UNIVERSITÉ DE REIMS CHAMPAGNE-ARDENNE
ÉCOLE DOCTORALE SCIENCES DU NUMÉRIQUE ET DE L'INGÉNIEUR n° 620

THÈSE

Pour obtenir le grade de
DOCTEUR DE L'UNIVERSITÉ DE REIMS CHAMPAGNE-ARDENNE
Spécialité: Automatique et Traitement du Signal
Présentée et soutenue publiquement par

ÁLAN CRÍSTOFFER E SOUSA

Enhancing Cyber-Physical System Security: Detection and Mitigation Strategies

---

Thése co-dirigée par **Nadhir Messai** et **Noureddine Manamanni**

JURY

Mme. Taous-Meriem Laleg Kirati	HDR	INRIA Paris-Saclay	Rapporteur
M. Valter Leite	Professeur	CEFET-MG	Rapporteur
M. Mohamed Djemai	Professeur	Université Polytechnique Hauts-de-France	Examinateur
M. Christophe Sueur	Professeur	Centrale Lille	Examinateur
M. Frédéric Hamelin	Professeur	Université de Lorraine	Examinateur
M. Noureddine Manamanni	Professeur	Université de Reims Champagne-Ardenne	Co-Directeur de Thèse
M. Nadhir Messai	Professeur	Université de Reims Champagne-Ardenne	Co-Directeur de Thèse

It’s not about what you have, it’s about what you make of what you have.

Hank Green

To Eduardo, Renan and Richard, who live in my heart rent-free.

Acknowledgements

to my parents for all the support and trust placed in me;

to Prof. Nadhir Messai and Noureddine Manamanni for all the dedication in guidance;

to Prof. Kevin Guelton for welcoming me into his team;

to my closest work colleagues: Adriano, Alex, Dylan, Fabien, Lucas, Noémi, Sara, Shurong, Soizic, and Tarek for the camaraderie and the exchange of experiences;

to the other colleagues from CReSTIC for the great coexistence;

to everyone who in some way contributed to my progress.

Abstract

Les systèmes cyber-physiques (SCP) se situent à l'intersection des opérations physiques, des technologies informatiques et des communications réseau. Ces systèmes sont fondamentaux pour l'infrastructure des environnements de fabrication intelligente, fournissant les moyens d'un contrôle amélioré, d'une optimisation et d'une adaptabilité au sein de divers processus de production. Cette thèse se penche sur les défis associés à la protection des SCP contre une large gamme de menaces cybernétiques qui posent des risques pour la sécurité de ces systèmes. À travers l'application de stratégies de contrôle de pointe telles que l'Observabilité Fonctionnelle et la Théorie des Échelles de Temps, ce travail introduit des méthodes innovantes visant la détection et l'atténuation de l'Injection de Données Fausses (IDF), des Attaques de Dynamique Nulle (ADN), et du Déni de Service (DoS). Ces menaces cybernétiques sont identifiées comme prévalentes et dommageables dans le contexte des SCP industriels. L'incorporation stratégique des *observateurs fonctionnels*, l'application du cadre analytique des Échelles de Temps, et l'utilisation du critère de stabilité de Lyapunov se démarquent comme contributions notables au domaine de la sécurité des SCP. Ces approches offrent de nouvelles perspectives et des solutions robustes aux défis multifacettes que rencontrent les systèmes cyber-physiques. De plus, l'efficacité des solutions proposées est testée et validée à travers une combinaison d'études expérimentales et de simulation. Ces efforts de validation soulignent la capacité des améliorations de sécurité proposées à élever les mesures protectrices au sein des environnements de fabrication intelligente. En conséquence, ce travail de thèse tente de repousser non seulement les limites des connaissances actuelles dans le domaine de la sécurité des SCP, mais d'établir également une base solide pour les recherches en cours et futures visant à renforcer l'infrastructure numérique essentielle au fonctionnement des industries intelligentes. À travers cette étude, la thèse aspire à contribuer de manière significative au développement de systèmes cyber-physiques plus sûrs, fiables et résilients, soutenant ainsi l'évolution continue des technologies de fabrication intelligente.

Mots-clés : Systèmes Cyber-Physiques, Cybersécurité, Injection de Données Fausses, Attaque de Dynamique Nulle, Déni de Service.

Abstract

Cyber-physical systems (CPS) stand at the intersection of physical operations, computing technologies, and network communications. These sophisticated systems are fundamental to the infrastructure of smart manufacturing environments, providing the means for improved control, optimization, and adaptability within various production processes. This thesis delves into addressing the significant challenges associated with safeguarding CPS from a wide range of cyber threats that pose risks to the operational integrity and overall safety of these systems. Through the application of cutting-edge control strategies like Functional Observability and Time-Scale Theory, this work introduces innovative methods aimed at the early detection and effective mitigation of False Data Injection (FDI), Zero-Dynamics Attacks (ZDA), and Denial of Service (DoS). These particular cyber threats are identified as especially prevalent and damaging within the context of industrial CPS. The strategic incorporation of *functional observers*, the application of the Time-Scale analytical framework, and the use of the Lyapunov stability criterion stand as notable contributions to the domain of CPS security. These approaches provide fresh insights and robust solutions to the multifaceted challenges that cyber-physical systems encounter. Moreover, the efficacy of the proposed solutions is tested and validated through a combination of simulation and hands-on experimental studies. These validation efforts underscore the capability of the proposed security enhancements to elevate the protective measures within smart manufacturing environments. Consequently, this thesis not only pushes the boundaries of current knowledge in the field of CPS security but also establishes a solid foundation for ongoing and future researches aimed at fortifying the digital infrastructure essential to the operation of smart industries. Through this study, the thesis aspires to contribute meaningfully to the development of more secure, reliable, and resilient cyber-physical systems, thereby supporting the continued evolution of smart manufacturing technologies.

Keywords: Cyber-Physical Systems, Cybersecurity, False Data Injection, Zero-Dynamics Attack, Denial of Service.

Contents

List of Figures

List of Algorithms

List of Abbreviations

CPS	Cyber-Physical Systems
DoS	Denial-of-Service
ECU	Electronic Control Unit
FDI	False Data Injection
IoT	Internet of Things
IT	Information Technology
LMI	Linear Matrix Inequality
LPV	Linear Parameter Varying
MitM	Man-in-the-Middle
MQTT	Message Queuing Telemetry Transport
OT	Operational Technology
PLC	Programmable Logic Controller
PMU	Phasor Measurement Unit
ROS	Robot Operating System
ZDA	Zero-Dynamics Attack

Chapter 1

Introduction

1.1 Cyber-Physical Systems

Cyber-Physical Systems (CPS) represent a groundbreaking approach in engineering, blending the physical and digital to create systems that are both interconnected, making use of different network kinds, and intelligent, being able to adapt to changes, to make use of different data sources and to understand its surroundings, if necessary. The strategic integration of sensors, actuators, computing units, and communication networks allows CPS to effectively monitor, analyze, and influence their environment.

This transformative paradigm not only enhances the capability to interact with the physical world but also opens new avenues for innovation, heralding a new era in the design and implementation of engineering systems. Through the lens of CPS, the future of engineering is envisioned as a realm where the barriers between the physical and digital are not just blurred but effectively dismantled, paving the way for unprecedented levels of interaction and control over the physical world (Foundation 2022).

CPS allow the real-time monitoring, analysis, and control of physical processes. This enables a wide range of applications where the dynamic and efficient control of physical processes is of importance. The real-time data exchange and processing enable CPS to anticipate changes, mitigate risks, and enhance operational efficiency, making it an interesting prospect for application in medical devices, traffic control, autonomous vehicles, avionics, defense and manufacturing, for example.

This development incurs a significant trade-off: the employment of more sophisticated computational devices and the deployment of broader, interconnected networks expand the vulnerability landscape of CPS. Conse-

quently, these systems become prone to attack vectors traditionally associated with computer systems, such as Denial-of-Service (DoS) attacks. Previously, such attacks were deemed impractical against conventional control systems owing to the absence of exploitable attack vectors.

Accordingly, the security of CPS remains an unresolved challenge, actively investigated by the scientific community. Advances are being achieved across various domains, attributable to the diverse array of potential attack methodologies. Because of the overlap with traditional IT defense needs, the informatics security side of the CPS developed more than its automation counterpart, an imbalance this thesis aims to help even.

1.2 Main Contributions

The main contributions of this thesis is the formulation of detection mechanisms for False Data Injection (FDI) and Zero-Dynamics Attack (ZDA), along with a mitigation mechanism for DoS attacks. The contributions are delineated as follows:

1. Development of a detection mechanism for FDI attacks that is robust against the mathematical challenges presented by ill-posed matrices, a common issue in power systems where the dynamics matrices are sparse and encompass hundreds of states.

2. Creation of a detection mechanism for zero-dynamics attacks employing the Time-Scale framework. This approach aims to derive a straightforward control law that is computationally efficient during online evaluation.

3. Design of a mitigation mechanism for DoS attacks by leveraging the Time-Scale framework. This mechanism intends to adjust the controller's sampling time in response to network package drop induced by the attack, thereby enhancing system resilience.

1.3 Document organization

This thesis is structured as follows:

Chapter 2 presents an overview of cyber-physical system vulnerabilities, focusing on attacks targeted towards these systems. It includes a comprehensive review of literature pertaining to the three categories of attacks

explored within this study. Furthermore, the objectives of this research are clearly defined in this segment.

Chapter 3 elaborates on the methodology for detecting False Data Injection (FDI) attacks, presenting an innovative approach that incorporates Functional Observers to address the challenge of ill-posed matrices identified in prior discussions. This method leverages both Functional Observers and the Lyapunov stability criterion to ensure the observer's convergence, thus circumventing the need for techniques that are susceptible to mathematical inaccuracies in the context of ill-posed matrices, such as zeros of the matrices and singular value decomposition.

Chapter 4 expounds on the detection strategy for Zero-Day Attacks (ZDA), introducing the fundamental concepts of Time-Scale theory. This theoretical framework enables the observer, designed for state estimation, to operate with a sampling interval that approximates continuous-time operation under normal circumstances, while facilitating the detection of ZDA through stochastic modifications to the sampling interval during its operation.

Chapter 5 discusses a mitigation approach for Denial of Service (DoS) attacks, utilizing time-scale controllers to bolster system robustness. The technique capitalizes on the adaptive sampling capability of time-scale controllers, ensuring system stability even under the strain of a weak DoS assault.

Finally, Chapter 6 provides conclusive remarks and outlines potential avenues for future investigations.

1.4 Publications

The following papers were published during the development of this work:

- Á. e Sousa, N. Messai, and N. Manamanni, "False Data Injection Detection in Cyber-Physical System," presented at the Symposium on Fault Detection, Supervision and Safety for Technical Processes, Cyprus: IFAC, 2022, p. 7. doi: 10.1016/j.ifacol.2022.07.165.

- Á. e Sousa, N. Messai, and N. Manamanni, "Load-altering attack detection on smart grid using functional observers," International Journal of Critical Infrastructure Protection, vol. 37, p. 100518, Jul. 2022, doi: 10.1016/j.ijcip.2022.100518.

- *Under review*: Á. e Sousa, N. Messai, and N. Manamanni, ”DoS attack mitigation using Time-Scale theory,” Control Engineering Practice.

Chapter 2

Overview

2.1 Contextualization

The American National Science Foundation defines Cyber-Physical Systems (CPS) as[1]:

> *"Cyber-physical Systems (CPS) are engineered systems that are built from, and depend upon, the seamless integration of computation and physical components. CPS tightly integrate computing devices, actuation and control, networking infrastructure, and sensing of the physical world. The system may include human interaction with or without human aided control. CPS may also include multiple integrated system components operating at wide varieties of spatial and temporal time scales. They can be characterized by architectures that may include distributed or centralized computing, multi-level hierarchical control and coordination of physical and organizational processes."*

Figure 2.1 provides a comprehensive depiction of the constituent elements within a CPS. This illustration delineates the actuators and sensors as components of the physical domain, whereas the remainder, inclusive of the interconnections, are categorized under the cyber domain of the CPS. The centralized cloud, depicted in the figure, may represent either the internet or a Large Area Network, serving as a pivotal hub for data and control flow. The database's primary function is the storage of system data, facilitating subsequent processing or logging activities. The server plays

[1] https://new.nsf.gov/funding/opportunities/cyber-physical-systems-cps/nsf21-551/solicitation

a crucial role in the orchestration of system operations, encompassing the management of set-points, user authentication, and overseeing other administrative and high-level control requisites. PLC are dedicated to low-level control tasks and actuator/sensor interfacing, potentially managing local control loops in adherence to the server-directed set-points or directly connecting with non-networkable equipment. Human-Machine Interfaces (HMIs) are versatile in nature, ranging from specialized hardware, such as control panels in nuclear facilities, to conventional computers employed by users for parameter adjustment, system command issuance, and the monitoring of system variables.

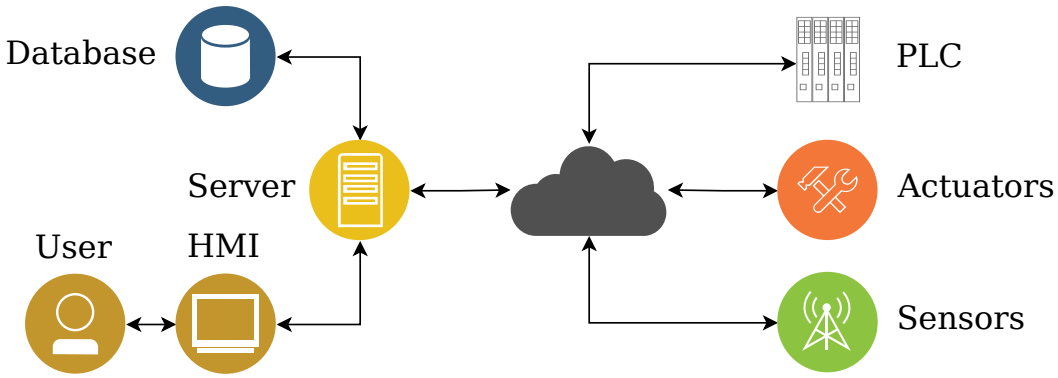


Figure 2.1 – Overview of a Cyber-Physical System

The architecture of CPS is predicated on the harmonious amalgamation of its constituent elements, spanning both tangible hardware and intangible software components. At the heart of the physical dimension lie sensors and actuators embedded within the environment. These components play a critical role in perceiving environmental conditions and implementing physical interventions based on computational directives. Such interactions underscore the tangible aspect of CPS, where physical actions are initiated and adjusted in real-time in accordance with the insights gained from data analysis (Krishnamurthy and Cecil 2018; Yajun Lu and J. Cecil 2015; Cecil 2017).

On the other hand, the cyber aspect of CPS encompasses the digital backbone of these systems, including computing units, software algorithms, communication networks, and sophisticated control mechanisms. This segment is pivotal in processing the data collected by physical sensors, making informed decisions, and orchestrating the actions of actuators (Bouheroum et al. 2022).

The seamless communication between the cyber and physical components is the cornerstone of CPS, enabling a continuous loop of data ex-

change, analysis, and action implementation. This interplay ensures that CPS can adapt to and influence their physical surroundings in real-time, optimizing performance and responding to emergent conditions.

The symbiotic relationship between the cyber and physical components of CPS fosters an ecosystem where real-time monitoring, analysis, and control of physical processes are achievable. This capability is vital in a wide array of applications, ranging from autonomous vehicle systems to smart energy grids, where the dynamic and efficient control of physical processes is paramount. The real-time data exchange and processing enable CPS to anticipate changes, mitigate risks, and enhance operational efficiency, embodying a paradigm shift in how systems interact with the physical world (E. A. Lee and Seshia 2017; L. D. Xu, E. L. Xu, and Li 2018).

CPS are extensively utilized across a diverse array of industries, underscoring their significant impact and broad applicability. The following list provides examples of their deployment in various sectors, illustrating the versatility and transformative potential of CPS technologies.

- Smart grids epitomize the vanguard of power grid evolution, leveraging CPS to revolutionize the conventional electricity infrastructure (IEEE 2013). This integration introduces an array of advantages and sophisticated features, fundamentally reshaping the operational landscape of energy systems. Of considerable significance is the empowerment of precise emission control facilitated by smart grid technologies, affording meticulous monitoring and intervention strategies to curtail emissions, thereby fostering environmental sustainability. Furthermore, this technological advancement empowers domestic consumers with enhanced control and visibility over their household energy consumption, yielding benefits that extend beyond individual households to encompass environmental conservation and economic efficiency on a broader scale (Bestehorn and Borsche 2014; Devi and Susmitha 2017).

- In the realm of healthcare, a notable trend involves the convergence of cyber and physical functionalities within medical devices, aimed at enhancing the landscape of healthcare services. A prevalent attribute among these devices is the integration of wireless capabilities, facilitating seamless updates and reconfigurations as required for optimal performance. Notably, wearable devices possess the potential to establish inter-device communication, not only among themselves but

also with an array of other devices, including smartphones or remote medical practitioners, leveraging the pervasive Bluetooth technology as a conduit for these interactions (Rushanan et al. 2014).

- Smart cars represent a new paradigm in the automotive industry, characterized by their eco-friendliness, heightened fuel efficiency, enhanced safety features, and an array of entertainment and convenience offerings, thereby elevating the driving experience. Central to enabling these advancements is a sophisticated network comprising around 70 computers known as Electronic Control Units (ECUs). These ECUs shoulder the responsibility of monitoring and governing a diverse spectrum of functions integral to smart car operations, encompassing tasks related to engine emission management, precision brake control mechanisms, as well as the orchestration of entertainment amenities such as radios and multimedia players, alongside the regulation of comfort features including window operation and cruise control functionalities (Humayed and Luo 2015).

Owing to the extensive deployment of CPS and their critical function in contemporary industrial applications, these systems have increasingly become the focus of malicious actors. Consequently, the capacity to identify and respond to these adversarial activities is of paramount importance. The subsequent section will delineate the current advancements in the domain of attack detection methodologies for CPS.

2.2 CPS security

Detecting attacks on CPS holds immense importance due to the critical role these systems play in our daily lives, infrastructure, and societal well-being. The interconnected nature of CPS, combining computational elements, physical machinery, and communication networks, makes them vulnerable to malicious cyber attacks. Understanding the significance of detecting such attacks is crucial in safeguarding against potentially devastating consequences (Giraldo et al. 2018; Zhenhua Wang, W. Xie, et al. 2021; D. Ding et al. 2018).

Firstly, CPS pervade various sectors, including energy, transportation, healthcare, and manufacturing. Attacks on these systems could lead to widespread disruptions, causing significant economic losses and impairing

essential services[2] For instance, a cyber attack on a smart grid could result in power outages affecting large populations, impacting businesses, hospitals, and everyday life[3]. Detecting attacks in CPS is crucial to prevent such disruptions, ensuring the continuity of vital services and minimizing the socio-economic impact[4].

Secondly, the integration of CPS into critical infrastructure exposes them to targeted cyber threats. Malicious actors often aim to exploit vulnerabilities within these systems to cause physical harm, financial loss, or manipulate data for their advantage. Detecting attacks on CPS becomes paramount to identify and thwart these threats before they escalate, mitigating potential damage to infrastructure, assets, and public safety (Miller et al. 2021).

Thirdly, the interconnectedness of CPS with the Internet of Things (IoT) introduces new attack surfaces. Vulnerabilities in one component of the system can cascade and affect the entire network, creating a domino effect of disruptions. Detecting attacks promptly helps in isolating affected areas, preventing the spread of attacks across the network, and reducing the overall impact on the system (Kimani, Oduol, and Langat 2019).

Finally, the potential for physical harm resulting from attacks on CPS cannot be overlooked. In sectors like healthcare and autonomous vehicles, a cyber attack causing manipulation or disruption in operations might directly jeopardize human lives (Djenna and Saidouni 2018; Haghighi et al. 2023). Timely detection of attacks is essential in ensuring the safety and well-being of individuals relying on these systems for critical services and support (Ozarar, Akansu, and Hasbay 2021).

Therefore, detecting attacks on cyber-physical systems is imperative to preserve the functionality, security, and resilience of these interconnected networks. Prioritizing robust detection mechanisms aids in thwarting malicious activities, safeguarding critical infrastructure, maintaining service continuity, and upholding public safety in an increasingly connected and technology-dependent world.

2.2.1 Exploits on CPS

Attacks against CPS encompass various strategies targeting the interconnected nature of these systems, aiming to disrupt, manipulate, or gain

[2] https://www.nytimes.com/2019/06/15/us/politics/trump-cyber-russia-grid.html

[3] https://www.bbc.com/news/technology-38573074

[4] https://www.independent.co.uk/news/long_reads/cyber-warfare-saudi-arabia-petrochemica
html

unauthorized access to critical infrastructure. The execution of such attacks may appear challenging; however, a variety of vulnerabilities can be exploited, enabling the attacker to conduct these assaults. Figure 2.2 reintroduces Figure 2.1, but with highlighting some vulnerabilities and attack locations, some of which are further explained in the sequence.

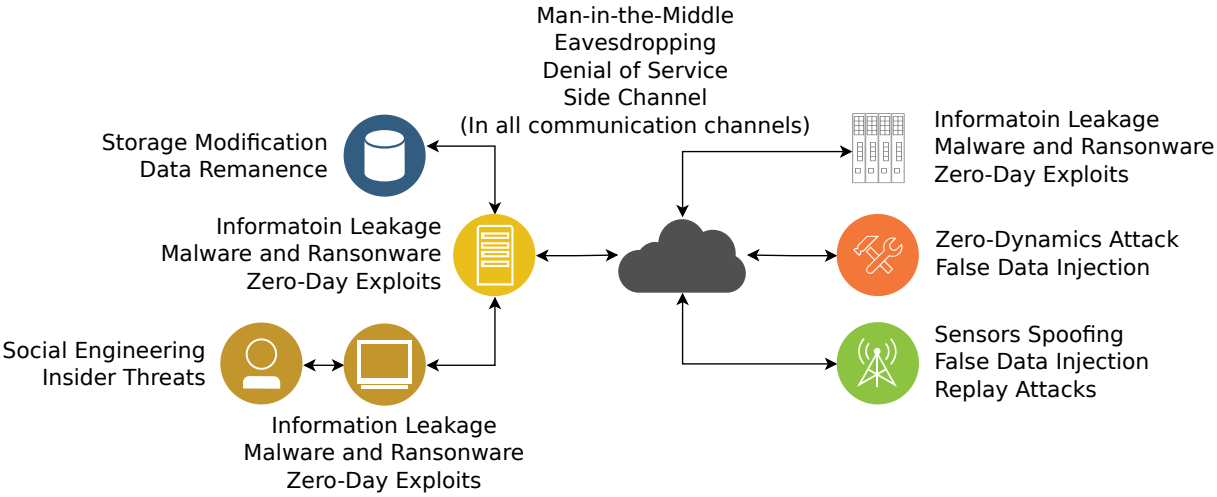


Figure 2.2 – Vulnerabilities and Attacks on CPS

- Man-in-the-Middle (MitM): Intercept communication between CPS components to eavesdrop, manipulate, or inject malicious content, compromising data integrity and confidentiality (Oliva, Cioaba, and Hadjicostis 2018; Jena, Padhy, and Guerrero 2023).

  A notable example of a MitM attack against a CPS is the Stuxnet worm, discovered in 2010[5]. Stuxnet specifically targeted Siemens Step7 software, which is used to program industrial control systems (ICS) that operate in various infrastructures, including nuclear facilities. Stuxnet's main goal was to sabotage Iran's nuclear program. It did so by first spreading via USB flash drives and then network shares, ultimately seeking out computers running the Siemens Step7 software. Once a suitable target was infected, Stuxnet would intercept the communication between the software and the PLC, which are crucial components in industrial control systems.

  The worm was designed to modify the code on the PLC to induce the centrifuges to spin at damaging speeds while simultaneously reporting to the monitoring systems that everything was operating as normal.

[5] `https://www.avast.com/c-stuxnet`

This deceptive reporting ensured that the system operators were unaware of the sabotage taking place, effectively placing Stuxnet ”in the middle” of the communication between the software and the physical machinery.

- Malware and Ransomware: Infiltrate the CPS with malicious software or ransomware, disrupting operations, stealing data, or encrypting critical systems, leading to data loss or financial extortion (Ibarra et al. 2019; Anand and Shanker 2023).

  A well-documented instance of ransomware impacting a cyber-physical system is the attack on the Colonial Pipeline by the DarkSide ransomware group in May 2021. The Colonial Pipeline is a crucial piece of infrastructure in the United States that transports gasoline, diesel, and jet fuel over a 5,500-mile (about 8,850 kilometers) pipeline system spanning from the Gulf Coast to the New York Harbor area. In May 2021, the DarkSide ransomware group launched a cyberattack that successfully infiltrated the Colonial Pipeline’s information technology (IT) systems.

  In response to the breach, Colonial Pipeline Company proactively shut down its operations to prevent the ransomware from spreading to the operational technology (OT) systems that control the pipeline, marking the first time a ransomware attack prompted the shutdown of a United States critical infrastructure facility. The shutdown led to significant disruptions, including fuel shortages across the Southeastern United States, increased gas prices, and widespread concern over the security of critical infrastructure against cyber threats. The attackers demanded and reportedly received a ransom payment of nearly 5 million dollars in cryptocurrency to provide a decryption tool and not to release stolen data.

- Zero-Day Exploits and Vulnerabilities: Exploit previously unknown vulnerabilities in CPS software or hardware, allowing unauthorized access or manipulation before patches or updates are available, posing security risks (Halabi and Zulkernine 2023; Gorbenko and Popov 2022).

  A prominent example of a zero-day exploit affecting a cyber-physical system is the case of the Triton malware (also known as Trisis), discov-

ered in 2017[6]. This malware targeted industrial control systems (ICS), specifically Safety Instrumented Systems (SIS), which are designed to monitor and ensure the safe operation of physical processes in industrial environments. The Triton malware exploited vulnerabilities in the Triconex SIS controllers manufactured by Schneider Electric, which are widely used in energy facilities, including nuclear plants.

The malware was first identified at a petrochemical plant in the Middle East, where it caused the plant's emergency shutdown systems to trigger unexpectedly. Further analysis revealed that Triton was designed to allow attackers to take control of the SIS controllers, modify their programming, and interfere with their operation. The ultimate goal of the malware appeared to be the ability to cause physical destruction to the facility or other catastrophic outcomes by disabling critical safety mechanisms.

- Tampering and Injection: Physically tamper with CPS components or inject faults into sensors or communication channels, disrupting normal system functionality, or causing hardware malfunctions (Jovanov and Pajic 2019; Garagad, Iyer, and Wali 2020).

  A noteworthy example of tampering and injection impacting cyber-physical systems involves the Maroochy Water Services incident in Queensland, Australia, in 2000[7]. This incident is one of the earliest known cyber-physical attacks that resulted in tangible environmental harm and public safety risks.

  A former employee of the company that provided the SCADA (Supervisory Control and Data Acquisition) system used by Maroochy Water Services executed the attack. They used a radio transmitter to gain unauthorized access to the SCADA system, which controlled sewage treatment and waste management processes. Over several weeks, they tampered with the system by injecting malicious commands that caused over 800,000 liters of raw sewage to spill into local parks, rivers, and even the grounds of a hotel. The sewage spill resulted in significant environmental damage, killed marine life, and posed serious health risks to the local population.

[6] https://cert.be/en/paper/trisis-malware
[7] https://www.industrialcybersecuritypulse.com/facilities/throwback-attack-an-insider-releases-265000-gallons-of-sewage-on-the-maroochy-shire/

This case of tampering and injection was particularly striking because it demonstrated how vulnerabilities in industrial control systems could be exploited to cause physical damage and public health hazards, highlighting the importance of securing such systems against unauthorized access and manipulation.

- Insider Threats: Pose threats from within an organization or system, where employees or trusted entities misuse access privileges, inadvertently or deliberately compromising CPS security (Ikany and Jazri 2019; Al Hammadi, Yeob Yeun, and Damiani 2020).

  The Maroochy Water Services case is also a great case of insider threat.

2.2.2 Attacks on CPS

Some prominent types of attacks are illustrated in Figure 2.3 and include:

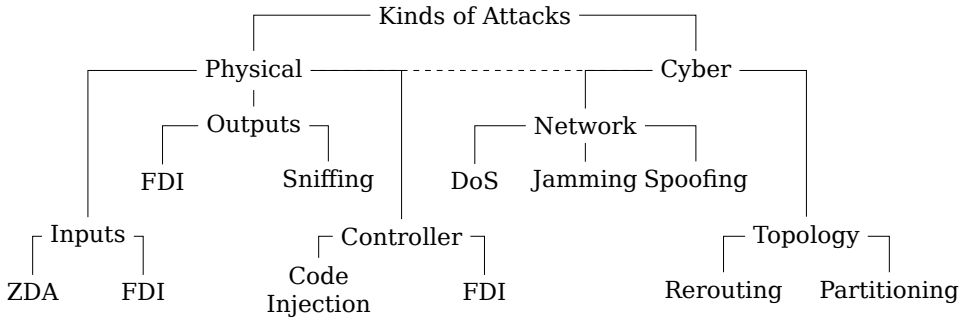


Figure 2.3 – Types of attacks on a CPS.

- Denial-of-Service (DoS): Overwhelm a CPS or its components with an influx of traffic, rendering the system inaccessible to legitimate users by exhausting system resources (G. He et al. 2021; P. Ding et al. 2017).

  A well-documented example of a Denial of Service (DoS) attack against a cyber-physical system is the attack on the Ukrainian power grid in December 2015[8]. This incident is notable as one of the first publicly acknowledged cyberattacks that successfully disrupted the operations of a power grid, resulting in significant real-world consequences.

  The attackers launched a sophisticated cyber operation against three regional electricity distribution companies in Ukraine. The primary goal was to cut off electricity to customers, and this was achieved by

[8] https://www.bbc.com/news/technology-38573074

remotely accessing the companies' Industrial Control Systems (ICS) and switching off electrical substations. As a part of the attack, the attackers also deployed malware that erased data on the companies' computers, further complicating recovery efforts.

In addition to directly manipulating the control systems to cut power, the attackers used a DoS attack against the phone systems of the affected companies. This was intended to prevent customers from reporting outages or receiving information, exacerbating the impact of the power disruption. The DoS attack on the customer service lines was a strategic move to increase the effect of the physical disruption by limiting communication and slowing down the response and recovery processes.

The Ukrainian power grid attack exemplifies how DoS tactics can be employed not just to disrupt digital services but also to enhance the impact of physical attacks on critical infrastructure. The incident underscored the vulnerability of critical infrastructure to cyberattacks and highlighted the need for enhanced cybersecurity measures, including robust incident response plans and improved resilience of industrial control systems against such threats.

- Spoofing Attacks: Impersonate a legitimate entity or device within the CPS to gain unauthorized access or privileges by falsifying IP addresses, MAC addresses, or digital signatures (Alsulami and Zein-Sabatto 2021; Kapoor, Vora, and Kang 2018).

  A notable example of a spoofing attack against a cyber-physical system is the incident involving the Iranian drone capture in 2011[9]. In this case, the U.S. claimed that Iran captured an RQ-170 Sentinel, a stealth surveillance drone, by spoofing its control signals. The Iranians reportedly intercepted the communication link of the drone, tricking it into landing in Iran under the belief that it was returning to its base in Afghanistan.

  The technique allegedly involved jamming the communication links to disrupt the control signals between the drone and its operators, then using a rogue transmission to spoof the drone's control channels, making it possible to redirect the drone to land in Iran. The exact technical details of the spoofing technique were not publicly disclosed, but the

[9] https://www.bbc.com/news/world-middle-east-21373353

incident highlighted the vulnerabilities in secure communication links for unmanned and remotely operated systems.

This event underscored the importance of securing communication channels and implementing robust authentication protocols to prevent unauthorized access and control of critical systems. It also highlighted the evolving landscape of cyber threats, where adversaries can exploit vulnerabilities in sophisticated technology to gain strategic advantages.

- False Data Injection: Inject false or manipulated data into the CPS sensors or control systems, causing erroneous decisions or actions based on corrupted information, impacting system reliability or safety (G. He et al. 2021; Garagad, Iyer, and Wali 2020).

  The biggest target of False Data Injection (FDI) attacks are critical infrastructures, like energy, finances and security, which do not disclose attacks, making it hard to find examples of such attacks.

- Side Channel Attack: Gain information from the physical implementation of a system, such as power consumption, electromagnetic leaks, or sound, to extract sensitive data (Gupta et al. 2019).

  A well-documented instance of a side channel attack against a cyber-physical system is the attack on the encryption system of the KeeLoq remote keyless entry system, widely used in car remote key fobs[10]. KeeLoq is a proprietary hardware-dedicated block cipher that implements a form of symmetric key encryption. It was used by various automobile manufacturers for remote keyless entry systems, allowing car owners to unlock their vehicles remotely. The security of KeeLoq was compromised through a side channel attack known as a ”power analysis attack.”

  In this attack, the attackers were able to deduce the secret encryption keys used by the KeeLoq system by observing the power consumption of the device during the encryption process. By analyzing the variations in power usage, the attackers could infer key bits and ultimately reconstruct the secret key used to encrypt and decrypt the signals sent between the key fob and the vehicle. With this information, an attacker could gain unauthorized access to a vehicle, locking or unlocking it at will.

[10] https://www.science20.com/news_releases/keeloq_remote_keyless_entry_system_for_cars_and_buildings_is_hacked

- Replay Attacks: Capturing data transmitted between devices and retransmitting it to create an unauthorized effect (Tang, Z. Zhang, and L. Xie 2023).

  A notable example of a replay attack against a cyber-physical system occurred in the context of automotive security systems, specifically keyless entry and ignition systems in cars[11]. In such an attack, malicious actors intercept and record the signal sent by a key fob to the vehicle to unlock the doors or start the engine. This recorded signal is then replayed at a later time to gain unauthorized access to the vehicle.

  Actual, publicly disclosed instances of replay attacks on cyber-physical systems, beyond research demonstrations or theoretical vulnerabilities, tend to be rare or less frequently reported in detailed public records. This scarcity is partly due to the sensitive nature of such incidents, where organizations might not disclose specifics to avoid revealing vulnerabilities or due to potential legal and reputational repercussions. However, the automotive industry has seen incidents that strongly suggest the use of replay or similar types of attacks, though these incidents often come to light through indirect sources such as insurance reports, law enforcement agencies, or media investigations.

- Jamming: Disrupting or blocking communication channels through the emission of interfering signals (Bout, Loscri, and Gallais 2020).

  A real-world example of a jamming attack against a cyber-physical system occurred with the use of GPS jamming devices to disrupt the tracking systems of commercial vehicles[12]. This type of attack has been observed in scenarios involving trucking and logistics companies, where thieves use GPS jammers to prevent vehicles from being tracked, facilitating the theft of the vehicle itself or the goods it is transporting.

- Zero-Dynamics Manipulation: Exploit the system's physical dynamics by manipulating actuators or control signals to induce unexpected behaviors or bypass control mechanisms, compromising system stability or safety (Zhenhua Wang, W. Xie, et al. 2021; Z. Yu et al. 2021).

  This kind of attack also sees no report in news outlets, making it hard to provide real-world examples.

[11] https://medium.com/codex/rollback-a-new-time-agnostic-replay-attack-against-the-autom
[12] https://www.theguardian.com/technology/2010/feb/22/car-thieves-using-gps-jammers

In this section we explored the different methods of exploiting and attacking a CPS system, showing how attacks are actually feasible and can have a significant impact not only on system, but also on the lives of people. In the next section we explore the existing defense techniques.

2.3 Defense against Cyber Attacks

The defense against cyber-attacks operates on dual fronts: Information Technology (IT) strategies encompass security measures like firewalls, encryption, and access control to thwart unauthorized network access. However, solely relying on these measures proves insufficient, given the potential for hackers to exploit human vulnerabilities within the system.

Therefore, an additional layer of defense emerges from an automation perspective, employing attack-resilient controllers, observers, and detection schemes to both identify and recover from cyber-attacks. Thus, while IT forms the initial protective layer, automation serves as a secondary line of defense (Mrabet et al. 2018).

CPS encompass a multitude of components, necessitating diverse defensive strategies to mitigate attacks. The scope of this dissertation is concentrated on assaults facilitated by the cybernetic aspect of the CPS yet have direct implications on its physical dimension, excluding considerations predominantly aligned with the (IT) facet of the CPS. Referring back to the schematic illustrated in Figure 2.1, the emphasis is placed on systems wherein the interaction between inputs (actuators) and outputs (sensors) can be delineated through differential equations. These interactions are subsequently modeled utilizing the state-space representation framework.

Considering the heterogeneity of attack strategies and their consequences, it is impractical to devise a universal defense mechanism capable of identifying all conceivable attacks. Consequently, it becomes imperative to concentrate on a specific category or class of attacks, to facilitate the development of effective attack detection schemes.

Therefore, emphasis has been placed on three attacks for focused investigation: FDI, Zero-Dynamics Attack (ZDA) and DoS. The following subsections present a literature review of attack detection and mitigation techniques for those attacks.

2.3.1 False Data Injection

FDI attacks exhibit a significant prevalence in power systems, particularly within smart grids, rendering these systems a primary focus for researchers engaged in the development of FDI detection mechanisms. Power systems refer to the networks of electrical components deployed to generate, transmit, and distribute electricity from power plants to consumers. These systems encompass generators, transmission lines, and distribution networks. Smart grids, an evolution of traditional power systems, integrate digital communications technology to monitor, predict, and intelligently control the flow of electricity, enhancing efficiency, reliability, and sustainability. They allow for two-way communication between utility providers and consumers, supporting real-time energy management and the integration of renewable energy sources.

FDI attacks pose a significant threat to power systems by compromising the integrity of information flow through tactics like eavesdropping and replay (Makar et al. 1975; Yohanandhan et al. 2020; J. Zhang et al. 2018). These attacks target the computation of state information within the system, particularly when the attacker possesses knowledge of the system's physical topology. This awareness grants the attacker the ability to evade conventional protective mechanisms such as static estimators and the bad data detector (Lakshminarayana et al. 2021). The success of FDI attacks can result in severe consequences, ranging from system interference and power overloads to widespread outages and potential risks to both physical infrastructure and human safety (Ahmed and Pathan 2020; J. Zhang et al. 2018).

The research by Zhao, Mili, and M. Wang (2018) introduced a concept termed as a flawless FDI attack, which posed challenges in detection while allowing the attacker control over altering state information. When an intruder manipulates power injection at a specific node, changes in power injections at the adjacent nodes become necessary for executing predefined state modifications. Assuming comprehensive knowledge of the system's topology, Zhao, Mili, and M. Wang (2018) demonstrated that the magnitude of the attack influenced the power system's topology and the construction of the attack vector, delineating constraints through Binary Decision Diagrams (BDD). Employing an autoregression model to predict Phasor Measurement Unit (PMU) measurements while ensuring measurement statistical consistency, they designed a model wherein the Huber-estimator minimized the summation of standardized residuals. Detection of an FDI attack

hinged upon the asymptotic approach of residuals between predicted and actual measurements towards zero.

Abusorrah et al. (2019) proposed a multi-tiered game-theoretic framework aimed at formulating a cost-effective defensive strategy to minimize operational costs amidst FDI instances. This strategy involved optimizing operational expenses under maximum regret in a multi-level optimization problem, accounting for interactions between defenders, attackers, and operators. The framework was reformulated into an equivalent bi-level mixed-integer linear programming problem. A technique of greedy implicit enumeration facilitated the discovery of a globally optimal secure set for this complex bi-level problem with fewer iterations. The conversion of detection into a more manageable problem was achieved by solving the Security Constrained Optimal Power Flow (SCOFF).

In the context of Cyber-Physical Power Systems (CPPS) comprising n nodes, the priority rested on securing a basic measurement set (BMS) of $(n-1)$ meters against FDI (Sreeram and Krishna 2019). A method employing a greedy optimal BMS was devised to ascertain an acceptable BMS. Given that executing FDI necessitated knowledge of the measurement matrix, an optimal subset of this matrix was chosen to augment the attack's cost. Sreeram and Krishna (2019) utilized a basic attack cost model known as multiple sparse attack vectors (MSAV). However, the enumeration method's limitations surfaced as Sreeram and Krishna (2019) failed to conclude validation in the IEEE 300 nodes model due to arithmetic constraints. In practical scenarios, more efficient search methodologies or compatible algorithms are poised to excel in handling increasingly intricate CPPS.

The construction of the FDI attack vector heavily relies on the CPPS topology. Real-time optimization of the topology enhances detection rates via Binary Decision Diagrams (BDD). System reconfiguration and power flow rerouting are chiefly implemented through load shedding, line switching, and node splitting. Zheng et al. (2021) devised a Deep Reinforcement Learning (DRL) model to resolve topology optimization issues and preemptively identify vulnerabilities to safeguard the grid's security. Employing a criticality-based adversarial perturbation model, misleading state vectors were sought when the DRL model's parameters were known, prompting erroneous power configurations and divergence of power flow due to FDI. Adversarial perturbations aimed at minimizing the DRL reward within minimal error magnitudes were constructed. Additionally, vulnerability indexes

for DRL models were proposed based on probability and gradient criteria.

Cheng and Chow (2020) addressed FDI attacks within the Energy Management System (EMS) for economic dispatch. The evaluation highlighted risks and assets pertinent to a consensus-based EMS. Various attack models were conceived from a risk-driven standpoint, introducing a majority voting mechanism for a unified reputation-based attack detection algorithm. Agents compared local estimates with data provided by neighboring agents and excessive errors affected an agent's reputation. To update global consensus, agents leaned towards information from highly reputable agents. When an agent lost trust due to reputation-based neighborhood monitoring, others substituted the agent's updated state data from FDI with local estimations. Diverse agents, including wind and solar power, along with different loads, were integrated into a comprehensive CPPS framework.

Saad et al. (2020) constructed a digital twin model for CPPS and assessed its resilience. A cloud-based digital twin platform formed the control center, selecting cluster leader and follower nodes in the secondary control based on grid topology. The Point of Common Coupling (PCC) node acted as the bridge between the grid's physical and informational layers in tertiary control. Integrity verification of every transmitted information to the PCC ensured maximum security. Utilizing residuals from local state estimation and transmitted values aided in identifying the source of FDI. The Kullback-Leibler divergence (a statistical method to measure the distance between two probability distributions) helped ascertain deviations among the follower's neighbors from the consensus, updating expectations for leader nodes.

M. Shi et al. (2021) proposed an observer-based resilient integrated distributed control strategy to mitigate FDI attacks on sensors and actuators within island-based AC Microgrids (MGs). Adhering to consensus theory, integrated distributed frequency control focused on transmitting normalized active power signals while ensuring global frequency stability. This approach significantly reduced communication burdens while preserving local values of frequency and voltage. By relying solely on local state variables to estimate and counteract potential attacks on each generator, the strategy ensured cyber-security from a sensor's standpoint.

However, traditional methodologies, notably those reliant on Kalman filters, encounter substantial challenges. These challenges are primarily attributed to the extensive state space associated with smart grids, often encompassing hundreds of states. This complexity can precipitate numeri-

cal instabilities and computational inefficiencies in conventional detection mechanisms (Chen 2013). Consequently, there is an escalating demand for detection strategies that are not only more effective but also scalable. This research is directed towards addressing this issue through the employment of LMI and Functional Observers, as delineated in Chapter 3.

2.3.2 Zero-Dynamics Attack

In the realm of control systems, the zero-dynamics attack poses a significant threat due to its inherent difficulty in detection. It exploits the system's zero dynamics to inject a signal which affects the system's states but not its outputs. Typically, detection occurs at the controller side, utilizing input data from the controller and output data from the plant. However, these data can be compromised across communication networks. The attacker strategically manipulates these data to ensure consistency with the plant's dynamics, complicating the detection process (Andre Teixeira et al. 2012).

The study in (Jihan Kim, Back, et al. 2020) employs a Generalized Hold (Yuz and Goodwin 2014) to alter the system's zeros, thereby hindering the attacker's ability to exploit them. This technique of relocating system zeros to a stable region through Generalized Hold effectively counters the zero-dynamics attack. Nevertheless, it alters the closed-loop system performance compared to that with Zero Order Hold, potentially degrading from the original design.

In the research presented in (D. Kim, Ryu, and Back 2020), a Generalized Sampler is utilized for a similar objective. The Generalized Sampler enables the placement of the zeros of the discrete-time plant representation within the unit circle. Although zero-dynamics attacks can remain stealthy, their disruptive potential is mitigated as attack signals diminish. The combined use of Generalized Hold and Generalized Sampler neutralizes the zero-dynamics attack. Notably, unlike Generalized Hold, the Generalized Sampler does not impact the continuous-time system's inter-sample behavior. However, the output $y_g(k)$ differs from the conventional output sample $y(k)$, which may be undesirable for certain applications.

The vulnerability of encrypted control systems to zero-dynamics attacks is investigated in (J. Lee, Junsoo Kim, and Shim 2020). Despite the implementation of homomorphic encryption, these systems are susceptible to manipulation via knowledge of the plant model and encryption properties. The study demonstrates through simulations how the attack can disrupt the plant state while remaining undetected at the controller's input and out-

put, emphasizing the need for heightened security against zero-dynamics attacks in encrypted control systems.

D. Kim, Ryu, J. H. Kim, et al. (2021) propose an innovative defense against zero-dynamics attacks in networked control systems, involving the use of a generalized sampler to modify the system's zeros. This method remains effective against attackers with complete system knowledge due to its capability for arbitrary zero assignment. The paper details the design process of the generalized sampler and confirms its efficacy through numerical simulations, marking an advancement over existing defense methods.

The article (Park et al. 2019) delves into a robust zero-dynamics attack strategy for uncertain cyber-physical systems, addressing the challenge of model uncertainty. The authors suggest a method involving the isolation and substitution of real zero-dynamics with auxiliary nominal zero-dynamics, maintaining stealth under uncertain conditions. Utilizing a disturbance observer to compensate for model inaccuracies, the strategy proves effective without precise model knowledge. Simulations with a hydro-turbine power system demonstrate the attack's robustness, underlining the necessity for resilient security strategies against such adaptable attacks.

Baniamerian, Khorasani, and Meskin (2020) introduce a methodology for detecting zero dynamics (ZD) cyber-attacks in cyber-physical systems (CPS). They propose an auxiliary system and detection filters effective against attackers knowledgeable of the CPS and defensive measures. This approach, distinct from current methods, does not depend on hidden information or extensive system alterations, highlighting the importance of robust, resilient defenses against sophisticated ZD attacks.

The study in (Mao and Akyol 2018) focuses on detecting cooperative zero-dynamics attacks in networked systems, particularly in coupled harmonic oscillators. The authors present a defense strategy encompassing strategic topology switching and output observation, independent of the attack's initiation time or targeted oscillators. This method effectively exposes cooperative zero-dynamics attacks, ensuring they do not alter the system's steady-state value or stability. The paper substantiates this approach with simulations, contributing notably to cyber-physical network security.

In (Jihan Kim and Shim 2019a), the authors address zero-dynamics sensor attacks in cyber-physical systems, proposing 'generalized hold feedback' as a novel countermeasure to stabilize the system. This technique shifts the system's poles to stable positions, neutralizing the attack. Demon-

strated using a magnetic levitation system, this strategy not only counteracts the attack but also enhances overall system security, offering a substantial advancement in cyber-physical systems' resilience against stealthy sensor attacks.

