Computers, Materials & Continua

Tech Science Press

DOI: 10.32604/cmc.2023.039466

ARTICLE

Stochastic Models to Mitigate Sparse Sensor Attacks in Continuous-Time Non-Linear Cyber-Physical Systems

Borja Bordel Sánchez[1,*], **Ramón Alcarria**[2] **and Tomás Robles**[1]

[1]Information Technologies Department, Universidad Politécnica de Madrid, Madrid, 28031, Spain

[2]Geospatial Information Department, Universidad Politécnica de Madrid, Madrid, 28031, Spain

*Corresponding Author: Borja Bordel Sánchez. Email: borja.bordel@upm.es

Received: 31 January 2023 Accepted: 14 June 2023 Published: 08 October 2023

ABSTRACT

Cyber-Physical Systems are very vulnerable to sparse sensor attacks. But current protection mechanisms employ linear and deterministic models which cannot detect attacks precisely. Therefore, in this paper, we propose a new non-linear generalized model to describe Cyber-Physical Systems. This model includes unknown multivariable discrete and continuous-time functions and different multiplicative noises to represent the evolution of physical processes and random effects in the physical and computational worlds. Besides, the digitalization stage in hardware devices is represented too. Attackers and most critical sparse sensor attacks are described through a stochastic process. The reconstruction and protection mechanisms are based on a weighted stochastic model. Error probability in data samples is estimated through different indicators commonly employed in non-linear dynamics (such as the Fourier transform, first-return maps, or the probability density function). A decision algorithm calculates the final reconstructed value considering the previous error probability. An experimental validation based on simulation tools and real deployments is also carried out. Both, the new technology performance and scalability are studied. Results prove that the proposed solution protects Cyber-Physical Systems against up to 92% of attacks and perturbations, with a computational delay below 2.5 s. The proposed model shows a linear complexity, as recursive or iterative structures are not employed, just algebraic and probabilistic functions. In conclusion, the new model and reconstruction mechanism can protect successfully Cyber-Physical Systems against sparse sensor attacks, even in dense or pervasive deployments and scenarios.

KEYWORDS

Cyber-physical systems; sparse sensor attack; non-linear models; stochastic models; security

1 Introduction

Cyber-Physical Systems (CPS) are seamless integrations of physical and computational processes [1]. Many different architectures and approaches to support these unions have been reported, from schemes based on the control theory [2] to feedback loops in computational systems [3]. But all proposed CPS implementations include a sensing platform to monitor the evolution of the physical world [1]. That platform is dense, including thousands of networked sensor nodes capable of capturing information through several different physical parameters [4]. Those data must be injected into computational processes, to ensure that the cybernetic and physical worlds evolve together in a feedback control loop [5].

Therefore, precise information about physical processes is essential to ensure that the loop is convergent and follows the expected evolution [6]. However, it is hardly possible to obtain precise information in real applications [4]. Many random effects have an impact on the behavior of CPS, such as noise, transmission errors, measurement, digitalization, and discretization processes [4]. In that way, information finally injected into computational processes is not the raw or authentic information acquired from the physical world, but a non-deterministic transformation of it. And this transformed information prevents the CPS from integrating the computational and physical processes with the expected synchronicity and showing the required behavior [7].

Furthermore, as Cyber-Physical Systems are used in more scenarios and applications, including critical infrastructures, they are more exposed to new risks. Eventual and unexpected cyberattacks are the main ones. Although innovative attack strategies have been reported to exploit specific vulnerabilities of CPS [8], nowadays the greatest risks for CPS are still associated with classic cyberattacks such as the Sparse Sensor Attack (SSA). In the SSA [9], attackers introduce false information and/or cause delays in the sensing platform monitoring the physical world at a low level, so the CPS behavior is altered or denied. It is the most common attack in control solutions, and new uncertainty about the physical information injected into the computational processes is to be handled.

In this context, reconstruction mechanisms to recover the original and real information extracted from the physical world are essential [10]. The state of any CPS may be described as a multidimensional vector, where each position represents a physical parameter. By establishing the analytical law that describes the trajectory of all those state variables in the phase space, the transformed information received may be corrected through a theoretically predicted CPS state [11]. However, in the general case, all physical parameters are not independent, but they are interrelated through complex physical laws [12]. The appearance of complex non-linear laws, together with the need for stochastic terms to describe random effects such as sparse sensor cyberattacks, turns quite difficult to find a general high-precision model. Thus, traditional reconstruction schemes are based on some basic assumptions, so the mathematical expressions describing the evolution of CPS are easier to manipulate and implement [13].

Our work is motivated by limitations and vulnerabilities caused by these simple assumptions, which make CPS weaker against cyberattacks than other state-of-the-art technological systems. Namely:

- First, Cyber-Physical Systems are assumed to evolve according to a linear law.
- Second, all terms are considered deterministic, including noise and attacking signals.
- Third, all physical variables are assumed to be fully independent of each other.
- And fourth, physical processes are assumed to be discrete, so digitalization and transmission processes do not have to be explicitly considered. Although those linear deterministic models present important advantages (for example, they can be manipulated to find analytical expressions for the detection and identification of SSA), their applicability is very limited [14].
- Only closed CPS based on a reduced number of physical variables with a smooth and invariant behavior (such as the temperature in a climatized space) are governed and can be secured and protected by such a simple model.

Therefore, more complex and general models are required to protect and mitigate SSA in multidimensional CPS with a continuous-time non-linear behavior. In this paper, we address this challenge.

Three innovative contributions are introduced in this paper:

- A complex non-linear model to describe the CPS behavior in a general situation.
- New signals and models for SSA and digitalization processes.
- The third and final contribution is an innovative reconstruction scheme.

The proposed model describes the behavior of CPS using unknown generic functions, which are developed as Taylor series. This model also includes stochastic terms to represent physical, transmission, and measurement noises. Besides, SS attacks are described as a new signal whose value follows a probabilistic behavior according to a given discrete random variable. Physical processes are represented by continuous-time signals that are discretized using an event-based scheme. The resulting multidimensional model injects discrete data into computational processes, but is too complex to generate analytic expressions to mitigate SSA in CPS. Finally, the proposed reconstruction scheme is supported by a weighted stochastic model where the error probability is estimated through different indicators commonly employed to describe non-linear dynamics (such as the Fourier transform, first-return maps, or the probability density function). A decision algorithm calculates the final reconstructed value considering the previous error probability.

The rest of the paper is organized as follows. Section 2 analyzes the state-of-the-art on cyberattacks and countermeasures in CPS. Section 3 describes the proposed solution, including the mathematical model to describe the behavior of the CPS and the reconstruction and protection scheme to mitigate SSA. Finally, Section 4 describes the experimental validation and the results obtained. Section 5 concludes the paper.

2 Related Works

Cyber-Physical Systems are one of the most promising technological revolutions nowadays. They are expected to govern all production, domestic, and critical digital systems. Due to this relevance, many authors have investigated how to protect CPS against various well-known and innovative attacks. In general, we can distinguish two different protection approaches: those based on control theory and those supported by Information Technologies (IT).

IT protection mechanisms for CPS are usually data processing and filtering modules to remove and correct malicious or corrupted data packets. Stochastic techniques and models [4], advanced filtering algorithms such as the Kalman filter [15], hardware-enabled algorithms such as parameter estimation [16], and pattern recognition techniques to identify unusual information [17] are the most common technologies. As well as game-theory and other common technologies for CPS protection, such as honeypots [18] or Software-Defined Networks [19]. However, a limited number of works supporting this vision may be found, as information theory techniques are high-level and agnostic concerning the underlying hardware platform [20]. And the most critical cyber risks in CPS nowadays are associated with sensor and actuator nodes [8]. Different authors have identified new attack vectors and strategies [8], so feedback loops in CPS can be used to magnify cyberattacks starting in a single hardware node and spreading throughout the entire system. Furthermore, these IT protection technologies are computationally heavy and require long processing times, so they are not effective against fast cyberattacks. Other low-level lightweight techniques are required.

Physical infrastructure protection is, then, a priority in CPS. And most works on CPS security employ control theory to design new hardware protection schemes. Globally, all these technologies follow the same strategy [21]: they estimate or predict a secure future state for the physical platform and/or control loop, which is used to mitigate different types of attacks. Although this paradigm could fully protect CPS [22], it is very difficult to implement in practice and the reported implementation presents different weaknesses. Techniques may be local (or decentralized), distributed, or centralized.

Decentralized state estimation techniques are handled by independent sensing nodes. They are sparse as individual sensors have very limited information and actuation capabilities, so the achieved protection level is poor. Continuous bidimensional linear models are employed to detect perturbations and attacks (typically Denial of Service attacks) and modify the behavior of nodes by, for example, increasing their computational resources [23]. The objective is to guarantee the local stability of the control loops by mitigating all perturbances [21]. In contrast, other decentralized CPS protection schemes use variance-based strategies (also known as 'secure control' [14]). This approach is more general and can be applied against a generic cyberattack. Using discrete bidimensional models, tuned filters and tuned control loops can be varied to reduce system errors, even while a cyberattack is running [24]. However, even if local control loops can operate normally, with variance-based techniques the global system is handling corrupted data, and that impacts the later global behavior. Some authors have shown that global system protection requires cooperation and information sharing among all agents [21]. Distributed techniques fill this gap.

Distributed secure state estimation is useful against systemic attacks such as Byzantine attacks [25]. System states are deducted through an optimization process where linear models represent the sensors' outputs and graphs [26], Markov chains [27], binary decision trees [28], and other mathematical paradigms (such as the Lipschitz continuity) [29] are used to represent the interconnections and transmissions among nodes. Custom quasi-linear models for specific applications, such as series-parallel systems, have been also reported [30,31]. However, these protection mechanisms are passive and cannot deploy countermeasures to mitigate the impact of cyberattacks. Then, they must be complemented with specific controllers [32,33] to apply active protection policies on the CPS. Anyway, the final performance of distributed protection techniques is highly dependent on the number of trusted nodes, not affected by the attack [21,34]. Furthermore, linear and quasi-linear models cannot represent the output of most complex sensing platforms [35]. Thus, reported schemes can only be applied to a reduced number of application scenarios, excluding critical risks such as massive or viral attacks and common nonlinear algorithms.

On the other hand, recently distributed artificial intelligent solutions, such as federated learning [36], Support Vector Machines [37], feature selection [38] or eXplainable Artificial Intelligence (XAI) [39], have also been applied to CPS securitization and intrusion detection. But performance must be enhanced through additional techniques such as reinforcement learning [40]. Intelligent solutions must be designed for very specific attacks, as they are usually focused on Denial-of-Service attacks. Although the final results are promising, the balance between cost and performance is still worse than the one observed in other distributed techniques, and they are preferred to be used for privacy preservation [41].

The main disadvantage of distributed protection mechanisms is the increase in system congestion, due to the large number of transmissions required to run the distributed algorithms. On the contrary, centralized approaches may handle global stability and attacks (as distributed techniques) but with a lower system overload. Most reported works follow this paradigm.

Centralized control is usual in CPS, as it is the traditional approach in legacy Supervisory Control And Data Acquisition (SCADA) systems. Different kinds of multi-dimensional models are employed to represent the state of every single node on the platform. These models can be analytically manipulated to define protection algorithms based on Orthogonal-Triangular (QR) decomposition [42] or Linear–Quadratic (LQ) control [43], mitigating the impact of attacks. Models can be deterministic [44] or include some stochastic terms to represent noises [45]. Besides, continuous [44] and discrete [46] models may be found. However, most of these models are linear and only consider the self-maintained evolution of the node output and the measurement errors (in line with traditional control theory models). While other relevant effects, such as the digitalization process or the transmission protocols, are not considered, although they can be relevant. On the other hand, nonlinear models are very rare [47] and they are only developed for specific use cases. This centralized approach is successful against false information attacks (also known as sparse sensor attacks or deception attacks [14]), as it handles a full picture of the CPS. However, current models are very limited, and analytical protection algorithms have a reduced impact in real applications.

Table 1 summarizes the main current approaches and their associated open challenges.

Table 1: State-of-the-art

Reference	Technology	Short description	Open challenges
[4]	Information theory and technologies	Clustering techniques and stochastics models to big data collections	These high-level technologies and solutions cannot represent precisely attacks at physical level, such as the sparse sensor attack
[15]		Numerical data filtering (Kalman)	
[16]		Parameter estimation for numerical models and time series	
[17]		Pattern recognition for numerical models and time series	
[18]		Honeypots for attacker capture using the game theory	
[19]		Software-Defined Network technologies to create intrusion detection systems	
[21,24,23]	Decentralized state estimation	Sensor nodes apply linear models to analyze and guarantee the local stability of feedback control loops	Models are simple and they can only protect against some few local (never global) attacks

(Continued)

Table 1 (continued)

Reference	Technology	Short description	Open challenges
[26–31]	Distributed state estimation	Linear models and optimization algorithms to represent the nodes, their relations, behavior, and outputs	Models are still linear, and they only enable passive protection (such as alerts). Mitigation action are not possible
[32,33]	Controllers for distributed state estimation	Modules monitoring specific situation and triggering mitigation actions	They are specifically tailored for some applications, attacks, number of nodes to protect . . . the performance decreases in other scenarios. They can cost network congestion because of the intense information exchange
[36–40]	Artificial intelligence	Federated learning, Support Vector Machines, reinforcement learning, feature selection and XAI to identify attacks to the hardware platform	
[42–46]	Centralized protection control	Analytical models to represent the CPS and attackers' behavior. They can be discrete, continuous, deterministic, or stochastic	Linear models cannot represent CPS behavior in a general case, and important processes, such as the digitalization stage, are not considered
[47]		Nonlinear models for state estimation	Specifically tailored for some applications, the performance decreases in other scenarios

In this paper, we address this challenge, with a continuous-time generic multidimensional nonlinear model, and a protection policy based on probabilistic decision-making schemes.

