IET Cyber-Physical Systems: Theory & Applications

Research Article

IET Journals
The Institution of Engineering and Technology

Covert Channels in Stochastic Cyber-Physical Systems

ISSN 1751-8644
doi: 0000000000
www.ietdl.org

Walter Lucia[1]*, *Amr Youssef*[1]

[1] *Concordia Institute for Information Systems Engineering (CIISE), Concordia Unviersity, Montreal, Canada*
* *E-mail: walter.lucia@concordia.ca*

Abstract: A covert channel is a communication channel that is not intended to exist, and that can be used to transfer information in a manner that violates the system security policy. Attackers can abuse such channel to exfiltrate sensitive information from cyber-physical systems (CPSs), e.g., to leak the confidential or proprietary parameters in a control system. Furthermore, attacks against CPSs can exploit the leaked information about the implementation of the control system, e.g., to determine optimal false data injection attack values that degrade the system performance while remaining undetected. In this paper, we present a control theoretic approach for establishing covert channels in stochastic CPSs. In particular, we consider a scenario where an attacker is able to inject malware into the networked controller and arbitrarily alter the control logic. By exploiting such capability, an attacker can establish an illegitimate communication channel, e.g., to transmit sensitive plant parameters, between the networked controller and an eavesdropper intercepting the sensor measurements. We show that such a channel can be established by exploiting the closed-loop system operations, a decoding mechanism based on an unknown input observer, and an error-correcting coding scheme that exploits the control loop to obtain an implicit acknowledgement. A simple proof of concept implementation of the covert channel is presented, and its performance is evaluated by resorting to a numerical example. Finally, we propose some defences and countermeasures against the proposed covert channel.

1 Introduction

The development of cyber-physical systems (CPSs) aims to improve the capabilities of traditional engineering systems by introducing advanced computational capacity and communications among system entities. On the other hand, the adoption of such technologies introduces a threat and exposes the system to cyber-attacks. Given the unique properties of CPSs, e.g., physically interacting with its environment, malicious parties might be interested in exploiting the physical properties of the system in the form of a cyber-physical attacks [1]. Attacks on CPSs have been investigated from different angles [2]. Since in a large class of CPSs the physical systems are controlled using a feedback control loop, the majority of previous research works have investigated control theoretic techniques to detect and mitigate cyber-attacks occurrences affecting the communication channels between the plant and the controller [3, 4]. Recently, the privacy of cyber-physical systems has received increasing attention and different privacy preserving solutions against eavesdropping attacks have been proposed [5–7]. On the other hand, the problem of securing CPSs against intruders targeting the control algorithm operations to covertly leak sensitive information from CPSs, such as confidential/proprietary gains in a control system, has not received sufficient attention. This leaked information can be exploited by attackers to launch further attacks against CPSs, e.g., to determine optimal false data injection attack values that degrade the system performance while remaining undetected.

In this paper, we take a step towards this direction and we investigate the existence of covert channels in stochastic CPSs. This problem is of particular interest for the CPS community because it can potentially enable an internal adversary (e.g., malware in the controller) to transmit sensible plant information to an external entities while bypassing existing attack detection mechanisms and privacy preserving solutions.

1.1 Related works

A covert channel is an evasion technique that aims to illicitly transfer information in an unauthorized and secretive manner that violates the existing security policy. While encryption can be used to protect communication from being decoded by an unauthorized party [8], a covert channel aims to hide the very existence of such communication. During the investigation of monolithic systems, Lampson [9] introduced covert channels as a mechanism by which a process at a high-security level can leak information to a process at a low-security level. In particular, Lampson defined covert channels as "channels not intended for information transfer at all". Similarly, the US Department of Defence (DoD) define covert channels as "any communication channels that can be exploited by a process to transfer information in a manner that violates the system's security policy" [10]. According to the "Orange Book" [11] of the US DoD, system developers shall conduct a thorough search for covert channels and make a determination, either by actual measurement or by engineering estimation, of the maximum bandwidth of each identified channel. Also, the continued existence of identified covert channels in the system must be justified.

In the context of Information Technology networks, covert channels can be established by abusing different communication protocols and shared computing/storage resources. Different methodologies have been proposed to establish covert channels, e.g. see the timing-based covert channels and storage-based covert channels in [12]. As the name entails, timing based channels utilize delay to separate bits of information shared between two malicious parties. On the other hand, storage-based channels utilize shared storage or memory resources that are not designed to transfer data. Covert channels have also been used to send information from air-gapped machines, by encoding information over physical infrastructures that cannot be noticed with naked senses such as inaudible speaker sounds, acoustical mesh, and optical emanations [13–18].

1.1.1 Covert Channels in CPSs: The existence of covert channels in CPSs and its associated vulnerabilities have been investigated by many authors. In [19, 20], Chhetri et al. demonstrated how CPSs are prone to covert information leakage from the physical domain. In particular, they explained how analog emissions such as vibration, acoustic, magnetic, and power could allow attackers (monitoring these analog emissions) to determine the correlation between the observed phenomena and the cyber-domain data. As a consequence, the attackers can leverage this relation to breach

