Secure Control of Cyber-Physical Systems

Ahmed Abdelwahab

A Thesis
in
The Concordia Institute
for
Information Systems Engineering

Presented in Partial Fulfillment of the Requirements
For the Degree of
Master of Applied Science (Information Systems Security) at
Concordia University
Montréal, Québec, Canada

November 2020

CONCORDIA UNIVERSITY
School of Graduate Studies

This is to certify that the thesis prepared

By: **Ahmed Abdelwahab**

Entitled: **Secure Control of Cyber-Physical Systems**

and submitted in partial fulfillment of the requirements for the degree of

Master of Applied Science (Information Systems Security)

complies with the regulations of this University and meets the accepted standards with respect to originality and quality.

Signed by the final examining committee:

_____ Chair
Dr. Mohammad Mannan

_____ External
Dr. Ahmed Kishk

_____ Examiner
Dr. Jun Yan

_____ Examiner
Dr. Mohammad Mannan

_____ Thesis Supervisor
Dr. Amr Youssef

_____ Thesis Supervisor
Dr. Walter Lucia

Approved by _____
Dr. Mohammad Mannan, Graduate Program Director

November 26, 2020 _____
Dr. Mourad Debbabi, Interim Dean
Gina Cody School of Engineering and Computer Science

Abstract

Secure Control of Cyber-Physical Systems

Ahmed Abdelwahab

Cyber-Physical Systems (CPS) are smart co-engineered interacting networks of physical and computational components. They refer to a large class of technologies and infrastructure in almost all life aspects including, for example, smart grids, autonomous vehicles, Internet of Things (IoT), advanced medical devices, and water supply systems. The development of CPS 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, i.e. 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 attack. In a large class of CPSs, the physical systems are controlled using a feedback control loop. In this thesis, we investigate, from many angles, how CPSs' control systems can be prone to cyber-physical attacks and how to defend them against such attacks using arguments drawn from control theory.

In our first contribution, by considering Smart Grid applications, we address the problem of designing a Denial of Service (DoS)-resilient controller for recovering the system's transient stability robustly. We propose a Model Predictive Control (MPC) controller based on the set-theoretic (ST) arguments, which is capable of dealing with both model uncertainties, actuator limitations, and DoS. Unlike traditional MPC solutions, the proposed controller has the capability of moving most of the required computations into an offline

phase. The online phase requires the solution of a quadratic programming problem, which can be efficiently solved in real-time. Then, stemming from the same ST based MPC controller idea, we propose a novel physical watermarking technique for the active detection of replay attacks in CPSs. The proposed strategy exploits the ST-MPC paradigm to design control inputs that, whenever needed, can be safely and continuously applied to the system for an apriori known number of steps. Such a control scheme enables the design of a physical watermarked control signal. We prove that, in the attack-free case, the generators' transient stability is achieved for all admissible watermarking signals and that the closed-loop system enjoys uniformly ultimately bounded stability.

In our second contribution, we address the attacker's ability to collect useful information about the control system in the reconnaissance phase of a cyber-physical attack. By using existing system identification tools, an attacker who has access to the control loop can identify the dynamics of the underlying control system. We develop a decoy-based moving target defense mechanism by leveraging an auxiliary set of virtual state-based decoy systems. Simulation results show that the provided solution degrades the attacker's ability to identify the underlying state-space model of the considered system from the intercepted control inputs and sensor measurements. It also does not impose any penalty on the control performance of the underlying system.

Finally, in our third contribution, we introduce a covert channel technique, enabling a compromised networked controller to leak information to an eavesdropper who has access to the measurement channel. We show that this can be achieved without establishing any additional explicit communication channels by properly altering the control logic and exploiting robust reachability arguments. A dual-mode receding horizon MPC strategy is used as an illustrative example to show how such an undetectable covert channel can be established.

Acknowledgments

I would like to express my deepest gratitude to my supervisors, Dr. Amr Youssef and Dr. Walter Lucia for their endless support and guidance throughout my master's degree. This work would not have seen the light without their continuous encouragement and persistent help. Anyone who would have the chance to work with them both is a very lucky person.

I would also like to extend my gratitude to the Concordia Institute for Information Systems Engineering (CIISE) and Concordia University for supporting me both technically and financially during my degree, and to my Professors, for their patience and valuable knowledge, especially Dr. Ayda Basyouni, Dr. Makan Pourzandi, and Dr. Amin Ranj Bar.

I wish to thank my colleagues at Concordia University: M. Tolba, M. Elsayed, M. Elshiekh, Hisham, Mahdi, Mounir, Mariam, Kian, Mohsen, and Shima. They did not hold back when I asked for help, no matter how hard the problem was.

Last but not least, I thank my parents, my wife, my siblings, and my son, for unconditional love and support. It is, without a doubt, because of them I am writing these words now, at the last mile of my degree.

Contents

List of Figures

List of Abbreviations

CPS	Cyber-Physical System
FDI	False Data Injection
DoS	Denial of Service
MPC	Model Predictive Control
ST	Set-Theoretic
ST-MPC	Set-Theoretic Model Predictive Control
QP	Quadratic Programming
DoA	Domain of Attraction
RPI	Robust Positively Invariant
UUB	Uniformly Ultimately Bounded
PMU	Phasor Measurement Unit
ST-WC	Set-Theoretic Watermarked Control
AGC	Automatic Generation Control
SCADA	Supervisory Control and Data Acquisition
PLC	Programmable Logic Controller
CC-NCS	Covert Channel in Networked Control System

Chapter 1

Introduction

As defined by the U.S. National Institute of Standards and Technology (NIST), Cyber-Physical Systems (CPS) refer to "smart" co-engineered interacting networks of physical and computational components. As such, the computing/communication components, and the physical parts are tightly integrated. CPS range from small devices, such as implantable medical devices, to large-scale systems such as smart grids, nuclear power plants, and water supply systems. In most CPS, embedded sensors and actuators are connected with distributed control systems through communication channels. The security of these *intelligent* infrastructures against cyber-physical attacks is a major concern and this problem has received increasing attention in the control community in the last decade (e.g., see [1–11]).

In this chapter, we will give a brief background on the field and the basic concepts related to it, e.g. attack classifications, system architecture, attacker resources. Later, at the beginning of each chapter, a more in-depth look into the related literature is provided.

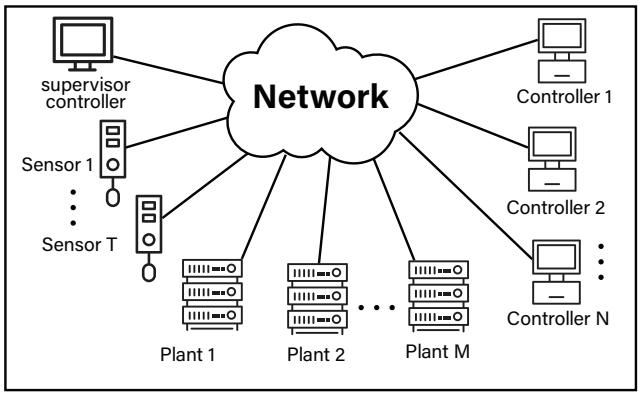


Figure 1: A Cyber-Physical System System conceptual model [12]

1.1 Background

1.1.1 Cyber-Physical Systems

Cyber-Physical Systems (CPSs) is a term used to describe physical systems equipped with communication capabilities, in addition to computational power. The key feature of CPSs is the flow of information (e.g. control input, plant measurements, reference points, etc.) in the network among system components, see Figure 1. The applications of CPSs, are endless. They include, for example, smart grids, medical robots, manufacturing plants, water distribution systems, and aircraft. The advantages of such systems stretch over a wide spectrum. CPSs allow the rapid deployment of various components, like sensors, actuators, and controllers without fundamental system change or restructuring. The efficiency of sharing data while being able to fuse it to make intelligent decisions over a large system such as smart grids is a clear example of such use. Although CPSs present lots of advantages, the same features might expose the system to cyber-physical attacks and an adversary might launch different attacks by violating confidentiality, integrity, or availability (CIA) properties of the communication channels and the information exchanged through it.

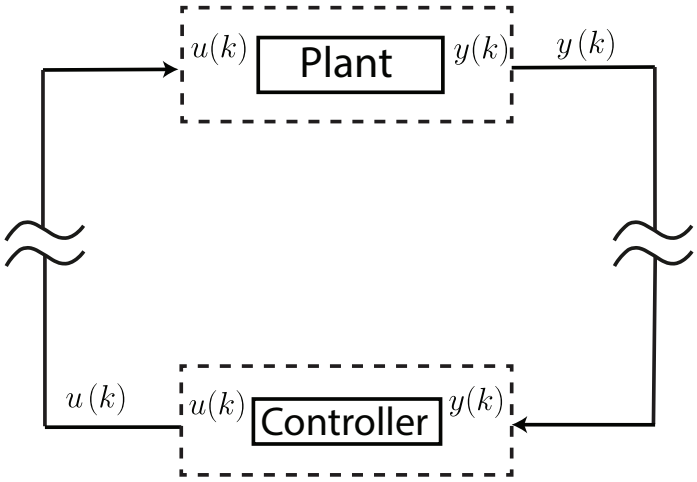


Figure 2: Networked Control System architecture.

1.1.2 Cyber-Physical Attacks Classifications

Consider the CPS architecture depicted in Figure 2, where the communication channels between the plant and controller are assumed to be insecure. The malicious party's goal is to destabilize the plant operations.

The shown system dynamics can be represented by the following state-space:

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

where $k \in \mathbb{Z}_+ := \{0, 1, \ldots\}$ is the discrete-time index, $x(k) \in \mathbb{R}^n$, $y(k) \in \mathbb{R}^p$, $u(k) \in \mathbb{R}^m$ are the state, measurements and control inputs vectors, respectively, and $f(\cdot,\cdot,\cdot) : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^n$ and $g(\cdot,\cdot) : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^p$ denote the system's dynamics and measurement functions, and $d_p \in \mathcal{N}(0_n, \Sigma_{d_p})$ and $d_m \in \mathcal{N}(0_m, \Sigma_{d_m})$ are independent and identically distributed (i.i.d.) Gaussian process and measurement noises, with zero mean and covariance matrix Σ_{d_p} and Σ_{d_m}. Matrices A, B, and C are assumed to be time independent with compatible dimensions.

The system (1) is assumed to be observable and controllable. Also, the system states

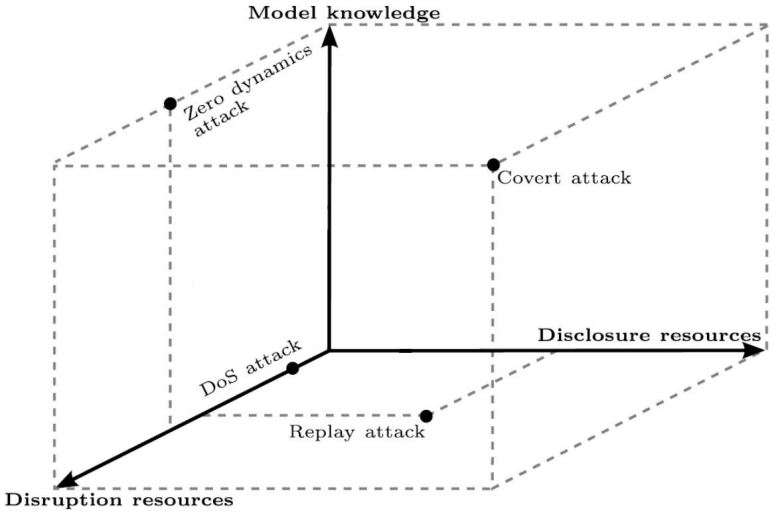


Figure 3: The attack-space in Cyber-Physical Systems [4]

and input can be subject to:

$$x(k) \in \mathcal{X},\ u(k) \in \mathcal{U} \tag{2}$$

where $\mathcal{X}$ and $\mathcal{U}$ are compact subsets of $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively.

Communication channels are considered secure if the confidentiality, integrity, and availability properties (CIA triad) are satisfied [13]. A channel is insecure if at least one of the CIA properties is not met. The ability of an attacker to disrupt the plant operations depends on the resources available to them. Moreover, the capability of the attacker to perform sophisticated attacks also depends on the available information on the system. Therefore, the attacker's resources determine the effect and severity of the cyber-physical attack.

To understand the concept of an attacker's resources, let us consider an insecure communication channel, where a data packet is transmitted at each sampling time. According to [4], the attacker's resources can be categorized as follows:

- *Disclosure Resources*: An attacker has disclosure resource on the the channel if he/she can violate the confidentiality property, i.e. intercept/read the data packet.

- *Disruptive Resources*: An attacker has disruptive resources on the channel if he/she can violate the authentication or integrity properties, i.e. the attacker can arbitrary change the transmitted data packet into a new compatible one.

- *Model Knowledge*: An attacker has model knowledge when the attacker has a subset $\mathcal{I}_{attacker}$ of the information characterizing the system dynamics, i.e.

$$\mathcal{I}_{attacker} \subseteq \{A, B, C, D, \mathcal{X}, \mathcal{U}, f(\cdot, \cdot, \cdot)\} \tag{3}$$

The model knowledge, disclosure, and disruptive resources are the basis to shape the attack space. This is shown in Fig. 3 where the different attacks are shown with respect to their required resources.

Denial-of-Service (DoS) attack is an attack that targets the availability of the information packets. DoS attacks require the attacker to have disruption resources on the attacked channels. It prevents the information from flowing to the destination either completely or by enforcing a delay.

False data injection (FDI) attack is a deception attack where the attacker adds a malicious signal on top of the transmitted packet in the compromised channel, and it requires disruption and disclosure resources. The attack can be modeled as an additive data injection, where the injected signal can be an arbitrary value or designed [13–15].

Subcategories or variations of the FDI attack present itself in the literature. Here, we will name a few:

- *Stealthy FDI attack* [4]: An FDI attack is considered stealthy when the attacker is capable of injecting a malicious signal in a communication channel for an arbitrary period of time without being detected by attack detection mechanisms.

- *Replay attack* [16]: Replay attacks require both the disclosure and disruptive resources to be successfully carried out. This type of attack can be stealthy if performed on a system in steady-state condition. It consists of two steps; first, the attacker records a feedback signal (disclosure resources) for an arbitrary period of time. Then, the attacker injects a malicious signal (disruptive resources) on the other channel while replaying the recorded channel instead of the legitimate feedback signal.

- *Zero dynamics attack* [17]: A zero dynamics attack requires full knowledge of the plant's model $\mathcal{I}_{attacker} = \{A, B, C, D\}$ and disruptive resources on the actuation channel. In particular, the attacker exploits the unstable transmission zeros of the system to inject a malicious input vector in the control signal which produces a zero response on the sensor measurements.

- *Covert attack* [18]: A covert attack is a sophisticated attack that requires plant's model $\mathcal{I}_{attacker} = \{A, B, C, D\}$, disruptive and disclosure resources on both actuation and feedback channels. In particular, the attacker injects a malicious vector on the actuation channel to deteriorate system performance and then injects another signal on the feedback channel to completely cancel the attack's effect in the feedback signal. This type of attack is undetectable by any detection mechanisms whose actions are performed on the controller-side of the networked control system.

Intelligence attacks are a type of cyber-physical attacks that target gathering some useful information about the system to enable the attacker of launching a more coordinated and sophisticated attacks. This is done as a part the reconnaissance phase of cyber-physical attack. One of the targets of this phase is to identify the plant model. This can be done by traditional intelligence operations, or by performing a System identification attack [19]. This attack requires disclosure resources on both input and output channels, and in some

instances disruptive resources as well. The attacker observes the input and output of the system to accurately identify the dynamics of the system by using system identification algorithms.

1.2 Thesis Motivation and Contributions

Observing the growing dependency on CPSs in our everyday life and the risk associated with such systems in the form of cyber-physical attacks, as well as the wide range of cyber-physical attacks that target such systems, in this work, we focus on securing CPSs by adding an extra level of defense in the form of what we call control layer security. The main motivation of this research is investigating attacks like FDI, and DoS attacks [8, 17, 20, 21] against CPSs and designing suitable mitigation and detection techniques. Starting from control theory arguments and existing solutions [3,22–24], we propose mitigation and detection strategies to counter malicious efforts to harm or violate CPSs' security. Also, we explore potential vulnerabilities in the typical CPSs architecture. In particular, we examine four problems:

- The resilient control in presence of DoS attacks targeting the measurement signal or the feedback channel of the control loop.
- The detection of replay attacks using active detection mechanisms while minimizing any performance loss as a result of the detection scheme.
- The prevention of malicious parties from disclosing system dynamics by performing system identification using the input and output signals of the control loop.
- The design for a covert channel technique that enables a compromised controller to leak information to an eavesdropper on the feedback channel.

1.3 Publications

- A. Abdelwahab, W. Lucia, and A. Youssef. “A DoS-resilient Set-Theoretic Controller for Smart Grid Applications." In The IEEE Power and Energy Society General Meeting (PESGM), 2020.
- A. Abdelwahab, W. Lucia, and A. Youssef. “Set-Theoretic Control for Active Detection of Replay Attacks with Applications to Smart Grid." In The IEEE Conference on Control Technology and Applications (CCTA), 2020.
- A. Abdelwahab, W. Lucia, and A. Youssef. “Decoy-based Moving Target defense Against Cyber-physical Attacks On Smart Grid." In The IEEE Electric Power and Energy Conference (EPEC), 2020.
- A. Abdelwahab, W. Lucia, and A. Youssef. “Covert Channels in Cyber-Physical Systems." In The IEEE control systems letters (L-CSS), 2020.

1.4 Outline

The rest of the thesis is organized as follows. In chapter 2, two different schemes are designed, one to mitigate DoS attacks and second to detect replay attacks in CPS, with application to smart grids. In chapter 3, another mitigation scheme is designed to prevent the identification of the control system dynamics by malicious parties. In chapter 4, a covert channel technique is designed to enable compromised controllers to leak information to an eavesdropper without triggering any detection mechanisms. Finally, chapter 5 concludes the thesis and highlights future research directions.

Chapter 2

Cyber-Physical Systems Security in Smart Grid Applications

The work in this chapter is published in two conference papers, PESGM 2020 and CCTA 2020, see [25, 26], respectively.

Over the past few years, CPSs have embraced communication and computation technologies to improve its reliability and efficiency. CPSs use data collection tools and sensors across the system to enable bidirectional communication between system entities. For instance, in a Smart Grid, the flow of information in the communication channels allows a high degree of freedom in using advanced control methods to control the power generation according to the consumption at any given moment. A properly designed smart power grid can cut down the possibility of major outages and pave the way for green energy.

Compromising the CPSs' communication channels threatens the functionality and security of the system. Also, disclosing the underlying information violates the privacy of the system and exposes it to greater risk, i.e. a sophisticated coordinated attack.

As a result, many security measures have to be put in place in order to defend the integrity, availability, and confidentiality of the information flowing in the CPS. It should also be noted that investigating the security of a system is not only concerned with deliberate

attacks but also with studying how the system acts under unexpected conditions such as natural disasters, system delays, or disturbances.

