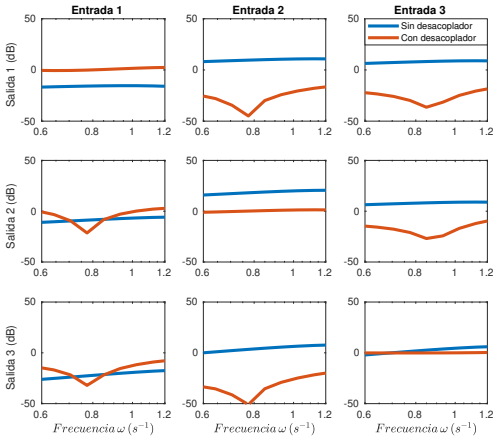

Figura 2: Desacoplador de orden uno.

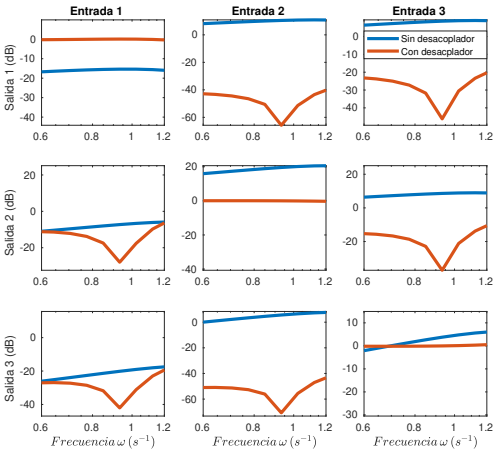

Figura 3: Desacoplador de orden dos.

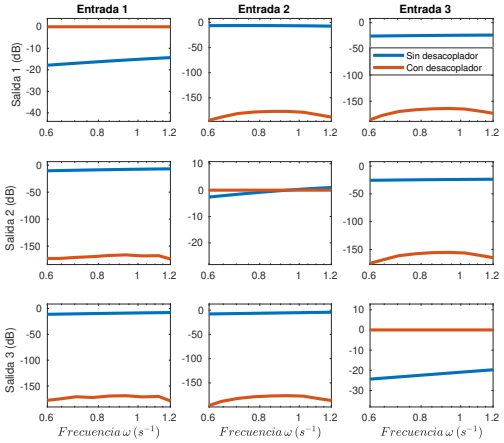

Figura 4: Desacoplador de orden tres.

Tabla 1: Comparación de J.

Orden desacoplador	$J(\omega^*)$	J media
1	455,25 797,75	626,5
2	101,25 272,75	187,0
3	- -	0,0

tección del ataque.

9 Resultados de simulación

A continuación, se plantean tres escenarios para simular. En el Escenario 1 se simula que ninguna salida esté siendo afectada por el atacante, en el Escenario 2 que todas las salidas estén siendo atacadas y en el Escenario 3 que solo la salida 1 esté siendo afectada.

9.1 Escenario 1

En el primer escenario el sistema no queda afectado por ataques de reproducción. En la Figura 5, las salidas de los filtros paso banda $z_{ij}(t), i = 1, 2, j = 1, 2$ se grafican junto al valor de la frecuencia de la señal sinusoidal correspondiente a dichos filtros. Se puede ver que cuando el valor de ω_{σ_l} es $\omega_1 = 0.6s^{-1}$, la señal predominante es z_{1l}, mientras que cuando el valor de ω_{σ_l} es $\omega_2 = 1.2s^{-1}$, la señal predominante es z_{2l}. Basándose en la ecuación (15), se pueden calcular $\hat{\sigma_1}$ y la $\hat{\sigma_2}$, tal y como se enseña en la Figura 6: las señales coinciden, lo cual proporciona la información que ninguna salida esté siendo atacada.

9.2 Escenario 2

En la Figura 7, se observa que a partir de los 200 segundos, el comportamiento de las salidas de los filtros paso banda (señales azul y verde) se repite debido a la acción del ataque. La Figura 8 enseña las señales σ_1 y σ_2 comparadas con $\hat{\sigma_1}$ y $\hat{\sigma_2}$ para cada salida. Debido al desajuste entre las señales, el sistema puede detectar de manera individual el ataque en cada una de las salidas.