© The Institution of Engineering and Technology 2015

the confidentiality of the system. The authors also presented a 3D-Printer proof of concept case study where they demonstrated how acoustic analog emissions could reveal parameters such as speed, direction, axis, and extrusion. Uluagac et al. [21] showed that using sensory channels such as light, temperature, and infrared, an adversary can trigger existing malware, transfer malware, or combine malicious use of different sensory channels to increase the impact of the attack on CPS devices. They also introduced the design of a sensory channel-aware intrusion detection system as a protection mechanism against sensory channel threats for CPSs. Another example of covert channels in CPS, that borrows the idea of an air-gapped receiver, is described in [22] where the adversary is assumed to be able to load a malicious code onto a Programmable Logic Controller (PLC) to change actuation signals being output to the motors. The actuation signal is then perturbed to transmit sensitive information covertly by creating analog acoustic channel signatures without changing the closed-loop process characteristics. In [23], a covert channel specifically designed against power grid cyber-physical critical infrastructures through physical substrates, e.g., line loads, is proposed. Using this approach, two compromised controllers that are miles apart can coordinate their efforts by manipulating relays to modify the power network's topology. In [24], Wendzel et al. studied the threat of covert channels in building automation systems (BAS) protocols. The authors presented network covert storage and network covert timing channels in the network and application layer of the BACnet protocol stack to show that protocol-level covert channels in BAS are feasible. In [25], Alcaraz et al. addressed the security issues related to covert channels applied to industrial networks, identifying new vulnerability points when ITs converge with operational technologies such as edge computing infrastructures. Specifically, the authors defined two signalling strategies where they exploit the Modbus/transmission control protocol (TCP) as the target to set up a covert channel. The authors also discussed some possible mitigation and defensive measures. In [26], [27], information flow analysis techniques were used to analyze covert channels in CPSs. In particular, Gamage et al. [26] presented a general theory of event compensation as an information flow security enforcement mechanism for CPSs. The fundamental research problem being investigated is that externally observable events in modern CPSs have the propensity to covertly divulge sensitive settings to adversaries, resulting in a confidentiality violation. To mitigate such violations, the authors proposed to use information flow security-based enforcement mechanisms since access control-based security models cannot impose restrictions on information propagation. The proposed framework unifies cyber and physical aspects of security through the shared semantics of information flow. Along a similar line of research, Akella et al. [27] applied classical models of non-deducibility and non-inference to CPSs to determine information flow in the coupled cyber and physical worlds. The presented results demonstrate that the combined physical and cyber properties of a CPS can both protect and divulge information. The authors also presented a semantic model for information flow analysis in a CPS and described an approach to perform the analysis, including both trace-based analysis and automated analysis through process algebra specification. In [28, 29], the authors turned upside-down the originally malicious concept of covert channels and utilized it to build defensive mechanisms. Specifically, Ying et al. [28] presented TACAN (Transmitter Authentication in insecure Controller Area Network (CAN)), which provides secure authentication of Electronic Control Units (ECUs) by exploiting the covert channels without introducing CAN protocol modifications or traffic overheads. Similarly, Taylor et al. [29] demonstrated the use of covert channels as a method of secure communication that would prevent a number of attacks, including man-in-the-middle., against the Modbus protocol.

Of particular interest for this work are covert channels solutions leveraging control theory models [30–32]. In [30], Herzberg and Kfir presented a unidirectional covert channel from a malicious sensor to a malicious actuator. The covert traffic is encoded within the output noise of the covertly transmitting sensor, whose distribution is indistinguishable from that of a benign sensor with comparable specifications. In [31], the same authors presented a malicious actuator that receives commands from a threshold controller. The corrupt actuator uses the response time to send signals to a corrupt sensor, by encoding the signals using different response times of the actuator. In [32], the authors presented a covert channel technique utilizing a robust control-theoretic approach for CPSs with bounded disturbances (also see [33]). The presented technique enables a compromised networked controller to leak information to an eavesdropper who has access to the measurement channel by properly altering the control logic and exploiting robust reachability arguments.

1.2 Contribution

The approach in [32] leverages robust reachability arguments to show the existence of covert channels for constrained CPSs subject to bounded disturbances. However, this approach cannot be applied to the class of stochastic CPSs where only the distribution of the disturbance is known. In the present work, we address this limitation and design a covert channel in stochastic CPSs by leveraging, as decoding mechanism, an Unknown Input Observer (UIO) coupled with an error correcting code. Moreover, differently from [30, 31], the covert channel considered in this work is established without any assumption on the sensors and actuators' hardware characteristics. To the best of the authors' knowledge, this paper shows for the first time that it is possible to establish a covert channel by combining an UIO and error-correcting code schemes. Finally, a proof-of-concept implementation of the proposed covert channel is presented with the aim to evaluate its information rate for a numerical testbed.

The rest of the paper is organized as follows. The system setup, adversary model and problem formulation are described in section 2. The proposed covert channel is described in section 3 where we also present an error-correcting coding scheme that exploits the control loop to obtain an implicit acknowledgement to improve the covert channel capacity. A proof of concept implementation and numerical example are described in in section 4 and section 5, respectively. Defences and countermeasures against the proposed covert channel are discussed in section 6. Finally, the paper is concluded in section 7.

2 Problem Formulation

In this section, first the considered CPS and the adversary model are presented. Then, the considered covert channel problem is formally stated.

2.1 System Setup

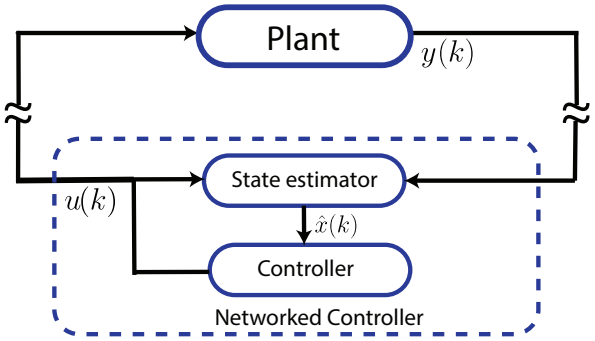

Fig. 1: Cyber-physical system setup.