3 Proposed Scheme

The proposed solution includes two different phases. First, a multidimensional stochastic model is employed to estimate or predict the future state of the CPS. Later, the obtained secure state estimation is compared to the real state produced by the physical platform. Both values are compared using a probabilistic model, where several indicators are considered. The system state may be replaced or corrected using the predicted secure state if the decision-making algorithm indicates the information is false (corrupted or caused by an SSA). This section describes in detail the entire scheme. Section 3.1 introduces the stochastic model, while Section 3.2 presents the decision-making and protection algorithms.

3.1 Secure State Estimation. Model Description

A CPS is supported by a dense sensing platform including N different sensor nodes n_m. These nodes monitor and control a catalogue of P different physical variables $x_i\,(t)$ (1). Each variable is monitored in K_i different geographical locations $g_s^{x_i}$ (2), so every node controls a different physical variable in a different location (3). If any node n_m monitors more than one variable, we are analyzing it as two independent nodes located in the same geographical position.

Hereinafter we are naming $x_i^s\ \ (t)$ the value of physical variable $x_i\,(t)$ in location $g_s^{x_i}$.

$$\{x_i\,(t)\ \ i=1,\ldots,P\} \tag{1}$$

$$\left\{g_s^{x_i}\ s=1,\ldots,K_i\right\} \tag{2}$$

$$N=\sum_{i=1}^{P}K_i \tag{3}$$

The information to be finally injected into the computational processes (or system state) is a set of M discrete state variables $y_j\,[k]$, related to the physical variables through a vector unknown function, $\mathcal{S}\,(\cdot)$, named as "system function" (4). This system function integrates five different processes: (i) the physical world's evolution, (ii) the transduction phase, (iii) the measurement scheme, (iv) the data transmission, and (v) the final processing stage.

$$\left\{y_j\,[k]\ j=1,\ldots,M\right\}=\ \mathcal{S}\left(x_i^s\ \ (t)\ \ i=1,\ldots,P\ s=1,\ldots,K_i\right) \tag{4}$$

The physical world (i) is considered a closed autonomous system, with no external intervention, so the future evolution of the physical variables is only determined by the past values of those same variables (5). The function relating the past and future values of the physical variables $\mathcal{F}^{x_i^s}\,(\cdot)$ is unknown and, in the general case, non-linear. For clarity, we are using vector $\vec{X}$ to represent the full ordered collection of physical variables (6). Although it is unknown, vector function $\mathcal{F}\,(\cdot)$ could be developed as Taylor's series.

$$x_i^s\ \ (t)\ \ t\ \geq t_0=\mathcal{F}^{x_i^s}\left(\left\{x_i^s\ \ (t)\ \ t<t_0\ i=1,..,P\ s=1,\ldots,K_i\right\}\right)\ \forall\ i,s \tag{5}$$

$$\vec{X}\ \ (t)=\left\{x_i^s\ \ (t)\ \ i=1,..,P\ s=1,\ldots,K_i\right\}\ \forall\ t \tag{6}$$

Any multidimensional function may be developed as Taylor's series using the partial derivation (7). For simplicity, we are using a McLaurin development around the origin. In this expression terms $\left(\frac{1}{k_1!\cdot\ldots\cdot k_N!}\frac{\partial^{k_1+\ldots+k_N}}{\partial\left(x_1^1\right)^{k_1}\cdot\ldots\cdot\partial\left(x_P^{K_P}\right)^{k_N}}\mathcal{F}^{x_i^s}\left(\vec{0}\right)\right)$ are unknown coefficients as function $\mathcal{F}^{x_i}(\cdot)$ is unknown too. We are representing them as λ_r (8). The purpose of our model is to estimate the future CPS state, in order to mitigate any potential SSA. Then, the model must be numerically implementable. Infinite series are not, and they must be limited to the R first terms (9). The error $E_{\mathcal{F}}$ (10) we are introducing because of this truncation is difficult to estimate as function $\mathcal{F}^{x_i^s}(\cdot)$ is unknown, but the Lagrange formula represents its analytical expression. Because this error is not easy to compute, the value for $R_{\mathcal{F}}$ parameter must be experimentally chosen, so the numerical model is precise enough to represent the behavior of CPS. However, in order to handle uncertainties in our model, we are proposing an estimation (maximum value) for error $E_{\mathcal{F}}$ (11). We are assuming function $\mathcal{F}^{x_i^s}$ grows exponentially (the maximum increasing speed) in all directions, so term $\left(\frac{\partial^{k_1+\ldots+k_N}}{\partial\left(x_1^1\right)^{k_1}\cdot\ldots\cdot\partial\left(x_P^{K_P}\right)^{k_N}}\mathcal{F}^{x_i^s}\left(\vec{0}\right)\right)$ is the unit. Besides, we choose the maximum value for term $\left(\frac{1}{k_1!\cdot\ldots\cdot k_N!}\right)$ which is achieved for $k_1=\ldots=k_{N-1}=1$.

$$\mathcal{F}^{x_i^s}\left(\vec{X}\ (t)\right)=\sum_{k_1=0}^{\infty}\cdots\sum_{k_N=0}^{\infty}\left(\frac{1}{k_1!\cdot\ldots\cdot k_N!}\frac{\partial^{k_1+\ldots+k_N}}{\partial\left(x_1^1\right)^{k_1}\cdot\ldots\cdot\partial\left(x_P^{K_P}\right)^{k_N}}\mathcal{F}^{x_i^s}\left(\vec{0}\right)\right)\cdot\left(x_1^1\right)^{k_1}\cdot\ldots\cdot\left(x_P^{K_P}\right)^{k_N} \tag{7}$$

$$\mathcal{F}^{x_i^s}\left(\vec{X}\ (t)\right)=\sum_{\substack{r=0\\ k_1+\ldots+k_N=r}}^{\infty}\lambda_r\cdot\left(x_1^1\right)^{k_1}\cdot\ldots\cdot\left(x_P^{K_P}\right)^{k_N} \tag{8}$$

$$\mathcal{F}^{x_i^s}\left(\vec{X}\ (t)\right)\approx\sum_{\substack{r=0\\ k_1+\ldots+k_N=r}}^{R_{\mathcal{F}}}\lambda_r\cdot\left(x_1^1\right)^{k_1}\cdot\ldots\cdot\left(x_P^{K_P}\right)^{k_N} \tag{9}$$

$$|E_{\mathcal{F}}|\leq max\left\{\lambda_{R_{\mathcal{F}}+1}\right\}=max\left\{\frac{1}{k_1!\cdot\ldots\cdot k_N!}\frac{\partial^{k_1+\ldots+k_N}}{\partial\left(x_1^1\right)^{k_1}\cdot\ldots\cdot\partial\left(x_P^{K_P}\right)^{k_N}}\mathcal{F}^{x_i^s}\left(\vec{0}\right)\ k_1+\ldots+k_N=R_{\mathcal{F}}+1\right\} \tag{10}$$

$$|E_{\mathcal{F}}|\leq\frac{1}{\left(R_{\mathcal{F}}-N+2\right)!} \tag{11}$$

The second process represented by the system function $\mathcal{S}(\cdot)$ is the transduction phase (ii). Physical variables $x_i^s(t)$ are transformed into electrical signals $v_i^s(t)$ through an unknown function $\mathcal{T}_m$ which is different for each sensor node n_m (12). Functions $\mathcal{T}_m$ are unidimensional (scalar) as the transduction process must be bijective to preserve the information. Besides, two different kinds of multiplicative noises affect the transduction phase. On the one hand, multiplicative physical noises $\varepsilon_r^{x_i^s}(t)$ (such as thermal noise or environmental radiation) are mixed with real physical variables $x_i^s(t)$ in the transformation function $\mathcal{T}_m$. On the other hand, multiplicative electrical noises $\xi_r^{x_i^s}(t)$ are added to the obtained electrical signals, because of the impact of electronic circuits. Each noise (electrical or physical) has a similar probability distribution $f[\cdot]$ (13). Noises are white (Gaussian), mutually uncorrelated with zero mean and unitary variances [42]. We are considering all these stochastic processes are stationary, and their probability distribution remains stable along time (14). Parameters $R^1_{v_i^s}$ and $R^2_{v_i^s}$ are positive integer numbers.

Unknown function $\mathcal{T}_m$ can be developed as Taylor's series as well, but in this case, expressions are simpler are unidimensional techniques can be applied (15). As done before, unknown coefficients $\left(\frac{1}{r!}\frac{d^r\mathcal{T}_m}{d\left(\widetilde{x}_i^s\right)^r}(0)\right)$ are represented by variables α_r. Besides, Taylor's series must be truncated too, to make our model computationally handleable (16), although an error $E_\mathcal{T}$ is introduced (17). An estimation for the maximum value of error $E_\mathcal{T}$ is proposed too (18), in order to enable the uncertainty management.

$$v_i^s(t) = \mathcal{T}_m\left(x_i^s(t) + \sum_{r=1}^{R^1_{v_i^s}} \varepsilon_r^{x_i^s}(t)\right) + \sum_{r=1}^{R^2_{v_i^s}} \xi_r^{x_i^s}(t) + \sum_{r=1}^{R^3_{v_i^s}} m_r^{x_i^s}(t) \tag{12}$$

$$P(a \le \varepsilon \le b) = \int_a^b f_\varepsilon(u)\ du \tag{13}$$

$$P\left(\xi_r^{x_i^s}(t) = u\right) \sim P\left(\varepsilon_r^{x_i^s}(t) = u\right) \sim \frac{1}{\sqrt{2\pi}} e^{\frac{-u^2}{2}}\ \forall t \tag{14}$$

$$\begin{gathered}\mathcal{T}_m\left(\widetilde{x}_i^s\right) = \sum_{r=0}^{\infty}\frac{1}{r!}\frac{d^r\mathcal{T}_m}{d\left(\widetilde{x}_i^s\right)^r}(0)\cdot\left(\widetilde{x}_i^s\right)^r \\ being\ \widetilde{x}_i^s(t) = x_i^s(t) + \sum_{r=1}^{R^1_{v_i^s}} \varepsilon_r^{x_i^s}(t)\end{gathered} \tag{15}$$

$$\mathcal{T}_m\left(\widetilde{x}_i^s\right) \approx \sum_{r=0}^{R_\mathcal{T}} \alpha_r \cdot \left(\widetilde{x}_i^s\right)^r \tag{16}$$

$$E_\mathcal{T} \le \alpha_{R_\mathcal{T}+1} = \frac{1}{(R_\mathcal{T}+1)!}\frac{d^{R_\mathcal{T}+1}\mathcal{T}_m}{d\left(\widetilde{x}_i^s\right)^{R_\mathcal{T}+1}}(0) \tag{17}$$

$$E_\mathcal{T}\ \le\ \frac{1}{(R_\mathcal{T}+1)!} \tag{18}$$

Although white noises $\varepsilon_r^{x_i^s}$ are affected by functions $\mathcal{T}_m$ and then, they are part of the Taylor's series (15)–(16) because of the fact they follow a Gaussian distribution, these expressions can be simplified, so only the physical variables are part of the Taylor's polynomial. Every term in the Taylor's series where a noise $\varepsilon_r^{x_i^s}$ is included may be considered as a non-monotonous transformation $T(\cdot)$ of a Gaussian random variable. The transformation theorem (19) shows that any transformed Gaussian distribution is a new Gaussian distribution with mean μ_t and variance σ_t (20). Later, all the transformed Gaussian distributions may be aggregated, and, because of the central limit theorem, the resulting random variable $\chi^{x_i^s}$ is a Gaussian distribution too. However, mean μ_{tt} and variance σ_{tt} (21) are unknown, as the final value for the mean and variance of the global distribution $\chi^{x_i^s}$ depends on the transformations and the value of $R_{\mathcal{T}}$ parameter. Taylor's series for function $\mathcal{T}_m$ may be finally rewritten (22).

$$f_{\mathcal{T}m\left(\varepsilon_r^{x_i^s}\right)}(u) = \sum_u \frac{f_{\varepsilon_r^{x_i^s}}(u)}{\left|\frac{d}{d\varepsilon_r^{x_i^s}} T\left(\varepsilon_r^{x_i^s}\right)\Big|_{\varepsilon=u}\right|} \tag{19}$$

$$\textit{being } \{u\} \textit{ the roots of } T\left(\varepsilon_r^{x_i^s}\right)$$

$$P\left(\mathcal{T}_m\left(\varepsilon_r^{x_i^s}\right) = u\right) \sim \frac{1}{\sigma_t\sqrt{2\pi}}\, e^{\frac{-(u-\mu_t)^2}{2\sigma_t^2}} \tag{20}$$

$$P\left(\chi^{x_i^s}(t) = u\right) \sim \frac{1}{\sigma_{tt}\sqrt{2\pi}}\, e^{\frac{-(u-\mu_{tt})^2}{2\sigma_{tt}^2}} \quad \forall t \tag{21}$$

$$\mathcal{T}_m\left(\widetilde{x}_i^s\right) \approx \mathcal{T}_m\left(x_i^s\right) = \sum_{r=0}^{R_{\mathcal{T}}} \alpha_r \cdot \left(x_i^s\right)^r + \chi^{x_i^s} \tag{22}$$