The first addressed topic in this chapter is Denial-of-Service (DoS) attacks on cyber-physical systems. More preciously, the effect of DoS attacks on the stability of a system. A prominent example of which this problem poses a serious issue is "smart grids", where DoS attacks can disrupt the "transient stability" of the grid and affect the dynamic performance of the power system [20].

Due to the real-time requirements of smart grids, packet delay and/or packet drop/loss can have severe effects on the grid, see e.g. [27–29]. Moreover, under attack conditions, for the system to be resilient, transient stability should be reached in the least time possible. To address such a problem, set-theoretic (ST) based Model Predictive Control (MPC) solutions [23, 30–32] are particularly appealing. Indeed, unlike other MPC approaches, ST-MPC is capable of addressing constrained control problems with a modest computational demand. In [33], the ST-MPC paradigm has been applied to solve the transient stability problem under constrained inputs, model uncertainties, and bounded measurement errors, but not packet delays or DoS attacks.

Introducing the ability to handle packet delays and packet drops to the characteristics of the ST-MPC motivated the solution of a different problem, which is the detection of replay attacks, see the definition in (1.1.2). Replay attacks have been proved to be undetectable by any passive detection mechanism if they are performed when the plant is in a steady-state condition [5]. As a consequence, to detect such attacks, active detection mechanisms must be used. The connection between the ST-MPC able to handle packet drops/delay and the detection of replay attacks is explained by looking at the work of Ozel *et al.* in [6], where they proposed a watermarking signal obtained using intentional packet-drops performed on the control signal sent to the actuator.

In [5], Mo and Sinopoli have proposed the use of watermarked input signals to actively

authenticate the system dynamics and detect replay attacks. Miao *et al.* [34] proposed a stochastic game approach to design a switching watermarking input signal which achieves the best trade-off between detection rate and control performance degradation. In [35], Romagnoli *et al.* introduced a model inversion-based physical watermark input signal to achieve a control scheme where the control performance is predictable during attack-free operation. It is clear in the detection of replay attacks that there is a trade-off between between the detection performance and the control cost. In the second part of this chapter will address this problem.

In what follows, we will set some definitions that will be used across the chapter. Then the problem of mitigating DoS attacks in smart grids will be discussed.Then in the second part of the chapter, a dynamic mechanism of detecting replay attacks without significantly degrading the control performance of the system is discussed.

2.1 Preliminaries and Definitions

Let us consider the discrete-time nonlinear system

$$x(k+1) = f(x(k), u(k), d(k)) \tag{4}$$

where $k \in \mathbb{Z}_+ := \{0,1, \dots\}$ denotes the sampling time instants, $x(k) \in \mathbb{R}^n$ denotes the plant state, $u(k) \in \mathbb{R}^m$ denotes the control input, and $f(\cdot,\cdot,\cdot) : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^d \to \mathbb{R}^n$ denotes the plant's dynamics. Moreover, assume that $d(k) \in \mathbb{R}^d$ is a bounded disturbance. More precisely,

$$d(k) \in \mathcal{D} \subset \mathbb{R}^d, 0_d \in \mathcal{D} \tag{5}$$

We also assume that (4) is subject to the following state and input constraints

$$u(k) \in \mathcal{U}, x(k) \in \mathcal{X}, \forall k \geq 0 \tag{6}$$

where $\mathcal{U} \subset \mathbb{R}^m$ and $\mathcal{X} \subset \mathbb{R}^n$ are compact subsets with $0_m \in \mathcal{U}$ and $0_n \in \mathcal{X}$, respectively. The following definitions [32, 36] are used in the rest of the chapter:

Definition 1. *(Robust Positively Invariant (RPI) region) A set $\mathcal{T}^0 \in \mathcal{X}$ is said to be robustly positive invariant for (4) under disturbance (5) and constraints (6) if*

$$\forall x(0) \in \mathcal{T}^0,\ \exists u \in \mathcal{U} : f(x(0), u, d) \in \mathcal{T}^0,\ \forall d \in \mathcal{D} \tag{7}$$

Definition 2. *Let $\mathcal{S} \subset \mathbb{R}^n$ be a neighborhood region of the origin. The autonomous system $x(k+1) = f(x(k), d(k))$ is said to be Uniformly Ultimately Bounded (UUB) in $\mathcal{S}$ if for all $\mu > 0$ there exists $T(\mu) > 0$ such that $\forall \|x(0)\| \leq \mu \ \rightarrow x(k) \in \mathcal{S}$, $\forall d(k) \in \mathcal{D}$ and $\forall\, k \geq T(\mu)$.*

Definition 3. *(One-step robust controllable set) Given (4)-(6) and a set $\mathcal{T} \subset \mathbb{R}^n$, the set of states robustly controllable to $\mathcal{T}$ in one-step, namely $\mathcal{T}^1$, is defined as*

$$\mathcal{T}^1 := \{x \in \mathbb{R}^n : \exists u \in \mathcal{U} \text{ s.t. } f(x, u, d) \in \mathcal{T}, \forall d \in \mathcal{D}\} \tag{8}$$

Definition 4. *(Pontryagin Set Difference and Minkowski Set Sum) Given two sets $\mathcal{P} \subset \mathbb{R}^n$ and $\mathcal{Q} \subset \mathbb{R}^n$, the Pontryagin Set difference $\mathcal{P} \sim \mathcal{Q}$ is*

$$\mathcal{P} \sim \mathcal{Q} := \{x \in \mathbb{R}^n : x + q \in \mathcal{P}, \forall q \in \mathcal{Q}\} \tag{9}$$

while the Minkowski Set Sum $\mathcal{P} \oplus \mathcal{Q}$ is

$$\mathcal{P} \oplus \mathcal{Q} := \{y + z \in \mathbb{R}^n : y \in \mathcal{P}, z \in \mathcal{Q}\} \tag{10}$$

2.2 A DoS-resilient Set-Theoretic Controller for Smart Grid Applications

In this section, we extend the ST framework in [23, 30–32] and the solution in [33] to deal with the transient stability problem under DoS attack occurrences. The main advantage of the proposed solution, compared to existing solution is that finite-time robust transient stability is guaranteed regardless of any admissible DoS occurrences and despite constraints and disturbances.

2.2.1 Problem Formulation

Consider a smart grid consisting of L agents configured according to the IEEE new England 39-bus power system architecture [37]. Each agent of this grid, as illustrated in Fig. 4, consists of a generator, a phasor measurement unit (PMU) that measures generators rotor angel and its angular speed, a local generator controller, and an actuation system equipped with a fast-acting energy storage system (e.g., see the fast-acting flywheel used in [38]).

Assumption 1. *We assume that a communication infrastructure is available for data exchange throughout the entire grid. By denoting with $y_i(k)$ the measurements obtained from the i^{th} PMU and with $y_i'(k)$ the measurements received by the i^{th} local controller. We assume that finite-time DoS attack occurrences, launched by a finite energy attacker, [2] could affect the measurement channels. Therefore, we model the receiver's handling of finite-time DoS attacks as follows:*

$$y_i'(k) = y_i(k - \tau) \tag{11}$$

where $\tau \leq \bar{\tau}$ and $\bar{\tau}$ is a finite upper bound on the delay generated by the DoS attack.

By exploiting the Kron reduction [39], the i^{th} synchronous generator is modeled as a

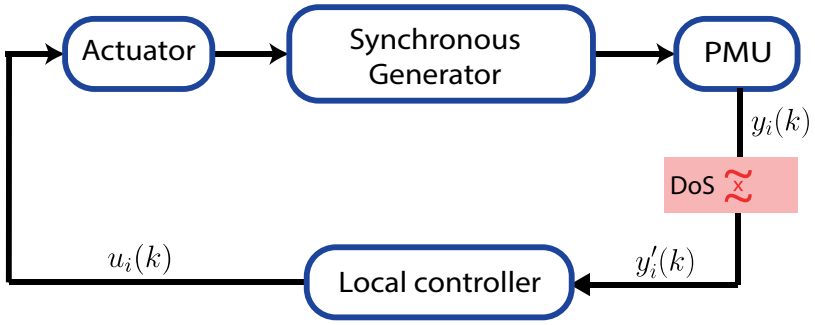


Figure 4: The i^{th} agent in the considered power grid model.

nonlinear continuous system centered around the nominal electrical frequency ω_0 as follows:

$$\begin{aligned} \dot{\delta}_i(t) &= \omega_i(t) \\ \dot{\omega}_i(t) &= \tfrac{\omega_0}{2H_i}(-\tfrac{D_i}{\omega_0}\omega_i + P_i^a(t) + u_i(t)) \end{aligned} \tag{12}$$

where $\delta_i(t)$ is the rotor angel and $\omega_i(t) = (\omega_i^{act} - \omega_0)$ denotes the angular speed deviation of the angular rotor speed ω_i^{act} w.r.t. the nominal speed ω_0, H_i denotes the generator inertia, D_i denotes the damping coefficient, and P_i^a denotes the difference between the mechanical power and the electrical power of the generator i.

To ensure phase cohesiveness, we assume that each generator i operates around an equilibrium state $x_i^{eq} = [\delta_i^*, 0]^T$, which satisfies the following requirement

$$|\delta_i^* - \delta_j^* \leq 100| \ , \ \forall \ (i,j)$$

Definition 5. *(Transient Stability) The power system is considered transiently stable if starting from a post-fault initial state $x_i(0)$, the state of each generator, i, converges to the equilibrium state x_i^{eq} [40].*

Let ϵ_ω denote the admissible frequency deviation for each generator, we describe the transient stability region as the following polyhedral set

$$\Xi_i^{ts} \ := \ \{\omega_i \ \in \ \mathbb{R} \ : \ |\omega_i| \ \leq \ \epsilon_\omega\} \tag{13}$$

Furthermore, as in [33], we extend the generator model (12) to take into account the physical limitations on the maximum power deliverable by the fast-acting power source, modeling errors and disturbances. The resulting discrete-time model is thus given by

$$\begin{aligned} x_i(k+1) &= A_i x_i(k) + B_i(u_i(k) + P_i^a(k)) + G_p d_i^p(k) \\ y_i(k) &= x_i(k) + G_m d_i^m(k) \end{aligned} \tag{14}$$

$$|u_i(k)| < \bar{P}_i^s, \;\; \bar{P}_i^s \subset \mathbb{R}^m \tag{15}$$

$$d_i^p(k) \in \mathcal{D}_i^p, \; d_i^m(k) \in \mathcal{D}_i^m \tag{16}$$

where $\bar{P}_i^s$ is the maximum deliverable power, $0_2 \in \mathcal{D}_i^p \in \mathbb{R}^2$, $0_2 \in \mathcal{D}_i^m \in \mathbb{R}^2$ disturbance sets, $x_i(k) = [\delta_i, \omega_i]^T$ is the state vector and $y(k) = [\delta_i, \omega_i]$ the measurement vector. The system dynamical matrices are given by:

$$A_i = T_s \begin{bmatrix} 1 & 1 \\ 0 & 1 - \frac{D_i}{2H_i} \end{bmatrix}, B_i = T_s \begin{bmatrix} 0 \\ \frac{\omega_0}{2H_i} \end{bmatrix} \tag{17}$$

$$G_p = T_s I_2, \; G_m = T_s I_3$$

where T_s is the sampling time.

The objective of this work can be stated as follows: given the above smart grid architecture, the constrained uncertain generators' model (14)-(16), the admissible transient region Ξ^{ts}, and an upper bound on the DoS attack duration $\bar{\tau}$, design a state-feedback control policy

$$u_i(k) := \eta_i(y_i'(k), x_i^{eq}, \bar{\tau})$$

that is capable of recovering, in a finite number of time-steps, the transient stability regardless of any admissible disturbance and DoS occurrences.

2.2.2 Proposed Controller

In traditional smart grids, a governor control scheme is used to control synchronous generators to ensure that undesired state perturbations are rejected [41]. This centralized scheme usually exhibits slow transient stability recovery time, and cannot efficiently deal with large perturbations [38]. We propose an alternative control scheme which takes advantage of the available PMUs measurements to design local decentralized robust controllers. In particular, the proposed control law is given by the sum of two contributions

$$u_i(k) \ := u_i^c(k) + u_i^f(k) \tag{18}$$

where $u_i^f(k)$ performs a partial feedback compensation for the dynamical coupling term P_i^a among the agents and $u_i^c(k)$ ensures robust transient stability in the presence of disturbances and DoS attacks.

In the next subsections, first $u_i^f(k)$ and $u_i^c(k)$ are designed, then the complete control algorithm is summarized.

Partial Coupling compensation Controller (u_i^f)

To dynamically decouple the generator dynamics, we resort to a a well-established parametric feedback linearization technique [42]. In particular, by following the idea in [43], we take advantage of the available PMU measurements to partially compensate the coupling term P_i^a in (14) as follows:

$$u_i^f(k) = -\hat{P}_i^a(k) \tag{19}$$

where $\hat{P}_i^a$ is an estimation of P_i^a at the time instant k given the information from the PMUs, i.e.,

$$\hat{P}_i{}^a(k) = P_i^a(k) + e_{P_i^a}(k), \ \ e_{P_i^a}(k) \in \mathcal{D}_i^{P_i^a}$$

and $\mathcal{D}_i^{P_i^a}$ is a bounded estimation error. As a result, $u_i^f(k)$ preforms partial compensation for the coupling term P_i^a.

By assuming that the flywheel can compensate $\hat{P}_i^a$, i.e., $\bar{P}_i^s \geq \hat{P}_i^a = \max \hat{P}_i^a$, and by substituting u_i^f into (14), we obtain the following decoupled generator dynamics

$$x_i(k+1) = A_i x_i(k) + B_i(u_i^c(k) + e_{P_i^a}(k)) + G_p d_i^p(k) \tag{20}$$

$$|u_i^c(k)| \leq \bar{\mathcal{U}}_i^c \tag{21}$$

where $\bar{\mathcal{U}}_i^c := \bar{P}_i^s - \hat{P}_i^a$.

Resilient Command Input (u_i^c)

The objective of the control action $u_i^c(k)$ is to achieve transient stability under DoS attacks, given the constrained system model (20), and in spite of imperfect coupling cancellation, disturbances and DoS attacks realizations. The proposed controller extends the robust set-theoretic MPC controller in [30]. In particular, we propose a dual-model set-theoretic controller based on the construction of *(i)* a terminal RPI region and associated terminal controller, and *(ii)* a family of robust one-step controllable sets.

The terminal RPI region, namely $\mathcal{T}_i^0$ is computed to ensure that when $x_i(k)$ enters $\mathcal{T}_i^0$, then the transient stability is preserved for any future time instant regardless any admissible disturbance and DoS occurrence. The latter is obtained by means of the following terminal control law applied for any $x_i(k) \in \mathcal{T}_i^0$:

$$u_i^c(k) = K_i^0(x_i(k) - x_i^{eq}) + u_i^{eq} \tag{22}$$

where the controller gain $K_i^0 \in \mathrm{R}^{m \times n}$ and $\mathcal{T}_i^0 \subseteq \Xi_i^{ts}$ are designed as prescribed in [44] to handle bounded uncertainties and finite-time delays produced by DoS attacks (see the red polyhedral region in Fig. 5).

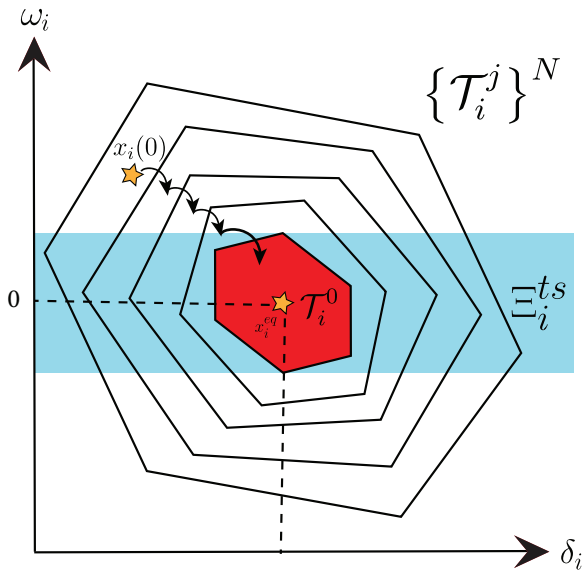


Figure 5: One-step controllable sets.

By construction, the controller (22) exhibits a small working domain ($\mathcal{T}_i^0 \subseteq \Xi_i^{ts}$). Therefore, to ensure recovery from large impulsive state perturbation, its domain must be properly enlarged. To this end, starting from $\mathcal{T}_i^0$, we construct a family of robust one-step controllable sets, namely $\{\mathcal{T}_i^j(\bar{\tau})\}_{j=1}^N$, as depicted in Fig. 5 and detailed in the following Lemma.

Lemma 1. *Consider the constrained and decoupled generator model (14)-(16), the terminal RPI region $\mathcal{T}_i^0$, and the upper bound $\bar{\tau}$ on the DoS attack duration. The family of $N > 0$ robust one-step controllable set $\left\{\mathcal{T}_i^j(\bar{\tau})\right\}_{j=1}^N$ can be recursively computed as follows:*

$$\begin{aligned}
&\mathcal{T}_i^0(\bar{\tau}) := \mathcal{T}_i^0 \subseteq \Xi_i^{ts} \\
&\mathcal{T}_i^j(\bar{\tau}) := \bigcap_{k=0}^{N} \Big\{ (x,u) \in (\mathcal{X},\mathcal{U}) : \overbrace{A_i^k}^{A(s)} x + \overbrace{\Big(\sum_{n=0}^{k-1} A_i^n B_i\Big)}^{B(s)} u \subseteq \tilde{\mathcal{T}}_i^{j-1} \Big\}
\end{aligned} \tag{23}$$

with

$$\tilde{\mathcal{T}}_i^j = \mathcal{T}_i^j \sim \bigcup_{k=1}^{\bar{\tau}} A_i^{k-1} \left(G_p \mathcal{D}_i^p \oplus B_i \mathcal{D}_i^{P_i^a} \oplus G_m \mathcal{D}_i^m \right) \tag{24}$$

Recursion (23) ensures that for any $y_i'(k) \in \mathcal{T}_i^j(\bar{\tau})$ there exists an admissible control command $u_i^c(k)$ that can be held constant for $\bar{\tau}$ time instants during DoS attack occurrences. Moreover, such control input bounds the state trajectory into the successive one-step controllable set, namely $\mathcal{T}_i^{j-1}(\bar{\tau})$. Given the family $\left\{\mathcal{T}_i^j(\bar{\tau})\right\}_{j=1}^N$ of robust one-step controllable sets, the control input $u_i^c(k)$ can be computed by means of the following convex optimization problem:

$$\begin{aligned} u_i^c(k) = \arg\min_u \ ||Ay_i'(k) + Bu||_2^2, \ \text{s.t.} \\ A(s)y_i'(k) + B(s)u \in \tilde{\mathcal{T}}_i^{j-1}, \ u \in \mathcal{U} \\ s = 1, ..., \bar{\tau} \end{aligned} \tag{25}$$

Consequently, for any initial perturbation $x_i(0) \in \{\mathcal{T}_i^j(\bar{\tau})\}_{j=1}^N$, the control law (25) ensures that the terminal region $\mathcal{T}_i^0$ can be reached in a finite number of steps in spite of any DoS realization with $\tau \leq \bar{\tau}$. Moreover, since $\mathcal{T}_i^0 \subseteq \Xi_i^{ts}$, transient stability recovery is also ensured in a finite number of steps which in the worst-case scenario is equal to $(N–1)\bar{\tau}$.