Consider the following Linear Time Invariant (LTI) stochastic system

$$\begin{aligned} x(k+1) &= Ax(k) + Bu(k) + w(k) \\ y(k) &= Cx(k) + v(k) \end{aligned} \tag{1}$$

© The Institution of Engineering and Technology 2015

where $k \in \mathbb{Z}_+ := \{0, 1, \ldots\}$ denotes the discrete sampling time instants, $x(k) \in \mathbb{R}^{n_x}$ the plant state vector, $u(k) \in \mathbb{R}^{m_u}$ the control input vector, $A \in \mathbb{R}^{n_x \times n_x}$, $B \in \mathbb{R}^{n_x \times m_u}$, $C \in \mathbb{R}^{p \times n_x}$, Moreover, $w(k) \sim \mathcal{N}(0, \mathcal{W})$ and $v(k) \sim \mathcal{N}(0, \mathcal{V})$ are mutually independent and identically distributed (i.i.d) Gaussian noises with zero mean and covariance matrix $\mathcal{W}$ and $\mathcal{V}$, respectively.
By referring to Fig. 1, we consider a CPS where the plant is regulated by a networked controller consisting of a steady-state Kalman predictor (state estimator) and a state-feedback controller. The Kalman predictor is described by the following dynamical system

$$\hat{x}(k) = A\hat{x}(k-1) + Bu(k-1) + L(y(k-1) - C\hat{x}(k-1)) \quad (2)$$

where $\hat{x}(k)$ is the estimated state vector. Moreover, $L = APC^T(CPC^T + \mathcal{W})^{-1}$ is the steady-state Kalman gain where $P = P^T \geq 0$ is the only positive semi-definite solution of the Riccati equation $P = APA^T + \mathcal{W} - APC^T(CPC^T + \mathcal{V})^{-1}CPA^T$.

The state feedback controller is designed to stabilize the closed-loop system and its actions are generically modeled as follows:

$$u(k) = f(\hat{x}(k)), \quad f(\cdot) : \mathbb{R}^{n_x} \to \mathbb{R}^{m_u} \quad (3)$$

2.2 *Adversary Model*

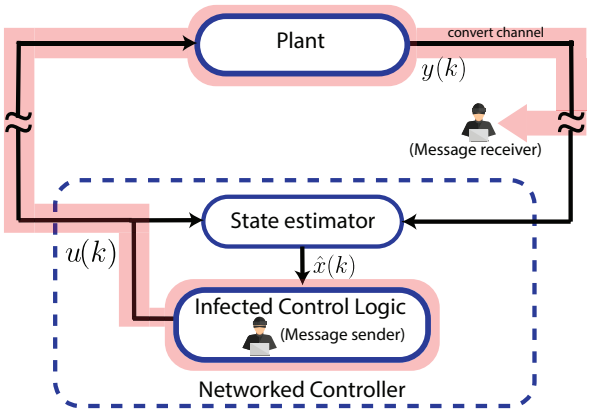

Fig. 2: Covert Channel.

We consider an attacker aiming to establish an illegitimate covert channel to leak information between a sender located inside the controller and a receiver having access to the transmitted measurements (see Fig. 2). To perform such an attack, we assume that the attacker possesses the following assets [34]:

- *Model Knowledge*: the attacker has knowledge of the plant's model (1);
- *Disruptive resources*: the attacker is capable of injecting malware in the networked controller and arbitrarily changing the control logic (3);
- *Disclosure resources*: the attacker can read the transmitted sensor measurement $y(k)$.

2.3 *Problem Formulation*

The covert channel design problem can be stated as follows:

Given the LTI stochastic plant model (1)*, design the sender and receiver actions such that a binary vector* $M = [m_1, \ldots, m_l]^T \in \mathbb{R}^l$, $m_i \in \{0, 1\}$, $1 \leq i \leq l$, *can be sequentially encoded in the control action* $u(k)$ *and decoded from the sensor measurements* $y(k)$.

In this paper, a solution to the above problem is given under the assumption that an unknown input observer for (1) can be defined to simultaneously estimate the state $x(k)$ and the input signal $u(k)$ from the sensor measurement $y(k)$.

3 Covert Channels Design

This section first shows how, in the considered stochastic setup, a simple covert channel can be established by exploiting an UIO as a decoding mechanism. It is then shown that the robustness of the obtained solution can be improved by adding an error-correcting code on top of the UIO operations. Finally, we show that the sender can exploit the control loop to obtain an implicit acknowledge (Ack) message about the decoding performed by the receiver and re-transmit the same bit message whenever the decoding operations are unsuccessful. Fig. 3 shows a graphical illustration of the proposed covert channel.

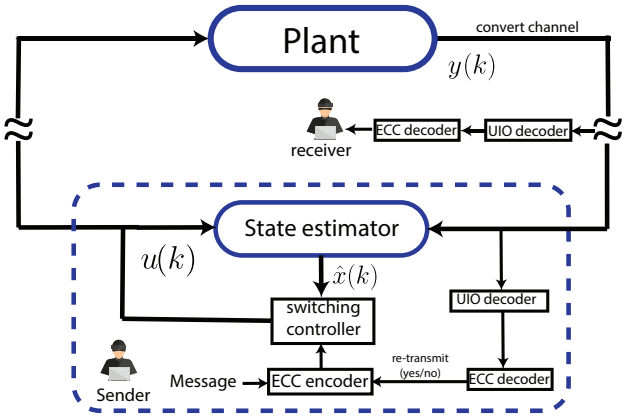

Fig. 3: Proposed covert channel implementation.

3.1 *UIO-based Covert Channel*

3.1.1 Sender: If the attacker (sender) can arbitrarily manipulate the control logic (3), then each bit of the message vector $M = [m_1, \ldots, m_l]^T$ can be sequentially encoded in the control signal $u(k)$ by resorting to the following switching control policy:

$$u(k) = \begin{cases} u_0(k) \triangleq f_0(\hat{x}(k)) & \text{if} \quad m_i = 0 \\ u_1(k) \triangleq f_1(\hat{x}(k)) & \text{else} \quad m_i = 1 \end{cases} \quad (4)$$

where $f_0(\hat{x}(k))$ and $f_1(\hat{x}(k))$ are two stabilizing control laws.

It should be noted that regardless of the switching logic used by the sender, the sensor measurements $y(k)$ and the control signal $u(k)$ evolution will be coherent with the system model (1). As a consequence, any standard anomaly detector leveraging the system dynamics (1)-(2) (e.g., χ^2 detector) will fail to detect the control logic alteration. The latter finds explanation in the fact that existing physics-based anomaly detectors focus their attention only on false data injection attacks affecting the communication channels, see e.g. the survey paper [3].