Finally, it is necessary to estimate the value for mean μ_{tt} and variance σ_{tt}, so error in the proposed model may be properly handled. In general, errors are bigger as values for the mean μ_t and variance σ_t go up. Then, a superior limit is a good approximation for both parameters (23), which may be easily calculated considering the reproducibility of the Gaussian random variables. In order to get the final values for the mean μ_{tt} and variance σ_{tt} it is enough to apply the same reproducibility law a second time (24).

$$f_{\mathcal{T}m\left(\varepsilon_r^{x_i^s}\right)}(u) = \sum_u \frac{f_{\varepsilon_r^{x_i^s}}(u)}{\left|\frac{d}{d\varepsilon_r^{x_i^s}} T\left(\varepsilon_r^{x_i^s}\right)\Big|_{\varepsilon=u}\right|} \leq \sum_u f_{\varepsilon_r^{x_i^s}}(u) \sim \frac{1}{\sqrt{2\pi \cdot R_{\mathcal{T}}}}\, e^{\frac{-\left(u-\sum_{r=1}^{R_{\mathcal{T}}}\left(x_i^s\right)^r\right)^2}{2R_{\mathcal{T}}}} \tag{23}$$

$$\begin{aligned} \sigma_{tt} &= \sqrt{R_{\mathcal{T}} \cdot R_{v_i^s}^1} \\ \mu_{tt} &= R_{v_i^s}^1 \cdot \sum_{r=1}^{R_{\mathcal{T}}} \left(x_i^s\right)^r \end{aligned} \tag{24}$$

The transduction phase is open, so it can be affected by SSA and malicious signals. In order to represent this risk, we are considering a set of additive malicious signals $m_r^{x_i^s}(t)$ whose value is determined by a stochastic process. This stochastic process is characterized by a Bernoulli distribution Γ_a, representing the existence of a running SSA. Parameter a is equal to the unit if the attack is running, or zero in the opposite case (25). The attack probability ρ_a varies with time (as Γ_a is a stochastic process). The estimation scheme for this probability is part of the attack detection algorithm (see Section 3.2). In our model, a SSA consists of adding a false data signal $z_r^{x_i^s}(t)$ to the data electrical signal $v_i^s(t)$, according to the previously described distribution (26). $R_{v_i^s}^3$ different uncorrelated attackers may be operating over the CPS at the same time. False data signals, in our model, are understood as unreported and unexpected perturbations. This is relevant in order to define a precise attack detection and mitigation strategy (see Section 3.2).

$$\Gamma_a(t) \;:\; P(a;t) = \begin{cases} 1-\rho_a(t) & \textit{if } a=0 \\ \rho_a(t) & \textit{if } a=1 \end{cases} \tag{25}$$

$$m_r^{x_i^s} = a_r^{x_i^s} \cdot z_r^{x_i^s}(t) \tag{26}$$

Finally, as every sensor node n_m has a different function $\mathcal{T}_m$, the whole CPS is represented by a set of N different Eq. (27), which can be represented in one vector expression (28).

$$\begin{cases} v_1^1(t) = \sum_{r=0}^{RJ} \alpha_r^{x_1^1} \cdot \left(x_i^s\right)^r + \mathcal{X}^{x_1^1} + \sum_{r=1}^{R^2_{v_1^1}} \xi_r^{x_1^1}(t) + \sum_{r=1}^{R^3_{v_1^1}} m_r^{x_1^1}(t) \\ \ldots \\ v_p^{kp}(t) = \sum_{r=0}^{RJ} a_r^{x_p^{kp}} \cdot \left(x_i^s\right) + \mathcal{X}^{x_p^{kp}} + \sum_{r=1}^{R^{2}_{v_p^{kp}}} \xi_r^{x_p^{kp}}(t) + \sum_{r=1}^{R^{3}_{v_p^{kp}}} m_r^{x_p^{kp}}(t) \end{cases} \tag{27}$$

$$\vec{V}(t) = \sum_{r=0}^{R_{\mathcal{T}}} \vec{\alpha}_r \cdot \left(\vec{X}\ (t)\right)^r + \vec{\chi} + \sum_{r=1}^{R^2_{v_1^1}} \vec{\xi}_r(t) + \sum_{r=1}^{R^3_{v_1^1}} \vec{m}_r(t) \tag{28}$$

The third subprocess to be represented in our model is the measurement scheme (iii). This, basically, is a digitalization scheme, developed internally by sensor nodes (see Fig. 1). Discrete signal $d_i^s[k]$ are obtained through an ideal sampling scheme (29), where electrical signals are multiplied by a Dirac comb or impulse train $\omega_{T_m}(t)$ with period T_m (30). This period T_m is different for each sensor node n_m.

$$\begin{aligned} d_i^s(t) &= v_i^s(t) \cdot \omega_{T_m}(t) + q_i^s(t) = \sum_{k=-\infty}^{\infty} v_i^s(k \cdot T_m) \cdot \delta(t - k \cdot T_m) + q_i^s(k \cdot T_m) \\ d_i^s[k] &= v_i^s(k \cdot T_m) + q_i^s(k \cdot T_m) \ \ k \in \mathbb{N} \end{aligned} \tag{29}$$

$$\omega_{T_m}(t) = \sum_{k=-\infty}^{\infty} \delta(t - k \cdot T_m) \tag{30}$$

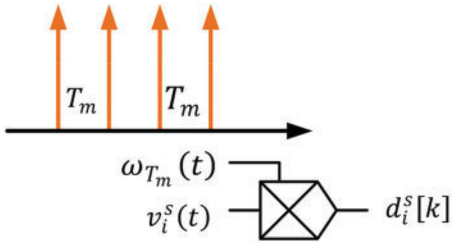

Figure 1: Ideal sampling scheme

In this digitalization process only the quantification noise $q_i^s(t)$ is relevant. This noise, as the digitalization scheme is invariant in time, is also time-invariant, and characterized by a uniform random variable (31).

$$P\left(q_i^s(t) = u\right) \sim \begin{cases} \frac{1}{\Delta_m} \ u \in [-\Delta_m, \ \Delta_m] \\ 0 \ otherwise \end{cases} \tag{31}$$

where Δ_m is the quantification step, fixed for every node n_m.

The fourth process to be represented is the data transmission (iv). In general, hardware platforms in CPS are low-energy, and they sleep most of the time. Being event-based, they only activate the transmission subsystem when an event is detected in the physical world. We are defining function $\phi_m[k]$

as event-triggering function (32). This function takes as value the unit in the discrete time instant a new event must be generated. Its value is zero otherwise. Function $u(\cdot)$ is the Heaviside step function, k_e is the last time instant where an event took place, and e_m is a parameter, different for each node n_m. If signal d_i^s changes its value more than e_m units, a new event is triggered.

$$\phi_m[k] = u\left(d_i^s[k]\, d_i^s[k_e] - (e_m)^2\right) \tag{32}$$

Data transmission is, once again, an open process, so it is vulnerable to attacks. Denial-of-Service (DoS) attacks in this case. But SSA too (as the transduction phase). Bernoulli distribution Γ_b represents the probability of a DoS attack to be running. Parameter b is equal to the unit if an attack is being performed, or zero if not (33). The attack probability ρ_b varies with time (as Γ_b is a stochastic process). The estimation scheme for this probability is part of the attack detection algorithm (see Section 3.2). Similarly, Bernoulli distribution Γ_c represents the probability of a SSA to be running at the data transmission stage (34).

$$\Gamma_b[k] \ : \ P[b;k] = \begin{cases} 1 - \rho_b[k] \ \ if \ b = 0 \\ \rho_b[k] \ \ if \ b = 1 \end{cases} \tag{33}$$

$$\Gamma_c[k] \ : \ P[c;k] = \begin{cases} 1 - \rho_c[k] \ \ if \ c = 0 \\ \rho_c[k] \ \ if \ c = 1 \end{cases} \tag{34}$$

All parameters and their meaning are equivalent to distributions Γ_a and Γ_b.

In our model, a DoS attack is represented by an arbitrary delay of k_d units in the data transmission, while an SSA is represented by injected false signals $h_i^s[k]$ (35). Then, the received signal by the remote central control platform $w_i^s[k]$ depends on function $\phi_m[k]$ and distributions Γ_b and Γ_c, but also is affected by transmission errors and noises.

$$w_i^s[k] = c \cdot h_i^s[k] + (1-c) \cdot \left(b \cdot d_i^s[k-k_d] + (1-b) \cdot d_i^s[k]\right) + \sum_{r=1}^{R_{w_i^s}} \varphi_r^{w_i^s}[k] \tag{35}$$

$$being\ k \vdots \phi_m[k] = 1$$

Our model considers $R_{w_i^s}$ multiplicative white Gaussian uncorrelated noises, $\varphi_r^{w_i^s}$ with zero mean and unitary variance, affecting the data transmission.

Finally, information injected into computational processes $y_j[k]$ may not be the raw physical information, but a transformation of it (mean, minimum or maximum values, for example). Then, the final step in the system function $\mathcal{S}(\cdot)$ is the processing stage (v). Processing processes may combine different transmitted signals $w_i^s[k]$ according to function $\mathcal{P}^{y_j}(\cdot)$ which, in general, is unknown and, even, may change with time (36). This is an internal process where only numerical errors may affect the final result. However, central control systems are usually computationally powerful, and numerical error are negligible.

$$y_j[k] = \mathcal{P}^{y_j}\left(\overrightarrow{W\ [k]}\right) = \mathcal{P}^{y_j}\left(w_i^s[k]\ \ i = 1, \ldots, P\ s = 1, \ldots, K_i\right) \tag{36}$$

Function $\mathcal{P}^{y_j}(\cdot)$ may be developed as Taylor's series, in a similar way as done for function $\mathcal{F}^{x_i^s}$. Considering unknown coefficients β_r, the final equation of our model may be described as a polynomial (37). Introducing an error $E_{\mathcal{P}}$, whose maximum value may be estimated using the same techniques described before (38), the model may be truncated and limited to $R_{\mathcal{P}}$ terms (39) so it can be managed by computational infrastructures.

$$\mathcal{P}^{y_j}\left(\overrightarrow{W\ [k]}\right) = \sum_{\substack{r=0 \\ k_1+\cdots+k_N=r}}^{\infty} \beta_r \cdot \left(w_1^1\right)^{k_1} \cdot \ldots \cdot \left(w_P^{K_P}\right)^{k_N} \tag{37}$$

$$being\ \beta_r = \left(\frac{1}{k_1! \cdot \ldots \cdot k_N!} \frac{\partial^{k_1+\ldots+k_N}}{\partial \left(x_1^1\right)^{k_1} \cdot \ldots \cdot \partial \left(x_P^{K_P}\right)^{k_N}} \mathcal{P}^{y_j}\left(\vec{0}\right)\right)$$

$$|E_{\mathcal{F}}| \leq max\left\{\beta_{R_{\mathcal{P}}+1}\right\} \leq \frac{1}{(R_{\mathcal{P}} - N + 2)\,!} \tag{38}$$

$$\mathcal{P}^{y_j}\left(\overrightarrow{W\ [k]}\right) \approx \sum_{\substack{r=0 \\ k_1+\ldots+k_N=r}}^{R_{\mathcal{P}}} \beta_r \cdot \left(w_1^1\right)^{k_1} \cdot \ldots \cdot \left(w_P^{K_P}\right)^{k_N} \tag{39}$$

Then, the final analytical model to describe the behavior of CPS includes five different Eq. (40). All parameters and coefficients are known (or may be estimated) but λ_r, α_r and β_r which must be calculated. The value for those parameters is obtained from an initial calibration process and an optimization algorithm based on the minimization of the Mean Square Error (MSE).

$$\begin{cases} x_i^s(t) = \displaystyle\sum_{\substack{r=0 \\ k_1+\cdots+k_N=r}}^{R_{\mathcal{F}}} \lambda_r \cdot \left(x_1^1\right)^{k_1} \cdot \ldots \cdot \left(x_P^{K_P}\right)^{k_N} \qquad t \geq t_0 \\ \vec{V}(t) = \displaystyle\sum_{r=0}^{R_{\mathcal{T}}} \vec{\alpha}_r \cdot \left(\vec{X}\ (t)\right)^r + \vec{\chi} + \sum_{r=1}^{R^2_{v_1^1}} \vec{\xi}_r(t) + \sum_{r=1}^{R^3_{v_1^1}} \vec{m}_r(t) \\ d_i^s[k] = v_i^s(k \cdot T_m) + q_i^s(k \cdot T_m) \qquad k \in \mathbb{N} \\ w_i^s[k] = c \cdot h_i^s[k] + (1-c) \cdot \left(b \cdot d_i^s[k-k_d] + (1-b) \cdot d_i^s[k]\right) + \displaystyle\sum_{r=1}^{R_{w_i^s}} \varphi_r^{w_i^s}[k] \\ y_j[k] = \displaystyle\sum_{\substack{r=0 \\ k_1+\ldots+k_N=r}}^{R_{\mathcal{P}}} \beta_r \cdot \left(w_1^1\right)^{k_1} \cdot \ldots \cdot \left(w_P^{K_P}\right)^{k_N} \end{cases} \tag{40}$$

3.2 Reconstruction and Protection Mechanisms

The proposed model (see Section 3.1) considers seven sources of perturbations. On the one hand, errors may be caused by four different phenomena: erratic behaviors in the physical variables, electrical noises, quantification noise, and transmission perturbations. And, on the other hand, three different potential attacks affect CPS in the general case: SSA at the transduction phase, and SSA and Denial-of-Service attacks at the transmission phase. Fig. 2 represents the proposed reconstruction and protection mechanisms.