Finally, all the above developments can be summarized in the following computational algorithm.

Set-theoretic MPC of the i^{th} generator

– Offline phase –

Input: $\bar{\tau}$, system parameters in $(14) - (16)$

Output: $\left\{\mathcal{T}_i^j(\bar{\tau})\right\}_{j=1}^N$

1: Compute the RPI region $\mathcal{T}_i^0$ satisfying $\mathcal{T}_i^0 \subset \Xi_i^{ts}$, $u_i^c(k) \in \bar{\mathcal{U}}_i^c$ and the associated control function $u_i^c(k)$, according to [44].

2: Compute the family of N one-step controllable sets according the recursion (23).

3: Store $\left\{\mathcal{T}_i^j(\bar{\tau})\right\}_{j=1}^N$

– Online phase –

Input: $\left\{\mathcal{T}_i^j(\bar{\tau})\right\}_{j=1}^N, x_i(0) \in \bigcup_{j=0}^N \mathcal{T}_i^j$

Output: $u_i(k)$

1: Find the smallest set index $j(k)$ containing $y_i'(k)$, i.e.

$$j(k) := \min_{j \in \{0,\dots,N\}} j \quad s.t.\ y_i'(k) \in \mathcal{T}_i^j(\bar{\tau})$$

2: **if** $j(k) = 0$ **then** ▷ **terminal region**

3: Compute $u_i^c(k) = K_i(y_i'(k) - x_i^{eq}) + u_i^{eq}$

4: **else** Solve the convex optimization problem (25)

5: **end if**

6: Apply $u_i(k) = u_i^f(k) + u_i^c(k)$

7: $k \leftarrow k + 1$ goto Step 1

2.2.3 Simulation Results

In this section, using simulations, we evaluate the proposed system performance and compare the results with the parametric feedback linearization (PFL) controller proposed in [43], in terms of the time required to achieve transient stability under DoS attacks. The New England 10-generator 39-bus system has been used as a test-case. This system has been simulated in the Matlab environment using the parameters given in [43]. Moreover, the MPT3 toolbox [45] has been used to implement the proposed set-theoretic controller.

By using a sampling time $T_s = 0.2\,\mathrm{sec}$ [43], the discrete-time dynamics (14) of the i^{th} generator, are defined by the following matrices:

$$A_i = \begin{bmatrix} 1 & 0.2 \\ 0 & 0.997 \end{bmatrix},\ B_i = \begin{bmatrix} 0.0075 \\ 0.0755 \end{bmatrix}$$

We assume that the available external fast acting power storage imposes the following constraints on the control input u_i^c:

$$|u_i^c| \leq 10\, p.u. \tag{26}$$

For simplicity, we use the Minkowski set sum to model the cumulative effect of the noises and disturbances

$$d(k) \in \mathcal{D} \in \mathbb{R}, \quad \mathcal{D} := G_p \mathcal{D}_i^p \oplus B_i \mathcal{D}_i^{P_i^a} \oplus G_m \mathcal{D}_i^m$$

The following bounds on the disturbance are then applied to $d(k) = [d_1(k), d_2(k)]^T$, $|d_1(k)| \leq 0.01$, $|d_2(k)| \leq 0.01$. The transient stability region is chosen to describe a 0.2% variation of the normalized rotor speed

$$\Xi_i^{ts} := \left\{ \omega_i \in \mathbb{R} : |\omega_i| \leq 0.8 \right\}$$

and the maximum delay on the measurement channel, caused by a DoS attack, is upper bounded by $\bar{\tau} = 3$ (i.e. packets drop for $0.6\,\mathrm{sec}$).

The proposed resilient set-theoretic controller has been designed according to (23). In particular, a family of $N = 37$ one-step controllable sets, with $\mathcal{T}_i^0 \subseteq \Xi_i^{ts}$, have been offline computed. The resulting controller domain of attraction (DoA $= \bigcup_{j=0}^{37} \{\mathcal{T}_i^j\}$) is shown in Fig. 6. On the other hand, the design parameters for the controller in [43] has been set to $\alpha_i = 2.5$ and $\beta_i = 0.8$. Moreover, to satisfy the input constraint (26), input saturation has been enforced.

In what follows, for the sake of simplicity and to better understand the properties of the proposed solution, the performance of a single generator is shown when 7 DoS attack instances, each of duration $\bar{\tau} = 3$, are simulated. To compare the time to transient stability

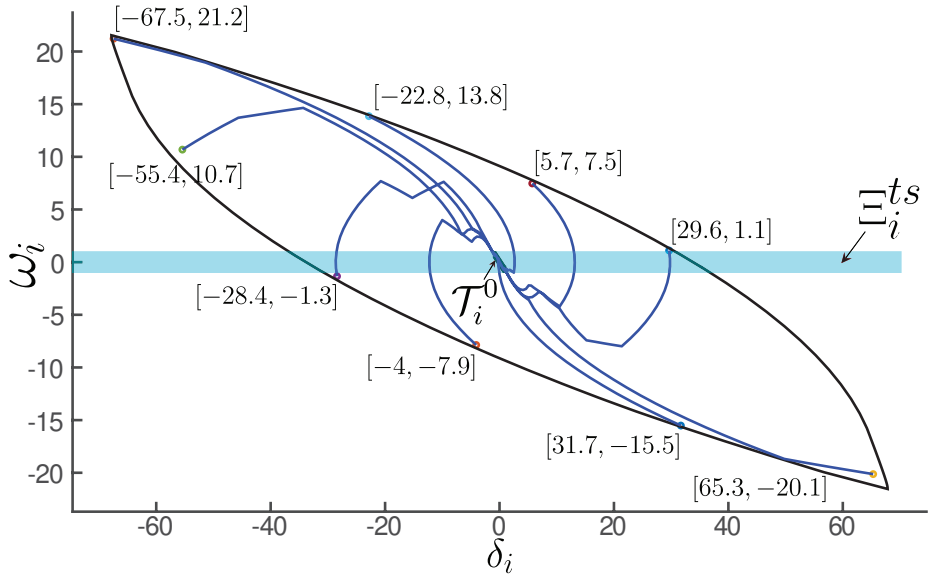


Figure 6: The Domain of Attraction of generator i. State trajectory (Blue) from outer sets to the RPI region.

obtained by the proposed controller and [43], 500 randomly generated initial perturbations have been simulated. The obtained results show that the worst-case transient stability time (t_s^{worst}) for [43] is $t_s^{worst} = 11.8\,\text{sec}$ while for the proposed resilient controller is $t_s^{worst} = 8.2\,\text{sec}$. The theoretical guaranteed worst-case time to recovery for the considered DoS attack is

$$t_s^{worst} = \min(T_s((N-1)\bar{\tau}),\ T_s((N-1)+\bar{\tau}z)) = 11.4\,\text{sec}$$

where $z = 7$ is the number of DoS occurrences used in our simulations.

The simulations results for 9 initial perturbations selected on the border the domain of attraction to depict different system responses, are also shown in Figs. 6-7.

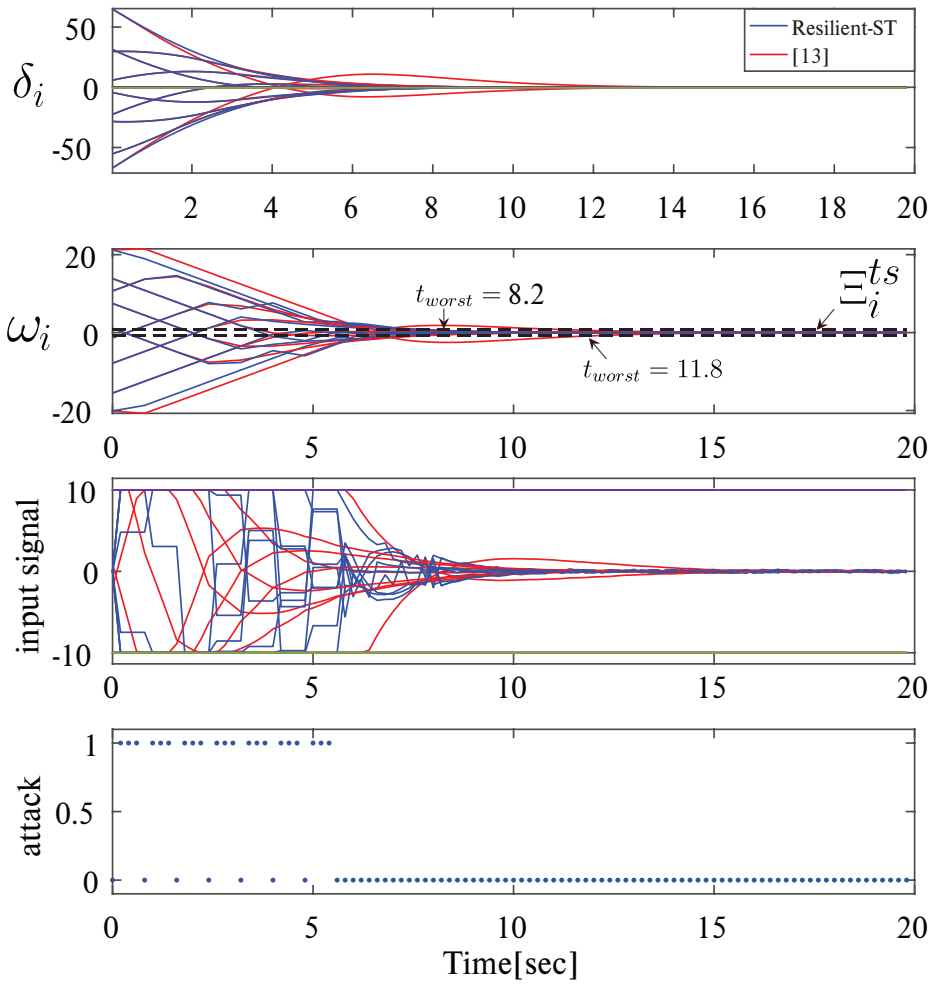


Figure 7: Rotor angle δ_i, Rotor angular speed ω_i, Command input u_i, and Attack status. Resilient Set-Theoretic (blue) vs [43] (red). In the sub-figure showing the attack, “1" represents a DoS packet drop instance, and “0" represents no attack.

2.3 Set-Theoretic Control for Active Detection of Replay Attacks with Applications to Smart Grid

Watermarking solutions usually achieve replay-attack detection at the expense of degraded control performance. In this regard, the second part of this chapter is inspired by the work in [6], proposes an improved watermarking-based detection mechanism with an a priori guaranteed control performance.

2.3.1 Problem Formulation

To explain the considered problem and our proposed solution, we consider, as an example, a smart grid consisting of L agents configured according to the IEEE new England 39-bus power system architecture [37]. As illustrated in Fig. 8, each agent of the grid consists of a generator, a phasor measurement unit (PMU) that measures the generator's rotor angle and its angular speed, a local generator controller, and an actuation system equipped with a fast-acting energy storage system (e.g., see the fast-acting flywheel used in [38]). Moreover, we assume that a communication infrastructure is available for data exchange throughout the entire grid and that a watermarking module is used to generate a watermarked control input (see Section 2.3.2).

We denote with $y_i(k)$ the measurements obtained from the i^{th} PMU and with $y_i'(k)$ the signal received by the state estimator and the detector module. $\hat{x}_i(k)$ denotes the estimated states of the system.

By exploiting the Kron reduction [39], the i^{th} synchronous generator is modeled as a continuous-time system centered around the nominal electrical frequency ω_0 as follows:

$$\begin{aligned} \dot{\delta}_i(t) &= \omega_i(t) \\ \dot{\omega}_i(t) &= \tfrac{\omega_0}{2H_i}(-\tfrac{D_i}{\omega_0}\omega_i + P_i^a(t) + u_i(t)) \end{aligned} \tag{27}$$

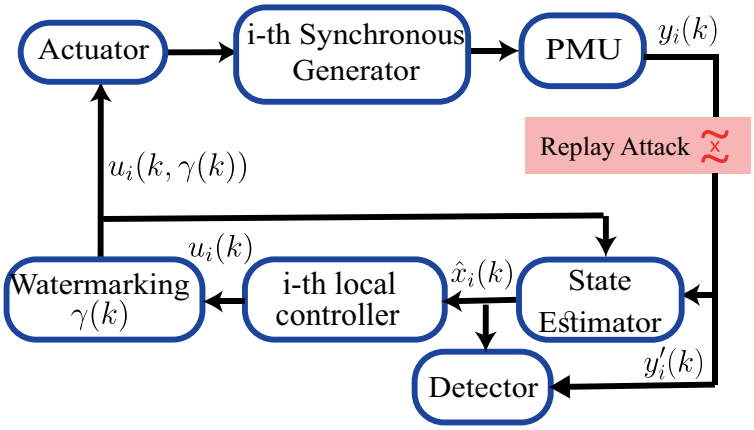


Figure 8: The i^{th} agent in the considered power grid model.

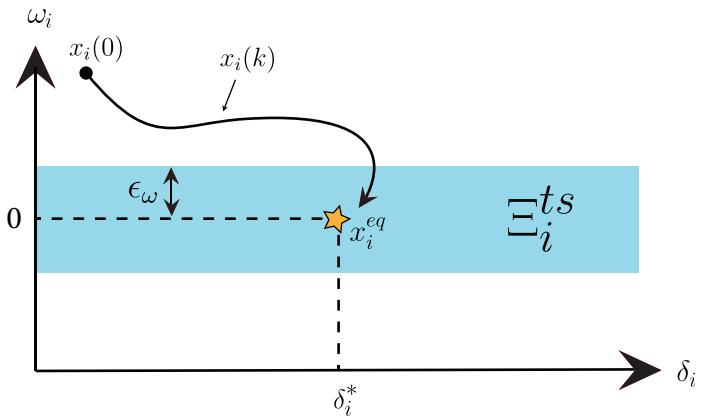


Figure 9: Transient stability region of generator i.

where $\delta_i(t)$ is the rotor angle and $\omega_i(t) = (\omega_i^{act} - \omega_0)$ denotes the angular speed deviation of the angular rotor speed ω_i^{act} w.r.t. the nominal speed ω_0, H_i denotes the generator inertia, D_i denotes the damping coefficient, and P_i^a denotes the difference between the mechanical power and the electrical power of the generator i.

To ensure phase cohesiveness, we assume that each generator i operates around an equilibrium state $x_i^{eq} = [\delta_i^*, 0]^T$, which satisfies the following requirement [40]:

$$|\delta_i^* - \delta_j^*| \leq 100 \ , \ \forall \ (i, j)$$

Definition 6. *A power system is considered transiently stable if starting from a post-fault*

initial state $x_i(0)$, the state of each $i-th$ generator (27) converges to the equilibrium state x_i^{eq} [40].

Let ϵ_ω denote the admissible frequency deviation for each generator. We describe the transient stability region as the following polyhedral set (see Fig. 9):

$$\Xi_i^{ts} \ := \ \left\{\omega_i \ \in \ \mathbb{R} \ : \ |\omega_i| \ \leq \ \epsilon_\omega\right\} \tag{28}$$

Furthermore, as in [22, 33], we extend the generator model (27) to take into account the physical limitations on the maximum power deliverable by the fast-acting power source, the modeling errors and the disturbances. The resulting discrete-time model is thus given by

$$\begin{aligned} x_i(k+1) &= A_i x_i(k) + B_i(u_i(k) + P_i^a(k)) + G_p d_i^p(k) \\ y_i(k) &= x_i(k) + G_m d_i^m(k) \end{aligned} \tag{29}$$

$$|u_i(k)| < \bar{P}_i^s, \ \ \bar{P}_i^s \in \mathbb{R} \tag{30}$$

$$d_i^p(k) \in \mathcal{D}_i^p, \ d_i^m(k) \in \mathcal{D}_i^m \tag{31}$$

where $\bar{P}_i^s$ is the maximum deliverable power, $0_2 \in \mathcal{D}_i^p \subset \mathbb{R}^2$, $0_2 \in \mathcal{D}_i^m \subset \mathbb{R}^2$ process and measurements disturbance sets, and $d_i^p(k)$, $d_i^m(k)$ are independent truncated Gaussian random variables ($d_i^p(k) \sim \mathcal{N}(0, Q)$, $d_i^m(k) \sim \mathcal{N}(0, R)$) conditional to the compact sets $\mathcal{D}_i^p$ and $\mathcal{D}_i^m$, respectively. Moreover, $x_i(k) = [\delta_i, \omega_i]^T$ is i^{th} generator's state vector while $y_i(k)$ the i^{th} measurement vector. The discrete-time dynamical matrices are given by:

$$A_i = T_s \begin{bmatrix} 1 & 1 \\ 0 & 1 - \frac{D_i}{2H_i} \end{bmatrix}, B_i = T_s \begin{bmatrix} 0 \\ \frac{\omega_0}{2H_i} \end{bmatrix} \tag{32}$$

$$G_p = T_s I_2, \ G_m = T_s I_3$$

where T_s is the sampling time.

We use a Kalman filter as a state estimator for the linear system in (29), which is assumed to be controllable and observable, to provide a minimum mean square error state estimate $\hat{x}_i(k)$ given the previous observations of $y_i(k-1)$. Since the considered system is linear, the Kalman filter provides the best linear state estimator despite the non-Gaussian noise [46]. Thus we have

$$\begin{aligned} \hat{x}_i(k+1) &= A\hat{x}_i(k) + B(u_i(k) + P_i^a(k)) + Lz(k) \\ z(k) &= y_i'(k) - C\hat{x}_i(k) \end{aligned} \tag{33}$$

where y_i' is the received measurement signal on the controller side, $\hat{x}_i(k-1)$ is the previously estimated state of the system and $L = PC^T(CPC^T + R)^{-1}$, with P being the solution of the Riccati equation:

$$P = APA^T + Q - APC^T(CPC^T + R)^{-1}CPA^T \tag{34}$$

Moreover, we assume that a $\mathcal{X}^2$ detector is used as an anomaly detector [47]. The $\mathcal{X}^2$ function is built on the so-called residual signal $r_i(k) := y_i'(k) - C\hat{x}_i(k)$ as follows:

$$g_k = \sum_{b=k-\mathscr{L}+1}^{k} r_i(b)^T \mathscr{P}^{-1} r_i(b) \tag{35}$$

where $\mathscr{L}$ is the window detection size and $\mathscr{P}$ is the covariance of the residual signal. By defining a threshold value $\alpha > 0$, a binary anomaly detector is designed as follows:

$$g_k \underset{\mathcal{H}_1}{\overset{\mathcal{H}_0}{\lessgtr}} \alpha \tag{36}$$

where the hypothesis $\mathcal{H}_0$ refers to normal operations (or no cyber-attacks) while the hypothesis $\mathcal{H}_1$ denotes an anomaly (or cyber-attacks).

