Detection of Covert Attacks on Cyber-Physical Systems using Markovian Jump Systems ⋆

Alexsandra C. Ribeiro * **Diego S. Carneiro** * **Eduardo F. Costa** ** **Vilma A. Oliveira** *

* *Departamento de Engenharia Elétrica e de Computação, Universidade de São Paulo (USP) em São Carlos, SP, Brasil (e-mail: alexsandracarmona@usp.br, diegocarneiro@usp.br, voliveira@usp.br)*
** *Departamento Departamento de Matemática Aplicada e Estatística, Universidade de São Paulo (e-mail: efcosta@icmc.usp.br)*

Abstract: In this paper we propose a strategy to detect covert attacks on cyber-physical systems by extending the original plant with an auxiliary system. This auxiliary system is designed as a Markovian jump system. A detection system composed by a Kalman filter is presented. The efficacy of the proposed method is illustrated by a simulation example.

Keywords: Cyber physical systems; Industrial control systems; Cyber attacks; Covert attack; Markovian jump systems; Markov chains.

1. INTRODUCTION

The advancement of researches in the areas of control, computing and communication has boost the emergence of complex and intelligent technologies. One of them are the so called cyber-physical systems (CPS) which are systems that integrate virtual networking and the physical world.

A class of CPS is industrial control systems (ICS), which are used for measure and control procedures in the manufacturing industry and in critical infrastructures, such as energy and water (Schellenberger and Zhang, 2017). The security and integrity of these systems must be guaranteed, since they are highly distributed and their communication is frequently performed by computer networks, making these systems be more sensitive to invasions by hackers.

Nowadays, the most usual application of CPSs is, undoubtedly, in an industrial environment. Industrial control systems include Supervisory Control and Data Acquisition (SCADA) systems, Distributed Control Systems (DCS), and other configurations such as Programmable Logic Controllers (PLCs).

The first relevant case of a cyber attack to a control system happened on 2010 and it was caused by the Stuxnet virus, whose target were Siemens PLCs. Once a computer was contaminated by the virus, the invaders would get access to the software used to program the PLC. Meanwhile, a Replay Attack was used to hide Stuxnet actions (Hoehn and Zhang, 2016). Since the events caused by Stuxnet, control systems have become an increasingly frequent target of invasions, thus the necessity for developing strategies for detecting cyber attacks on CPS.

It is possible to find in the literature several works that address detection methods for CPS exposed to different types of virtual attacks, such as in Góes et al. (2017), Li et al. (2018), Pessim and Lacerda (2021), Abbadi and Jamouli (2019), Cao et al. (2020) and Zhao et al. (2020).

The most sophisticated attack nowadays in terms of control systems is the covert attack. This type attack requires complete knowledge of the plant as well as full access to input control signals and data collected by sensors transmitted over the network (Hoehn and Zhang, 2016). The idea behind this attack is to exploit the linearity of the plant so that the attack's input and output effects cancel each other out. In other words, the additive attack signal at the input is canceled by calculating its influence on the plant output and the resulting value is, then, subtracted from the sensor readings, which makes this type of attack almost impossible to detect.

One approach to detect covert attacks is by avoiding that the invaders obtain total knowledge of the system dynamics. In Hoehn and Zhang (2016), the authors utilize a time-varying modulation matrix to alter the behavior of the system's input preventing the hacker from having perfect knowledge of the plant and therefore allowing their covert attack to be detectable. The strategy proposed in Schellenberger and Zhang (2017) is based on implementing a switched auxiliary system as an extension of the original plant, whose dynamics changes at random time instants to prevent the invader from obtaining a perfect model of the auxiliary system.

In this work, we use a Markovian jump system (MJS) as an auxiliary system to the original plant. Markov chains are discrete stochastic processes that describe the evolution of memoryless random dynamical systems, that is, the future of the process, once the present state is known, is independent of the past (Oliveira et al., 2017). Markovian jump systems can be applied to several different

⋆ This research was supported by the Brazilian National Research Council (CNPq) under grants 311959/2021-0 and by The São Paulo Research Foundation (FAPESP) under grants 2014/50851-0, 2016/21120-2 and 2019/25530-9.

situations in real life such as economic systems (Svensson and Williams, 2009), aeronautics (Gray et al., 2000) and also power systems (Li and Ugrinovskii, 2006).

A Markov chain with time-homogeneous transition probabilities is responsible for changing the mode of operation of the auxiliary system at each time instant. The addition of the MJS does not interfere with the original process dynamics. By doing this, an invader will not be able to appropriately assess the system to obtain an accurate plant model and generate attack signals. Since the hacker can no longer calculate the attack signals that cancel each other effects, these signals will then generate detectable disturbances on the auxiliary system outputs. A Kalman filter for a MJS is used to calculate the signal estimate. Deviations on the auxiliary system estimate indicate an attack.