However, a shared vulnerability among these detection techniques is evident: if attackers learn about the detection strategy, they can tamper with the signals used for detection, effectively concealing the attack. Pasha and Ayub (2021) have investigated several attack detection methods, all based on the same structure, which remains at risk if attackers uncover the detection approach. This weakness stems from the process of sending a signal back to the controller and detection system, which an attacker could alter. In Chapter 4, a new approach using Time-Scale calculus is presented. This method builds on the proven principles of previous techniques but uses a different structure and accommodates variable sampling times. This combination makes it more challenging for attackers to succeed, especially when paired with game-theory techniques.

2.3.3 Denial of Service

DoS attacks are a significant threat to the reliability and functionality of the cyber components in cyber-physical systems. These attacks primarily target and disrupt the communication channels, exploit vulnerabilities in networking protocols, and inundate the network with excessive traffic. The primary consequence of these attacks is the failure of communication links and the introduction of excessive delays, which severely impede the critical and timely exchange of vital information between sensors, actuators, and the central control systems, thereby undermining the overall system integrity.

In the realm of DoS attacks, three predominant and distinct strategies are commonly observed. Firstly, the Random DoS strategy is characterized by an unpredictable and sporadic loss of system measurement data, which can be quantitatively analyzed using the statistical Bernoulli model, as described by K. Ding et al. (2019) in their seminal paper. Secondly, the Periodic DoS strategy, a more structured form of attack, results in complete incapacitation of communication channels during the onslaught of the attack, effectively prohibiting any form of information transmission during these periods. Lastly, Non-periodic DoS attacks are marked by their irregular attack timing and intervals of dormancy, making them unpredictable and challenging to mitigate.

The concept of a Self-organizing architecture (SOA) has been proposed as a defensive structure against such attacks. SOA is strategically organized into three hierarchical tiers: the foundational tier consists of customer/generation agents, the intermediate level comprises a local controller known as the Aggregate Agent (AA), and the upper level is formed by the local control center. This structure was elaborately detailed in the study conducted by Cameron et al. (2019). They introduced a novel Cyber-Physical Power System (CPPS) platform that integrates SOA, represented in the Java Agent Development Framework (JADE), with a multilayer power system simulation using Matlab's Matpower toolbox.

This integration was highlighted as a critical advancement over previous research, which Cameron et al. (2019) critiqued for its lack of comprehensive cyber attack modeling and vulnerability assessments in power systems, thereby limiting the ability to quantitatively assess the impacts of cyber attacks. In the event of a DoS attack, the affected agent within this architecture is promptly isolated, and the control center strategically redirects communication to the most efficient available channel. The efficiency and response time of SOA are further optimized by deactivating agents deemed non-critical and isolating those compromised.

However, the effectiveness of SOA is inherently limited by the intensity and variability of the DoS attack, necessitating ongoing research to further enhance its decision-making engine, potentially through the incorporation of deep reinforcement learning techniques (Cameron et al. 2019).

Moreover, the role of Load Frequency Control (LFC) is critically examined in the context of auxiliary services in power systems. LFC plays an essential role in maintaining the balance of short-term energy and frequency, thereby stabilizing the overall power quality and facilitating efficient energy exchange. This is achieved by dynamically adjusting the rotor angles of adjacent generators in response to power deviation signals. Traditional multi-area power systems have historically relied on time-triggered mechanisms for this purpose. However, these mechanisms are now recognized as inadequate in the face of DoS attacks, primarily due to their bandwidth-intensive nature. Addressing this issue, Hossain et al. (2022) proposed a novel distributed event-trigger mechanism (DETM) to bolster the LFC system's real-time defense capabilities while concurrently minimizing the usage of communication resources.

Additionally, the application of Lyapunov function (LF) methods in assessing the stability of nonlinear systems has been explored in recent re-

search. For instance, J. Liu, Lu, and Jianhui Wang (2019) employed LF in developing a feedback-control-based time-variant direct current (DC) microgrid (MG), converting the Lyapunov candidate function solution into a convex optimization problem. In a similar vein, Briat and Seuret (2012) introduced a functional-based approach for the stability analysis of linear impulsive systems. They presented a looped Lyapunov function, which has demonstrated enhanced efficiency in conserving communication resources compared to the conventional Lyapunov-functional method, thus easing constraints related to sampling periods. Further extending this research, a Takagi-Sugeno (T-S) fuzzy observer-based control model was developed for doubly-fed induction generator (DFIG) based wind parks (WP), providing a robust framework for ensuring system resilience (Z. Hu et al. 2022).

In the domain of supervised learning, Decision Tree (DT) and Support Vector Machine (SVM) are recognized as prominent classification methods. P. Wang and Govindarasu (2020) developed an innovative Support Vector Machine embedded Layered Decision Tree (SVMLDT) model, specifically tailored for anomaly detection within power systems. This model has shown efficacy in classifying operational statuses of power systems and optimizing detection rates for DoS attacks.

Furthermore, innovative solutions such as the distributed finite-time control (DFTC) scheme proposed by Ziqiang Wang and Jie Wang (2019) integrate a Distributed Energy Storage System (DESS) to enhance the stability of power systems under DoS attacks. Chlela et al. (2018) utilized an Energy Storage System (ESS) as an active isochronous distributed generator in islanded microgrids, effectively mitigating the impacts of DoS attacks. Additionally, Farraj, Hammad, and Kundur (2018) introduced a parametric feedback linearization control (PFL) approach, which enhances the power system's delay tolerance and maintains transient stability in the aftermath of a DoS attack.

Seeking a simpler yet effective approach, we suggest adopting Time-Scale theory. By applying the continuous to Time-Scale transformation as outlined in Chapter 4, we develop a controller and observer that maintain the stability of the closed-loop system under DoS attacks, treating these attacks as delays. This method's flexibility in adjusting the sampling time for both controller and observer within a certain range ensures accurate system responses upon packet arrival, provided the delay does not exceed a predefined maximum. Consequently, this strategy effectively shields the system from less severe DoS attacks.

Chapter 3

Detection of False Data Injection using Functional Observers

3.1 Introduction

The advent of smart grids heralds a new era in energy management. Smart grids are a dynamic, adaptive power grid, capable of integrating a variety of energy sources, including renewable ones like solar and wind power. This integration is crucial for a sustainable energy future, as it reduces reliance on fossil fuels and lowers carbon emissions (Office of the National Coordinator for Smart Grid Interoperability 2012). Nonetheless, the inherent variability and unpredictability of renewable energy sources pose unique challenges in grid management, requiring advanced control systems to balance supply and demand.

However, the transition to smart grids also brings significant cybersecurity challenges (Knowles et al. 2015). The interconnected and complex nature of these grids makes them vulnerable to cyber-attacks. Such attacks can range from stealing consumer data to destabilizing the entire power system. As smart grids become more prevalent, ensuring their security against cyber threats is of paramount importance.

Several types of cyber-attacks can target smart grids, each with its unique methods and impacts. Load-altering attacks, for instance, involve attackers feeding false information to the grid's control systems, leading to incorrect decisions in power distribution. Such attacks can cause widespread power outages and have severe economic and safety implications (Miller et al. 2021; Kimani, Oduol, and Langat 2019).

More generally, False Data Injection attacks are another concern. In

these attacks, hackers subtly alter the data being transmitted within the grid, making it difficult to detect any anomalies (Pedramnia and Shojaei 2020; Xiong et al. 2020; H. Shi, L. Xie, and Peng 2021; Zhiwen Wang, J. Hu, and Sun 2020; Khazaei and Amini 2021).

Topology attacks, where attackers mislead operators about the grid's physical layout, can lead to inefficient or even dangerous responses to grid conditions (Zhenhua Wang, H. He, et al. 2021; Liberati, Garone, and Giorgio 2021; Liberati, Garone, and Giorgio 2021). Load redistribution attacks and attacks targeting market operations, like manipulating pricing information sent to smart meters, also pose threats to the stability and financial integrity of the grid (Choeum and Choi 2021; Z. Liu and L. Wang 2021; Kaviani and Hedman 2021).

Addressing these cyber-security threats requires a multi-faceted approach.

- From an Information Technology (IT) perspective, standard security measures like firewalls, encryption, and access control are essential. However, these are not sufficient on their own, as attackers can find ways to circumvent such defenses, including exploiting insider access (Mrabet et al. 2018).

- From an automation perspective, the development of resilient control and observation systems is crucial. These systems can detect anomalies in grid operation and respond effectively to mitigate the impact of attacks. Advanced detection techniques, such as using observers and residual generators, are vital in identifying and responding to cyber-attacks (Pham, Oo, and Hieu Trinh 2021; Islam, Lim, and P. Shi 2020; H. M. Tran and H. Trinh 2016; H. M. Tran, H. Trinh, and Nam 2015; H. Trinh et al. 2013).

Traditional techniques for detecting attacks, such as those based on Kalman filters, face limitations in the context of smart grids. The large number of states in a smart grid, sometimes in the hundreds, can lead to numerical issues and computational inefficiency in these techniques (Chen 2013). As a result, there is a growing need for more effective and scalable detection methods. Other methods, as those presented in Section 2.3.1, present potential numerical issues from the usage of techniques such as Singular Value Decomposition and Null Space calculation on large, sparse matrices (Arthur N. Montanari and Aguirre 2020).

Graph theory presents an avenue for enhancing smart grid security. By analyzing the grid's topology, graph theory can provide insights into the stability, controllability, and observability of the system. This approach is particularly well-suited to large and sparse systems like smart grids, avoiding the numerical problems associated with traditional methods (Aguirre, Portes, and Letellier 2018; Cowan et al. 2012). When applying through Functional Observers, they offer a balance between accuracy and computational efficiency, essential for real-time grid management.

In this chapter, we first delineate the models for both the system and the False Data Injection (FDI) attack. Then we introduce the concept of Functional Observers and do a detailed exposition on the synthesis of the functional observer, articulating the novel approach proposed. We follow by illustrating the development of a comprehensive suite of residual generators, essential for the detection and isolation of cyber-attacks. We then present validation through simulation, demonstrating the practical application and effectiveness of the proposed detection methodology.

In the subsequent section, the models pertaining to the smart grid and False Data Injection (FDI) are delineated. These models will be subsequently employed in the design of the detection system.

3.2 System and Attack Modeling

In modeling the dynamics of power networks, a prevalent approach involves using coupled second-order Kuramoto oscillators, as detailed in studies by Dorfler, Chertkov, and Bullo (2013) and Nishikawa and Motter (2015). This model effectively represents the oscillators' phases and frequencies, assuming an equilibrium frequency across the network, a concept applicable to electrical systems. Each oscillator within this framework is defined by the equation:

$$\frac{2H_i}{\omega_R}\ddot{\phi}_i + \frac{D_i}{\omega_R}\dot{\phi}_i = A_i + \sum_{j=1, j\neq i}^{N} K_{ij}\sin(\phi_j - \phi_i + \gamma_{ij}), \tag{3.1}$$

In this equation, N denotes the total number of nodes; $\phi_i(t)$ represents the phase angle of the i^{th} oscillator, adjusted to a frame rotating at the reference frequency ω_R; H_i and D_i are the inertia and damping constants, respectively; A_i is linked to the power injection at node i; K_{ij} corresponds to the coupling weight associated with the maximum power transfer capacity between nodes i and j; and γ_{ij} is the phase shift.

When expressed in a state-space form, the system is represented as:

$$\begin{bmatrix} \dot{\phi}_G \\ \ddot{\phi}_G \\ \dot{\phi}_L \end{bmatrix} = \begin{bmatrix} \dot{\phi}_G \\ \frac{\omega_R}{2H(A-\frac{D}{\omega_R})\dot{\phi}_G+\mathcal{S}} \\ \frac{\omega_R(A+\mathcal{S})}{D} \end{bmatrix}, \tag{3.2}$$

$$\mathcal{S} = \sum_{j=1,j\neq i}^{N} K_{ij}\sin(\phi_j - \phi_i + \gamma_{ij}), \tag{3.3}$$

In this representation, ϕ_G relates to generators, ϕ_L to loads and A to the power generation in the entire network. The matrices are partitioned to align with this generator-load division. A linearized version of this system is utilized in our study[1].

To identify load-altering attacks, the model encompasses three variations:

$$\tilde{\phi}_j = \phi_i, \tag{3.4}$$

$$\tilde{\phi}_j = \phi_j + \delta, \tag{3.5}$$

$$\tilde{\phi}_j = \phi_j \cdot \alpha, \tag{3.6}$$

In these equations, ϕ denotes the actual state value, and $\tilde{\phi}$ represents the measured value. The first attack type involves replacing a measurement with another state's value, as in Eq. (3.4). The second type adds a constant bias to the measurement (Eq. (3.5)), while the third multiplies the measurement by a constant (Eq. (3.6)).

Power networks comprise numerous nodes, typically numbering in the hundreds, exhibiting weak inter-connectivity. Transformers, for example, are capable of forming partially isolated sub-networks where information flow is unidirectional. This results in a dynamic, sparse graph structure. Such characteristics pose challenges for traditional methods, as delineated in Section 2.3.1, which led us to explore the applicability of Functional Observers, to which a comprehensive overview is given in the following section.

[1] The code used for linearizing the system is available at `https://github.com/acristoffers/SmartGrid/blob/master/ca_pg.m#L40`

3.3 Functional Observer

Luenberger (1966) initially proposed the concept of functional observers. These observers are characterized by their ability to estimate a selected linear combination of the state variables of a system with dynamic behavior. The primary challenge in designing such observers lies in determining the essential states for observation and constructing the corresponding system matrices.

Consider a dynamic system defined by the following set of equations:

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t) + Lf(t), \\ y(t) &= Cx(t), \\ z(t) &= Fx(t), \end{aligned} \tag{3.7}$$

In these equations, the matrices $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times p}$, and $C \in \mathbb{R}^{q\times n}$ represent the system parameters. The matrix $L \in \mathbb{R}^{n\times r}$ maps the external influences $f(t) \in \mathbb{R}^r$ to the state vector $x \in \mathbb{R}^n$. Furthermore, the matrix $F \in \mathbb{R}^{s\times n}$ transforms the state vector into the estimated state vector $z(t) \in \mathbb{R}^s$. Notably, matrices L and F are structured such that they contain a solitary non-zero entry in each column and row, respectively, corresponding to the affected and desired states.

It is pertinent to note that from the perspective of the system, $z(t)$ is merely an output, akin to $y(t)$. However, there is a distinct semantic difference: $y(t)$ represents a tangible output, directly correlating to physical sensors within the system, whereas $z(t)$ is a theoretical construct, a linear combination of states intended for estimation by the observer. For the purposes of this analysis, both $y(t)$ and $z(t)$ are direct mappings of $x(t)$, implying that matrices C and F possess a singular non-zero entry per row, each of unit value.

To accurately estimate $z(t)$ from $y(t)$, it is imperative to observe a subset of $x(t)$ that encompasses $z(t)$. This necessity stems from the inherent dynamics of the system. An observer designed solely around $z(t)$ is likely to lack the requisite dynamics for precise estimation, leading us to Definition 1, which defines functional observability.

Definition 1. *A system characterized by equation* (3.7) *is deemed functionally observable with respect to the triplet (A, C, F) if, for any initial state $x(0)$, the knowledge of $u(t)$ and $y(t)$ suffices to estimate $z(0) = Fx(0)$ over a finite time interval $t > 0$. In contrast, the system is considered functionally*

unobservable if this condition is not met (Jennings, Fernando, and H. M. Trinh 2011).

This definition of functional observability parallels the general concept of observability in systems theory, albeit with a narrower focus. Specifically, a functional observer cannot discern any state that is unobservable through the system's output. This leads to the enunciation of Theorem 1, which provides a rank condition to ascertain the functional observability of the triplet (A, C, F).

Theorem 1. *The triplet (A, C, F) satisfies the condition of functional observability if and only if (Jennings, Fernando, and H. M. Trinh 2011)*

$$rank \begin{bmatrix} C \\ CA \\ F \\ FA \end{bmatrix} = rank \begin{bmatrix} C \\ CA \\ F \end{bmatrix}. \tag{3.8}$$

The rationale behind this theorem is that a functional observer can, at best, observe the states discernible through the system's output. Therefore, the observable states must form a subset of the outputs, manifesting as linear combinations thereof. The rank condition ensures that every state in $z(t)$ is observable through $y(t)$. A more extensive mathematical proof is presented in (Jennings, Fernando, and H. M. Trinh 2011).

Identifying the requisite states for observation necessitates defining the sensor set and the desired states. In the context of developing a residual generator, it is optimal for $y(t)$ to be a subset of $z(t)$, thereby enabling the calculation of estimation errors. Nevertheless, $z(t)$ may need to be slightly more extensive than $y(t)$ to prevent the observer from merely replicating $y(t)$ into $z(t)$.

In our knowledge, there exists no algebraic method to identify the necessary states. Consequently, graph theory algorithms have become prevalent for this purpose, particularly due to their scalability in large systems. Algorithm 1, as proposed by Arthur Noronha Montanari (2021), outlines a methodology for determining the set of states that must be observed for a given $z(k)$.

Broadly speaking, the algorithm operates as follows:

1. Transforms the system into a graphical representation, wherein each node symbolizes a distinct state and the interconnecting edges reflect the dynamics inherent to the system.

2. Pinpoints a pathway linking the state that requires estimation to a corresponding output, thereby elucidating the route through which information flows.

3. Incorporates the states integral to this identified pathway into the compilation of states slated for estimation, thereby expanding the scope of the system's observability.

Algorithm 1 Method for Determining the Functionally Observable Set of States

input: triplet (A,C,S_0)
output: set S of states needed to observe S_0
let $F \leftarrow$ matrix for S_0, $\mathcal{M}_1 \leftarrow \emptyset$, $\mathcal{M}_2 \leftarrow \emptyset$, $r_0 \leftarrow \text{rank}(F)$
repeat
    let $G \leftarrow \begin{bmatrix} C^\top & (CA)^\top & F^\top \end{bmatrix}^\top$
    build a bipartite graph $\mathcal{B}(\mathcal{V}, \mathcal{X}, \mathcal{E}_\mathcal{V}, \mathcal{X})$, where $\mathcal{V} = \{v_1, \ldots, v_{2q+r_0}\}$ is a set of nodes where each element corresponds to a row of $\mathcal{G}$, $\mathcal{X} = \{x_1, \ldots, x_n\}$ is the set of state nodes (where each element also corresponds to a column of $\mathcal{G}$), and (v_i, x_j) is an undirected edge in $\mathcal{E}_\mathcal{V}$ if $\mathcal{G}_{ij}$ is a non-zero entry;
    find the maximum matching set $\mathcal{E}_m$ associated with $\mathcal{B}(\mathcal{V}, \mathcal{X}, \mathcal{E}_\mathcal{V}, \mathcal{X})$ (e.g., via the Hopcroft-Karp algorithm);
    $\forall x_i \in \mathcal{X}$, if x_i is connected to an edge in $\mathcal{E}_m$, then update the set of right matched nodes $\mathcal{M}_1 \leftarrow \mathcal{M}_1 \cup \{x_i\}$;
    define the set of candidate nodes $\mathcal{C} = \mathcal{M}_2 \setminus \mathcal{M}_1$, where $x_j \in \mathcal{M}_2$ if $[FA]_{ij}$ is a non-zero entry;
    draw an element $x_k \in \mathcal{C}$ and update $F \leftarrow \begin{bmatrix} F^\top & (F')^{top} \end{bmatrix}$ and $r_0 = r_0 + 1$, where $F' \in \mathbb{R}^{1 \times n}$ and $[F']_{ij} = 1$ if $j = k$ and 0 otherwise;
until $\mathcal{C} \neq \emptyset$

Utilizing the concepts delineated in this section, a design for a functional observer is introduced in the subsequent section.

3.4 Observer Synthesis

In addressing the challenge of detecting load-altering attacks within smart grids, our research introduces a novel approach: a bank of functional observers (e Sousa, Messai, and Manamanni 2022). These observers are distinct in their reduced-order design, which eases computational demands.

The proposed framework comprises r functional observers, each aligned with a specific attacked node within the smart grid. While these observers share a common structural foundation, they differ in the application of the

system matrix L. This matrix is pivotal in mapping the specific attack to the corresponding state. The unique configuration of each observer, tailored to individual attacked sensors, enables precise attack isolation. Additionally, each observer is equipped with its own residual generator. This generator is critical for determining the estimation error, which in turn is instrumental in identifying the presence of an attack. Figure 3.1 illustrates the schematic representation of the complete assembly of observers and their associated residual generators.

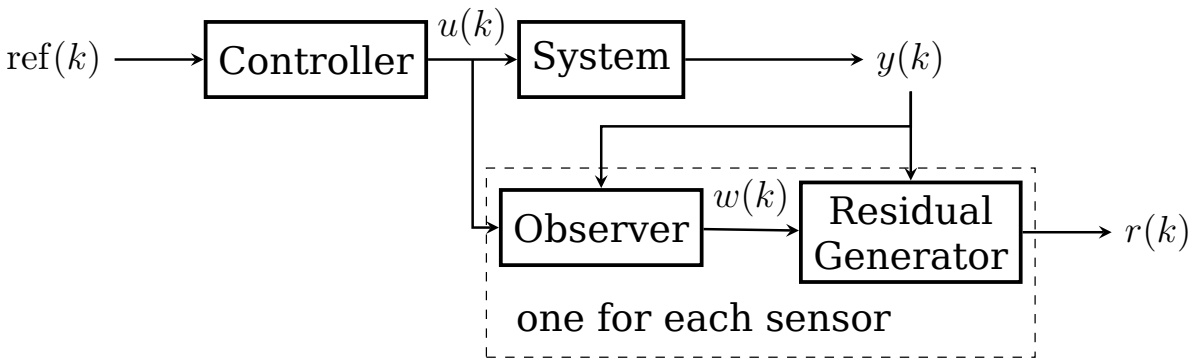


Figure 3.1 – Schematic of the Observer and Residual Generator System

We advocate for an observer design based on LMI. This formulation is advantageous due to its inherent flexibility, allowing the integration of additional constraints into the system merely by expanding the LMI framework.

Alternative approaches for the design of Functional Observers exist within the literature. Islam, Lim, and P. Shi (2020) introduces a methodology tailored to fuzzy systems, providing a nuanced perspective on handling systems characterized by uncertainty and imprecision. Similarly, Rios-Ruiz et al. (2019) delineates a design strategy for descriptor systems, which are distinguished by their algebraic constraints alongside differential equations. In the realm of systems influenced by time delays, Hieu Manh Tran and Hieu Trinh (2019) offers a formulation specifically for time-delayed systems, addressing the complexities introduced by such delays. Haes Alhelou, Golshan, and Hatziargyriou (2019) contributes to the discourse by proposing a robust observer design, focusing on resilience and adaptability without directly addressing attack identification. Finally, Pham, Oo, and Hieu Trinh (2021) ventures into the domain of electrical vehicles, presenting a technique for the detection and isolation of anomalies, thereby underscoring the relevance of Functional Observers in ensuring the operational integrity of increasingly prevalent electric mobility solutions.

Nevertheless, the methodologies discussed share a foundational princi-

ple: the partitioning and decomposition of systems. This process frequently employs techniques such as Singular Value Decomposition and the calculation of matrix null spaces. Despite their utility, these methods are acknowledged to encounter numerical challenges when applied to large, sparse matrices, as highlighted by Chen (2013). Moreover, the adoption of LMI in the proposed approach facilitates the integration of additional constraints within the system. This capability grants designers enhanced control over system performance, allowing for a more tailored and efficient observer design.

The design of our observer follows the subsequent theorem (e Sousa, Messai, and Manamanni 2022):

Theorem 2. *Given a system as defined in equation* (3.7)*, and considering the triplet (A, C, F) functionally observable in accordance with Definition 1 and Theorem 1, an observer can be formulated as:*

$$\begin{aligned} \dot{w}(t) &= Nw(t) + Jy(t) + Hu(t), \\ \hat{z}(t) &= w(t) + Ey(t), \end{aligned} \tag{3.9}$$

This observer is capable of estimating $\hat{z}(t) \approx z(t)$ given $y(t)$ and $u(t)$, provided that, for the Lyapunov candidate function

$$V(x) = x^\top P x, \tag{3.10}$$

where P is a definite positive matrix, a solution exists for the following LMI:

$$\begin{aligned} \textit{arg min} \ & \|P\|_2 \\ \textit{s.t.} \ & \dot{V} < 0 \\ & P \succ 0, \end{aligned} \tag{3.11}$$

where

$$\dot{V} < 0 \implies \begin{bmatrix} X & W \\ W^\top & -I \end{bmatrix} \prec 0, \tag{3.12}$$

with

$$X = \hat{A}^\top F^\top P - \hat{A}^\top C^\top \hat{E}^\top - \hat{C}^\top \hat{K}^\top + PF\hat{A} - \hat{E}C\hat{A} - \hat{K}\hat{C} - \lambda I, \tag{3.13}$$

$$W = \sqrt{\lambda}(PF - \hat{E}C), \tag{3.14}$$

$$\lambda \in \mathbb{R}^+ \textit{ is a free constant related to the attack magnitude}, \tag{3.15}$$

$$\tag{3.16}$$

where

$$\hat{A} = AF^{+}, \tag{3.17}$$
$$\hat{C} = CF^{+}, \tag{3.18}$$
$$\hat{E} = PE = PU + \hat{Y}V, \tag{3.19}$$
$$\hat{K} = PK, \tag{3.20}$$
$$\hat{Y} = PY. \tag{3.21}$$

The known values are then the system matrices and λ, and all other variables are optimization variables.

The observer's matrices are recovered as

$$K = P^{-1}\hat{K}, \tag{3.22}$$
$$Y = P^{-1}\hat{Y}, \tag{3.23}$$
$$E = U + YV, \tag{3.24}$$
$$R = F - EC, \tag{3.25}$$
$$N = (RA - KC)F^{+}, \tag{3.26}$$
$$J = K + NE, \tag{3.27}$$
$$H = RB, \tag{3.28}$$

Proof. Start by defining the estimation error e as the discrepancy between the estimated state $\hat{z}$ and the actual state z. For simplification, temporal dependencies in the vector notation are omitted:

$$e = \hat{z} - z = w + Ey - Fx = w + ECx - Fx. \tag{3.29}$$

Subsequently, the dynamics of this estimation error are calculated as follows:

$$\begin{aligned}
\dot{e} &= \dot{w} + (EC - F)\dot{x} \\
&= Nw + Jy + Hu + (EC - F)(Ax + Bu + Lf) \\
&= N(e + Fx - ECx) + JCx + Hu + ECAx \\
&+ ECBu + ECLf - FAx - FBu - FLf \\
&= Ne + (NF - NEC + ECA - FA + JC)x \\
&+ (H + ECB - FB)u + (ECL - FL)f.
\end{aligned} \tag{3.30}$$

For the convergence of $\hat{z}$ to z, N must be Hurwitz-stable, and the following

conditions must be satisfied:

$$N(F - EC) - (F - EC)A + JC = 0, \tag{3.31}$$

$$H - (F - EC)B = 0, \tag{3.32}$$

$$(F - EC)L = 0. \tag{3.33}$$

To adapt the observer to be selectively insensitive to specific sensor attacks, we modify the latter condition. Let $L = [\, L_i \; L_n \,]$, with $L_i \in \mathbb{R}^{n\times 1}$ representing the mapping of the insensitive attack, and $L_n \in \mathbb{R}^{n\times l-1}$ corresponding to the sensitive attacks. Consequently, we establish the following criteria for each observer:

$$(F - EC)L_i = 0, \tag{3.34}$$

$$(F - EC)L_n \neq 0. \tag{3.35}$$

The subsequent step involves determining the matrices N, J, H, and E. This is achieved through the optimization of a LMI. We introduce the Lyapunov candidate function:

$$V = e^\top Pe. \tag{3.36}$$

Incorporating the aforementioned restrictions, the dynamics of the error function become:

$$\dot{e} = Ne - (F - EC)L_n f. \tag{3.37}$$

To conceptualize the attack in terms of error, we assume a proportionality relation between the error and the variation introduced by the attack:

$$e \propto L_n f, \tag{3.38}$$

leading to the equation:

$$\|L_n f\| = \lambda \|e\|, \tag{3.39}$$

where $\lambda \in \mathbb{R}^+$ is a freely adjustable parameter that scales the attack in proportion to the error.

With this formulation, and defining

$$R = F - EC, \tag{3.40}$$

the derivative of the Lyapunov candidate function is expressed as:

$$\begin{aligned}
\dot{V} &= \dot{e}^\top Pe + e^\top P\dot{e} \\
&= (Ne - \lambda R\|e\|)^\top Pe + e^\top P(Ne - \lambda R\|e\|) \\
&= e^\top (N^\top P + PN)e - 2\lambda \|e^\top PR\| \cdot \|e\| \\
&\leq e^\top (N^\top P + PN)e - \lambda(\|e^\top PR\|^2 + \|e\|^2) \\
&= e^\top (N^\top P + PN - \lambda PRR^\top P - \lambda I)e,
\end{aligned} \tag{3.41}$$

where I denotes the identity matrix of appropriate dimensions.

This Bilinear Matrix Inequality (BMI) assures the stability of the matrix N. To integrate the constraint delineated in equation (3.31), we define:

$$N(F - EC) = RA - JC, \tag{3.42}$$

$$NF = RA - (J - NE)C, \tag{3.43}$$

$$K = J - NE, \tag{3.44}$$

$$N = RAF^{+} - KCF^{+}, \tag{3.45}$$

where $K \in \mathbb{R}^{s \times q}$ is a matrix determined through optimization and F^{+} represents the Moore-Penrose inverse (also known as pseudo-inverse) of matrix F.

To accommodate the condition in (3.34), the following equations are posited:

$$(F - EC)L_i = 0, \tag{3.46}$$

$$ECL_i = FL_i, \tag{3.47}$$

$$E = FL_i(CL_i)^{+} + Y(I - (CL_i)(CL_i)^{+}), \tag{3.48}$$

$$U = ECL_iL_i^{+}, \tag{3.49}$$

$$V = I - L_iL_i^{+}, \tag{3.50}$$

$$E = U + YV, \tag{3.51}$$

revisiting the equations from (3.17) to (3.21).

Consequently, equation (3.41) transforms into:

$$\dot{V} = e^{\top}((R\hat{A} - EC\hat{A} - K\hat{C})^{\top}P + P(R\hat{A} - EC\hat{A} - K\hat{C}) - \lambda PRR^{\top}P - \lambda I)e. \tag{3.52}$$

After comprehensive expansion and substitution of variables, the derivative of the candidate function ultimately resolves to:

$$\dot{V} = e^{\top}(\hat{A}^{\top}F^{\top}P - \hat{A}^{\top}C^{\top}\hat{E}^{\top} - \hat{C}^{\top}K^{\top} + PF\hat{A} - \hat{E}C\hat{A} - K\hat{C} - \lambda PRR^{\top}P - \lambda I)e. \tag{3.53}$$

Finally, to address the bilinear nature of $\dot{V}$, resulting from the variable R, the Schur complement technique is employed, as referenced in Boyd et al. (1994).

$$X = \hat{A}^{\top}F^{\top}P - \hat{A}^{\top}C^{\top}\hat{E}^{\top}\hat{C}^{\top}K^{\top} + PF\hat{A} - \hat{E}C\hat{A} - K\hat{C} - \lambda I, \tag{3.54}$$

$$W = \sqrt{\lambda}(PF - \hat{E}C), \tag{3.55}$$

$$\dot{V} = e^{\top}\begin{bmatrix} X & W \\ W^{\top} & -I \end{bmatrix} e. \tag{3.56}$$

□

The estimation error within this observer system is defined as:

$$e = w + ECx - Fx. \tag{3.57}$$

The dynamics of this error are then elucidated, leading to a set of conditions that ensure convergence of the estimated state $\hat{z}$ to the actual state z. These conditions include the Hurwitz stability of matrix N and other specific constraints outlined in equations (3.31), (3.32), and (3.33). Notably, to achieve insensitivity to a particular sensor attack while remaining responsive to others, we redefine condition (3.33) as illustrated in equations (3.34) and its complement.

Subsequently, the matrices N, J, H, and E are determined through an LMI optimization process, incorporating the Lyapunov candidate function and considering the necessary restrictions. This results in the final derivative of the candidate function as presented in equation (3.41) and its expanded form.

Possessing an observer that exhibits insensitivity to attacks on a single sensor, the subsequent section introduces the design of a residual generator. This generator is devised to offer a numerical method for assessing the presence of an attack, while simultaneously accommodating disturbances within the system.

3.5 Residual generators

A residual refers to the discrepancy between an estimated value and the actual value. Under normal conditions, without any external interference, it is anticipated that this residual will approximate zero. Conversely, in the event of an attack, a marked deviation from zero is expected. A binary function, based on a set threshold, can use this residual to indicate the presence of an attack in the system. This subsection focuses on the dynamics of the residual rather than the binary function.

We introduced a residual generator that employs a specifically designed observer. This observer is insensitive to one type of attack and generates an output reflective of the estimation error, termed as the residual. This residual is essentially a measure of the observer's accuracy in state estimation. During an attack, which is an external, unmeasured signal, the observer's state estimation becomes inaccurate, leading to a noticeable change in the residual. The following lemma encapsulates this concept:

Lemma 1. *A residual signal can be generated as*

$$r(t) = Gw(t) + My(t), \tag{3.58}$$

with

$$M = (C(1 - L_i))^\top, \tag{3.59}$$
$$G = -M(I - CF^+E)^{-1}CF^+, \tag{3.60}$$

where $1 - L_i$ *signifies an entry-wise operation. This forms a residual generator for observer* (3.9)*, as outlined in Lemma 2.*

Proof. The error equation for the residual generator, utilizing the observer's matrices, is given by

$$\begin{aligned} r &= Gw + My \\ &= G(e + Fx - ECx) + MCx \\ &= Ge + (G(F - EC) + MC)x, \end{aligned} \tag{3.61}$$

For the residual generator to converge to zero in the absence of error, the following condition must be met:

$$G(F - EC) + MC = 0. \tag{3.62}$$

Expressing the error equation in terms of the output, we have:

$$\begin{aligned} r &= Gw + My \\ &= Q(y - Cx) \\ &= Q(y - CF^{-1}\hat{z}) \\ &= Q(y - CF^{-1}(w + Ey)) \\ &= Q((I - CF^{-1}E)y - CF^{-1}w), \end{aligned} \tag{3.63}$$
$$M = Q(I - CF^{-1}E), \tag{3.64}$$
$$G = -QCF^{-1}. \tag{3.65}$$

Substituting Equations (3.64) and (3.65) into (3.62) confirms that the condition is satisfied regardless of the value of Q. Thus, Q can be utilized to adjust the residual's magnitude. Hence, M is defined using matrix L as in (3.66):

$$M = (C(1 - L_i))^\top, \tag{3.66}$$

representing the sum of all outputs $y(t)$ prone to error, excluding the insensitive one. By replacing Eq. (3.66) into Eq. (3.64) and eliminating Q from

Eq. (3.65), we derive

$$M = (C(1 - L_i))^\top, \tag{3.67}$$

$$G = -M(I - CF^+E)^{-1}CF^+. \tag{3.68}$$

Note that Equations (3.64) and (3.65) may also be directly applied by identifying a matrix Q with desirable properties, such as generating a residual with sufficiently large magnitude during expected attacks. □

The merit of our proposed technique lies in the bespoke nature of observer design, offering more control to designers compared to prevalent algebraic methods in literature, such as the one in Emami et al. (2013). These conventional methods, while straightforward, do not support the integration of constraints like $\mathcal{D}$-stability or considerations for saturations. Our LMI approach is straightforward and adaptable, allowing for the incorporation of various LMI techniques into observer design.

Several existing solutions suffer from numerical instability, as exemplified in Arthur Noronha Montanari (2021), which relies on the inversion of a segment of the system matrix A. This approach is problematic in large, sparse systems as the inversion of sparse matrices is inherently unstable. While there are algorithms to compute such inversions, their performance is context-dependent and not universally reliable, thus best avoided. Operations like singular value decomposition and null-space calculation also pose stability concerns in large, sparse systems.

The issue of numerical stability underscores the preference for functional observers over traditional methods like the Kalman filter (Chen 2013). Established techniques, effective in identifying similar attacks, do not scale well with large, sparse systems, which is a limitation for most Kalman-based methods. Conversely, these techniques might not be as relevant for systems with fewer states.

In the following section, a simulation is presented to demonstrate the efficacy of the proposed methodology.

3.6 Simulation Example

To demonstrate the effectiveness of our methodology, let's explore the IEEE 118-bus test system, a model representing the electrical power network of the U.S. Midwest. This system comprises 19 generators, 35 synchronous condensers, 177 transmission lines, nine transformers, and 91

loads. The schematic of this network is depicted in Figure 3.2. The corresponding data file, a MATLAB file encapsulating the constants for (3.1) in matrix form, is accessible on GitHub[2]. Additionally, the data in IEEE Common Data Format can be found online[3], with the base KV levels being estimated due to their absence in the original data-set.

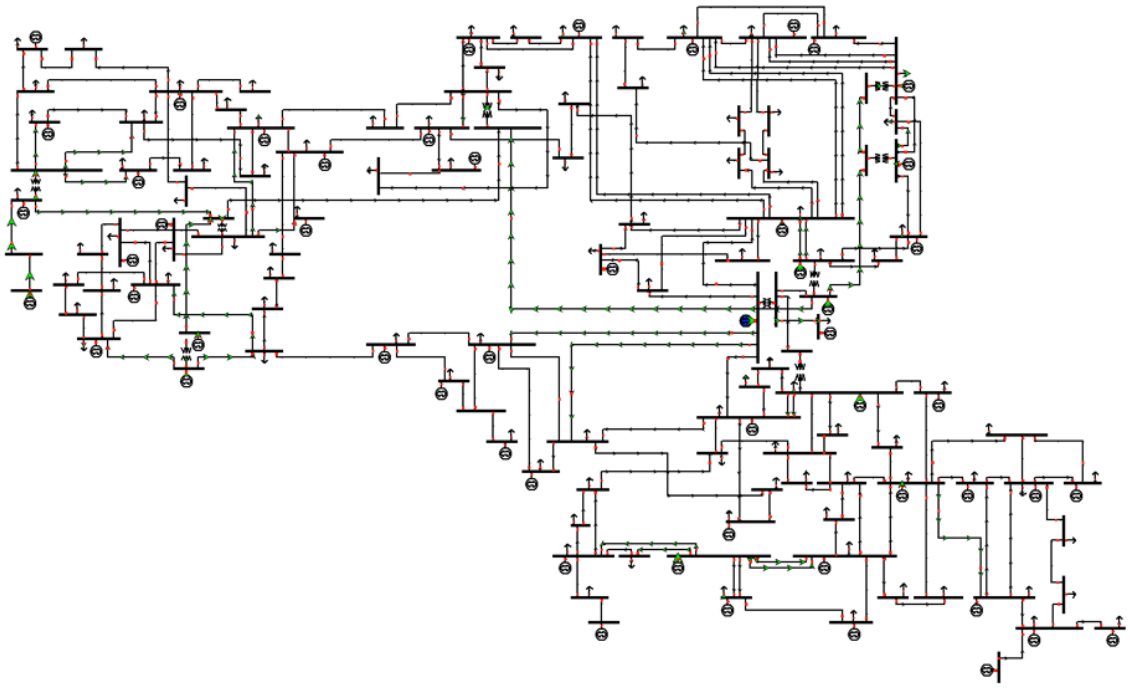


Figure 3.2 – Schematic of the IEEE 118-bus network

We simulated the IEEE 118 power grid model using a linearized version of the dynamic system described in (3.3). Figure 3.3 presents the dynamic graph of the system, highlighting generators, loads, and sensors. The graph reveals sparse connectivity with an average node degree of 2 and high betweenness centrality, implying longer paths between sensors and desired states and the criticality of intermediary states for effective observation. To monitor the targeted states, observers need to estimate approximately 150 out of 226 total states, which is about 66%, underscoring the utility of a functional observer.

Sensor placement was randomized, with PMUs installed at 30% of the terminals of generators and loads, totaling 35 sensors. Arthur Noronha Montanari (2021)'s algorithm identified a set S of 126 states. Employing

[2] `https://github.com/acristoffers/SmartGrid/blob/master/powergrid/IEEE118pg.mat`
[3] `https://labs.ece.uw.edu/pstca/pf118/pg_tca118bus.htm`

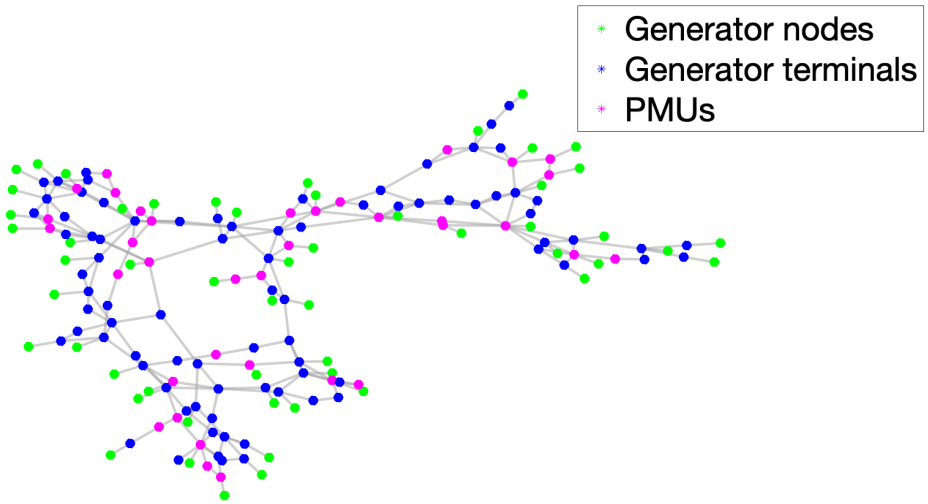


Figure 3.3 – Dynamic graph representation of the IEEE 118-bus system

Theorem 2 and Lemma 1, we created a suite of observers and residual generators for the system, capable of identifying load-altering attacks.

Three types of load-altering attacks were simulated: the first duplicates the value of another state at the attacked node, the second adds a constant to the existing signal, and the third multiplies the state value by a constant. Figure 3.4, Figure 3.5, and Figure 3.6 display the output of the residual generators for each type of attack.

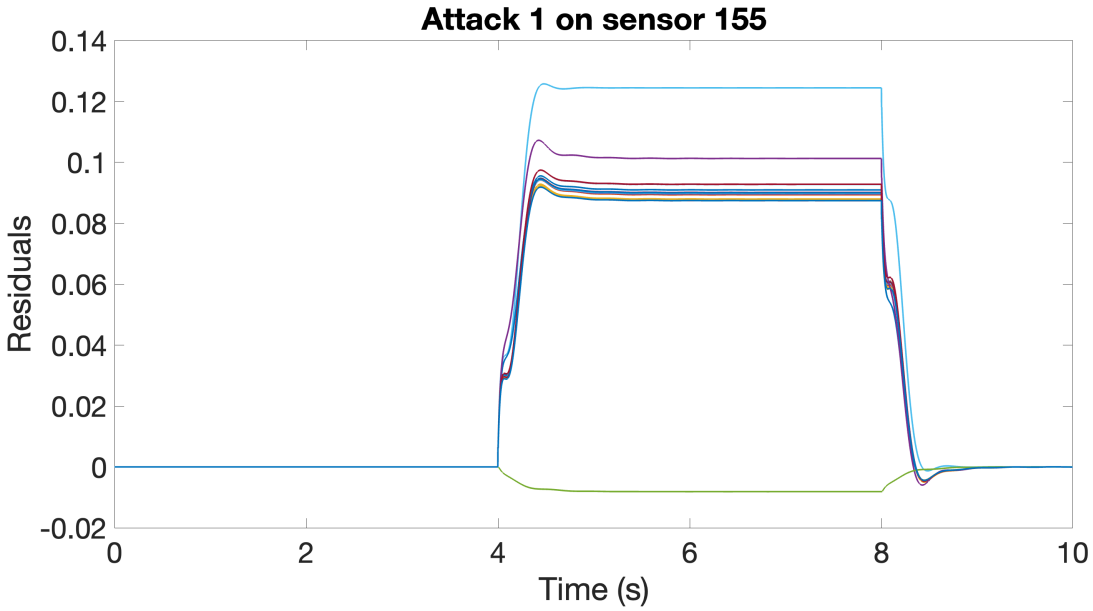


Figure 3.4 – Residuals for state-value-copy attack on state 155

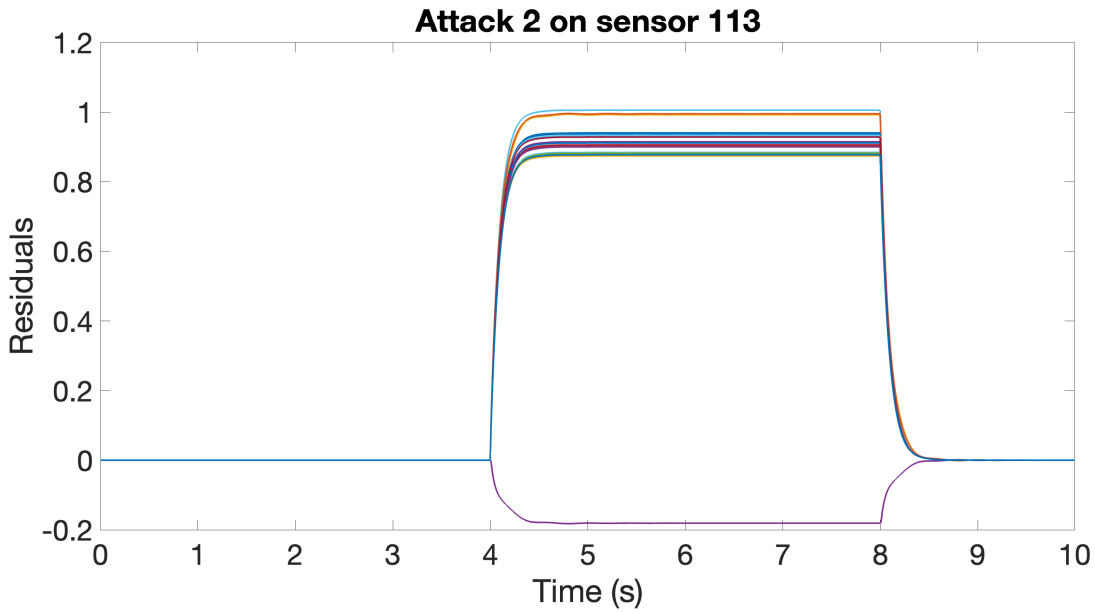


Figure 3.5 – Residuals for additive attack on state 113

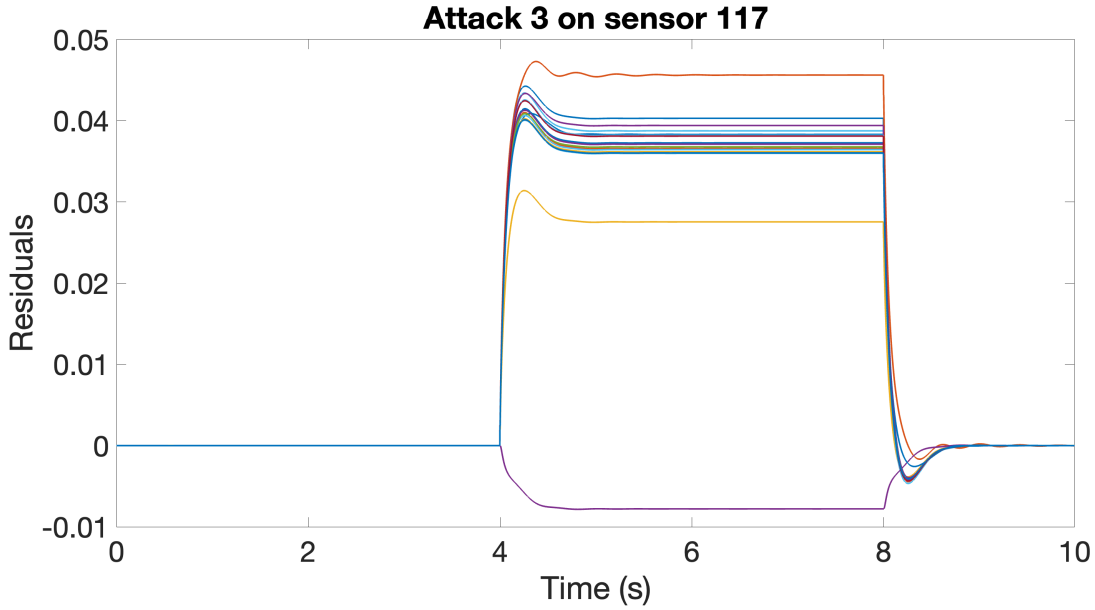


Figure 3.6 – Residuals for multiplicative attack on state 221

The implemented residual generators successfully detected each attack type. The residuals were zero before the attack, shifted from zero during the attack, and returned to zero post-attack. One residual remained consistently near zero, as its corresponding observer was insensitive to the attack on its sensor, facilitating the identification of attack occurrences and isolation of the compromised node.