We assume that a replay attack can affect the sensors measurement vector where the

attacker can replay previously recorded data $y_i(k-\theta)$ for the attack duration $t_{attack} = k_{end} - k_{start}$, $k_{end} > k_{start} > 0$ where $\theta >$ is the delay in the replayed signal

$$y_i'(k) = y_i(k-\theta), \forall k \in \{k_{start}, \cdots, k_{end}\} \tag{37}$$

while injecting a malicious signal on the actuation channel that disrupts the stability of the system.

The objective of this work can be stated as follow:
Given the system architecture in Fig. 8, the constrained uncertain generators model (29)-(31), and the anomaly detector (36), design a state-feedback control policy

$$u_i(k) := \eta_i(y_i'(k), x_i^{eq})$$

equipped with an embedded watermarking feature, which is capable of detecting replay attack occurrences while assuring guaranteed and a-priori defined control performance in the attack-free scenario.

2.3.2 Proposed Watermarked Controller

In CPS literature, it is well-established that watermarked control signals render the detector (36) capable of detecting the presence of advanced stealthy replay attacks, e.g., see [48] for an extensive discussion and formal proofs. In its first adoption in CPS [5], the watermarking signal is assumed to be randomly generated and added on top of the optimal control input. Over the past few years, other solutions have been proposed either to try to mitigate the performance drawback [35] or to incorporate a watermarking behavior in the control signal itself [6]. In particular, following the solution in [6], a watermarked control signal for (29)

can be obtained by imposing the following dynamical evolution

$$x_i(k+1) = A_i x_i(k) + B_i(u_i(k,\gamma(k)) + P_i^a(k)) + G_p d_i^p(k) \tag{38}$$

where $\gamma(k) \in \{0,1\}$ imposes an independent and identically distributed packet-drop process on the control input signal u_i, i.e.

$$u_i(k,\gamma(k)) \ := \begin{cases} u_i(k) & \text{if } \gamma(k)=1 \\ \tilde{u}_i & \text{if } \gamma(k)=0 \end{cases} \tag{39}$$

where $\tilde{u}_i$ denotes the last successfully transmitted control input. Such technique has been proved in [6] to have an effect similar to a standard additive watermarking [5], hence enabling the detection of replay attacks (37) using the χ^2-based detector in (36).

In this section, we adopt the packets-drop idea in [6] but we design a different control policy that takes into account, in the design stage, the possibility of input packets drop. The latter allows us to derive a control strategy that, under watermarking, does not affect the system's transient stability and constraints satisfaction. In particular, at the end of the developments, we formally prove that the proposed closed-loop control system (38), under intentional drops, is subject to limited and a-priori known control performance loss that, in the attack-free scenario, leads to uniformly ultimately bounded stability.

The control strategy design proceeds as follows. First, we use standard technicalities to decouple the generators' dynamics, then we design a model predictive control (MPC) scheme for transient stability that contemplates in its design the possibility of input packets drop.

The proposed control law is given by the sum of two contributions

$$u_i(k,\gamma(k)) \ := \begin{cases} u_i^f(k) + u_i^c(k) & \text{if } \gamma(k)=1 \\ u_i^f(k) + \tilde{u}_i^c & \text{if } \gamma(k)=0 \end{cases} \tag{40}$$

where $u_i^f(k)$ performs a partial feedback compensation for the dynamical coupling term P_i^a among the agents, and $u_i^c(k)$ ensures robust transient stability in spite of imperfect coupling cancellation, bounded disturbance and input packets drop. As a consequence, according to (40), random packet drops can be performed only on $u_i^c(k)$.

Partial Coupling of compensation Controller (u_i^f)

Following [38], for the IEEE new England 39-bus power system considered in our work, we take advantage of the available PMU measurements to partially compensate the coupling term P_i^a in (29). In particular, the term $u_i^f(k)$ is designed as follows:

$$u_i^f(k) = -\hat{P}_i^a(k) \tag{41}$$

where $\hat{P}_i^a$ is an estimation of P_i^a at the time instant k given the information from the PMUs, i.e.,

$$\hat{P}_i^{\,a}(k) = P_i^a(k) + e_{P_i^a}(k), \;\; e_{P_i^a}(k) \in \mathcal{D}_i^{P_i^a}$$

and $\mathcal{D}_i^{P_i^a}$ is a bounded estimation error. As a result, $u_i^f(k)$ preforms partial compensation for the coupling term P_i^a.

By assuming that the flywheel can compensate $\hat{P}_i^a$, i.e., $\bar{P}_i^s \geq \hat{P}_i^a = \max \hat{P}_i^a$, and by substituting u_i^f into (29), we obtain the following decoupled generator dynamics

$$x_i(k+1) = A_i x_i(k) + B_i(u_i^c(k) + e_{P_i^a}(k)) + G_p d_i^p(k) \tag{42}$$

$$|u_i^c(k)| \leq \bar{\mathcal{U}}_i^c \tag{43}$$

where $\bar{\mathcal{U}}_i^c := \bar{P}_i^s - \hat{P}_i^a$.

Resilient Command Input (u_i^c)

The objective of the control action $u_i^c(k)$ is to achieve transient stability in spite of imperfect coupling cancellation, bounded disturbance and input-packet drops. To this end, we propose a dual-mode MPC controller based on the solutions proposed in [22,30]. In particular, we extend the controller in [22] by including the possibility of intentional command inputs packet drops.

According to dual-mode MPC control paradigm [30, 32], we need to build: (i) a terminal RPI region and associated terminal controller, and (ii) a family of robust one-step controllable sets.

In [22], the terminal RPI region, namely $\mathcal{T}_i^0$ is computed to ensure that when $x_i(k)$ enters $\mathcal{T}_i^0$, the transient stability is preserved for any future time instant regardless any admissible disturbance. This is obtained by first computing a stabilizing state-feedback controller

$$u_i^c(k) = K_i^0(x_i(k) - x_i^{eq}), \tag{44}$$

$K_i^0 \in \mathrm{R}^{m\times n}$, for the disturbance and constraint-free generator model. Then, the minimal terminal RPI region $\mathcal{T}_i^0 \subset \Xi_i^{ts}$ is computed utilizing the algorithm outlined in [24]. In this work, since input packets drop might be desired, the above terminal region is not assured to be an RPI w.r.t. the packet drops. To overcome such a drawback, first we analyze the maximum number of consecutive packet drops, $\bar{\tau}$, that starting from an initial condition $x_i(0) \in \mathcal{T}_i^0 \subset \Xi_i^{ts}$ does not affect the generator's transient stability, i.e. $x_i(\bar{\tau}) \in \Xi_i^{ts}$. In particular, this is computed by off-line solving the following worst-case forward reachability optimization problem

$$\bar{\tau} = \max_{\tau}, \;\; s.t. \;\; \mathcal{X}_i^{\tau} \subseteq \Xi_i^{ts} \tag{45}$$

where $\mathcal{X}_i^\tau$ is the $\tau-$steps robust forward reachable set given by the recursive definition

$$\begin{aligned}
&\mathcal{X}_i^0 = \mathcal{T}_i^0 \\
&\mathcal{X}_i^\tau = (A_i + B_i K_i^0)\mathcal{X}_i^{\tau-1} \oplus G_p \mathcal{D}_i^p \oplus B_i \mathcal{D}_i^{P_i^a} \oplus A_i G_m \mathcal{D}_i^m
\end{aligned}$$

Given the computed regions $\mathcal{T}_i^0$ and $\mathcal{X}_i^{\bar{\tau}}$, if the terminal controller (44) domain is not sufficient to either cover any initial admissible transient stability perturbations or the predicted generator's state evolution under packet drops ($\mathcal{T}_i^0 \subset \mathcal{X}_i^{\bar{\tau}}$), we compute a family of one-step controllable sets to enlarge its domain.

To this end, starting from $\mathcal{T}_i^0$, we construct a family of robust one-step controllable sets, namely $\{\mathcal{T}_i^j(\bar{\tau})\}_{j=1}^N$, taking into account possible input packet drops sequences of maximum duration $\bar{\tau}$. Such family is illustrated in Fig. 10 and computed by means of the following recursion [3]:

$$\begin{aligned}
&\mathcal{T}_i^0(\bar{\tau}) := \mathcal{T}_i^0 \\
&\mathcal{T}_i^j(\bar{\tau}) := \bigcap_{k=0}^{\bar{\tau}} \Big\{ (x,u) \in (\mathcal{X},\mathcal{U}) : \overbrace{A_i^k}^{\text{A(s)}} x + \overbrace{(\sum_{n=0}^{k-1} A_i^n B_i)u}^{\text{B(s)}} \subseteq \tilde{\mathcal{T}}_i^{j-1} \Big\}
\end{aligned} \tag{46}$$

with

$$\tilde{\mathcal{T}}_i^j = \mathcal{T}_i^j \sim \bigcup_{k=1}^{\bar{\tau}} A_i^{k-1} \big(G_p \mathcal{D}_i^p \oplus B_i \mathcal{D}_i^{P_i^a} \oplus A_i G_m \mathcal{D}_i^m \big) \tag{47}$$

Notice that, accordingly to the previous discussion, recursion (46) must guarantee that the following condition is satisfied

$$\mathcal{X}_i^{\bar{\tau}} \subseteq \bigcup_{j=0}^{N} \mathcal{T}_i^j(\bar{\tau}) \tag{48}$$

Proposition 1. *Consider the constrained and decoupled generator model (29)-(31), the terminal RPI region $\mathcal{T}_i^0$, the upper bound $\bar{\tau}$ on the packet drops sequences, the $\bar{\tau}$-step forward reachable set $\mathcal{X}_i^{\bar{\tau}}$, and the family of $N > 0$ robust one-step controllable set $\big\{\mathcal{T}_i^j(\bar{\tau})\big\}_{j=1}^N$. If*

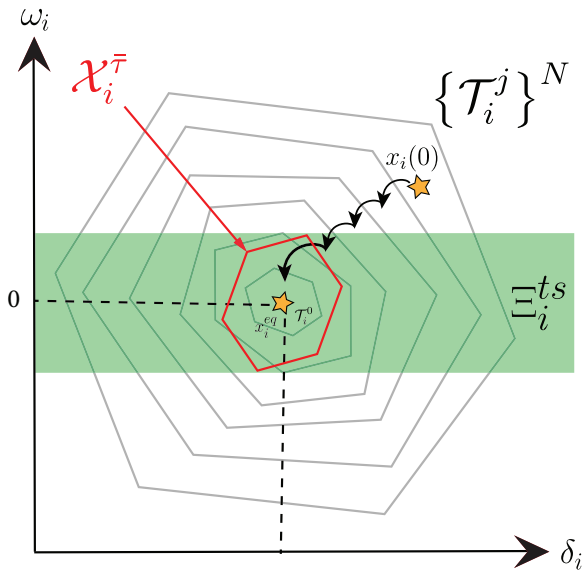


Figure 10: One-step controllable sets.

the control input is computed according to ***ST-WC*** *algorithm below, then there always exists, by construction, an admissible control command* $u_i^c(k)$ *that can be held constant for* $\bar{\tau}$ *time instants without violating the* i^{th} *generator input constraints. Moreover, the generator state trajectory* $x_i(k)$ *is uniformly ultimately bounded in* $\mathcal{X}_i^{\bar{\tau}}$ *regardless of any admissible input packets drop of duration less than* $\bar{\tau}$.

Proof. By referring to the offline construction of the family of one-step controllable set (46), and to the inclusion condition in (48), simple arguments can be given to prove that the ***ST-WC*** scheme fulfills the claim in *Proposition* 1. In particular, if the current measurement vector $y(k)$ belongs to the family of one-step controllable set $\mathcal{T}_i^j(\bar{\tau})$, $i > 0$, (see Step 4), by construction (see eq. (46)), the optimization problem (49) always admits a solution that can be applied for a number of steps equals to $\bar{\tau}$. As a consequence, the generator's input constraints are fulfilled and the trajectory is confined, in the worst-case scenario, in $\mathcal{T}_{i-1}^j(\bar{\tau})$. Moreover, by interatively applying the same arguments, in a finite number of steps, the state trajectory will reach the terminal RPI region $\mathcal{T}_i^0$. On the other hand, when $y(k) \in \mathcal{T}_i^0$ (see Step 3), by construction (see eq. (48)), the terminal controller (44) assures that state trajectory is jailed either in $\mathcal{T}_i^0 \subseteq \mathcal{X}_i^{\bar{\tau}}$ (in the absence of input packet drops) or in $\mathcal{X}_i^{\bar{\tau}}$ (in

the presence of packet drops of duration less than $\bar{\tau}$). As a consequence, the generator's trajectory is in the worst-case scenario uniformly ultimately bounded in an region $\mathcal{X}_i^{\bar{\tau}} \subseteq \Xi_i^{ts}$ contained in the desired transient stability region. □

Set-Theoretic Watermarked Control (ST-WC) algorithm

1: Find the smallest set index $j(k)$ containing $y(k)$:

$$j(k) = \min j \;\; s.t. \;\; y_i'(k) \in \mathcal{T}_i^j(\bar{\tau})$$

2: **if** $j(k) = 0$ **then**

3: Compute $u_i^c(k) = K_i(y_i'(k) - x_i^{eq})$

4: **else** Solve the convex optimization problem:

$$\begin{aligned} u_i^c(k) = \arg\min_u \;\; & ||Ay_i'(k) + Bu||_2^2, \text{ s.t.} \\ A(s)y_i'(k) + B(s)u \in \tilde{\mathcal{T}}(\bar{\tau})_i^{j-1}, \;\; & u \in \mathcal{U} \\ s = 1, ..., \bar{\tau} & \end{aligned} \tag{49}$$

5: **end if**

6: **if** $\gamma(k) == 1$ **then** ▷ *watermarking*

7: $u_i(k, \gamma(k)) = u_i^f(k) + u_i^c(k)$

8: $\tilde{u}_i^c = u_i^c(k)$

9: **else** $u_i(k, \gamma(k)) = u_i^f(k) + \tilde{u}_i^c$

10: **end if**

11: Apply $u_i(k, \gamma(k)), \; k \leftarrow k + 1$, and go to Step 1

2.3.3 Simulation Results

In this section, using simulations, we evaluate the performance of the proposed detection and control strategy under replay attacks. The results are also compared with the solution proposed in [6]. We have used, as testbed scenario, the New England 10-generator 39-bus system described in [38]. The system has been coded in Matlab where the MPT3 toolbox [45] has been used to implement the proposed set-theoretic controller. The discrete-time dynamics (29) of the i^{th} generator (29) have been obtained using a sampling time $T_s = 0.2\,\text{sec}$. The system matrices are given by:

$$A_i = \begin{bmatrix} 1 & 0.2 \\ 0 & 0.997 \end{bmatrix},\ B_i = \begin{bmatrix} 0.0075 \\ 0.0755 \end{bmatrix}$$

We assume that the available external fast acting power storage imposes the following constraints on the control input u_i^c:

$$|u_i^c| \leq 10\,p.u. \tag{50}$$

For simplicity, we use the Minkowski set sum to model the cumulative effect of the noises and disturbances

$$d(k) \in \mathcal{D} \in \mathbb{R}, \quad \mathcal{D} := G_p \mathcal{D}_i^p \oplus B_i \mathcal{D}_i^{P_i^a} \oplus A_i G_m \mathcal{D}_i^m$$

and the following upper bounds have been considered $d(k) = [d_1(k), d_2(k)]^T$

$$|d_1(k)| \ \leq 0.01,\ |d_2(k)| \ \leq 0.01$$

An extensive simulation was carried out to compare the performance of both controllers, examine the relationship between the packet drop rate and the probability of detection, and compare the probability of detection for both systems. By considering a replay attack occurrence of duration $25\,\mathrm{sec}$, we ran patches of 500 hundred simulations, each one is of 500 time steps (100 seconds), and with different packet drops $\gamma(k)$ realization and false alarm rates.

According to the ***ST-WC*** algorithm, a family $\{\mathcal{T}(3)\}_{j=0}^{39}$ of $N = 40$ robust one-step controllable sets has been computed by assuming a maximum number of consecutive input-packet drops equal to $\bar{\tau} = 3$. Moreover, the packet drop process $\gamma(k)$ has been implemented as a Bernoulli process where we enforced that the maximum number of consecutive packet drops is equal to the maximum delay $\bar{\tau} = 3$ tolerated by the proposed controller. In what follows, we denote by P_d the effective rate of packet drops after enforcing this constraint.

The detector performance have been evaluated by computing the probability of detection as a function the false alarm rate. On the other hand, the control performance degradation (due to the packet-drops) have been measured using the following cost index

$$J_e = \frac{\sum_{k=1}^{N_T} \|x_i(k) - x_{eq}\|^2 + \|u_i(k)\|^2}{N_T} \tag{51}$$

where N_T denotes the number of steps across which the system performance are evaluated. In particular, the cost (51) takes into account both the state-deviation from the equilibrium configuration and the control effort.

The obtained numerical results are depicted in Figs. 11-12. As shown in the figures, increasing the packet drop rate, namely P_d, improves the detection rate (Fig. 12) at the expense of degrading the system performance J_e (Fig. 11). Moreover, Fig. 11 contrasts, in terms of control performance loss in the attack-free scenario, the proposed watermarked solution and the one in [6] at a false alarm rate of 0.1. As shown in the figure, the proposed watermarked controller exhibits performance degradation lower than the standard LQG

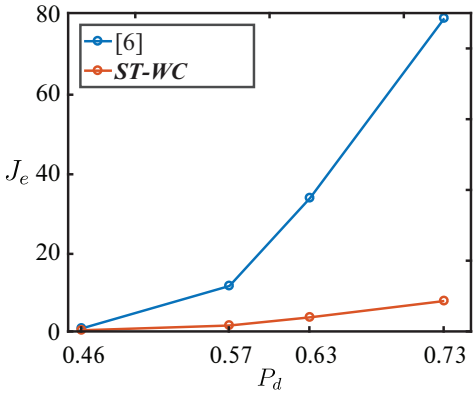


Figure 11: System performance J_e vs Packet drop rate P_d.

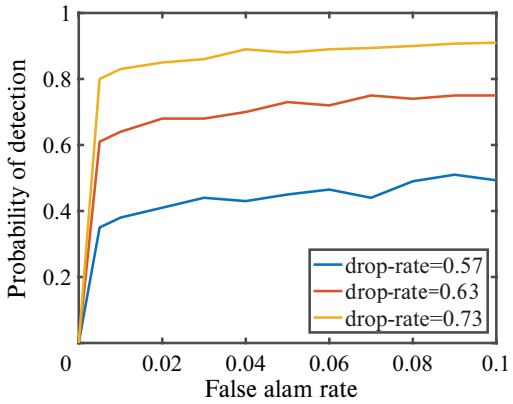


Figure 12: Detection rate vs False alarm rate

strategy used in [6]. The latter finds mainly justification in the nature of the proposed set-theoretic controller that takes into account into the design stage of the possibility of input-packet drops. On the other hand, the LQG controller in [6] is not built to assure graceful performance degradation.