3.1.2 Receiver: If the receiver is aware of the switching control law (4), then, it can treat $u(k)$ as an unmeasurable disturbance and exploit an unknown input observer [35–37] to estimate its value, namely $\hat{u}_r(k)$, and decode the transmitted bit.

Let us abstractly model the UIO operations by means of the following recursive function

$$[\hat{u}_r(k-1), \hat{x}_r(k)] = UIO(y(k), \hat{u}_r(k-2), \hat{x}_r(k-1)) \quad (5)$$

where the pair $(\hat{u}_r(k-1), \hat{x}_r(k))$ are the estimated input and state vectors at the time k, respectively, and $(\hat{u}_r(k-2), \hat{x}_r(k-1))$ the estimations obtained at the previous iteration.

Then, at each time instant k, given $\hat{x}_r(k)$, the receiver can predict the future admissible inputs, namely $\hat{u}_0(k)$ and $\hat{u}_1(k)$, as

$$\begin{aligned} \hat{u}_0(k) &= f_0(\hat{x}_r(k)), \\ \hat{u}_1(k) &= f_1(\hat{x}_r(k)), \end{aligned} \quad (6)$$

© The Institution of Engineering and Technology 2015

and estimate the previously transmitted bit according to the following rule

$$\hat{bit}(k-1) = \begin{cases} 0 & \text{if } d_0 < d_1 \\ 1 & \text{otherwise} \end{cases} \quad (7)$$

where

$$\begin{aligned} d_0 &= \|\hat{u}_r(k-1) - \hat{u}_0(k-1)\|_2, \\ d_1 &= \|\hat{u}_r(k-1) - \hat{u}_1(k-1)\|_2 \end{aligned} \quad (8)$$

are the distances between the predicted and estimated inputs.

3.2 Using ECC to improve the covert channel capacity

The above-proposed encoding (using (4)) and decoding (using (7)) mechanisms show the existence of a covert channel in stochastic CPSs. However, such implementation is not robust against bit decoding errors that might arise due to noise realizations. To mitigate such a drawback, an error correcting code scheme can be utilized to correct a finite number of decoding errors that might be obtained using (7). For example, an (n_c, k_c, d_c) block error correcting code C where k_c message bits are encoded to n_c code bits with minimum distance d_c between code words (i.e., it can correct up to $(d_c-1)/2$ errors) [38] can be used on top of the UIO-based scheme (4)-(7) following the procedure below:

Algorithm 1 Basic Covert Channel: Sender and Receiver

—Sender (S)—

1: The original binary message M is divided into blocks $\{m\}$ of length k_c.
2: Each message block m is encoded by the ECC encoder into a codeword c of length n_c.
3: Let $c = [c_1, \ldots, c_{n_c}] \in \mathcal{C}$ denote the current codeword. Each bit c_i is encoded in the control input according to (4).
4: When all the bit of the current codeword are transmitted, the next message block is considered and the same procedure is re-applied.

—Receiver (R)—

1: At each time instant, the UIO-based bit decoder (5)-(8) is used to obtain an estimation of the transmitted bit, namely $\hat{c}_i$. Such bit is appended in the received codeword $\bar{c}$.
2: Every time a string $\bar{c}$ of length n_c is obtained, the ECC decoder is used to estimate $\hat{m}$ and $\bar{c}$ is reset, i.e. $\bar{c} = \emptyset$, to receive a new codeword.

3.3 Implicit Acknowledgment channel

The effectiveness of the ECC scheme, and hence the capacity of the underlying covert channe, can be further improved by noting that the networked control systems' closed-loop structure can be leveraged to establish an implicit acknowledgment message between the sender and the receiver. In particular, the feedback loop allows the controller (sender) to deterministically emulate the receiver operations by locally implementing the UIO (5) and the ECC decoding operations (steps 1-2 of Algorithm 1-*Receiver*), also see Fig. 3. Therefore, both the sender and receiver can estimate the number of bit errors during the transmission of each coding block. Then, the receiver, in accordance with the sender, can accept a received word only if it can make a very reliable decision about the codeword sent by the networked controller. For instance, the latter can be achieved by accepting codewords $\hat{c}$ whose Hamming distance d_c from the closest codeword in $\mathcal{C}$ is much smaller than the number of errors correctable by an optimal decoder for $\mathcal{C}$.

4 Proof-of-concept implementation

This section provides a simple proof of concept covert channel implementation that will be used in the next section to numerically evaluated the covert channel effectiveness.

4.1 Infected control logic

By assuming that the controller logic (3) is a stabilizing linear state-feedback controller with control gain $K \in \mathbb{R}^{m_u \times n_x}$, i.e.

$$u(k) = -K\hat{x}(k) \quad (9)$$

the sender can obtain a switching control logic as in (4) by simply introducing a shift $\delta > 0$ in the current state, i.e.,

$$u(k) = \begin{cases} u_0(k) = -K(\hat{x}(k) - [1,\ 1,\ 1]^T\delta) & \text{if} \quad c_i = 0 \\ u_1(k) = -K(\hat{x}(k) + [1,\ 1,\ 1]^T\delta) & \text{else} \quad c_i = 1 \end{cases} \quad (10)$$

A lower bound on the the probability of error at the receiver as function of δ can be analytically derived if the entire state vector can be measured or estimated perfectly, i.e. $\hat{x}(k) = x(k),\ \forall\, k$. Let's consider, for simplicity, the case where no ECC is implemented, and denote with $y_0(k+1)$ and $y_1(k+1)$ the measurements obtained when applying the controller input $u_0(k)$, and $u_0(k)$, respectively. Then, the distance D between $y_0(k+1)$ and $y_1(k+1)$ is given by $D = \|y_1(k+1) - y_0(k+1)\| = \|2CB(u_1(k) - u_0(k))\| = \|2CBK\delta\|$. Thus the probability of bit error at the receiver is lower bounded by

$$p_e = Q(D/\sqrt{2N_0}), \quad (11)$$

where $N_0 = CWC^T + V$, and $Q(\cdot)$ denotes the Q function given by $Q(z) = \frac{1}{\sqrt{(2\pi)}}\int_z^\infty e^{\frac{u^2}{2}}\, du$ [39]. This shows that increasing δ can improve the decoding performance of the underlying covert channel.