The simulation code and all generated matrices are available online[4].

[4] `https://github.com/acristoffers/SmartGrid`

We do not present direct comparisons with other methods because the value of the proposed technique lies in its formulation, which does not suffer from numerical problems on large, sparse systems, and straightforward formulation. Comparing performance would not make sense, as all methods have some tuning parameter which would lead to claims that the performance of any particular method would be better with a different choice of parameters, not providing an objective assessment.

To access the computational cost of the proposed solution, the number of variables and lines in the LMIs are of interest:

- Observer design ($N \in \mathbb{R}^{n\times n}, \hat{K} \in \mathbb{R}^{n\times q}, \hat{Y} \in \mathbb{R}^{n\times s}$):
  - Number of variables: $\frac{n(n+1)}{2} + n(q+s) + 1$
  - Number of lines: $2n^2 + n$

The proposed residual generators take a few milliseconds to detect the attacks. This is, however, by chance, since the formulation does not enforce any kind of timing. That's not to say that this cannot be controlled, though: the N matrix of the observer drives its dynamics and can be used to control the convergence time of the observer and, therefore, the attack detection time. A way of doing so is to add constraints to the LMI that specify a region for pole allocation, with methods like those found in Duan and H.-H. Yu (2013).

3.7 Conclusion

In this chapter, we introduced a straightforward LMI-based design for a suite of functional observer residual generators, primarily aimed at detecting load-altering attacks. The simplicity of our approach is a noteworthy aspect, as it not only facilitates the understanding and implementation of the design but also offers considerable flexibility in extending the design to incorporate various constraints within the observer's dynamics. This flexibility is also advantageous in applications beyond attack detection, such as in the realms of state observation with disturbance rejection, fault detection, and potentially in thwarting other kinds of cyber-physical attacks.

Our method was evaluated on a model of a power grid, where it demonstrated its efficacy by enabling a significant reduction in the number of states required for observation. This reduction translates into a leaner

observer system, which in turn leads to faster computation times and improved efficiency. Such an outcome is particularly valuable in real-world applications where computational resources and time are critical factors.

In summary, the straightforwardness and adaptability of our LMI-based design for functional observer residual generators not only make it a potent tool for detecting load-altering attacks but also open up a broad spectrum of applications in enhancing the robustness and reliability of various systems. The potential for future developments in this domain is vast, promising even greater advancements in the field of cyber-physical system security and efficiency.

In the subsequent chapter, the emphasis transitions to the Zero-Dynamics Attack (ZDA) and the introduction of our proposed solution predicated upon Time-Scale Theory.

Chapter 4

Detection of Zero-Dynamics Attacks using observers on time-scale

4.1 Introduction

In the preceding chapter, an examination was conducted on the False Data Injection (FDI) attack, culminating in the proposition of a detection methodology. This chapter pivots attention towards the Zero-Dynamics Attack (ZDA), for which a novel detection solution is proposed.

A particularly insidious type of attack within this realm is the ZDA. ZDA, requiring comprehensive system knowledge, is notoriously difficult to detect due to its stealthy nature. It exploits the system's transmission-zero dynamics to manipulate the control signal in a manner that impacts the system states (potentially destabilizing the system) without altering the system output. Consequently, the sensors within the system remain unaffected by the attack, complicating its detection (André Teixeira et al. 2015).

Traditional detection techniques prove ineffective against ZDA due to its covert impact on system outputs, as discussed in Section 2.3.2, which led us to the exploration of the Time-Scale Theory as an alternative. Distinct from the two-time-scale and three-time-scale systems documented in existing literature, Time-Scale Theory represents a comprehensive framework for the analysis of dynamic systems, accommodating arbitrary time domains. This framework encapsulates both continuous and discrete-time systems as particular instances within its broader applicability (Agarwal et al. 2002), and allows for arbitrary changes in sampling time, surpassing the limitations of

a bank of discrete observers which are confined to a fixed set of sampling time values. The flexibility of the time-scale system in selecting any value for the sampling time enhances its effectiveness in attack detection.

Despite its potential, the time-scale approach presents challenges in direct application or adaptation of classic continuous and discrete-time techniques, particularly due to its time-dependent model elements. The practicality of verifying observability in such systems remains an unsolved issue and is the subject of ongoing research (Ben Nasser et al. 2022). This work won't delve into the verification of system observability, assuming that the designer did verify it.

This chapter's primary contribution is the development of a novel observer capable of dynamically adjusting its sampling time through simple parameter variations. The observer's gain is designed as a convex function, involving only additions and multiplications, thereby facilitating its implementation even in resource-constrained environments. This adaptability in sampling time effectively renders the ZDA visible. To the best of our knowledge, this innovative application of the time-scale approach has not been previously explored in the literature. Other significant features of our approach include:

- The use of a Linear Parameter Varying (LPV) observer with a mere four vertices, independent of the system's size, which avoids exponential growth of the number of vertices.
- The number of restrictions and variables in the derived Linear Matrix Inequality (LMI) grows linearly with the system, and does not exhibit exponential growth.
- The observer's capability to alter its sampling time at any point introduces a random element in attack detection, significantly complicating an attacker's efforts to conceal the attack.

Thus, the proposed technique notably improves upon existing methods in two critical aspects: 1) it does not necessitate alterations to the system's model dynamics, and 2) it does not require changes in the network topology, presenting a simpler yet effective alternative for zero-dynamics attack detection.

In this chapter, we first delve into the specifics of ZDA, outlining its execution and impact on system states while maintaining a stealthy profile. Then we provide an introduction to Time-Scale Theory, highlighting its most

important elements. Afterwards we introduce a novel observer design utilizing Time-Scale Theory, offering a comprehensive framework for dynamic system analysis before presenting simulation results to validate the effectiveness of the proposed detection technique.

The subsequent section elucidates the ZDA and delineates its application to impact a system.

4.2 Zero-Dynamics Attack

ZDA is executed by injecting energy into the system in such a manner that the system states are either driven to a perilous zone or forced to diverge, all the while maintaining a stealthy profile by not affecting the observable output. The critical point of vulnerability lies in the network, where the attack manipulates the control signals transmitted from the controller to the physical system. This manipulation remains undetected by the control side due to the unchanged output, hence obscuring the presence of the attack.

Given the assumption of a compromised network, the conventional detection schemes located at the control side become ineffective, as the attacker could easily falsify any transmitted information. As a countermeasure, it is advised to position the detection mechanism directly within the system's framework, integrating it with the alarm systems, as illustrated in Figure 4.1.

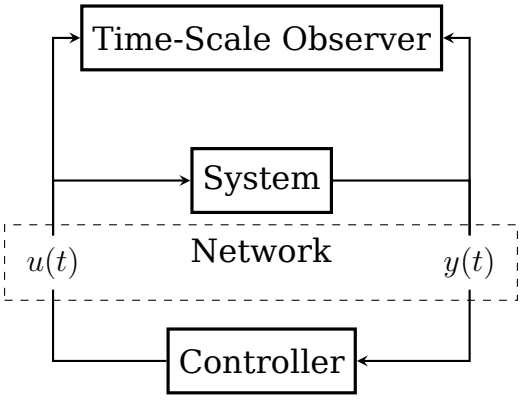


Figure 4.1 – Schematic representation of the observer's integration

Despite necessitating comprehensive knowledge of the system, the formulation of the attack is relatively straightforward. Consider a continuous-

time system characterized by

$$\dot{x}(t) = \mathcal{A}x(t) + \mathcal{B}u_a(t), \tag{4.1}$$

$$y(t) = \mathcal{C}x(t), \tag{4.2}$$

where $u_a(t)$ represents the attack-induced control signal. The system's transfer zeros are defined as the roots of $\det(P(s))$, with $P(s)$ being the Rosenbrock matrix, as defined in Equation (4.3). The intricate relationship between the Rosenbrock matrix and the system's transmission zeros is elucidated in Lemma 2.

Lemma 2. *Given the Rosenbrock system matrix*

$$P(s) = \begin{bmatrix} sI - \mathcal{A} & -\mathcal{B} \\ \mathcal{C} & D \end{bmatrix}, \tag{4.3}$$

its determinant represents the numerator of the system's transfer-function realization. Specifically, the transfer function is expressed as $G(s) = \frac{\det(P(s))}{\det(sI-\mathcal{A})}$*, where* $\det(P(s))$ *is a polynomial, potentially containing cancelled zeros (Murdoch 1973).*

The zero direction z_0 corresponding to the zero s_0 is determined by

$$P(s_0)z_0 = 0. \tag{4.4}$$

Driving the system along z_0 ensures that any alterations in the states are not reflected in the output, as highlighted by Baniamerian, Khorasani, and Meskin (2020). The attack control signal can be formulated as

$$z_0 = \begin{bmatrix} x_0 \\ a_0 \end{bmatrix}, \tag{4.5}$$

$$a(t) = a_0 e^{s_0 t}. \tag{4.6}$$

The altered control signal $u_a(t)$ is computed as $u_a(t) = u(t)+a(t)$, effectively preserving the original output as if it were influenced by $u(t)$ alone.

In the subsequent section, the focus shifts to elucidate the fundamentals of Time-Scale Theory, the methodology employed for the detection of ZDA.

4.3 Time-Scale Theory

In this section, we offer a foundational overview of the essential tools required for working with time-scale systems. For those seeking a more

comprehensive understanding of time-scale calculus, we recommend consulting the work of Agarwal et al. (2002).

Time-scale calculus serves as a unifying framework that amalgamates the principles of continuous and discrete calculi. It generalizes the concepts of differentials ($\dot{x}(t) = f(x(t))$) and differences ($\Delta x(k) = f(x(k))$) into a singular formulation ($x^{\Delta}(t) = f(x(t))$). The most pivotal aspect of this formulation lies in the non-uniformity and arbitrariness of the time domain ($\mathbb{T}$). To illustrate, consider Figure 4.2, where the first line depicts the initial 10 seconds of the time domain for a continuous differential equation, corresponding to the set of real numbers ($\mathbb{T} = \mathbb{R}$); the second line represents the time domain of a discrete difference equation ($\mathbb{T} = h\mathbb{Z}$) with a sampling time of $h = 0.5\,\mathrm{s}$; and the third line demonstrates the time domain of a time-scale equation, comprising a mixture of continuous, discrete, and other arbitrary points, thereby forming a valid time-scale domain.

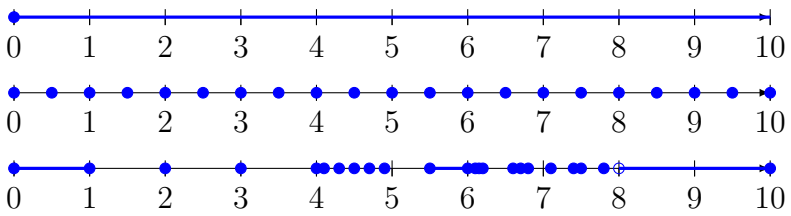


Figure 4.2 – Illustration of time domains for the continuous time (first line), discrete time (second line) and time scale (last line).

A common operation in dealing with differential or difference equations involves navigating forward and backward to valid time points. Expressions like $x(k+1)$ and $x(t+\delta)$ frequently appear in equation manipulations. The forward and backward operators in time-scale calculus serve as equivalents to these, identifying valid time points immediately before and after a specified time point. These operators are defined as follows:

$$\text{forward: } \sigma(t) = \inf\{\tau \in \mathbb{T} | \tau > t\}, \tag{4.7}$$

$$\text{backward: } \rho(t) = \sup\{\tau \in \mathbb{T} | \tau < t\}. \tag{4.8}$$

The concept of graininess, denoted as $\mu(t) = \sigma(t) - t$, is crucial in time-scale calculus. It specifies the distance between a given time point t and the next valid point in time. In a continuous domain ($\mathbb{T} = \mathbb{R}$), graininess is zero ($\mu(t) = 0$), while in a discrete domain ($\mathbb{T} = h\mathbb{Z}$), it equals the sampling time ($\mu(t) = h$). In a time-scale equation, however, graininess is not constant and varies as time progresses. This variability is illustrated in Figure 4.2, where the graininess assumes different values within a ten-second span.

With these definitions of operators and graininess, we can now introduce the Δ-derivative, the time-scale equivalent of the derivative. For a function $f(t)$ defined within the time scale, the Δ-derivative is given by:

$$f^{\Delta}(t) = \lim_{\delta \to \mu(t)} \frac{f(t+\delta) - f(t)}{\delta}, \tag{4.9}$$

Here, $f^{\Delta}(t)$ represents the Δ-derivative of $f(t)$. This formulation encompasses both the differential and difference derivatives as special cases, thereby unifying these concepts under a single, versatile framework.

Physical systems can be represented through direct modeling employing time-scale equations. An alternative approach involves transforming a pre-existing continuous-time system into a time-scale system, as delineated in Lemma 3 (Agarwal et al. 2002). This manuscript focuses exclusively on the latter methodology due to its inherent properties that facilitate certain simplifications.

Lemma 3. *Consider a continuous-time state-space system represented by*

$$\begin{aligned} \dot{x}(t) &= \mathcal{A}x(t) + \mathcal{B}u(t), \\ y(t) &= \mathcal{C}x(t), \end{aligned} \tag{4.10}$$

where $\mathcal{A} \in \mathbb{R}^{n\times n}$, $\mathcal{B} \in \mathbb{R}^{n\times m}$, $\mathcal{C} \in \mathbb{R}^{p\times n}$, $x(t) \in \mathbb{R}^n$, and $y(t) \in \mathbb{R}^p$. Then, the equivalent time-scale system is given by

$$\begin{aligned} x^{\Delta}(t) &= A(\mu(t))x(t) + B(\mu(t))u(t), \\ y(t) &= \mathcal{C}x(t), \end{aligned} \tag{4.11}$$

where

$$A(\mu(t)) = \frac{e^{\mathcal{A}\mu(t)} - I}{\mu(t)}, \tag{4.12}$$

$$B(\mu(t)) = \int_0^{\mu(t)} \frac{e^{(\mu(t)-s)\mathcal{A}}}{\mu(t)} \mathcal{B} ds, \tag{4.13}$$

Proof. The solution to the ordinary differential equation

$$\dot{x}(t) = \mathcal{A}x(t) + \mathcal{B}u(t) \tag{4.14}$$

with respect to time yields

$$x(t) = e^{(t-t_0)\mathcal{A}}x(t_0) + \int_{t_0}^{t} e^{(t-\tau)\mathcal{A}}\mathcal{B}u(\tau)d\tau. \tag{4.15}$$

By altering the interval from (t_0, t) to $(t, \sigma(t))$, the equation transforms to

$$\begin{aligned} x(\sigma(t)) &= e^{(\sigma(t)-t)\mathcal{A}}x(t) + \int_t^{\sigma(t)} e^{(\sigma(t)-\tau)\mathcal{A}}\mathcal{B}u(\tau)d\tau \\ &= e^{\mathcal{A}\mu(t)}x(t) + \int_0^{\mu(t)} e^{(\mu(t)-\tau)\mathcal{A}}\mathcal{B}u(t)d\tau. \end{aligned} \tag{4.16}$$

Applying (4.9) to (4.16), we obtain

$$x^{\Delta}(t) = \frac{x(\sigma(t)) - x(t)}{\mu(t)} \tag{4.17}$$

$$= \underbrace{\frac{e^{\mu(t)\mathcal{A}} - I}{\mu(t)}}_{A(\mu(t))} x(t) + \underbrace{\int_0^{\mu(t)} \frac{e^{(\mu(t)-\tau)\mathcal{A}}}{\mu(t)}\mathcal{B}d\tau}_{B(\mu(t))} u(t). \tag{4.18}$$

□

Remark: in (4.17), if $\mu(t) = 0$, then it is necessary to take the limit, which becomes the continuous-time derivative by definition, implying $A(0) = \mathcal{A}$, $B(0) = \mathcal{B}$ and $x^{\Delta}(t) = \dot{x}(t)$, which can be verified by applying the limit to (4.18).

Time-scale calculus offers a unified framework for derivatives, analogous to how time-scale systems provide an alternative methodology for stability analysis. The Hilger criterion, integral for assessing the stability of time-scale systems, is delineated in (Agarwal et al. 2002). It mandates that all eigenvalues of $A(\mu(t))$ reside within a circle of radius $\mu(t)^{-1}$ centered at $(-\mu(t)^{-1}, 0)$, expressed as:

$$\mathcal{H}_\mu \coloneqq \left\{ z \in \mathbb{C}_\mu : |z + \frac{1}{\mu}| < \frac{1}{\mu} \right\}. \tag{4.19}$$

In cases where $\mu(t) = 1$, this criterion converges to that of the discrete case, specifically the difference equation scenario ($\Delta x(k) = f(x(k))$) rather than the advance operator format ($x(k+1) = f(x(k))$). Conversely, for $\mu(t) = 0$, the circle expands infinitely, essentially encompassing the entire left semi-plane of the complex plane as discussed in (Davis et al. 2010). Figure 4.3 illustrates the regions of stability for three distinct scenarios, each corresponding to different values of $\mu(t)$.

For brevity, subsequent sections of this document will omit the explicit mention of the μ and t dependencies in variables. For instance, (4.11) is simplified to:

$$\begin{aligned} x^{\Delta} &= Ax + Bu, \\ y &= \mathcal{C}x. \end{aligned} \tag{4.20}$$

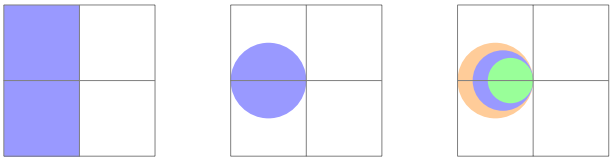


Figure 4.3 – Illustration of Different Stability Regions

Adherence to the Hilger criterion is both a necessary and sufficient condition for time-scale systems. However, it is feasible to deduce additional criteria that comply with the Hilger criterion while presenting alternative formulations. One such example is the Lyapunov stability condition, elaborated in Lemma 4 (Davis et al. 2010). This condition will later be employed in the formulation of the observer.

It is imperative for time-scale systems to adhere to the Hilger criterion, as it constitutes a necessary and sufficient condition for stability. However, additional criteria can be derived to fulfill this foundational requirement while offering alternative analytical frameworks. One such example is the Lyapunov stability condition, elucidated in Lemma 4 (Davis et al. 2010), which will subsequently be employed in the development of an observer framework.

Lemma 4. *The system described by* (4.20) *is deemed Lyapunov-stable if there exists a positive definite matrix* P *satisfying the following linear matrix inequality:*

$$PA + A^\top P + \mu A^\top P A \prec 0. \tag{4.21}$$

Proof. Considering the classic candidate function $V = x^\top P x$, and employing (4.9), we find:

$$\begin{aligned} V^\Delta &= \frac{V(\sigma) - V}{\mu} \\ &= \frac{x(\sigma)^\top P x(\sigma) - x^\top P x}{\mu} \\ &= \frac{\mu x^\top P x^\Delta + \mu x^{\Delta\top} P x + \mu^2 x^{\Delta\top} P x^\Delta}{\mu} \\ &= x^\top (PA + A^\top P + \mu A^\top P A) x. \end{aligned} \tag{4.22}$$

□

Having elucidated the fundamentals of Time-Scale Theory, the following section concentrates on observer synthesis.

4.4 Observer Synthesis

This section delineates the procedural framework for observer synthesis within time-scale systems, structured into three distinct subsections. The initial subsection elucidates the construction of an observer for a static parameter μ, establishing the foundational concepts pivotal for the subsequent discourse. Subsequently, Section 4.4.2 delineates several pertinent lemmas that underpin the advanced theoretical constructs developed later. The culminating subsection, as detailed in Section 4.4.3, introduces the principal contribution of this chapter: the observer for parameter-varying time scales which we will use to detect ZDA attacks. Through this chapter, the system is supposed to be observable, as it makes no sense to develop an observer for an unobservable system. For more information on verifying the observability of a time-scale system, we invite the reader to refer to (Ben Nasser et al. 2022).

4.4.1 Observer Synthesis for Fixed μ

In this subsection, we present a LMI based formulation for the synthesis of an observer with a fixed parameter μ. The assignment of a specific value to μ allows the delineation of the observer as either a continuous-time or discrete-time variant. This foundational framework is instrumental in the development of the LPV observer.

The parameter μ in (4.23) is treated as a constant entity. Prior to engaging in the resolution of the LMI, it is imperative to fix the value of μ, which results in a constant observer gain L, as opposed to a variable one.

Lemma 5. *Given a system characterized by* (4.20)*, it can be demonstrated that for a fixed μ, there exists a matrix L which ensures the asymptotic stability of the observer delineated by*

$$\hat{x}^{\Delta} = A(\mu)\hat{x} + B(\mu)u - L(y - C\hat{x}) \tag{4.23}$$

provided there exists a positive definite matrix P and a matrix Z of appropriate dimensions that satisfy the following LMI:

$$\begin{bmatrix} PA + A^{\top}P - ZC - C^{\top}Z^{\top} & (PA - ZC) \\ \star & -\mu^{-1}P \end{bmatrix} \prec 0. \tag{4.24}$$

Here, the observer gain is derived as $L = P^{-1}Z$.

Proof. Let us define the estimation error e as

$$e = \hat{x} - x, \tag{4.25}$$

where its delta derivative, representing the dynamics of the error, is expressed as

$$e^{\Delta} = (A - LC)e. \tag{4.26}$$

Substituting the error dynamics into (4.21), we derive

$$\begin{aligned} V^{\Delta} &= P(A - LC) + (A - LC)^{\top} P \\ &\quad + \mu(A - LC)^{\top} P(A - LC) \\ &= P(A - LC) + (A - LC)^{\top} P \\ &\quad + \mu[(A - LC)^{\top} P P^{-1} P(A - LC)]. \end{aligned} \tag{4.27}$$

Utilizing the Schur complement, the above equation can be equivalently represented as

$$V^{\Delta} \equiv \begin{bmatrix} P(A - LC) + (A - LC)^{\top} P & P(A - LC) \\ \star & -\mu^{-1} P \end{bmatrix}. \tag{4.28}$$

Further expansion and substitution yield

$$V^{\Delta} \equiv \begin{bmatrix} PA + A^{\top} P - ZC - C^{\top} Z^{\top} & PA - ZC \\ \star & -\mu^{-1} P \end{bmatrix}. \tag{4.29}$$

□

It should be noted that in the scenario where $\mu = 0$ (the continuous-time case), the LMI simplifies to $PA + A^{\top} P - ZC - C^{\top} Z^{\top} \prec 0$ as the employment of the Schur complement becomes unnecessary. See equation (4.27) for a clear indication that this is true.

Given that the time-scale system exhibits parameter variability, the design of an observer with a fixed μ necessitates the re-solving of the LMI upon every alteration in μ. To augment computational efficiency and practical applicability, this study introduces an LPV variant of the observer. In this modified framework, the LMI requires resolution only once, and the resultant observer gain $L(\mu)$ evolves as a function of μ, facilitating real-time computation.

4.4.2 Sector Decomposition: A quasi-LPV Approach

This subsection elucidates a well-established yet possibly unfamiliar result in the existing literature pertaining to the sector decomposition of matrix exponential functions.

The governing equation of the time-scale system, as delineated in (4.11), exhibits intrinsic non-linear characteristics. To facilitate its transformation into a polytopic representation, this study proposes the application of a sector decomposition technique, inspired by the Takagi-Sugeno system modeling approach. This method has its roots in quasi-LPV systems (Tanaka and H. O. Wang 2001).

Employing the sector decomposition methodology allows for the expression of a non-linear function as a convex polytope within a specified interval, a concept elucidated in Lemma 6 (Tanaka and H. O. Wang 2001):

Lemma 6. *Consider a function $f(\rho) : \mathbb{R} \to \mathbb{R}$ with the parameter ρ confined within the interval $[\underline{\rho}, \bar{\rho}]$. It can be represented as*

$$f(\rho) = \alpha_0(\rho)\underline{f} + \alpha_1(\rho)\bar{f}, \tag{4.30}$$

where

$$\bar{f} = \sup f(\rho), \tag{4.31}$$

$$\underline{f} = \inf f(\rho), \tag{4.32}$$

$$\alpha_0(\rho) = \frac{\bar{f} - f(\rho)}{\bar{f} - \underline{f}}, \tag{4.33}$$

$$\alpha_1(\rho) = \frac{f(\rho) - \underline{f}}{\bar{f} - \underline{f}}. \tag{4.34}$$

Proof. Refer to (Tanaka and H. O. Wang 2001) for the proof. □

In essence, the function $f(\rho)$ is reformulated as a simplex, with the membership functions α_i mapping the bounded input ρ to the bounded output range $[\underline{f}, \bar{f}]$.

However, a direct application of Lemma 6 to Equation (4.24) is not straightforward, due to its scalar function specificity, whereas the term $e^{\mathcal{A}\mu} - I$ in (4.12) is a matrix. Typically, sector decomposition is applied to each non-linearity within the matrix. In this particular case, though, explicitly solving for the algebraic form of the matrix exponential entails significant computational overhead, even for modestly sized systems. Furthermore, the non-linearities in each matrix element would result in an exponential increase in the number of vertices of the polytope with the system size.

Nevertheless, in this unique scenario, where the non-linear function is a matrix exponential, the subsequent lemma is applicable (Kraus 1936):

Lemma 7. *For a matrix $A \in \mathbb{R}^{n\times n}$, the exponential e^A is convex provided that A is self-adjoint.*

Proof. A self-adjoint matrix A is characterized by exclusively real-valued eigenvalues and is thus diagonalizable. It can be represented as

$$A = VDV^{-1}, \tag{4.35}$$

with D being a diagonal matrix with real entries, and V being the matrix effecting the similarity transformation. Consequently,

$$e^A = e^{VDV^{-1}} = Ve^DV^{-1}, \tag{4.36}$$

where e^D is convex. The preservation of convexity through similarity transformation thus ensures the convexity of e^A. □

Lemma 7 establishes that for a self-adjoint matrix A, e^A is convex, leading to the following corollary:

Corollary 2.1. *For a self-adjoint matrix $\mathcal{A}$, the following inequality holds:*

$$e^{\sum_i \alpha_i \mathcal{A}} \preceq \sum_i \alpha_i e^{\mathcal{A}_i}. \tag{4.37}$$

The convexity of a function and the satisfaction of the Jensen inequality (as delineated in (4.37)) are concomitant, as detailed in the works of Bernstein (2009), Malamud (2001), Antezana, Massey, and Stojanoff (2004), and Kraus (1936). The pre-established convexity of the function thus validates the inequality.

Now that an observer has been established for a fixed μ value, the subsequent subsection will expand this concept to accommodate a variable μ value within a specified interval.

4.4.3 Linear Parameter-Varying Observer Synthesis

This subsection delineates the principal result of this chapter, Theorem 3, which extends and refines the concepts presented in Lemma 5. The focal point is the development of an LPV observer for the time-scale system, as characterized in Lemma 3.

Theorem 3. *In consideration of the time-scale system shown in Lemma 3, with a self-adjoint $\mathcal{A}$ matrix, then there exists an asymptotically stable LPV observer for μ within the interval $[\underline{\mu}, \bar{\mu}]$ for the observer described by:*

$$\hat{x}^{\Delta} = A(\mu)\hat{x} + B(\mu)u - L(\mu)(y - C\hat{x}), \tag{4.38}$$

where

$$L(\mu) = \sum_{i=0}^{3} \alpha_i(\mu) L_i, \tag{4.39}$$

$$\alpha_0(\mu) = \frac{\underline{\mu}^{-1} - \mu^{-1}}{\underline{\mu}^{-1} - \bar{\mu}^{-1}} \frac{\bar{\mu} - \mu}{\bar{\mu} - \underline{\mu}}, \tag{4.40}$$

$$\alpha_1(\mu) = \frac{\mu^{-1} - \bar{\mu}^{-1}}{\underline{\mu}^{-1} - \bar{\mu}^{-1}} \frac{\bar{\mu} - \mu}{\bar{\mu} - \underline{\mu}}, \tag{4.41}$$

$$\alpha_2(\mu) = \frac{\underline{\mu}^{-1} - \mu^{-1}}{\underline{\mu}^{-1} - \bar{\mu}^{-1}} \frac{\mu - \underline{\mu}}{\bar{\mu} - \underline{\mu}}, \tag{4.42}$$

$$\alpha_3(\mu) = \frac{\mu^{-1} - \bar{\mu}^{-1}}{\underline{\mu}^{-1} - \bar{\mu}^{-1}} \frac{\mu - \underline{\mu}}{\bar{\mu} - \underline{\mu}}, \tag{4.43}$$

provided there exist a positive definite matrix P and matrices Z_i satisfying the following LMI:

$$\begin{bmatrix} \Psi_0(\bar{\mu}, \bar{\mu}^{-1}) + \Psi_0(\bar{\mu}, \bar{\mu}^{-1})^{\top} & \Psi_0(\bar{\mu}, \bar{\mu}^{-1}) \\ \Psi_0(\bar{\mu}, \bar{\mu}^{-1})^{\top} & -\bar{\mu}^{-1} P \end{bmatrix} \prec 0, \tag{4.44}$$

$$\begin{bmatrix} \Psi_1(\underline{\mu}, \bar{\mu}^{-1}) + \Psi_1(\underline{\mu}, \bar{\mu}^{-1})^{\top} & \Psi_1(\underline{\mu}, \bar{\mu}^{-1}) \\ \Psi_1(\underline{\mu}, \bar{\mu}^{-1})^{\top} & -\bar{\mu}^{-1} P \end{bmatrix} \prec 0, \tag{4.45}$$

$$\begin{bmatrix} \Psi_2(\bar{\mu}, \underline{\mu}^{-1}) + \Psi_2(\bar{\mu}, \underline{\mu}^{-1})^{\top} & \Psi_2(\bar{\mu}, \underline{\mu}^{-1}) \\ \Psi_2(\bar{\mu}, \underline{\mu}^{-1})^{\top} & -\underline{\mu}^{-1} P \end{bmatrix} \prec 0, \tag{4.46}$$

$$\begin{bmatrix} \Psi_3(\underline{\mu}, \underline{\mu}^{-1}) + \Psi_3(\underline{\mu}, \underline{\mu}^{-1})^{\top} & \Psi_3(\underline{\mu}, \underline{\mu}^{-1}) \\ \Psi_3(\underline{\mu}, \underline{\mu}^{-1})^{\top} & -\underline{\mu}^{-1} P \end{bmatrix} \prec 0, \tag{4.47}$$

where

$$\Psi_i(\lambda, \omega) = P\omega(e^{\mathcal{A}\lambda} - I) - Z_i C \tag{4.48}$$

$$L_i = P^{-1} Z_i. \tag{4.49}$$

Proof. Starting with (4.29), we define μ and μ^{-1} as non-linearities to facilitate the application of Corollary 2.1 to matrix A. Although μ is linear, this definition aids in delineating its pertinence function:

$$A = \frac{e^{\mathcal{A}\mu} - I}{\mu} \preceq \sum_{i=0}^{1} \sum_{j=0}^{1} \beta_i \gamma_j \mu_i^{-1} (e^{A\mu_j} - I). \tag{4.50}$$

Here, β_i and γ_j conform to the definitions outlined in Lemma 6.

Subsequently, $\Psi_{i,j}$ is defined as:

$$\Psi_{i,j} = P\mu_i^{-1}(e^{\mathcal{A}\mu_j} - I) - Z_{i+2j}C. \tag{4.51}$$

Utilizing this definition, V^{Δ} can be represented as:

$$V^{\Delta} \preceq \sum_{i=0}^{1}\sum_{j=0}^{1} \beta_i\gamma_j \begin{bmatrix} \Psi_{i,j} + \Psi_{i,j}^{\top} & \Psi_{i,j} \\ \star & \mu_i^{-1}P \end{bmatrix} \tag{4.52}$$

By defining $\alpha_k = \beta_i\gamma_j$ with $k = i + 2j$ (resulting in the pertinence functions elucidated in (4.39)) and substituting the simplex with its vertices, we derive the expressions encompassed in Equations (4.44) to (4.47). □

In summation, Section 4.4.3 introduces an LPV observer synthesis for time-scale systems. It underlines the criticality of the self-adjoint nature of matrix $\mathcal{A}$ for the effective estimation of system states across the specified μ range. The theorem assures the asymptotic stability of the observer via the satisfaction of the stipulated LMI, thereby guaranteeing precise state estimation in time-scale systems. It is noteworthy that the constraint on the matrix dynamics in the continuous-time domain might appear restrictive; however, any system with only real-valued eigenvalues can undergo a similarity transformation to become self-adjoint.

The forthcoming section introduces a simulation designed to illustrate the detection technique.

4.5 Simulation Example

This section presents the simulation validation of the proposed detection technique through a series of simulations conducted on a continuous-time system. These simulations aim to demonstrate the efficacy of the technique in detecting system intrusions.

For the purpose of these simulations, we consider a continuous-time system as delineated by (4.10), characterized by the following matrix parameters:

$$\mathcal{A} = \begin{bmatrix} -1 & -1 & 0 & 0 & 0 \\ -1 & -1 & 0 & 0 & 0 \\ 0 & 0 & -2 & 0 & 0 \\ 0 & 0 & 0 & -3 & 0 \\ 0 & 0 & 0 & 0 & -5 \end{bmatrix}, \tag{4.53}$$

$$\mathcal{B} = \begin{bmatrix} -7.318 \\ 0.563 \\ 20.78 \\ 21.33 \\ 8.053 \end{bmatrix}, \tag{4.54}$$

$$C = \begin{bmatrix} -1.38 & 0.69 & -1.59 & 1.24 & -0.48 \end{bmatrix}, \tag{4.55}$$

$$D = 0. \tag{4.56}$$

Figure 4.4 illustrates the system's behavior under standard operating conditions and during an attack. The system possesses a zero at $s_0 = 1$, oriented in the direction of

$$z_0 = \begin{bmatrix} 2.81 \\ -1.69 \\ -6.93 \\ -5.33 \\ -1.34 \end{bmatrix}. \tag{4.57}$$

The simulation results reveal that the system's output remains constant in both scenarios, despite the divergence in states during the attack.

Subsequently, a time-scale system was developed using (4.11) and an observer was synthesized based on the methodologies described in Section 4.4.3, setting $\mu = 1$. The same simulation was then conducted on the observer, as depicted in Figure 4.5. A notable difference is that the attack initiation is delayed until $10\,\mathrm{s}$. The divergence in the state is apparent irrespective of the value of μ, highlighting the presence of an attack.

As demonstrated in Figure 4.5, the observer exhibits sensitivity to zero-dynamics attacks, manifesting divergence in both output and states during the attack. This deviation can be effectively detected using methods such as

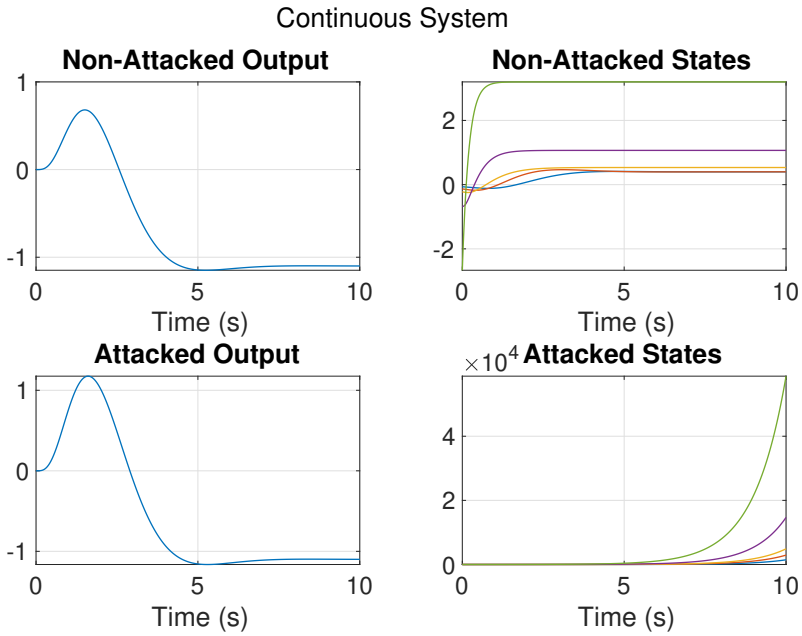


Figure 4.4 – Simulation of continuous-time system with and without an attack.

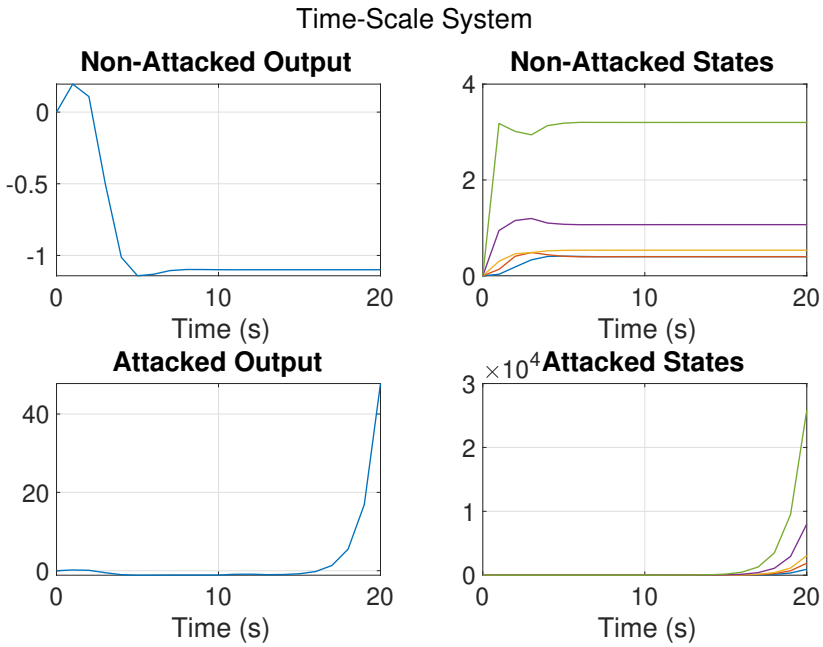


Figure 4.5 – Simulation of time-scale system's observer with and without an attack.

Bayes classification, as suggested by (Isermann 2006), to trigger an alarm system.

4.6 Conclusion

This chapter has delved into the intricate realm of detection techniques against ZDA, presenting a novel approach for its detection. The landscape of current methodologies primarily bifurcates into two streams: generalized holders (Jihan Kim and Shim 2019b; D. Kim, Ryu, and Back 2021), characterized by their straightforward implementation yet vulnerability to new attack vectors, and topology reconfiguration strategies (Mao, Jafarnejadsani, et al. 2020), which, despite their efficacy in attack detection, are hampered by the complexity of control reconfiguration.

The implementation of the proposed technique does not necessitate the random alteration of the parameter μ, a choice that streamlines the operational aspect. However, introducing randomness in the value of μ can further fortify the detection system, particularly against sophisticated attacks that simultaneously target both continuous- and time-scale systems. The integration of game-theoretical approaches, such as the Moving Target Technique, can optimize this randomness, striking a balance between the frequency of changes in μ and the maximization of attack detection probability (Umsonst et al. 2022).

To the best of the authors' knowledge, the application of time-scale systems in addressing ZDA is an unexplored territory. This novel approach opens a new chapter in the field, offering a unique perspective and methodology. It is important to note that while the comparison of observers' response times across different systems may not be particularly revealing due to the tunability of their dynamics, the proposed technique stands out in its inherent flexibility and adaptability.

The research presented in this chapter has introduced a pioneering technique for detecting zero-dynamics attacks using time-scale systems. This technique not only enhances the detection of attacks but also significantly raises the bar for attackers to maintain stealth. By eliminating the need for signal transmission to the control side for processing, the proposed system effectively closes a critical vulnerability. Additionally, the introduction of a variable dynamic in the observer, through the potential alteration of μ, necessitates a higher degree of system knowledge for an attacker, thereby increasing the complexity and risk of detection for any attempted intrusion.

In conclusion, this chapter has contributed a novel approach to the detection of zero-dynamics attacks, providing a robust, flexible, and efficient solution. The implications of this research extend far beyond its immediate

application, paving the way for future innovations in secure control system design and implementation. As the landscape of cyber-physical systems continues to evolve and expand, the methodologies and principles explored in this thesis will serve as a foundation for the development of increasingly sophisticated and resilient security measures.

In the forthcoming chapter, attention is redirected towards an alternative form of cyber threat: the Denial-of-Service (DoS) attack. The application of Time-Scale Theory will be reiterated, facilitating the development of an observer and controller designed to ameliorate the impact of such attacks on the system.

Chapter 5

Denial of Service attack mitigation using time-scale

5.1 Introduction

In the preceding chapter, Time-Scale Theory was employed to design an observer specifically for Zero-Dynamics Attack (ZDA) detection. This chapter shifts emphasis towards Denial-of-Service (DoS) attacks, applying Time-Scale Theory once more to formulate both a controller and an observer that exhibit resilience against such attacks.

The dependency of Cyber-Physical Systems (CPS) on communication networks engenders a spectrum of security vulnerabilities. Particularly disconcerting among these threats are DoS attacks, owing to their capacity to disrupt the communication networks of CPS. By obstructing the transmission of control and measurement data packets, DoS attacks directly impinge on the functionality of CPS.

These assaults can be orchestrated via diverse methodologies such as communication channel jamming, device compromise to impede data transmission, and inundating the network with excessive traffic (Cetinkaya, Ishii, and Hayakawa 2019; Huang et al. 2009). The severity of DoS attacks varies, encompassing relatively innocuous attacks causing minor delays and packet losses to severe assaults capable of completely incapacitating communication networks (Pang et al. 2019). A literature review of DoS attacks was given in Section 2.3.3.

This chapter proposes an innovative solution based on Time-Scale Theory. Utilizing the continuous to Time-Scale transformation delineated in Chapter 4, a controller and observer for a system are devised to ensure the

stability of the closed-loop system in the presence of DoS attacks, treating the attack as a delay. The possibility of varying the sampling time of the controller and observer to any value within an interval allows them to adjust themselves to provide accurate values upon the receipt of a packet. As long as the data arrives within a delay shorter than the maximum delay threshold, the system is pratically immune to weak DoS attacks.

In this chapter we provide a comprehensive overview of DoS attacks, establishing a foundational understanding. Then we introduce a design methodology for developing a controller and observer with inherent resilience to DoS attacks. Afterwards we demonstrate the effectiveness of the proposed technique through its application to a TurtleBot3 robot.

The subsequent section offers a concise review of DoS attacks.

5.2 Denial of Service Attack

DoS attacks manifest in diverse forms, with the Volume-Based Attack being predominant. These attacks primarily aim to inundate the bandwidth of the target site by generating copious amounts of traffic, including web page requests, emails, and spurious packets. Such an onslaught overburdens the target, leading to significant slowdowns or complete obstruction of legitimate traffic. A prevalent tactic involves the deployment of botnets—networks of compromised computers or devices, remotely manipulated to dispatch an overwhelming volume of traffic towards the target.

Additionally, there are Protocol Attacks, strategically designed to exhaust server resources or those of intermediate network components, such as firewalls and load balancers. These attacks exploit vulnerabilities within the layer 3 and layer 4 protocol stacks, typically by dispatching packets that necessitate partial processing by the target system, culminating in system overload. A notable example is the SYN flood attack, where the assailant initiates a TCP connection but abstains from completing the handshake, resulting in an accumulation of half-open connections.

The Application Layer Attacks represent a more sophisticated breed, targeting specific functionalities of an application or service. These attacks are executed by sending seemingly legitimate, yet intricately crafted requests that exploit vulnerabilities or inefficiencies in the application's code, leading to crashes or significant slowdowns. Tactics may include bombarding a database with complex queries or leveraging inefficient code segments.

From the perspective of system automation, a DoS attack poses signif-

icant risks to system stability. This is primarily because automated systems rely on the prompt and reliable delivery of information to modify and maintain the system's state. Unfortunately, these systems are generally ill-equipped to cope with interruptions in data flow, leading to difficulties in adapting to absent information. In the context of continuous-time systems, this interruption means that, over a certain time interval t, the system's measurements $y(t)$ will not be available.

Similarly, discrete-time systems experience interruptions in data transmission, resulting in some values of $y(k)$ not being transmitted. Consequently, this scenario can be likened to a system operating with variable sampling-time, characterized by the unpredictable arrival times of information. Such variability complicates the system's ability to process and respond to incoming data effectively, thereby undermining the system's overall stability and performance.

In the present chapter, the DoS attack is perceived as a variable sampling-time. This perspective holds particular pertinence for robotic systems which depend on networked sensor or reference signals. This is notably applicable in industrial settings, where robots may rely on networked tracking systems to acquire positional data. This configuration often involves a network connection that is susceptible to interference or attack, particularly in the segment linking the controller/observer and the external sensor or reference generator.

The forthcoming section details the design of the controller and observer utilizing Time-Scale Theory to mitigate DoS attacks.

5.3 Controller and Observer Synthesis

In order to synthesize a controller that maintains functionality under weak DoS attacks, this research adopts the time-scale theory outlined in Chapter 4. A controller based on this theory would modify its operational time domain in response to the attack, thereby generating control signals in accordance with the passage of packets. This approach enables system stability to be preserved, albeit with a reduction in performance, until a predetermined maximum time-delay threshold is reached. The schematic of the control system is depicted in Figure 5.1. There is no network between the system and external sensors, so it is not considered attack-able. The network between the sensors and observer, on the hand, is susceptible to DoS attacks. To mitigate it, both the observer and controller must be able

to change their sampling-time to accommodate for the loss of information.

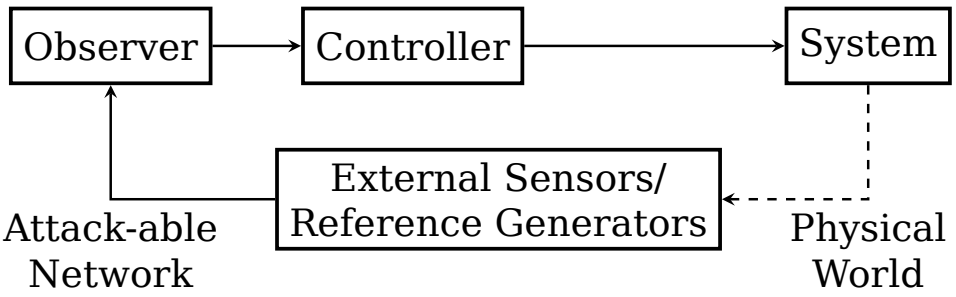


Figure 5.1 – System schematic for DoS attack.