This paper's main contribution is the use of a well known class of stochastic systems as an auxiliary system to the plant which allows for a more structured approach. The MJS constant switching provides an uninterrupted attack detection method. The controller is specifically designed for the MJS, but the standard Kalman filter for time-varying systems can be used here, as the controller has access to the variable θ_k at time k. One interpretation is that, for previous time instants $k-1, k-2, \ldots, 0$, the MJS is "equivalent" to a time-varying system from the perspective of the controller.

2. PRELIMINARIES

2.1 Cyber-Physical Systems

Cyber-physical systems (CPSs) are networked intelligent systems embedded with sensors, controllers and actuators which are designed to interact with the physical world and human users with the aim of supporting real-time management and guaranteed performance in safety critical applications (Sun et al., 2018). These systems have several applications such as urban water cycle management (Sun et al., 2018), surgeries performed by robots (D'Auria and Persia, 2017), aircraft fuel management system (Sun et al., 2014), more interactive and realistic video games (Wu et al., 2010) and production lines (Herwan et al., 2018). A CPS can be structured as shown in Figure 1.

This paper considers that the attack happens through the network and in a subsystem of a large ICS, such as a SCADA system. A plant assumed to be a linear discrete-time system can be mathematically represented as follows:

$$\begin{aligned} x_{k+1} &= Ax_k + Bu_k + Dw_k \\ y_k &= Cx_k + Ev_k \end{aligned} \tag{1}$$

for all $k = 0, 1, ...$, where $x_k \in \mathbb{R}^n$ is the state vector, $u_k \in \mathbb{R}^{m1}$ is the control input, $y_k \in \mathbb{R}^p$ is the measure vector (observations), $w_k \in \mathbb{R}^{m2}$ e $v_k \in \mathbb{R}^t$ are random Gaussian noise with zero mean and A, B, C, D and E are matrices of compatible dimensions.

2.2 Markov chains

In terms of mathematical expression, consider that a Markov chain is a stochastic process $(\theta_k)_{k\in\mathbb{N}}$ associated with a probability space where its conditional probability,

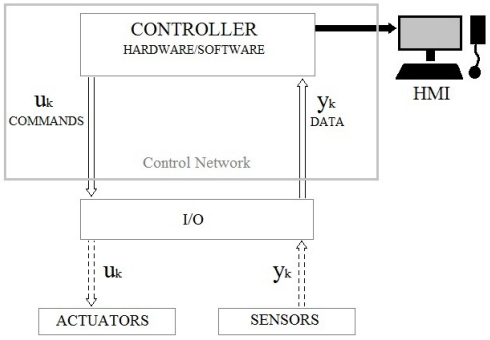

Figure 1. Basic architecture of a CPS. The dotted arrows in the figure represent analog or fieldbus connections and the double solid arrows represent the connections in the control network, which are communication protocols like Modbus or DNP3, for example. The black solid line is for video/audio transmission as the human-machine-interface (HMI) exists so one may visualize some parts of the ICS software as well as manually control a few chosen variables.

for all $k = 0, 1, ...$ and for all sequences $i_0, i_1, ..., i_{n+1}$ of elements in a finite set $\mathcal{S} = \{1, 2, ..., s\}$ satisfies:

$$\begin{aligned} &Prob(\theta_{k+1} = i_{k+1} | \theta_0 = i_0, ..., \theta_k = i_k] = \\ &Prob[\theta_{k+1} = i_{k+1} | \theta_k = i_k). \end{aligned} \tag{2}$$

where $Prob(\cdot|\cdot)$ denotes the conditional probability function. We consider a time-homogeneous Markov chain with a transition probability matrix $\mathbb{P}$ defined by

$$\text{Prob}[\theta_{k+1} = j | \theta_k = i] = [\mathbb{P}]_{i,j} \tag{3}$$

where $[\mathbb{P}]_{i,j}$ is the element at the i-th row and j-th column of matrix $\mathbb{P}$.

2.3 Markovian jump systems

A discrete linear MJS is a special class of hybrid and stochastic systems that exhibits a parameter-switching behaviour and is modeled by a set of linear or nonlinear systems with the transitions between the models determined by a Markov chain taking values in a finite set Shi and Li (2015). The MJS can be mathematically represented as:

$$\begin{aligned} x_{k+1} &= A_{\theta_k,k}x_k + B_{\theta_k,k}u_k + D_{\theta_k,k}w_k \\ y_k &= C_{\theta_k,k}x_k + E_{\theta_k,k}v_k, \end{aligned} \tag{4}$$

for all $k = 0, 1, ...$, where $x_k \in \mathbb{R}^n$ is the state vector, $u_k \in \mathbb{R}^{m_1}$ is the control input, $y_k \in \mathbb{R}^p$ is the measure vector (observations), $w_k \in \mathbb{R}^{m_2}$ and $v_k \in \mathbb{R}^t$ are random Gaussian noise with zero mean. The parameter matrices $A_{\theta_k,k} \in \mathbb{R}^{n\times n}$, $B_{\theta_k,k} \in \mathbb{R}^{n\times m_1}$, $C_{\theta_k,k} \in \mathbb{R}^{p\times n}$, $D_{\theta_k,k} \in \mathbb{R}^{n\times m_2}$, $E_{\theta_k,k} \in \mathbb{R}^{p\times t}$ are known for each (θ_k, k), where $\{\theta_k\}$ is a discrete-time finite-state Markov chain with transition probabilities matrix $\mathbb{P} = [p_{ij}] \in \mathbb{R}^{s\times s}$ whose entries satisfy:

$$\begin{aligned} &Prob[\theta_{k+1} = j | \theta_k = i] = [p_{ij}] \\ &\sum_{j=1}^{s} p_{ij} = 1,\ 0 \le p_{ij} \le 1. \end{aligned} \tag{5}$$

It is assumed that x_0, w_k e v_k are mutually independent, the initial operation modes θ_0 and the initial state x_0 are usually known and y_k and θ_k are observed at every instant of time k.