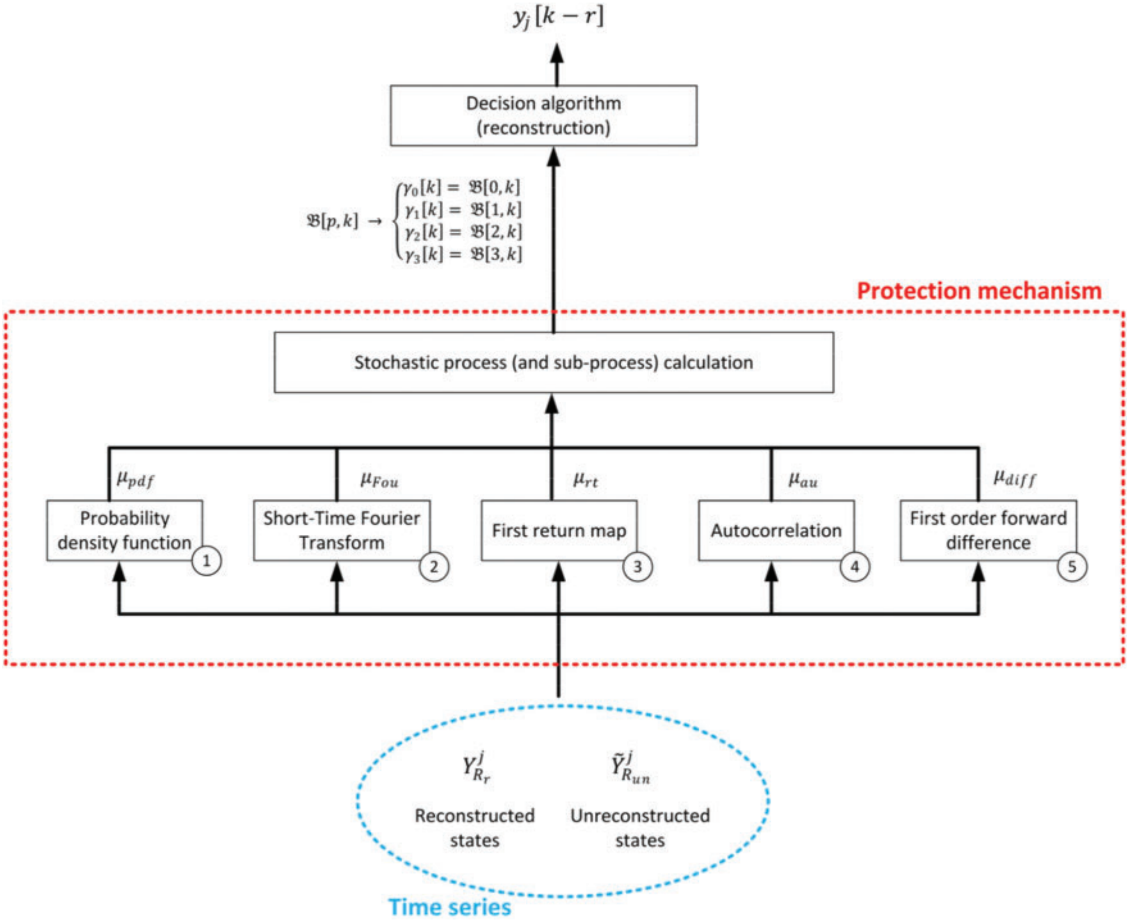
$y_j[k-r]$
Decision algorithm (reconstruction)
$\mathfrak{B}[p,k] \rightarrow \begin{cases} \gamma_0[k] = \mathfrak{B}[0,k] \\ \gamma_1[k] = \mathfrak{B}[1,k] \\ \gamma_2[k] = \mathfrak{B}[2,k] \\ \gamma_3[k] = \mathfrak{B}[3,k] \end{cases}$
Protection mechanism
Stochastic process (and sub-process) calculation
μ_{pdf}
μ_{Fou}
μ_{rt}
μ_{au}
μ_{diff}
Probability density function
1
Short-Time Fourier Transform
2
First return map
3
Autocorrelation
4
First order forward difference
5
$Y^j_{R_r}$
$\tilde{Y}^j_{R_{un}}$
Reconstructed states
Unreconstructed states
Time series

Figure 2: Protection and reconstruction mechanism

As a novelty, the proposed reconstruction and protection mechanism evaluates all these potential perturbations to build a global stochastic process (contrary to traditional deterministic models). This stochastic process $\mathfrak{B}[p,\ k]$ (41) is discrete. p is a discrete variable representing the four possible situations a CPS state may achieve: unperturbed ($p = 0$), noisy ($p = 1$), SSA-attacked ($p = 2$) and DoS-attacked ($p = 3$). While k is a variable representing the discrete time. Similarly, we can define M stochastic sub-processes $\mathfrak{B}_j[p,\ k]$ for each one of the M state variables y_j considered in the CPS.

$$\mathfrak{B}[p,\ k] \rightarrow \begin{cases} \gamma_0[k] = \mathfrak{B}[0,\ k] \\ \gamma_1[k] = \mathfrak{B}[1,\ k] \\ \gamma_2[k] = \mathfrak{B}[2,\ k] \\ \gamma_3[k] = \mathfrak{B}[3,\ k] \end{cases} \tag{41}$$

In the proposed protection mechanism, five indicators are employed to evaluate the probability distribution of the stochastic process $\mathfrak{B}[p,\ k]$ at every time instant k: ① the probability density function, ② the Short-Time Fourier transform, ③ the first-return map, ④ the autocorrelation and ⑤ the first order forward difference. To evaluate all these indicators, the protection algorithm operates with two numerical series. $Y^j_{R_r}$ (42) represents the series of the last R_r reconstructed states for the j-th state variable, and $\widetilde{Y}^j_{R_{un}}$ (43) represents the series of the last R_{un} unreconstructed states for the j-th state variable (being k the current time instant).

$$Y^j_{R_r} = \left\{y_j\left[k-r-R_{un}\right]\ r=1,\ldots,R_r\right\} \tag{42}$$

$$\widetilde{Y}^j_{R_{un}} = \left\{\widetilde{y}_j\left[k-r\right]\ r=1,\ldots,R_{un}\right\} \tag{43}$$

In order to make feasible the calculation of all these indicators from time series $Y^j_{R_r}$ and $\widetilde{Y}^j_{R_{un}}$, in this work we propose a specifically tailored definition. The calculation process for every indicator is described below.

Regarding the probability density function ①, the probability μ^1_j of any unreconstructed state $\widetilde{y}_j[k]$ for the j-th state variable to happen in a given CPS, may be evaluated considering the previous reconstructed states $Y^j_{R_r}$ achieved by that CPS and the Laplace definition of probability (44), and being $\delta[\cdot]$ the Kronecker's delta function. The probability μ^j_{pdf} for the entire $\widetilde{Y}^j_{R_{un}}$ series of unreconstructed states may be obtained as the mean value of all the individual probabilities (45). And, finally, the global probability μ_{pdf} for all the M state variables may be calculated as the average value (46).

$$\mu^1_j = \frac{\sum_{r=1}^{R_r}\delta\left[y_j\left[k-r-R_{un}\right]-\widetilde{y}_j\left[k\right]\right]}{R_r} \tag{44}$$

$$\mu^j_{pdf} = \frac{\sum_{m=1}^{R_{un}}\sum_{r=1}^{R_r}\delta\left[y_j\left[k-r-R_{un}\right]-\widetilde{y}_j\left[k-m\right]\right]}{R_{un}\cdot R_r} \tag{45}$$

$$\mu_{pdf} = \frac{1}{M}\sum_{j=1}^{M}\mu^j_{pdf} \tag{46}$$

But even if the unreconstructed CPS state has a relevant probability, it can still be manipulated and not be coherent with the system evolution. This situation may be detected through two different indicators: the Short-Time Fourier Transform (STFT) and the first-return map. Considering the Short-Time Fourier Transform (STFT) ②, the Fourier spectrum tends to be stable in a CPS, so any abrupt change may indicate an attack. The STFT (47) is equivalent to the traditional Fourier transform, but only considering a limited number of samples (instead of the usual infinite sum) through a window function $\Omega\left[k,\ R_{sam}\right]$, typically the Hann (Hanning) window (48) with a width of R_{sam} samples. Then, the STFT $\mathcal{Y}^j_{R_r}$ for the reconstructed states $Y^j_{R_r}$ (49) may be calculated using a numerical algorithm, as well as the STFT $\widetilde{\mathcal{Y}}^j_{R_{un}}$ for the reconstructed states $\widetilde{Y}^j_{R_{un}}$ (50).

$$STFT\left\{y\left[k\right]\right\} = \mathcal{Y}\left(m,\nu\right) = \sum_{k=-\infty}^{\infty} y\left[k\right]\cdot\Omega\left[k-m,\ R_{sam}\right]\cdot e^{-jk\nu} \tag{47}$$

$$\Omega\left[k,\ R_{sam}\right] = \frac{1}{2}-\frac{1}{2}cos\left(\frac{2\pi k}{R_{sam}}\right) \tag{48}$$

$$\mathcal{Y}^j_{R_r} = STFT\left\{Y^j_{R_r}\right\} = \sum_{k=-\infty}^{\infty} y_j\left[k\right]\cdot\Omega\left[k-\frac{R_r}{2}-R_{un},\ R_r\right]\cdot e^{-jk\nu} \tag{49}$$

$$\widetilde{\mathcal{Y}}^j_{R_{un}} = STFT\left\{\widetilde{Y}^j_{R_{un}}\right\} = \sum_{k=-\infty}^{\infty} \widetilde{y}_j\left[k\right]\cdot\Omega\left[k-\frac{R_{un}}{2},\ R_{un}\right]\cdot e^{-jk\nu} \tag{50}$$

Then, using the Euclidean definition for distance, we can analyze how different $\mathcal{Y}^j_{R_r}$ and $\widetilde{\mathcal{Y}}^j_{R_{un}}$ are (51). As the distance μ^j_{Fou} gets bigger, the probability of unreconstructed states $\widetilde{Y}^j_{R_{un}}$ to be manipulated increases. As done before, the global distance μ_{Fou} for all the M state variables may be calculated as the average value of all partial distances μ^j_{Fou} (52).

$$\mu^j_{Fou} = \| \mathcal{Y}^j_{R_r} - \widetilde{\mathcal{Y}}^j_{R_{un}} \| = \sqrt{\left(\mathcal{Y}^j_{R_r} - \widetilde{\mathcal{Y}}^j_{R_{un}}\right) \cdot \left(\mathcal{Y}^j_{R_r} - \widetilde{\mathcal{Y}}^j_{R_{un}}\right)} \tag{51}$$

$$\mu_{Fou} = \frac{1}{M} \sum_{j=1}^{M} \mu^j_{Fou} \tag{52}$$

Another indicator we can use to identify situations where the unreconstructed CPS state is manipulated is the first-return map ③. The first return map is a function $\Pi(\cdot)$, which can be obtained numerically, and shows the relation between consecutive (reconstructed) CPS states (53). The minimum Euclidean distance μ^2_j between every ordered pair of unreconstructed states $\pi[m]$ (54) and the first return map $\Pi(\cdot)$ represents how close the unreconstructed states are to the expected behavior (55). The global distance μ^j_{rt} for the entire $\widetilde{Y}^j_{R_{un}}$ series may be obtained as the mean value (56), and the global distance μ_{rt} for all the M state variables may be calculated as the average value of all partial distances μ^j_{rt} (57).

$$y_j[k+1] = \Pi\left(y_j[k]\right) \tag{53}$$

$$\pi[m] = \left(\tilde{y}_j[k-m], \tilde{y}_j[k-m+1]\right) \tag{54}$$

$$\begin{aligned} \mu^2_j &= \min_{\eta(r) \in \Pi(\cdot)} \| \pi[m] - \eta(r) \| \\ &= \min_{r \in (2, R_r)} \|\pi[m] - \left(y_j[k - R_{un} - r], y_j[k - R_{un} - r + 1]\right)\| \end{aligned} \tag{55}$$

$$\mu^j_{rt} = \frac{1}{R_{un}} \sum_{m=2}^{R_{un}} \left(\min_{r \in (2, R_r)} \|\pi[m] - \left(y_j[k - R_{un} - r], y_j[k - R_{un} - r + 1]\right)\| \right) \tag{56}$$

$$\mu_{rt} = \frac{1}{M} \sum_{j=1}^{M} \mu^j_{rt} \tag{57}$$

But in some situations, very noisy states are difficult to distinguish from attacks. To clarify and separate these two situations we use the autocorrelation ④. Noise is a random effect, so autocorrelation tend to the null value very quickly. While planned attacks follow a certain structure, and autocorrelation oscillates but not disappears because of these patterns. But autocorrelation cannot be directly applied to series $\widetilde{Y}^j_{R_{un}}$ or $Y^j_{R_r}$, as they contain actual information, and it would be always non-null. Then, before calculating the autocorrelation we are using a stop-band filter to remove the legitimate information (see Fig. 3).