4.2 UIO algorithm

Since the considered plant model (1) does not contain a direct feedthrough term, the attacker is capable of implementing the UIO solution developed in [35].
Under standard UIO assumptions:

- $rank(C) = p, rank(B) = m_u, m_u \leq p, rank(CB) = m_u$,
- $C(zI - A)^{-1}B$ is left invertible and strictly minimum phase, i.e.

$$rank\left(\begin{bmatrix} zI - A & -zB \\ C & 0 \end{bmatrix}\right) = n_x + m_u, \forall z \in \mathbb{C}, |z| \geq 1, \quad (12)$$

both the sender and receiver can obtain the optimal input $\hat{u}_r(k-1)$ and state $\hat{x}_r(k)$ estimations by means of the stable recursive algorithm described by the equations (13)-(22) :

$$\hat{u}_r(k-1) = K_k^d\left(y(k) - CA\hat{x}_r(k-1)\right) \quad (13)$$

$$\hat{x}_r(k) = \hat{x}_r^+(k-1) + K_k^x\left(y(k) - C\hat{x}_r^+(k-1)\right) \quad (14)$$

with

$$\hat{x}_r^+(k-1) = A\hat{x}_r(k-1) + B\hat{u}_r(k-1) \quad (15)$$

$$K_k^x = \left(\bar{P}_{k-1/k-1}^{-1} + C^T\mathcal{V}C\right)^{-1} C^T\mathcal{V}^{-1} \quad (16)$$

$$K_k^d = P_{k/k}^{dx}C^T\mathcal{V}^{-1} \quad (17)$$

where

$$\bar{P}_{k-1/k-1} = AP_{k-1/k-1}^x A^T + \mathcal{W} \quad (18)$$

© The Institution of Engineering and Technology 2015

$$P^{d}_{k-1/k} = \left(B^T C^T (\mathcal{V} + C\bar{P}_{k-1/k-1}C^T)^{-1}CB\right)^{-1} \quad (19)$$

$$P^{xd}_{k/k} = P^{x}_{k/k}\bar{P}^{-1}_{k-1/k-1}B\left(B^T\bar{P}^{-1}_{k-1/k-1}F\right)^{-1} \quad (20)$$

$$P^{dx}_{k/k} = P^{d}_{k-1/k}B^T\bar{P}^{-1}_{k-1/k-1}\left(\bar{P}^{-1}_{k-1/k-1} + C^T\mathcal{V}^{-1}C\right)^{-1} \quad (21)$$

$$P^{x}_{k/k} = \left(\bar{P}^{-1}_{k-1/k-1} + C^T\mathcal{V}^{-1}C\right)^{-1} + P^{xd}_{k/k}(P^{d}_{k-1/k})^{-1}P^{dx}_{k/k} \quad (22)$$

Remark 1. *If the plant model contains a feedthrough term, then the UIO developed in [37] can be used instead, and the same methodology still holds.*

4.3 ECC scheme

In what follows, we utilize a repetition code where every bit of the message m_i of M is encoded n_c times in the channel before moving to the next bit. The receiver decodes the message bits accordingly. Note that the repetition code is the simplest form of ECC, which we is chosen in here for ease of exposition but in practice a more efficient coding scheme, i.e., one with a better coding rate, can be utilized.

When using repetition code, the implicit acknowledgment can thus be implemented as follows. Let τ denote the number of error bits within each block as estimated by both the sender and receiver. The sender and receiver offline agree to accept $\hat{m}_r$ if and only if $\tau << (d_c - 1)/2 = (n_c - 1)/2$, otherwise the decoded message $\hat{m}_r$ is discarded by the receiver and the sub-string m_r re-transmitted by the sender.

5 Numerical Results

In this section, the performance of the covert channel is evaluated by considering a single area Automatic Generation Control (AGC) system. Such a choice is mainly motivated by the fact that AGC systems are core components of any smart grid. Moreover, differently from other CPS applications (e.g., smart transportation's system), smart grids have already a great diffusion in our society, and their security is of a great concern. It should be noted, however, that the proposed covert channel is based on a control-theoretical approach. Therefore, the proposed solution does not depend on the specific example, but instead on its mathematical abstraction. In other words, as long as the mathematical model of the CPS can be abstracted as a stochastic linear time-invariant system, see (1), then the proposed approach can be used.

The AGC dynamics are here approximated by means of a LTI model. By considering a sampling time $T_s = 0.02\,\text{sec}$, the AGC system matrices (1) are [40]:

$$A = \begin{bmatrix} 0.8664 & -0.1928 & -0.3840 \\ 0.0186 & 0.998 & -0.0039 \\ 0.0002 & 0.02 & 1 \end{bmatrix}$$

$$B = \begin{bmatrix} 0.0186 \\ 0.0002 \\ 0 \end{bmatrix}$$

$$C = \begin{bmatrix} 0.1 & 0.7 & 1 \end{bmatrix}$$

with $x(k) \in \mathbb{R}^3$, $u(k) \in \mathbb{R}^1$, $y(k) \in \mathbb{R}$, $\mathcal{W} = 10^{-9}I_3$ and $\mathcal{V} = 10^{-9}$. Moreover, the controller (9) is a Linear Quadratic (LQ) controller , where the gain

$$K = [0.0211,\ 0.0542,\ 0.003]$$

has been obtained by using $Q = 0.2I_3$ and $R = 1$ as state and input LQ weight matrices, respectively.

To evaluate the effectiveness of the proposed covert channel, the bit error rate (BER) characterizing the fraction of errors in the bits decoded by the receiver, as a function of the number of transmitted bits, is of interest. In particular, the BER is evaluated for $\delta \in \{0,\ 0.5,\ 0.6,\ 0.7,\ 0.8,\ 0.9,\ 1\}$.