Chapter 3

Decoy-based Moving Target Defense Against Cyber-physical Attacks on Smart Grid

The decoy-based scheme proposed in this chapter is published as a conference paper in EPEC 2020, see [49].

An attacker seeking to create a high impact cyber-physical attack against a smart grid must undertake planning and research to conduct an effective attack on its power system target. The steps an attacker must undertake can be described by the industrial control system (ICS) cyber kill chain [50]. In the first stage of this cyber kill chain, the attacker conducts reconnaissance to understand as much as possible about the target. Advanced undetectable attacks, such as covert attacks [18, 51], zero-dynamics attacks [21], can be launched only if a good model of the target system is available. Consequently, in the reconnaissance phase of such cyber-physical attacks on a power system, the attacker usually needs to perform an accurate identification of the dynamics of the underlying control system.

A possible way for the attacker to perform the system identification process is to intercept the communications (command inputs and sensor measurements) between the plant and the controller in a SCADA system and then apply a system identification [52] procedure on the available input/output data. Therefore, preventing the attacker from recovering the model parameters of smart grid power systems can potentially help improve the resistance of such systems to a multitude of sophisticated cyber-attacks [8, 53].

In this work, we design a novel moving target defense solution that mitigates the drawbacks of competitor schemes by resorting to the concept of decoy systems. In particular, a finite number of decoy subsystems, in parallel to the real plant, is used to deceive an eavesdropper about the real dynamical behavior of the system. Moreover, contrary to currently existing proposals, the outputs of the decoy and sensor measurements are not coupled, but they are randomly permuted and sent to the controller. The controller in return, computes the appropriate control action for each received set of measurements and send them to the plant. As a consequence, if the decoy subsystems are designed to be indistinguishable from the real system, it will be significantly hard to determine the corresponding pairs of measurements/control output that correspond to the real plant.

3.1 System Setup and Problem Formulation

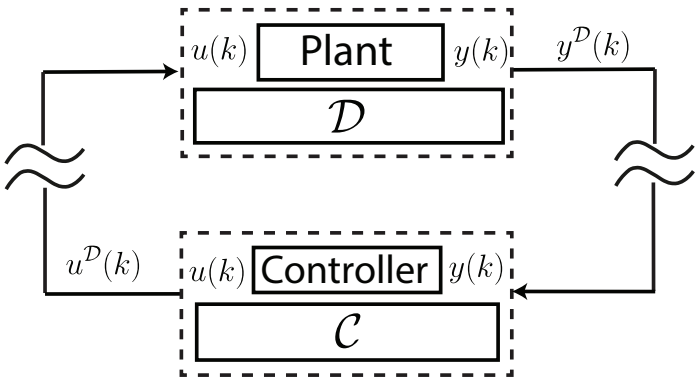


Figure 13: Considered Control System Setup

In this section, first, the considered setup is described (see Fig. 13), and then the problem formulation is stated.

The plant is described by a discrete-time dynamical system

$$\begin{aligned} x(k+1) &= f(x(k), u(k), d_p(k)) \\ y(k) &= g(x(k), d_m(k)) \end{aligned} \qquad x(0) = x^0 \tag{52}$$

where $k \in \mathbb{Z}_+ := \{0, 1, \ldots\}$ is the discrete-time index, $x(k) \in \mathbb{R}^n$, $y(k) \in \mathbb{R}^p$, $u(k) \in \mathbb{R}^m$ are the state, measurements and control inputs vectors, respectively, and x^0 is the plant initial state. Moreover, $f(\cdot,\cdot,\cdot) : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}^n$ and $g(\cdot,\cdot) : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^p$ denote the system's dynamics and measurement functions, and $d_p \in \mathcal{N}(0_n, \Sigma_p)$ and $d_m \in \mathcal{N}(0_m, \Sigma_m)$ are independent and identically distributed (i.i.d.) Gaussian process and measurement noises, with zero mean and covariance matrix Σ_p and Σ_m.

Assumption 2. *Inspired by the moving-target ideas in [10, 54, 55], we assume that an auxiliary system, namely $\mathcal{D}$, can be locally added to the plant. Such a system intercepts the received control vector, namely $u^{\mathcal{D}}(k) \in \mathbb{R}^{m_d}$, and the plant's measurement vector $y(k)$, and capable of performing the following operations*

$$\mathcal{D}: \begin{array}{ll} u(k) = \gamma_d(u^{\mathcal{D}}(k)), & \gamma_d(\cdot) : \mathbb{R}^{m_d} \to \mathbb{R}^m \\ y^{\mathcal{D}}(k) = \eta_d(y(k), u^{\mathcal{D}}(k)), & \eta_d(\cdot,\cdot) : \mathbb{R}^p \times \mathbb{R}^{m_d} \to \mathbb{R}^{p_d} \end{array} \tag{53}$$

where $\gamma_d(\cdot)$ is a function extracting the plant control input $u(k)$ from $u^{\mathcal{D}}(k)$ and $\eta_d(\cdot,\cdot)$ a function embedding the plant's measurement vector $y(k)$ into a new vector $y^{\mathcal{D}}(k)$ which is transmitted to the controller.

□

Assumption 3. *We assume that the plant (52) is stabilized by the controller's logic*

$$u(k) = h(y(k)), \qquad h(\cdot) : \mathbb{R}^p \to \mathbb{R}^m \tag{54}$$

Moreover, according to the operations performed by $\mathcal{D}$, the subsystem $\mathcal{C}$ intercepts the received measurements $y^{\mathcal{D}}(k)$ and the control input vector $u(k)$ computed by (54), and perform the following operations

$$\mathcal{C}: \begin{array}{ll} u^{\mathcal{D}}(k)) = \gamma_c(u(k), y^{\mathcal{D}}(k)), & \gamma_c(\cdot) : \mathbb{R}^m \times \mathbb{R}^{p_d} \to \mathbb{R}^{m_d} \\ y(k) = \eta_c(y^{\mathcal{D}}(k)), & \eta_c(\cdot,\cdot) : \mathbb{R}^{p_d} \to \mathbb{R}^p \end{array} \tag{55}$$

where $\eta_c(\cdot)$ is a function extracting the sensor measurement $y(k)$ from $y^{\mathcal{D}}(k)$ and $\gamma_c(\cdot,\cdot)$ is a function embedding the control input vector $u(k)$ into a new vector $u^{\mathcal{D}}(k)$ which is transmitted to the plant. □

The objective of this work can be formally stated as follow: Given the control system in Fig. 13, the system model (52), design the auxiliary systems $\mathcal{D}$ and $\mathcal{C}$ such that:

(O1): the closed-loop system performance are not affected by the operations in (53) and (55);

(O2): an attacker, intercepting the transmitted measurement and control signals:

$$(u^{\mathcal{D}}(k),\ y^{\mathcal{D}}(k)),\ \forall\, k$$

is not able to accurately reconstruct the system model (52).

3.2 Proposed Decoy-Based Solution

In this section, a decoy-based solution to achieve (O1) and (O2) is presented. First, a set of $l > 0$ decoy dynamical systems is defined

$$S_j: \begin{array}{rcll} x_j(k+1) & = & f_j(x_j(k), u_j(k), d_{p_j}(k)) & \\ y_j(k) & = & g_j(x_j(k), d_{m_j}(k)) & , x_j(0) = x_j^0 \\ & & j = 1, ..., l & \end{array} \tag{56}$$

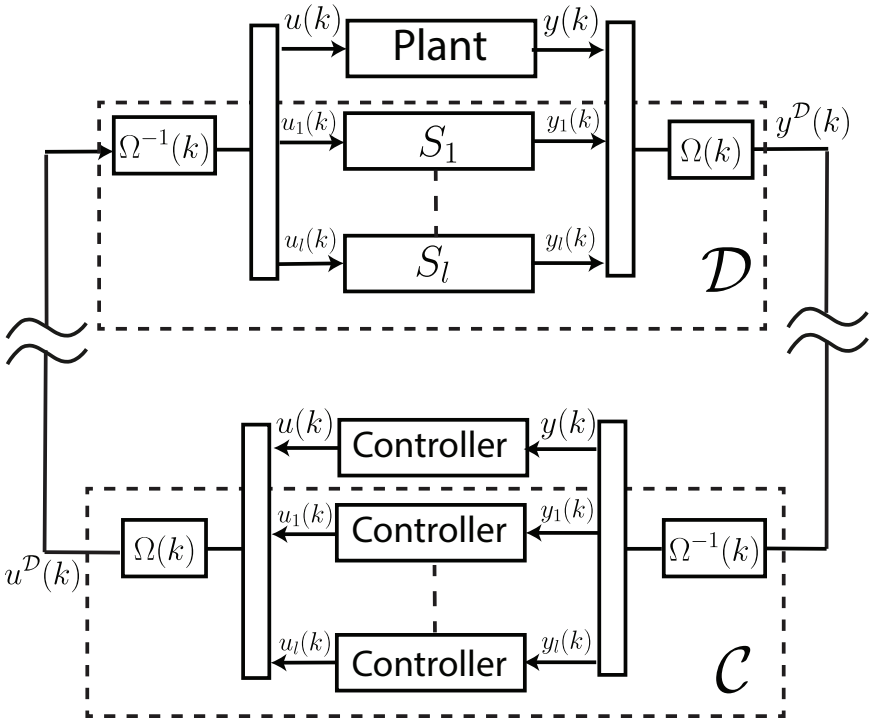


Figure 14: Decoy-based proposed solution

where all the variables have the same meaning as in (52). According to the perfect signal believability concept, each $j-th$ decoy model resembles the plant model (52), therefore we assume that the controller (54) is capable of stabilizying each decoy $\mathcal{S}_j,\ j=1,\ldots,l$.

Such dynamical systems are placed in parallel with (52), and the command inputs and the sensor measurements of the real plant are randomly combined and transmitted with the decoy input and output signals, see Fig. 14. The following auxiliary systems $\mathcal{D}$ and $\mathcal{C}$ behavior are obtained:

$$\mathcal{D}: \begin{array}{l} \left[u^T(k), u_1^T(k), \ldots, u_l^T(k)\right]^T = \Omega_c^{-1}(k) u^{\mathcal{D}}(k) \\ y_j(k) \text{ is obtained from (56), } \ j=1,\ldots,l \\ y^{\mathcal{D}}(k) = \Omega_p(k)[y^T(k),\ {y_1}^T(t),\ \ldots,\ {y_l}^T(t)]^T \end{array} \tag{57}$$

$$\mathcal{C}: \begin{array}{l} \left[y^T(k), y_1^T(k), \ldots, y_l^T(k)\right]^T = \Omega_p^{-1}(k) y^{\mathcal{D}}(k) \\ u_j(k) = h(y_j(k)),\ j = 1, \ldots, l \\ u^{\mathcal{D}}(k) = \Omega_c(k)[u^T(k), u_1^T(k), \ldots, u_l^T(k)]^T \end{array} \tag{58}$$

with $\Omega_p(k)$ and $\Omega_c(k)$ random permutation matrices. A good decoy system should make it difficult for adversaries, with access to both the measurements and control input channels to extract $y(k)$ from $y^{\mathcal{D}}(k)$, and $u(k)$ from $u^{\mathcal{D}}(k)$, to determine whether they are looking at an authentic measurement signal from the actual system or if they are indeed looking at the output of a decoy system. The indistinguishability of any particular decoy system can be measured by the adversary's failure to discern from the outputs one from the other. We formalize this by defining the following decoy system indistinguishability experiment. The experiment is defined for the measurement space Y_m with the set of decoy system measurements' $Y^{\mathcal{D}}$ such that $Y^{\mathcal{D}} \subseteq Y_m$ and $Y_m \setminus Y^{\mathcal{D}}$ is the set of authentic measurements. We note that the same rationale can be applied to the control input space U_m.

Decoy indistinguishability experiment: $\mathbf{Exp}^{ind}_{M,Y^{\mathcal{D}},Y_m}$

- For any $y_j \in Y^{\mathcal{D}}(k)$, choose two measurement signals $y_0, y_1 \in Y_m$ such that $y_0 = y_j$ or $y_1 = y_j$, and $y_0 \neq y_1$; that is, one is a decoy signal, and the second is chosen at random from the set of authentic signals (i.e., signals corresponding to the actual underlying physical system).
- Adversary M obtains y_0, y_1 and attempts to choose $\hat{y} \neq y_j$, using only information intrinsic to y_0, y_1.
- The output of the experiment is 1 if $\hat{y} \neq y_j$ and 0 otherwise.

We build upon the definition of "perfect secrecy" proposed in the cryptography community [56] and define a "perfect decoy signal" when:

$$Pr[\text{Exp}^{ind}_{M,Y^{\mathcal{D}},Y_m} = 1] = 1/2$$

That is, a perfect decoy signal is one that is completely indistinguishable from one produced by the real system.

Remark 1. *The proposed solution presents the following properties:*

- *The control system performance is not affected by the decoy-based auxiliary systems $\mathcal{D}$ and $\mathcal{C}$;*
- *An eavesdropper intercepting $y^{\mathcal{D}}(k)$ and $u^{\mathcal{D}}(k)$ cannot discriminated between the real plant pair $(u(k), y(k))$ and the decoy pairs $(u_j(k), y_j(k))$, $j = 1, \ldots, l$.*
- *The proposed decoy mechanism has two benefits to the security of the system. First, it prevents system model identification; second, it enables the detection of False Data Injection (FDI) attacks on the measurement and actuation channel since the controller can deterministically calculate the expected received measurements corresponding to the decoys, and any mismatch would indicate an FDI attack on the system.*

3.3 Decoy Defense Strategy Against AGC Model Identification

In this section, first, we present an Automatic Generation Control (AGC) system as an application model. Then, the decoy defense strategy is customized for the considered application.

3.3.1 AGC Model

We consider a single area AGC system [57], Fig. 15, where the transfer function between $u = \Delta P_{ref}$ (control input) and Δ_f (frequency deviation) is

$$T(s) = \frac{(1+\tau_g s)(1+\tau_T s)}{(2Hs+D)(1+\tau_g s)(1+\tau_T s)+1/R} \tag{59}$$

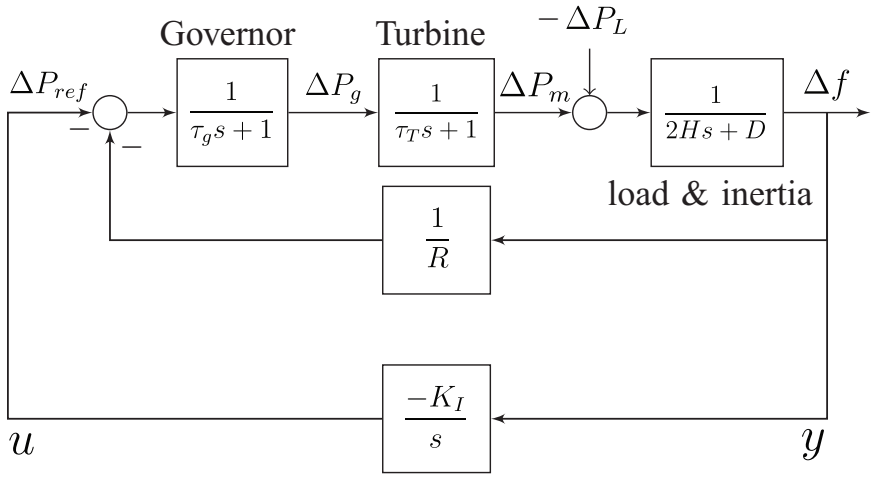


Figure 15: Block diagram of the considered AGC system.

and where ΔP_L is the load change, D is the frequency sensitivity load coefficient, H is the governor inertia constant, τ_g is the governor time constant, τ_T is the turbine time constant, and R is the governor speed regulation.

By resorting to the Controllable Canonical Form (CCF) and to a zero-order hold discretization method, the following state-space representation of (59) is considered

$$\begin{aligned} x(k+1) &= Ax(k) + Bu(k) + d_p(k) \\ y(k) &= Cx(k) + d_m(k) \end{aligned}, \quad x(0) = x^0 \tag{60}$$

where $d_p \in \mathcal{N}(0_3, \Sigma_p)$ and $d_m \in \mathcal{N}(0, \Sigma_m)$ model independent and identically distributed (i.i.d.) Gaussian process and measurement noises, with zero mean and covariance matrix Σ_p and Σ_m, respectively.

As shown in Fig. 15, the AGC controller is characterized by an integral action [57] that in the discrete-domain becomes:

$$u(k) = -K_I e(k) \tag{61}$$

where K_I is the control gain and $e(k)$ is the state of the discrete-time integrator $e(k+1) = e(k) + T_s \Delta f$, with T_s the used sampling time.

We assume that a SCADA infrastructure manages the controller logic (61) and the actuation and measurement channels. Therefore, an intruder, exploiting the absence of basic security mechanisms in the SCADA communication protocols [58], is capable of eavesdropping $u(k)$ and $y(k)$.

3.3.2 Decoy Systems

In this section, we apply the proposed decoy-based solution in section 3.2 to the AGC example model identification. In particular, each decoy has the same dynamical model of (60) but with different process and measurement noise covariance matrices and initial state vectors:

$$S_j : \begin{array}{rcll} x_j(k+1) & = & Ax(k) + Bu(k) + d_{p_j}(k) & \\ y_j(k) & = & Cx(k) + d_{m_j}(k) & , x_j(0) = x_j^0 \\ & & j = 1, ..., l & \end{array} \tag{62}$$

where $d_{p_j} \sim \mathcal{N}(0, \Sigma_{p_j})$, and $d_{m_j} \sim \mathcal{N}(0, \Sigma_{m_j})$.

3.4 Simulation Results

For simulation purpose, a single area AGC system model (59) with $D = 0.6$, $H = 5$, $R = 0.05$, $\tau_g = 0.2$, and $\tau_T = 0.5$ is considered as a case study. The discrete-time dynamics state-space model (60) of the AGC system is obtained, as explained in section

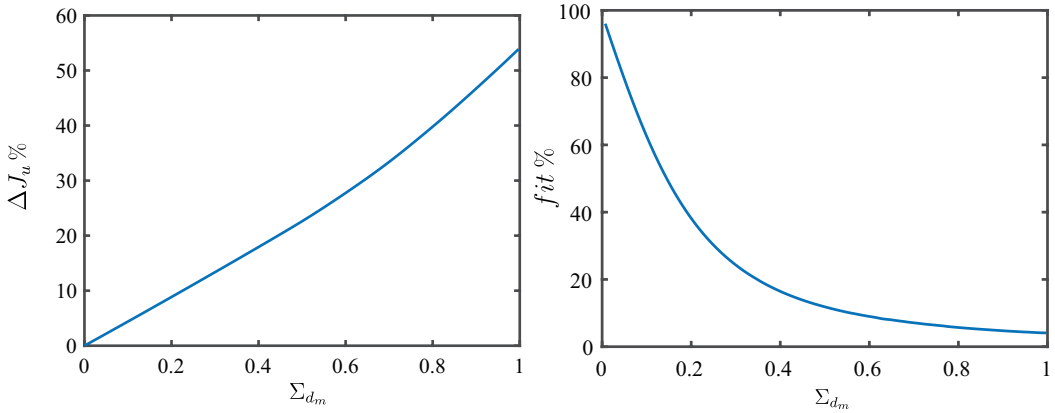
3.3.1, by using a sampling time $T_s = 0.02$, $\Sigma_p = 0.001\, I_3$, where I_3 denotes a 3×3 identity matrix, and $\Sigma_m = 0.001$. The used integral controller gain in (61) is $K_I = 0.3$.