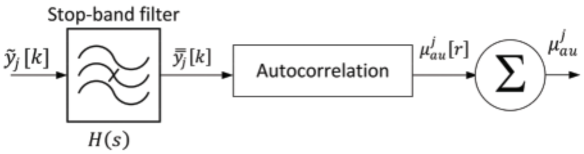

Figure 3: Stop-band filtering for autocorrelation calculation

From the STFT $y^j_{R_r}$ we can obtain the central frequency Ω_0 and the bandwidth Ω_c of the j-th information signal (state variable). And then, the stop-band filter in the Laplace domain may be described as a quotient function (58). And the filtering process as a product (59). The resulting filtered signal $\overline{\overline{y}}_j[k]$ (60), thus, only contains information about the noise and/or attacks affecting the CPS. The autocorrelation $\mu^j_{au}[r]$ can be now obtained (61).

$$H(s) = \frac{s + \Omega_0^2}{s^2 + \Omega_c \cdot s + \Omega_0^2} \tag{58}$$

$$\overline{\overline{y}}^j_{R_{un}} = \widetilde{y}^j_{R_{un}} \cdot H(s) \tag{59}$$

$$\overline{\overline{y}}_J[k] = STFT^{-1}\left\{\overline{\overline{y}}^j_{R_{un}}\right\} \tag{60}$$

$$\begin{aligned} \mu^j_{au}[r] &= \frac{\sum_{m=1}^{R_{un}} \left(\overline{\overline{y}}_j[k-m] - \varpi_j\right) \cdot \left(\overline{\overline{y}}_j[k-m+r] - \varpi_j\right)}{\sum_{m=1}^{R_{un}} \left(\overline{\overline{y}}_j[k-m] - \varpi_j\right)^2} \quad r \in \left[0, \frac{R_{un}}{2}\right] \\ \varpi_j &= \frac{1}{R_{un}} \sum_{r=1}^{R_{un}} \overline{\overline{y}}_j[k-r] \end{aligned} \tag{61}$$

This autocorrelation $\mu^j_{au}[r]$ should disappear as r parameter increases if the CPS is just noisy. To get that confirmation but avoid possible transitory effects, we are aggregating the last R_{cor} samples in the autocorrelation function $\mu^j_{au}[r]$ (62). The resulting indicator μ^j_{au} will be lower as the perturbations in the unreconstructed state are more similar to Gaussian white noise. As in all the previous indicators, the global autocorrelation μ_{au} for all the M state variables may be calculated as the average value of all partial distances μ^j_{au} (63).

$$\mu^j_{au} = \sum_{r=\frac{R_{un}}{2}-R_{cor}}^{\frac{R_{un}}{2}} \mu^j_{au}[r] \tag{62}$$

$$\mu_{au} = \frac{1}{M}\sum_{j=1}^{M} \mu^j_{au} \tag{63}$$

However, some attacks may use perturbations within the information signals' bandwidth, and autocorrelation may not generate a conclusive result. To analyze this situation, we are using our last indicator, the first order forward difference ⑤. The first order forward difference $\mu^j_{diff}[k]$ (60) represents the tendency, evolution or growing of the j-th unreconstructed state variable. In general, CPS states fluctuate but do not increase or decrease in a monotonous manner. Even less if the evolution is divergent (for example, exponential). Then, the sum μ^j_{diff} of all $(R_{un} - 1)$ samples in the first order forward difference $\mu^j_{diff}[m]$ is typically very small (61), as growing periods are cancelled by decreasing phases and vice versa. But if the CPS state is manipulated and it does not oscillate but increases or decreases monotonously and diverges, the sum μ^j_{diff} will take very extreme values (positive or negative). The global tendency (aggregated first order differences) μ_{diff} for all the M state variables may be calculated as the average value (62).

$$\mu^j_{diff}[m] = \widetilde{y}_j[m+1] - \widetilde{y}_j[m] \;\; m \in [k-2,\; k-R_{un}] \tag{64}$$

$$\mu^j_{diff} = \sum_{m=k-2}^{k-R_{un}} \mu^j_{diff}[m] \tag{65}$$

$$\mu_{diff} = \frac{1}{M}\sum_{j=1}^{M}\mu_{diff}^{j} \tag{66}$$

Using these five indicators, we can now estimate the probability distribution of the stochastic process $\mathfrak{B}[p,\ k]$. Mathematical models for all this probability distribution are a genuine contribution of this work. The unperturbed state ($p = 0$) is only probable when the probability μ_{pdf} is very high (close to the unit), and distances μ_{Fou} and μ_{rt} are very low. Any other value is indicating a noisy state ($p = 1$), which is still probable even for smaller values of probability μ_{pdf} and bigger values of distances μ_{Fou} and μ_{rt}. But noisy states require a low value for autocorrelation μ_{au} to be probable (on the contrary, the CPS may be under a cyberattack). Because of this sensitivity, exponential and power laws are the most adequate ones to represent the probability of the unperturbed state $\gamma_0[k]$ (63), while linear evolutions and slower exponential laws fit the more tolerant behavior of the probability law $\gamma_1[k]$ of the noisy state (64).

$$\gamma_0[k] = \left(1 - \left(\mu_{pdf} - 1\right)^{2\tau_0^1}\right)\cdot exp\left(-\frac{\mu_{Fou}}{\tau_0^2}\right)\cdot exp\left(-\frac{\mu_{rt}}{\tau_0^3}\right) \tag{67}$$

$$\gamma_1[k] = \mu_{pdf}\cdot exp\left(-\frac{\mu_{Fou}}{\tau_1^2}\right)\cdot exp\left(-\frac{\mu_{rt}}{\tau_1^3}\right)\cdot exp\left(-\frac{\mu_{au}}{\tau_1^4}\right) \tag{68}$$

Being τ_0^1 a positive integer number and τ_0^2, τ_0^3, τ_1^2, τ_1^3 and τ_1^4 positive real numbers (weights). They are used to control the sensitivity of the stochastic process.

On the other hand, SSA-attacked state ($p = 2$) is characterized by very a low probability μ_{pdf} but very high distances μ_{Fou} and μ_{rt}. As well as a relevant non-null value in the aggregated autocorrelation μ_{au} and the aggregated first order forward differences μ_{diff}. On the contrary, DoS-attacked states ($p = 3$) are usually associated to moderate values for the probability μ_{pdf} (states are delayed but not manipulated) while still very high distances μ_{Fou} and μ_{rt} (as they are delayed, states are not coherent with the historical series). The aggregated autocorrelation μ_{au} and the aggregated first order forward differences μ_{diff} tend also to be quite reduced. Following a similar philosophy to employed before, we can define the evolution laws for the probabilities $\gamma_2[k]$ (65) and $\gamma_3[k]$ (66).

$$\gamma_2[k] = \left(1 - \left(\mu_{pdf}\right)^{2\tau_2^1}\right)\cdot\left(1 - exp\left(-\frac{\mu_{Fou}}{\tau_2^2}\right)\right)\cdot\left(1 - exp\left(-\frac{\mu_{rt}}{\tau_2^3}\right)\right)\cdot\left(1 - exp\left(-\frac{\mu_{au}}{\tau_2^4}\right)\right) \cdot\left(1 - exp\left(-\frac{\left(\mu_{diff}\right)^2}{\tau_2^5}\right)\right) \tag{69}$$

$$\gamma_3[k] = \mu_{pdf}\cdot\left(1 - exp\left(-\frac{\mu_{Fou}}{\tau_3^2}\right)\right)\cdot\left(1 - exp\left(-\frac{\mu_{rt}}{\tau_3^3}\right)\right)\cdot exp\left(-\frac{\mu_{au}}{\tau_3^4}\right)\cdot exp\left(-\frac{\left(\mu_{diff}\right)^2}{\tau_3^5}\right) \tag{70}$$

Being τ_2^1 a positive integer number and τ_2^2, τ_2^3, τ_2^4, τ_2^5, τ_3^2, τ_3^3, τ_3^4 and τ_3^5 positive real numbers (weights).

In an equivalent manner we may calculate the probability distribution for all stochastic subprocesses $\mathfrak{B}_j[p,\ k]$ (67)

$$\begin{aligned}
\gamma_0^j[k] &= \left(1-\left(\mu_{pdf}^j-1\right)^{2\tau_0^1}\right)\cdot exp\left(-\frac{\mu_{Fou}^j}{\tau_0^2}\right)\cdot exp\left(-\frac{\mu_{rt}^j}{\tau_0^3}\right)\\
\gamma_1^j[k] &= \mu_{pdf}^j\cdot exp\left(-\frac{\mu_{Fou}^j}{\tau_1^2}\right)\cdot exp\left(-\frac{\mu_{rt}^j}{\tau_1^3}\right)\cdot exp\left(-\frac{\mu_{au}^j}{\tau_1^4}\right)\\
\gamma_2^j[k] &= \left(1-\left(\mu_{pdf}^j\right)^{2\tau_2^1}\right)\ \cdot\left(1-exp\left(-\frac{\mu_{Fou}^j}{\tau_2^2}\right)\right)\cdot\left(1-exp\left(-\frac{\mu_{rt}^j}{\tau_2^3}\right)\right)\cdot\left(1-exp\left(-\frac{\mu_{au}^j}{\tau_2^4}\right)\right)\\
&\quad\cdot\left(1-exp\left(-\frac{\left(\mu_{diff}^j\right)^2}{\tau_2^5}\right)\right)\\
\gamma_3^j[k] &= \mu_{pdf}^j\cdot\ \left(1-exp\left(-\frac{\mu_{Fou}^j}{\tau_3^2}\right)\right)\cdot\left(1-exp\left(-\frac{\mu_{rt}^j}{\tau_3^3}\right)\right)\cdot exp\left(-\frac{\mu_{au}^j}{\tau_3^4}\right)\cdot exp\left(-\frac{\left(\mu_{diff}^j\right)^2}{\tau_3^5}\right)
\end{aligned}\tag{71}$$

Based on this stochastic process $\mathfrak{B}[p,\ k]$, and all subprocesses $\mathfrak{B}_j[p,\ k]$, we propose a decision function with different thresholds to identify and trigger the proper protection and/or reconstruction mechanism at every time instant k. At this point we are also considering the series $\hat{Y}_{R_{un}}^j$ of predicted states (68), according to the proposed model (see Section 3.1). Time instants are exactly the same to the ones observed in series $\widetilde{Y}_{R_{un}}^j$ of unreconstructed states. For the calculation of this series of estimated states, probabilities $\rho_a,\ \rho_b$ and ρ_c are obtained probabilities γ_2 and γ_3 (69).

$$\hat{Y}_{R_{un}}^j=\left\{\widehat{y_j}[k-r]\ \ r=1,\ldots,R_{un}\right\}\tag{72}$$

$$\begin{aligned}\rho_a&=\rho_c=\gamma_2\\ \rho_b&=\gamma_3\end{aligned}\tag{73}$$

Fig. 4 shows the proposed decision algorithm. This is an original contribution firstly presented in this work. In the first step it is evaluated if any global probability γ_i is θ_{init} units higher than any other probability (70). If that is the case, the situation represented by that probability γ_i is considered to be the actual situation of the last received unreconstructed states $\widetilde{Y}_{R_{un}}^j\forall j$. If γ_0 is the highest probability, states are unperturbed, and they are added with no modification to the series of reconstructed secure states (71). If γ_1 is the highest probability, states are noisy. The reconstruction action depends on how noisy the unreconstructed states are (72). If the Mean Square Error (MSE) between series $\widetilde{Y}_{R_{un}}^j$ and $\hat{Y}_{R_{un}}^j$ (73) is lower than threshold θ_{noise}^{low}, noise is negligible and unreconstructed states $\widetilde{Y}_{R_{un}}^j$ are added with no modification to the series of reconstructed secure states. On the contrary, if the MSE is higher than threshold θ_{noise}^{high}, noise is considered too invasive and next R_{un} reconstructed secure states are taken from the predicted series $\hat{Y}_{R_{un}}^j$. In any other situation, noise is relevant but not dominant, and reconstructed states are calculated as the average between unreconstructed $\widetilde{Y}_{R_{un}}^j$ and predicted $\hat{Y}_{R_{un}}^j$ series. For MSE calculation, predicted values $\hat{Y}_{R_{un}}^j$ are obtained considering all possible perturbation sources (for examples, parameters $a,\ b$ and c takes the most probable value). Finally, if probability γ_2 or probability γ_3 is the highest, the CPS is under an attack (SSA or DoS respectively). In both circumstances, unreconstructed states are not secure and next R_{un} reconstructed secure states are taken from the predicted series $\hat{Y}_{R_{un}}^j$(74). In this last situation, predicted values are obtained in absent of attacks of any kind (i.e., $a,\ b$ and c are null).