By numerical substitution the different values of δ in (11), we obtain BER is lower bounded by $\{0.50,\ 1.62 \times 10^{-2},\ 5.1 \times 10^{-3},\ 1.4 \times 10^{-3},\ 3.11 \times 10^{-04},\ 5.91 \times 10^{-5}, 9.45 \times 10^{-6}\}$, for $\delta = \{0,\ 0.5,\ 0.6,\ \cdots,\ 1.0\}$, respectively. For the proposed UIO-based covert channel implementation, to obtain a more accurate estimate of the BER, it is instead evaluated numerically. In particular, we have conducted a set of simulations to compute the BER for different initial state conditions and values of $n_c \in \{1,\ 3,\ 5,\ 7,\ 9,\ 11\}$ in the repetition code. Moreover, a randomly generated message M of $100,000$ bits is considered and a threshold $\tau = 0$ is used for the ECC with implicit acknowledgment. The obtained simulations results are depicted in the Tables 1-2, and Figs. 4-6.

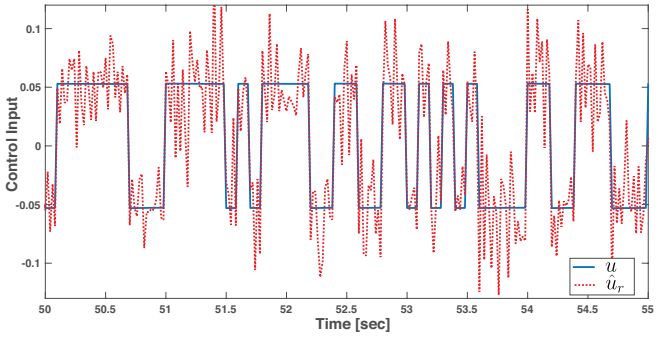

Fig. 4: Control input $\hat{u}_r(k)$ estimated by the UIO vs input $u(k)$ used by the attacker for $\delta = 0.7$, $n_c = 3$.

In Fig. 4, by considering the time interval $t \in [50,\ 55]\,\text{sec}$, it is possible to observe how the UIO is capable of estimating the input used by the attacker from the available sensor measurements.

Table 1 reports the bit error rate (BER) characterizing the fraction of errors in the bits decoded by the receiver, as a function of the number of transmitted bits. The obtained results indicate that either increasing the state shift δ or the number of repetition n_c, the BER decreases. Repeating the experiment with the same parameters shown in Table 1, all the bits were decoded correctly when using the implicit Ack scheme with $\tau = 0$.

In Fig. 6, we report the percentage of accepted codewords satisfying the ECC decoding rule. If only the majority rule is considered, then any received codeword will satisfy the decoding rule in spite of the number of repetitions and state shift δ. On the other hand, if the ECC with consensus is considered, then the number of accepted codewords decreases with both the number of repetitions n_c and state shift δ.

δ \ n_c	*1*	*3*	*5*	*7*	*9*	*11*
0.5	0.102	0.017	0.003	0.001	0	0
0.6	0.064	0.006	0.001	0	0	0
0.7	0.038	0.002	0	0	0	0
0.8	0.021	0	0	0	0	0
0.9	0.011	0	0	0	0	0
1	0.005	0	0	0	0	0

Table 1 BER using repetition code with majority decoding.

As discussed in section 4.1 and confirmed by the numerical results in Fig. 6 and Table 1, the bit error rate of the considered covert channel improves (i.e., decreases) when we increase δ. However, it is also important that the covert channel does not alter, in a significant way, the performance of the plant. To this end, the following normalized

© The Institution of Engineering and Technology 2015

LQ cost has been evaluated

$$J = \frac{\sum_{k=0}^{N} \left(x^T(k) Q x(k) + u^T(k) R u(k) \right)}{N+1}$$

where N is the number of simulation steps. From Table 2, it is possible to notice that although the control cost remains relatively small, its value increases with δ. The latter finds justification in the fact that state trajectory oscillations around the equilibrium are directly proportional to the state shift δ. This finds confirmation in Fig. 5, where, by considering the time interval $t \in [0, 50]$ sec, the state components evolution is reported in the presence and in the absence of the covert channel.

	- δ -						
	0	*0.5*	*0.6*	*0.7*	*0.8*	*0.9*	*1*
J	3.37 $\times 10^{-8}$	0.14 $\times 10^{-3}$	0.20 $\times 10^{-3}$	0.28 $\times 10^{-3}$	0.37 $\times 10^{-3}$	0.46 $\times 10^{-3}$	0.57 $\times 10^{-3}$

Table 2 Normalize LQ cost for different values of δ.

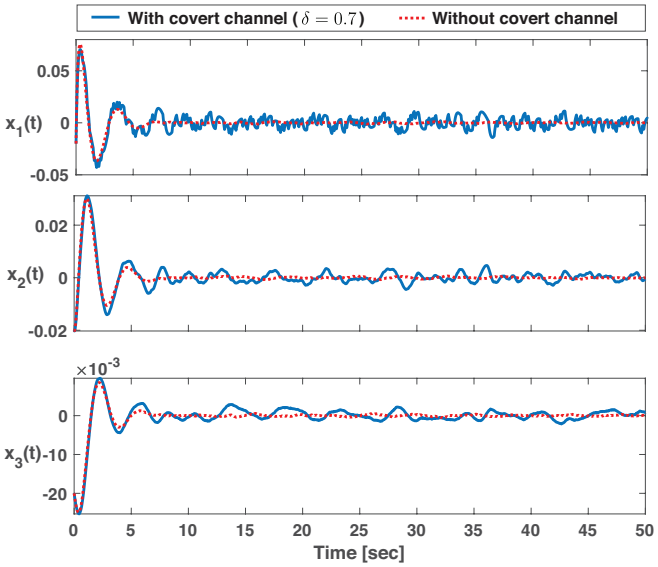

Fig. 5: State components evolution: with covert channel vs without covert channel.