2.4 Design of the auxiliary system

The auxiliary system is designed as a linear MJS with $s \geq 3$ operation modes, one of them being the original plant. Consequently, the other new operation modes have to be calculated. The auxiliary system should be stable and mimic the behaviour of the plant.

A discrete-time, linear MJL is said to be stochastically stable if $\sum_{k=0}^{\infty} E[\|x_k\|^2] < \infty$ when u_k, w_k and v_k are all equal to zero, where $E[\|.\|]$ denotes the expected value (Costa et al., 2005, Theorem 9). As a testable condition for stability, one can check if there exists symmetric, positive definite matrices $V_1, \ldots, V_s$ such that $V_j > \sum_{i\in\mathcal{S}} p_{ij} A_i V_i V_i', \quad \forall i \in \mathcal{S}$.

The matrices $A_{\theta,aux}$, $B_{\theta,aux}$ and $C_{\theta,aux}$ should have similar elements, i.e. elements with similar numerical values, as A_{sys}, B_{sys} and C_{sys}. Here we choose the matrices of the auxiliary system with the same dimensions as their counterparts in the original plant. In addition, $A_{\theta,aux}$ and A_{sys} should have eigenvalues in the same range. To generate the elements of the matrices, a truncated normal distribution (TND) as given in Schellenberger and Zhang (2017) is used.

2.5 System controller and detection system

The controller used in this work was obtained using a linear quadratic regulator for nominal MJS algorithm developed by Cerri (2013). The optimal control law is determined by minimizing the expected value of a quadratic cost function constrained to the dynamic model in a finite horizon. The solution is recursive, given by a set of coupled Ricatti equations and the steps are presented in Algorithm 1, where $Q_k \in \mathbb{R}^{m_2\times m_2}$ and $R_k \in \mathbb{R}^{t\times t}$ are positive definite

Algorithm 1 Linear quadratic regulator for nominal MJSs

Initial Conditions: Define x_0, θ_0, $\mathbb{P}$, and $P_{i,N} \geq 0, \forall_i \in \Theta$.

Step 1: *(Backwards)* Calculate for all $k = N-1, ..., 0$:

$$\begin{aligned}
\Psi_{i,k+1} &:= \sum_{j=1}^{s} P_{j,k+1} p_{ij}, \\
L_{i,k} &= A_{i,k} + B_{i,k} K_{i,k}, \\
K_{i,k} &= -\left(R_{i,k} + B_{i,k}^T \Psi_{i,k+1} B_{i,k}\right)^{-1} B_{i,k}^T \Psi_{i,k+1} A_{i,k}, \\
P_{i,k} &= A_{i,k}^T \big(\Psi_{i,k+1} - \Psi_{i,k+1} B_{i,k} \big(R_{i,k} \\
&\quad + B_{i,k}^T \Psi_{i,k+1} B_{i,k}\big)^{-1} G_{i,k}^T \Psi_{i,k+1}\big) A_{i,k} + Q_{i,k}.
\end{aligned}$$

Step 2: *(Forward)* Calculate for all $k = 0, ..., N-1$:

$$\begin{bmatrix} x_{k+1}^* \\ u_k^* \end{bmatrix} = \begin{bmatrix} L_{\theta_k,k} \\ K_{\theta_k,k} \end{bmatrix} x_k^*.$$

weighting matrices, respectively. It is important to mention that, from the point of view of the attack designer, knowledge of the controller is not necessary for the success of the attack Smith (2011). The detection system proposed in this paper consists on analysing the difference between the real output of the auxiliary system and an estimation given by a Kalman Filter for MJS developed in Cerri (2013), which the structure is is shown in Algorithm 2. The presence of significant deviation indicates an attack. State estimates are obtained from the minimization of quadratic functionals using dynamic programming. The procedure consists of combining the solution of weighted least squares and penalty functions.

Algorithm 2 Optimal state estimate in predictive and filtered forms

Initial conditions: $P_{0|-1} > 0$, $\hat{x}_{0|-1} = 0$.