We assume that an eavesdropper is capable of intercepting both the control and measurement signals, and intends to identify the state-space model (60) using a grey-box linear identification method [52, 59]. In the performed simulation, system identification has been performed using the built-in method provided by the Matlab system identification toolbox [60].

3.4.1 Artificial Noise on the AGC Measurements

In this subsection, we show how the accuracy of the identification AGC model (60) and the control performance are affected if the defense mechanism simply prescribes to add an artificial i.i.d. Gaussian noise $\bar{d}_m(k) \sim \mathcal{N}(0, \Sigma_{d_m})$ on top of the AGC measurements, without resorting to the proposed decoy solution, i.e.

$$\begin{aligned} x(k+1) &= Ax(k) + Bu(k) + d_p(k) \\ y(k) &= Cx(k) + d_m(k) + \bar{d}_m(k) \end{aligned} \tag{63}$$



(a) Control performance cost.

(b) Fit percentage.

Figure 16: The effect of an artificial noise on the system performance and the system identification accuracy.

In Fig 16, we have evaluated such a solution for different artificial noise variance Σ_{d_m}, and averaged the results over 500 simulation runs. For each value of Σ_{d_m}, we have evaluated the accuracy of the identified model by calculating the Normalized Root Mean Squared Error (NRMSE) fit index [61] between the outputs predicted by the identified model, namely $\bar{y}(k)$, and the AGC measurements signal $y(k)$:

$$fit = 100 \times \left(1 - \frac{\|y - \bar{y}\|}{\|y - mean(y)\|}\right)\% \tag{64}$$

where N_s is the number of steps for which the index is evaluated and $mean(y)$ denotes the average value of $y(k)$ for $0 \leq k \leq N_s$. Moreover, to evaluate the control performance loss associated with the confidentiality preserving mechanism, the following parameter is used

$$\Delta J_u = 100 \times \big(\frac{J_u - \bar{J}_u}{\bar{J}_u}\big)\% \tag{65}$$

where $\bar{J}_u$ is the control cost without adding any artificial noise, and

$$J_u = \frac{1}{N_s}\sum_{k=1}^{N_s} \|u_i(k)\| \tag{66}$$

Subplot (a) shows how increasing Σ_{d_m} degrades the control performance. On the other hand, subplot (b) shows that by increasing Σ_{d_m}, the accuracy of the identified model decreases. Based on the obtained results, it is clear that the classical solution of adding artificial noise to the measurement signal, as indicated in (63), suffers from an undesired trade-off between control performance and accuracy of the identified model.

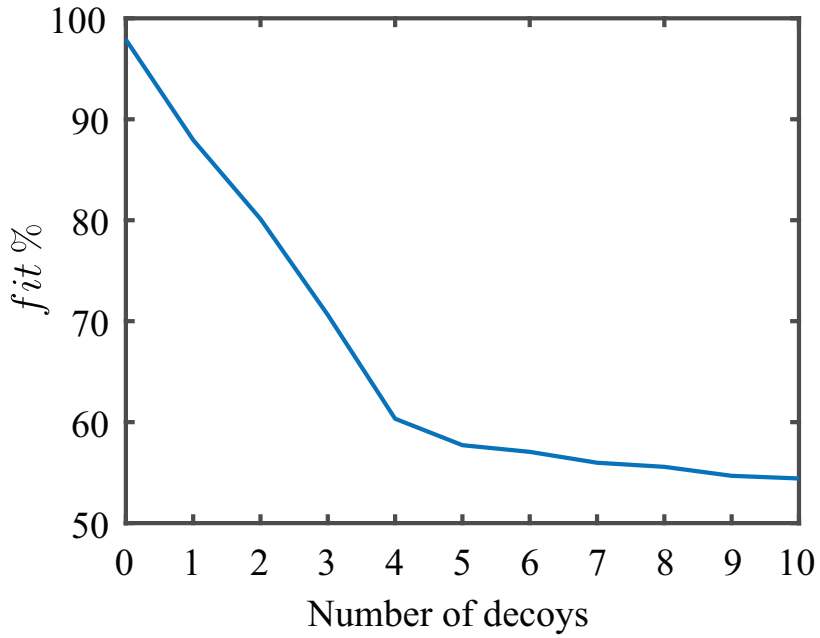


Figure 17: The decline of attacker's identification performance with respect to the number of decoy used by the defender.

3.4.2 Artificial AGC Decoys

In this subsection, the effectiveness of the proposed decoy-based strategy is evaluated. All the simulated decoy subsystems (62) are characterized by the same measurement and process noise variance $\Sigma_{m_j} = 0.025$ and $\Sigma_{p_j} = 0.002$, where $0 < j \leq 10$ and where each decoy has a unique initial state. The decoy pairs $(u_j, y_j),\ j = 1, 2, 3, ..., 10$ are, in principle, indistinguishable from $(u(k), y(k))$. Thus, at each time step, the probability of guessing the correct signal pair with l decoys is $\frac{1}{(l+1)^2}$. Fig. 17 shows how the identified model's accuracy decreases with the number of deployed decoys. Finally, contrary to the solution shown in 3.4.1, the control system performance is, by design, not affected at all by the defense mechanism.

Chapter 4

Covert Channels in Cyber-Physical Systems

The proposed covert channel technique in this chapter is published in IEEE L-CSS 2020, see [62].

A covert channel is a type of attack that allows adversaries to transfer information between entities that, according to the defined security policy, are not allowed/supposed to communicate. Lampson [63] introduced this term in 1973 while addressing the problem of confining a program during its execution so that it cannot transmit information to any other program except its caller. Since then, the covert channel problem has attracted a lot of attention in the cyber security community. For example, the "Orange Book" [64] of the US Department of Defence specifies that the system developer 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, with the objective of reducing covert channel bandwidths. The continued existence of identified covert channels in the system must be justified.

Previous research has focused on how covert channels can be established in the context

of Information Technology (IT) networks, e.g., by abusing different communication protocols and shared resources. In [65–69], timing-based covert channels have been designed utilizing time-delays to separate bits of information shared between two malicious parties. In [68, 69], storage-based covert channels are implemented exploiting shared storage or memory resources that are not designed to transfer data.Covert channels have also been designed for air-gapped machines by encoding information over a physical infrastructure that cannot be noticed with naked senses such as inaudible speaker sounds, acoustical mesh, and optical emanations (e.g., see [70, 71]).

In [72], Wendzel *et al.* demonstrated that hidden messages can be stored in CPS environments and presented two approaches for such data storage, namely by modifying unused registers of devices, and by modifying actuator states.

Recently, some works focused on establishing covert channels between devices in networked CPSs [73–77]. For example, in [73], the authors present 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 [74], the same authors present 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. Another example of covert channels in CPS, that borrows the idea of an air-gapped receiver, is described in [76]. In this work, the adversary loads malicious code onto a PLC to change actuation signals being output to the motors. The actuation signal is perturbed to transmit sensitive information covertly by creating analog acoustic channel signatures without changing the closed-loop process characteristics. In [77], a covert channel specifically designed against power grid

cyber-physical critical infrastructures through physical substrates, e.g., line loads, is proposed. Using their approach, two compromised controllers that are miles apart can coordinate their efforts by manipulating relays to modify the power network's topology.

This chapter's contribution can be summarized as follows: (i) We present a covert channel technique, enabling a compromised networked controller to leak information to an eavesdropper who has access to the measurement channel, see Fig. 19. We demonstrate that this can be achieved by properly altering the control logic, without establishing any additional explicit communication channels. Unlike [73, 74], our approach does not require any special hardware such as low response time actuators or high-quality sensors, (ii) We utilize a receding horizon set-theoretic model predictive control strategy as an illustrative example to show how an undetectable covert channel can be established, and (iii) We provide numerical simulation results to evaluate the information rate of the studied covert channel.

4.1 Problem Formulation

In this section, first, the definitions used along the chapter are given and the considered networked control system is presented. Then, the adversary model is described and the chapter objectives are stated.

4.1.1 Preliminaries and Definitions

Consider the discrete-time nonlinear systems

$$x(k+1) = f(x(k), u(k), d_p(k)) \tag{67}$$

where $k \in \mathbb{Z}_+ := \{0, 1, \ldots\}$ denotes the discrete sampling time instants, $x(k) \in \mathbb{R}^{n_x}$ is the plant state vector, $u(k) \in \mathbb{R}^{n_u}$ is the control input vector, $d_p(k)$ an unknown exogenous

bounded plant disturbance

$$d_p(k) \in \mathcal{D}_p \subset \mathbb{R}^{n_d} \tag{68}$$

with $\mathcal{D}_p$ a compact sets with $0_{n_d} \in \mathcal{D}_p$, and $f(\cdot,\cdot,\cdot) : \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \times \mathbb{R}^{n_d} \to \mathbb{R}^{n_x}$ is a continuous function describing the system's dynamics. Moreover, we assume that (67) is subject to the following state and input set-membership constraints:

$$u(k) \in \mathcal{U},\ x(k) \in \mathcal{X}, \quad \forall k \geq 0 \tag{69}$$

where $\mathcal{U} \subset \mathbb{R}^{n_u}$ and $\mathcal{X} \subset \mathbb{R}^{n_x}$ are compact subsets with $0_{n_u} \in \mathcal{U}$ and $0_{n_x} \in \mathcal{X}$, respectively.

Definition 7. *Given the system (67)-(69) and a set $\mathcal{S} \subset \mathbb{R}^{n_x}$, the set of state vectors $x^+ \in \mathbb{R}^n$ one-step reachable from $\mathcal{S}$, namely $Reach(\mathcal{S})$, is defined as*

$$\begin{aligned} Reach(\mathcal{S}) \triangleq \{x^+ \in \mathbb{R}^{n_x} : \exists\, x \in \mathcal{S},\ u \in \mathcal{U},\, d_p \in \mathcal{D}_p \text{ s.t.} \\ x^+ = f(x,u,d_p)\} \end{aligned} \tag{70}$$

□

4.1.2 Networked Control System Model

We consider the class of networked control systems shown in Fig. 18, where the plant's dynamics are described by (67)-(69) and the state space vector is observable. As in [3], we model the non perfect knowledge of $x(k)$ or the presence of bounded measurement noise by means of a bounded unknown exogenous disturbance vector $d_m(k)$, i.e.,

$$y(k) = x(k) + d_m(k), \quad d_m(k) \in \mathcal{D}_m \subset \mathbb{R}^{n_x} \tag{71}$$

with $y(t) \in \mathbb{R}^{n_x}$ the measurement vector and $\mathcal{D}_m$ a compact set with $0_n \in \mathcal{D}_m$.

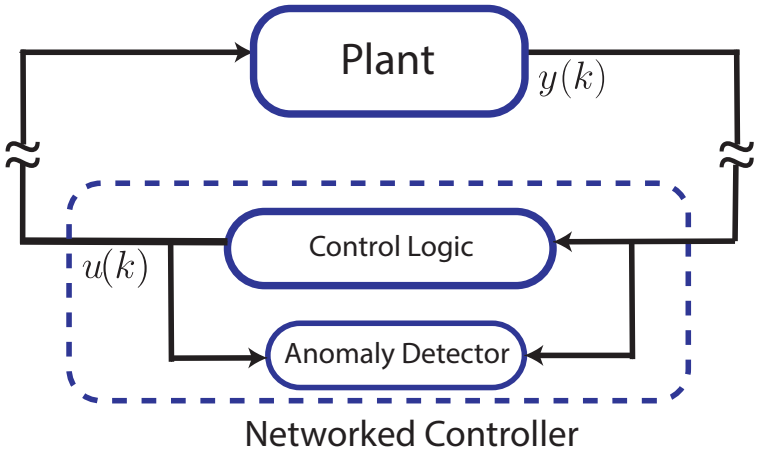


Figure 18: Networked control system.

Without loss of generality, we assume the networked controller consists of two modules: the control logic and the anomaly detector. The control logic is in charge of computing the control input $u(k)$ and its policy is described by the control law

$$u(k) = g(y(k)) \tag{72}$$

where $g(y(k))$ denotes a robustly stabilizing feedback controller complying with the plant's state and input constraints (69), see e.g. [78, 79].

On the other hand, given the uncertain bounded plant dynamics (67)-(69), (71), we assume that the anomaly detector, namely $\mathcal{D}(y(k))$, is a binary detector, $\mathcal{D}(y(k)) \in \{\text{normal, anomaly}\}$, leveraging the expected robust one-step ahead evolution of the measurement vector [3, 32], i.e.

$$\mathcal{D}(y(k)) = \begin{cases} \text{normal} & \text{If } y(k) \in \mathcal{Y}^+(y(k-1), u(k-1)) \\ \text{anomaly} & \text{Otherwise} \end{cases} \tag{73}$$

where $\mathcal{Y}^+(y(k-1), u(k-1))$ is the robust one-step output evolution predicted at $k-1$,

and $\mathcal{Y}^+(y(k), u(k))$ defined as follows

$$\begin{aligned}\mathcal{Y}^+(y(k), u(k)) \triangleq \{y^+ \in \mathbb{R}^{n_x} : \\ y^+ = f(y(k) - d_{m_1}, u(k), d_p) + d_{m_2}, \\ \forall\, d_p \in \mathcal{D}_p,\ d_{m_1}, d_{m_2} \in \mathcal{D}_m\}\end{aligned} \tag{74}$$

Remark 2. *We assume a detector module based on set-theoretic arguments [3, 32] only for the sake of clarity. However, since in this work, we consider the design of a particular cyber-attack altering the control logic itself, then the obtained results are also valid for any residual-based detector [80].*

4.1.3 Adversary Model

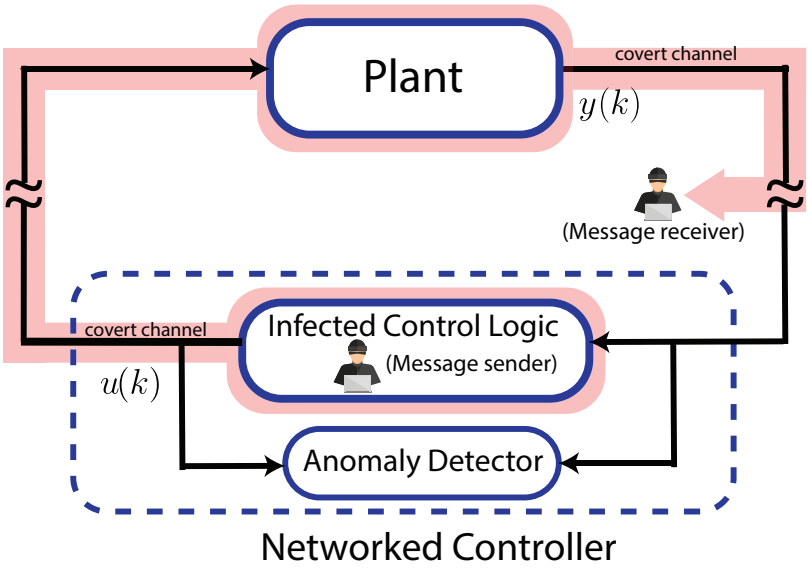


Figure 19: Covert channel in networked control systems.

We consider an attacker aiming to establish a covert channel between an intruder (sender) within the networked controller and an eavesdropper (receiver) with access to the measurement channel to exfiltrate sensitive information available within the networked control system or send secret messages possibly to coordinate for successive attacks.

To this ends, following the attacks classification given in [4], the considered cyber-attacker is assumed to possess the following assets: (i) *Model Knowledge*: The attacker is aware of the uncertain plant dynamics and constraints (67)-(69), and (71); (ii) *Disruptive Resources*: The attacker is capable of injecting malware in the controller logic and arbitrarily changing the control law, and (iii) *Disclosure Knowledge*: The attacker can read the transmitted sensor measurement $y(k)$. It should be noted that modifying the controller logic might also be achieved through a supply chain attack, e.g., by a malicious manufacturer or supplier of CPSs' equipment [81].

4.1.4 Objectives and Covert Channel Design Problem

The objective of this work is to show how a covert channel can be established in CPSs by exploiting the set of resources described above.

The attacker design problem can be formally stated as follows:

Given the plant model (67)-(69), (71), the anomaly detector rule (73), and the adversary model described in Section 4.1.3, show the existence of a covert channel such that:

- *A binary vector message of length* $p > 0$, *namely* $M = [m_1, \ldots, m_p] \in \mathbb{R}^p$, $m_i \in \{0, 1\},\ 1 \leq i \leq p$, *can be sequentially encoded in the control action* $u(k)$;
- *The control action* $u(k)$ *does not trigger the anomaly detection rule (73);*
- *The encoded message can be correctly decoded, without ambiguity, by a receiver reading the sensor measurement* $y(k)$.

4.2 Covert Channel Design

In this section, first, we show how a covert channel can be designed exploiting reachability arguments, then the operations of the covert channel transmitter and receiver are summarized into a computational algorithm and the correctness of the utilized encoding/decoding scheme is proved.

The intuition to establish a covert channel is the following. If the attacker knows the plant model (67)-(69), (71) and he/she can arbitrarily change the control logic, then the attacker can replace the legitimate control law (72) with the following switching logic embedding the binary messages m_i, $1 \leq i \leq p$

$$u(k) = \begin{cases} u_0(k) \triangleq g_0(y(k)) & \text{if} \quad m_i = 0 \\ u_1(k) \triangleq g_1(y(k)) & \text{else} \quad m_i = 1 \end{cases} \tag{75}$$

where $g_0(y(k))$ and $g_1(y(k))$ are two robustly stabilizing control laws.

Remark 3. *In principle, the switching control law (75) can be arbitrarily chosen. However, in practice, any industrial process is equipped with embedded protection mechanisms to prevent reaching hazardous plant conditions. Therefore, the switching control law must be designed to prevent instability and, as a consequence, avoid that safety mechanisms (besides the anomaly detector (73)) could shut-down the plant operations and, as a consequence, interrupt the covert channel. A possible way to address this design problem is to assure that a common robust Lyapunov function exists, see, e.g., the seminal papers [82–84] or by resorting to the robust set-theoretic model predictive framework proposed in [30]. The reader is referred to section 4.3 for a practical implementation of the switching law.* □

If the switching controller (75) is known to the eavesdropper with access to the sensor measurement $y(k+1)$, then robust output reachable set arguments, can be used to determine if $u_0(k)$ or $u_1(k)$ has been applied to the plant.