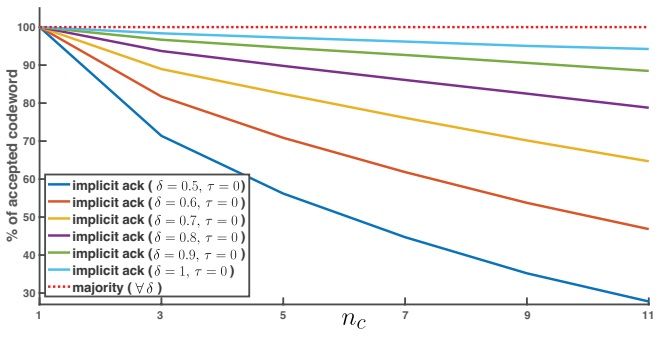

Fig. 6: Percentage of accepted codewords: ECC with majority vs ECC with implicit feedback acknowledgment and $\tau = 0$.

6 Defences and Countermeasures

Given our threat model, CPSs can be vulnerable to such covert channels if the attacker is able to inject malware into the networked controller and arbitrarily alter the control logic. Thus the root cause of such covert channel can be prevented by ensuring the security of the supply-chain of the CPS equipment and software (e.g., see [41]) in order to avoid the risk of malicious, compromised, or infected suppliers who alter the control logic of the controller. In order to prevent malware infection after installing the controller, a defence technique would be to use behavioural monitoring with anomaly detection to detect reconnaissance by adversaries prior to data exfiltration. These intrusion detection systems should provide alarms whenever it detects any unauthorized controller programming, or suspicious traffic going to/from ICS devices. The use of Trusted Execution Environment (TEE) such as Trusted Platform Module (TPM), Intel Software Guard Extensions (SGX) and ARM TrustZone can also help ensure the integrity of the code execution within the controller (e.g., see [42]).

In addition to the above defence techniques, the specific covert channel we present may be foiled by specific countermeasures. For example, an effective solution is to deploy a smart actuator that applies, to the plant, a randomized version of the controller's output signal, hence reducing the attacker's capability to predict the expected plant's evolution given the available model of the plant. However, such solution should be carefully implemented in order to avoid reducing the performance of the control-loop.

7 Conclusions and Future Works

In this paper, a covert channel for stochastic cyber-physical systems was presented. In particular, first, an unknown input observer is used to allow the receiver to decode, from the sensor measurement, the binary messages encoded in the control signal by a sender manipulating the control logic. Then, such a channel's reliability is improved jointly leveraging the control system feedback loop and standard ECC schemes. Finally, by considering a numerical testbed and a simple ECC implementation, the covert channel's capacity has been investigated. Future works include the development of methods for detection, capacity limitation, and elimination of such covert channels in CPSs.

8 References

1 F. Pasqualetti, F. Dörfler, and F. Bullo, "Attack detection and identification in cyber-physical systems," *IEEE transactions on automatic control*, vol. 58, no. 11, pp. 2715–2729, 2013.

2 A. Cardenas, S. Amin, B. Sinopoli, A. Giani, A. Perrig, S. Sastry *et al.*, "Challenges for securing cyber physical systems," in *Workshop on future directions in cyber-physical systems security*, vol. 5, no. 1. Citeseer, 2009.

3 J. Giraldo, D. Urbina, A. Cardenas, J. Valente, M. Faisal, J. Ruths, N. O. Tippenhauer, H. Sandberg, and R. Candell, "A survey of physics-based attack detection in cyber-physical systems," *ACM Computing Surveys (CSUR)*, vol. 51, no. 4, pp. 1–36, 2018.

4 S. M. Dibaji, M. Pirani, D. B. Flamholz, A. M. Annaswamy, K. H. Johansson, and A. Chakrabortty, "A systems and control perspective of cps security," *Annual Reviews in Control*, 2019.

5 M. S. Chong, H. Sandberg, and A. M. Teixeira, "A tutorial introduction to security and privacy for cyber-physical systems," in *2019 18th European Control Conference (ECC)*. IEEE, 2019, pp. 968–978.

6 F. Farokhi and H. Sandberg, "Ensuring privacy with constrained additive noise by minimizing fisher information," *Automatica*, vol. 99, pp. 275–288, 2019.

7 A. Abdelwahab, W. Lucia, and A. Youssef, "Decoy-based moving target defense against cyber-physical attacks on smart grid," in *2020 IEEE Electric Power and Energy Conference (EPEC)*. IEEE, 2020, pp. 1–5.

8 A. J. Menezes, P. C. Van Oorschot, and S. A. Vanstone, *Handbook of applied cryptography*. CRC press, 2018.

9 B. W. Lampson, "A note on the confinement problem," *Communications of the ACM*, vol. 16, no. 10, pp. 613–615, 1973.

10 B. Carrara and C. Adams, "Out-of-band covert channelsâĂŤa survey," *ACM Computing Surveys (CSUR)*, vol. 49, no. 2, pp. 1–36, 2016.

11 D. C. Latham, "Department of defense trusted computer system evaluation criteria," *Department of Defense*, 1986.

12 S. Zander, G. Armitage, and P. Branch, "A survey of covert channels and countermeasures in computer network protocols," *IEEE Communications Surveys & Tutorials*, vol. 9, no. 3, pp. 44–57, 2007.