The design process considers the independence between the controller and the observer. Initially, the controller is formulated. Subsequently, the observer is constructed with consideration of the previously designed controller to ensure that it operates sufficiently rapidly to prevent destabilization of the controller.

As in the preceding chapter, the design is segmented into two distinct sections. Initially, the synthesis is conducted with a constant value of μ. Subsequently, the approach is extended to accommodate a variable μ.

5.3.1 Synthesis for a Fixed μ

The mathematical model for the system under study is expressed as follows:

$$x^{\Delta} = Ax + Bu, \tag{5.1}$$

$$y = Cx, \tag{5.2}$$

$$u = -K\hat{x}, \tag{5.3}$$

where x^{Δ} denotes the change in the state vector x, A and B represent the system matrices, u is the control input, y is the output, and K is the controller gain matrix. The observer, which estimates the state vector $\hat{x}$, is described by the equation

$$\hat{x}^{\Delta} = A\hat{x} + Bu + L(y - \hat{y}). \tag{5.4}$$

Here, $\hat{x}^{\Delta}$ signifies the estimated state vector's change, L is the observer gain matrix, and $\hat{y}$ represents the estimated output.

Lemma 8 establishes the Lyapunov condition to design the controller and observer for a fixed value of μ.

Lemma 8. *Considering a system governed by the dynamics presented in* (5.1) *and assuming the system is both observable and controllable. Then, for a constant value of* μ*, the existence of a solution to the following LMI guarantees the Hilger stability criteria of the closed-loop controlled system with estimated states, as well as calculates the values of the controller's and observer's gains:*

$$AP_1 - BZ_1 = \Omega_1, \tag{5.5}$$

$$\begin{bmatrix} \Omega_1 + \Omega_1^\top & \Omega_1 \\ \Omega_1^\top & -\mu^{-1}P_1 \end{bmatrix} \prec 0, \tag{5.6}$$

$$P_1 \succ 0, \tag{5.7}$$

This LMI serves as the basis for deriving a controller on a designated time scale. Subsequently, an observer can be designed by employing the following LMI:

$$P_2A - Z_2C = \Omega_2, \tag{5.8}$$

$$\begin{bmatrix} \Omega_1 + \Omega_1^\top & \Omega_1 & P_1BK & P_1BK \\ \Omega_1^\top & -\mu^{-1}P_1 & 0 & 0 \\ K^\top B^\top P_1 & 0 & \Omega_2 + \Omega_2^\top & \Omega_2 \\ K^\top B^\top P_1 & 0 & \Omega_2^\top & -\mu^{-1}P_2 \end{bmatrix} \prec 0, \tag{5.9}$$

$$P_2 \succ 0 \tag{5.10}$$

where A*,* B *and* C *are system matrices,* P_1 *and* P_2 *are positive definite matrices,* K *is the controller gain,* L *is the observer gain, and* Z_1 *and* Z_2 *are full variable matrices.*

The gains are recovered as $K = Z_1P_1^{-1}$ *and* $L = P_2^{-1}Z_2$*.*

Proof. The estimation error for the controlled system is defined as:

$$e = x - \hat{x}, \tag{5.11}$$

where e represents the estimation error, x is the actual state, and $\hat{x}$ denotes the estimated state. The temporal derivative of the estimation error is given by:

$$
\begin{aligned}
e^{\Delta} &= \frac{e(\sigma) - e}{\mu} && (5.12)\\
&= \frac{x(\sigma) - \hat{x}(\sigma) - (x - \hat{x})}{\mu} && (5.13)\\
&= (x^{\Delta} - \hat{x}^{\Delta}) && (5.14)\\
&= Ax - BK\hat{x} - A\hat{x} + BK\hat{x} - L(Cx - C\hat{x}) && (5.15)\\
&= A(x - \hat{x}) - LC(x - \hat{x}) && (5.16)\\
&= (A - LC)e. && (5.17)
\end{aligned}
$$

The dynamics of the system, incorporating the estimation error, can be reformulated as:

$$
\begin{aligned}
x^{\Delta} &= Ax - BK\hat{x} && (5.18)\\
&= Ax - BK(x - e) && (5.19)\\
&= (A - BK)x + BKe. && (5.20)
\end{aligned}
$$

Combining the dynamics of both the controller and the observer, the augmented system can be described by the following matrix equation:

$$
\underbrace{\begin{bmatrix} x^{\Delta} \\ e^{\Delta} \end{bmatrix}}_{z^{\Delta}} = \underbrace{\begin{bmatrix} A - BK & BK \\ 0 & A - LC \end{bmatrix}}_{\Phi} \underbrace{\begin{bmatrix} x \\ e \end{bmatrix}}_{z}, \tag{5.21}
$$

which represents the combined evolution of the system state and the estimation error over time.

For the Lyapunov candidate function, defined as:

$$
V = z^{\top} P z, \tag{5.22}
$$

where V represents the Lyapunov function and P is a positive definite matrix, the subsequent condition ensures the stability of both the controller and observer for a specified value of μ:

$$V^{\Delta} = \frac{z(\sigma)^{\top} P z(\sigma) - z^{\top} P z}{\mu} \tag{5.23}$$

$$= z^{\top} P(\mu z^{\Delta}) + (\mu z^{\Delta})^{\top} P z + \mu^2 (z^{\Delta})^{\top} P z^{\Delta} \tag{5.24}$$

$$= z^{\top} P z^{\Delta} + (z^{\Delta})^{\top} P z + \mu (z^{\Delta})^{\top} P z^{\Delta} \tag{5.25}$$

$$= P\Phi + \Phi^{\top} P + \mu \Phi^{\top} P \Phi \tag{5.26}$$

$$\equiv \begin{bmatrix} P\Phi + \Phi^{\top} P & P\Phi \\ \Phi^{\top} P & -\frac{1}{\mu} P \end{bmatrix} \prec 0, \tag{5.27}$$

$$P \succ 0. \tag{5.28}$$

The LMI formulation requires a similarity transformation and substitution of variables with $Z_1 = KP$ and $Z_2 = PL$. It is advisable to select P in the form of $\begin{bmatrix} P_1 & 0 \\ 0 & P_2 \end{bmatrix}$, to reduce the number of bilinear elements within the resulting matrix. Consequently, this leads to:

$$\Omega_1 = AP_1 - BZ_1, \tag{5.29}$$

$$\Omega_2 = P_2 A - Z_2 C, \tag{5.30}$$

$$V^{\Delta} = \begin{bmatrix} \Omega_1 + \Omega_1^{\top} & \Omega_1 & P_1 BK & P_1 BK \\ \Omega_1^{\top} & \mu^{-1} P_1 & 0 & 0 \\ K^{\top} B^{\top} P_1 & 0 & \Omega_2 + \Omega_2^{\top} & \Omega_2 \\ K^{\top} B^{\top} P_1 & 0 & \Omega_2^{\top} & \mu^{-1} P_2 \end{bmatrix}. \tag{5.31}$$

Although this matrix includes unsubstituted variables and bilinear terms, addressing the controller portion of the LMI first by solving

$$\begin{bmatrix} \Omega_1 + \Omega_1^{\top} & \Omega_1 \\ \Omega_1^{\top} & \mu^{-1} P_1 \end{bmatrix} \prec 0 \tag{5.32}$$

and subsequently substituting constants for P_1 and K in the derived system, allows for the resolution of the overall LMI. This results in an observer configuration that maintains the stability of the closed-loop system under control. □

The upcoming subsection extends the design from a fixed μ value to a variable μ over a specified interval.

5.3.2 Synthesis for a Variable μ

This subsection delineates the principal result of this chapter, Theorem 4, which extends and refines the concepts presented in Lemma 8.

Theorem 4. *In consideration of the time-scale system explicated in Lemma 3 for a system described as in* (5.1) *and a Luenberguer observer as in* (5.4)*, and assuming that the matrix $\mathcal{A}$ is self-adjoint, there exists an asymptotically stable Linear Parameter Varying (LPV) controller and observer for μ within the interval $[\underline{\mu}, \bar{\mu}]$ if there exist positive definite matrices P_i and matrices $Z_{i,j}$ satisfying the following LMIs:*

Firstly, for the controller:

$$A_i P_1 - B_i Z_{1,i} = \Omega_{1,i}, \tag{5.33}$$

$$\begin{bmatrix} \Omega_{1,1} + \Omega_{1,1}^\top & \Omega_{1,1} \\ \Omega_{1,1}^\top & \underline{\mu}^{-1} P_1 \end{bmatrix} \prec 0, \tag{5.34}$$

$$\begin{bmatrix} \Omega_{1,2} + \Omega_{1,2}^\top & \Omega_{1,2} \\ \Omega_{1,2}^\top & \bar{\mu}^{-1} P_1 \end{bmatrix} \prec 0 \tag{5.35}$$

$$\begin{bmatrix} \Omega_{1,3} + \Omega_{1,3}^\top & \Omega_{1,3} \\ \Omega_{1,3}^\top & \underline{\mu}^{-1} P_1 \end{bmatrix} \prec 0 \tag{5.36}$$

$$\begin{bmatrix} \Omega_{1,4} + \Omega_{1,4}^\top & \Omega_{1,4} \\ \Omega_{1,4}^\top & \bar{\mu}^{-1} P_1 \end{bmatrix} \prec 0 \tag{5.37}$$

$$P_1 \succ 0, \tag{5.38}$$

Then, for the observer:

$$P_2 A_i - Z_{2,i} C_i = \Omega_{2,i}, \tag{5.39}$$

$$\begin{bmatrix} \Omega_{1,1} + \Omega_{1,1}^\top & \Omega_{1,1} & P_1 B_1 K_1 & P_1 B_1 K_1 \\ \Omega_{1,1}^\top & \underline{\mu}^{-1} P_1 & 0 & 0 \\ K_1^\top B_1^\top P_1 & 0 & \Omega_{2,1} + \Omega_{2,1}^\top & \Omega_{2,1} \\ K_1^\top B_1^\top P_1 & 0 & \Omega_{2,1}^\top & \underline{\mu}^{-1} P_2 \end{bmatrix} \prec 0, \tag{5.40}$$

$$\tag{5.41}$$

$$\begin{bmatrix} \Omega_{1,2}+\Omega_{1,2}^\top & \Omega_{1,2} & P_1B_2K_2 & P_1B_2K_2 \\ \Omega_{1,2}^\top & \bar{\mu}^{-1}P_1 & 0 & 0 \\ K_2^\top B_2^\top P_1 & 0 & \Omega_{2,2}+\Omega_{2,2}^\top & \Omega_{2,2} \\ K_2^\top B_2^\top P_1 & 0 & \Omega_{2,2}^\top & \bar{\mu}^{-1}P_2 \end{bmatrix} \prec 0, \tag{5.42}$$

$$\begin{bmatrix} \Omega_{1,3}+\Omega_{1,3}^\top & \Omega_{1,3} & P_1B_3K_3 & P_1B_3K_3 \\ \Omega_{1,3}^\top & \underline{\mu}^{-1}P_1 & 0 & 0 \\ K_3^\top B_3^\top P_1 & 0 & \Omega_{2,3}+\Omega_{2,3}^\top & \Omega_{2,3} \\ K_3^\top B_3^\top P_1 & 0 & \Omega_{2,3}^\top & \underline{\mu}^{-1}P_2 \end{bmatrix} \prec 0, \tag{5.43}$$

$$\begin{bmatrix} \Omega_{1,4}+\Omega_{1,4}^\top & \Omega_{1,4} & P_1B_4K_4 & P_1B_4K_4 \\ \Omega_{1,4}^\top & \bar{\mu}^{-1}P_1 & 0 & 0 \\ K_4^\top B_4^\top P_1 & 0 & \Omega_{2,4}+\Omega_{2,4}^\top & \Omega_{2,4} \\ K_4^\top B_4^\top P_1 & 0 & \Omega_{2,4}^\top & \bar{\mu}^{-1}P_2 \end{bmatrix} \prec 0, \tag{5.44}$$

$$P_2 \succ 0. \tag{5.45}$$

where

$$A_1 = A(\underline{\mu}), A_2 = \frac{A(\underline{\mu})\underline{\mu}}{\bar{\mu}}, A_3 = \frac{A(\bar{\mu})\bar{\mu}}{\underline{\mu}}, A_4 = A(\bar{\mu}) \tag{5.46}$$

$$B_1 = B(\underline{\mu}), B_2 = \frac{B(\underline{\mu})\underline{\mu}}{\bar{\mu}}, B_3 = \frac{B(\bar{\mu})\bar{\mu}}{\underline{\mu}}, B_4 = B(\bar{\mu}) \tag{5.47}$$

$$K_i = Z_{1,i}P^{-1}, L_i = P^{-1}Z_{2,i}, \tag{5.48}$$

$$L(\mu) = \sum_{i=0}^{3} \alpha_i(\mu)L_i, K(\mu) = \sum_{i=0}^{3} \alpha_i(\mu)K_i, \tag{5.49}$$

$$\alpha_0(\mu) = \frac{\underline{\mu}^{-1}-\mu^{-1}}{\underline{\mu}^{-1}-\bar{\mu}^{-1}}\frac{\bar{\mu}-\mu}{\bar{\mu}-\underline{\mu}}, \tag{5.50}$$

$$\alpha_1(\mu) = \frac{\mu^{-1}-\bar{\mu}^{-1}}{\underline{\mu}^{-1}-\bar{\mu}^{-1}}\frac{\bar{\mu}-\mu}{\bar{\mu}-\underline{\mu}}, \tag{5.51}$$

$$\alpha_2(\mu) = \frac{\underline{\mu}^{-1}-\mu^{-1}}{\underline{\mu}^{-1}-\bar{\mu}^{-1}}\frac{\mu-\underline{\mu}}{\bar{\mu}-\underline{\mu}}, \tag{5.52}$$

$$\alpha_3(\mu) = \frac{\mu^{-1}-\bar{\mu}^{-1}}{\underline{\mu}^{-1}-\bar{\mu}^{-1}}\frac{\mu-\underline{\mu}}{\bar{\mu}-\underline{\mu}}, \tag{5.53}$$

Proof. The proof is similar to that of Theorem 3. □

In summation, Section 5.3.2 introduces an LPV controller and observer synthesis for time-scale systems. The theorem assures the asymptotic stability of the controller and observer via the satisfaction of the stipulated

LMI, thereby guaranteeing precise state estimation and control in time-scale systems. The fact that the designed controller and observers are able to change their sampling time online in the interval $[\underline{\mu}, \bar{\mu}]$ allows for the system to adapt to DoS attacks, as shown in the next section.

5.4 Experimental Validation

This section elucidates the validation methodology employed to evaluate the efficacy of the proposed controller and observer in mitigating DoS attacks. The experimental setup involved the deployment of an autonomous mobile robot, which derived its positional data from an external optical sensor. The DoS attack was orchestrated within the communication link between the sensor and the robot, thereby obstructing the transmission of sensor readings to the robot.

This obstruction was configured such that only a single reading was allowed to transmit after a predetermined interval, with an added element of randomness to simulate a real-world attack scenario more accurately. This experimental framework was designed to test the resilience of the controller and observer against targeted disruptions in sensor data flow, critical for maintaining operational integrity in autonomous robotic systems.

The autonomous robot used was the TurtleBot3 Waffle Pi[1], shown in Figure 5.2, which is a highly versatile and customizable robot platform designed for education, research, and development in the field of robotics. It is part of the TurtleBot series, which are open-source robotic platforms that provide out-of-the-box support for various robotic applications, including autonomous navigation, object recognition, and manipulation.

Nevertheless, the majority of the aforementioned hardware capabilities and support mechanisms were not utilized, with a preference given towards our bespoke controller designs and sensing configurations. The sole feature of the robot that was employed in this study was its odometer. This instrument facilitated the verification of the robot's displacement in accordance with the control signals, thereby obviating the necessity for low-level direct current (DC) motor control. This approach enabled a more streamlined integration of our control algorithms, focusing on high-level motion accuracy and efficiency.

[1] https://www.turtlebot.com/turtlebot3

[2] Source: https://www.elektor.com/cdn/shop/files/robotis-turtlebot3-waffle-pi-incl-raspb jpg, accessed on 2024-04-08T1400

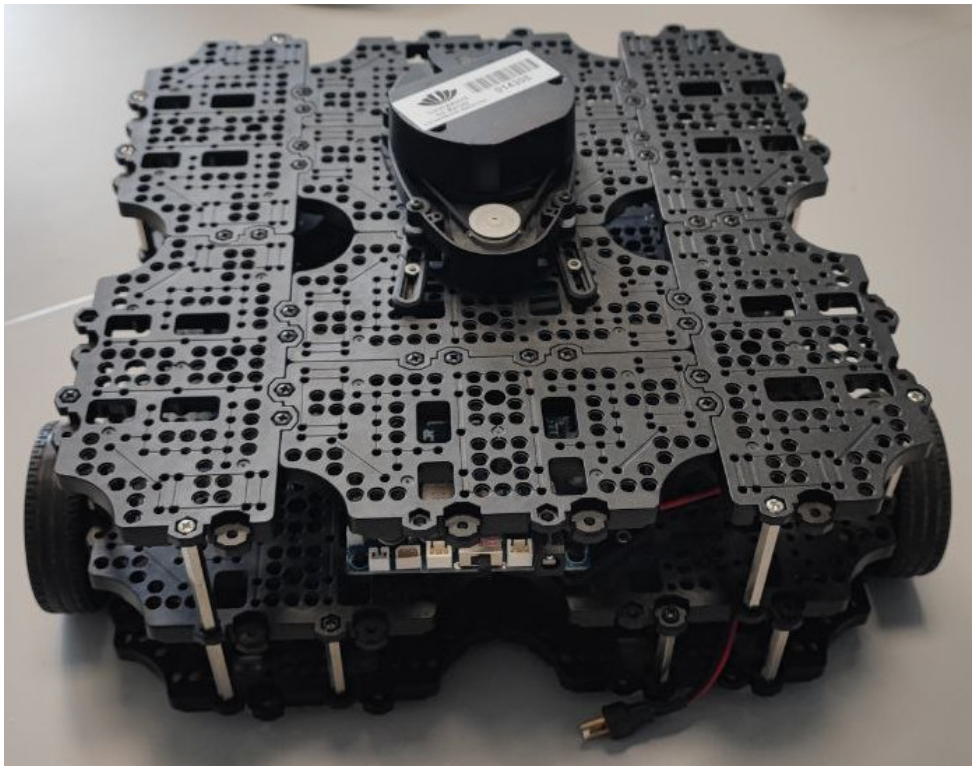


Figure 5.2 – TurtleBot3 Waffle Pi.[2]

Equipped with a Raspberry Pi as its main controller, the robot can be programmed using the Robot Operating System (ROS) in both `C++` and `Python`. ROS is an open-source, flexible framework for writing robot software. It is not an operating system in the traditional sense of managing hardware resources but rather a collection of software frameworks for robot software development. ROS provides a structured communications layer above the host operating systems of a mixed compute cluster, offering services designed for a heterogeneous computer cluster such as hardware abstraction, low-level device control, implementation of commonly-used functionality, message-passing between processes, and package management.

Designed to facilitate the development of complex and robust robotic applications, ROS enables programmers to use `C++` and `Python` to write applications, which significantly enhances the accessibility and versatility of programming robots. This is particularly beneficial in academic and research settings, where robots like the one equipped with a Raspberry Pi serve as experimental platforms for testing theories and algorithms in real-world scenarios.

ROS's architecture is modular, allowing for the reuse of code in various projects, which accelerates the development process. This modularity is achieved through the use of packages, which are collections of nodes (pro-

cesses that perform computation), libraries, and tools. Nodes communicate with one another over topics using a publisher-subscriber model, services with a request-response model, or through action servers for long-running tasks. Figure 5.3 show a graph representation of interconnected nodes and topics.

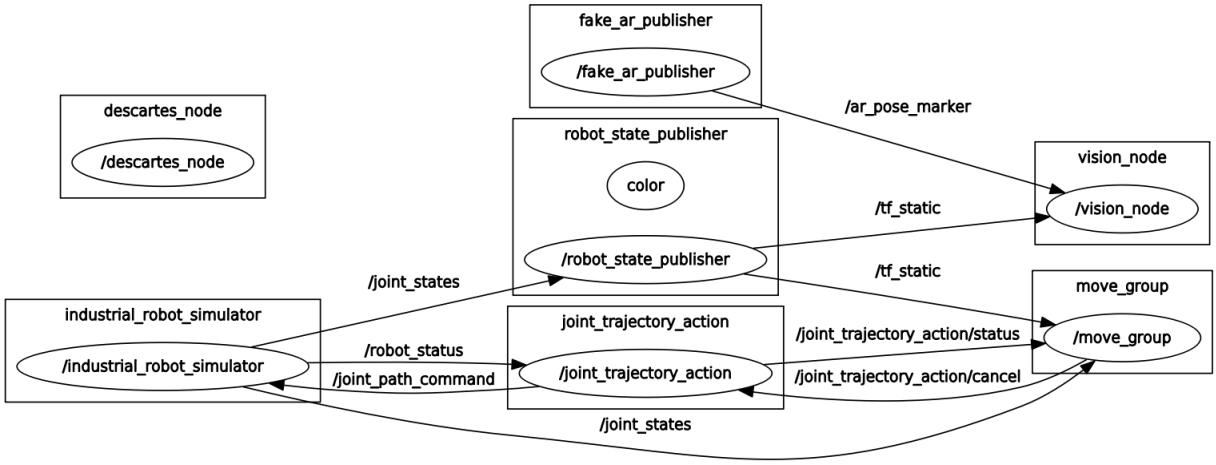


Figure 5.3 – ROS nodes and topics.

To monitor the robot's position in space, the Optitrack system was used. The Optitrack[3] system is a sophisticated motion capture solution designed to accurately track the position and orientation of objects in three-dimensional space. Utilized extensively in animation, sports science, biomechanics, robotics, and virtual reality applications, it employs a series of high-speed cameras equipped with infrared sensors to detect specially designed reflective markers or LEDs placed on the object of interest. By capturing the spatial position of these markers at high frequencies, the system can reconstruct the precise movement of the object with remarkable accuracy and minimal latency.

The Optitrack system intrinsically lacks a direct mechanism for network-based dissemination of tracking data. Nonetheless, a collaborative effort within the institution has led to the development of an intermediary system designed to bridge this gap. This system captures the positional data generated by the Optitrack system and transmits it to a Message Queuing Telemetry Transport (MQTT) server, facilitating network-wide accessibility of the data.

MQTT operates on a publish-subscribe model that is highly efficient for various applications due to its lightweight protocol design. At the heart of MQTT's communication strategy lies the concept of "topics". A topic is

[3] `https://optitrack.com/`

a `UTF-8` string that the broker uses to filter messages for each connected client. Topics are structured in a hierarchy, similar to a file-system path, using slashes (/) as separators. This hierarchical structure allows for precise and flexible routing of messages. Both MQTT and ROS topics operate on the same way, following the same philosophy, but using different underlying technologies.

When a publisher wants to send data, it publishes a message to a specific topic. For example, a temperature sensor might publish its readings to a topic named "`home/livingroom/temperature`". Subscribers, on the other hand, indicate their interest in receiving messages by subscribing to one or more topics. The MQTT broker, which is the central server facilitating the publish-subscribe communication, is responsible for distributing the messages from publishers to the subscribers based on the topic subscriptions. The ROS version 1 works the same way, with the broker being called "master", however version 2 makes use of a distributed system that does not make use of a broker.

To integrate this data stream with the TurtleBot3's operational framework, a dedicated ROS node was developed. This node serves not only as a conduit for routing the Optitrack positional data to the TurtleBot3 but also acts as a strategic point for the implementation of the DoS attack, as depicted in Figure 5.4. This effectively creates a closed-loop system, in which the robot's movement is captured by the tracking system, sent to the MQTT server to which the robot itself subscribes and can therefore use the information for decision-making related to its movement.

In the delineation of the system's inputs and outputs — namely, the robot's steering commands as inputs and the sensor readings as outputs — a dynamic model was conceptualized to interface seamlessly with the proposed technique. Consistent with prevalent conventions observed in ROS implementations for robotic platforms, the model delineates the inputs as the robot's linear and angular velocities (v and $\dot{\theta}$, respectively). Concurrently, the outputs are defined as the robot's position and orientation, with x and y representing the Cartesian coordinates of the position, and θ denoting the orientation. The state vector of the system is identical to its output vector.

The dynamic behavior of the system is characterized predominantly by integrative components, predicated on the principle that the robot remains stationary in the absence of any external inputs. The system's input-output relationship exhibits state-dependency, wherein the linear velocity is intri-

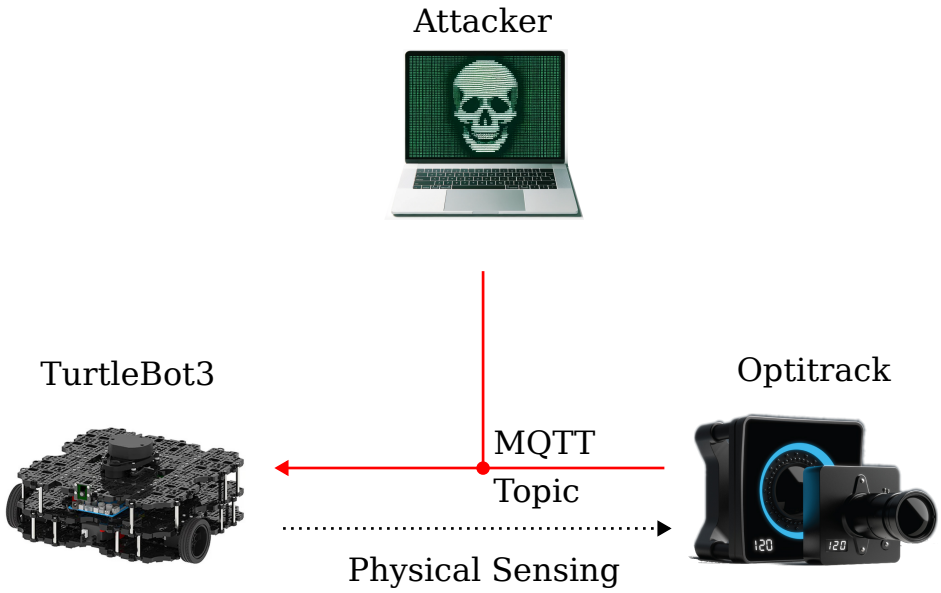


Figure 5.4 – Communication scheme of the TurtleBot3, Optitrack and the attacker. The dotted line represent a physical, non-network connection.

cately mapped to the robot's positional change as $\dot{x} = \cos(\theta)$ and $\dot{y} = \sin(\theta)$, with the angular velocity being directly correlated. This intrinsic state-dependency necessitates the adoption of a linearization process to simplify the system's dynamics for analysis and control.

To facilitate this linearization, an initial state was established at $\theta = -170$ degrees, and the process was executed through sequential increments of $10°$. Subsequently, for each linearized model, a unique pair comprising a controller and an observer was synthesized. The corresponding gains for each pair are accessible via an online repository[4]. The ROS node is programmed to select the appropriate pair for the current state of the system.

The parameters $\underline{\mu}$ and $\bar{\mu}$ were determined to be 0.1 and 2, respectively. This selection was aimed at facilitating seamless functioning under normal conditions, and enabling the execution of a significantly aggressive attack, lasting up to $2\,\mathrm{s}$, in the absence of feedback mechanisms. Subsequently, the DoS attack was executed with varying durations, specifically $0.5\,\mathrm{s}$, $1.5\,\mathrm{s}$, and $1.9\,\mathrm{s}$, to assess the system's resilience under different levels of attack intensity.

The graphical representations in Figure 5.5 delineate the position and orientation of the robotic unit, with waypoints indicated by blue dots. The

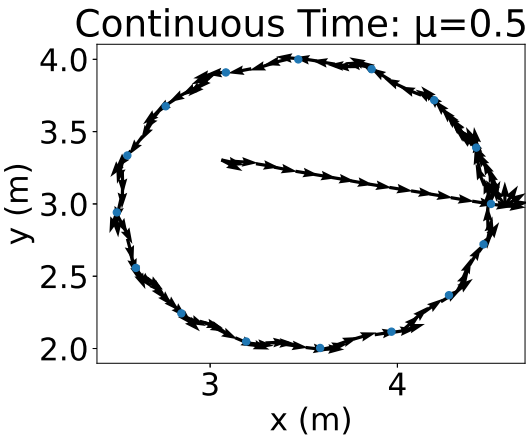
[4] `https://github.com/acristoffers/phd-turtlebot/blob/main/workspace/gains.json`

intended trajectory for the robot was to navigate through each waypoint, ultimately delineating a circular path.

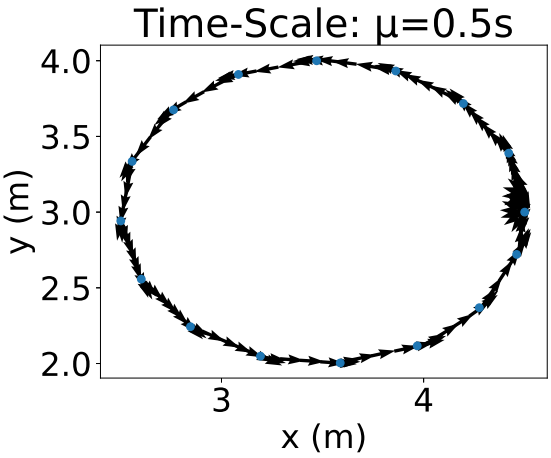
The operational protocol of the robot encompasses an initial angular realignment to direct itself towards the subsequent waypoint, followed by the commencement of linear movement in the direction of the waypoint. During this transit, the robot may undertake additional angular adjustments as required to maintain the correct heading.

Observations from the plots reveal that both the continuous-time and the time-scale controllers successfully execute the navigation task under a mild DoS attack lasting $0.5\,\mathrm{s}$ (Figure 5.5(a) and Figure 5.5(b)). However, during a more prolonged attack of $1.5\,\mathrm{s}$, the continuous-time controller (Figure 5.5(c)) fails to align with the initial waypoint due to persistent overshooting, while the time-scale controller (Figure 5.5(d)) experiences only partial trajectory performance degradation.

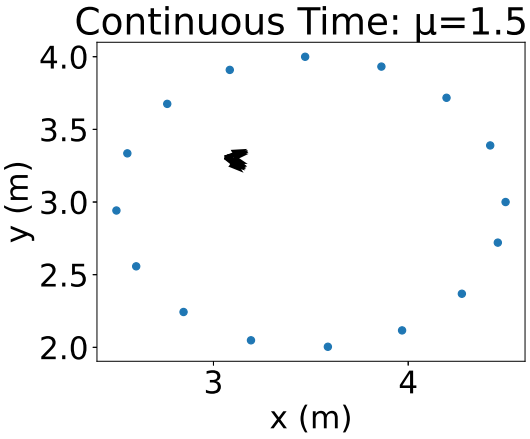
Under an attack duration of $1.9\,\mathrm{s}$, the time-scale controller (Figure 5.5(f)) maintains its capability to navigate the robot through the waypoints, albeit with reduced efficiency. Notably, the continuous-time controller (Figure 5.5(e)) exhibits instability when the attack duration extends to approximately $0.8\,\mathrm{s}$, whereas the time-scale controller remains stable up to approximately $2.3\,\mathrm{s}$, surpassing its expected operational threshold.



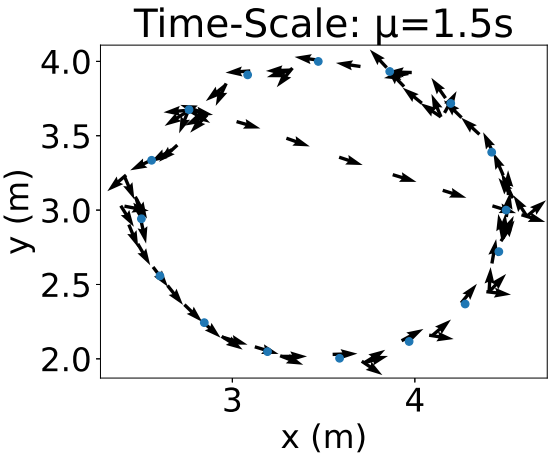
((a)) Continuous-time controller with a DoS attack of $0.5\,\mathrm{s}$



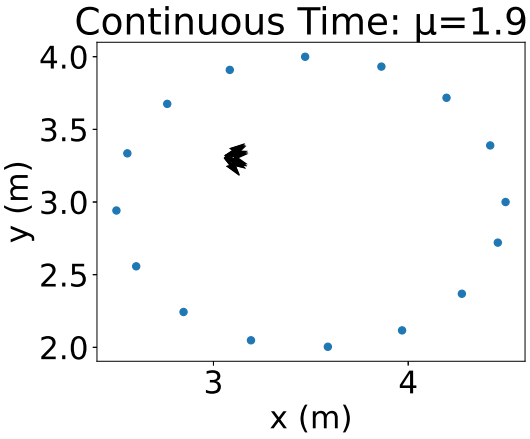
((b)) Time-Scale controller with a DoS attack of $0.5\,\mathrm{s}$



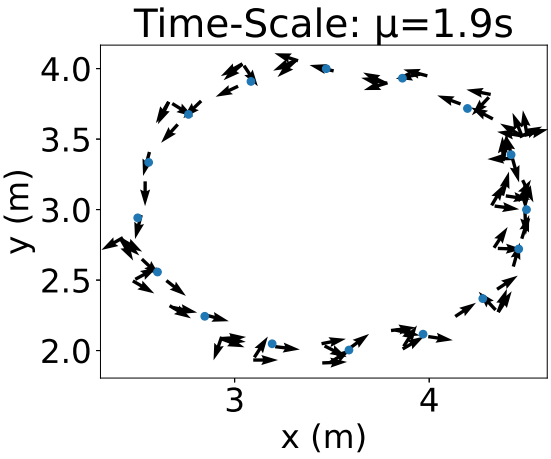
((c)) Continuous-time controller with a DoS attack of $1.5\,\mathrm{s}$



((d)) Time-Scale controller with a DoS attack of $1.5\,\mathrm{s}$



((e)) Continuous-time controller with a DoS attack of $1.9\,\mathrm{s}$



((f)) Time-Scale controller with a DoS attack of $1.9\,\mathrm{s}$

Figure 5.5 – DoS attacks with different strengths on the Continous-Time and Time-Scale controllers.

5.5 Conclusion

In this chapter, we explored the ramifications and mitigation strategies concerning DoS attacks within CPS. These attacks represent a significant threat due to their ability to disrupt communication networks, a critical component underpinning the functionality of CPS. By obstructing the flow of control and measurement data, DoS attacks impair the operational integrity of such systems.

The scholarly investigation into the susceptibility of CPS to DoS assaults has yielded a plethora of strategies aimed at understanding and neutralizing these threats. This includes the deployment of Intrusion Detection Systems (IDS), utilization of unused queue resources for attack detection, and the implementation of advanced network architectures designed to enhance security and resilience. Additionally, machine learning pipelines and frameworks leveraging fog computing and Software Defined Networking (SDN) have been proposed to detect and mitigate these attacks effectively.

Furthermore, the chapter delves into the application of Time-Scale Theory to design controllers and observers capable of maintaining system stability under DoS attack conditions. By treating the attack as a form of delay and adjusting the operational time domain of the controller and observer, the system's resilience is significantly enhanced. The mathematical foundation underpinning this approach, including the synthesis of fixed and variable μ systems, provides a robust framework for ensuring system stability despite the presence of DoS attacks.

The validation on a real system, conducted through a detailed case study involving a TurtleBot3 Waffle Pi, underscores the practical applicability of the proposed solutions. The experimental setup demonstrated the effectiveness of the time-scale controllers and observers in navigating the robot through predetermined waypoints, even under varying intensities of DoS attacks. This not only exemplifies the robustness of the designed control strategies but also highlights the potential for their application in real-world scenarios.

To the authors' understanding, employing time-scale systems in counteracting DoS attacks remains an uninvestigated field. Conventionally, DoS mitigation strategies are implemented within the domain of Information Technology (IT), which directly corresponds to the locus of these attacks. Nevertheless, the integration of a DoS-resilient controller and observer enables designers to ensure continued system functionality, even in instances

where IT-based attack mitigation techniques fall short of complete effectiveness. Thus, this approach serves as an additional layer of defense, complementing existing security measures effectively.

In conclusion, this chapter contributes to the burgeoning body of knowledge regarding the defense against DoS attacks in CPS. By integrating insights from information technology and system automation, and by leveraging advanced theoretical frameworks such as Time-Scale Theory, this research delineates a holistic approach towards enhancing the resilience of CPS against DoS attacks.

The ensuing, concluding chapter furnishes a summation of this thesis, alongside offering insights and projections for prospective research endeavors.

Chapter 6

Conclusion

As we conclude this thesis, it is evident that the susceptibility of cyber-physical systems to attacks poses significant risks to their integrity and functionality. These systems, integral to modern infrastructure, face threats through unauthorized access and manipulation of network layers, potentially leading to the disruption of physical devices. The initiation of an attack can compromise the entire network, enabling adversaries to fabricate data and undermine detection methodologies. Thus, the development of robust detection mechanisms becomes crucial to preserve the security, functionality, and resilience of these systems in a technology-driven world.

In addressing these challenges, cyber-defense strategies must encompass both information technology (IT) and automation. IT-based defenses, including firewalls, encryption, and access control, form the first layer of protection against unauthorized access. However, to counteract the sophisticated tactics employed by attackers, including the exploitation of human vulnerabilities, a secondary layer of defense emerges through automation. This involves the deployment of resilient controllers and advanced detection schemes designed to identify and mitigate the impacts of cyber-attacks. This dual-layered approach highlights the complexity of securing cyber-physical systems and underscores the necessity for focused defense mechanisms tailored to specific types of attacks.

In Chapter 3 we focused on the False Data Injection attack. The contribution of this chapter is the development of a method based on Functional Observers and the Lyapunov stability criterion, thereby circumventing the dependency on methods susceptible to mathematical inaccuracies and addressing the mathematical challenges posed by ill-conditioned matrices frequently encountered in power systems. Additionally, the adoption of Functional Observers offers an alternative perspective on observability,

grounded in graph theory, which is less constraining than conventional approaches.

In Chapter 4, we switched focused to the Zero-Dynamics Attack. The contribution of this chapter is the use of the Time-Scale framework to develop an observer that is straightforward and computationally efficient for real-time applications. This methodology provides an innovative approach to the detection of zero-dynamics attacks, by making use of a technique not yet explored. The working principle of the proposed methodology is akin to that of Zero Order Holder-based ones, making it so the new solution have roots on already established methods.

In Chapter 5, we apply the Time-Scale technique to mitigate Denial Of Service attacks. The contribution of this chapter is the design of an observer and controller which adapt to the attack, making it so the system remains operational despite it. This strategy allows the adjustment of the observer's and controller's sampling interval in reaction to package drops caused by such attacks, enabling the control system to adapt the control signal to the attack, deteriorating performance but remaining functional. This method enhances system resilience and opens a new pathway for the mitigation of Denial-of-Service (DoS) attacks.

There exist certain limitations within the proposed methodologies. Specifically, the techniques predicated on the time-scale framework necessitate the system's dynamic matrix $\mathcal{A}$ to be self-adjoint, a condition not commonly satisfied in practical scenarios. Fortuitously, a transformation can be applied to systems with solely real-valued eigenvalues to endow them with this requisite property. Nonetheless, further research aimed at eliminating this condition while retaining the simplicity inherent in Linear Matrix Inequalities (LMIs) would significantly enhance the applicability of the approach. Additionally, the method for determining the parameter $\bar{\mu}$, which is dependent on specific system properties and presently approximated based on the Nyquist criterion to avoid aliasing, remains underexplored. Establishing a definitive methodology for this parameter's calculation could substantially refine the proposed techniques.

Regarding the Functional Observer approach, opportunities for advancement are discernible. An avenue for exploration is the development of observers that monitor varying paths leading to the same output, thereby embedding redundancy within the system and enhancing its reliability and resistance to disruptions or cyber-assaults. Furthermore, employing disparate sets of outputs to monitor a single sensor could present a method

for increasing the complexity faced by attackers employing stealth strategies. Such approaches exploit the predictable nature of standard observation models. By introducing variability in the observation dynamics, the system's exposure to stealth attacks could be significantly diminished, thereby strengthening its defense mechanisms.

In conclusion, this thesis has demonstrated the significant advancements and potential avenues for further research in the realm of cyber-physical system security, particularly through the focused lens of False Data Injection (FDI), Zero-Dynamics Attack (ZDA), and DoS attacks. The proposed methodologies not only address the current challenges posed by these attacks but also pave the way for innovative approaches to enhance system robustness and resilience. The integration of Functional Observers, the Time-Scale framework, and the application of the Lyapunov stability criterion, among others, represent significant contributions to the field, offering new perspectives and solutions to the complex problems faced by cyber-physical systems.

Looking forward, it is evident that the journey toward fully securing these systems is ongoing and requires continuous effort. As this research contributes to the foundation of knowledge, it also highlights the necessity for interdisciplinary approaches that encompass both Information Technology (IT) and Operational Technology (OT) perspectives. The dynamic nature of cyber threats demands adaptable, intelligent, and multifaceted defense strategies. Therefore, this work not only serves as a testament to the progress made but also as a call to action for future researchers to build upon these findings, explore uncharted territories, and fortify the defenses of our increasingly interconnected world. By advancing our understanding and methodologies, we can aspire to a future where cyber-physical systems operate securely and reliably, safeguarding the critical infrastructures and services that underpin modern society.

Part II

Traduction en Français

Chapter 7

Introduction

7.1 Systèmes Cyber-Physiques

Les Cyber-Physical Systems (CPS) représentent une approche révolutionnaire dans l'ingénierie, fusionnant le physique et le numérique pour créer des systèmes qui sont à la fois interconnectés, utilisant différents types de réseaux, et intelligents, capables de s'adapter aux changements, d'utiliser différentes sources de données et de comprendre leur environnement, si nécessaire. L'intégration stratégique de capteurs, actionneurs, unités informatiques et réseaux de communication permet aux CPS de surveiller, analyser et influencer efficacement leur environnement.

Ce paradigme transformateur améliore non seulement la capacité d'interagir avec le monde physique mais ouvre également de nouvelles voies pour l'innovation, annonçant une nouvelle ère dans la conception et la mise en œuvre de systèmes d'ingénierie. À travers le prisme des CPS, l'avenir de l'ingénierie est envisagé comme un domaine où les barrières entre le physique et le numérique ne sont pas seulement estompées mais effectivement démantelées, ouvrant la voie à des niveaux sans précédent d'interaction et de contrôle sur le monde physique (Foundation 2022).

Les CPS permettent la surveillance, l'analyse et le contrôle en temps réel des processus physiques. Cela permet une large gamme d'applications où le contrôle dynamique et efficace des processus physiques est important. L'échange et le traitement des données en temps réel permettent aux CPS d'anticiper les changements, de réduire les risques et d'améliorer l'efficacité opérationnelle, ce qui en fait une perspective intéressante pour une application dans les dispositifs médicaux, le contrôle du trafic, les véhicules autonomes, l'aviation, la défense et la fabrication, par exemple.

Ce développement entraîne un compromis significatif : l'emploi de dispositifs informatiques plus sophistiqués et le déploiement de réseaux plus larges et interconnectés élargissent le paysage de vulnérabilité des CPS. En conséquence, ces systèmes deviennent susceptibles à des vecteurs d'attaque traditionnellement associés aux systèmes informatiques, tels que les attaques Denial-of-Service (DoS). Auparavant, de telles attaques étaient considérées comme impraticables contre les systèmes de contrôle conventionnels en raison de l'absence de vecteurs d'attaque exploitables.

Par conséquent, la sécurité des CPS reste un défi non résolu, activement étudié par la communauté scientifique. Des progrès sont réalisés dans divers domaines, attribuables une gamme diversifiée de méthodologies d'attaque potentielles. En raison de la superposition avec les besoins traditionnels de défense informatique, le côté sécurité informatique des CPS s'est développé plus que le côté Automatique, un déséquilibre que cette thèse vise à aider à équilibrer.

7.2 Principales Contributions

Les principales contributions de cette thèse sont la formulation de mécanismes de détection pour les attaques False Data Injection (FDI) et Zero-Dynamics Attack (ZDA), ainsi qu'un mécanisme d'atténuation pour les attaques DoS. Les contributions sont délimitées comme suit :

1. Développement d'un mécanisme de détection des attaques FDI qui est robuste face aux défis mathématiques présentés par les matrices mal posées, un problème courant dans les systèmes électriques où les matrices dynamiques sont éparses et englobent des centaines d'états.

2. Création d'un mécanisme de détection des attaques à dynamique nulle en utilisant le cadre Temporel-Échelle. Cette approche vise à dériver une loi de commande simple qui est efficace en termes de calcul lors de l'évaluation en ligne.

3. Conception d'un mécanisme d'atténuation des attaques DoS en exploitant le cadre Temporel-Échelle. Ce mécanisme a pour but d'ajuster le temps d'échantillonnage du contrôleur en réponse aux retards de réseaux induits par l'attaque, améliorant ainsi la résilience du système.

7.3 Organisation du document

Cette thèse est structurée comme suit :

Le Chapter 8 présente un aperçu des vulnérabilités des systèmes cyber-physiques, en se concentrant sur les attaques ciblées contre ces systèmes. Il inclut une revue complète de la littérature concernant les trois catégories d'attaques explorées dans cette étude. De plus, les objectifs de cette recherche sont clairement définis dans ce segment.

Le Chapter 9 détaille la méthodologie pour détecter les attaques par injection de fausses données (FDI), présentant une approche novatrice qui intègre des Observateurs Fonctionnels pour relever le défi des matrices mal posées identifié dans les discussions précédentes. Cette méthode utilise à la fois les Observateurs Fonctionnels et le critère de stabilité de Lyapunov pour assurer la convergence de l'observateur, évitant ainsi le besoin de techniques susceptibles d'imprécisions mathématiques dans le contexte de matrices mal posées, telles que les zéros des matrices et la décomposition en valeurs singulières.

Le Chapter 10 expose la stratégie de détection pour les attaques Zero-Day (ZDA), introduisant les concepts fondamentaux de la théorie des Échelles Temporelles. Ce cadre théorique permet à l'observateur, conçu pour l'estimation d'état, de fonctionner avec un interval d'échantillonnage qui approxime l'opération en temps continu dans des circonstances normales, tout en facilitant la détection de ZDA par des modifications stochastiques de l'interval d'échantillonnage pendant son fonctionnement.

Le Chapter 11 discute d'une approche d'atténuation pour les attaques par déni de service (DoS), utilisant des contrôleurs à échelles temporelles pour renforcer la robustesse du système. La technique tire parti de la capacité d'échantillonnage adaptatif des contrôleurs à échelles temporelles, garantissant la stabilité du système même sous la contrainte d'une faible attaque DoS.

Enfin, le Chapter 12 fournit des remarques conclusives et esquisse des pistes potentielles pour des investigations futures.

7.4 Publications

Les articles suivants ont été publiés durant le développement de ce travail:

- Á. e Sousa, N. Messai, and N. Manamanni, "False Data Injection Detec-

tion in Cyber-Physical System," presented at the Symposium on Fault Detection, Supervision and Safety for Technical Processes, Cyprus: IFAC, 2022, p. 7. doi: 10.1016/j.ifacol.2022.07.165.