Step $k \geq 0$: Calculate for all $\{\hat{x}_{k+1|k}; P_{k+1|k}\}$ and $\{\hat{x}_{k|k}; P_{k|k}\}$ according to:

$$\begin{aligned}
\hat{x}_{k|k} &= \hat{x}_{k|k-1} + P_{k|k-1} C_{\theta_k,k}^T \big(R_k \\
&\quad + C_{\theta_k,k} P_{k|k-1} C_{\theta_k,k}^T\big)^{-1} \big(y_k - C_{\theta_k,k} \hat{x}_{k|k-1}\big), \\
P_{k|k} &= P_{k|k-1} - P_{k|k-1} C_{\theta_k,k}^T \big(R_k \\
&\quad + C_{\theta_k,k} P_{k|k-1} C_{\theta_k,k}^T\big)^{-1} C_{\theta_k,k} P_{k|k-1} \\
\hat{x}_{k+1|k} &= A_{\theta_k,k} \hat{x}_{k|k} + B_{\theta_k,k} u_{k|k}, \\
P_{k+1|k} &= Q_k + A_{\theta_k,k} P_{k|k} A_{\theta_k,k}^T.
\end{aligned}$$

3. MAIN RESULTS

The system architecture is structured with an auxiliary system as shown in Figure 2 including the connections between the controller and the detection system. The system dynamics is given as follows:

$$\begin{aligned}
\begin{bmatrix} x_{sys,k+1} \\ x_{aux,k+1} \end{bmatrix} &= \begin{bmatrix} A_{sys} & 0 \\ 0 & A_{\theta_k,k,aux} \end{bmatrix} \begin{bmatrix} x_{sys,k} \\ x_{aux,k} \end{bmatrix} \\
&+ \begin{bmatrix} B_{sys} & 0 \\ 0 & B_{\theta_k,k,aux} \end{bmatrix} \begin{bmatrix} u_{sys,k} \\ u_{aux,k} \end{bmatrix} \\
&+ \begin{bmatrix} D_{sys} & 0 \\ 0 & D_{\theta_k,k,aux} \end{bmatrix} \begin{bmatrix} w_{sys,k} \\ w_{aux,k} \end{bmatrix}
\end{aligned} \tag{6}$$

$$\begin{aligned}
\begin{bmatrix} y_{sys,k} \\ y_{aux,k} \end{bmatrix} &= \begin{bmatrix} C_{sys} & 0 \\ 0 & C_{\theta_k,k,aux} \end{bmatrix} \begin{bmatrix} x_{sys,k} \\ x_{aux,k} \end{bmatrix} \\
&+ \begin{bmatrix} E_{sys} & 0 \\ 0 & E_{\theta_k,k,aux} \end{bmatrix} \begin{bmatrix} v_{sys,k} \\ v_{aux,k} \end{bmatrix}
\end{aligned} \tag{7}$$

where $[A_{sys}, B_{sys}, C_{sys}, D_{sys}]$ is the original plant and $[A_{\theta_k,k,aux}, B_{\theta_k,k,aux}, C_{\theta_k,k,aux}, D_{\theta_k,k,aux}]$ is the auxiliary MJS.

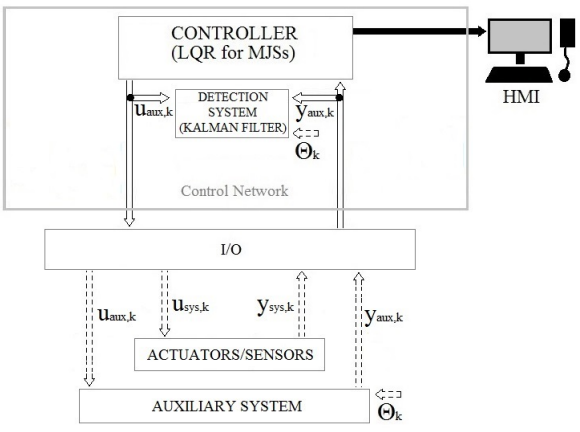

Figure 2. System architecture for the detection method proposed.

A covert attack applied to the system considered in this work can be represented as shown in Figure 3. The attack model used here is the one presented in Smith (2011) where the hacker can only measure and add to existing control or measurement signals and the calculation of the attack signals is done through a reference tracking controller. The hacker is considered to have full knowledge of the system including the information being transmitted via control network except for the Markov chain θ_k. Attack signals are obtained from the following equations:

$$\begin{bmatrix} x_{a,sys,k+1} \\ x_{a,aux,k+1} \end{bmatrix} = \begin{bmatrix} A_{a,sys} & 0 \\ 0 & A_{a,aux} \end{bmatrix} \begin{bmatrix} x_{a,sys,k} \\ x_{a,aux,k} \end{bmatrix} + \begin{bmatrix} B_{a,sys} & 0 \\ 0 & B_{a,aux} \end{bmatrix} \begin{bmatrix} u_{a,sys,k} \\ u_{a,aux,k} \end{bmatrix} \tag{8}$$

$$\begin{bmatrix} y_{a,sys,k} \\ y_{a,aux,k} \end{bmatrix} = \begin{bmatrix} C_{a,sys} & 0 \\ 0 & C_{a,aux} \end{bmatrix} \begin{bmatrix} x_{a,sys,k} \\ x_{a,aux,k} \end{bmatrix} \tag{9}$$