9.3 Escenario 3

En el caso en que solo una salida está siendo afectada (Figura 9) la estructura de la señal del primer filtro paso banda se repite a partir de los 200 segundos. La ω_{σ_1} no falla justo en el momento en que se empieza a repetir la estructura base por cuestiones de azar, eso implica que en este caso el ataque tardaría unos segundos más en detectarse. Conclusiones parecidas se pueden sacar de

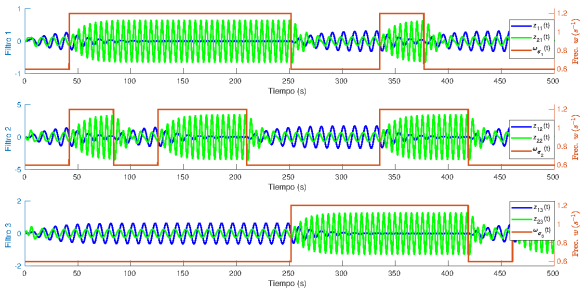

Figura 5: Señales $z_{il}(t)$ y $\omega_{\sigma_l(t)}$ (Escenario 1).

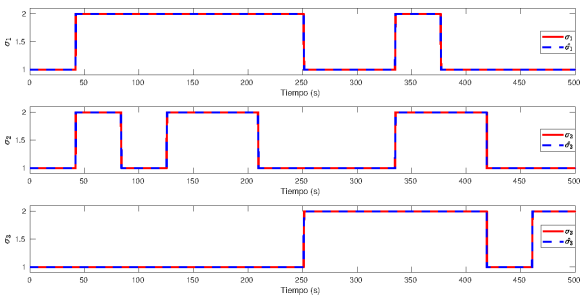

Figura 6: σ_1, σ_2 y sus estimaciones $\hat{\sigma_1}$,$\hat{\sigma_2}$ (Esc. 1).

la Figura 10 que muestra la relación entre las σ reales y las $\hat{\sigma}$ estimadas, de manera que se puede detectar el ataque de reproducción aparece y aislar el canal que está siendo afectado.

10 Conclusiones

En este trabajo, se ha presentado un método de detección de ataques cibernéticos denominados de reproducción. El método de detección consiste en aplicar una señal aditiva sinusoidal con frecuencia variable, y analizar la firma frecuencial en la salida. Se ha realizado un estudio frecuencial para poder encontrar un buen desacoplador para una banda de frecuencias, que pueda ser utilizado de manera eficaz por el detector. En particular, se ha enseñado que el orden del desacoplador es un parámetro muy importante a la hora de encontrar una solución efectiva en relación al desacople obtenido. El método de detección se ha simulado en tres escenarios, en todos los cuales se ha demostrado su buen funcionamiento y efectividad.

Agradecimientos

Este trabajo ha sido realizado gracias al apoyo de MINECO y FEDER a través de los proyectos CICYT HARCRICS (ref.DPI2014-58104-R) y SCAV (ref.DPI2017-88403-R) y de la Agencia Estatal de Investigación (AEI) mediante el sello de excelencia científica María de Maetzu al IRI (ref. MDM-2016-0656) y la ayuda Juan de la Cierva-Formación (ref. FJCI-2016-29019).

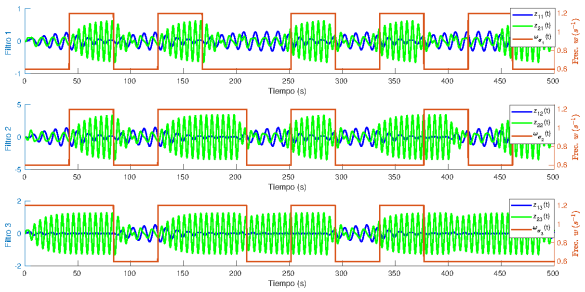

Figura 7: Señales $z_{il}(t)$ y $\omega_{\sigma_l(t)}$ (Escenario 2).

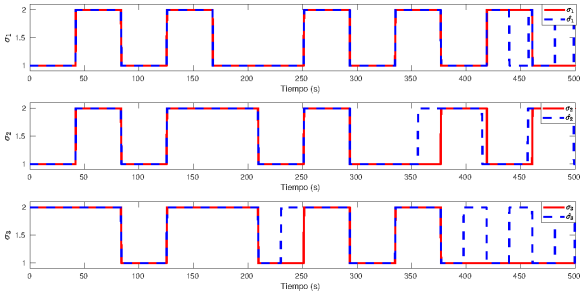

Figura 8: σ_1, σ_2 y sus estimaciones $\hat{\sigma_1}$,$\hat{\sigma_2}$ (Esc. 2).