- Á. e Sousa, N. Messai, and N. Manamanni, "Load-altering attack detection on smart grid using functional observers," International Journal of Critical Infrastructure Protection, vol. 37, p. 100518, Jul. 2022, doi: 10.1016/j.ijcip.2022.100518.
- *Under review*: Á. e Sousa, N. Messai, and N. Manamanni, "DoS attack mitigation using Time-Scale theory," Control Engineering Practice.

Chapter 8

Aperçu

8.1 Contextualisation

La Fondation Nationale des Sciences Américaine définit les Cyber-Physical Systems (CPS) comme suit[1]:

> *"Les systèmes cyber-physiques (CPS) sont des systèmes conçus qui sont constitués de, et dépendent de, l'intégration transparente de composants informatiques et physiques. Les CPS intègrent étroitement les dispositifs informatiques, l'actionnement et le contrôle, l'infrastructure de réseau, et la détection du monde physique. Le système peut inclure une interaction humaine avec ou sans contrôle aidé par l'humain. Les CPS peuvent également inclure plusieurs composants de système intégrés opérant à une grande variété d'échelles spatiales et temporelles. Ils peuvent être caractérisés par des architectures qui peuvent inclure des calculs distribués ou centralisés, un contrôle hiérarchique multiniveau et la coordination des processus physiques et organisationnels."*

La Figure 8.1 fournit une représentation complète des éléments constitutifs d'un CPS. Cette illustration délimite les actionneurs et les capteurs comme composants du domaine physique, tandis que le reste, y compris les interconnexions, est classé sous le domaine cybernétique du CPS. Le nuage centralisé, représenté dans la figure, peut représenter soit internet soit un réseau étendu, servant de centre névralgique pour le flux de données et de contrôle. La fonction principale de la base de données est le

[1] `https://new.nsf.gov/funding/opportunities/cyber-physical-systems-cps/nsf21-551/solicitation`

stockage des données systèmes, facilitant le traitement ou les activités de consignation ultérieurs. Le serveur joue un rôle crucial dans l'orchestration des opérations du système, comprenant la gestion des points de consigne, l'authentification des utilisateurs, et la supervision d'autres exigences administratives et de contrôle de haut niveau. Les PLC sont dédiés aux tâches de contrôle de bas niveau et à l'interface avec les actionneurs/capteurs, gérant potentiellement des boucles de contrôle locales conformément aux points de consigne dirigés par le serveur ou en se connectant directement avec des équipements non réseautables. Les Interfaces Homme-Machine (IHM) sont polyvalentes par nature, allant de matériel spécialisé, tel que des panneaux de contrôle dans des installations nucléaires, à des ordinateurs conventionnels utilisés par les utilisateurs pour l'ajustement des paramètres, l'émission de commandes systèmes, et la surveillance des variables systèmes.

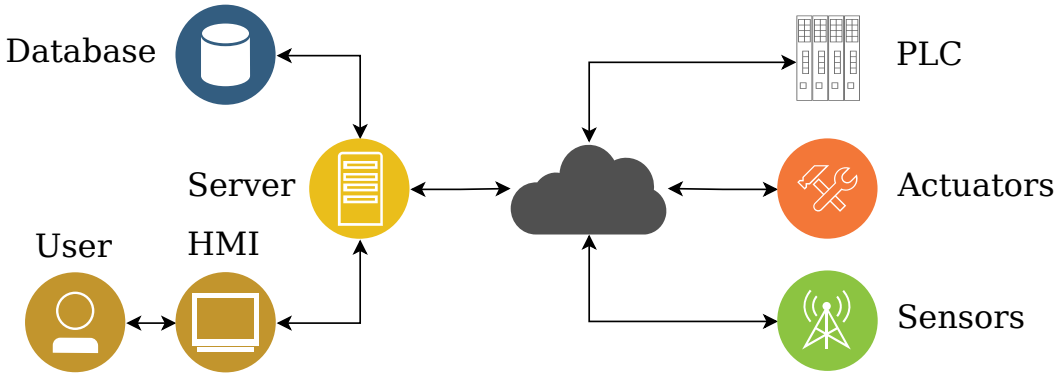


Figure 8.1 – Vue d'ensemble d'un système cyber-physique

L'architecture des CPS repose sur l'amalgame harmonieux de ses éléments constitutifs, couvrant à la fois le matériel tangible et les composants logiciels intangibles. Au cœur de la dimension physique se trouvent des capteurs et des actionneurs intégrés dans l'environnement. Ces composants jouent un rôle crucial dans la perception des conditions environnementales et la mise en œuvre d'interventions physiques basées sur des directives informatiques. Ces interactions soulignent l'aspect tangible des CPS, où des actions physiques sont initiées et ajustées en temps réel en fonction des insights obtenus de l'analyse des données (Krishnamurthy and Cecil 2018; Yajun Lu and J. Cecil 2015; Cecil 2017).

D'autre part, l'aspect cyber des CPS englobe la colonne vertébrale numérique de ces systèmes, incluant les unités de calcul, les algorithmes logiciels, les réseaux de communication et les mécanismes de contrôle sophistiqués. Ce segment est essentiel pour traiter les données collectées

par les capteurs physiques, prendre des décisions éclairées et orchestrer les actions des actionneurs (Bouheroum et al. 2022).

La communication sans faille entre les composants cyber et physiques est la pierre angulaire des CPS, permettant une boucle continue d'échange de données, d'analyse et de mise en œuvre des actions. Cette interaction garantit que les CPS peuvent s'adapter à leur environnement physique en temps réel et l'influencer, optimisant la performance et répondant aux conditions émergentes.

La relation symbiotique entre les composants cyber et physiques des CPS favorise un écosystème où le suivi, l'analyse et le contrôle en temps réel des processus physiques sont réalisables. Cette capacité est vitale dans un large éventail d'applications, allant des systèmes de véhicules autonomes aux réseaux d'énergie intelligents, où le contrôle dynamique et efficace des processus physiques est primordial. L'échange et le traitement de données en temps réel permettent aux CPS d'anticiper les changements, de réduire les risques et d'améliorer l'efficacité opérationnelle, incarnant un changement de paradigme dans la façon dont les systèmes interagissent avec le monde physique (E. A. Lee and Seshia 2017; L. D. Xu, E. L. Xu, and Li 2018).

Les CPS sont largement utilisés à travers une diversité de secteurs, soulignant leur impact significatif et leur applicabilité étendue. La liste suivante fournit des exemples de leur déploiement dans divers secteurs, illustrant la polyvalence et le potentiel transformateur des technologies CPS.

- Les réseaux intelligents incarnent l'avant-garde de l'évolution du réseau électrique, exploitant les CPS pour révolutionner l'infrastructure électrique conventionnelle (IEEE 2013). Cette intégration introduit une gamme d'avantages et de fonctionnalités sophistiquées, redéfinissant fondamentalement le paysage opérationnel des systèmes énergétiques. D'une importance considérable est l'habilitation d'un contrôle précis des émissions facilité par les technologies de réseau intelligent, permettant une surveillance et des stratégies d'intervention méticuleuses pour réduire les émissions, favorisant ainsi la durabilité environnementale. De plus, cette avancée technologique donne aux consommateurs domestiques un contrôle et une visibilité accrus sur leur consommation d'énergie domestique, offrant des avantages qui dépassent les ménages individuels pour englober la conservation de l'environnement et l'efficacité économique à plus grande échelle (Bestehorn and Borsche 2014; Devi and Susmitha 2017).

- Dans le domaine de la santé, une tendance notable implique la convergence des fonctionnalités cyber et physiques au sein des dispositifs médicaux, visant à améliorer le paysage des services de santé. Un attribut prévalent parmi ces dispositifs est l'intégration de capacités sans fil, facilitant les mises à jour et les reconfigurations sans couture nécessaires pour des performances optimales. Notamment, les dispositifs portables possèdent le potentiel d'établir une communication inter-dispositifs, non seulement entre eux mais aussi avec un éventail d'autres dispositifs, y compris les smartphones ou les praticiens médicaux à distance, exploitant la technologie Bluetooth omniprésente comme conduit pour ces interactions (Rushanan et al. 2014).

- Les voitures intelligentes représentent un nouveau paradigme dans l'industrie automobile, caractérisé par leur respect de l'environnement, une efficacité énergétique accrue, des caractéristiques de sécurité améliorées, et un éventail d'offres de divertissement et de commodité, élevant ainsi l'expérience de conduite. Au cœur de l'activation de ces avancées se trouve un réseau sophistiqué comprenant environ 70 ordinateurs connus sous le nom de Electronic Control Units (ECUs). Ces ECUs assument la responsabilité de surveiller et de gouverner un large spectre de fonctions essentielles aux opérations des voitures intelligentes, englobant des tâches liées à la gestion des émissions du moteur, des mécanismes de contrôle précis des freins, ainsi que l'orchestration des commodités de divertissement telles que les radios et les lecteurs multimédias, en plus de la régulation des fonctionnalités de confort telles que l'opération des fenêtres et le contrôle de croisière (Humayed and Luo 2015).

En raison du déploiement étendu des CPS et de leur fonction critique dans les applications industrielles contemporaines, ces systèmes sont de plus en plus devenus le centre d'intérêt des acteurs malveillants. Par conséquent, la capacité à identifier et à répondre à ces activités adverses est d'une importance capitale. La section suivante délimitera les avancées actuelles dans le domaine des méthodologies de détection d'attaques pour les CPS.

8.2 Sécurité des CPS

Détecter les attaques sur les CPS est d'une importance capitale en raison du rôle crucial que ces systèmes jouent dans notre quotidien, nos infrastructures et le bien-être sociétal. La nature interconnectée des CPS, combinant des éléments informatiques, des machines physiques et des réseaux de communication, les rend vulnérables aux attaques cybernétiques malveillantes. Comprendre l'importance de détecter de telles attaques est crucial pour se protéger contre des conséquences potentiellement dévastatrices (Giraldo et al. 2018; Zhenhua Wang, W. Xie, et al. 2021; D. Ding et al. 2018).

Premièrement, les CPS sont omniprésents dans divers secteurs, y compris l'énergie, les transports, la santé et la fabrication. Les attaques contre ces systèmes pourraient entraîner des perturbations généralisées, causant d'importantes pertes économiques et compromettant des services essentiels[2]. Par exemple, une attaque cybernétique contre un réseau électrique intelligent pourrait provoquer des coupures de courant affectant de grandes populations, impactant les entreprises, les hôpitaux et la vie quotidienne[3]. Détecter les attaques dans les CPS est crucial pour prévenir de telles perturbations, assurant la continuité des services vitaux et minimisant l'impact socio-économique[4].

Deuxièmement, l'intégration des CPS dans les infrastructures critiques les expose à des menaces cybernétiques ciblées. Les acteurs malveillants cherchent souvent à exploiter les vulnérabilités de ces systèmes pour causer des dommages physiques, des pertes financières ou manipuler des données à leur avantage. Détecter les attaques sur les CPS devient primordial pour identifier et contrer ces menaces avant qu'elles ne s'aggravent, atténuant les dommages potentiels aux infrastructures, aux actifs et à la sécurité publique (Miller et al. 2021).

Troisièmement, l'interconnexion des CPS avec l'Internet of Things (IoT) introduit de nouvelles surfaces d'attaque. Les vulnérabilités dans un composant du système peuvent se propager et affecter l'ensemble du réseau, créant un effet domino de perturbations. Détecter les attaques rapidement aide à isoler les zones affectées, empêchant la propagation des attaques à travers le réseau et réduisant l'impact global sur le système (Kimani, Oduol,

[2] https://www.nytimes.com/2019/06/15/us/politics/trump-cyber-russia-grid.html

[3] https://www.bbc.com/news/technology-38573074

[4] https://www.independent.co.uk/news/long_reads/cyber-warfare-saudi-arabia-petrochemica
html

and Langat 2019).

Enfin, le potentiel de dommages physiques résultant des attaques sur les CPS ne peut être négligé. Dans des secteurs tels que la santé et les véhicules autonomes, une attaque cybernétique causant une manipulation ou une interruption des opérations pourrait mettre directement en danger la vie humaine (Djenna and Saidouni 2018; Haghighi et al. 2023). La détection rapide des attaques est essentielle pour garantir la sécurité et le bien-être des individus qui dépendent de ces systèmes pour des services et un soutien critiques (Ozarar, Akansu, and Hasbay 2021).

Ainsi, détecter les attaques sur les systèmes cyber-physiques est impératif pour préserver la fonctionnalité, la sécurité et la résilience de ces réseaux interconnectés. Prioriser des mécanismes de détection robustes aide à contrer les activités malveillantes, à protéger les infrastructures critiques, à maintenir la continuité des services et à sauvegarder la sécurité publique dans un monde de plus en plus connecté et dépendant de la technologie.

8.2.1 Exploits sur les CPS

Les attaques contre les CPS comprennent diverses stratégies ciblant la nature interconnectée de ces systèmes, visant à perturber, manipuler ou obtenir un accès non autorisé à des infrastructures critiques. L'exécution de telles attaques peut sembler difficile ; cependant, une variété de vulnérabilités peut être exploitée, permettant à l'attaquant de mener ces assauts. Figure 8.2 réintroduit Figure 8.1, mais en mettant en évidence certaines vulnérabilités et emplacements d'attaque, certains étant expliqués plus en détail dans la séquence.

- Man-in-the-Middle (MitM) : Intercepter la communication entre les composants des CPS pour écouter, manipuler ou injecter du contenu malveillant, compromettant l'intégrité et la confidentialité des données (Oliva, Cioaba, and Hadjicostis 2018; Jena, Padhy, and Guerrero 2023).

  Un exemple notable d'une attaque MitM contre un CPS est le ver Stuxnet, découvert en 2010[5]. Stuxnet visait spécifiquement le logiciel Siemens Step7, utilisé pour programmer les systèmes de contrôle industriel (ICS) opérant dans diverses infrastructures, y compris les installations nucléaires. Le principal objectif de Stuxnet était de saboter

[5] https://www.avast.com/c-stuxnet

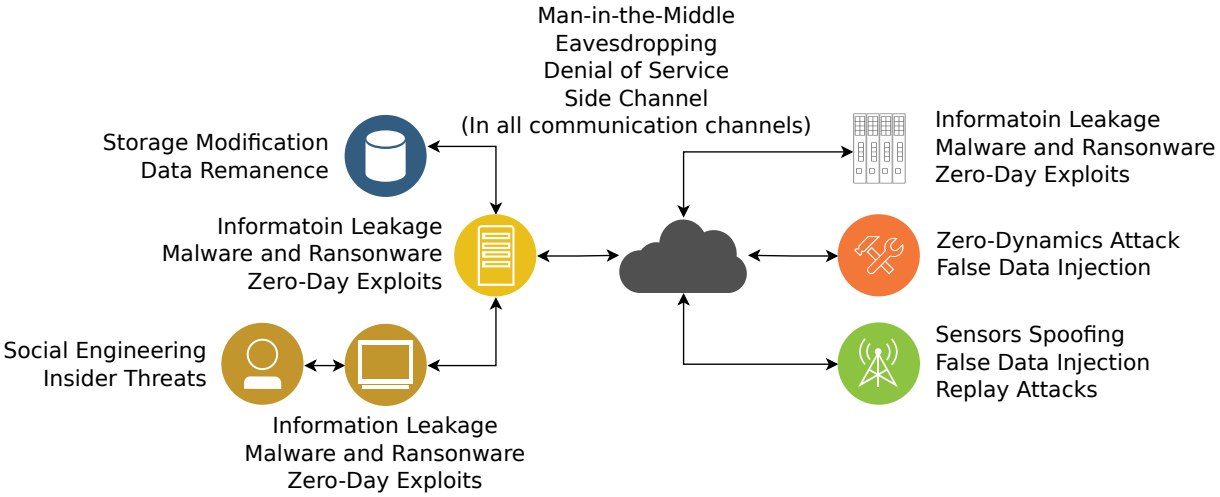


Figure 8.2 – Vulnérabilités et attaques sur les CPS

le programme nucléaire iranien. Il se propageait d'abord via des clés USB puis des partages réseau, cherchant finalement des ordinateurs exécutant le logiciel Siemens Step7. Une fois une cible appropriée infectée, Stuxnet interceptait la communication entre le logiciel et les PLC, qui sont des composants cruciaux dans les systèmes de contrôle industriel.

Le ver était conçu pour modifier le code sur les PLC afin d'induire les centrifugeuses à tourner à des vitesses endommageant tout en signalant simultanément aux systèmes de surveillance que tout fonctionnait normalement. Ce rapport trompeur assurait que les opérateurs du système n'étaient pas conscients du sabotage en cours, plaçant efficacement Stuxnet "au milieu" de la communication entre le logiciel et la machinerie physique.

- Logiciels malveillants et rançongiciels : Infiltrer les CPS avec un logiciel malveillant ou un rançongiciel, perturbant les opérations, volant des données ou chiffrant des systèmes critiques, entraînant une perte de données ou une extorsion financière (Ibarra et al. 2019; Anand and Shanker 2023).

  Un exemple bien documenté d'un rançongiciel impactant un système cyber-physique est l'attaque contre le Colonial Pipeline par le groupe de rançongiciels DarkSide en mai 2021. Le Colonial Pipeline est une pièce cruciale de l'infrastructure aux États-Unis qui transporte de l'essence, du diesel et du carburant pour jets sur un système de pipeline de 5,500 miles (environ 8,850 kilomètres) allant de la côte du Golfe à la région du port de New York. En mai 2021, le groupe

DarkSide a lancé une cyberattaque qui a réussi à infiltrer les systèmes informatiques (IT) du Colonial Pipeline.

En réponse à la violation, la Colonial Pipeline Company a proactivement arrêté ses opérations pour empêcher le rançongiciel de se propager aux systèmes de technologie opérationnelle (OT) qui contrôlent le pipeline, marquant la première fois qu'une attaque de rançongiciel a provoqué l'arrêt d'une installation d'infrastructure critique aux États-Unis. L'arrêt a conduit à des perturbations significatives, y compris des pénuries de carburant dans le sud-est des États-Unis, une augmentation des prix du gaz et une préoccupation généralisée concernant la sécurité des infrastructures critiques contre les menaces cybernét

iques. Les attaquants ont exigé et auraient reçu un paiement de rançon de près de 5 millions de dollars en cryptomonnaie pour fournir un outil de déchiffrement et ne pas divulger les données volées.

- Exploits et vulnérabilités Zero-Day : Exploiter des vulnérabilités auparavant inconnues dans le logiciel ou le matériel des CPS, permettant un accès ou une manipulation non autorisés avant que des correctifs ou des mises à jour ne soient disponibles, posant des risques de sécurité (Halabi and Zulkernine 2023; Gorbenko and Popov 2022).

  Un exemple marquant d'un exploit zero-day affectant un système cyber-physique est le cas du logiciel malveillant Triton (également connu sous le nom de Trisis), découvert en 2017[6]. Ce malware ciblait les systèmes de contrôle industriel (ICS), spécifiquement les systèmes instrumentés de sécurité (SIS), conçus pour surveiller et garantir le fonctionnement sûr des processus physiques dans des environnements industriels. Le malware Triton a exploité des vulnérabilités dans les contrôleurs SIS Triconex fabriqués par Schneider Electric, largement utilisés dans les installations énergétiques, y compris les centrales nucléaires.

  Le malware a été identifié pour la première fois dans une usine pétrochimique au Moyen-Orient, où il a causé le déclenchement inattendu des systèmes d'arrêt d'urgence de l'usine. Des analyses ultérieures ont révélé que Triton était conçu pour permettre aux attaquants de prendre le contrôle des contrôleurs SIS, de modifier leur

[6] https://cert.be/en/paper/trisis-malware

programmation et d'interférer avec leur fonctionnement. L'objectif ultime du malware semblait être la capacité de causer des destructions physiques à l'installation ou d'autres résultats catastrophiques en désactivant des mécanismes de sécurité critiques.

- Altération et injection : Altérer physiquement les composants des CPS ou injecter des défauts dans les capteurs ou les canaux de communication, perturbant la fonctionnalité normale du système ou causant des défaillances matérielles (Jovanov and Pajic 2019; Garagad, Iyer, and Wali 2020).

  Un exemple notable d'altération et d'injection impactant les systèmes cyber-physiques implique l'incident des services d'eau de Maroochy dans le Queensland, en Australie, en 2000[7]. Cet incident est l'une des premières attaques cyber-physiques connues ayant entraîné des dommages environnementaux tangibles et des risques pour la sécurité publique.

  Un ancien employé de l'entreprise qui fournissait le système SCADA (Supervisory Control and Data Acquisition) utilisé par les services d'eau de Maroochy a exécuté l'attaque. Ils ont utilisé un émetteur radio pour accéder sans autorisation au système SCADA, qui contrôlait les processus de traitement des eaux usées et de gestion des déchets. Pendant plusieurs semaines, ils ont altéré le système en injectant des commandes malveillantes qui ont causé le déversement de plus de 800,000 litres d'eaux usées brutes dans des parcs locaux, des rivières et même sur les terrains d'un hôtel. La fuite d'eaux usées a entraîné des dommages environnementaux importants, tué la vie marine et posé de sérieux risques pour la santé de la population locale.

  Ce cas d'altération et d'injection était particulièrement frappant car il a démontré comment les vulnérabilités dans les systèmes de contrôle industriels pouvaient être exploitées pour causer des dommages physiques et des dangers pour la santé publique, soulignant l'importance de sécuriser de tels systèmes contre les accès et les manipulations non autorisés.

- Menaces internes : Poser des menaces de l'intérieur d'une organisation ou d'un système, où des employés ou des entités de confiance abusent de leurs privilèges d'accès, compromettant involontairement

[7] https://www.industrialcybersecuritypulse.com/facilities/throwback-attack-an-insider-releases-265000-gallons-of-sewage-on-the-maroochy-shire/

ou délibérément la sécurité des CPS (Ikany and Jazri 2019; Al Hammadi, Yeob Yeun, and Damiani 2020).

Le cas des services d'eau de Maroochy est également un excellent exemple de menace interne.

8.2.2 Attaques sur les CPS

Certains types d'attaques notables sont illustrés dans Figure 8.3 et incluent:

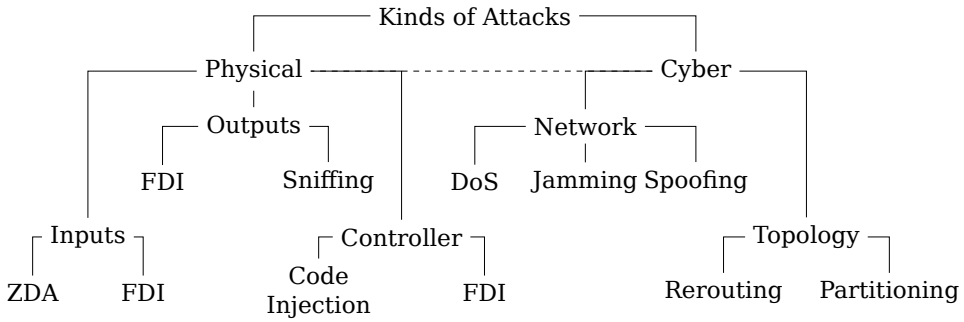


Figure 8.3 – Types d'attaques sur un CPS.

- Denial-of-Service (DoS) : Submerger un CPS ou ses composants avec un afflux de trafic, rendant le système inaccessible aux utilisateurs légitimes en épuisant les ressources du système (G. He et al. 2021; P. Ding et al. 2017).

  Un exemple bien documenté d'une attaque par déni de service (DoS) contre un système cyber-physique est l'attaque contre le réseau électrique ukrainien en décembre 2015[8]. Cet incident est notable comme l'une des premières attaques cybernétiques publiquement reconnues ayant réussi à perturber les opérations d'un réseau électrique, entraînant des conséquences réelles significatives.

  Les attaquants ont lancé une opération cybernétique sophistiquée contre trois compagnies de distribution d'électricité régionales en Ukraine. L'objectif principal était de couper l'électricité aux clients, et cela a été réalisé en accédant à distance aux systèmes de contrôle industriels (ICS) des compagnies et en éteignant les sous-stations électriques. Dans le cadre de l'attaque, les attaquants ont également déployé un malware qui a effacé les données sur les ordinateurs des compagnies, compliquant les efforts de récupération.

[8] https://www.bbc.com/news/technology-38573074

En plus de manipuler directement les systèmes de contrôle pour couper le courant, les attaquants ont utilisé une attaque DoS contre les systèmes téléphoniques des compagnies affectées. Cela avait pour but d'empêcher les clients de signaler les pannes ou de recevoir des informations, exacerbant l'impact de la perturbation de l'alimentation électrique. L'attaque DoS sur les lignes de service clientèle était un mouvement stratégique pour augmenter l'effet de la perturbation physique en limitant la communication et en ralentissant les processus de réponse et de récupération.

L'attaque contre le réseau électrique ukrainien illustre comment les tactiques de DoS peuvent être utilisées non seulement pour perturber les services numériques mais aussi pour renforcer l'impact des attaques physiques sur les infrastructures critiques. L'incident a souligné la vulnérabilité des infrastructures critiques aux cyberattaques et a mis en lumière la nécessité de renforcer les mesures de cybersécurité, y compris des plans de réponse robustes aux incidents et une amélioration de la résilience des systèmes de contrôle industriel contre de telles menaces.

- Attaques par usurpation : Se faire passer pour une entité ou un dispositif légitime au sein du CPS pour obtenir un accès ou des privilèges non autorisés en falsifiant des adresses IP, des adresses MAC ou des signatures numériques (Alsulami and Zein-Sabatto 2021; Kapoor, Vora, and Kang 2018).

  Un exemple notable d'une attaque par usurpation contre un système cyber-physique est l'incident impliquant la capture d'un drone iranien en 2011[9]. Dans ce cas, les États-Unis ont affirmé que l'Iran avait capturé un RQ-170 Sentinel, un drone de surveillance furtif, en usurpant ses signaux de contrôle. Les Iraniens auraient intercepté le lien de communication du drone, le trompant pour qu'il atterrisse en Iran sous la croyance qu'il retournait à sa base en Afghanistan.

  La technique aurait impliqué le brouillage des liens de communication pour perturber les signaux de contrôle entre le drone et ses opérateurs, puis l'utilisation d'une transmission voyou pour usurper les canaux de contrôle du drone, rendant possible la redirection du drone pour qu'il atterrisse en Iran. Les détails techniques précis de la technique d'usurpation n'ont pas été divulgués publiquement, mais

[9] https://www.bbc.com/news/world-middle-east-21373353

l'incident a mis en évidence les vulnérabilités dans les liens de communication sécurisés pour les systèmes sans pilote et télécommandés.

Cet événement a souligné l'importance de sécuriser les canaux de communication et de mettre en place des protocoles d'authentification robustes pour empêcher l'accès et le contrôle non autorisés de systèmes critiques. Il a également mis en lumière le paysage évolutif des menaces cybernétiques, où les adversaires peuvent exploiter les vulnérabilités dans la technologie sophistiquée pour obtenir des avantages stratégiques.

- Injection de fausses données : Injecter des données fausses ou manipulées dans les capteurs ou les systèmes de contrôle des CPS, provoquant des décisions ou des actions erronées basées sur des informations corrompues, affectant la fiabilité ou la sécurité du système (G. He et al. 2021; Garagad, Iyer, and Wali 2020).

  La plus grande cible des attaques par injection de fausses données sont les infrastructures critiques, comme l'énergie, les finances et la sécurité, qui ne divulguent pas les attaques, rendant difficile la recherche d'exemples de telles attaques.

- Attaque par canal auxiliaire : Obtenir des informations à partir de l'implémentation physique d'un système, telles que la consommation d'énergie, les fuites électromagnétiques ou le son, pour extraire des données sensibles (Gupta et al. 2019).

  Un exemple bien documenté d'une attaque par canal auxiliaire contre un système cyber-physique est l'attaque contre le système de chiffrement du système KeeLoq de verrouillage à distance sans clé, largement utilisé dans les télécommandes de clés de voiture[10]. KeeLoq est un chiffreur à blocs dédié au matériel qui implémente une forme de chiffrement par clé symétrique. Il était utilisé par divers fabricants automobiles pour les systèmes de verrouillage à distance sans clé, permettant aux propriétaires de voitures de déverrouiller leurs véhicules à distance. La sécurité de KeeLoq a été compromise par une attaque par canal auxiliaire connue sous le nom d'« attaque d'analyse de puissance ».

  Dans cette attaque, les attaquants ont pu déduire les clés de chiffrement secrètes utilisées par le système KeeLoq en observant la consom-

[10] https://www.science20.com/news_releases/keeloq_remote_keyless_entry_system_for_cars_and_buildings_is_hacked

mation d'énergie du dispositif pendant le processus de chiffrement. En analysant les variations de consommation d'énergie, les attaquants pouvaient inférer les bits de clé et finalement reconstruire la clé secrète utilisée pour chiffrer et déchiffrer les signaux envoyés entre la télécommande et le véhicule. Avec ces informations, un attaquant pouvait obtenir un accès non autorisé à un véhicule, le verrouillant ou le déverrouillant à volonté.

- Attaques par rejeu : Capturer des données transmises entre des dispositifs et les retransmettre pour créer un effet non autorisé (Tang, Z. Zhang, and L. Xie 2023).

  Un exemple notable d'une attaque par rejeu contre un système cyber-physique s'est produit dans le contexte des systèmes de sécurité automobile, spécifiquement les systèmes d'entrée et d'allumage sans clé dans les voitures[11]. Dans une telle attaque, des acteurs malveillants interceptent et enregistrent le signal envoyé par une télécommande de clé de voiture au véhicule pour déverrouiller les portes ou démarrer le moteur. Ce signal enregistré est ensuite retransmis ultérieurement pour obtenir un accès non autorisé au véhicule.

  Les instances réelles, divulguées publiquement, d'attaques par rejeu sur des systèmes cyber-physiques, au-delà des démonstrations de recherche ou des vulnérabilités théoriques, tendent à être rares ou moins fréquemment rapportées dans des archives publiques détaillées. Cette rareté est en partie due à la nature sensible de tels incidents, où les organisations pourraient ne pas divulguer des détails pour éviter de révéler des vulnérabilités ou en raison de potentielles répercussions légales et réputationnelles. Cependant, l'industrie automobile a vu des incidents qui suggèrent fortement l'utilisation de rejeu ou de types d'attaques similaires, bien que ces incidents soient souvent révélés par des sources indirectes telles que des rapports d'assurance, des agences d'application de la loi ou des enquêtes médiatiques.

- Brouillage : Perturber ou bloquer des canaux de communication par l'émission de signaux interférents (Bout, Loscri, and Gallais 2020).

  Un exemple réel d'une attaque par brouillage contre un système cyber-physique s'est produit avec l'utilisation de dispositifs de brouillage GPS pour perturber les systèmes de suivi de véhicules commerciaux[12].

[11] https://medium.com/codex/rollback-a-new-time-agnostic-replay-attack-against-the-autom
[12] https://www.theguardian.com/technology/2010/feb/22/car-thieves-using-gps-jammers

Ce type d'attaque a été observé dans des scénarios impliquant des entreprises de camionnage et de logistique, où des voleurs utilisent des brouilleurs GPS pour empêcher les véhicules d'être suivis, facilitant le vol du véhicule lui-même ou des marchandises qu'il transporte.

- Manipulation des dynamiques à zéro : Exploiter la dynamique physique du système en manipulant des actionneurs ou des signaux de contrôle pour induire des comportements inattendus ou contourner des mécanismes de contrôle, compromettant la stabilité ou la sécurité du système (Zhenhua Wang, W. Xie, et al. 2021; Z. Yu et al. 2021).

  Ce type d'attaque ne voit également aucun rapport dans les médias, rendant difficile la fourniture d'exemples concrets.

Dans cette section, nous avons exploré les différentes méthodes d'exploitation et d'attaque d'un système CPS, montrant comment les attaques sont réellement réalisables et peuvent avoir un impact significatif non seulement sur le système, mais aussi sur la vie des personnes. Dans la section suivante, nous explorons les techniques de défense existantes.

8.3 Défense contre les cyberattaques

La défense contre les cyberattaques opère sur deux fronts : les stratégies Information Technology (IT) englobent des mesures de sécurité telles que les pare-feu, le chiffrement et le contrôle d'accès pour contrecarrer l'accès non autorisé au réseau. Cependant, se fier uniquement à ces mesures s'avère insuffisant, compte tenu de la possibilité pour les pirates d'exploiter les vulnérabilités humaines au sein du système.

Par conséquent, une couche supplémentaire de défense émerge du point de vue de l'automatisation, employant des contrôleurs résilients aux attaques, des observateurs et des schémas de détection pour à la fois identifier et se remettre des cyberattaques. Ainsi, alors que les IT forment la première couche de protection, l'automatisation sert de deuxième ligne de défense (Mrabet et al. 2018).

Les CPS englobent de multiples composants, nécessitant diverses stratégies de défense pour atténuer les attaques. Le champ de cette dissertation se concentre sur les assauts facilités par l'aspect cybernétique des CPS mais ayant des implications directes sur sa dimension physique, excluant les considérations principalement alignées avec la facette IT des CPS. En

se référant au schéma illustré dans Figure 8.1, l'accent est mis sur les systèmes où l'interaction entre les entrées (actionneurs) et les sorties (capteurs) peut être délimitée par des équations différentielles. Ces interactions sont ensuite modélisées en utilisant le cadre de représentation d'espace d'état.

Étant donné l'hétérogénéité des stratégies d'attaque et leurs conséquences, il est impraticable de concevoir un mécanisme de défense universel capable d'identifier toutes les attaques envisageables. Par conséquent, il devient impératif de se concentrer sur une catégorie ou une classe spécifique d'attaques, pour faciliter le développement de schémas de détection d'attaques efficaces.

L'accent a donc été mis sur trois attaques pour une enquête ciblée : False Data Injection (FDI), Zero-Dynamics Attack (ZDA) et DoS. Les sous-sections suivantes présentent une revue de la littérature sur les techniques de détection et d'atténuation de ces attaques.

8.3.1 Injection de Fausses Données

Les attaques par FDI présentent une prévalence significative dans les systèmes électriques, notamment au sein des réseaux intelligents, rendant ces systèmes un principal centre d'intérêt pour les chercheurs engagés dans le développement de mécanismes de détection des FDI. Les systèmes électriques désignent les réseaux de composants électriques déployés pour générer, transmettre et distribuer l'électricité des centrales électriques aux consommateurs. Ces systèmes englobent des générateurs, des lignes de transmission et des réseaux de distribution. Les réseaux intelligents, une évolution des systèmes électriques traditionnels, intègrent la technologie de communications numériques pour surveiller, prédire et contrôler intelligemment le flux d'électricité, améliorant l'efficacité, la fiabilité et la durabilité. Ils permettent une communication bidirectionnelle entre les fournisseurs de services et les consommateurs, soutenant la gestion de l'énergie en temps réel et l'intégration des sources d'énergie renouvelables.

Les attaques par FDI posent une menace significative aux systèmes électriques en compromettant l'intégrité du flux d'informations à travers des tactiques telles que l'écoute clandestine et la répétition (Makar et al. 1975; Yohanandhan et al. 2020; J. Zhang et al. 2018). Ces attaques ciblent le calcul des informations d'état au sein du système, particulièrement lorsque l'attaquant possède une connaissance de la topologie physique du système. Cette connaissance permet à l'attaquant d'éviter les mécanismes de pro-

tection conventionnels tels que les estimateurs statiques et le détecteur de mauvaises données (Lakshminarayana et al. 2021). Le succès des attaques FDI peut entraîner de graves conséquences, allant de l'interférence du système et des surcharges de puissance à des pannes généralisées et des risques potentiels pour l'infrastructure physique et la sécurité humaine (Ahmed and Pathan 2020; J. Zhang et al. 2018).

La recherche de Zhao, Mili, and M. Wang (2018) a introduit un concept nommé une attaque FDI parfaite, qui pose des défis en matière de détection tout en permettant à l'attaquant de contrôler la modification des informations d'état. Lorsqu'un intrus manipule l'injection de puissance à un nœud spécifique, des changements dans les injections de puissance aux nœuds adjacents deviennent nécessaires pour exécuter des modifications d'état prédéfinies. En supposant une connaissance complète de la topologie du système, Zhao, Mili, and M. Wang (2018) a démontré que l'ampleur de l'attaque influençait la topologie du système électrique et la construction du vecteur d'attaque, délimitant les contraintes à travers les diagrammes de décision binaire (BDD). En employant un modèle d'autorégression pour prédire les mesures de Phasor Measurement Unit (PMU) tout en assurant la consistance statistique des mesures, ils ont conçu un modèle où l'estimateur Huber minimisait la somme des résidus standardisés. La détection d'une attaque FDI reposait sur l'approche asymptotique des résidus entre les mesures prédites et réelles tendant vers zéro.

Abusorrah et al. (2019) ont proposé un cadre théorique multi-niveaux visant à formuler une stratégie de défense rentable pour minimiser les coûts opérationnels en présence d'instances de FDI. Cette stratégie impliquait l'optimisation des dépenses opérationnelles sous un regret maximal dans un problème d'optimisation multi-niveaux, tenant compte des interactions entre les défenseurs, les attaquants et les opérateurs. Le cadre a été reformulé en un problème de programmation linéaire mixte bi-niveau équivalent. Une technique d'énumération implicite avide facilitait la découverte d'un ensemble sécurisé globalement optimal pour ce problème bi-niveau complexe avec moins d'itérations. La conversion de la détection en un problème plus gérable était réalisée en résolvant le flux optimal de puissance sécurisé (SCOFF).

Dans le contexte des systèmes de puissance cyber-physiques (CPPS) comprenant n nœuds, la priorité reposait sur la sécurisation d'un ensemble de mesures de base (BMS) de $(n-1)$ compteurs contre les FDI (Sreeram and Krishna 2019). Une méthode employant un BMS optimal avide a été

conçue pour déterminer un BMS acceptable. Étant donné que l'exécution d'une FDI nécessitait une connaissance de la matrice de mesure, un sous-ensemble optimal de cette matrice était choisi pour augmenter le coût de l'attaque. Sreeram and Krishna (2019) ont utilisé un modèle de coût d'attaque de base connu sous le nom de vecteurs d'attaque parcimonieux multiples (MSAV). Cependant, les limites de la méthode d'énumération ont émergé lorsque Sreeram and Krishna (2019) n'ont pas réussi à conclure la validation dans le modèle à 300 nœuds de l'IEEE en raison de contraintes arithmétiques. Dans des scénarios pratiques, des méthodologies de recherche plus efficaces ou des algorithmes compatibles sont prêts à exceller dans la gestion de CPPS de plus en plus complexes.

La construction du vecteur d'attaque FDI repose fortement sur la topologie du CPPS. L'optimisation en temps réel de la topologie améliore les taux de détection via les diagrammes de décision binaire (BDD). La reconfiguration du système et le reroutage du flux de puissance sont principalement mis en œuvre par le délestage, la commutation de ligne et la division de nœud. Zheng et al. (2021) ont conçu un modèle d'apprentissage par renforcement profond (DRL) pour résoudre les problèmes d'optimisation de la topologie et identifier de manière préventive les vulnérabilités pour sauvegarder la sécurité du réseau. En utilisant un modèle de perturbation adversaire basé sur la criticité, des vecteurs d'état trompeurs étaient recherchés lorsque les paramètres du modèle DRL étaient connus, provoquant des configurations de puissance erronées et une divergence du flux de puissance due à FDI. Des perturbations adverses visant à minimiser la récompense DRL dans des magnitudes d'erreur minimales étaient construites. De plus, des indices de vulnérabilité pour les modèles DRL étaient proposés sur la base de critères de probabilité et de gradient.

Cheng and Chow (2020) ont abordé les attaques FDI au sein du système de gestion de l'énergie (EMS) pour la distribution économique. L'évaluation a mis en évidence les risques et les actifs pertinents pour un EMS basé sur le consensus. Divers modèles d'attaque ont été conçus d'un point de vue axé sur le risque, introduisant un mécanisme de vote majoritaire pour un algorithme unifié de détection d'attaques basé sur la réputation. Les agents comparaient les estimations locales avec les données fournies par les agents voisins et des erreurs excessives affectaient la réputation d'un agent. Pour mettre à jour le consensus global, les agents se tournaient vers des informations provenant d'agents hautement réputés. Lorsqu'un agent perdait la confiance en raison de la surveillance de voisinage basée

sur la réputation, les autres substituaient les données d'état mises à jour de l'agent FDI par des estimations locales. Divers agents, y compris l'énergie éolienne et solaire, ainsi que différentes charges, étaient intégrés dans un cadre CPPS complet.

Saad et al. (2020) ont construit un modèle de jumeau numérique pour CPPS et ont évalué sa résilience. Une plateforme de jumeau numérique basée sur le cloud formait le centre de contrôle, sélectionnant les nœuds leader et suiveur du cluster dans le contrôle secondaire basé sur la topologie du réseau. Le nœud de point de couplage commun (PCC) agissait comme le pont entre les couches physiques et informationnelles du réseau dans le contrôle tertiaire. La vérification de l'intégrité de chaque information transmise au PCC assurait une sécurité maximale. L'utilisation des résidus de l'estimation d'état locale et des valeurs transmises aidait à identifier la source de FDI. La divergence de Kullback-Leibler (une méthode statistique pour mesurer la distance entre deux distributions de probabilité) aidait à déterminer les écarts parmi les voisins du suiveur par rapport au consensus, mettant à jour les attentes pour les nœuds leader.

M. Shi et al. (2021) ont proposé une stratégie de contrôle distribué intégré résiliente basée sur un observateur pour atténuer les attaques FDI sur les capteurs et les actionneurs dans les micro-réseaux AC insulaires (MG). Adhérant à la théorie du consensus, le contrôle de fréquence distribué intégré se concentrait sur la transmission de signaux de puissance active normalisés tout en assurant la stabilité de fréquence globale. Cette approche réduisait considérablement les charges de communication tout en préservant les valeurs locales de fréquence et de tension. En s'appuyant uniquement sur les variables d'état locales pour estimer et contrer les attaques potentielles sur chaque générateur, la stratégie assurait la cybersécurité du point de vue d'un capteur.

Cependant, les méthodologies traditionnelles, notamment celles qui dépendent des filtres de Kalman, rencontrent encore des défis. Ces défis sont principalement attribués à l'espace d'état étendu associé aux réseaux intelligents, souvent comprenant des centaines d'états. Cette complexité peut provoquer des instabilités numériques et des inefficacités computationnelles dans les mécanismes de détection conventionnels (Chen 2013). Par conséquent, il existe une demande croissante pour des stratégies de détection non seulement plus efficaces mais également évolutives. Cette recherche vise à aborder cette question par l'emploi d'LMI et d'observateurs fonctionnels, comme détaillé dans Chapter 9.

8.3.2 Attaque à Zéro-Dynamique

Dans le domaine des systèmes de contrôle, l'attaque à zéro-dynamique représente une menace significative en raison de sa difficulté inhérente à la détection. Elle exploite la dynamique zéro du système pour injecter un signal qui affecte les états du système mais pas ses sorties. Typiquement, la détection se fait du côté du contrôleur, en utilisant les données d'entrée du contrôleur et les données de sortie de l'usine. Cependant, ces données peuvent être compromises à travers les réseaux de communication. L'attaquant manipule stratégiquement ces données pour assurer la cohérence avec la dynamique de l'usine, compliquant le processus de détection (Andre Teixeira et al. 2012).

L'étude dans (Jihan Kim, Back, et al. 2020) emploie un Generalized Hold (Yuz and Goodwin 2014) pour altérer les zéros du système, entravant ainsi la capacité de l'attaquant à les exploiter. Cette technique de déplacement des zéros du système vers une région stable grâce au Generalized Hold contrarie efficacement l'attaque à zéro-dynamique. Néanmoins, elle modifie la performance du système en boucle fermée par rapport à celle avec Zero Order Hold, pouvant dégrader par rapport à la conception originale.

Dans la recherche présentée dans (D. Kim, Ryu, and Back 2020), un Generalized Sampler est utilisé dans un objectif similaire. Le Generalized Sampler permet de placer les zéros de la représentation en temps discret de la plante à l'intérieur du cercle unitaire. Bien que les attaques à zéro-dynamique puissent rester furtives, leur potentiel perturbateur est atténué car les signaux d'attaque diminuent. L'utilisation combinée du Generalized Hold et du Generalized Sampler neutralise l'attaque à zéro-dynamique. Notamment, contrairement au Generalized Hold, le Generalized Sampler n'affecte pas le comportement inter-échantillon du système en temps continu. Cependant, la sortie $y_g(k)$ diffère de l'échantillon de sortie conventionnel $y(k)$, ce qui peut être indésirable pour certaines applications.

La vulnérabilité des systèmes de contrôle chiffrés aux attaques à zéro-dynamique est examinée dans (J. Lee, Junsoo Kim, and Shim 2020). Malgré la mise en œuvre du chiffrement homomorphe, ces systèmes sont susceptibles de manipulation via la connaissance du modèle de l'usine et des propriétés du chiffrement. L'étude démontre par des simulations comment l'attaque peut perturber l'état de l'usine tout en restant non détectée à l'entrée et à la sortie du contrôleur, soulignant la nécessité d'une sécurité accrue contre les attaques à zéro-dynamique dans les systèmes de contrôle chiffrés.

D. Kim, Ryu, J. H. Kim, et al. (2021) proposent une défense innovante contre les attaques à zéro-dynamique dans les systèmes de contrôle en réseau, impliquant l'utilisation d'un échantillonneur généralisé pour modifier les zéros du système. Cette méthode reste efficace contre les attaquants possédant une connaissance complète du système en raison de sa capacité d'attribution arbitraire des zéros. L'article détaille le processus de conception de l'échantillonneur généralisé et confirme son efficacité à travers des simulations numériques, marquant un progrès par rapport aux méthodes de défense existantes.

L'article (Park et al. 2019) se penche sur une stratégie d'attaque à zéro-dynamique robuste pour les systèmes cyber-physiques incertains, abordant le défi de l'incertitude du modèle. Les auteurs suggèrent une méthode impliquant l'isolement et la substitution de la zéro-dynamique réelle par une zéro-dynamique nominale auxiliaire, maintenant la furtivité sous des conditions incertaines. Utilisant un observateur de perturbation pour compenser les inexactitudes du modèle, la stratégie se révèle efficace sans connaissance précise du modèle. Des simulations avec un système de puissance hydro-turbine démontrent la robustesse de l'attaque, soulignant la nécessité de stratégies de sécurité résilientes contre de telles attaques adaptables.

Baniamerian, Khorasani, and Meskin (2020) introduisent une méthodologie pour détecter les attaques à zéro dynamique (ZD) dans les systèmes cyber-physiques (CPS). Ils proposent un système auxiliaire et des filtres de détection efficaces contre les attaquants informés du CPS et des mesures défensives. Cette approche, distincte des méthodes actuelles, ne dépend pas d'informations cachées ou de modifications étendues du système, mettant en évidence l'importance de défenses robustes et résilientes contre des attaques ZD sophistiquées.

L'étude dans (Mao and Akyol 2018) se concentre sur la détection d'attaques à zéro dynamique coopératives dans les systèmes en réseau, en particulier dans les oscillateurs harmoniques couplés. Les auteurs présentent une stratégie de défense englobant la commutation stratégique de topologie et l'observation des sorties, indépendamment du moment d'initiation de l'attaque ou des oscillateurs ciblés. Cette méthode expose efficacement les attaques à zéro dynamique coopératives, garantissant qu'elles n'altèrent ni la valeur d'équilibre ni la stabilité du système. L'article étaye cette approche avec des simulations, contribuant notablement à la sécurité des réseaux cyber-physiques.