The matrices $A_{a,aux}$, $B_{a,aux}$ and $C_{a,aux}$ are the attacker estimate of $A_{\theta,aux}$, $B_{\theta,aux}$ and $C_{\theta,aux}$, respectively. Therefore, the dynamics of the augmented system under attack are:

$$\begin{bmatrix} x_{sys,k+1} \\ x_{aux,k+1} \end{bmatrix} = \begin{bmatrix} A_{sys} & 0 \\ 0 & A_{\theta_k,k,aux} \end{bmatrix} \begin{bmatrix} x_{sys,k} \\ x_{aux,k} \end{bmatrix} + \begin{bmatrix} B_{sys} & 0 \\ 0 & B_{\theta_k,k,aux} \end{bmatrix} \begin{bmatrix} u_{sys,k} + u_{a,sys,k} \\ u_{aux,k} + u_{a,aux,k} \end{bmatrix} + \begin{bmatrix} D_{sys} & 0 \\ 0 & D_{\theta_k,k,aux} \end{bmatrix} \begin{bmatrix} w_{sys,k} \\ w_{aux,k} \end{bmatrix} \tag{10}$$

$$\begin{bmatrix} y^*_{sys,k} \\ y^*_{aux,k} \end{bmatrix} = \begin{bmatrix} C_{sys} & 0 \\ 0 & C_{\theta_k,k,aux} \end{bmatrix} \begin{bmatrix} x_{sys,k} \\ x_{aux,k} \end{bmatrix} - \begin{bmatrix} y_{a,sys,k} \\ y_{a,aux,k} \end{bmatrix} + \begin{bmatrix} E_{sys} & 0 \\ 0 & E_{\theta_k,k,aux} \end{bmatrix} \begin{bmatrix} v_{sys,k} \\ v_{aux,k} \end{bmatrix} \tag{11}$$

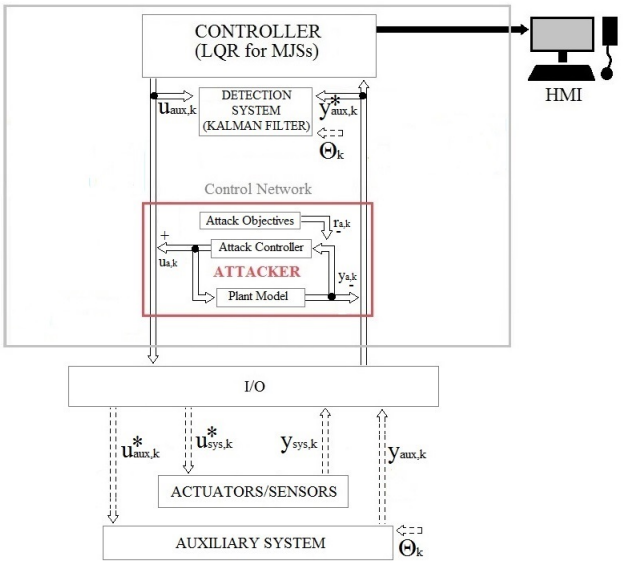

Figure 3. System under the action of a covert attack with $u^*_{aux,k} = u_{aux,k} + u_{a,aux,k}$, $u^*_{sys,k} = u_{sys,k} + u_{a,sys,k}$.

Since the attack is undetectable with the original plant estimate, only the auxiliary system dynamics is used in the detection system. Assuming that all operation modes of the auxiliary system are observable and asymptotically stable, for the augmented system under attack, the auxiliary system estate estimate is given as follows

$$\begin{aligned} \hat{x}_{aux,k+1|k} &= A_{\theta_k,k,aux}\hat{x}_{aux,k|k} + B_{\theta_k,k,aux}(u_{aux,k} + u_{a,aux,k}) \\ &= A_{\theta_k,k,aux}\big(\hat{x}_{aux,k|k-1} \\ &\quad + P_{k|k-1,aux}C^T_{\theta_k,k,aux}\big(R_{k,aux} \\ &\quad + C_{\theta_k,k,aux}P_{k|k-1,aux}C^T_{\theta_k,k,aux}\big)^{-1}\big(y^*_{aux,k} \\ &\quad - C_{\theta_k,k,aux}\hat{x}_{aux,k|k-1}\big)\big) \\ &\quad + B_{\theta_k,k,aux}(u_{aux,k} + u_{a,aux,k}) \\ \hat{y}_{aux,k} &= C_{\theta_k,k,aux}\hat{x}_{aux,k|k-1} \end{aligned} \tag{12}$$

where

$$\begin{aligned} y^*_{aux,k} &= y_{aux,k} - y_{a,aux,k} \\ &= C_{\theta_k,k,aux}x_{aux,k} - C_{a,aux}x_{a,aux,k} \\ &= C_{\theta_k,k,aux}\big(A_{\theta_k,k-1,aux}x_{aux,k-1} \\ &\quad + B_{\theta_k,k-1,aux}\big(u_{aux,k-1} + u_{a,aux,k-1}\big) \\ &\quad + D_{\theta_k,k-1,aux}w_{aux,k-1}\big) \\ &\quad - C_{a,aux}\big(A_{a,aux}x_{a,aux,k-1} + B_{a,aux}u_{a,aux,k-1}\big). \end{aligned} \tag{13}$$