$$\gamma_i>\ \gamma_j+\theta_{init}\ \forall\ j\ \neq i\ j,i\ \in[1,3]\tag{74}$$

$$y_j\,[k-r] = \widetilde{y}_j\,[k-r]\ \forall\ r \in [1, R_{un}]\ \forall\ j\ \in [1, M] \tag{75}$$

$$y_j\,[k-r] = \begin{cases} \widetilde{y}_j\,[k-r]\ \ if\ \ MSE <\ \theta_{noise}^{low} \\ \widehat{y}_j\,[k-r]\ \ if\ \ MSE > \theta_{noise}^{high} \qquad \forall r \in [1, R_{un}]\ \forall\ j\ \in [1, M] \\ \dfrac{1}{2}\left(\widetilde{y}_j\,[k-r] + \widehat{y}_j\,[k-r]\right)\ \ otherwise \end{cases} \tag{76}$$

$$MSE = \frac{1}{R_{un} \cdot M} \sum_{j=1}^{M} \sum_{r=1}^{R_{un}} \left(\widetilde{y}_j\,[k-r] - \widehat{y}_j\,[k-r]\right)^2 \tag{77}$$

$$y_j\,[k-r] = \widehat{y}_j\,[k-r]\ \forall\ r \in [1, R_{un}]\ \forall\ j\ \in [1, M] \tag{78}$$

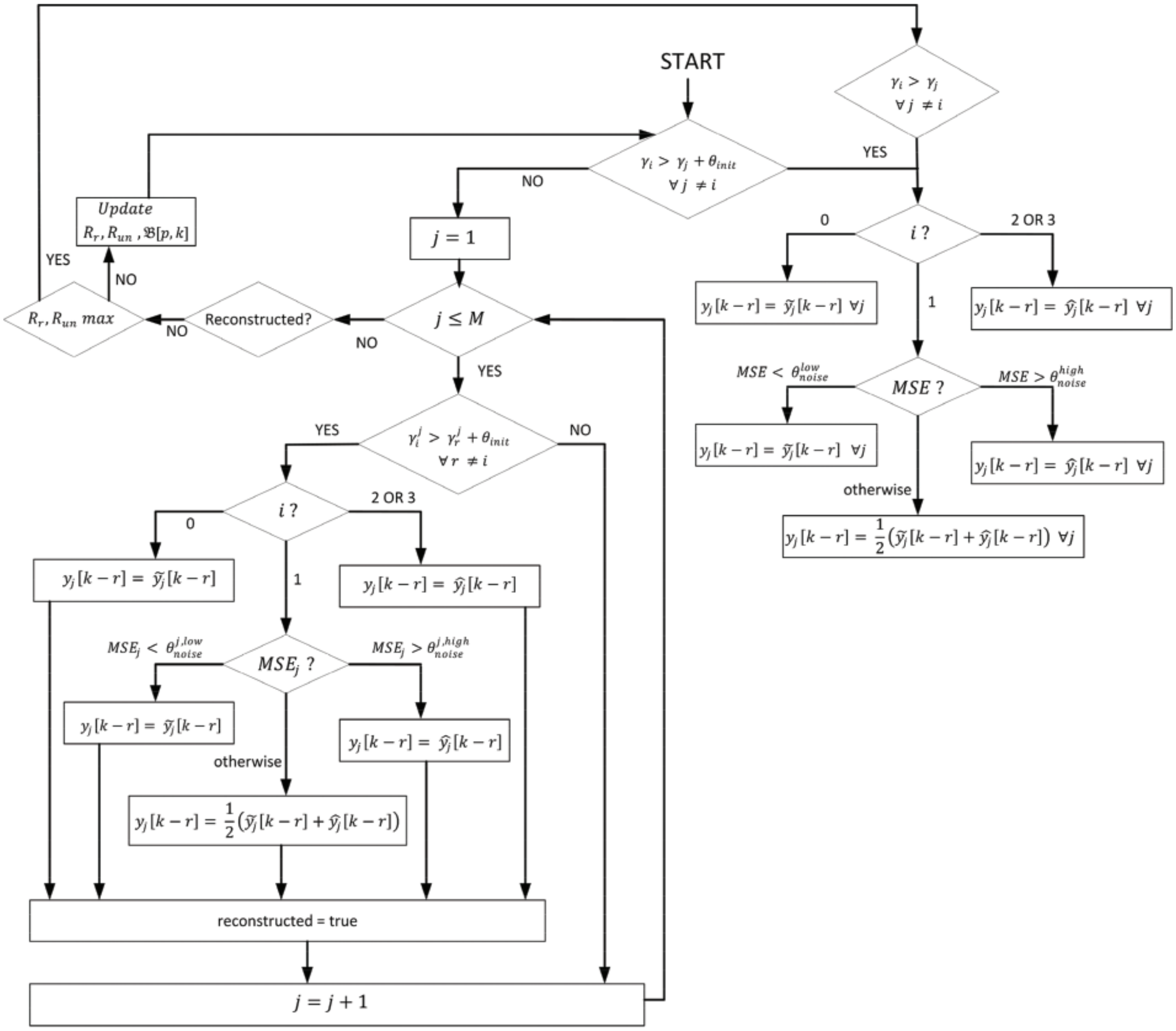
START
$\gamma_i > \gamma_j + \theta_{init}$ $\forall\ j\ \neq i$
NO
YES
$\gamma_i > \gamma_j$ $\forall\ j\ \neq i$
$j = 1$
$Update$ $R_r, R_{un}, \mathfrak{B}[p,k]$
YES
NO
$R_r, R_{un}\ max$
NO
Reconstructed?
NO
$j \leq M$
YES
$\gamma_i^j > \gamma_r^j + \theta_{init}$ $\forall\ r\ \neq i$
YES
NO
$i\ ?$
0
2 OR 3
1
$y_j\,[k-r] = \widetilde{y}_j\,[k-r]$
$y_j\,[k-r] = \widehat{y}_j\,[k-r]$
$MSE_j < \theta_{noise}^{j,low}$
$MSE_j\ ?$
$MSE_j > \theta_{noise}^{j,high}$
$y_j\,[k-r] = \widetilde{y}_j\,[k-r]$
$y_j\,[k-r] = \widehat{y}_j\,[k-r]$
otherwise
$y_j\,[k-r] = \frac{1}{2}(\widetilde{y}_j[k-r] + \widehat{y}_j[k-r])$
reconstructed = true
$j = j + 1$
$i\ ?$
0
2 OR 3
1
$y_j\,[k-r] = \widetilde{y}_j\,[k-r]\ \forall j$
$y_j\,[k-r] = \widehat{y}_j\,[k-r]\ \forall j$
$MSE < \theta_{noise}^{low}$
$MSE\ ?$
$MSE > \theta_{noise}^{high}$
$y_j\,[k-r] = \widetilde{y}_j\,[k-r]\ \ \forall j$
$y_j\,[k-r] = \widehat{y}_j\,[k-r]\ \ \forall j$
otherwise
$y_j\,[k-r] = \frac{1}{2}(\widetilde{y}_j[k-r] + \widehat{y}_j[k-r])\ \forall j$

Figure 4: Reconstruction and protection algorithm

If no global probability γ_i is θ_{init} units higher than any other probability, the same algorithm described above is applied to every $j-th$ state variable. If any probability γ_i^j is θ_{init}^j units higher than any other, this is considered to be the actual situation in the CPS for this state variable (75). The next steps in the algorithm are equivalent to the description above, just using specific thresholds $\theta_{noise}^{j,low}$ and $\theta_{noise}^{j,high}$ for the situation when γ_2^j is the dominant probability (76). The objective, in this case, is to reconstruct the CPS state, variable by variable. This approach is slower and computationally more costly, so it is only triggered when the global analysis is not conclusive.

$$\gamma_i^j > \gamma_r^j + \theta_{init}^j \; \forall \; r \neq i \; r, i \in [1,3] \tag{79}$$

$$y_j[k-r] = \begin{cases} \widetilde{y}_j[k-r] \;\; if \;\; MSE_j < \theta_{noise}^{j,low} \\ \widehat{y}_j[k-r] \;\; if \;\; MSE_j > \theta_{noise}^{j,high} & \forall \; r \in [1, R_{un}] \\ \frac{1}{2}\left(\widetilde{y}_j[k-r] + \widehat{y}_j[k-r]\right) \;\; otherwise \end{cases} \tag{80}$$

$$being$$

$$MSE_j = \frac{1}{R_{un}} \sum_{r=1}^{R_{un}} \left(\widetilde{y}_j[k-r] - \widehat{y}_j[k-r]\right)^2$$

If no probability γ_i^j is θ_{init}^j units higher than any other for any $j-th$ state variable, the stochastic process $\mathfrak{B}[p, \; k]$ is not precise enough. Then, all the algorithm and calculations are repeated for larger values of R_r and R_{un} sizes. Then, results may be more precise when operating with more samples. But, if the proper reconstruction actions could not be selected before the maximum values for R_r and R_{un} sizes are reached, the global algorithm is run one last time. In this last case, the situation represented by the highest probability γ_i (with no restriction) determines the reconstruction action, according to the algorithm described before. In any case, R_r and R_{un} sizes are always returned to the initial values.

Sizes for parameters R_{un} and R_r are actually very important and sensible. Large values for those sizes avoid most spurious numerical and transitory effects, but they reduce the precision and sensitivity of the protection and reconstruction algorithm to detect short-term attacks and high-frequency noises. While reduced values for parameters R_{un} and R_r behave exactly the opposite. The balance cannot be generalized and therefore must be found for every specific application.

4 Experimental Validation

To validate the proposed mechanisms for the protection of Cyber-Physical Systems against Sparse Sensor and Denial of Service attacks, an experimental validation was conducted. Section 4.1 describes the experimental methodology, while Section 4.2 presents the obtained results.

4.1 Experimental Methodology and Environment

The experiments were based on an emulated industrial scenario with real hardware devices (microcontrollers). The experimental works were divided into two different phases. First, we focused on analyzing the precision and attack detection capacity of the proposed technology. The second phase focused on studying the performance and scalability of the proposed model and the reconstruction and protection mechanism.

For all the experiments, the proposed CPS was supported by a collection of ESP-32 microcontrollers. Its number is variable depending on the experiment. ESP-32 microcontrollers are low-cost System-on-Chip provided with Wireless Fidelity (WiFi) and Bluetooth capabilities. It is based on a

Tensilica Xtensa LX6 processor, and it includes several peripheral interfaces (Universal Asynchronous Receiver-Transmitter-UART-, Pulse Width Modulation-PWM-, Serial Peripheral Interface-SPI-, etc.), so it can handle a large catalog of different sensors. In our experiments, each ESP-32 node was provided with two sensors, monitoring four physical variables in total. The first sensor was a CCS811 sensor to monitor air quality. It can provide two different variables: carbon dioxide equivalent (eCO2) and organic volatile compounds concentration (TVOC). The second sensor is a DTH-11 device, which generates measurements for the environmental humidity and temperature. The measurement periodicity is variable and depends on the experiment.

All these sensors employed a WiFi connection to send all the collected information to a cloud server, located within the same building. Hypertext Transfer Protocol (HTTP) messages and Representational State Transfer (REST) interfaces were employed to support these communications. The server was a Linux-based machine (Ubuntu 18.04 LTS) with the following hardware characteristics: Dell R540 Rack 2U, 96 GB RAM, two processors Intel Xeon Silver 4114 2.2G, HD 2 TB SATA 7,2K rpm. In this server, both the proposed model and the reconstruction and protection algorithm were hosted and executed. A Node.js server was deployed to collect all data from the sensor nodes and send them to the computational process executing our proposal. A supervisory process was continuously evaluating the evolution and performance of the proposed algorithms and model. The acquired information was employed to carry out a statistical analysis using the MATLAB 2022a software, to validate our hypotheses. All experiments were repeated twelve times to remove possible spurious effects. The results for every measurement are obtained as the average of all these individual twelve realizations.

In the first phase, we performed two different experiments. The first experiment was aimed at analyzing the precision of the proposed model (Section 3.1) by comparing (using the Mean Square Error metric) the information received by the computational processes in the real CPS deployment and the samples predicted by the proposed model. Data were collected for 24 h, and the relative (percentage) Mean Square Error was calculated for all the acquired samples. The experiment was repeated for different values of parameters $R_{\mathcal{P}}$, $R_{\mathcal{T}}$, and $R_{\mathcal{F}}$, which control the complexity of the proposed model. In this experimental phase, these three parameters are considered to have the same value.

The second experiment in this first phase was aimed at analyzing the probability of the proposed reconstruction and protection algorithm to successfully detect the real situation that occurs in the CPS. Some additional ESP32 nodes were deployed to increment the electrical noise in the environment and/or perform Sparse Sensor and Denial of Service attacks. Different situations were generated, with a duration of ten minutes. It was monitored if the proposed algorithm was able to identify them properly. The second experiment had a duration of 24 h too. Results were processed to generate a confusion matrix representing the algorithm's behavior. The experiment was repeated for different values of R_{un}, and R_r parameters. During these experiments, both parameters had the same value.