Dans (Jihan Kim and Shim 2019a), les auteurs abordent les attaques de capteurs à zéro dynamique dans les systèmes cyber-physiques, proposant un « retour de rétroaction généralisé » comme nouvelle contre-mesure pour stabiliser le système. Cette technique déplace les pôles du système vers des positions stables, neutralisant l'attaque. Démontrée à l'aide d'un système de lévitation magnétique, cette stratégie contrarie non seulement l'attaque mais améliore également la sécurité globale du système, offrant un progrès substantiel dans la résilience des systèmes cyber-physiques contre des attaques de capteurs furtives.

Cependant, une vulnérabilité partagée parmi ces techniques de détection est évidente : si les attaquants apprennent la stratégie de détection, ils peuvent altérer les signaux utilisés pour la détection, dissimulant efficacement l'attaque. Pasha and Ayub (2021) ont enquêté sur plusieurs méthodes de détection d'attaques, toutes basées sur la même structure, qui reste à risque si les attaquants découvrent l'approche de détection. Cette faiblesse provient du processus d'envoi d'un signal au contrôleur et au système de détection, que l'attaquant pourrait modifier. Dans Chapter 10, une nouvelle approche utilisant le calcul à l'échelle temporelle est présentée. Cette méthode s'appuie sur les principes éprouvés des techniques précédentes mais utilise une structure différente et s'adapte aux temps d'échantillonnage variables. Cette combinaison rend plus difficile la réussite des attaquants, surtout lorsqu'elle est associée à des techniques de théorie des jeux.

8.3.3 Denial of Service

Les attaques par déni de service (DoS) constituent une menace significative pour la fiabilité et la fonctionnalité des composants cybernétiques dans les systèmes cyber-physiques. Ces attaques ciblent principalement et perturbent les canaux de communication, exploitent les vulnérabilités dans les protocoles de réseautage et inondent le réseau avec un trafic excessif. La conséquence principale de ces attaques est l'échec des liaisons de communication et l'introduction de retards excessifs, ce qui entrave sévèrement l'échange critique et opportune d'informations vitales entre les capteurs, les actionneurs et les systèmes de contrôle central, compromettant ainsi l'intégrité globale du système.

Dans le domaine des attaques DoS, trois stratégies prédominantes et distinctes sont couramment observées. Premièrement, la stratégie DoS aléatoire se caractérise par une perte imprévisible et sporadique de données

de mesure du système, qui peut être analysée quantitativement en utilisant le modèle statistique de Bernoulli, comme décrit par K. Ding et al. (2019) dans leur article fondateur. Deuxièmement, la stratégie DoS périodique, une forme d'attaque plus structurée, entraîne l'incapacitation complète des canaux de communication pendant l'assaut de l'attaque, empêchant efficacement toute forme de transmission d'informations pendant ces périodes. Enfin, les attaques DoS non périodiques se distinguent par leur minutage irrégulier et des intervalles de dormance, les rendant imprévisibles et difficiles à atténuer.

Le concept d'une architecture auto-organisatrice (SOA) a été proposé comme structure défensive contre de telles attaques. La SOA est stratégiquement organisée en trois niveaux hiérarchiques : le niveau de base comprend des agents clients/générateurs, le niveau intermédiaire comprend un contrôleur local appelé Agent Agrégé (AA), et le niveau supérieur est formé par le centre de contrôle local. Cette structure a été élaborée de manière détaillée dans l'étude menée par Cameron et al. (2019). Ils ont introduit une nouvelle plateforme de système de puissance cyber-physique (CPPS) qui intègre la SOA, représentée dans le Java Agent Development Framework (JADE), avec une simulation de système électrique multiniveau utilisant la boîte à outils Matpower de Matlab.

Cette intégration a été mise en avant comme une avancée critique par rapport aux recherches précédentes, que Cameron et al. (2019) ont critiquée pour son manque de modélisation complète des attaques cybernétiques et d'évaluations de la vulnérabilité dans les systèmes électriques, limitant ainsi la capacité à évaluer quantitativement les impacts des attaques cybernétiques. En cas d'attaque DoS, l'agent affecté au sein de cette architecture est promptement isolé, et le centre de contrôle redirige stratégiquement la communication vers le canal disponible le plus efficace. L'efficacité et le temps de réponse de la SOA sont davantage optimisés en désactivant les agents jugés non critiques et en isolant ceux compromis.

Cependant, l'efficacité de la SOA est intrinsèquement limitée par l'intensité et la variabilité de l'attaque DoS, nécessitant des recherches continues pour améliorer encore son moteur de prise de décision, potentiellement par l'incorporation de techniques d'apprentissage par renforcement profond (Cameron et al. 2019).

De plus, le rôle du contrôle de la fréquence de charge (LFC) est examiné de manière critique dans le contexte des services auxiliaires dans les systèmes électriques. Le LFC joue un rôle essentiel dans le maintien de

l'équilibre de l'énergie à court terme et de la fréquence, stabilisant ainsi la qualité globale de l'énergie et facilitant l'échange efficace d'énergie. Cela est réalisé en ajustant dynamiquement les angles de rotor des générateurs adjacents en réponse aux signaux de déviation de puissance. Les systèmes électriques multi-zones traditionnels ont historiquement compté sur des mécanismes déclenchés par le temps à cet effet. Cependant, ces mécanismes sont désormais reconnus comme inadéquats face aux attaques DoS, principalement en raison de leur nature gourmande en bande passante. Abordant cette question, Hossain et al. (2022) ont proposé un nouveau mécanisme de déclenchement d'événements distribués (DETM) pour renforcer les capacités de défense en temps réel du système LFC tout en minimisant simultanément l'utilisation des ressources de communication.

De plus, l'application des méthodes de fonction Lyapunov (LF) dans l'évaluation de la stabilité des systèmes non linéaires a été explorée dans des recherches récentes. Par exemple, J. Liu, Lu, and Jianhui Wang (2019) ont employé LF dans le développement d'un micro-réseau à courant continu (MG) à commande directe basée sur le retour d'information en temps variant, convertissant la solution de la fonction candidate de Lyapunov en un problème d'optimisation convexe. Dans une veine similaire, Briat and Seuret (2012) ont introduit une approche basée sur des fonctions pour l'analyse de la stabilité des systèmes impulsifs linéaires. Ils ont présenté une fonction Lyapunov en boucle, qui a démontré une efficacité accrue dans la conservation des ressources de communication par rapport à la méthode fonctionnelle Lyapunov conventionnelle, facilitant ainsi les contraintes liées aux périodes d'échantillonnage. Étendant davantage cette recherche, un modèle de contrôle basé sur un observateur flou Takagi-Sugeno (T-S) a été développé pour les parcs éoliens (WP) basés sur des générateurs à induction double alimentation (DFIG), fournissant un cadre robuste pour garantir la résilience du système (Z. Hu et al. 2022).

Dans le domaine de l'apprentissage supervisé, l'arbre de décision (DT) et la machine à vecteurs de support (SVM) sont reconnus comme des méthodes de classification éminentes. P. Wang and Govindarasu (2020) ont développé un modèle innovant de machine à vecteurs de support intégré à un arbre de décision multicouche (SVMLDT), spécifiquement adapté à la détection d'anomalies dans les systèmes électriques. Ce modèle a montré son efficacité dans la classification des états opérationnels des systèmes électriques et l'optimisation des taux de détection pour les attaques DoS.

De plus, des solutions innovantes telles que le schéma de contrôle

à temps fini distribué (DFTC) proposé par Ziqiang Wang and Jie Wang (2019) intègrent un système de stockage d'énergie distribué (DESS) pour améliorer la stabilité des systèmes électriques sous attaques DoS. Chlela et al. (2018) ont utilisé un système de stockage d'énergie (ESS) comme générateur distribué isochrone actif dans les micro-réseaux insulaires, atténuant efficacement les impacts des attaques DoS. De plus, Farraj, Hammad, and Kundur (2018) ont introduit une approche de contrôle de linéarisation de retour paramétrique (PFL), qui améliore la tolérance au retard du système électrique et maintient la stabilité transitoire après une attaque DoS.

Cherchant une approche plus simple mais efficace, nous suggérons d'adopter la théorie des échelles temporelles. En appliquant la transformation continue vers l'échelle temporelle comme décrit dans Chapter 10, nous développons un contrôleur et un observateur qui maintiennent la stabilité du système en boucle fermée sous attaques DoS, traitant ces attaques comme des retards. La flexibilité de cette méthode dans l'ajustement du temps d'échantillonnage pour le contrôleur et l'observateur dans une certaine plage garantit des réponses précises du système à l'arrivée des paquets, à condition que le retard ne dépasse pas un maximum prédéfini. Par conséquent, cette stratégie protège efficacement le système contre les attaques DoS moins sévères.

Chapter 9

Détection de l'injection de fausses données à l'aide d'observateurs fonctionnels

9.1 Introduction

L'avènement des réseaux intelligents annonce une nouvelle ère dans la gestion de l'énergie. Les réseaux intelligents sont un réseau électrique dynamique et adaptable, capable d'intégrer une variété de sources d'énergie, y compris des sources renouvelables telles que le solaire et l'éolien. Cette intégration est cruciale pour un avenir énergétique durable, car elle réduit la dépendance aux combustibles fossiles et diminue les émissions de carbone (Office of the National Coordinator for Smart Grid Interoperability 2012). Néanmoins, la variabilité inhérente et l'imprévisibilité des sources d'énergie renouvelables posent des défis uniques dans la gestion du réseau, nécessitant des systèmes de contrôle avancés pour équilibrer l'offre et la demande.

Cependant, la transition vers les réseaux intelligents apporte également d'importants défis en matière de cybersécurité (Knowles et al. 2015). La nature interconnectée et complexe de ces réseaux les rend vulnérables aux cyberattaques. Ces attaques peuvent aller du vol de données de consommation à la déstabilisation de l'ensemble du système énergétique. À mesure que les réseaux intelligents deviennent plus répandus, garantir leur sécurité contre les menaces cybernétiques est d'une importance capitale.

Plusieurs types de cyberattaques peuvent cibler les réseaux intelligents, chacun avec ses méthodes et impacts uniques. Les attaques modifiant la

charge, par exemple, impliquent des attaquants qui fournissent de fausses informations aux systèmes de contrôle du réseau, conduisant à des décisions incorrectes dans la distribution de l'énergie. De telles attaques peuvent provoquer des pannes de courant généralisées et avoir de graves implications économiques et de sécurité (Miller et al. 2021; Kimani, Oduol, and Langat 2019).

Plus généralement, les attaques par injection de fausses données sont également préoccupantes. Dans ces attaques, les pirates modifient subtilement les données transmises au sein du réseau, rendant difficile la détection de toute anomalie (Pedramnia and Shojaei 2020; Xiong et al. 2020; H. Shi, L. Xie, and Peng 2021; Zhiwen Wang, J. Hu, and Sun 2020; Khazaei and Amini 2021).

Les attaques de topologie, où les attaquants induisent en erreur les opérateurs concernant la disposition physique du réseau, peuvent conduire à des réponses inefficaces ou même dangereuses aux conditions du réseau (Zhenhua Wang, H. He, et al. 2021; Liberati, Garone, and Giorgio 2021; Liberati, Garone, and Giorgio 2021). Les attaques de redistribution de charge et les attaques ciblant les opérations de marché, comme la manipulation des informations de tarification envoyées aux compteurs intelligents, posent également des menaces à la stabilité et à l'intégrité financière du réseau (Choeum and Choi 2021; Z. Liu and L. Wang 2021; Kaviani and Hedman 2021).

Aborder ces menaces de cybersécurité nécessite une approche multifacette.

- D'un point de vue technologique de l'information (IT), les mesures de sécurité standard telles que les pare-feu, le chiffrement et le contrôle d'accès sont essentielles. Cependant, celles-ci ne sont pas suffisantes seules, car les attaquants peuvent trouver des moyens de contourner de telles défenses, y compris en exploitant l'accès interne (Mrabet et al. 2018).

- D'un point de vue de l´Automatique, le développement de systèmes de contrôle et d'observation résilients est crucial. Ces systèmes peuvent détecter les anomalies dans l'opération du réseau et répondre efficacement pour atténuer l'impact des attaques. Les techniques de détection avancées, telles que l'utilisation d'observateurs et de générateurs de résidus, sont vitales pour identifier et répondre aux cyberattaques (Pham, Oo, and Hieu Trinh 2021; Islam, Lim, and P. Shi 2020;

H. M. Tran and H. Trinh 2016; H. M. Tran, H. Trinh, and Nam 2015; H. Trinh et al. 2013).

Les techniques traditionnelles de détection des attaques, telles que celles basées sur les filtres de Kalman, rencontrent des limitations dans le contexte des réseaux intelligents. Le grand nombre d'états dans un réseau intelligent, parfois des centaines, peut entraîner des problèmes numériques et une inefficacité de calcul dans ces techniques (Chen 2013). Par conséquent, il y a un besoin croissant pour des méthodes de détection plus efficaces et évolutives. D'autres méthodes, comme celles présentées dans Section 8.3.1, présentent des problèmes numériques potentiels dus à l'utilisation de techniques telles que la décomposition en valeurs singulières et le calcul de l'espace nul sur de grandes matrices creuses (Arthur N. Montanari and Aguirre 2020).

La théorie des graphes présente une piste interessante pour renforcer la sécurité des réseaux intelligents. En analysant la topologie du réseau, la théorie des graphes peut fournir des informations sur la stabilité, la contrôlabilité et l'observabilité du système. Cette approche est particulièrement bien adaptée aux systèmes grands et creux comme les réseaux intelligents, évitant les problèmes numériques associés aux méthodes traditionnelles (Aguirre, Portes, and Letellier 2018; Cowan et al. 2012). Lorsqu'elle est appliquée à travers des observateurs fonctionnels, elle offre un équilibre entre précision et efficacité de calcul, essentiel pour la gestion en temps réel du réseau.

Dans ce chapitre, nous définissons d'abord les modèles pour le système et l'attaque par injection de fausses données. Ensuite, nous introduisons le concept d'observateurs fonctionnels et faisons une exposition détaillée sur la synthèse de l'observateur fonctionnel, articulant l'approche novatrice proposée. Nous poursuivons en illustrant le développement d'une suite complète de générateurs de résidus, essentiels pour la détection et l'isolation des cyberattaques. Nous présentons ensuite une validation par simulation, démontrant l'application pratique et l'efficacité de la méthodologie de détection proposée.

Dans la section suivante, les modèles relatifs au réseau intelligent et à l'injection de fausses données (FDI) sont détaillés. Ces modèles seront par la suite utilisés dans la conception du système de détection.

9.2 Modélisation du système et des attaques

Dans la modélisation de la dynamique des réseaux électriques, une approche répandue implique l'utilisation d'oscillateurs de Kuramoto à second ordre couplés, comme détaillé dans les études de Dorfler, Chertkov, and Bullo (2013) and Nishikawa and Motter (2015). Ce modèle représente efficacement les phases et les fréquences des oscillateurs, en supposant une fréquence d'équilibre à travers le réseau, un concept applicable aux systèmes électriques. Chaque oscillateur dans ce cadre est défini par l'équation :

$$\frac{2H_i}{\omega_R}\ddot{\phi}_i + \frac{D_i}{\omega_R}\dot{\phi}_i = A_i + \sum_{j=1,j\neq i}^{N} K_{ij}\sin(\phi_j - \phi_i + \gamma_{ij}), \tag{9.1}$$

Dans cette équation, N désigne le nombre total de nœuds ; $\phi_i(t)$ représente l'angle de phase du $i^{\text{ème}}$ oscillateur, ajusté à un cadre tournant à la fréquence de référence ω_R ; H_i et D_i sont les constantes d'inertie et d'amortissement, respectivement ; A_i est lié à l'injection de puissance au nœud i ; K_{ij} correspond au poids de couplage associé à la capacité de transfert de puissance maximale entre les nœuds i et j ; et γ_{ij} est le décalage de phase.

Lorsqu'exprimé sous forme d'espace d'états, le système est représenté comme suit:

$$\begin{bmatrix} \dot{\phi}_G \\ \ddot{\phi}_G \\ \dot{\phi}_L \end{bmatrix} = \begin{bmatrix} \dot{\phi}_G \\ \frac{\omega_R}{2H(A-\frac{D}{\omega_R})\dot{\phi}_G + \mathcal{S}} \\ \frac{\omega_R(A+\mathcal{S})}{D} \end{bmatrix}, \tag{9.2}$$

$$\mathcal{S} = \sum_{j=1,j\neq i}^{N} K_{ij}\sin(\phi_j - \phi_i + \gamma_{ij}), \tag{9.3}$$

Dans cette représentation, ϕ_G est relatif aux générateurs, ϕ_L aux charges et A à la production d'énergie dans l'ensemble du réseau. Les matrices sont partitionnées pour s'aligner avec cette division générateur-charge. Une version linéarisée de ce système est utilisée dans notre étude[1].

Pour identifier les attaques modifiant la charge, le modèle comprend trois variations :

[1] Le code utilisé pour linéariser le système est disponible à l'adresse `https://github.com/acristoffers/SmartGrid/blob/master/ca_pg.m#L40`

$$\tilde{\phi}_j = \phi_i, \tag{9.4}$$
$$\tilde{\phi}_j = \phi_j + \delta, \tag{9.5}$$
$$\tilde{\phi}_j = \phi_j \cdot \alpha, \tag{9.6}$$

Dans ces équations, ϕ désigne la valeur réelle de l'état, et $\tilde{\phi}$ représente la valeur mesurée. Le premier type d'attaque implique le remplacement d'une mesure par la valeur d'un autre état, comme dans l'Eq. (9.4). Le deuxième type ajoute un biais constant à la mesure (Eq. (9.5)), tandis que le troisième multiplie la mesure par une constante (Eq. (9.6)).

Les réseaux électriques comprennent de nombreux nœuds, généralement en nombre de centaines, présentant une inter-connectivité faible. Les transformateurs, par exemple, sont capables de former des sous-réseaux partiellement isolés où le flux d'informations est unidirectionnel. Cela résulte en une structure de graphe dynamique et clairsemée. De telles caractéristiques posent des défis pour les méthodes traditionnelles, comme détaillé dans la Section 8.3.1, ce qui nous a conduit à explorer l'applicabilité des Observateurs Fonctionnels, dont un aperçu complet est donné dans la section suivante.

9.3 Observateur Fonctionnel

Luenberger (1966) a initialement proposé le concept d'observateurs fonctionnels. Ces observateurs se caractérisent par leur capacité à estimer une combinaison linéaire sélectionnée des variables d'état d'un système à comportement dynamique. Le principal défi dans la conception de tels observateurs réside dans la détermination des états essentiels à observer et la construction des matrices système correspondantes.

Considérons un système dynamique défini par l'ensemble des équations suivantes :

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t) + Lf(t), \\ y(t) &= Cx(t), \\ z(t) &= Fx(t), \end{aligned} \tag{9.7}$$

Dans ces équations, les matrices $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times p}$, et $C \in \mathbb{R}^{q\times n}$ représentent les paramètres du système. La matrice $L \in \mathbb{R}^{n\times r}$ mappe les influences

externes $f(t) \in \mathbb{R}^r$ au vecteur d'état $x \in \mathbb{R}^n$. De plus, la matrice $F \in \mathbb{R}^{s \times n}$ transforme le vecteur d'état en le vecteur d'état estimé $z(t) \in \mathbb{R}^s$. Notamment, les matrices L et F sont structurées de manière à contenir une seule entrée non nulle par colonne et rangée, respectivement, correspondant aux états affectés et souhaités.

Il est pertinent de noter que du point de vue du système, $z(t)$ est simplement une sortie, semblable à $y(t)$. Cependant, il existe une différence sémantique distincte : $y(t)$ représente une sortie tangible, directement liée aux capteurs physiques au sein du système, tandis que $z(t)$ est une construction théorique, une combinaison linéaire d'états destinée à être estimée par l'observateur. Pour les besoins de cette analyse, $y(t)$ et $z(t)$ sont des correspondances directs de $x(t)$, impliquant que les matrices C et F possèdent une seule entrée non nulle par rangée, chacune de valeur unitaire.

Pour estimer précisément $z(t)$ à partir de $y(t)$, il est impératif d'observer un sous-ensemble de $x(t)$ qui englobe $z(t)$. Cette nécessité découle de la dynamique inhérente du système. Un observateur conçu uniquement autour de $z(t)$ est susceptible de manquer la dynamique requise pour une estimation précise, ce qui nous amène à la Definition 2, qui définit l'observabilité fonctionnelle.

Definition 2. *Un système caractérisé par l'équation* (9.7) *est considéré comme fonctionnellement observable par rapport au triplet (A, C, F) si, pour tout état initial $x(0)$, la connaissance de $u(t)$ et $y(t)$ suffit pour estimer $z(0) = Fx(0)$ sur un intervalle de temps fini $t > 0$. En revanche, le système est considéré comme fonctionnellement inobservable si cette condition n'est pas remplie (Jennings, Fernando, and H. M. Trinh 2011).*

Cette définition de l'observabilité fonctionnelle est parallèle au concept général d'observabilité dans la théorie des systèmes, bien qu'avec un focus plus étroit. Spécifiquement, un observateur fonctionnel ne peut pas discerner un état qui est inobservable à travers la sortie du système. Cela conduit à l'énonciation du Theorem 5, qui fournit une condition de rang pour cerifier l'observabilité fonctionnelle du triplet (A, C, F).

Theorem 5. *Le triplet (A, C, F) satisfait la condition d'observabilité fonc-*

tionnelle si et seulement si (Jennings, Fernando, and H. M. Trinh 2011)

$$rank \begin{bmatrix} C \\ CA \\ F \\ FA \end{bmatrix} = rank \begin{bmatrix} C \\ CA \\ F \end{bmatrix}. \tag{9.8}$$

La raison d'être de ce théorème est qu'un observateur fonctionnel peut, au mieux, observer les états discernables à travers la sortie du système. Par conséquent, les états observables doivent former un sous-ensemble des sorties, se manifestant comme des combinaisons linéaires de celles-ci. La condition de rang garantit que chaque état dans $z(t)$ est observable à travers $y(t)$. Une preuve mathématique plus approfondie est présentée dans (Jennings, Fernando, and H. M. Trinh 2011).

Identifier les états requis pour l'observation nécessite de définir l'ensemble de capteurs et les états désirés. Dans le contexte du développement d'un générateur de résidus, il est optimal que $y(t)$ soit un sous-ensemble de $z(t)$, permettant ainsi le calcul des erreurs d'estimation. Néanmoins, $z(t)$ peut devoir être légèrement plus étendu que $y(t)$ pour empêcher l'observateur de simplement répliquer $y(t)$ dans $z(t)$.

À notre connaissance, il n'existe aucune méthode algébrique pour identifier les états nécessaires. Par conséquent, les algorithmes de théorie des graphes sont devenus prévalents à cet effet, particulièrement en raison de leur évolutivité dans les grands systèmes. L'Algorithm 2, telle que proposée par Arthur Noronha Montanari (2021), décrit une méthodologie pour déterminer l'ensemble des états qui doivent être observés pour un $z(k)$ donné.

En termes généraux, l'algorithme fonctionne comme suit :

1. Transforme le système en une représentation graphique, où chaque nœud symbolise un état distinct et les arêtes interconnectées reflètent la dynamique inhérente au système.

2. Identifie un chemin reliant l'état qui nécessite une estimation à une sortie correspondante, clarifiant ainsi la route par laquelle les informations circulent.

3. Intègre les états essentiels à ce chemin identifié dans la compilation des états destinés à l'estimation, élargissant ainsi la portée de l'observabilité du système.

Algorithm 2 Méthode pour Déterminer l'Ensemble d'États Fonctionnellement Observable

input: triplet (A,C,S_0)
output: ensemble S des états nécessaires pour observer S_0
let $F \leftarrow$ matrice pour S_0, $\mathcal{M}_1 \leftarrow \emptyset$, $\mathcal{M}_2 \leftarrow \emptyset$, $r_0 \leftarrow \text{rang}(F)$
repeat
    let $G \leftarrow \begin{bmatrix} C^\top & (CA)^\top & F^\top \end{bmatrix}^\top$
    construire un graphe biparti $\mathcal{B}(\mathcal{V}, \mathcal{X}, \mathcal{E}_\mathcal{V}, \mathcal{X})$, où $\mathcal{V} = \{v_1, \ldots, v_{2q+r_0}\}$ est un ensemble de nœuds où chaque élément correspond à une rangée de $\mathcal{G}$, $\mathcal{X} = \{x_1, \ldots, x_n\}$ est l'ensemble des nœuds d'état (où chaque élément correspond également à une colonne de $\mathcal{G}$), et (v_i, x_j) est une arête non dirigée dans $\mathcal{E}_\mathcal{V}$ si $\mathcal{G}_{ij}$ est une entrée non nulle;
    trouver l'ensemble de correspondance maximale $\mathcal{E}_m$ associé à $\mathcal{B}(\mathcal{V}, \mathcal{X}, \mathcal{E}_\mathcal{V}, \mathcal{X})$ (par exemple, via l'algorithme de Hopcroft-Karp);
    $\forall x_i \in \mathcal{X}$, si x_i est connecté à une arête dans $\mathcal{E}_m$, alors mettre à jour l'ensemble des nœuds appariés à droite $\mathcal{M}_1 \leftarrow \mathcal{M}_1 \cup \{x_i\}$;
    définir l'ensemble des nœuds candidats $\mathcal{C} = \mathcal{M}_2 \setminus \mathcal{M}_1$, où $x_j \in \mathcal{M}_2$ si $[FA]_{ij}$ est une entrée non nulle;
    sélectionner un élément $x_k \in \mathcal{C}$ et mettre à jour $F \leftarrow \begin{bmatrix} F^\top & (F')^{top} \end{bmatrix}$ et $r_0 = r_0 + 1$, où $F' \in \mathbb{R}^{1 \times n}$ et $[F']_{ij} = 1$ si $j = k$ et 0 sinon;
until $\mathcal{C} \neq \emptyset$

En utilisant les concepts détaillés dans cette section, une conception pour un observateur fonctionnel est introduite dans la section suivante.

9.4 Synthèse de l'observateur

Dans le défi que nous nous sommes fixé de détecter des attaques modifiant la charge dans les réseaux électriques intelligents, notre recherche introduit une approche novatrice : un ensemble d'observateurs fonctionnels (e Sousa, Messai, and Manamanni 2022). Ces observateurs se distinguent par leur conception d'ordre réduit, ce qui réduit les exigences de calcul.

Le cadre proposé comprend r observateurs fonctionnels, chacun aligné avec un nœud spécifiquement attaqué au sein du réseau intelligent. Bien que ces observateurs partagent une fondation structurelle commune, ils diffèrent dans l'application de la matrice système L. Cette matrice est cruciale pour lier l'attaque spécifique à l'état correspondant. La configuration unique de chaque observateur, adaptée aux capteurs attaqués individuellement, permet une isolation précise de l'attaque. De plus, chaque observateur est équipé de son propre générateur de résidu. Ce générateur est essentiel pour déterminer l'erreur d'estimation, qui à son tour est cruciale pour identifier la présence d'une attaque. La Figure 9.1 illustre la représentation schématique de l'assemblage complet des observateurs et de leurs générateurs de résidus associés.

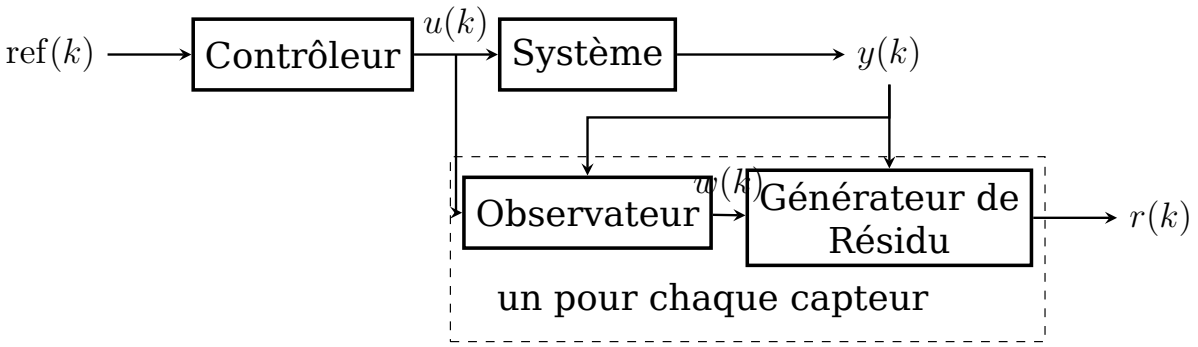


Figure 9.1 – Schéma de l'observateur et du système générateur de résidus

Nous préconisons une conception d'observateur basée sur les LMI. Cette formulation est avantageuse en raison de sa flexibilité inhérente, permettant l'intégration de contraintes supplémentaires dans le système simplement en élargissant le cadre LMI.

D'autres approches pour la conception d'observateurs fonctionnels existent dans la littérature. Islam, Lim, and P. Shi (2020) introduit une méthodologie adaptée aux systèmes flous, offrant une perspective nuancée sur la gestion des systèmes caractérisés par l'incertitude et l'imprécision. De même, Rios-Ruiz et al. (2019) délimite une stratégie de conception

pour les systèmes descripteurs, qui se distinguent par leurs contraintes algébriques en plus des équations différentielles. Dans le domaine des systèmes influencés par des retards temporels, Hieu Manh Tran and Hieu Trinh (2019) propose une formulation spécifiquement pour les systèmes à retards temporels, abordant les complexités introduites par de tels retards. Haes Alhelou, Golshan, and Hatziargyriou (2019) contribue au discours en proposant une conception robuste d'observateur, axée sur la résilience et l'adaptabilité sans aborder directement l'identification des attaques. Enfin, Pham, Oo, and Hieu Trinh (2021) s'aventure dans le domaine des véhicules électriques, présentant une technique pour la détection et l'isolement des anomalies, soulignant ainsi la pertinence des observateurs fonctionnels pour assurer l'intégrité opérationnelle des solutions de mobilité électrique de plus en plus répandues.

Néanmoins, les méthodologies discutées partagent un principe fondamental : la partition et la décomposition des systèmes. Ce processus emploie fréquemment des techniques telles que la décomposition en valeurs singulières et le calcul des espaces nuls de matrices. Malgré leur utilité, ces méthodes rencontrent des défis numériques lorsqu'appliquées à de grandes matrices creuses, comme souligné par Chen (2013). De plus, l'adoption de LMI dans l'approche proposée facilite l'intégration de contraintes supplémentaires au sein du système. Cette capacité offre aux concepteurs un contrôle accru sur la performance du système, permettant une conception d'observateur plus adaptée et efficace.

La conception de notre observateur suit le théorème suivant (e Sousa, Messai, and Manamanni 2022):

Theorem 6. *Étant donné un système tel que défini dans l'équation* (9.7)*, et en considérant le triplet (A, C, F) observable fonctionnellement conformément à la Definition 2 et au Theorem 5, un observateur peut être formulé comme suit :*

$$\begin{aligned} \dot{w}(t) &= Nw(t) + Jy(t) + Hu(t), \\ \hat{z}(t) &= w(t) + Ey(t), \end{aligned} \tag{9.9}$$

Cet observateur est capable d'estimer $\hat{z}(t) \approx z(t)$ étant donné $y(t)$ et $u(t)$, à condition que, pour la fonction candidate de Lyapunov

$$V(x) = x^\top Px, \tag{9.10}$$

où P est une matrice définie positive, une solution existe pour les LMI suiv-

ant:

$$\begin{aligned} &\textit{arg min } \|P\|_2 \\ &\textit{s.t. } \dot{V} < 0 \\ &P \succ 0, \end{aligned} \tag{9.11}$$

où

$$\dot{V} < 0 \implies \begin{bmatrix} X & W \\ W^\top & -I \end{bmatrix} \prec 0, \tag{9.12}$$

avec

$$X = \hat{A}^\top F^\top P - \hat{A}^\top C^\top \hat{E}^\top - \hat{C}^\top \hat{K}^\top + PF\hat{A} - \hat{E}C\hat{A} - \hat{K}\hat{C} - \lambda I, \tag{9.13}$$

$$W = \sqrt{\lambda}(PF - \hat{E}C), \tag{9.14}$$

$$\lambda \in \mathbb{R}^+ \textit{ est une constante libre liée à l'ampleur de l'attaque}, \tag{9.15}$$

$$\tag{9.16}$$

où

$$\hat{A} = AF^+, \tag{9.17}$$

$$\hat{C} = CF^+, \tag{9.18}$$

$$\hat{E} = PE = PU + \hat{Y}V, \tag{9.19}$$

$$\hat{K} = PK, \tag{9.20}$$

$$\hat{Y} = PY. \tag{9.21}$$

Les valeurs connues sont alors les matrices du système et λ, et toutes les autres variables sont des variables d'optimisation.

Les matrices de l'observateur sont récupérées comme suit :

$$K = P^{-1}\hat{K}, \tag{9.22}$$

$$Y = P^{-1}\hat{Y}, \tag{9.23}$$

$$E = U + YV, \tag{9.24}$$

$$R = F - EC, \tag{9.25}$$

$$N = (RA - KC)F^+, \tag{9.26}$$

$$J = K + NE, \tag{9.27}$$

$$H = RB, \tag{9.28}$$

Proof. Commencez par définir l'erreur d'estimation e comme l'écart entre l'état estimé $\hat{z}$ et l'état réel z. Dans un soucis de simplification, les dépendances temporelles dans la notation vectorielle sont omises :

$$e = \hat{z} - z = w + Ey - Fx = w + ECx - Fx. \tag{9.29}$$

Ensuite, la dynamique de cette erreur d'estimation est calculée comme suit:

$$\begin{aligned}\dot{e} &= \dot{w} + (EC - F)\dot{x} \\ &= Nw + Jy + Hu + (EC - F)(Ax + Bu + Lf) \\ &= N(e + Fx - ECx) + JCx + Hu + ECAx \\ &+ ECBu + ECLf - FAx - FBu - FLf \\ &= Ne + (NF - NEC + ECA - FA + JC)x \\ &+ (H + ECB - FB)u + (ECL - FL)f. \end{aligned} \tag{9.30}$$

Pour que l'état estimé $\hat{z}$ converge vers l'état réel z, N doit être stable au sens de Hurwitz, et les conditions suivantes doivent être satisfaites :

$$N(F - EC) - (F - EC)A + JC = 0, \tag{9.31}$$

$$H - (F - EC)B = 0, \tag{9.32}$$

$$(F - EC)L = 0. \tag{9.33}$$

Pour adapter l'observateur à être sélectivement insensible à certaines attaques de capteurs, nous modifions cette dernière condition. Laissons $L = \begin{bmatrix} L_i & L_n \end{bmatrix}$, avec $L_i \in \mathbb{R}^{n\times 1}$ représentant la cartographie de l'attaque insensible, et $L_n \in \mathbb{R}^{n\times l-1}$ correspondant aux attaques sensibles. Par conséquent, nous établissons les critères suivants pour chaque observateur :

$$(F - EC)L_i = 0, \tag{9.34}$$

$$(F - EC)L_n \neq 0. \tag{9.35}$$

L'étape suivante consiste à déterminer les matrices N, J, H et E. Ceci est réalisé par l'optimisation d'un LMI. Nous introduisons la fonction candidate de Lyapunov :

$$V = e^\top Pe. \tag{9.36}$$

Intégrant les restrictions mentionnées ci-dessus, la dynamique de la fonction d'erreur devient :

$$\dot{e} = Ne - (F - EC)L_n f. \tag{9.37}$$

Pour conceptualiser l'attaque en termes d'erreur, nous supposons une relation de proportionnalité entre l'erreur et la variation introduite par l'attaque:

$$e \propto L_n f, \tag{9.38}$$

menant à l'équation :

$$\|L_n f\| = \lambda \|e\|, \tag{9.39}$$

où $\lambda \in \mathbb{R}^+$ est un paramètre librement ajustable qui échelonne l'attaque proportionnellement à l'erreur.

Avec cette formulation, et en définissant

$$R = F - EC, \tag{9.40}$$

la dérivée de la fonction candidate de Lyapunov s'exprime comme suit :

$$\begin{aligned}
\dot{V} &= \dot{e}^\top P e + e^\top P \dot{e} \\
&= (Ne - \lambda R\|e\|)^\top P e + e^\top P (Ne - \lambda R\|e\|) \\
&= e^\top (N^\top P + PN) e - 2\lambda \|e^\top P R\| \cdot \|e\| \\
&\leq e^\top (N^\top P + PN) e - \lambda (\|e^\top P R\|^2 + \|e\|^2) \\
&= e^\top (N^\top P + PN - \lambda P R R^\top P - \lambda I) e,
\end{aligned} \tag{9.41}$$

où I désigne la matrice identité de dimensions appropriées.

Cette inégalité matricielle bilinéaire (BMI) assure la stabilité de la matrice N. Pour intégrer la contrainte énoncée dans l'équation (9.31), nous définissons :

$$N(F - EC) = RA - JC, \tag{9.42}$$

$$NF = RA - (J - NE)C, \tag{9.43}$$

$$K = J - NE, \tag{9.44}$$

$$N = RAF^+ - KCF^+, \tag{9.45}$$

où $K \in \mathbb{R}^{s \times q}$ est une matrice déterminée par optimisation et F^+ représente l'inverse de Moore-Penrose (également connu sous le nom d'inverse pseudo) de la matrice F.

Pour accommoder la condition de (9.34), les équations suivantes sont posées :

$$(F - EC)L_i = 0, \tag{9.46}$$

$$ECL_i = FL_i, \tag{9.47}$$

$$E = FL_i(CL_i)^+ + Y(I - (CL_i)(CL_i)^+), \tag{9.48}$$

$$U = ECL_iL_i^+, \tag{9.49}$$

$$V = I - L_iL_i^+, \tag{9.50}$$

$$E = U + YV, \tag{9.51}$$

revisitant les équations de (9.17) à (9.21).

En conséquence, l'équation (9.41) se transforme en :

$$\dot{V} = e^\top ((R\hat{A} - EC\hat{A} - K\hat{C})^\top P + P(R\hat{A} - EC\hat{A} - K\hat{C}) - \lambda P R R^\top P - \lambda I) e. \tag{9.52}$$

Après une expansion et une substitution complètes des variables, la dérivée de la fonction candidate se résout finalement à :

$$\dot{V} = e^\top(\hat{A}^\top F^\top P - \hat{A}^\top C^\top \hat{E}^\top - \hat{C}^\top K^\top + PF\hat{A} - \hat{E}C\hat{A} - K\hat{C} - \lambda PRR^\top P - \lambda I)e. \tag{9.53}$$

Enfin, pour répondre à la nature bilinéaire de $\dot{V}$, résultant de la variable R, la technique du complément de Schur est employée, comme référencé dans Boyd et al. (1994).

$$X = \hat{A}^\top F^\top P - \hat{A}^\top C^\top \hat{E}^\top \hat{C}^\top K^\top + PF\hat{A} - \hat{E}C\hat{A} - K\hat{C} - \lambda I, \tag{9.54}$$

$$W = \sqrt{\lambda}(PF - \hat{E}C), \tag{9.55}$$

$$\dot{V} = e^\top \begin{bmatrix} X & W \\ W^\top & -I \end{bmatrix} e. \tag{9.56}$$

□

L'erreur d'estimation au sein de ce système d'observateur est définie comme suit:

$$e = w + ECx - Fx. \tag{9.57}$$

La dynamique de cette erreur est ensuite élucidée, menant à un ensemble de conditions qui garantissent la convergence de l'état estimé $\hat{z}$ vers l'état réel z. Ces conditions incluent la stabilité de Hurwitz de la matrice N et d'autres contraintes spécifiques détaillées dans les équations (9.31), (9.32), et (9.33). Notamment, pour atteindre une insensibilité à une attaque de capteur particulière tout en restant réactif aux autres, nous redéfinissons la condition (9.33) comme illustré dans les équations (9.34) et son complément.

Par la suite, les matrices N, J, H et E sont déterminées par un processus d'optimisation LMI, intégrant la fonction candidate de Lyapunov et tenant compte des restrictions nécessaires. Ceci résulte en la dérivée finale de la fonction candidate présentée dans l'équation (9.41) et sa forme étendue.

Possédant un observateur qui manifeste une insensibilité aux attaques sur un seul capteur, la section suivante introduit la conception d'un générateur de résidu. Ce générateur est conçu pour offrir une méthode numérique pour évaluer la présence d'une attaque, tout en accommodant les perturbations au sein du système.

9.5 Générateurs de résidus

Un résidu se réfère à l'écart entre une valeur estimée et la valeur réelle. Dans des conditions normales, sans aucune interférence externe, il est anticipé que ce résidu se rapprochera de zéro. Au contraire, en cas d'attaque, un écart significatif par rapport à zéro est attendu. Une fonction binaire, basée sur un seuil défini, peut utiliser ce résidu pour indiquer la présence d'une attaque dans le système. Cette sous-section se concentre sur la dynamique du résidu plutôt que sur la fonction binaire.

Nous avons introduit un générateur de résidus qui utilise un observateur spécifiquement conçu. Cet observateur est insensible à un type d'attaque et génère une sortie reflétant l'erreur d'estimation, appelée résidu. Ce résidu est essentiellement une mesure de la précision de l'observateur dans l'estimation de l'état. Lors d'une attaque, qui est un signal externe non mesuré, l'estimation de l'état de l'observateur devient inexacte, entraînant un changement notable dans le résidu. Le lemme suivant encapsule ce concept :

Lemma 9. *Un signal résiduel peut être généré comme*

$$r(t) = Gw(t) + My(t), \tag{9.58}$$

avec

$$M = (C(1 - L_i))^{\top}, \tag{9.59}$$

$$G = -M(I - CF^{+}E)^{-1}CF^{+}, \tag{9.60}$$

où $1 - L_i$ *signifie une opération entrée par entrée. Cela forme un générateur de résidus pour l'observateur* (9.9)*, comme décrit dans Lemma 6.*

Proof. L'équation d'erreur pour le générateur de résidus, utilisant les matrices de l'observateur, est donnée par

$$\begin{aligned} r &= Gw + My \\ &= G(e + Fx - ECx) + MCx \\ &= Ge + (G(F - EC) + MC)x, \end{aligned} \tag{9.61}$$

Pour que le générateur de résidus converge vers zéro en l'absence d'erreur, la condition suivante doit être remplie :

$$G(F - EC) + MC = 0. \tag{9.62}$$

En exprimant l'équation d'erreur en termes de la sortie, nous avons :

$$\begin{aligned} r &= Gw + My \\ &= Q(y - Cx) \\ &= Q(y - CF^{-1}\hat{z}) \\ &= Q(y - CF^{-1}(w + Ey)) \\ &= Q((I - CF^{-1}E)y - CF^{-1}w), \end{aligned} \tag{9.63}$$

$$M = Q(I - CF^{-1}E), \tag{9.64}$$

$$G = -QCF^{-1}. \tag{9.65}$$

En substituant Equations (9.64) and (9.65) dans (9.62) confirme que la condition est satisfaite quel que soit la valeur de Q. Ainsi, Q peut être utilisé pour ajuster l'amplitude du résidu. Par conséquent, M est défini en utilisant la matrice L comme dans (9.66) :

$$M = (C(1 - L_i))^\top, \tag{9.66}$$

représentant la somme de toutes les sorties $y(t)$ sujettes à l'erreur, à l'exception de celle insensible. En remplaçant l'Eq. (9.66) dans Eq. (9.64) et en éliminant Q de Eq. (9.65), nous dérivons

$$M = (C(1 - L_i))^\top, \tag{9.67}$$

$$G = -M(I - CF^{+}E)^{-1}CF^{+}. \tag{9.68}$$

Notez que les équations (9.64) et (9.65) peuvent aussi être directement appliquées en identifiant une matrice Q avec des propriétés souhaitables, telles que générer un résidu d'amplitude suffisamment grande pendant les attaques attendues. □

Le principal avantage de la technique proposée réside dans sa conception sur mesure de l'observateur, qui offre davantage de flexibilité aux concepteurs par rapport à certaines méthodes algébriques couramment utilisées dans la littérature, comme celle décrite dans Emami et al. (2013). Ces méthodes classiques, bien que souvent directes et efficaces dans certains cas, présentent des limitations lorsqu'il s'agit d'intégrer des contraintes telles que la $\mathcal{D}$-stabilité ou de prendre en compte des saturations. Notre approche basée sur les LMI est conçue pour être simple et adaptable, permettant d'inclure différentes techniques LMI dans la conception de l'observateur.

Certaines solutions existantes rencontrent des difficultés liées à la stabilité numérique, comme illustré dans Arthur Noronha Montanari (2021),

qui utilise l'inversion d'une partie de la matrice système A. Cette démarche peut devenir problématique pour les grands systèmes épars, car l'inversion de matrices éparses est techniquement complexe et peut engendrer des instabilités numériques. Bien qu'il existe des algorithmes pour effectuer ces inversions, leur performance est souvent sensible au contexte spécifique et peut ne pas être toujours fiable. Par ailleurs, des opérations comme la décomposition en valeurs singulières ou le calcul de l'espace nul peuvent poser des défis similaires de stabilité numérique pour les grands systèmes épars.

Ces considérations mettent en lumière l'intérêt potentiel des observateurs fonctionnels par rapport à certaines méthodes traditionnelles comme le filtre de Kalman (Chen 2013). Si les techniques établies restent très performantes dans l'identification de certains types d'attaques, elles montrent toutefois des limites lorsqu'il s'agit de grands systèmes épars. À l'inverse, ces méthodes pourraient être moins adaptées à des systèmes plus réduits, illustrant ainsi la nécessité de choisir une approche en fonction des caractéristiques spécifiques du système étudié.

Dans la section suivante, une simulation est présentée pour démontrer l'efficacité de la méthodologie proposée.

9.6 Exemple de Simulation

Pour démontrer l'efficacité de notre méthodologie, explorons le système de test IEEE 118 bus, un modèle représentant le réseau électrique du Midwest des États-Unis. Ce système comprend 19 générateurs, 35 condenseurs synchrones, 177 lignes de transmission, neuf transformateurs et 91 charges. Le schéma de ce réseau est illustré dans la Figure 9.2. Le fichier de données correspondant, un fichier MATLAB encapsulant les constantes pour (9.1) sous forme matricielle, est accessible sur GitHub[2]. De plus, les données au format de données commun IEEE peuvent être trouvées en ligne[3], les niveaux de base KV étant estimés en raison de leur absence dans le jeu de données original.