As the operation of the Markovian jump auxiliary system starts, we have $C_{a,aux} \neq C_{\theta_k,k,aux}$ and $B_{a,aux} \neq B_{\theta_k,k-1,aux}$ and the disturbance generated by the attack on $\hat{x}_{aux,k+1|k}$ is

$$d_{aux,k} = y^*_{aux,k} - \hat{y}_{aux,k}. \tag{14}$$

Therefore, a decision logic signal σ for the attack detection is defined as follows

$$\sigma = \begin{cases} 1, & \text{if } \|d_{aux,k}\| > J_{th} \\ 0, & \text{if } \|d_{aux,k}\| \leq J_{th} \end{cases} \tag{15}$$

where $\sigma = 1$ indicates that the system is under attack and $\sigma = 0$ indicates that the system is attack free with a threshold J_{th} taken as the estimate of the maximum fluctuation of the disturbance in the attack free case, that is, $J_{th} = \max_{u_a=0, y_a=0} \|d_{aux,k}\|$. The symbol $\|\cdot\|$ denotes the euclidean norm.

4. SIMULATION RESULTS

In order to show the efficacy of the detection method, the attack results on the auxiliary system state estimate are given. An adaptation of a simple numerical example presented in Do Val et al. (2003) is used, with the matrices of the plant and the auxiliary systems:

$$\mathbf{A_{sys}} = \begin{bmatrix} 2 & 1 \\ -2.5 & 3.2 \end{bmatrix}, \ \mathbf{B_{sys}} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \ \mathbf{C_{sys}} = [\,0.5 \ 1\,], \ \mathbf{D_{sys}} = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 0.5 \end{bmatrix}, \ \mathbf{E_{sys}} = [\,0.1\,]$$

$$\mathbf{A_{1,aux}} = \begin{bmatrix} 2 & 1 \\ -4.3 & 4.5 \end{bmatrix}, \ \mathbf{B_{1,aux}} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \ \mathbf{C_{1,aux}} = \begin{bmatrix} 0.4 \\ 0.9 \end{bmatrix}^T, \ \mathbf{D_{1,aux}} = \begin{bmatrix} 0.5 & 2.5 \\ 1.5 & 0.5 \end{bmatrix}, \ \mathbf{E_{1,aux}} = [\,-0.5\,]$$

$$\mathbf{A_{2,aux}} = \begin{bmatrix} 1 & 1 \\ 5.3 & 5.2 \end{bmatrix}, \mathbf{B_{2,aux}} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \mathbf{C_{2,aux}} = \begin{bmatrix} 0.45 \\ 0.95 \end{bmatrix}^T,$$
$$\mathbf{D_{2,aux}} = \begin{bmatrix} 0.75 & 1 \\ 1 & 0.5 \end{bmatrix}, \mathbf{E_{2,aux}} = [\,0.3\,].$$
$$\mathbf{A_{3,aux}} = \begin{bmatrix} 2 & 1 \\ -2.5 & 3.2 \end{bmatrix}, \mathbf{B_{3,aux}} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \mathbf{C_{3,aux}} = [\,0.5\ 1\,],$$
$$\mathbf{D_{3,aux}} = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 0.5 \end{bmatrix}, \mathbf{E_{3,aux}} = [\,0.1\,].$$

Note that mode 3 corresponds to the actual plant model. The attacker is assumed to have perfect system knowledge at $k = 0$. The attack starts at $k = 50$ and ends at $k = 200$, its input signal u_a is calculated as:

$$\begin{bmatrix} u_{a,sys,k} \\ u_{a,aux,k} \end{bmatrix} = \begin{bmatrix} K_{a,1} & 0 \\ 0 & K_{a,1} \end{bmatrix} \begin{bmatrix} x_{sys,k} \\ x_{aux,k} \end{bmatrix} + \begin{bmatrix} K_{a,2} & 0 \\ 0 & K_{a,2} \end{bmatrix} \begin{bmatrix} r_{sys,k} \\ r_{aux,k} \end{bmatrix} \tag{16}$$

where $K_{a,1}$ and $K_{a,2}$ are the gain matrices of the attacker reference tracking controller and $r_{sys,k}$ and $r_{aux,k}$ are the reference signals bounded by 0.001. The initial conditions selected are $\theta_0 = 3, x_{sys,0} = x_{aux,0} = [\,1\ 0.5\,], x_{a,sys,0} = x_{a,aux,0} = [\,0.001\ 0.0005\,]$. The threshold is $J_{th} = 0.001$.