In the second experimental phase, we evaluated the performance and scalability. We measured the computational time needed for the proposed model and the reconstruction and protection algorithm to obtain a final and stable output. The first experiment focused on the mathematical model. The calculation time was analyzed for different values of parameters $R_{\mathcal{P}}$, $R_{\mathcal{T}}$, and $R_{\mathcal{F}}$ (all three had the same value) and different quantities of state variables (M). To allow this experiment, the number of sensor nodes in the CPS was increased with each realization. Each configuration was operating for 24 h. The result was obtained as the average of all measurements collected.

Finally, the second experiment in this second phase evaluated the computational time required by the reconstruction and protection algorithm. The experiment was repeated for different values of

R_{un}, and R_r parameters. During these experiments, both parameters had the same value. Different quantities of state variables (M) were also considered. Each configuration was operating for 24 h. The result was obtained as the average of all collected measurements.

4.2 Results

To evaluate the behavior of the proposed technology, first, we analyze the precision of our model (Section 3.1), comparing the predicted future CPS states and the actual state finally achieved. Fig. 5 shows the results. As can be seen, the evolution is exponential, as expected from the error in the Taylor series, as the number of terms increases. In general, all configurations show good behavior, although models with only two terms introduce an error of 12% (which may be too high for some applications). The minimum error (2%) may be achieved for models with more than 12 terms. This error is caused by the truncation of the Taylor series, so they can be numerically computed. But, as a counterpart, the resulting finite series does not perfectly represent the original function and we are introducing a numerical error.

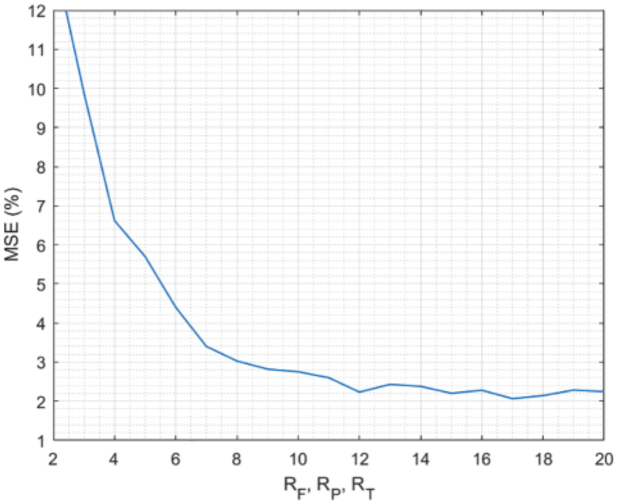

Figure 5: Precision of the proposed model

Other limitations in the proposed model (such as the numerical precision of the underlying hardware platform) are also affecting, so, only by increasing the number of terms in Taylor's series cannot reduce the global error as much as desired. But for a very large catalog of applications, an error of 2% is acceptable and can be tolerated. Even, for those scenarios where computationally lightweight solutions are preferred, schemes with four or five terms generate an error of around 6%, which is a standard error for mass non-critical applications. In common applications, errors below 10% can be handled. From these results, we can conclude that the proposed model represents with good precision the physical processes in CPS.

Similarly, we need to analyze the capacity of the proposed reconstruction algorithm to successfully detect the real situation that is happening in the CPS. Fig. 6 shows the results of this experiment. For all possible situations, three regions are identified. First, for low values of R_{un} and R_r parameters (below 20), transitory effects are dominant, indicators do not capture properly the CPS behavior, and sensitivity (rate of situation correctly classified) decreases. Random natural variations may be relevant when very short periods are analyzed. To focus on global tendencies, larger collections of data samples are needed. In that way, later, in the central region, R_{un} and R_r parameters present good enough values (between 20 and 60), and the proposed algorithm works with a very satisfactory behavior (sensitivity is between 91% and 98%). In this region, short-time transitory effects are not dominant because larger time series are employed in the reconstruction mechanism, and global tendencies are easily detected. But when the values for R_{un} and R_r parameters increase beyond a certain limit (60 samples in this case), real fluctuations effects and high-frequency perturbations are ignored when they are aggregated in a large operation. Then, sensitivity decreases. In this region, even natural fluctuations and changes are not significant compared to long-term tendencies. Relevant changes are ignored, because we are integrating too many samples in the same series, and calculation algorithms do not have enough sensitivity.

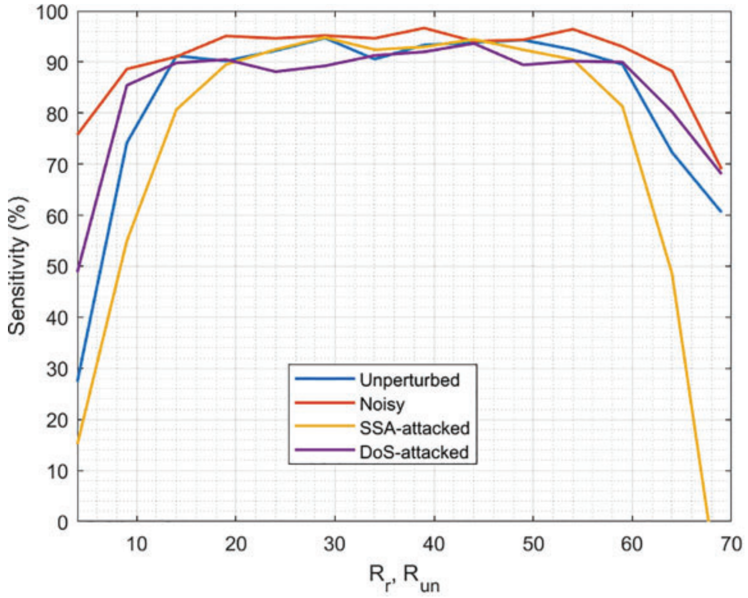

Figure 6: Precision of the proposed reconstruction and protection

In conclusion, R_{un} and R_r parameters must be balanced: small values cause instabilities, while too big values generate a loss of sensitivity. Values between 20 and 60 are the most appropriate region, as shown above (Fig. 5).

On the other hand, the proposed algorithm does not show the same sensitivity when detecting the different situations in a CPS. In general, situations whose probability is calculated using functions with a higher growth rate (such as the exponential) are more sensitive to the quality of indicators representing the CPS (first return maps, STFT, etc.) and then more sensible to the value of R_{un} and R_r parameters. This is because small changes in the exponents may generate big variations in the final function. Values must be selected very carefully and according to previous observations. For example, the probability for the SSA-attacked situation is only supported by an exponential function, so it is the one with the most relevant fluctuations. The noisy situation, which includes a linear term, is much flatter. That means SSA-attacked situations are much more difficult to detect, and probably a heuristic calibration process is required in real deployments and scenarios.

Anyway, the sensitivity of the proposed algorithm (in balanced values of R_{un} and R_r parameters) is very satisfactory. Noisy situations are detected on 98% of the occasions, while unperturbed and SSA-attacked situations are correctly identified on 92% of the times. The DoS-attacked situation is the one with the worst behavior, but its sensitivity is just slightly lower: 91%. With these results, we can conclude that the proposed algorithm can reconstruct and protect Cyber-Physical Systems against attacks and perturbations.

To go deeper into the analysis of these data, we present the complete confusion matrix (Table 2) for the configuration $R_{un} = R_r = 40$.

Table 2: Confusion matrix for the configuration $\boldsymbol{R_{un}} = \boldsymbol{R_r} = 40$

Detected situation	Real situation			
	Unperturbed	Noisy	SSA-attacked	DoS-attacked
Unperturbed	92.89	2.43	0.84	1.43
Noisy	3.31	96.66	3.98	5.85
SSA-attacked	1.44	0.43	92.96	0.80
DoS-attacked	2.36	0.48	2.22	91.92

As can be seen, most errors when identifying the situation in the CPS are false detections of the noisy situation. Probably, that is caused by the linear term in its probability function, which does not reduce its value as much as the exponential function. If this sensitivity needs to be improved, that linear term should be enriched with new indicators and functions.

It is also important to evaluate the performance and scalability of the proposed solution to identify its limitations. Fig. 7 shows the computational time required for the proposed model to operate.

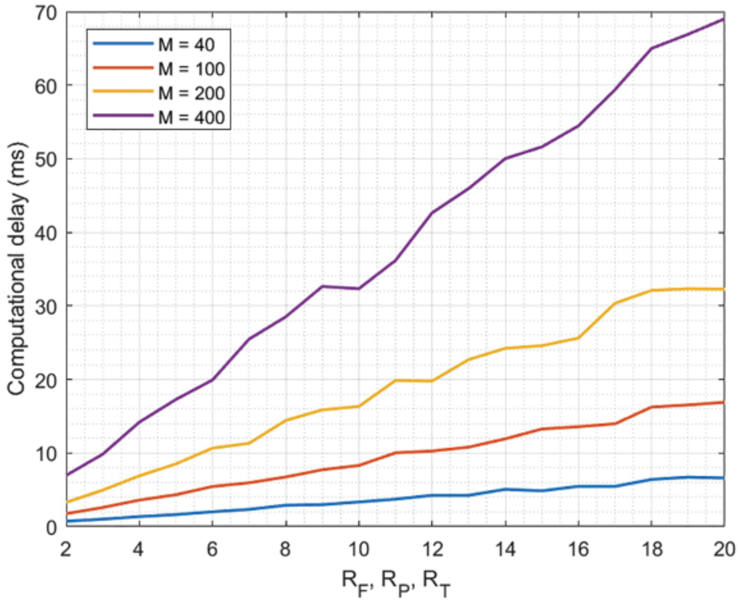

Figure 7: Computational delay and scalability (mathematical model)

As can be seen, the evolution of the computational delay is linear. This is because our model consists of additions and multiplications, without loops or recursive problems. Besides, each new state variable is independent of the others, so the increase is linear. This facilitates the employment of this protection and reconstruction solution in future dense and pervasive scenarios, where up to ten million sensors per square kilometer could be deployed.

Moreover, the delay is always in the range of milliseconds. Additions and multiplications are performed very efficiently on modern computers, and they require a short time to perform millions of operations. In this case, even for a CPS that includes 100 devices (i.e., 400 state variables) and very complex models (with almost 20 terms in the Taylor series), the computational time required to operate the model is below 100 milliseconds (70 milliseconds, to be precise). Considering the most usual Cyber-Physical Systems capture information from the environment every few seconds, this delay is satisfactory.

Finally, the same scalability and performance analysis must be applied to the reconstruction and protection algorithm. Fig. 8 shows the results. Here, again, the evolution is almost linear, because all the proposed computational procedures do not require any loop or recursive processing. In this case, for the largest deployment (one hundred devices) and a typical value for R_{un} and R_r parameters, the delay is above one second. This may be slightly above the acceptable maximum for certain critical real-time applications. For smaller deployments and the same configuration, the delay is below one second (between 100 and 600 milliseconds). But even delays above one second are acceptable in mass non-critical applications, where data are acquired every few seconds. Due to linear evolution, scalability is guaranteed (even in future scenarios) as consumed resources grow at the same rate as the number of sensor nodes in the physical platform.

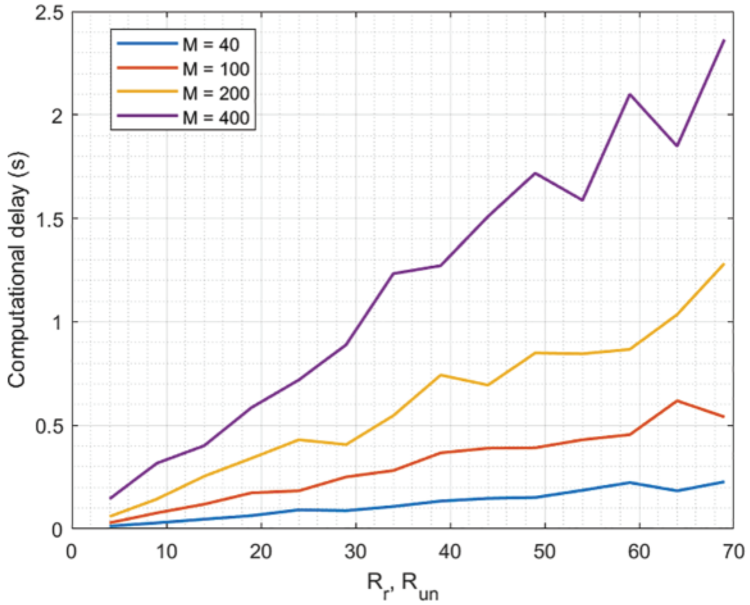

Figure 8: Computational delay and scalability (reconstruction and protection algorithm)

In conclusion, considering the limitations that may arise in critical real-time applications, the performance of the proposed reconstruction and protection mechanism is satisfactory.

5 Conclusions and Future Works

This paper presents a new stochastic model to represent the behavior of Cyber-Physical Systems precisely. This model includes unknown multivariate discrete and continuous-time functions and different multiplicative noises to represent the evolution of physical processes and random effects in the physical and computational worlds. As a novelty, in this model, engineered processes such as the digitalization stage are represented too. Additionally, and contrary to the commonly employed deterministic attackers, in this new model attackers are described through a stochastic process. Standard error sources are estimated through different indicators and non-linear techniques (such as the Fourier transform, first-return maps, or the probability density function). Finally, the reconstruction mechanism consists of a weighted stochastic model combining all error sources. The actual reconstructed value is generated as the output from a decision algorithm.