Nous avons simulé le modèle de réseau électrique IEEE 118 en utilisant une version linéarisée du système dynamique décrit dans (9.3). La Figure 9.3 présente le graphique dynamique du système, mettant en évidence

[2] `https://github.com/acristoffers/SmartGrid/blob/master/powergrid/IEEE118pg.mat`

[3] `https://labs.ece.uw.edu/pstca/pf118/pg_tca118bus.htm`

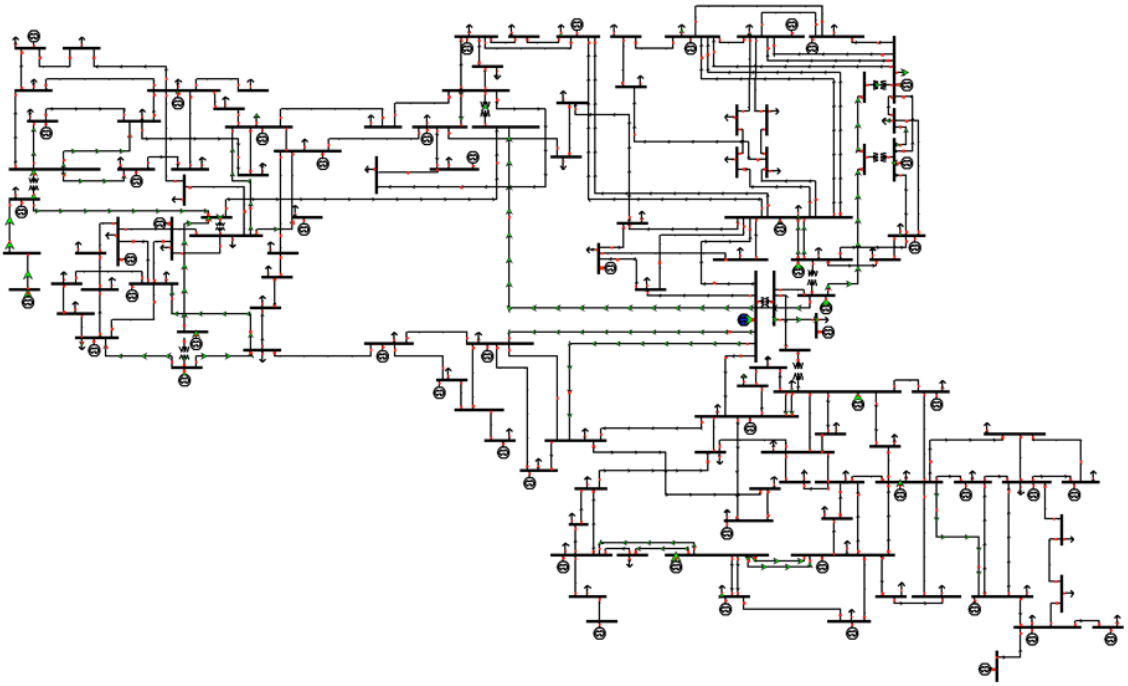


Figure 9.2 – Schéma du réseau IEEE 118 bus

les générateurs, les charges et les capteurs. Le graphique révèle une connectivité clairsemée avec un degré moyen de nœud de 2 et une centralité d'intermédiation élevée, impliquant des chemins plus longs entre les capteurs et les états désirés et la criticité des états intermédiaires pour une observation efficace. Pour surveiller les états ciblés, les observateurs doivent estimer environ 150 des 226 états totaux, ce qui représente environ 66

Le placement des capteurs a est aléatoire, avec des PMU installés à 30% des terminaux des générateurs et des charges, totalisant 35 capteurs. L'algorithme de Arthur Noronha Montanari (2021) a identifié un ensemble S de 126 états. En utilisant le Theorem 6 et le Lemma 9, nous avons créé une suite d'observateurs et de générateurs de résidus pour le système, capables d'identifier les attaques modifiant les charges.

Trois types d'attaques modifiant les charges ont été simulés : la première duplique la valeur d'un autre état au nœud attaqué, la deuxième ajoute une constante au signal existant, et la troisième multiplie la valeur de l'état par une constante. Les Figure 9.4, Figure 9.5, et Figure 9.6 affichent la sortie des générateurs de résidus pour chaque type d'attaque.

Les générateurs de résidus mis en œuvre ont réussi à détecter chaque type d'attaque. Les résidus étaient nuls avant l'attaque, se sont déplacés de zéro pendant l'attaque, et sont revenus à zéro après l'attaque. Un résidu est resté constamment proche de zéro, car son observateur correspondant

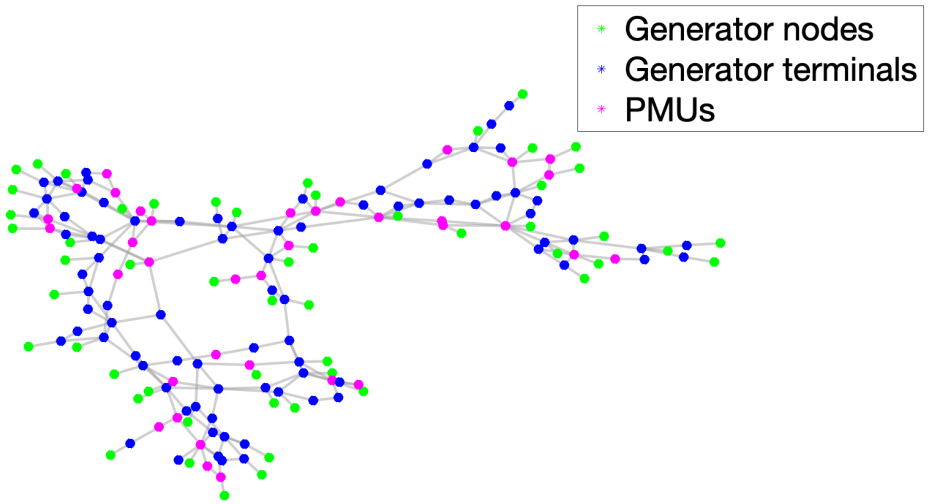


Figure 9.3 – Représentation graphique dynamique du système bus IEEE 118

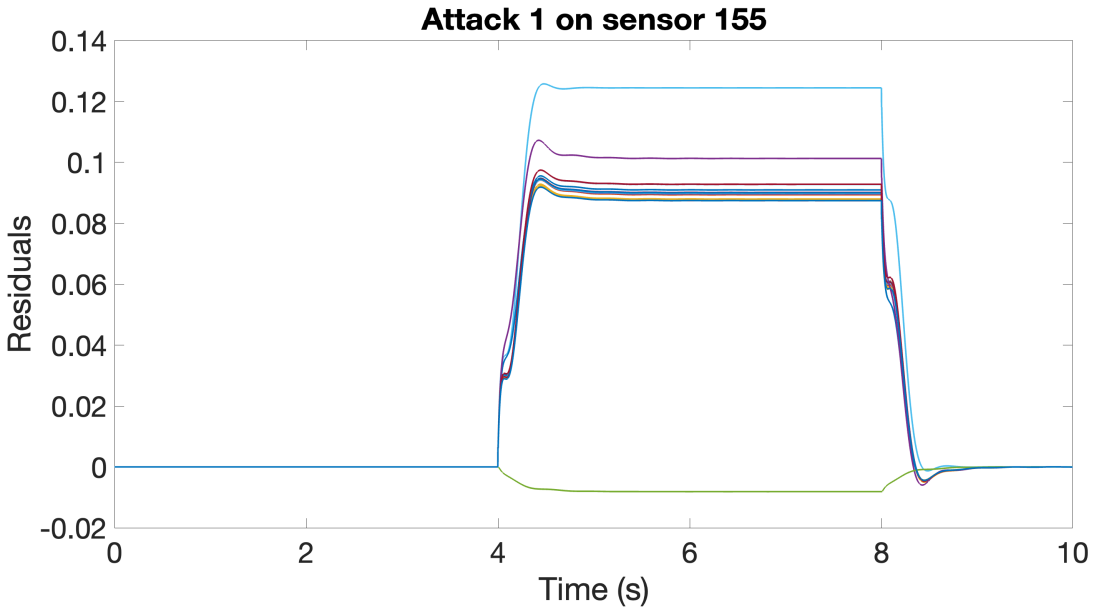


Figure 9.4 – Résidus pour une attaque de copie de valeur d'état sur l'état 155

était insensible à l'attaque sur son capteur, facilitant l'identification des occurrences d'attaque et l'isolation du nœud compromis.

Le code de simulation et toutes les matrices générées sont disponibles en ligne[4].

Nous ne présentons pas de comparaisons directes avec d'autres méth-

[4] `https://github.com/acristoffers/SmartGrid`

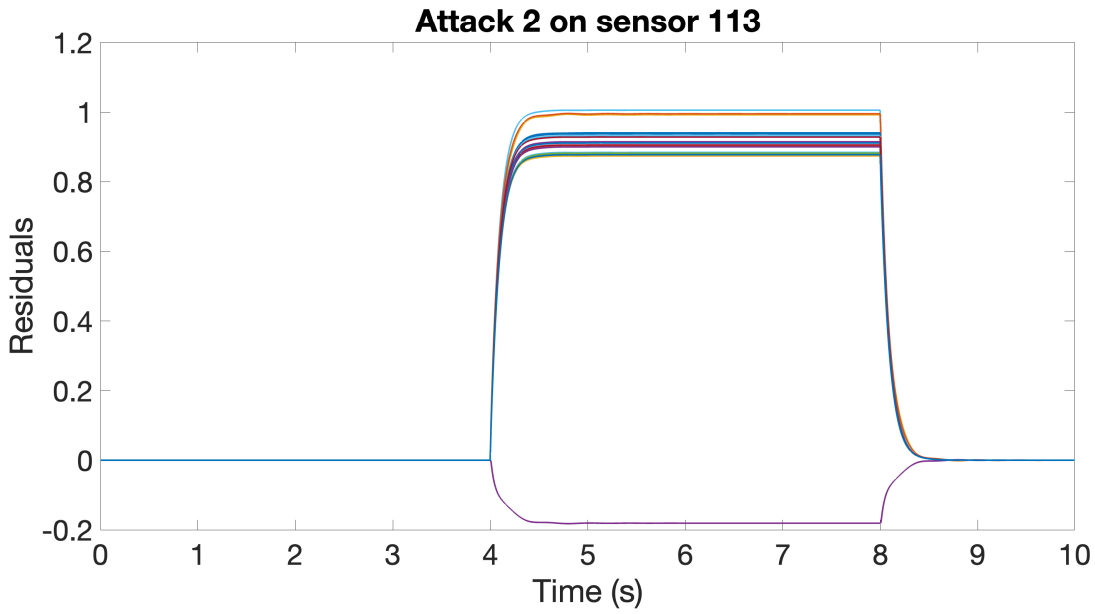


Figure 9.5 – Résidus pour une attaque additive sur l'état 113

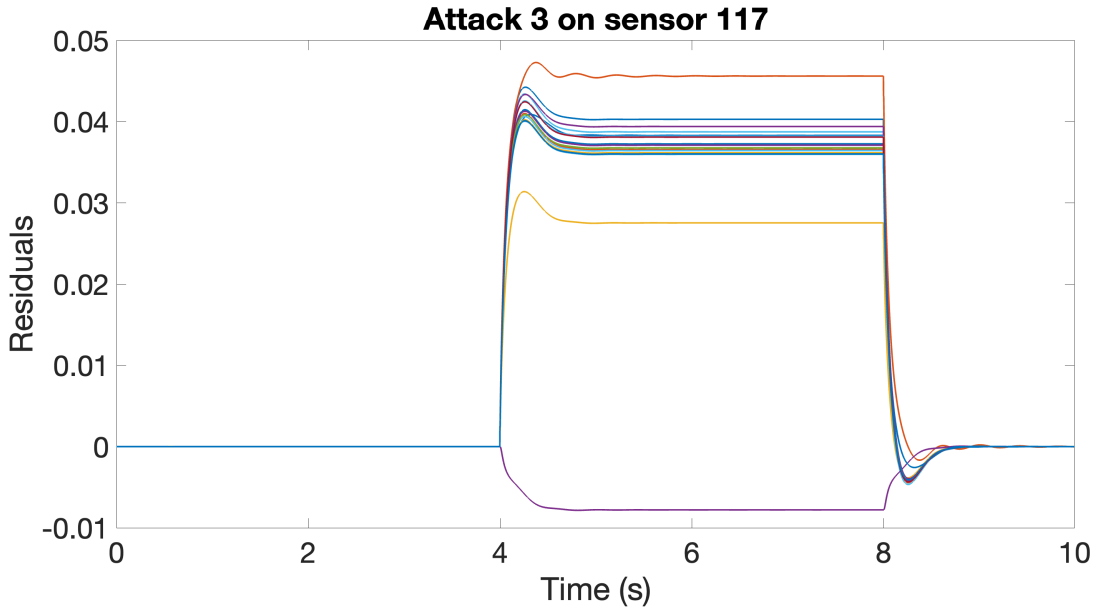


Figure 9.6 – Résidus pour une attaque multiplicative sur l'état 221

odes car la valeur de la technique proposée réside dans sa formulation, qui ne souffre pas de problèmes numériques sur des systèmes grands et clairsemés, et dans sa formulation directe. Comparer les performances n'aurait pas de sens, car toutes les méthodes ont un paramètre de réglage qui conduirait à des affirmations selon lesquelles la performance de toute méthode particulière serait meilleure avec un choix différent de paramètres, ne fournissant pas une évaluation objective.

Pour évaluer le coût computationnel de la solution proposée, le nombre

de variables et de lignes dans les LMIs est d'intérêt :

- Conception de l'observateur ($N \in \mathbb{R}^{n\times n}, \hat{K} \in \mathbb{R}^{n\times q}, \hat{Y} \in \mathbb{R}^{n\times s}$) :
  - Nombre de variables : $\frac{n(n+1)}{2} + n(q+s) + 1$
  - Nombre de lignes : $2n^2 + n$

Les générateurs résiduels proposés mettent quelques millisecondes à détecter les attaques. Cela est, cependant, dû au hasard, car la formulation n'impose aucun type de synchronisation. Cela ne veut pas dire que cela ne peut pas être contrôlé : la matrice N de l'observateur détermine sa dynamique et peut être utilisée pour contrôler le temps de convergence de l'observateur et, par conséquent, le temps de détection des attaques. Une manière de le faire est d'ajouter des contraintes à la LMI qui spécifient une région pour l'allocation des pôles, avec des méthodes telles que celles trouvées dans Duan and H.-H. Yu (2013).

9.7 Conclusion

Dans ce chapitre, nous avons introduit une conception simple basée sur les LMI pour une suite de générateurs de résidus d'observateurs fonctionnels, principalement destinés à détecter les attaques modifiant les charges. La simplicité de notre approche est un aspect remarquable, car elle facilite non seulement la compréhension et la mise en œuvre de la conception, mais offre également une flexibilité considérable pour étendre la conception afin d'incorporer diverses contraintes dans la dynamique de l'observateur. Cette flexibilité est également avantageuse dans des applications au-delà de la détection d'attaques, telles que l'observation d'état avec rejet des perturbations, la détection de défauts et potentiellement pour contrer d'autres types d'attaques cyber-physiques.

Notre méthode a été évaluée sur un modèle de réseau électrique, où elle a démontré son efficacité en permettant une réduction significative du nombre d'états nécessaires à l'observation. Cette réduction se traduit par un système d'observateur plus épuré, qui à son tour conduit à des temps de calcul plus rapides et une efficacité améliorée. Un tel résultat est particulièrement précieux dans les applications réelles où les ressources informatiques et le temps sont des facteurs critiques.

En résumé, la simplicité et l'adaptabilité de notre conception basée sur l'LMI pour les générateurs de résidus d'observateurs fonctionnels en font

non seulement un outil puissant pour détecter les attaques modifiant les charges, mais ouvrent également un large éventail d'applications pour renforcer la robustesse et la fiabilité de divers systèmes. Le potentiel de développements futurs dans ce domaine est vaste, promettant des avancées encore plus grandes dans le domaine de la sécurité et de l'efficacité des systèmes cyber-physiques.

Dans le chapitre suivant, l'accent est mis sur le Zero-Dynamics Attack (ZDA) et l'introduction de notre solution proposée basée sur la Théorie des Échelles de Temps.

Chapter 10

Détection des attaques par dynamiques nulles à l'aide d'observateurs à échelle de temps

10.1 Introduction

Dans le chapitre précédent, une étude a été menée sur l'attaque par False Data Injection (FDI), aboutissant à la proposition d'une méthodologie de détection. Ce chapitre se concentre sur la Zero-Dynamics Attack (ZDA), pour laquelle une nouvelle solution de détection est proposée.

Les ZDA nécessitent une connaissance complète du système, et est notoirement difficile à détecter en raison de sa nature furtive. Elle exploite les dynamiques des zéros de transmission du système pour manipuler le signal de commande de manière à impacter les états du système (pouvant déstabiliser le système) sans modifier la sortie du système. Par conséquent, les capteurs au sein du système restent non affectés par l'attaque, ce qui complique sa détection (André Teixeira et al. 2015).

Les techniques traditionnelles de détection présentent des limites face aux ZDA, en raison de leur impact souvent discret sur les sorties du système, comme discuté dans la Section 8.3.2. Ces limitations ont motivé notre exploration de la Théorie des Échelles Temporelles comme une approche alternative. Distincte des systèmes à deux échelles et à trois échelles documentés dans la littérature existante, la Théorie des Échelles Temporelles représente un cadre complet pour l'analyse des systèmes dynamiques, accommodant des domaines temporels arbitraires. Ce cadre encapsule les systèmes à temps continu et à temps discret comme instances particulières

au sein de son applicabilité plus large (Agarwal et al. 2002), et permet des changements arbitraires dans le temps d'échantillonnage, dépassant les limitations d'un groupe d'observateurs discrets qui sont limités à un ensemble fixe de valeurs de temps d'échantillonnage. La flexibilité du système d'échelles temporelles dans la sélection de toute valeur pour le temps d'échantillonnage améliore son efficacité dans la détection d'attaques.

Malgré son potentiel, l'approche par échelles temporelles présente des défis dans l'application directe ou l'adaptation des techniques classiques continues et discrètes, particulièrement en raison de ses éléments de modèle dépendants du temps. La praticabilité de la vérification de l'observabilité dans de tels systèmes reste un problème non résolu et fait l'objet de recherches en cours (Ben Nasser et al. 2022). Ce travail ne se penchera pas sur la vérification de l'observabilité du système, en supposant que le concepteur l'a vérifiée.

La contribution principale de ce chapitre est le développement d'un nouvel observateur capable d'ajuster dynamiquement son temps d'échantillonnage grâce à de simples variations de paramètres. Le gain de l'observateur est conçu comme une fonction convexe, impliquant uniquement des additions et des multiplications, facilitant ainsi sa mise en œuvre même dans des environnements à ressources limitées. Cette adaptabilité dans le temps d'échantillonnage rend effectivement les ZDA visibles. À notre connaissance, cette application innovante de l'approche par échelles temporelles n'a pas été explorée auparavant dans la littérature. D'autres caractéristiques significatives de notre approche incluent :

- L'utilisation d'un observateur Linear Parameter Varying (LPV) avec seulement quatre sommets, indépendamment de la taille du système, ce qui évite une croissance exponentielle du nombre de sommets.
- Le nombre de contraintes et de variables dans les Linear Matrix Inequality (LMI) dérivées croît linéairement avec le système, et ne présente pas de croissance exponentielle.
- La capacité de l'observateur à modifier son temps d'échantillonnage à tout moment introduit un élément aléatoire dans la détection des attaques, compliquant considérablement les efforts d'un attaquant pour dissimuler l'attaque.

Ainsi, la technique proposée améliore de manière notable les méthodes existantes sous deux aspects critiques : 1) elle ne nécessite pas de modifi-

cations des dynamiques du modèle du système, et 2) elle ne requiert pas de changements dans la topologie du réseau, présentant une alternative plus simple mais efficace pour la détection d'attaques à dynamique nulle.

Dans ce chapitre, nous commençons par explorer les spécificités des ZDA, en détaillant leur exécution et leur impact sur les états du système tout en maintenant un profil furtif. Ensuite, nous présentons une introduction à la Théorie des Échelles Temporelles, en soulignant ses éléments les plus importants. Après, nous introduisons une conception d'observateur novatrice utilisant la Théorie des Échelles Temporelles, offrant un cadre complet pour l'analyse des systèmes dynamiques avant de présenter des résultats de simulation pour valider l'efficacité de la technique de détection proposée.

La section suivante élucide les ZDA et délimite son application pour impacter un système.

10.2 Zero-Dynamics Attack

La ZDA est exécutée en injectant de l'énergie dans le système de manière à ce que les états du système soient soit conduits vers une zone dangereuse, soit forcés à diverger, tout en maintenant un profil furtif en ne modifiant pas la sortie observable. Le point critique de vulnérabilité se trouve dans le réseau, où l'attaque manipule les signaux de contrôle transmis du contrôleur au système physique. Cette manipulation reste indétectée par le côté contrôle en raison de la sortie inchangée, masquant ainsi la présence de l'attaque.

Étant donné l'hypothèse d'un réseau compromis, les schémas de détection conventionnels situés du côté du contrôle deviennent inefficaces, car l'attaquant pourrait facilement falsifier toute information transmise. En contre-mesure, il est conseillé de positionner le mécanisme de détection directement dans le cadre du système, en l'intégrant aux systèmes d'alarme, comme illustré dans Figure 10.1.

Malgré la nécessité d'une connaissance approfondie du système, la formulation de l'attaque est relativement simple. Considérons un système en temps continu caractérisé par

$$\dot{x}(t) = \mathcal{A}x(t) + \mathcal{B}u_a(t), \tag{10.1}$$

$$y(t) = \mathcal{C}x(t), \tag{10.2}$$

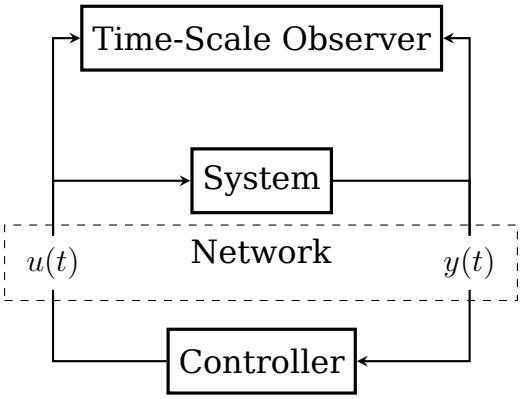


Figure 10.1 – Représentation schématique de l'intégration de l'observateur

où $u_a(t)$ représente le signal de contrôle induit par l'attaque. Les zéros de transfert du système sont définis comme les racines de $\det(P(s))$, avec $P(s)$ étant la matrice de Rosenbrock, comme défini dans Equation (10.3). La relation complexe entre la matrice de Rosenbrock et les zéros de transmission du système est élucidée dans Lemma 10.

Lemma 10. *Étant donné la matrice système de Rosenbrock*

$$P(s) = \begin{bmatrix} sI - \mathcal{A} & -\mathcal{B} \\ \mathcal{C} & D \end{bmatrix}, \tag{10.3}$$

son déterminant représente le numérateur de la réalisation de la fonction de transfert du système. Spécifiquement, la fonction de transfert est exprimée comme $G(s) = \frac{\det(P(s))}{\det(sI-\mathcal{A})}$*, où* $\det(P(s))$ *est un polynôme, potentiellement contenant des zéros annulés (Murdoch 1973).*

La direction du zéro z_0 correspondant au zéro s_0 est déterminée par

$$P(s_0)z_0 = 0. \tag{10.4}$$

Conduire le système le long de z_0 assure que toute modification des états n'est pas reflétée dans la sortie, comme souligné par Baniamerian, Khorasani, and Meskin (2020). Le signal de contrôle de l'attaque peut être formulé comme

$$z_0 = \begin{bmatrix} x_0 \\ a_0 \end{bmatrix}, \tag{10.5}$$

$$a(t) = a_0 e^{s_0 t}. \tag{10.6}$$

Le signal de contrôle modifié $u_a(t)$ est calculé comme $u_a(t) = u(t) + a(t)$, préservant efficacement la sortie originale comme si elle était influencée par $u(t)$ seul.

Dans la section suivante, l'accent est mis sur l'élucidation des fondamentaux de la Théorie des Échelles de Temps, la méthodologie employée pour la détection de ZDA.

10.3 Time-Scale Theory

Dans cette section, nous proposons un aperçu fondamental des outils essentiels nécessaires pour travailler avec les systèmes à échelles de temps. Pour ceux qui cherchent une compréhension plus complète du calcul à échelles de temps, nous recommandons de consulter le travail de Agarwal et al. (2002).

Le calcul à échelles de temps sert de cadre unificateur qui amalgames les principes des calculs continus et discrets. Il généralise les concepts de différentiels ($\dot{x}(t) = f(x(t))$) et de différences ($\Delta x(k) = f(x(k))$) en une formulation singulière ($x^{\Delta}(t) = f(x(t))$). L'aspect le plus crucial de cette formulation réside dans la non-uniformité et l'arbitraire du domaine temporel ($\mathbb{T}$). Pour illustrer, considérons Figure 10.2, où la première ligne dépeint les premières 10 secondes du domaine temporel pour une équation différentielle continue, correspondant à l'ensemble des nombres réels ($\mathbb{T} = \mathbb{R}$); la seconde ligne représente le domaine temporel d'une équation aux différences discrète ($\mathbb{T} = h\mathbb{Z}$) avec un temps d'échantillonnage de $h = 0.5\,\mathrm{s}$; et la troisième ligne démontre le domaine temporel d'une équation à échelles de temps, comprenant un mélange de points continus, discrets et autres points arbitraires, formant ainsi un domaine à échelles de temps valide.

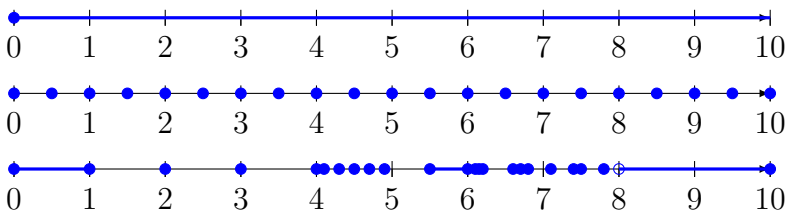


Figure 10.2 – Illustration des domaines temporels pour le temps continu (première ligne), le temps discret (deuxième ligne) et l'échelle de temps (dernière ligne).

Une opération courante dans le traitement des équations différentielles ou des équations aux différences implique de naviguer vers des points temporels valides, en avant ou en arrière. Des expressions comme $x(k+1)$ et

$x(t+\delta)$ apparaissent fréquemment dans les manipulations d'équations. Les opérateurs avance et recul dans le calcul sur échelles de temps servent d'équivalents à ceux-ci, identifiant les points temporels valides immédiatement avant et après un point temporel spécifié. Ces opérateurs sont définis comme suit :

$$\text{avance} : \sigma(t) = \inf\{\tau \in \mathbb{T} | \tau > t\}, \tag{10.7}$$

$$\text{recul} : \rho(t) = \sup\{\tau \in \mathbb{T} | \tau < t\}. \tag{10.8}$$

Le concept de granularité, noté $\mu(t) = \sigma(t) - t$, est crucial dans le calcul sur échelles de temps. Il spécifie la distance entre un point temporel donné t et le prochain point valide dans le temps. Dans un domaine continu ($\mathbb{T} = \mathbb{R}$), la granularité est nulle ($\mu(t) = 0$), tandis que dans un domaine discret ($\mathbb{T} = h\mathbb{Z}$), elle équivaut au temps d'échantillonnage ($\mu(t) = h$). Cependant, dans une équation sur échelles de temps, la granularité n'est pas constante et varie au fur et à mesure que le temps progresse. Cette variabilité est illustrée dans Figure 10.2, où la granularité prend différentes valeurs sur une période de dix secondes.

Avec ces définitions des opérateurs et de la granularité, nous pouvons maintenant introduire la Δ-dérivée, l'équivalent sur échelles de temps de la dérivée. Pour une fonction $f(t)$ définie dans l'échelle de temps, la Δ-dérivée est donnée par :

$$f^{\Delta}(t) = \lim_{\delta \to \mu(t)} \frac{f(t+\delta) - f(t)}{\delta}, \tag{10.9}$$

Ici, $f^{\Delta}(t)$ représente la Δ-dérivée de $f(t)$. Cette formulation englobe à la fois les dérivées différentielles et les dérivées de différence comme cas particuliers, unifiant ainsi ces concepts sous un seul cadre polyvalent.

Les systèmes physiques peuvent être représentés par une modélisation directe utilisant des équations sur échelles de temps. Une approche alternative implique la transformation d'un système en temps continu existant en un système sur échelles de temps, comme détaillé dans le Lemma 11 (Agarwal et al. 2002). Ce manuscrit se concentre exclusivement sur cette dernière méthodologie en raison de ses propriétés inhérentes facilitant certaines simplifications.

Lemma 11. *Considérons un système d'état dans l'espace continu représenté par*

$$\begin{aligned} \dot{x}(t) &= \mathcal{A}x(t) + \mathcal{B}u(t), \\ y(t) &= \mathcal{C}x(t), \end{aligned} \tag{10.10}$$

où $\mathcal{A} \in \mathbb{R}^{n\times n}$, $\mathcal{B} \in \mathbb{R}^{n\times m}$, $\mathcal{C} \in \mathbb{R}^{p\times n}$, $x(t) \in \mathbb{R}^n$, et $y(t) \in \mathbb{R}^p$. Alors, le système équivalent à l'échelle temporelle est donné par

$$\begin{aligned} x^{\Delta}(t) &= A(\mu(t))x(t) + B(\mu(t))u(t), \\ y(t) &= \mathcal{C}x(t), \end{aligned} \tag{10.11}$$

où

$$A(\mu(t)) = \frac{e^{\mathcal{A}\mu(t)} - I}{\mu(t)}, \tag{10.12}$$

$$B(\mu(t)) = \int_0^{\mu(t)} \frac{e^{(\mu(t)-s)\mathcal{A}}}{\mu(t)}\mathcal{B}ds, \tag{10.13}$$

Proof. La solution de l'équation différentielle ordinaire

$$\dot{x}(t) = \mathcal{A}x(t) + \mathcal{B}u(t) \tag{10.14}$$

par rapport au temps donne

$$x(t) = e^{(t-t_0)\mathcal{A}}x(t_0) + \int_{t_0}^{t} e^{(t-\tau)\mathcal{A}}\mathcal{B}u(\tau)d\tau. \tag{10.15}$$

En modifiant l'intervalle de (t_0, t) à $(t, \sigma(t))$, l'équation se transforme en

$$\begin{aligned} x(\sigma(t)) &= e^{(\sigma(t)-t)\mathcal{A}}x(t) + \int_t^{\sigma(t)} e^{(\sigma(t)-\tau)\mathcal{A}}\mathcal{B}u(\tau)d\tau \\ &= e^{\mathcal{A}\mu(t)}x(t) + \int_0^{\mu(t)} e^{(\mu(t)-\tau)\mathcal{A}}\mathcal{B}u(t)d\tau. \end{aligned} \tag{10.16}$$

En appliquant (10.9) à (10.16), nous obtenons

$$x^{\Delta}(t) = \frac{x(\sigma(t)) - x(t)}{\mu(t)} \tag{10.17}$$

$$= \underbrace{\frac{e^{\mu(t)\mathcal{A}} - I}{\mu(t)}}_{A(\mu(t))} x(t) + \underbrace{\int_0^{\mu(t)} \frac{e^{(\mu(t)-\tau)\mathcal{A}}}{\mu(t)}\mathcal{B}d\tau}_{B(\mu(t))} u(t). \tag{10.18}$$

□

Remarque: dans (10.17), si $\mu(t) = 0$, il est nécessaire de prendre la limite, qui devient la dérivée en temps continu par définition, impliquant $A(0) = \mathcal{A}$, $B(0) = \mathcal{B}$ et $x^{\Delta}(t) = \dot{x}(t)$, ce qui peut être vérifié en appliquant la limite à (10.18).

Le calcul sur les échelles de temps offre un cadre unifié pour les dérivées, de manière analogue à la manière dont les systèmes sur échelles de temps fournissent une méthodologie alternative pour l'analyse de stabilité. Le

critère de Hilger, essentiel pour évaluer la stabilité des systèmes sur échelles de temps, est délimité dans (Agarwal et al. 2002). Il exige que toutes les valeurs propres de $A(\mu(t))$ se trouvent à l'intérieur d'un cercle de rayon $\mu(t)^{-1}$ centré en $(-\mu(t)^{-1}, 0)$, exprimé comme suit :

$$\mathcal{H}_\mu \coloneqq \left\{ z \in \mathbb{C}_\mu : |z + \frac{1}{\mu}| < \frac{1}{\mu} \right\}. \tag{10.19}$$

Dans les cas où $\mu(t) = 1$, ce critère converge vers celui du cas discret, spécifiquement le scénario de l'équation aux différences ($\Delta x(k) = f(x(k))$) plutôt que le format de l'opérateur d'avancement ($x(k+1) = f(x(k))$). À l'inverse, pour $\mu(t) = 0$, le cercle s'étend à l'infini, englobant essentiellement tout le demi-plan gauche du plan complexe comme discuté dans (Davis et al. 2010). Figure 10.3 illustre les régions de stabilité pour trois scénarios distincts, chacun correspondant à différentes valeurs de $\mu(t)$.

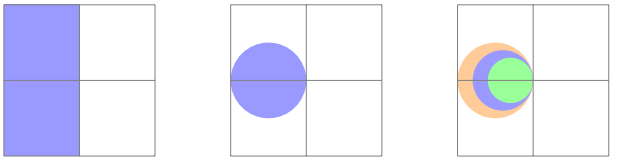


Figure 10.3 – Illustration des différentes régions de stabilité

Pour plus de concision, les sections suivantes de ce document omettront la mention explicite des dépendances en μ et t dans les variables. Par exemple, (10.11) est simplifiée en :

$$\begin{aligned} x^\Delta &= Ax + Bu, \\ y &= \mathcal{C}x. \end{aligned} \tag{10.20}$$

L'adhésion au critère de Hilger est à la fois une condition nécessaire et suffisante pour les systèmes à échelles de temps. Cependant, il est possible de déduire des critères supplémentaires qui respectent le critère de Hilger tout en présentant des formulations alternatives. Un tel exemple est la condition de stabilité de Lyapunov, développée dans le Lemma 12 (Davis et al. 2010). Cette condition sera plus tard utilisée dans la formulation de l'observateur.

Il est impératif pour les systèmes à échelles de temps de respecter le critère de Hilger, car il constitue une condition nécessaire et suffisante pour la stabilité. Cependant, des critères supplémentaires peuvent être dérivés pour répondre à cette exigence fondamentale tout en offrant des cadres analytiques alternatifs. Un tel exemple est la condition de stabilité

de Lyapunov, élucidée dans Lemma 12 (Davis et al. 2010), qui sera ensuite employée dans le développement d'un cadre pour l'observateur.

Lemma 12. *Le système décrit par* (10.20) *est considéré comme stable dans le sens de Lyapunov si il existe une matrice définie positive* P *qui satisfait à l'inégalité matricielle linéaire suivante :*

$$PA + A^\top P + \mu A^\top PA \prec 0. \tag{10.21}$$

Proof. En considérant la fonction candidate classique $V = x^\top Px$, et en utilisant (10.9), nous trouvons :

$$\begin{aligned} V^\Delta &= \frac{V(\sigma) - V}{\mu} \\ &= \frac{x(\sigma)^\top Px(\sigma) - x^\top Px}{\mu} \\ &= \frac{\mu x^\top Px^\Delta + \mu x^{\Delta\top} Px + \mu^2 x^{\Delta\top} Px^\Delta}{\mu} \\ &= x^\top (PA + A^\top P + \mu A^\top PA)x. \end{aligned} \tag{10.22}$$

□

Ayant élucidé les fondamentaux de la Théorie des Échelles de Temps, la section suivante se concentre sur la synthèse d'observateurs.

10.4 Synthèse d'observateurs

Cette section décrit le cadre procédural pour la synthèse d'observateurs au sein des systèmes à échelles de temps, structuré en trois sous-sections distinctes. La première sous-section éclaire la construction d'un observateur pour un paramètre statique μ, établissant les concepts fondamentaux essentiels pour les discours ultérieurs. Ensuite, Section 10.4.2 délimite plusieurs lemmes pertinents qui sous-tendent les constructions théoriques avancées développées plus tard. La sous-section culminante, comme détaillé dans Section 10.4.3, introduit la contribution principale de ce chapitre : l'observateur pour les échelles de temps à paramètres variables que nous utiliserons pour détecter les attaques ZDA. À travers ce chapitre, le système est supposé être observable, car il est inutile de développer un observateur pour un système non observable. Pour plus d'informations sur la vérification de l'observabilité d'un système à échelles de temps, nous invitons le lecteur à se référer à (Ben Nasser et al. 2022).

10.4.1 Synthèse d'observateur pour μ fixé

Dans cette sous-section, nous présentons une formulation basée sur LMI pour la synthèse d'un observateur avec un paramètre μ fixé. L'attribution d'une valeur spécifique à μ permet de délimiter l'observateur comme une variante en temps continu ou en temps discret. Ce cadre fondamental est instrumental dans le développement de l'observateur LPV.

Le paramètre μ dans (10.23) est traité comme une entité constante. Avant de s'engager dans la résolution de les LMI, il est impératif de fixer la valeur de μ, ce qui résulte en un gain d'observateur constant L, par opposition à un variable.

Lemma 13. *Étant donné un système caractérisé par* (10.20)*, il peut être démontré que pour un μ fixé, il existe une matrice L qui assure la stabilité asymptotique de l'observateur délimité par*

$$\hat{x}^{\Delta} = A(\mu)\hat{x} + B(\mu)u - L(y - C\hat{x}) \tag{10.23}$$

à condition qu'il existe une matrice définie positive P et une matrice Z de dimensions appropriées qui satisfont la LMI suivante :

$$\begin{bmatrix} PA + A^{\top}P - ZC - C^{\top}Z^{\top} & (PA - ZC) \\ \star & -\mu^{-1}P \end{bmatrix} \prec 0. \tag{10.24}$$

Ici, le gain de l'observateur est dérivé comme $L = P^{-1}Z$.

Proof. Définissons l'erreur d'estimation e comme

$$e = \hat{x} - x, \tag{10.25}$$

où sa dérivée delta, représentant la dynamique de l'erreur, est exprimée comme

$$e^{\Delta} = (A - LC)e. \tag{10.26}$$

En substituant la dynamique de l'erreur dans (10.21), nous dérivons

$$\begin{aligned} V^{\Delta} &= P(A - LC) + (A - LC)^{\top}P \\ &\quad + \mu(A - LC)^{\top}P(A - LC) \\ &= P(A - LC) + (A - LC)^{\top}P \\ &\quad + \mu[(A - LC)^{\top}PP^{-1}P(A - LC)]. \end{aligned} \tag{10.27}$$

En utilisant le complément de Schur, l'équation ci-dessus peut être représentée de manière équivalente comme

$$V^{\Delta} \equiv \begin{bmatrix} P(A - LC) + (A - LC)^{\top}P & P(A - LC) \\ \star & -\mu^{-1}P \end{bmatrix}. \tag{10.28}$$

Une expansion et substitution ultérieures donnent

$$V^{\Delta} \equiv \begin{bmatrix} PA + A^{\top}P - ZC - C^{\top}Z^{\top} & PA - ZC \\ \star & -\mu^{-1}P \end{bmatrix}. \tag{10.29}$$

□

Il convient de noter que dans le scénario où $\mu = 0$ (le cas en temps continu), l'LMI se simplifie à $PA + A^{\top}P - ZC - C^{\top}Z^{\top} \prec 0$ car l'utilisation du complément de Schur devient inutile. Voir l'équation (4.27) pour une indication claire que cela est vrai.

Étant donné que le système à échelles de temps présente une variabilité de paramètres, la conception d'un observateur avec un μ fixé nécessite la résolution de l'LMI à chaque modification de μ. Pour augmenter l'efficacité de calcul et l'applicabilité pratique, cette étude introduit une variante LPV de l'observateur. Dans ce cadre modifié, les LMI nécessitent une résolution unique, et le gain d'observateur $L(\mu)$ évolue en fonction de μ, facilitant ainsi le calcul en temps réel.

10.4.2 Décomposition sectorielle : Une approche quasi-LPV

Cette sous-section éclaire un résultat bien établi mais peut-être peu familier dans la littérature existante concernant la décomposition sectorielle des fonctions exponentielles de matrices.

L'équation régissant le système à échelle de temps, telle que décrite dans (10.11), présente des caractéristiques non linéaires intrinsèques. Pour faciliter sa transformation en une représentation polytopique, cette étude propose l'application d'une technique de décomposition sectorielle, inspirée par l'approche de modélisation des systèmes Takagi-Sugeno. Cette méthode prend ses racines dans les systèmes quasi-LPV (Tanaka and H. O. Wang 2001).

L'utilisation de la méthodologie de décomposition sectorielle permet d'exprimer une fonction non linéaire comme un polytope convexe dans un intervalle spécifié, un concept explicité dans Lemma 14 (Tanaka and H. O. Wang 2001) :

Lemma 14. *Considérez une fonction $f(\rho) : \mathbb{R} \to \mathbb{R}$ avec le paramètre ρ confiné dans l'intervalle $[\underline{\rho}, \bar{\rho}]$. Elle peut être représentée comme*

$$f(\rho) = \alpha_0(\rho)\underline{f} + \alpha_1(\rho)\bar{f}, \tag{10.30}$$

où

$$\bar{f} = \sup f(\rho), \tag{10.31}$$

$$\underline{f} = \inf f(\rho), \tag{10.32}$$

$$\alpha_0(\rho) = \frac{\bar{f} - f(\rho)}{\bar{f} - \underline{f}}, \tag{10.33}$$

$$\alpha_1(\rho) = \frac{f(\rho) - \underline{f}}{\bar{f} - \underline{f}}. \tag{10.34}$$

Proof. Référez-vous à (Tanaka and H. O. Wang 2001) pour la preuve. □

En essence, la fonction $f(\rho)$ est reformulée comme un simplexe, avec les fonctions de membership α_i qui mappent l'entrée bornée ρ à la plage de sortie bornée $[\underline{f}, \bar{f}]$.

Cependant, une application directe du Lemme 14 à l'Équation (10.24) n'est pas évidente, en raison de sa spécificité aux fonctions scalaires, tandis que le terme $e^{\mathcal{A}\mu} - I$ dans (10.12) est une matrice. Typiquement, la décomposition sectorielle est appliquée à chaque non-linéarité au sein de la matrice. Dans ce cas particulier, cependant, résoudre explicitement la forme algébrique de l'exponentielle matricielle implique un surcroît de calcul important, même pour des systèmes de taille modeste. De plus, les non-linéarités dans chaque élément de la matrice entraîneraient une augmentation exponentielle du nombre de sommets du polytope avec la taille du système.

Néanmoins, dans ce scénario unique, où la fonction non linéaire est une exponentielle matricielle, le lemme suivant est applicable (Kraus 1936) :

Lemma 15. *Pour une matrice $A \in \mathbb{R}^{n\times n}$, l'exponentielle e^A est convexe à condition que A soit auto-adjointe.*

Proof. Une matrice auto-adjointe A est caractérisée par des valeurs propres exclusivement réelles et est donc diagonalisable. Elle peut être représentée comme

$$A = VDV^{-1}, \tag{10.35}$$

avec D étant une matrice diagonale avec des entrées réelles, et V étant la matrice effectuant la transformation de similarité. Par conséquent,

$$e^A = e^{VDV^{-1}} = Ve^DV^{-1}, \tag{10.36}$$

où e^D est convexe. La préservation de la convexité à travers la transformation de similarité assure ainsi la convexité de e^A. □

Lemma 15 établit que pour une matrice auto-adjointe A, e^A est convexe, menant au corollaire suivant :

Corollary 6.1. *Pour une matrice auto-adjointe $\mathcal{A}$, l'inégalité suivante est valable :*

$$e^{\sum_i \alpha_i \mathcal{A}} \preceq \sum_i \alpha_i e^{\mathcal{A}_i}. \tag{10.37}$$

La convexité d'une fonction et la satisfaction de l'inégalité de Jensen (comme détaillé dans (10.37)) sont concomitantes, comme détaillé dans les travaux de Bernstein (2009), Malamud (2001), Antezana, Massey, and Stojanoff (2004), and Kraus (1936). La convexité pré-établie de la fonction valide ainsi l'inégalité.

Maintenant qu'un observateur a été établi pour une valeur fixe de μ, la sous-section suivante étendra ce concept pour accommoder une valeur de μ variable dans un intervalle spécifié.

10.4.3 Synthèse d'observateur à paramètres variant linéairement

Cette sous-section délimite le principal résultat de ce chapitre, Theorem 7, qui étend et affine les concepts présentés dans Lemma 13. Le point central est le développement d'un observateur LPV pour le système à échelles de temps, tel que caractérisé dans Lemma 11.

Theorem 7. *En considération du système à échelles de temps présenté dans Lemma 11, avec une matrice $\mathcal{A}$ auto-adjointe, il existe alors un observateur LPV asymptotiquement stable pour μ dans l'intervalle $[\underline{\mu}, \bar{\mu}]$ pour l'observateur décrit par :*

$$\hat{x}^{\Delta} = A(\mu)\hat{x} + B(\mu)u - L(\mu)(y - C\hat{x}), \tag{10.38}$$

où

$$L(\mu) = \sum_{i=0}^{3} \alpha_i(\mu) L_i, \tag{10.39}$$

$$\alpha_0(\mu) = \frac{\underline{\mu}^{-1} - \mu^{-1}}{\underline{\mu}^{-1} - \bar{\mu}^{-1}} \frac{\bar{\mu} - \mu}{\bar{\mu} - \underline{\mu}}, \tag{10.40}$$

$$\alpha_1(\mu) = \frac{\mu^{-1} - \bar{\mu}^{-1}}{\underline{\mu}^{-1} - \bar{\mu}^{-1}} \frac{\bar{\mu} - \mu}{\bar{\mu} - \underline{\mu}}, \tag{10.41}$$

$$\alpha_2(\mu) = \frac{\underline{\mu}^{-1} - \mu^{-1}}{\underline{\mu}^{-1} - \bar{\mu}^{-1}} \frac{\mu - \underline{\mu}}{\bar{\mu} - \underline{\mu}}, \tag{10.42}$$

$$\alpha_3(\mu) = \frac{\mu^{-1} - \bar{\mu}^{-1}}{\underline{\mu}^{-1} - \bar{\mu}^{-1}} \frac{\mu - \underline{\mu}}{\bar{\mu} - \underline{\mu}}, \tag{10.43}$$

à condition qu'il existe une matrice positive définie P et des matrices Z_i satisfaisant les LMI suivantes :

$$\begin{bmatrix} \Psi_0(\bar{\mu},\bar{\mu}^{-1}) + \Psi_0(\bar{\mu},\bar{\mu}^{-1})^\top & \Psi_0(\bar{\mu},\bar{\mu}^{-1}) \\ \Psi_0(\bar{\mu},\bar{\mu}^{-1})^\top & -\bar{\mu}^{-1}P \end{bmatrix} \prec 0, \tag{10.44}$$

$$\begin{bmatrix} \Psi_1(\underline{\mu},\bar{\mu}^{-1}) + \Psi_1(\underline{\mu},\bar{\mu}^{-1})^\top & \Psi_1(\underline{\mu},\bar{\mu}^{-1}) \\ \Psi_1(\underline{\mu},\bar{\mu}^{-1})^\top & -\bar{\mu}^{-1}P \end{bmatrix} \prec 0, \tag{10.45}$$

$$\begin{bmatrix} \Psi_2(\bar{\mu},\underline{\mu}^{-1}) + \Psi_2(\bar{\mu},\underline{\mu}^{-1})^\top & \Psi_2(\bar{\mu},\underline{\mu}^{-1}) \\ \Psi_2(\bar{\mu},\underline{\mu}^{-1})^\top & -\underline{\mu}^{-1}P \end{bmatrix} \prec 0, \tag{10.46}$$

$$\begin{bmatrix} \Psi_3(\underline{\mu},\underline{\mu}^{-1}) + \Psi_3(\underline{\mu},\underline{\mu}^{-1})^\top & \Psi_3(\underline{\mu},\underline{\mu}^{-1}) \\ \Psi_3(\underline{\mu},\underline{\mu}^{-1})^\top & -\underline{\mu}^{-1}P \end{bmatrix} \prec 0, \tag{10.47}$$

où

$$\Psi_i(\lambda,\omega) = P\omega(e^{\mathcal{A}\lambda} - I) - Z_i C \tag{10.48}$$

$$L_i = P^{-1}Z_i. \tag{10.49}$$

Proof. En partant de (10.29), nous définissons μ et μ^{-1} comme des non-linéarités pour faciliter l'application du corollaire 6.1 à la matrice A. Bien que μ soit linéaire, cette définition aide à délimiter sa fonction de pertinence :

$$A = \frac{e^{\mathcal{A}\mu} - I}{\mu} \preceq \sum_{i=0}^{1}\sum_{j=0}^{1} \beta_i\gamma_j\mu_i^{-1}(e^{A\mu_j} - I). \tag{10.50}$$

Ici, β_i et γ_j sont conformes aux définitions décrites dans Lemma 14.