The auxiliary system switches only between operation modes 1 and 2. The transition probability matrix and the expression for the next operation modes θ_{k+1} are :

$$\mathbb{P} = \begin{bmatrix} 0.77 & 0.23 \\ 0.36 & 0.64 \end{bmatrix} \tag{17}$$

$$\theta_{k+1} = 1 \times I(0 \leq R_k \leq \mathbb{S}_{\theta_k,1}) + 2 \times I(\mathbb{S}_{\theta_k,1} < R_k \leq 1) \tag{18}$$

where R_k is a random number $0 \leq R_k \leq 1$ generated at instant k, the seed of R_k is the same for the auxiliary system and the detection system, $\mathbb{S}$ is the matrix of cumulative probabilities obtained from $\mathbb{P}$ and I is the indicator function defined by

$$I(c) = \begin{cases} 1, & \text{if } c \text{ is true} \\ 0, & \text{if } c \text{ is false} \end{cases} \tag{19}$$

The attack results are shown in Figure 4. At $k = 100$ a Markov chain starts running and it is possible to detect the influence of the attack on the auxiliary system, while the effect on the original plant remains stealthy. The operation modes of the Markov chain is shown in Figure 5. The disturbance generated by the attack on the auxiliary system state estimate $d_{aux,k}$ is shown and Figure 6. It is possible to see that, even if the hacker has full information of the plant parameters, the attack is still detected since the disturbance is greater than the established threshold.

5. CONCLUSION

This work presented a method to successfully detect covert attacks in discrete-time linear systems by using an observer-based MJS structure. For a hacker with an initially perfect system model, the attack detection is only possible after the Markovian jump auxiliary system starts switching. On the other hand, if the operation modes start switching at $k = 0$, it is less likely that the attacker will be able to assess a viable system model due to the fact that the system dynamics is always changing and the hacker can not predict the behaviour of the system with exactitude.

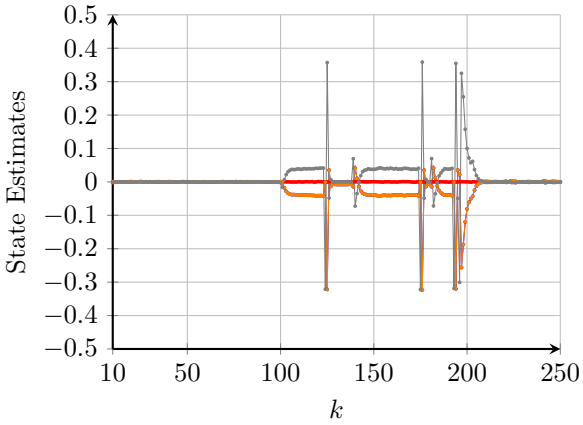

Figure 4. State estimate for the original plant and the auxiliary system under the presence of a covert attack. The state estimate for the original plant in red and blue lines (blue lines barely visible behind the red line) do not show any signs of the covert attack, neither do the auxiliary system state estimate (orange and grey lines) before the operation modes start switching.

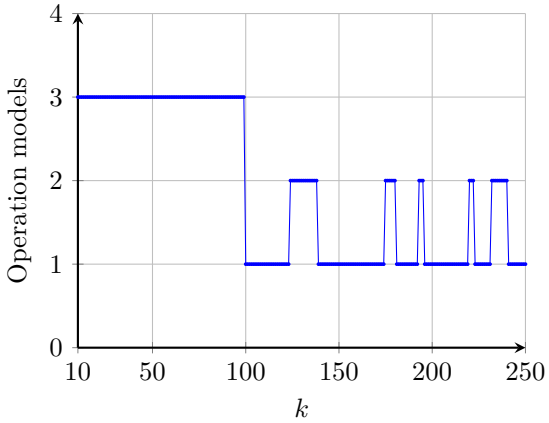

Figure 5. Switching of the operation modes of a Markov chain realization with $\theta_0 = 3$.

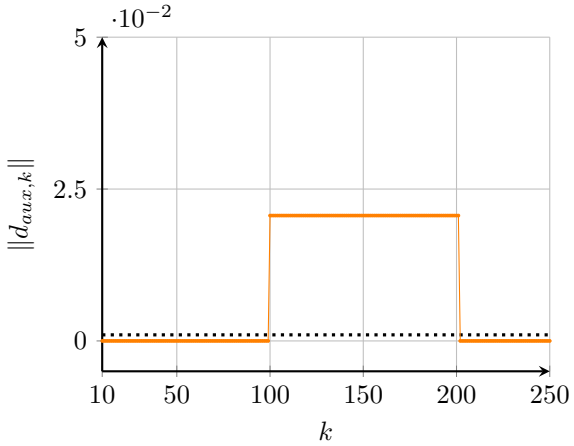

Figure 6. Norm of the disturbance on the auxiliary system $d_{aux,k}$ obtained by (14).

REFERENCES

Abbadi, R.E. and Jamouli, H. (2019). Stabilization of cyber physical system exposed to a random replay attack

modeled by markov chains. In *2019 6th International Conference on Control, Decision and Information Technologies (CoDIT)*, 528–533. IEEE, Paris, France.

Cao, Z., Niu, Y., and Song, J. (2020). Finite-time sliding-mode control of Markovian jump cyber-physical systems against randomly occurring injection attacks. *IEEE Transactions on Automatic Control*, 65(3), 1264–1271.