Experimental results show that the precision of the proposed model is above 90%, with a residual error between 6% and 2% for the most common configurations. Additionally, the sensitivity of the proposed reconstruction and protection algorithm is up to 92%. Considering all this, the proposed solution is a valid security scheme for CPS.

Future works will analyze new indicators and probability functions to improve the sensitivity, especially in noisy situation. In addition, the solution will be deployed in real industrial scenarios with legacy systems, to study the impact of second-order effects such as reduced connectivity or human accidents and manipulations.

Acknowledgement: The authors also gratefully acknowledge the helpful comments and suggestions of the reviewers, which have improved the presentation.

Funding Statement: This work is supported by Comunidad de Madrid within the framework of the Multiannual Agreement with Universidad Politécnica de Madrid to encourage research by young doctors (PRINCE).

Author Contributions: The authors confirm contribution to the paper as follows: study conception and design: Borja Bordel; data collection: Ramón Alcarria, Borja Bordel; analysis and interpretation of results: Ramón Alcarria, Tomás Robles; draft manuscript preparation: Borja Bordel, Tomás Robles. All authors reviewed the results and approved the final version of the manuscript.

Availability of Data and Materials: Data sharing is not applicable to this article as no new data were created or analyzed in this study.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

[1] S. Zanero, “Cyber-physical systems,” *Computer*, vol. 50, no. 4, pp. 14–16, 2017.

[2] S. Yin, J. J. Rodriguez-Andina and Y. Jiang, “Real-time monitoring and control of industrial cyberphysical systems: With integrated plant-wide monitoring and control framework,” *IEEE Industrial Electronics Magazine*, vol. 13, no. 4, pp. 38–47, 2019.

[3] F. C. Delicato, A. Al-Anbuky, I. Kevin and K. Wang, “Smart cyber–physical systems: Toward pervasive intelligence systems,” *Future Generation Computer Systems*, vol. 107, pp. 1134–1139, 2020.

[4] B. Bordel, R. Alcarria, T. Robles and A. Sánchez-Picot, “Stochastic and information theory techniques to reduce large datasets and detect cyberattacks in ambient intelligence environments,” *IEEE Access*, vol. 6, pp. 34896–34910, 2018.

[5] E. A. Lee, “Cyber-physical systems-are computing foundations adequate,” *Position Paper for NSF Workshop on Cyber-Physical Systems: Research Motivation, Techniques and Roadmap*, vol. 2, pp. 1–9, 2006.

[6] N. Negi and A. Chakrabortty, "Co-design of delays and sparse controllers for bandwidth-constrained cyber-physical systems," in *2020 American Control Conf. (ACC)*, Denver, CO, USA, IEEE, pp. 987–992, 2020.

[7] N. Wang and X. Li, "Secure synchronization control for a class of cyber-physical systems with unknown dynamics," *IEEE/CAA Journal of Automatica Sinica*, vol. 7, no. 5, pp. 1215–1224, 2020.

[8] V. Bolbot, G. Theotokatos, L. M. Bujorianu, E. Boulougouris and D. Vassalos, "Vulnerabilities and safety assurance methods in cyber-physical systems: A comprehensive review," *Reliability Engineering & System Safety*, vol. 182, pp. 179–193, 2019.

[9] Y. Shoukry and P. Tabuada, "Event-triggered state observers for sparse sensor noise/attacks," *IEEE Transactions on Automatic Control*, vol. 61, no. 8, pp. 2079–2091, 2015.

[10] H. Wang, X. Wen, Y. Xu, B. Zhou, J. Peng *et al.,* "Operating state reconstruction in cyber physical smart grid for automatic attack filtering," *IEEE Transactions on Industrial Informatics*, vol. 18, no. 5, pp. 2909–2922, 2020.

[11] M. Al-Sharman, D. Murdoch, D. Cao, C. Lv, Y. Zweiri *et al.,* "A sensorless state estimation for a safety-oriented cyber-physical system in urban driving: Deep learning approach," *IEEE/CAA Journal of Automatica Sinica*, vol. 8, no. 1, pp. 169–178, 2020.

[12] A. Akbarzadeh, P. Pandey and S. Katsikas, "Cyber-physical interdependencies in power plant systems: A review of cyber security risks," in *2019 IEEE Conf. on Information and Communication Technology (CICT)*, Allahabad, India, IEEE, pp. 1–6, 2019.

[13] H. Yang, S. Yin, H. Han and H. Sun, "Sparse actuator and sensor attacks reconstruction for linear cyber-physical systems with sliding mode observer," *IEEE Transactions on Industrial Informatics*, vol. 18, no. 6, pp. 3873–3884, 2021.

[14] D. Zhang, Q. G. Wang, G. Feng, Y. Shi and A. V. Vasilakos, "A survey on attack detection, estimation and control of industrial cyber–physical systems," *ISA Transactions*, vol. 116, pp. 1–16, 2021.

[15] H. Karimipour and H. Leung, "Relaxation-based anomaly detection in cyber-physical systems using ensemble kalman filter," *IET Cyber-Physical Systems: Theory & Applications*, vol. 5, no. 1, pp. 49–58, 2020.

[16] T. D. Memon, "FPGA implementation of extended kalman filter for parameters estimation of railway wheelset," *Computers, Materials & Continua*, vol. 74, no. 2, pp. 3351–3370, 2022.

[17] Z. Lv, D. Chen, R. Lou and A. Alazab, "Artificial intelligence for securing industrial-based cyber–physical systems," *Future Generation Computer Systems*, vol. 117, pp. 291–298, 2021.

[18] M. S. Miah, M. Zhu, A. Granados, N. Sharmin, I. Anjum *et al.,* "Optimizing honey traffic using game theory and adversarial learning," in *Cyber Deception: Techniques, Strategies, and Human Aspects*. Cham: Springer, pp. 97–124, 2022.

[19] A. Wani and R. Khaliq, "SDN-based intrusion detection system for IoT using deep learning classifier (IDSIoT-SDL)," *CAAI Transactions on Intelligence Technology*, vol. 6, no. 3, pp. 281–290, 2021.

[20] J. P. A. Yaacoub, O. Salman, H. N. Noura, N. Kaaniche, A. Chehab *et al.,* "Cyber-physical systems security: Limitations, issues and future trends," *Microprocessors and Microsystems*, vol. 77, pp. 103201–103234, 2020.

[21] D. Ding, Q. L. Han, X. Ge and J. Wang, "Secure state estimation and control of cyber-physical systems: A survey," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 51, no. 1, pp. 176–190, 2020.

[22] H. Fawzi, P. Tabuada and S. Diggavi, "Secure estimation and control for cyber-physical systems under adversarial attacks," *IEEE Transactions on Automatic Control*, vol. 59, no. 6, pp. 1454–1467, 2014.

[23] A. Y. Lu and G. H. Yang, "Observer-based control for cyber-physical systems under denial-of-service with a decentralized event-triggered scheme," *IEEE Transactions on Cybernetics*, vol. 50, no. 12, pp. 4886–4895, 2019.

[24] L. Ma, Z. Wang, Q. L. Han and H. K. Lam, "Variance-constrained distributed filtering for time-varying systems with multiplicative noises and deception attacks over sensor networks," *IEEE Sensors Journal*, vol. 17, no. 7, pp. 2279–2288, 2017.

[25] M. S. Mahmoud, M. M. Hamdan and U. A. Baroudi, “Modeling and control of cyber-physical systems subject to cyber attacks: A survey of recent advances and challenges,” *Neurocomputing*, vol. 338, pp. 101–115, 2019.

[26] L. An and G. H. Yang, “Distributed secure state estimation for cyber–physical systems under sensor attacks,” *Automatica*, vol. 107, pp. 526–538, 2019.

[27] X. Xie, Z. Yang and X. Mu, “Observer-based consensus control of nonlinear multiagent systems under semi-Markovian switching topologies and cyber attacks,” *International Journal of Robust and Nonlinear Control*, vol. 30, no. 14, pp. 5510–5528, 2020.

[28] V. S. Gaur, V. Sharma and J. McAllister, “Abusive adversarial agents and attack strategies in cyber-physical systems,” *CAAI Transactions on Intelligence Technology*, vol. 8, no. 1, pp. 149–165, 2023.

[29] W. He, Z. Mo, Q. L. Han and F. Qian, “Secure impulsive synchronization in Lipschitz-type multi-agent systems subject to deception attacks,” *IEEE/CAA Journal of Automatica Sinica*, vol. 7, no. 5, pp. 1326–1334, 2020.

[30] M. U. Danjuma, B. Yusuf and I. Yusuf, “Reliability, availability, maintainability, and dependability analysis of cold standby series-parallel system,” *Journal of Computational and Cognitive Engineering*, vol. 1, no. 4, pp. 193–200, 2022.

[31] A. S. Maihulla, I. Yusuf and S. I. Bala, “Reliability and performance analysis of a series-parallel system using Gumbel–Hougaard family copula,” *Journal of Computational and Cognitive Engineering*, vol. 1, no. 2, pp. 74–82, 2022.

[32] H. Modares, B. Kiumarsi, F. L. Lewis, F. Ferrese and A. Davoudi, “Resilient and robust synchronization of multiagent systems under attacks on sensors and actuators,” *IEEE Transactions on Cybernetics*, vol. 50, no. 3, pp. 1240–1250, 2019.

[33] R. Moghadam and H. Modares, “Resilient autonomous control of distributed multiagent systems in contested environments,” *IEEE Transactions on Cybernetics*, vol. 49, no. 11, pp. 3957–3967, 2018.

[34] B. Bordel, R. Alcarria, D. Martín and D. Sánchez-de-Rivera, “An agent-based method for trust graph calculation in resource constrained environments,” *Integrated Computer-Aided Engineering*, vol. 27, no. 1, pp. 37–56, 2020.

[35] L. Zhang, X. Chen, F. Kong and A. A. Cardenas, “Real-time attack-recovery for cyber-physical systems using linear approximations,” in *2020 IEEE Real-Time Systems Symp. (RTSS)*, Houston, TX, USA, IEEE, pp. 205–217, 2020.

[36] Z. Guo, K. Yu, Z. Lv, K. K. R. Choo, P. Shi *et al.,* “Deep federated learning enhanced secure POI microservices for cyber-physical systems,” *IEEE Wireless Communications*, vol. 29, no. 2, pp. 22–29, 2022.

[37] A. Chakraborty, M. Alam, V. Dey, A. Chattopadhyay and D. Mukhopadhyay, “A survey on adversarial attacks and defences,” *CAAI Transactions on Intelligence Technology*, vol. 6, no. 1, pp. 25–45, 2021.

[38] I. Hidayat, M. Z. Ali and A. Arshad, “Machine learning-based intrusion detection system: An experimental comparison,” *Journal of Computational and Cognitive Engineering*, vol. 2, pp. 88–97, 2022.

[39] B. A. Alqaralleh, F. Aldhaban, E. A. AlQarallehs and A. H. Al-Omari, “Optimal machine learning enabled intrusion detection in cyber-physical system environment,” *Computers, Materials & Continua*, vol. 72, no. 3, pp. 4691–4707, 2022.

[40] P. Dai, W. Yu, H. Wang, G. Wen and Y. Lv, “Distributed reinforcement learning for cyber-physical system with multiple remote state estimation under DoS attacker,” *IEEE Transactions on Network Science and Engineering*, vol. 7, no. 4, pp. 3212–3222, 2020.

[41] Y. Lu, X. Huang, Y. Dai, S. Maharjan and Y. Zhang, “Federated learning for data privacy preservation in vehicular cyber-physical systems,” *IEEE Network*, vol. 34, no. 3, pp. 50–56, 2020.

[42] J. Zhou, B. Chen, T. Li and L. Yu, “Secure estimation against non-fixed channel attacks in cyber-physical systems,” *International Journal of Robust and Nonlinear Control*, vol. 33, no. 3, pp. 2496–2507, 2023.

[43] L. Zhang, Y. Chen and M. Li, “Resilient predictive control for cyber–physical systems under denial-of-service attacks,” *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 69, no. 1, pp. 144–148, 2021.

[44] M. Zhang and C. Lin, "Secure state estimation for cyber physical systems with state delay and sparse sensor attacks," *Systems Science & Control Engineering*, vol. 9, pp. 71–80, 2021.

[45] A. Y. Lu and G. H. Yang, "Detection and identification of sparse sensor attacks in cyber physical systems with side information," *IEEE Transactions on Automatic Control*, vol. 2022, pp. 1–15, 2022.

[46] D. Ding, Z. Wang, G. Wei and F. E. Alsaadi, "Event-based security control for discrete-time stochastic systems," *IET Control Theory & Applications*, vol. 10, no. 15, pp. 1808–1815, 2016.

[47] S. Nateghi, Y. Shtessel, R. J. Rajesh and S. S. Das, "Control of nonlinear cyber-physical systems under attack using higher order sliding mode observer," in *2020 IEEE Conf. on Control Technology and Applications (CCTA)*, Montreal, Canada, IEEE, pp. 1–6, 2020.