© The Institution of Engineering and Technology 2015

13 M. Hanspach and M. Goetz, "On covert acoustical mesh networks in air," *arXiv preprint arXiv:1406.1213*, 2014.
14 M. Guri, Y. Solewicz, A. Daidakulov, and Y. Elovici, "Fansmitter: Acoustic data exfiltration from (speakerless) air-gapped computers," *arXiv preprint arXiv:1606.05915*, 2016.
15 M. Guri, M. Monitz, Y. Mirski, and Y. Elovici, "Bitwhisper: Covert signaling channel between air-gapped computers using thermal manipulations," in *IEEE Computer Security Foundations Symposium*. IEEE, 2015, pp. 276–289.
16 L. Deshotels, "Inaudible sound as a covert channel in mobile devices," in *USENIX Workshop on Offensive Technologies (WOOT)*, 2014.
17 D. B. Bartolini, P. Miedl, and L. Thiele, "On the capacity of thermal covert channels in multicores," in *European Conference on Computer Systems*, 2016, pp. 1–16.
18 J. Loughry and D. A. Umphress, "Information leakage from optical emanations," *ACM Transactions on Information and System Security (TISSEC)*, vol. 5, no. 3, pp. 262–289, 2002.
19 S. R. Chhetri and M. A. Al Faruque, "Side channels of cyber-physical systems: Case study in additive manufacturing," *IEEE Design & Test*, vol. 34, no. 4, pp. 18–25, 2017.
20 S. R. Chhetri, A. Canedo, and M. A. A. Faruque, "Confidentiality breach through acoustic side-channel in cyber-physical additive manufacturing systems," *ACM Transactions on Cyber-Physical Systems*, vol. 2, no. 1, pp. 1–25, 2017.
21 A. S. Uluagac, V. Subramanian, and R. Beyah, "Sensory channel threats to cyber physical systems: A wake-up call," in *2014 IEEE Conference on Communications and Network Security*. IEEE, 2014, pp. 301–309.
22 P. Krishnamurthy, F. Khorrami, R. Karri, D. Paul-Pena, and H. Salehghaffari, "Process-aware covert channels using physical instrumentation in cyber-physical systems," *IEEE Transactions on Information Forensics and Security*, vol. 13, no. 11, pp. 2761–2771, 2018.
23 L. Garcia, H. Senyondo, S. McLaughlin, and S. Zonouz, "Covert channel communication through physical interdependencies in cyber-physical infrastructures," in *Int. Conference on Smart Grid Communications (SmartGridComm)*, 2014, pp. 952–957.
24 S. Wendzel, B. Kahler, and T. Rist, "Covert channels and their prevention in building automation protocols: A prototype exemplified using bacnet," in *2012 IEEE International Conference on Green Computing and Communications*. IEEE, 2012, pp. 731–736.
25 C. Alcaraz, G. Bernieri, F. Pascucci, J. Lopez, and R. Setola, "Covert channels-based stealth attacks in industry 4.0," *IEEE Systems Journal*, vol. 13, no. 4, pp. 3980–3988, 2019.
26 T. T. Gamage, B. M. McMillin, and T. P. Roth, "Enforcing information flow security properties in cyber-physical systems: A generalized framework based on compensation," in *2010 IEEE 34th Annual Computer Software and Applications Conference Workshops*. IEEE, 2010, pp. 158–163.
27 R. Akella, H. Tang, and B. M. McMillin, "Analysis of information flow security in cyber–physical systems," *International Journal of Critical Infrastructure Protection*, vol. 3, no. 3-4, pp. 157–173, 2010.
28 X. Ying, G. Bernieri, M. Conti, and R. Poovendran, "Tacan: Transmitter authentication through covert channels in controller area networks," in *Proceedings of the 10th ACM/IEEE International Conference on Cyber-Physical Systems*, 2019, pp. 23–34.
29 J. M. Taylor and H. R. Sharif, "Enhancing integrity of modbus tcp through covert channels," in *2017 11th International Conference on Signal Processing and Communication Systems (ICSPCS)*. IEEE, 2017, pp. 1–6.
30 A. Herzberg and Y. Kfir, "The chatty-sensor: A provably-covert channel in cyber physical systems," in *Annual Computer Security Applications Conference*, ser. ACSAC'19, 2019, pp. 638–649.
31 ——, "The leaky actuator: A provably-covert channel in cyber physical systems," in *ACM Workshop on Cyber-Physical Systems Security & Privacy*, ser. CPS-SPC'19, 2019, pp. 87–98.
32 A. Abdelwahab, W. Lucia, and A. Youssef, "Covert channels in cyber-physical systems," *IEEE Control Systems Letters*, vol. 5, no. 4, pp. 1273–1278, 2021.
33 W. Lucia and A. Youssef, "Wyner wiretap-like encoding scheme for cyber-physical systems," *IET Cyber-Physical Systems: Theory Applications*, vol. 5, no. 4, pp. 359–365, 2020.
34 A. Teixeira, I. Shames, H. Sandberg, and K. H. Johansson, "A secure control framework for resource-limited adversaries," *Automatica*, vol. 51, pp. 135–148, 2015.
35 M. Darouach, M. Zasadzinski, A. B. Onana, and S. Nowakowski, "Kalman filtering with unknown inputs via optimal state estimation of singular systems," *International journal of systems science*, vol. 26, no. 10, pp. 2015–2028, 1995.
36 S. Z. Yong, M. Zhu, and E. Frazzoli, "Simultaneous input and state estimation for linear discrete-time stochastic systems with direct feedthrough," in *IEEE Conference on Decision and Control*. IEEE, 2013, pp. 7034–7039.
37 ——, "A unified filter for simultaneous input and state estimation of linear discrete-time stochastic systems," *Automatica*, vol. 63, pp. 321–329, 2016.
38 W. C. Huffman and V. Pless, *Fundamentals of error-correcting codes*. Cambridge university press, 2010.
39 B. P. Lathi, *Modern digital and analog communication systems*. Oxford University Press, Inc., 1995.
40 H. Saadat *et al.*, *Power system analysis*. McGraw-Hill, 1999, vol. 2.
41 S. Peisert, B. Schneier, H. Okhravi, F. Massacci, T. Benzel, C. Landwehr, M. Mannan, J. Mirkovic, A. Prakash, and J. B. Michael, "Perspectives on the solarwinds incident," *IEEE Security & Privacy*, vol. 19, no. 02, pp. 7–13, 2021.
42 A. M. Naseri, W. Lucia, M. Mannan, and A. Youssef, "On securing cloud-hosted cyber-physical systems using trusted execution environments," in *The IEE International Conference on Autonomous Systems (ICAS 2021)*. IEEE, 2021, pp. 1–5.

© The Institution of Engineering and Technology 2015