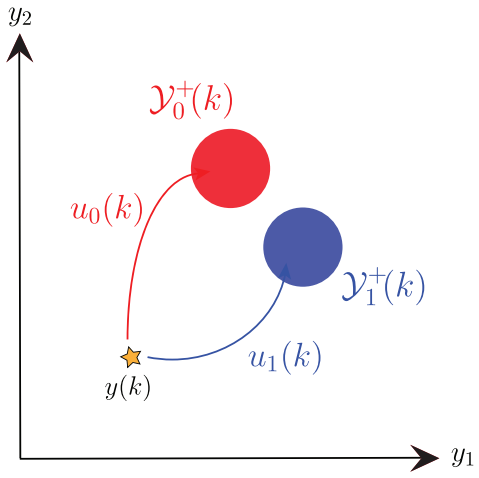


Figure 20: Robust one-step output reachable sets $\mathcal{Y}_0^+(k)$ and $\mathcal{Y}_1^+(k)$ associated to the switching control law (75).

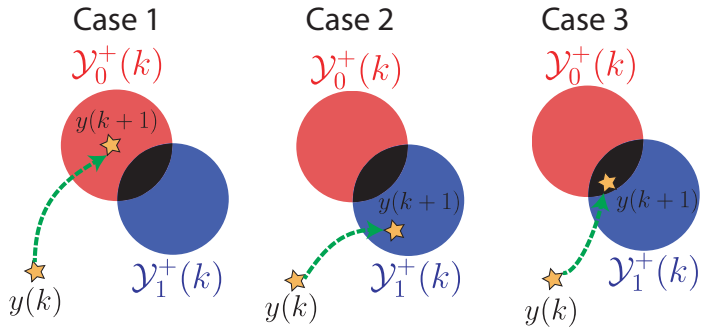


Figure 21: The eavesdropper can decode the message m_i by leveraging $y(k+1)$, $\mathcal{Y}_0^+(k)$, and $\mathcal{Y}_1^+(k)$: Case 1: $m_i = 0$, Case 2: $m_i = 1$, and Case 3: m_i undecided.

Specifically, following the definition of robust one-step output reachable set in (74), both the intruder (sender) and the eavesdropper (receiver) can compute, at each time step k, the following sets (see Fig. 20)

$$\mathcal{Y}_0^+(k) \triangleq \mathcal{Y}^+(y(k), u_0(k)), \quad \mathcal{Y}_1^+(k) \triangleq \mathcal{Y}^+(y(k), u_1(k)) \tag{76}$$

Given (76) and the sensor measurement $y(k+1)$, the following cases (see Fig. 21) can arise:

- *Case 1:* $y(k+1) \in (\mathcal{Y}_0^+(k) \setminus \mathcal{Y}_1^+(k))$. In this case, the eavesdropper determines that

the received bit is "0";

- *Case 2:* $y(k+1) \in (\mathcal{Y}_1^+(k) \setminus \mathcal{Y}_0^+(k))$. In this case, the eavesdropper determines that the received bit is "1";

- *Case 3:* $y(k+1) \in (\mathcal{Y}_1^+(k) \cap \mathcal{Y}_0^+(k))$. In this case, the eavesdropper is uncertain about the received bit.

The utilized encoding/decoding mechanism can take care of the above three cases by re-transmitting m_i whenever case 3 arises. The following algorithms summarize the infected controller (sender) encoding steps and the eavesdropper (receiver) decoding steps. It is interesting to note how the sensor measurement channel, in addition to its role in transmitting the information from the sender to the receiver, also serves as an acknowledgment feedback channel for the transmitter to ensure that the information is correctly decoded by the eavesdropper.

Covert Channel in Networked Control System (CC-NCS)

– Infected control logic (Sender) $(\forall\, k)$ *–*

Initialization: $y(0)$, the feedback control laws $g_0(\cdot),\ g_1(\cdot)$, the binary message vector $M = [m_1, \ldots, m_p]$, auxiliary index $i = 1$,

1: **if** $k > 0$ **then**
2: **if** $y(k) \notin (\mathcal{Y}_1^+(k-1) \cap \mathcal{Y}_0^+(k-1))$ **then**
3: $i = i + 1$ ▷ previous bit successfully transmitted
4: **end if**
5: **end if**
6: Compute $u_0(k) = g_0(y(k)),\ u_1(k) = g_1(y(k))$

7: Determine $\mathcal{Y}_0^+(k)$ and $\mathcal{Y}_1^+(k)$ as in (76)

8: **if** $m_i == 0$ **then**

9: $\quad u(k) = u_0(k)$

10: **else** $u(k) = u_1(k)$

11: **end if**

12: Send $u(k)$

– Receiver logic $(\forall\, k)$ *–*

Initialization: $y(0)$, $\mathcal{Y}_0^+(0)$, $\mathcal{Y}_1^+(0)$, the received message $M_r = [m_1, \ldots, m_p]$, $m_i = -1, \forall\, i$, auxiliary index $i = 1$

1: Read $y(k+1)$

2: **if** $y(k+1) \in (\mathcal{Y}_0^+(k) \setminus \mathcal{Y}_1^+(k))$ **then**

3: $\quad m_i = 0, i = i + 1$ ▷ bit 0 decoded

4: **else**

5: $\quad$ **if** $y(k+1) \in (\mathcal{Y}_1^+(k) \setminus \mathcal{Y}_0^+(k))$ **then**

6: $\quad\quad m_i = 1, i = 1 + 1$ ▷ bit 1 decoded

7: $\quad$ **else** $m_i = -1$ ▷ bit uncertain

8: $\quad$ **end if**

9: $\quad$ Determine $\mathcal{Y}_0^+(k+1)$ and $\mathcal{Y}_1^+(k+1)$ as in (76) for use at the next time step

10: **end if**

Proposition 2. *Given the networked control system shown in Fig. 18, the adversary resources detailed in Section 4.1.3, the CC-NCS algorithm described above allows the establishment of a covert channel for the transmission of binary messages from the networked controller to an eavesdropper with access to the measurement channel.*

Proof: To prove the proposition, it is sufficient to collect all the above developments and prove both the correctness and undetectability as follows:

- *Correctness:* The proposed *CC-NCS* algorithm exploits robust worst-case arguments (with respect to both the plant disturbance d_p and measurement noise d_m) to design the output reachable sets $\mathcal{Y}_0^+(k)$ and $\mathcal{Y}_1^+(k)$, see (76) and (74), associated to the feedback controller laws $u_0(y(k))$ and $u_1(y(k))$. As a consequence, the set-membership tests performed to decode the received messages are robust. Moreover, since each bit m_i is considered successfully transmitted (see Step 2 of the sender and Steps 2, 5 of the receiver) if and only if the output evolution $y(k+1)$ belongs, in a mutually exclusive fashion, to the output predicted sets ($\mathcal{Y}_0^+(k)$, $\mathcal{Y}_1^+(k)$), the encoding/decoding operations are correct.

- *Undetectability:* The purposed channel cannot be detected by monitoring system performance. This follows by noting that the attacker encodes a binary message into the control logic. As a consequence, the assumed set-based detector (73) is not able to detect the presence of anomalies since $y(k) \in \mathcal{Y}^+(y(k-1), u(k-1)),\ \forall\, k$. Along the same lines, it is straightforward to show that passive residual-based anomaly detectors [80] are ineffective. Therefore, the designed channel is by design undetectable, i.e., it is covert.

4.3 Proposed Implementation and Simulation Results

In this section, we propose a concrete implementation of the ***CC-NCS*** algorithm to establish a covert channel when the plant dynamics (67)-(69), (71) are described by the following constrained linear discrete-time system subject to bounded additive disturbances

$$\begin{aligned} x(k+1) &= Ax(k) + Bu(k) + d_p(k) \\ y(k) &= x(k) + d_m \end{aligned} \tag{77}$$

Moreover, numerical simulation results are presented to show the capability of the designed covert channel.

4.3.1 Infected Receding Horizon Model Predictive Controller

For our simulations, the networked controller is designed following the prescriptions of the dual-mode set-theoretic model predictive controller (MPC) in [30]. In particular, by following the computational scheme in [30, Sec. 4], first, a terminal stabilizing controller for the disturbance and constraint-free system model is designed as

$$u(k) = K_0 y(k), \quad K_0 \in \mathrm{R}^{m \times n} \text{ the controller gain} \tag{78}$$

and the associated minimal robust positively invariant (RPI) region $\mathcal{T}_0$ is computed [24]. Then, $\mathcal{T}_0$ is enlarged by computing a family or robust one-step controllable sets according to the following recursive definition

$$\mathcal{T}_j := \left\{ x \in \mathcal{X} : \exists u \in \mathcal{U} : \; A\,x + B\,u \in \tilde{\mathcal{T}}_{j-1} \right\}, \; j > 0 \tag{79}$$

with $\tilde{\mathcal{T}}_j = \mathcal{T}_j \sim \left(\mathcal{D}_p \oplus A\mathcal{D}_m \right)$.

Recursion (79) ensures that for any $y(k) \in \mathcal{T}_j$, there exists an admissible control command $u(k) \in \mathcal{U}$ that robustly steers the state trajectory into the successive one-step controllable set, namely $\mathcal{T}_{j-1}$, see Fig. 22. Such a control input is computed according to the following optimization problem:

$$\begin{aligned} u(k) &= \arg\min_{u} J(k), \quad \text{s.t.} \\ Ay(k) + Bu &\in \tilde{\mathcal{T}}_{j-1}, \; u \in \mathcal{U} \end{aligned} \tag{80}$$

where $J(k)$ is a cost function that can be changed at any time instant without affecting the uniformly ultimately bounded stability of the system, but it is desirable for $J(k)$ to

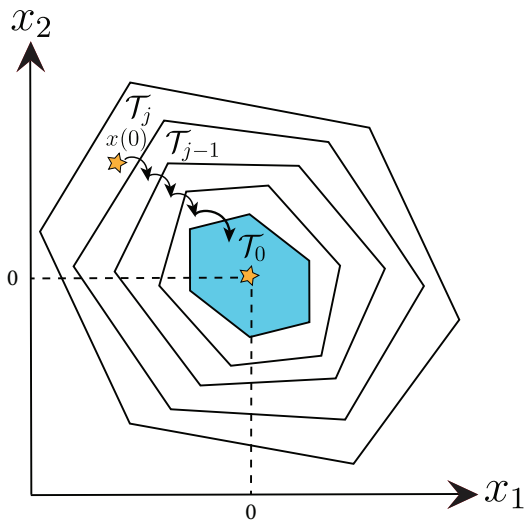


Figure 22: An illustration of the terminal region and family of one-step controllable sets.

be quadratic to have a convex optimisation problem to compute the control action online, see [30]. On the other hand, when $x(k) \in \mathcal{T}_0$, then $u(k)$ is computed as in (78).

The switching infected controller logic (75) can be obtained from (80) by using any two suitable cost functions, $J_0(k)$ and $J_1(k)$, for $u_0(k)$ and $u_1(k)$, respectively. By recalling that (80) is a Quadratic Programming (QP) optimization problem and that the existence of the solution is ensured by construction, the proposed switching controller can be computed in polynomial time.

However, in our implementation, to test the capacity of the designed covert channel (i.e., the average possible transmitted bits/sec), we compute both $u_0(k)$ and $u_1(k)$ by means of the following concave optimization problem

$$\begin{aligned} [u_0(k), u_1(k)] = \arg\max_{u_0,u_1} ||B(u_0 - u_1)||_2^2 \ \text{ s.t.} \\ Ay(k) + Bu_0 \in \tilde{\mathcal{T}}_{j-1}, \ u_0 \in \mathcal{U} \\ Ay(k) + Bu_1 \in \tilde{\mathcal{T}}_{j-1}, \ u_1 \in \mathcal{U} \end{aligned} \tag{81}$$

Although the above optimization does not guarantee that the two output evolution sets are different, the used cost function maximizes the distance between the centers of the

attacker's one-step output evolution sets $\mathcal{Y}_0^+$ and $\mathcal{Y}_1^+$, see (76). Moreover, for the linear model (77), the robust output reachable $\mathcal{Y}^+(y(k), u(k))$ computations simplify as follows

$$\mathcal{Y}^+(y(k), u(k)) = (Ay(k) + Bu(k)) \oplus (\mathcal{D}_p \oplus \mathcal{D}_m \sim A\mathcal{D}_m)$$

4.3.2 Simulation Results

The performed numerical simulation is coded in Matlab where the MPT3 toolbox [45] has been used to compute the required controllable (79) and reachable (76) sets.

The used system matrices (77), disturbance and constraint sets are:

$$A = \begin{bmatrix} 1 & 0.2 \\ 0 & 0.997 \end{bmatrix}, \; B = \begin{bmatrix} 0.0075 \\ 0.0755 \end{bmatrix}$$
$$d_p(k), d_m(k) \in \mathcal{D} = \{d \in \mathbb{R}^2 : \; |d_i| \leq 0.05, \; i = 1, 2\}$$
$$u(k) \in \mathcal{U} = \{u \in \mathbb{R} : \; |u| \leq 10\}$$

A set of simulation runs have been carried out to assess the achievable transmission rate of the covert channel. Offline, we have constructed the controller gain K_0, the terminal region $\mathcal{T}_0$, and a family of $N = 100$ one-step controllable sets $\{\mathcal{T}_j\}_{j=1}^{100}$. Then, online, each run is initiated with a random initial state $x(0)$ (belonging to the outer set of the computed family, i.e. $x(0) \in \mathcal{T}_{100}$) and carried for 70 time steps. Moreover, the disturbances $d_p(k)$, $d_m(k) \in \mathcal{D}$ are generated in Matlab by means of a uniformly distributed random generator whose initial seed is changed in each run. The infected controller binary logic (75) is calculated as in (81). The simulations results are summarized in Fig. 23 which shows a histogram for the probability of successful decoding from the first transmission trial over the 500 simulation runs (hereafter, we call it probability of successful decoding for short). The average probability is equal to 72.5%. In other words, the average number of bits that are transmitted over the covert channel during each one of these simulation runs, with 70

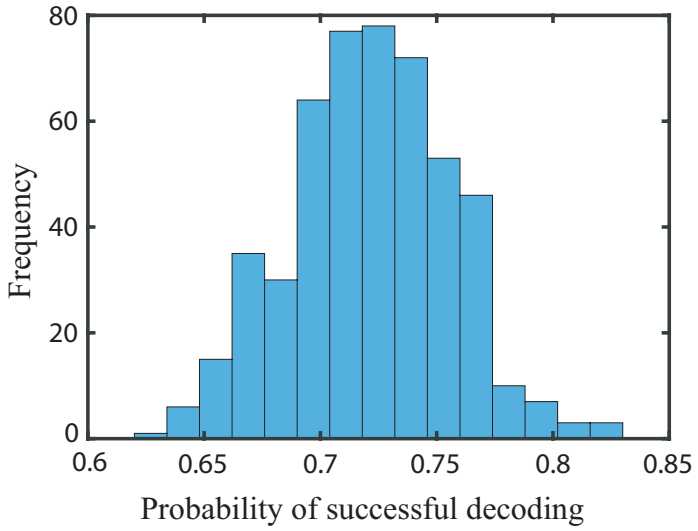
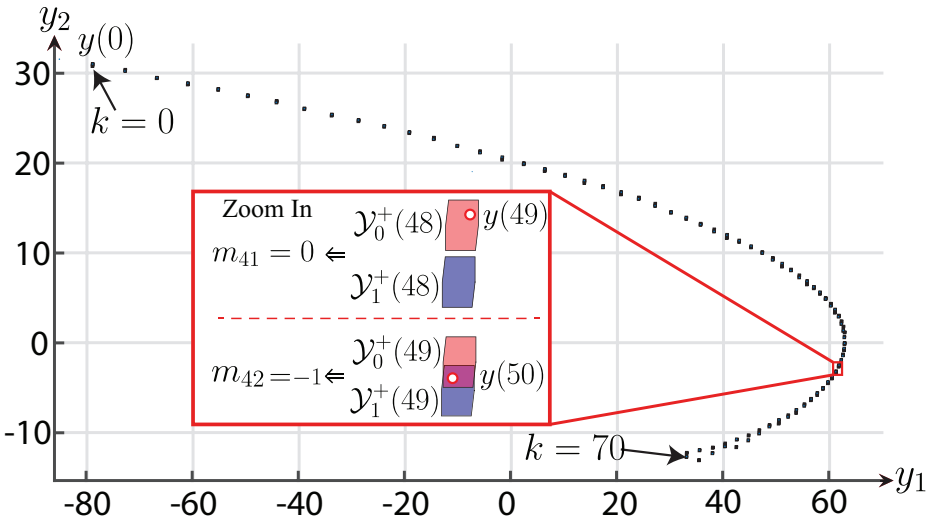


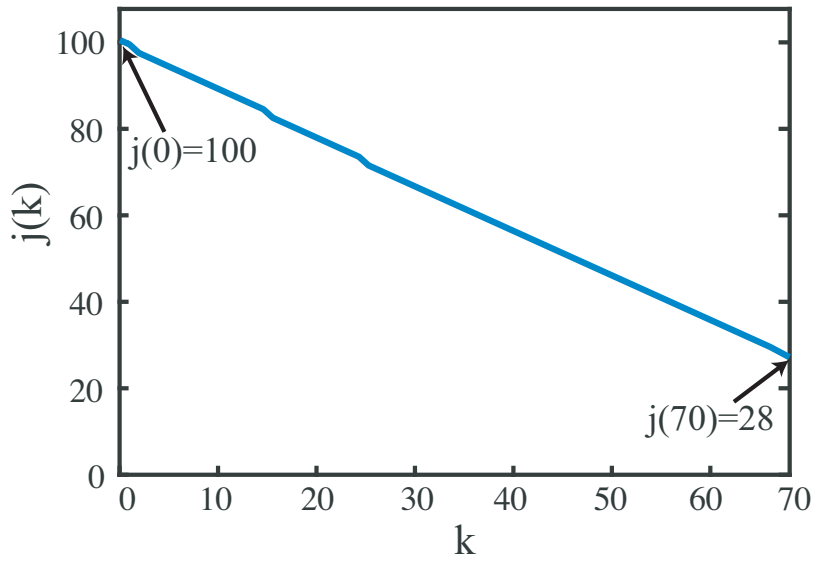
Figure 23: A histogram for the probability of successful decoding over 500 simulation runs.

times steps each, is about $0.725 \times 70 = 50.75$ bits.