Cerri, J.P. (2013). *Controle e Filtragem para Sistemas Lineares Discretos Incertos sujeitos a Saltos Markovianos*. Ph.D. thesis, Universidade de São Paulo.

Costa, O.L.V., Fragoso, M.D., and Marques, R.P. (2005). *Discrete-Time Markov Jump Linear Systems*. Springer, New York, NY, USA.

D'Auria, D. and Persia, F. (2017). A collaborative robotic cyber physical system for surgery applications. In *2017 IEEE International Conference on Information Reuse and Integration (IRI)*, 79–83. IEEE, San Diego, CA.

Do Val, J.B., Nespoli, C., and Cáceres, Y.R. (2003). Stochastic stability for Markovian jump linear systems associated with a finite number of jump times. *Journal of Mathematical Analysis and Applications*, 285(2), 551–563.

Gray, W.S., Gonzalez, O.R., and Dogan, M. (2000). Stability analysis of digital linear flight controllers subject to electromagnetic disturbances. *IEEE Transactions on Aerospace and Electronic Systems*, 36(4), 1204–1218.

Góes, R.M., Kang, E., Kwong, R., and Lafortune, S. (2017). Stealthy deception attacks for cyber-physical systems. In *2017 IEEE 56th Annual Conference on Decision and Control (CDC)*, 4224–4230. IEEE, Melbourne, Victoria, Australia.

Herwan, J., Kano, S., Oleg, R., Sawada, H., and Kasashima, N. (2018). Cyber-physical system architecture for machining production line. In *2018 IEEE Industrial Cyber-Physical Systems (ICPS)*, 387–391. IEEE, St. Petersburg.

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, Boston, MA, USA.

Li, H., He, X., Zhang, Y., and Guan, W. (2018). Attack detection in cyber-physical systems using particle filter: An illustration on three-tank system. In *2018 IEEE 8th Annual International Conference on Cyber Technology in Automation, Control, and Intelligent Systems (CYBER)*, 504–509. IEEE, Tianjin, China. doi:10.1109/CYBER.2018.8688281.

Li, L. and Ugrinovskii, V.A. (2006). Decentralized robust control of uncertain Markov jump parameter systems via output feedback. In *2006 American Control Conference*, 6 pp.–. IEEE, Minneapolis, MN, USA. doi:10.1109/ACC.2006.1656403.

Oliveira, A., Ribeiro, T., and Borges da Silva, F. (2017). Cadeia de Markov: modelo probabilístico e convergência das distribuições de probabilidade. *CQD Revista Eletrônica Paulista de Matemática*, 11ic, 49–61. doi:10.21167/cqdvol11ic201723169664aslotsgrfbs4961.

Pessim, P.S.P. and Lacerda, M.J. (2021). State-feedback control for cyber-physical lpv systems under dos attacks. *IEEE Control Systems Letters*, 5(3), 1043–1048.

Schellenberger, C. and Zhang, P. (2017). Detection of covert attacks on cyber-physical systems by extending the system dynamics with an auxiliary system. In *2017 IEEE 56th Annual Conference on Decision and Control (CDC)*, 1374–1379. IEEE, Melbourne, Victoria.

Shi, P. and Li, F. (2015). A survey on Markovian jump systems: Modeling and design. *Int. J. Control Autom. Syst.*, 13, 1–16.

Smith, R.S. (2011). A decoupled feedback structure for covertly appropriating networked control systems. *IFAC Proceedings Volumes*, 44(1), 90–95. 18th IFAC World Congress.

Sun, C., Cembrano, G., Puig, V., and Meseguer, J. (2018). Cyber-physical systems for real-time management in the urban water cycle. In *2018 International Workshop on Cyber-physical Systems for Smart Water Networks (CySWater)*, 5–8. IEEE, Porto.

Sun, X., Huang, N., Wang, B., and Zhou, J. (2014). Reliability of cyber physical systems assessment of the aircraft fuel management system. In *The 4th Annual IEEE International Conference on Cyber Technology in Automation, Control and Intelligent*, 424–428. IEEE, Hong Kong, China.

Svensson, L.E. and Williams, N. (2009). Optimal Monetary Policy under Uncertainty in DSGE Models: A Markov Jump-Linear-Quadratic Approach. In K. Schmidt-Hebbel, C.E. Walsh, N.L.S. Editor), and K.S.H. (Series (eds.), *Monetary Policy under Uncertainty and Learning*, volume 13 of *Central Banking, Analysis, and Economic Policies Book Series*, chapter 3, 077–114. Central Bank of Chile, Chile.

Wu, C.H., Chang, Y.T., and Tseng, Y. (2010). Multi-screen cyber-physical video game: An integration with body-area inertial sensor networks. In *2010 8th IEEE International Conference on Pervasive Computing and Communications Workshops (PERCOM Workshops)*, 832–834. IEEE, Mannheim, Germany.

Zhao, D., Wang, Z., Wei, G., and Han, Q.L. (2020). A dynamic event-triggered approach to observer-based pid security control subject to deception attacks. *Automatica*, 120, 109128.