English summary

DESIGN OF A DYNAMICAL DECOUPLER FOR THE DETECTION OF REPLAY ATTACKS

Abstract

This paper presents the design of a dynamical decoupler in a predetermined range of frequencies, for the detection of replay attacks. At first, the behavior of the decoupler is studied by changing its order. Later, a quantitative comparison is performed by using a cost function, which allows comparing the decoupling performance at a given frequency with respect to the desired value needed for a satisfactory performance of the decoupler. Finally, this technique is illustrated by means of a numerical example in which the decoupling performance and effectiveness in detecting replay attacks are deemed to be satisfactory.

Keywords: Cyber attacks, replay attacks, decouplers, frequency-based signature.

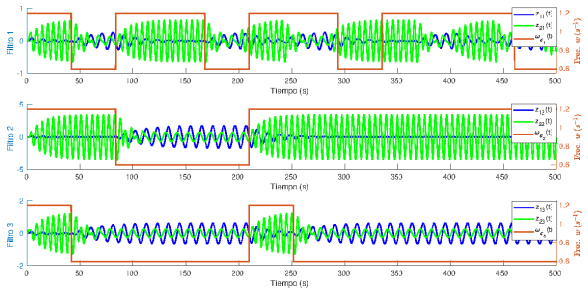

Figura 9: Señales $z_{il}(t)$ y $\omega_{\sigma_l(t)}$ (Escenario 3).

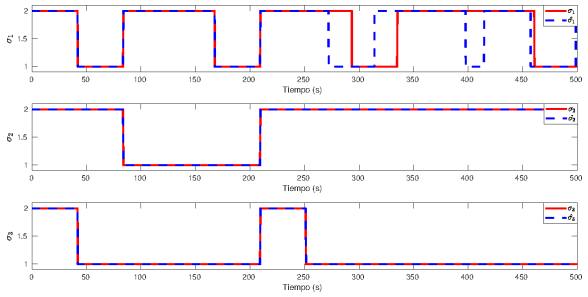

Figura 10: σ_1, σ_2 y sus estimaciones $\hat{\sigma_1}$,$\hat{\sigma_2}$ (Esc. 3).

Referencias

[1] R. Costa-Castelló, J. M. Olm, H. Vargas, and G. A. Ramos. An educational approach to the internal model principle for periodic signals. *International Journal of Innovative Computing, Information and Control*, 8(8):5591–5606, 2012.

[2] G. F. Franklin, J. D. Powell, and M. L. Workman. *Digital Control of Dynamic Systems*. Addison Wesley Longman, 3rd. edition, 1997.

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

[4] K. Kashima and D. Inoue. Replay attack detection in control systems with quantized signals. In *Control Conference (ECC), 2015 European*, pages 782–787. IEEE, 2015.

[5] 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.

[6] F. Miao, M. Pajic, and G. J. Pappas. Stochastic game approach for replay attack detection. In *Decision and control (CDC), 2013 IEEE 52nd annual conference on*, pages 1854–1859. IEEE, 2013.

[7] Y. Mo and B. Sinopoli. Secure control against replay attacks. In *Communication, Control, and Computing, 2009. Allerton 2009. 47th Annual Allerton Conference on*, pages 911–918. IEEE, 2009.

[8] F. Pasqualetti, F. Dörfler, and F. Bullo. Attack detection and identification in cyber-physical systems. *IEEE Transactions on Automatic Control*, 58(11):2715–2729, 2013.

[9] H. Sánchez, D. Rotondo, T. Escobet, V. Puig, and J. Quevedo. Frequency-based detection of replay attacks: application to a multiple tank system. *10th IFAC Symposium on Fault Detection, Supervision and Safety for Technical Processes, SAFEPROCESS 2018*.

[10] S. Skogestad and I. Postlethwaite. *Multivariable Feedback Control: Analysis and Design*. Wiley, 2005.

[11] B. Tang, L. D. Alvergue, and G. Gu. Secure networked control systems against replay attacks without injecting authentication noise. In *American Control Conference (ACC), 2015*, pages 6028–6033. IEEE, 2015.

[12] A. Teixeira, D. Pérez, H. Sandberg, and K. H. Johansson. Attack models and scenarios for networked control systems. In *Proceedings of the 1st international conference on High Confidence Networked Systems*, pages 55–64. ACM, 2012.

[13] Q.-G. Wang. *Decoupling Control*. Lecture Notes in Control and Information Sciences, Vol. 285, Springer-Verlag Berlin Heidelberg, 2003.

[14] 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.

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

© 2018 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution CC-BY-NC 3.0 license (http://creativecommons.org/licenses/by-nc/3.0/).