In Fig. 24-a, we show the system output evolution $y(k)$ and the attacker robust output prediction sets $\mathcal{Y}_0^+$, $\mathcal{Y}_1^+$ for $0 \leq k \leq 70$, for a single arbitrarily chosen simulation run. In particular, from the provided zoom in, it is possible to appreciate the decoding operations of the designed covert channel. At $k = 49$, the output vector $y(49)$ belongs exclusively to the robust output prediction set associated to the control law $u_0(48)$, i.e., $y(49) \in (\mathcal{Y}_0^+(48) \setminus \mathcal{Y}_1^+(48))$. Therefore, according to the Step 2 of the *CC-NCS* (receiver) algorithm, a bit equal to zero is successfully decoded and stored into the message vector M_r, i.e., $m_{41} = 0$ by both the sender and receiver. On the other hand, at $k = 50$, the output vector $y(50)$ belongs to the intersection of the two output prediction sets, i.e., $y(50) \in (\mathcal{Y}_0^+(49) \cap \mathcal{Y}_1^+(49))$. As a consequence, according to the Step 7 of *CC-NCS* (receiver) algorithm, the transmitted bit is labeled as uncertain ($m_{42} = -1$), discarded by the receiver and then re-transmitted in the next time step by the sender. Finally, in Fig. 24-b, the index $j(k)$, representing the order of set $\mathcal{T}_{j(k)}$, to which the current state belongs to, is shown for $0 \leq k \leq 70$. In particular, it is possible to notice that the infected control law (81) preserves the expected monotonically decreasing behaviour of $j(k)$, see [30]. As a consequence, the designed covert channel



(a) An illustration of system evolution and prediction sets for a single simulation run, $0 \leq k \leq 70$. The red rectangles represent $\mathcal{Y}_0^+$, the blue rectangles represent $\mathcal{Y}_1^+$ and the white dots represent the actual measurement y.



(b) $j(k)$ for $0 \leq k \leq 70$.

Figure 24: System evolution and signal $j(k)$

does not alter the "*normal*" closed-loop control system operations (see Remark 3), i.e., finite-time convergence into the terminal region as well as ultimately uniformly bounded stability are preserved.

Chapter 5

Conclusion and Future Work

5.1 Conclusion

In this thesis, the security dimension of controlling CPSs was examined. In particular, we proposed several mitigation and detection techniques for cyber-physical attacks like DoS, replay attacks, and system identification attacks against CPSs. We also investigated a design for covert channel technique that relies on a control-theoretic approach.

In chapter 2, a resilient set-theoretic controller that can deal with the transient stability control problem and mitigate DoS attacks and/or packet drops in smart grid applications was proposed. Simulation results confirmed the effectiveness of the proposed approach and showed that the proposed controller achieves better transient stability recovery time compared to another recently proposed scheme. Moreover, we proved that our controller has bounded, offline computed, worst-case transient stability time. Then, a novel physical watermarking technique for the detection of replay attacks in CPS was presented. The proposed strategy exploits the ST-MPC paradigm to design control inputs that, whenever needed, can be safely and continuously applied to the system for an apriori known number of steps. This control scheme is coupled with a designed random watermarking signal in the form of packet drops. We proved that the proposed watermarked input signal does not

affect the desired transient stability and that the closed-loop system enjoys, in the worst-case scenario, uniformly ultimately bounded stability.

In chapter 3, we presented a decoy-based defense strategy against eavesdroppers whose objective is the identification of a power system model from intercepted control inputs and sensor measurements. We showed that the proposed solution has the capability of degrading the accuracy of the identified model without affecting the control system performance. The proposed decoy solution, by design, does not affect the control system performance; it allows to arbitrarily degrade the system identification process by increasing the number of decoy systems.

In chapter 4, we showed the possibility of establishing a covert channel in CPSs which is capable of transmitting sensitive information from an infected networked controller to an eavesdropper with access to the feedback channel. The presented approach is unique to the CPSs as it encodes the message the control systems dynamics without creating any additional explicit communication channels. We used a dual-mode ST-MPC controller as an example to show how the encoding/decoding scheme can be performed over such channels. The achievable rate of the constructed covert channel was evaluated using numerical simulation results.

5.2 Future Work

In what follows, we provide a few suggestions for future work in the security of CPS are:

- In section 2.2, a DoS-resilient controller is designed to mitigate the effect of attacks on the feedback channel for a finite time steps. A possible extension of this work is to investigate how such controller can also deal with DoS attacks on the actuation channel.

- Research related to covert channels in CPSs is still in its infancy, and different challenging questions remain to be explored in future works. In chapter 4, we presented a design for a covert channel between a compromised controller and an eavesdropper that has access to the feedback channel.
  - This setup assumes that the attacker is very resourceful. Future research can be directed towards establishing a covert channel with less resources available to the attacker.
  - From the attacker's perspective, particularly relevant questions are the design of the infected control logic to maximize the covert channel's transmission rate and the study of the relation between noise magnitude and transmission rate capacity.
  - On the other hand, from a defender point of view, the design of control architectures and detection algorithms capable of preventing/detecting covert channels is of interest.

Bibliography

[1] 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.

[2] C. Peng, H. Sun, M. Yang, and Y. Wang, "A survey on security communication and control for smart grids under malicious cyber attacks," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 49, no. 8, pp. 1554–1569, 2019.

[3] W. Lucia, B. Sinopoli, and G. Franze, "A set-theoretic approach for secure and resilient control of cyber-physical systems subject to false data injection attacks," in *Science of Security for Cyber-Physical Systems Workshop (SOSCYPS)*. IEEE, 2016, pp. 1–5.

[4] 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.

[5] Y. Mo and B. Sinopoli, "Secure control against replay attacks," in *Allerton conference on communication, control, and computing (Allerton)*. IEEE, 2009, pp. 911–918.

[6] O. Ozel, S. Weerakkody, and B. Sinopoli, "Physical watermarking for securing cyber physical systems via packet drop injections," in *IEEE International Conference on Smart Grid Communications (SmartGridComm)*. IEEE, 2017, pp. 271–276.

[7] B. Satchidanandan and P. R. Kumar, "Secure control of networked cyber-physical systems," in *Conference on Decision and Control (CDC)*. IEEE, 2016, pp. 283–289.

[8] Y. Liu, P. Ning, and M. K. Reiter, "False data injection attacks against state estimation in electric power grids," *ACM Trans. Inf. Syst. Secur.*, vol. 14, no. 1, Jun. 2011.

[9] W. Lucia, K. Gheitasi, and M. Ghaderi, "A command governor based approach for detection of setpoint attacks in constrained cyber-physical systems," *IEEE Conference on Decision and Control (CDC)*, pp. 4529–4534, 2018.

[10] M. Ghaderi, K. Gheitasi, and W. Lucia, "A novel control architecture for the detection of false data injection attacks in networked control systems," in *American Control Conference (ACC)*. IEEE, 2019, pp. 139–144.

[11] K. Gheitasi, M. Ghaderi, and W. Lucia, "A novel networked control scheme with safety guarantees for detection and mitigation of cyber-attacks," *European Control Conference (ECC)*, pp. 1449–1454, 2019.

[12] D. Liu, *Networked Control Systems*. Springer, 01 2008.

[13] A. Shostack, *Threat modeling: Designing for security*. John Wiley & Sons, 2014.

[14] N. Forti, G. Battistelli, L. Chisci, and B. Sinopoli, "A bayesian approach to joint attack detection and resilient state estimation," *IEEE Conference on Decision and Control (CDC)*, pp. 1192–1198, 2016.

[15] 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.

[16] Y. Mo and B. Sinopoli, "Secure control against replay attacks," in *2009 47th annual Allerton conference on communication, control, and computing (Allerton)*. IEEE, 2009, pp. 911–918.

[17] A. Teixeira, I. Shames, H. Sandberg, and K. H. Johansson, "Revealing stealthy attacks in control systems," *Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, pp. 1806–1813, 2012.

[18] R. S. Smith, "Covert misappropriation of networked control systems: Presenting a feedback structure," *IEEE Control Systems Magazine*, vol. 35, no. 1, pp. 82–92, 2015.

[19] A. O. de Sá, L. F. R. da Costa Carmo, and R. C. Machado, "Covert attacks in cyber-physical control systems," *IEEE Transactions on Industrial Informatics*, vol. 13, no. 4, pp. 1641–1651, 2017.

[20] S. Liu, X. P. Liu, and A. El Saddik, "Denial-of-service (dos) attacks on load frequency control in smart grids," in *IEEE PES Innovative Smart Grid Technologies Conference (ISGT)*, 2013, pp. 1–6.

[21] G. Park, H. Shim, C. Lee, Y. Eun, and K. H. Johansson, "When adversary encounters uncertain cyber-physical systems: Robust zero-dynamics attack with disclosure resources," in *IEEE Conference on Decision and Control (CDC)*, 2016, pp. 5085–5090.

[22] M. Bagherzadeh and W. Lucia, "A set-theoretic model predictive control approach for transient stability in smart grid," *IET Control Theory & Applications*, vol. 14, no. 5, pp. 700–707, 2020.

[23] W. Lucia, D. Famularo, and G. Franze, "A set-theoretic reconfiguration feedback control scheme against simultaneous stuck actuators," *IEEE Transactions on Automatic Control*, vol. 63, no. 8, pp. 2558–2565, 2017.

[24] S. V. Rakovic, E. C. Kerrigan, K. I. Kouramas, and D. Q. Mayne, "Invariant approximations of the minimal robust positively invariant set," *IEEE Transactions on Automatic Control*, vol. 50, no. 3, pp. 406–410, 2005.

[25] A. Abdelwahab, W. Lucia, and A. Youssef, "A dos-resilient set-theoretic controller for smart grid applications," in *2020 IEEE Power and Energy Society General Meeting (PESGM)*. IEEE, 2020.

[26] ——, "Set-theoretic control for active detection of replay attacks with applications to smart grid," in *2020 IEEE Conference on Control Technology and Applications (CCTA)*. IEEE, 2020.

[27] H. Ali and D. Dasgupta, "Effects of time delays in the electric power grid," in *Critical Infrastructure Protection VI*, J. Butts and S. Shenoi, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 139–154.

[28] M. E. Bento, "A hybrid procedure to design a wide-area damping controller robust to permanent failure of the communication channels and power system operation uncertainties," *International Journal of Electrical Power & Energy Systems*, vol. 110, pp. 118–135, 2019.

[29] K. Mahapatra, M. Ashour, N. R. Chaudhuri, and C. M. Lagoa, "Malicious corruption resilience in pmu data and wide-area damping control," *IEEE Transactions on Smart Grid*, 2019.

[30] D. Angeli, A. Casavola, G. Franzé, and E. Mosca, "An ellipsoidal off-line mpc scheme for uncertain polytopic discrete-time systems," *Automatica*, vol. 44, pp. 3113–3119, 2008.

[31] D. P. Bertsekas and I. B. Rhodes, "On the minimax reachability of target sets and target tubes," *Automatica*, vol. 7, no. 2, pp. 233–247, 1971.

[32] F. Blanchini and S. Miani, *Set-Theoretic Methods in Control*. Springer, 2007.

[33] W. Lucia, K. Gheitasi, and M. Bagherzadeh, "A low computationally demanding model predictive control strategy for robust transient stability in smart grid," in *IEEE Conference on Decision and Control (CDC)*. IEEE, 2018, pp. 6013–6018.

[34] F. Miao, M. Pajic, and G. J. Pappas, "Stochastic game approach for replay attack detection," in *IEEE conference on decision and control*. IEEE, 2013, pp. 1854–1859.

[35] R. Romagnoli, S. Weerakkody, and B. Sinopoli, "A model inversion based watermark for replay attack detection with output tracking," in *2019 American Control Conference (ACC)*, 2019, pp. 384–390.

[36] F. Borrelli, A. Bemporad, and M. Morari, *Predictive Control for Linear and Hybrid Systems*. Cambridge University Press, 2017.

[37] L. Lima, R. Ramos, I. Hiskens, C. Cañizares, T. C. Fernandes, E. Jr, L. Gein-Lajoie, M. Gibbard, J. Kersulis, R. Kuiava, F. De Marco, N. Martins, B. Pal, A. Piardi, J. Santos, D. Silva, A. K. Singh, B. Tamimi, and D. Vowles, "Benchmark systems for small-signal stability analysis and control," *PES-TR*, 08 2015.

[38] A. Farraj, E. Hammad, and D. Kundur, "A cyber-enabled stabilizing control scheme for resilient smart grid systems," *IEEE Transactions on Smart Grid*, vol. 7, pp. 1–1, 2015.

[39] F. Dorfler and F. Bullo, "Kron reduction of graphs with applications to electrical networks," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 60, no. 1, pp. 150–163, 2013.

[40] H.-D. Chiang, F. F. Wu, and P. P. Varaiya, "A bcu method for direct analysis of power system transient stability," *IEEE Transactions on Power Systems*, vol. 9, no. 3, pp. 1194–1208, 1994.

[41] A. Bose, "Smart transmission grid applications and their supporting infrastructure," *IEEE Transactions on Smart Grid*, vol. 1, no. 1, pp. 11–19, 2010.

[42] A. Isidori, *Nonlinear control systems*. Springer Science & Business Media, 2013.

[43] A. Farraj, E. Hammad, and D. Kundur, "A cyber-enabled stabilizing control scheme for resilient smart grid systems," *IEEE Transactions on Smart Grid*, vol. 7, no. 4, pp. 1856–1865, 2016.

[44] G. Franzè, F. Tedesco, and D. Famularo, "Model predictive control for constrained networked systems subject to data losses," *Automatica*, vol. 54, pp. 272 – 278, 2015.

[45] M. Herceg, M. Kvasnica, C. N. Jones, and M. Morari, "Multi-parametric toolbox 3.0," in *2013 European Control Conference (ECC)*, 2013, pp. 502–510.

[46] D. Simon, "Kalman filtering with state constraints: a survey of linear and nonlinear algorithms," *IET Control Theory Applications*, vol. 4, no. 8, pp. 1303–1318, 2010.

[47] R. K. Mehra and J. Peschon, "An innovations approach to fault detection and diagnosis in dynamic systems," *Automatica*, vol. 7, no. 5, pp. 637–640, 1971.

[48] Y. Mo, S. Weerakkody, and B. Sinopoli, "Physical authentication of control systems: Designing watermarked control inputs to detect counterfeit sensor outputs," *IEEE Control Systems Magazine*, vol. 35, no. 1, pp. 93–109, 2015.

[49] 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.

[50] C. Glenn, D. Sterbentz, and A. Wright, "Cyber threat and vulnerability analysis of the us electric sector," Idaho National Lab.(INL), Idaho Falls, US, Tech. Rep., 2016.

[51] R. S. Smith, "A decoupled feedback structure for covertly appropriating networked control systems," *IFAC Proceedings Volumes*, vol. 44, no. 1, pp. 90–95, 2011.

[52] L. Ljung, "System identification," *Wiley encyclopedia of electrical and electronics engineering*, pp. 1–19, 1999.

[53] S. Amin, X. Litrico, S. Sastry, and A. M. Bayen, "Cyber security of water scada systems—part i: Analysis and experimentation of stealthy deception attacks," *IEEE Transactions on Control Systems Technology*, vol. 21, no. 5, pp. 1963–1970, 2012.

[54] S. Weerakkody and B. Sinopoli, "Detecting integrity attacks on control systems using a moving target approach," in *IEEE Conference on Decision and Control (CDC)*. IEEE, 2015, pp. 5820–5826.

[55] C. Schellenberger and P. Zhang, "Detection of covert attacks on cyber-physical systems by extending the system dynamics with an auxiliary system," in *IEEE Conference on Decision and Control (CDC)*. IEEE, 2017, pp. 1374–1379.

[56] J. Katz and Y. Lindell, *Introduction to modern cryptography*. CRC press, 2014.

[57] H. Saadat *et al.*, *Power system analysis*. McGraw-Hill, 1999, vol. 2.

[58] V. M. Igure, S. A. Laughter, and R. D. Williams, "Security issues in scada networks," *computers & security*, vol. 25, no. 7, pp. 498–506, 2006.

[59] T. Bohlin, "Grey-box model calibrator and validator," *IFAC Proceedings Volumes*, vol. 36, no. 16, pp. 1477 – 1482, 2003, iFAC Symposium on System Identification (SYSID 2003), Rotterdam, The Netherlands, 27-29 August, 2003.

[60] L. Ljung, "System identification toolbox for use with matlab," The Math Works, Tech. Rep., 1995.

[61] Mathworks, https://www.mathworks.com/help/ident/ref/goodnessoffit.html.

[62] A. Abdelwahab, W. Lucia, and A. Youssef, "Covert channels in cyber-physical systems," *IEEE control systems letters*, 2020.

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

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

[65] S. H. Sellke, C. . Wang, S. Bagchi, and N. Shroff, "Tcp/ip timing channels: Theory to implementation," in *IEEE INFOCOM 2009*, 2009, pp. 2204–2212.

[66] X. Luo, E. W. W. Chan, and R. K. C. Chang, "Tcp covert timing channels: Design and detection," in *IEEE International Conference on Dependable Systems and Networks With FTCS and DCC (DSN)*, 2008, pp. 420–429.

[67] Y. Liu, D. Ghosal, F. Armknecht, A.-R. Sadeghi, S. Schulz, and S. Katzenbeisser, "Robust and undetectable steganographic timing channels for iid traffic," in *Int. Workshop on Information Hiding*. Springer, 2010, pp. 193–207.

[68] R. A. Kemmerer, "Shared resource matrix methodology: An approach to identifying storage and timing channels," *ACM Transactions on Computer Systems (TOCS)*, vol. 1, no. 3, pp. 256–277, 1983.

[69] 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.

[70] 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*, 2015, pp. 276–289.

[71] L. Deshotels, "Inaudible sound as a covert channel in mobile devices," in $USENIX$ *Workshop on Offensive Technologies (*$WOOT$*)*, 2014.

[72] S. Wendzel, W. Mazurczyk, and G. Haas, "Don't you touch my nuts: Information hiding in cyber physical systems," in *2017 IEEE Security and Privacy Workshops (SPW)*. IEEE, 2017, pp. 29–34.

[73] 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, p. 638–649.

[74] ——, "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, p. 87–98.

[75] X. Ying, G. Bernieri, M. Conti, and R. Poovendran, "Tacan: Transmitter authentication through covert channels in controller area networks," in *ACM IEEE International Conference on Cyber-Physical Systems*, ser. ICCPS '19, 2019, p. 23–34.

[76] 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.

[77] L. Garcia, H. Senyondo, S. McLaughlin, and S. Zonouz, "Covert channel communication through physical interdependencies in cyber-physical infrastructures," in *SmartGridComm*, 2014, pp. 952–957.

[78] K. Zhou and J. C. Doyle, *Essentials of robust control*. Prentice hall Upper Saddle River, NJ, 1998, vol. 104.

[79] A. Bemporad and M. Morari, "Robust model predictive control: A survey," in *Robustness in identification and control*. Springer, 1999, pp. 207–226.

[80] 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.

[81] L. J. Wells, J. A. Camelio, C. B. Williams, and J. White, "Cyber-physical security challenges in manufacturing systems," *Manufacturing Letters*, vol. 2, no. 2, pp. 74–77, 2014.

[82] M. S. Branicky, "Multiple lyapunov functions and other analysis tools for switched and hybrid systems," *IEEE Transactions on automatic control*, vol. 43, no. 4, pp. 475–482, 1998.

[83] H. Lin and P. J. Antsaklis, "Stability and stabilizability of switched linear systems: a survey of recent results," *IEEE Transactions on automatic control*, vol. 54, no. 2, pp. 308–322, 2009.

[84] D. Liberzon and A. S. Morse, "Basic problems in stability and design of switched systems," *IEEE control systems magazine*, vol. 19, no. 5, pp. 59–70, 1999.