Par la suite, $\Psi_{i,j}$ est défini comme :

$$\Psi_{i,j} = P\mu_i^{-1}(e^{\mathcal{A}\mu_j} - I) - Z_{i+2j}C. \tag{10.51}$$

En utilisant cette définition, V^Δ peut être représenté comme :

$$V^\Delta \preceq \sum_{i=0}^{1}\sum_{j=0}^{1} \beta_i\gamma_j \begin{bmatrix} \Psi_{i,j} + \Psi_{i,j}^\top & \Psi_{i,j} \\ \star & \mu_i^{-1}P \end{bmatrix} \tag{10.52}$$

En définissant $\alpha_k = \beta_i\gamma_j$ avec $k = i+2j$ (résultant dans les fonctions de pertinence élucidées dans (10.39)) et en substituant le simplexe avec ses sommets, nous dérivons les expressions englobées dans les Equations (10.44) to (10.47). □

En somme, la Section 10.4.3 introduit une synthèse d'observateur LPV pour les systèmes à échelles de temps. Elle souligne l'importance de la

nature auto-adjointe de la matrice $\mathcal{A}$ pour l'estimation efficace des états du système sur la plage spécifiée de μ. Le théorème assure la stabilité asymptotique de l'observateur via la satisfaction des LMI stipulées, garantissant ainsi une estimation précise des états dans les systèmes à échelles de temps. Il est à noter que la contrainte sur la dynamique des matrices dans le domaine du temps continu pourrait sembler restrictive; cependant, tout système avec uniquement des valeurs propres réelles peut subir une transformation de similitude pour devenir auto-adjoint.

La section suivante introduit une simulation conçue pour illustrer la technique de détection.

10.5 Exemple de Simulation

Cette section présente la validation par simulation de la technique de détection proposée à travers une série de simulations réalisées sur un système en temps continu. Ces simulations visent à démontrer l'efficacité de la technique dans la détection des intrusions systémiques.

Pour ces simulations, nous considérons un système en temps continu tel que défini par (10.10), caractérisé par les paramètres matriciels suivants :

$$\mathcal{A} = \begin{bmatrix} -1 & -1 & 0 & 0 & 0 \\ -1 & -1 & 0 & 0 & 0 \\ 0 & 0 & -2 & 0 & 0 \\ 0 & 0 & 0 & -3 & 0 \\ 0 & 0 & 0 & 0 & -5 \end{bmatrix}, \tag{10.53}$$

$$\mathcal{B} = \begin{bmatrix} -7.318 \\ 0.563 \\ 20.78 \\ 21.33 \\ 8.053 \end{bmatrix}, \tag{10.54}$$

$$C = \begin{bmatrix} -1.38 & 0.69 & -1.59 & 1.24 & -0.48 \end{bmatrix}, \tag{10.55}$$

$$D = 0. \tag{10.56}$$

Figure 10.4 illustre le comportement du système en conditions normales d'exploitation et pendant une attaque. Le système possède un zéro à $s_0 = 1$, orienté dans la direction de

$$z_0 = \begin{bmatrix} 2.81 \\ -1.69 \\ -6.93 \\ -5.33 \\ -1.34 \end{bmatrix}. \tag{10.57}$$

Les résultats de la simulation révèlent que la sortie du système reste constante dans les deux scénarios, malgré la divergence des états pendant l'attaque.

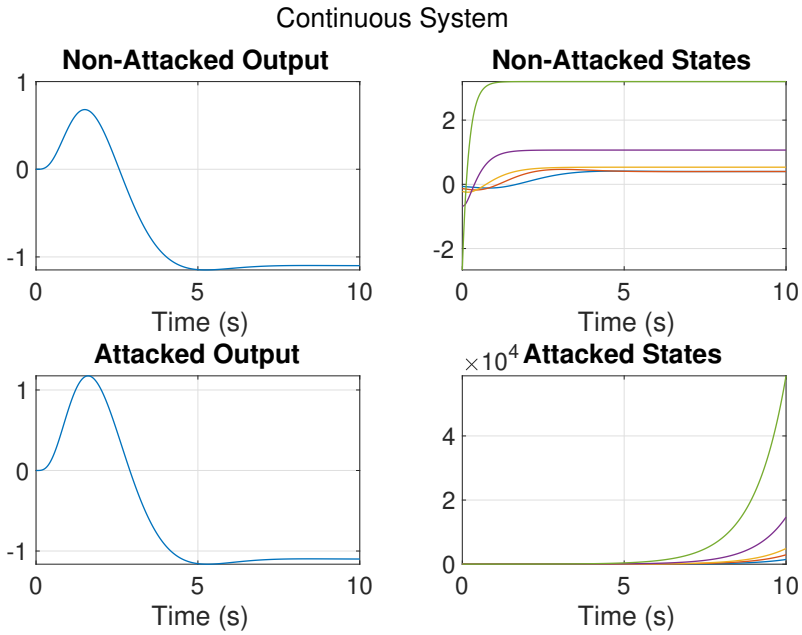


Figure 10.4 – Simulation du système en temps continu avec et sans attaque.

Par la suite, un système à échelles temporelles a été développé en utilisant (10.11) et un observateur a été synthétisé en se basant sur les méthodologies décrites dans Section 10.4.3, en fixant $\mu = 1$. La même simulation a ensuite été réalisée sur l'observateur, comme illustré dans la Figure 10.5. Une différence notable est que l'initiation de l'attaque est retardée jusqu'à $10\,\mathrm{s}$. La divergence de l'état est apparente indépendamment de la valeur de μ, soulignant la présence d'une attaque.

Comme le démontre la Figure 10.5, l'observateur manifeste une sensibilité aux attaques à dynamiques nulles, montrant une divergence tant dans la sortie que dans les états durant l'attaque. Cette déviation peut être détectée efficacement en utilisant des méthodes telles que la classification de

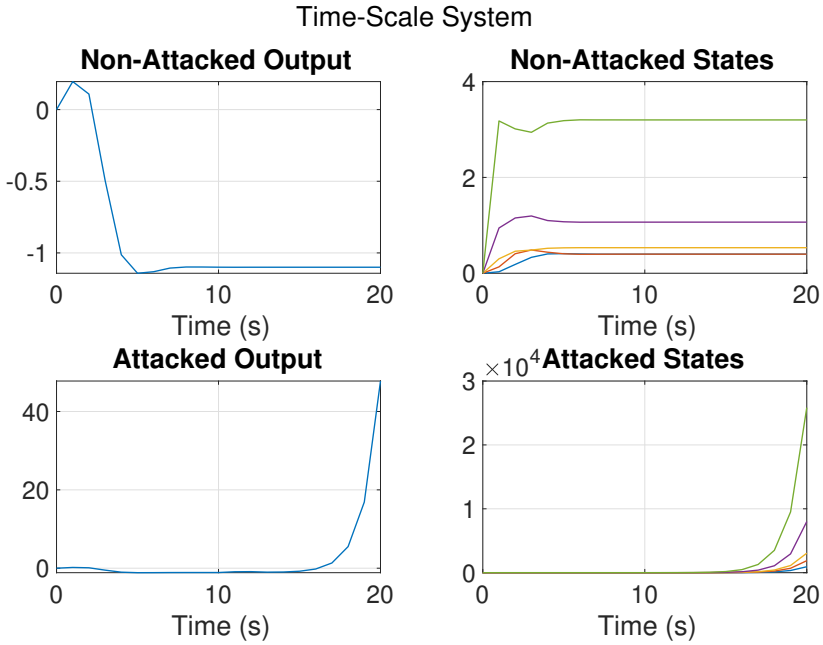


Figure 10.5 – Simulation de l'observateur du système à échelles temporelles avec et sans attaque.

Bayes, comme suggéré par (Isermann 2006), pour déclencher un système d'alarme.

10.6 Conclusion

Ce chapitre a exploré le domaine complexe des techniques de détection contre les ZDA, présentant une approche novatrice pour sa détection. Le paysage des méthodologies actuelles se divise principalement en deux courants : les détenteurs généralisés (Jihan Kim and Shim 2019b; D. Kim, Ryu, and Back 2021), caractérisés par leur mise en œuvre directe mais vulnérables à de nouveaux vecteurs d'attaque, et les stratégies de reconfiguration de topologie (Mao, Jafarnejadsani, et al. 2020), qui, malgré leur efficacité dans la détection des attaques, sont entravées par la complexité de la reconfiguration du contrôle.

L'implémentation de la technique proposée ne nécessite pas la modification aléatoire du paramètre μ, un choix qui simplifie l'aspect opérationnel. Cependant, introduire de l'aléatoire dans la valeur de μ peut renforcer davantage le système de détection, particulièrement contre des attaques sophistiquées qui ciblent simultanément les systèmes à échelles continues et temporelles. L'intégration d'approches théoriques des jeux, telles que la Technique de la Cible Mobile, peut optimiser cette aléatoire, trouvant un

équilibre entre la fréquence des changements de μ et la maximisation de la probabilité de détection des attaques (Umsonst et al. 2022).

À la connaissance des auteurs, l'application des systèmes à échelles temporelles dans l'adressage des ZDA est un territoire inexploré. Cette approche novatrice ouvre un nouveau chapitre dans le domaine, offrant une perspective et une méthodologie uniques. Il est important de noter que, bien que la comparaison des temps de réponse des observateurs à travers différents systèmes puisse ne pas être particulièrement révélatrice en raison de la réglabilité de leurs dynamiques, la technique proposée se distingue par sa flexibilité et son adaptabilité inhérentes.

La recherche présentée dans ce chapitre a introduit une technique pionnière pour la détection des attaques à dynamiques nulles utilisant des systèmes à échelles temporelles. Cette technique ne renforce pas seulement la détection des attaques mais élève également considérablement la barre pour les attaquants afin de maintenir leur furtivité. En éliminant le besoin de transmission de signal vers le côté de contrôle pour traitement, le système proposé ferme efficacement une vulnérabilité critique. De plus, l'introduction d'une dynamique variable dans l'observateur, à travers la modification potentielle de μ, nécessite un degré de connaissance du système plus élevé pour un attaquant, augmentant ainsi la complexité et le risque de détection pour toute tentative d'intrusion.

En conclusion, ce chapitre a contribué une approche novatrice à la détection des attaques à dynamiques nulles, fournissant une solution robuste, flexible et efficace. Les implications de cette recherche s'étendent bien au-delà de son application immédiate, ouvrant la voie à de futures innovations dans la conception et la mise en œuvre de systèmes de contrôle sécurisés. Alors que le paysage des systèmes cyber-physiques continue d'évoluer et de s'étendre, les méthodologies et principes explorés dans cette thèse serviront de fondation pour le développement de mesures de sécurité de plus en plus sophistiquées et résilientes.

Dans le chapitre à venir, l'attention est redirigée vers une autre forme de menace cybernétique : l'attaque Denial-of-Service (DoS). L'application de la Théorie des Échelles Temporelles sera réitérée, facilitant le développement d'un observateur et d'un contrôleur conçus pour atténuer l'impact de telles attaques sur le système.

Chapter 11

Atténuation des attaques par déni de service utilisant la théorie des échelles de temps

11.1 Introduction

Dans le chapitre précédent, la Théorie des Échelles de Temps a été employée pour concevoir un observateur spécifiquement pour la détection de Zero-Dynamics Attack (ZDA). Ce chapitre déplace l'accent vers les attaques Denial-of-Service (DoS), appliquant une fois de plus la Théorie des Échelles de Temps pour formuler à la fois un contrôleur et un observateur qui manifestent une résilience contre de telles attaques.

La dépendance des Cyber-Physical Systems (CPS) aux réseaux de communication engendre un spectre de vulnérabilités sécuritaires. Particulièrement préoccupantes parmi ces menaces sont les attaques DoS, en raison de leur capacité à perturber les réseaux de communication des CPS. En obstruant la transmission des paquets de données de contrôle et de mesure, les attaques DoS affectent directement la fonctionnalité des CPS.

Ces assauts peuvent être orchestrés via diverses méthodologies telles que le brouillage des canaux de communication, le compromis d'appareils pour empêcher la transmission des données, et l'inondation du réseau avec un trafic excessif (Cetinkaya, Ishii, and Hayakawa 2019; Huang et al. 2009). La gravité des attaques DoS varie, allant d'attaques relativement inoffensives causant des retards mineurs et des pertes de paquets à des assauts graves capables de paralyser complètement les réseaux de communication (Pang et al. 2019). Une revue de la littérature sur les attaques DoS

a été présentée dans le Section 8.3.3.

Ce chapitre propose une solution innovante basée sur la théorie des échelles de temps. En utilisant la transformation de continu à échelle de temps délimitée dans Chapter 10, un contrôleur et un observateur pour un système sont conçus pour assurer la stabilité du système en boucle fermée en présence d'attaques DoS, traitant l'attaque comme un délai. La possibilité de varier le temps d'échantillonnage du contrôleur et de l'observateur à n'importe quelle valeur dans un intervalle leur permet de s'ajuster pour fournir des valeurs précises lors de la réception d'un paquet. Tant que les données arrivent dans un délai plus court que le seuil de délai maximum, le système est pratiquement immunisé contre les attaques DoS faibles.

Dans ce chapitre, nous fournissons un aperçu complet des attaques DoS, établissant une compréhension fondamentale. Ensuite, nous introduisons une méthodologie de conception pour développer un contrôleur et un observateur avec une résilience inhérente aux attaques DoS. Par la suite, nous démontrons l'efficacité de la technique proposée à travers son application à un robot TurtleBot3.

La section suivante offre une revue concise des attaques DoS.

11.2 Attaque par Déni de Service

Les attaques DoS se manifestent sous diverses formes, l'attaque basée sur le volume étant prédominante. Ces attaques visent principalement à inonder la bande passante du site cible en générant de copieuses quantités de trafic, incluant des demandes de pages web, des courriels, et des paquets fallacieux. Une telle avalanche surcharge la cible, menant à des ralentissements significatifs ou à l'obstruction complète du trafic légitime. Une tactique répandue implique le déploiement de botnets—des réseaux d'ordinateurs ou d'appareils compromis, manipulés à distance pour envoyer un volume écrasant de trafic vers la cible.

De plus, il y a des attaques de protocole, stratégiquement conçues pour épuiser les ressources des serveurs ou celles des composants de réseau intermédiaires, tels que les pare-feux et les équilibreurs de charge. Ces attaques exploitent les vulnérabilités au sein des piles de protocoles de couche 3 et 4, typiquement en envoyant des paquets qui nécessitent un traitement partiel par le système cible, aboutissant à une surcharge du système. Un exemple notable est l'attaque par inondation SYN, où l'assaillant initie une connexion TCP mais s'abstient de compléter la poignée de main,

résultant en une accumulation de connexions à demi ouvertes.

Les attaques de la couche applicative représentent une forme plus sophistiquée, ciblant des fonctionnalités spécifiques d'une application ou d'un service. Ces attaques sont exécutées en envoyant des requêtes qui semblent légitimes, mais qui sont minutieusement conçues pour exploiter les vulnérabilités ou les inefficacités du code de l'application, entraînant des plantages ou des ralentissements significatifs. Les tactiques peuvent inclure le bombardement d'une base de données avec des requêtes complexes ou l'exploitation de segments de code inefficaces.

Du point de vue de l'Automatique des systèmes, une attaque par déni de service (DoS) présente des risques importants pour la stabilité du système. Cela est principalement dû au fait que les systèmes automatisés dépendent de la livraison rapide et fiable des informations pour modifier et maintenir l'état du système. Malheureusement, ces systèmes sont généralement mal équipés pour faire face aux interruptions de flux de données, ce qui entraîne des difficultés à s'adapter à l'absence d'informations. Dans le contexte des systèmes en temps continu, cette interruption signifie que, sur un certain intervalle de temps t, les mesures du système $y(t)$ ne seront pas disponibles.

De même, les systèmes en temps discret subissent des interruptions dans la transmission des données, ce qui entraîne que certaines valeurs de $y(k)$ ne sont pas transmises. Par conséquent, ce scénario peut être assimilé à un système fonctionnant avec un temps d'échantillonnage variable, caractérisé par les heures d'arrivée imprévisibles des informations. Cette variabilité complique la capacité du système à traiter et répondre efficacement aux données entrantes, compromettant ainsi la stabilité et la performance globales du système.

Dans le présent chapitre, l'attaque DoS est perçue comme un temps d'échantillonnage variable. Cette perspective est particulièrement pertinente pour les systèmes robotiques qui dépendent de signaux de capteurs ou de référence en réseau. Cela est notamment applicable dans les environnements industriels, où les robots peuvent dépendre de systèmes de suivi en réseau pour acquérir des données de position. Cette configuration implique souvent une connexion réseau susceptible d'interférences ou d'attaques, particulièrement dans le segment reliant le contrôleur/observateur au capteur externe ou au générateur de référence.

La section suivante détaille la conception du contrôleur et de l'observateur en utilisant la Théorie des Échelles de Temps pour atténuer les attaques DoS.

11.3 Synthèse du Contrôleur et de l'Observateur

Afin de synthétiser un contrôleur qui maintient la fonctionnalité sous des attaques DoS faibles, cette recherche adopte la théorie des échelles de temps décrite dans le Chapter 10. Un contrôleur basé sur cette théorie modifierait son domaine temporel opérationnel en réponse à l'attaque, générant ainsi des signaux de contrôle conformément au passage des paquets. Cette approche permet de préserver la stabilité du système, bien que avec une réduction de performance, jusqu'à ce qu'un seuil maximal de délai temporel prédéterminé soit atteint. Le schéma du système de contrôle est représenté dans la Figure 11.1. Il n'y a pas de réseau entre le système et les capteurs externes, donc il n'est pas considéré comme attaquable. Le réseau entre les capteurs et l'observateur, en revanche, est susceptible aux attaques DoS. Pour y remédier, l'observateur et le contrôleur doivent tous deux être capables de changer leur temps d'échantillonnage pour compenser la perte d'informations.

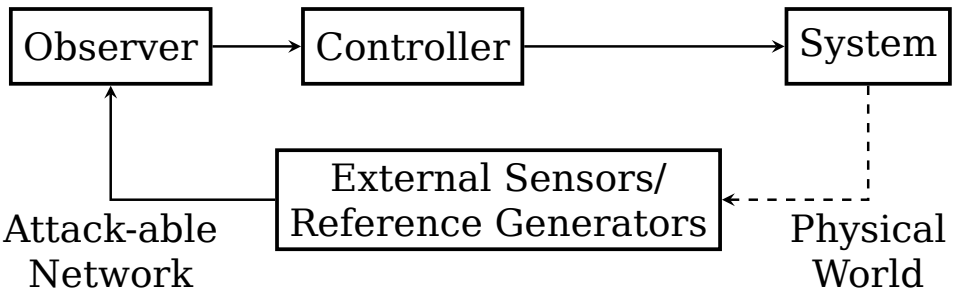


Figure 11.1 – Schéma du système pour une attaque DoS.

Le processus de conception prend en compte l'indépendance entre le contrôleur et l'observateur. Initialement, le contrôleur est formulé. Par la suite, l'observateur est construit en tenant compte du contrôleur préalablement conçu pour s'assurer qu'il fonctionne suffisamment rapidement afin de prévenir la déstabilisation du contrôleur.

Comme dans le chapitre précédent, la conception est segmentée en deux sections distinctes. Initialement, la synthèse est réalisée avec une valeur constante de μ. Par la suite, l'approche est étendue pour accommoder un μ variable.

11.3.1 Synthèse pour un μ fixe

Le modèle mathématique pour le système étudié est exprimé comme suit :

$$x^{\Delta} = Ax + Bu, \tag{11.1}$$
$$y = Cx, \tag{11.2}$$
$$u = -K\hat{x}, \tag{11.3}$$

où x^{Δ} dénote le changement dans le vecteur d'état x, A et B représentent les matrices du système, u est l'entrée de contrôle, y est la sortie, et K est la matrice de gain du contrôleur. L'observateur, qui estime le vecteur d'état $\hat{x}$, est décrit par l'équation

$$\hat{x}^{\Delta} = A\hat{x} + Bu + L(y - \hat{y}). \tag{11.4}$$

Ici, $\hat{x}^{\Delta}$ signifie le changement du vecteur d'état estimé, L est la matrice de gain de l'observateur, et $\hat{y}$ représente la sortie estimée.

Le Lemma 16 établit la condition de Lyapunov pour concevoir le contrôleur et l'observateur pour une valeur fixe de μ.

Lemma 16. *En considérant un système régi par la dynamique présentée dans* (5.1) *et en supposant que le système est à la fois observable et contrôlable. Alors, pour une valeur constante de μ, l'existence d'une solution à la LMI suivante garantit le critère de stabilité de Hilger du système en boucle fermée avec des états estimés, ainsi que calcule les valeurs des gains du contrôleur et de l'observateur :*

$$AP_1 - BZ_1 = \Omega_1, \tag{11.5}$$
$$\begin{bmatrix} \Omega_1 + \Omega_1^{\top} & \Omega_1 \\ \Omega_1^{\top} & -\mu^{-1}P_1 \end{bmatrix} \prec 0, \tag{11.6}$$
$$P_1 \succ 0, \tag{11.7}$$

Cette LMI sert de base pour dériver un contrôleur sur une échelle de temps désignée. Par la suite, un observateur peut être conçu en utilisant la LMI suivante :

$$P_2 A - Z_2 C = \Omega_2, \tag{11.8}$$

$$\begin{bmatrix} \Omega_1 + \Omega_1^\top & \Omega_1 & P_1 BK & P_1 BK \\ \Omega_1^\top & -\mu^{-1} P_1 & 0 & 0 \\ K^\top B^\top P_1 & 0 & \Omega_2 + \Omega_2^\top & \Omega_2 \\ K^\top B^\top P_1 & 0 & \Omega_2^\top & -\mu^{-1} P_2 \end{bmatrix} \prec 0, \tag{11.9}$$

$$P_2 \succ 0 \tag{11.10}$$

où A, B et C sont les matrices du système, P_1 et P_2 sont des matrices définies positives, K est le gain du contrôleur, L est le gain de l'observateur, et Z_1 et Z_2 sont des matrices de variables complètes.

Les gains sont récupérés comme $K = Z_1 P_1^{-1}$ et $L = P_2^{-1} Z_2$.

Proof. L'erreur d'estimation pour le système contrôlé est définie comme suit :

$$e = x - \hat{x}, \tag{11.11}$$

où e représente l'erreur d'estimation, x est l'état réel, et $\hat{x}$ désigne l'état estimé. La dérivée temporelle de l'erreur d'estimation est donnée par :

$$\begin{aligned} e^\Delta &= \frac{e(\sigma) - e}{\mu} && (11.12)\\ &= \frac{x(\sigma) - \hat{x}(\sigma) - (x - \hat{x})}{\mu} && (11.13)\\ &= (x^\Delta - \hat{x}^\Delta) && (11.14)\\ &= Ax - BK\hat{x} - A\hat{x} + BK\hat{x} - L(Cx - C\hat{x}) && (11.15)\\ &= A(x - \hat{x}) - LC(x - \hat{x}) && (11.16)\\ &= (A - LC)e. && (11.17) \end{aligned}$$

La dynamique du système, incorporant l'erreur d'estimation, peut être reformulée comme suit :

$$\begin{aligned} x^\Delta &= Ax - BK\hat{x} && (11.18)\\ &= Ax - BK(x - e) && (11.19)\\ &= (A - BK)x + BKe. && (11.20) \end{aligned}$$

En combinant les dynamiques du contrôleur et de l'observateur, le système augmenté peut être décrit par l'équation matricielle suivante :

$$\underbrace{\begin{bmatrix} x^{\Delta} \\ e^{\Delta} \end{bmatrix}}_{z^{\Delta}} = \underbrace{\begin{bmatrix} A - BK & BK \\ 0 & A - LC \end{bmatrix}}_{\Phi} \underbrace{\begin{bmatrix} x \\ e \end{bmatrix}}_{z}, \tag{11.21}$$

ce qui représente l'évolution combinée de l'état du système et de l'erreur d'estimation dans le temps.

Pour la fonction candidate de Lyapunov, définie comme :

$$V = z^{\top} P z, \tag{11.22}$$

où V représente la fonction de Lyapunov et P est une matrice définie positive, la condition suivante assure la stabilité du contrôleur et de l'observateur pour une valeur spécifiée de μ :

$$\begin{aligned} V^{\Delta} &= \frac{z(\sigma)^{\top} P z(\sigma) - z^{\top} P z}{\mu} && (11.23)\\ &= z^{\top} P(\mu z^{\Delta}) + (\mu z^{\Delta})^{\top} P z + \mu^2 (z^{\Delta})^{\top} P z^{\Delta} && (11.24)\\ &= z^{\top} P z^{\Delta} + (z^{\Delta})^{\top} P z + \mu (z^{\Delta})^{\top} P z^{\Delta} && (11.25)\\ &= P\Phi + \Phi^{\top} P + \mu \Phi^{\top} P \Phi && (11.26)\\ &\equiv \begin{bmatrix} P\Phi + \Phi^{\top} P & P\Phi \\ \Phi^{\top} P & -\frac{1}{\mu} P \end{bmatrix} \prec 0, && (11.27)\\ P &\succ 0. && (11.28) \end{aligned}$$

La formulation LMI nécessite une transformation de similarité et une substitution de variables avec $Z_1 = KP$ et $Z_2 = PL$. Il est conseillé de sélectionner P sous la forme de $\begin{bmatrix} P_1 & 0 \\ 0 & P_2 \end{bmatrix}$, pour réduire le nombre d'éléments bilinéaires dans la matrice résultante. En conséquence, cela mène à :

$$\Omega_1 = AP_1 - BZ_1, \tag{11.29}$$

$$\Omega_2 = P_2A - Z_2C, \tag{11.30}$$

$$V^{\Delta} = \begin{bmatrix} \Omega_1 + \Omega_1^\top & \Omega_1 & P_1BK & P_1BK \\ \Omega_1^\top & \mu^{-1}P_1 & 0 & 0 \\ K^\top B^\top P_1 & 0 & \Omega_2 + \Omega_2^\top & \Omega_2 \\ K^\top B^\top P_1 & 0 & \Omega_2^\top & \mu^{-1}P_2 \end{bmatrix}. \tag{11.31}$$

Bien que cette matrice inclue des variables non substituées et des termes bilinéaires, s'attaquer d'abord à la partie du contrôleur de l'LMI en résolvant

$$\begin{bmatrix} \Omega_1 + \Omega_1^\top & \Omega_1 \\ \Omega_1^\top & \mu^{-1}P_1 \end{bmatrix} \prec 0 \tag{11.32}$$

et en substituant ensuite des constantes pour P_1 et K dans le système dérivé, permet la résolution de l'LMI global. Cela résulte en une configuration de l'observateur qui maintient la stabilité du système en boucle fermée sous contrôle. □

La sous-section suivante étend la conception à partir d'une valeur de μ fixe vers un μ variable sur un intervalle spécifié.

11.3.2 Synthèse pour un μ Variable

Cette sous-section délimite le principal résultat de ce chapitre, le Theorem 8, qui étend et affine les concepts présentés dans le Lemma 16.

Theorem 8. *En considération du système à échelles de temps exposé dans Lemma 11 pour un système décrit comme dans* (11.1) *et un observateur de Luenberger comme dans* (11.4)*, et en supposant que la matrice $\mathcal{A}$ est auto-adjointe, il existe un contrôleur et un observateur Linear Parameter Varying (LPV) asymptotiquement stables pour μ dans l'intervalle $[\underline{\mu}, \bar{\mu}]$ s'il existe des matrices définies positives P_i et des matrices $Z_{i,j}$ satisfaisant les inégalités matricielles linéaires (LMIs) suivantes :*

Tout d'abord, pour le contrôleur :

$$A_i P_1 - B_i Z_{1,i} = \Omega_{1,i}, \tag{11.33}$$

$$\begin{bmatrix} \Omega_{1,1} + \Omega_{1,1}^\top & \Omega_{1,1} \\ \Omega_{1,1}^\top & \underline{\mu}^{-1} P_1 \end{bmatrix} \prec 0, \tag{11.34}$$

$$\begin{bmatrix} \Omega_{1,2} + \Omega_{1,2}^\top & \Omega_{1,2} \\ \Omega_{1,2}^\top & \bar{\mu}^{-1} P_1 \end{bmatrix} \prec 0 \tag{11.35}$$

$$\begin{bmatrix} \Omega_{1,3} + \Omega_{1,3}^\top & \Omega_{1,3} \\ \Omega_{1,3}^\top & \underline{\mu}^{-1} P_1 \end{bmatrix} \prec 0 \tag{11.36}$$

$$\begin{bmatrix} \Omega_{1,4} + \Omega_{1,4}^\top & \Omega_{1,4} \\ \Omega_{1,4}^\top & \bar{\mu}^{-1} P_1 \end{bmatrix} \prec 0 \tag{11.37}$$

$$P_1 \succ 0, \tag{11.38}$$

Ensuite, pour l'observateur :

$$P_2 A_i - Z_{2,i} C_i = \Omega_{2,i}, \tag{11.39}$$

$$\begin{bmatrix} \Omega_{1,1} + \Omega_{1,1}^\top & \Omega_{1,1} & P_1 B_1 K_1 & P_1 B_1 K_1 \\ \Omega_{1,1}^\top & \underline{\mu}^{-1} P_1 & 0 & 0 \\ K_1^\top B_1^\top P_1 & 0 & \Omega_{2,1} + \Omega_{2,1}^\top & \Omega_{2,1} \\ K_1^\top B_1^\top P_1 & 0 & \Omega_{2,1}^\top & \underline{\mu}^{-1} P_2 \end{bmatrix} \prec 0, \tag{11.40}$$

$$\tag{11.41}$$

$$\begin{bmatrix} \Omega_{1,2}+\Omega_{1,2}^\top & \Omega_{1,2} & P_1B_2K_2 & P_1B_2K_2 \\ \Omega_{1,2}^\top & \bar{\mu}^{-1}P_1 & 0 & 0 \\ K_2^\top B_2^\top P_1 & 0 & \Omega_{2,2}+\Omega_{2,2}^\top & \Omega_{2,2} \\ K_2^\top B_2^\top P_1 & 0 & \Omega_{2,2}^\top & \bar{\mu}^{-1}P_2 \end{bmatrix} \prec 0, \tag{11.42}$$

$$\begin{bmatrix} \Omega_{1,3}+\Omega_{1,3}^\top & \Omega_{1,3} & P_1B_3K_3 & P_1B_3K_3 \\ \Omega_{1,3}^\top & \underline{\mu}^{-1}P_1 & 0 & 0 \\ K_3^\top B_3^\top P_1 & 0 & \Omega_{2,3}+\Omega_{2,3}^\top & \Omega_{2,3} \\ K_3^\top B_3^\top P_1 & 0 & \Omega_{2,3}^\top & \underline{\mu}^{-1}P_2 \end{bmatrix} \prec 0, \tag{11.43}$$

$$\begin{bmatrix} \Omega_{1,4}+\Omega_{1,4}^\top & \Omega_{1,4} & P_1B_4K_4 & P_1B_4K_4 \\ \Omega_{1,4}^\top & \bar{\mu}^{-1}P_1 & 0 & 0 \\ K_4^\top B_4^\top P_1 & 0 & \Omega_{2,4}+\Omega_{2,4}^\top & \Omega_{2,4} \\ K_4^\top B_4^\top P_1 & 0 & \Omega_{2,4}^\top & \bar{\mu}^{-1}P_2 \end{bmatrix} \prec 0, \tag{11.44}$$

$$P_2 \succ 0. \tag{11.45}$$

où

$$A_1 = A(\underline{\mu}), A_2 = \frac{A(\underline{\mu})\underline{\mu}}{\bar{\mu}}, A_3 = \frac{A(\bar{\mu})\bar{\mu}}{\underline{\mu}}, A_4 = A(\bar{\mu}) \tag{11.46}$$

$$B_1 = B(\underline{\mu}), B_2 = \frac{B(\underline{\mu})\underline{\mu}}{\bar{\mu}}, B_3 = \frac{B(\bar{\mu})\bar{\mu}}{\underline{\mu}}, B_4 = B(\bar{\mu}) \tag{11.47}$$

$$K_i = Z_{1,i}P^{-1}, L_i = P^{-1}Z_{2,i}, \tag{11.48}$$

$$L(\mu) = \sum_{i=0}^{3} \alpha_i(\mu)L_i, K(\mu) = \sum_{i=0}^{3} \alpha_i(\mu)K_i, \tag{11.49}$$

$$\alpha_0(\mu) = \frac{\underline{\mu}^{-1}-\mu^{-1}}{\underline{\mu}^{-1}-\bar{\mu}^{-1}}\frac{\bar{\mu}-\mu}{\bar{\mu}-\underline{\mu}}, \tag{11.50}$$

$$\alpha_1(\mu) = \frac{\mu^{-1}-\bar{\mu}^{-1}}{\underline{\mu}^{-1}-bar\mu^{-1}}\frac{\bar{\mu}-mu}{bar\mu-\underline{\mu}}, \tag{11.51}$$

$$\alpha_2(\mu) = \frac{\underline{\mu}^{-1}-mu^{-1}}{\underline{\mu}^{-1}-\bar{\mu}^{-1}}\frac{\mu-\underline{\mu}}{\bar{\mu}-\underline{\mu}}, \tag{11.52}$$

$$\alpha_3(\mu) = \frac{\mu^{-1}-\bar{\mu}^{-1}}{\underline{\mu}^{-1}-bar\mu^{-1}}\frac{\mu-\underline{\mu}}{bar\mu-\underline{\mu}}, \tag{11.53}$$

Proof. La preuve est similaire à celle du Theorem 7. □

En résumé, la Section 11.3.2 introduit une synthèse de contrôleur et d'observateur LPV pour les systèmes à échelles de temps. Le théorème

assure la stabilité asymptotique du contrôleur et de l'observateur via la satisfaction des LMI stipulées, garantissant ainsi une estimation et un contrôle précis de l'état dans les systèmes à échelles de temps. Le fait que le contrôleur et les observateurs conçus puissent changer leur temps d'échantillonnage en ligne dans l'intervalle $[\underline{\mu}, \bar{\mu}]$ permet au système de s'adapter aux attaques par DoS, comme le montre la section suivante.

11.4 Validation Expérimentale

Cette section éclaire la méthodologie de validation utilisée pour évaluer l'efficacité du contrôleur et de l'observateur proposés dans l'atténuation des attaques par DoS. La configuration expérimentale impliquait le déploiement d'un robot mobile autonome, qui obtenait ses données de position d'un capteur optique externe. L'attaque par DoS était orchestrée au sein du lien de communication entre le capteur et le robot, bloquant ainsi la transmission des lectures du capteur au robot.

Cette obstruction était configurée de manière à ne permettre la transmission d'une seule lecture après un intervalle prédéterminé, avec un élément d'aléatoire ajouté pour simuler un scénario d'attaque plus réaliste. Ce cadre expérimental était conçu pour tester la résilience du contrôleur et de l'observateur face aux perturbations ciblées dans le flux de données des capteurs, essentielles pour maintenir l'intégrité opérationnelle dans les systèmes robotiques autonomes.

Le robot autonome utilisé était le TurtleBot3 Waffle Pi[1], illustré dans la Figure 11.2, qui est une plateforme robotique très polyvalente et personnalisable conçue pour l'éducation, la recherche et le développement dans le domaine de la robotique. Il fait partie de la série TurtleBot, qui sont des plateformes robotiques open-source offrant un support clé en main pour diverses applications robotiques, incluant la navigation autonome, la reconnaissance d'objets, et la manipulation.

Néanmoins, la majorité des capacités matérielles mentionnées précédemment et des mécanismes de support n'ont pas été utilisés, avec une préférence donnée à nos conceptions de contrôleurs sur mesure et configurations de détection. La seule fonctionnalité du robot qui a été employée dans cette étude était son odomètre. Cet instrument a facilité

[1] https://www.turtlebot.com/turtlebot3

[2] Source : https://www.elektor.com/cdn/shop/files/robotis-turtlebot3-waffle-pi-incl-raspb jpg, consulté le 2024-04-08T1400

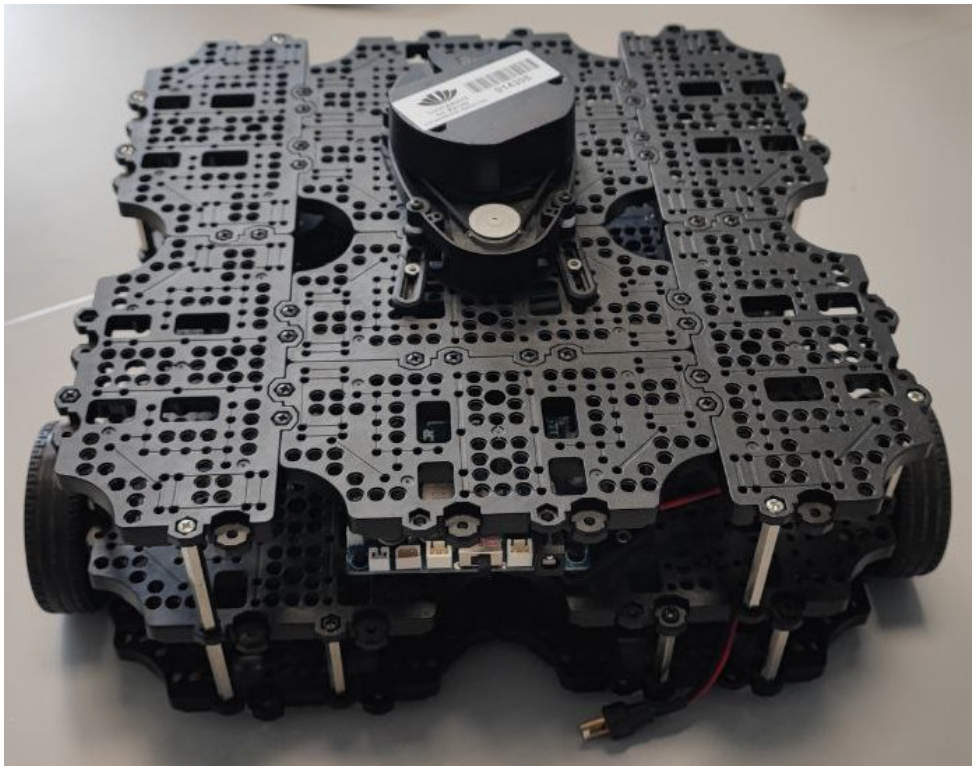


Figure 11.2 – TurtleBot3 Waffle Pi.[2]

la vérification du déplacement du robot conformément aux signaux de contrôle, éliminant ainsi la nécessité d'une commande directe des moteurs à courant continu (CC). Cette approche a permis une intégration plus rationalisée de nos algorithmes de contrôle, en se concentrant sur la précision et l'efficacité des mouvements de haut niveau.

Équipé d'un Raspberry Pi comme contrôleur principal, le robot peut être programmé en utilisant le ROS en `C++` et `Python`. Le ROS est un cadre de travail open-source et flexible pour l'écriture de logiciels robotiques. Il ne s'agit pas d'un système d'exploitation au sens traditionnel de la gestion des ressources matérielles, mais plutôt d'une collection de cadres logiciels pour le développement de logiciels robotiques. Le ROS fournit une couche de communication structurée au-dessus des systèmes d'exploitation hôtes d'un cluster informatique mixte, offrant des services conçus pour un cluster informatique hétérogène tels que l'abstraction matérielle, le contrôle des dispositifs de bas niveau, la mise en œuvre de fonctionnalités couramment utilisées, le passage de messages entre processus et la gestion de paquets.

Conçu pour faciliter le développement d'applications robotiques complexes et robustes, le ROS permet aux programmeurs d'utiliser le `C++` et le `Python` pour écrire des applications, ce qui améliore considérablement l'accessibilité et la polyvalence de la programmation des robots. Ceci est

particulièrement bénéfique dans les milieux académiques et de recherche, où des robots comme celui équipé d'un Raspberry Pi servent de plateformes expérimentales pour tester des théories et des algorithmes dans des scénarios réels.

L'architecture du ROS est modulaire, permettant la réutilisation du code dans divers projets, ce qui accélère le processus de développement. Cette modularité est atteinte grâce à l'utilisation de paquets, qui sont des collections de nœuds (processus qui effectuent des calculs), de bibliothèques et d'outils. Les nœuds communiquent entre eux sur des sujets en utilisant un modèle éditeur-abonné, des services avec un modèle de requête-réponse, ou par le biais de serveurs d'actions pour des tâches de longue durée.

Pour surveiller la position du robot dans l'espace, le système Optitrack a été utilisé. Le système Optitrack est une solution de capture de mouvement sophistiquée conçue pour suivre avec précision la position et l'orientation des objets dans l'espace tridimensionnel. Utilisé largement dans l'animation, les sciences du sport, la biomécanique, la robotique et les applications de réalité virtuelle, il emploie une série de caméras à grande vitesse équipées de capteurs infrarouges pour détecter des marqueurs réfléchissants ou des LED spécialement conçus placés sur l'objet d'intérêt. En capturant la position spatiale de ces marqueurs à haute fréquence, le système peut reconstruire le mouvement précis de l'objet avec une remarquable précision et une latence minimale.

Le système Optitrack manque intrinsèquement d'un mécanisme direct pour la diffusion en réseau des données de suivi. Néanmoins, un effort collaboratif au sein de l'institution a conduit au développement d'un système intermédiaire conçu pour combler cette lacune. Ce système capture les données positionnelles générées par le système Optitrack et les transmet à un serveur MQTT, facilitant ainsi l'accessibilité des données à l'échelle du réseau.

MQTT fonctionne selon un modèle de publication-abonnement qui est très efficace pour diverses applications grâce à sa conception de protocole légère. Au cœur de la stratégie de communication de MQTT se trouve le concept de « sujets ». Un sujet est une chaîne UTF-8 que le courtier utilise pour filtrer les messages pour chaque client connecté. Les sujets sont structurés en hiérarchie, semblable à un chemin de système de fichiers, en utilisant des barres obliques (/) comme séparateurs. Cette structure hiérarchique permet un routage précis et flexible des messages. Les sujets de MQTT et de ROS fonctionnent de la même manière, suivant la même

philosophie, mais utilisant des technologies sous-jacentes différentes.

Lorsqu'un éditeur souhaite envoyer des données, il publie un message sur un sujet spécifique. Par exemple, un capteur de température peut publier ses lectures sur un sujet nommé « home/livingroom/temperature ». Les abonnés, d'autre part, indiquent leur intérêt à recevoir des messages en s'abonnant à un ou plusieurs sujets. Le courtier MQTT, qui est le serveur central facilitant la communication de publication-abonnement, est responsable de la distribution des messages des éditeurs aux abonnés en fonction des abonnements aux sujets. La version 1 de ROS fonctionne de la même manière, avec le courtier appelé « master », cependant la version 2 utilise un système distribué qui ne fait pas usage de courtier.

Pour intégrer ce flux de données avec le cadre opérationnel du TurtleBot3, un nœud ROS dédié a été développé. Ce nœud sert non seulement de conduit pour acheminer les données positionnelles Optitrack vers le TurtleBot3 mais agit également comme un point stratégique pour la mise en œuvre de l'attaque DoS, comme représenté dans la figure ci-dessous. Cela crée effectivement un système en boucle fermée, dans lequel le mouvement du robot est capturé par le système de suivi, envoyé au serveur MQTT auquel le robot lui-même est abonné et peut donc utiliser les informations pour prendre des décisions relatives à son mouvement.

Dans la délimitation des entrées et des sorties du système, à savoir les commandes de direction du robot en entrée et les lectures des capteurs en sortie, un modèle dynamique a été conceptualisé pour s'interfacer de manière transparente avec la technique proposée. Conformément aux conventions prévalentes observées dans les implémentations ROS pour les plateformes robotiques, le modèle délimite les entrées comme les vitesses linéaire et angulaire du robot (v et $\dot{\theta}$, respectivement). Parallèlement, les sorties sont définies comme la position et l'orientation du robot, avec x et y représentant les coordonnées cartésiennes de la position, et θ dénotant l'orientation. Le vecteur d'état du système est identique à son vecteur de sortie.

Le comportement dynamique du système est principalement caractérisé par des composants intégratifs, basés sur le principe selon lequel le robot reste stationnaire en l'absence de toute entrée externe. La relation entrée-sortie du système montre une dépendance à l'état, où la vitesse linéaire est finement mappée au changement de position du robot comme $\dot{x} = \cos(\theta)$ et $\dot{y} = \sin(\theta)$, avec une corrélation directe de la vitesse angulaire. Cette dépendance intrinsèque à l'état nécessite l'adoption d'un processus de linéarisa-

tion pour simplifier la dynamique du système pour l'analyse et le contrôle.

Pour faciliter cette linéarisation, un état initial a été établi à $\theta = -170$ degrés, et le processus a été exécuté par incréments séquentiels de $10°$. Par la suite, pour chaque modèle linéarisé, une paire unique comprenant un contrôleur et un observateur a été synthétisée. Les gains correspondants pour chaque paire sont accessibles via un dépôt en ligne[3]. Le nœud Robot Operating System (ROS) est programmé pour sélectionner la paire appropriée pour l'état actuel du système.

Les paramètres $\underline{\mu}$ et $\bar{\mu}$ ont été déterminés à 0.1 et 2, respectivement. Cette sélection visait à faciliter un fonctionnement sans heurts dans des conditions normales, et à permettre l'exécution d'une attaque significativement agressive, durant jusqu'à $2\,\mathrm{s}$, en l'absence de mécanismes de rétroaction. Par la suite, l'attaque DoS a été exécutée avec des durées variables, spécifiquement $0.5\,\mathrm{s}$, $1.5\,\mathrm{s}$ et $1.9\,\mathrm{s}$, pour évaluer la résilience du système sous différents niveaux d'intensité d'attaque.

Les représentations graphiques dans la Figure 11.3 délimitent la position et l'orientation de l'unité robotique, avec des points de passage indiqués par des points bleus. La trajectoire prévue pour le robot était de naviguer à travers chaque point de passage, délimitant finalement un chemin circulaire.

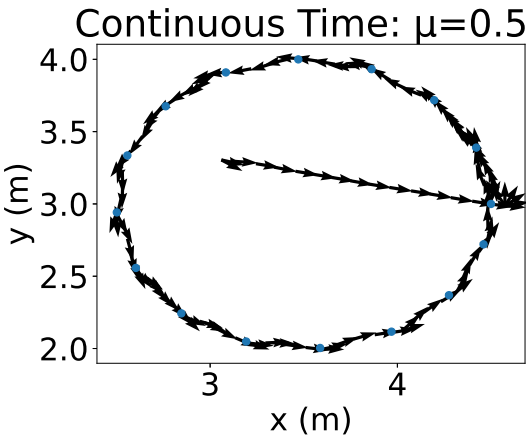
Le protocole opérationnel du robot englobe un réalignement angulaire initial pour se diriger vers le point de passage suivant, suivi par le début d'un mouvement linéaire dans la direction du point de passage. Durant ce transit, le robot peut entreprendre des ajustements angulaires supplémentaires au besoin pour maintenir le cap correct.

Les observations des graphiques révèlent que les contrôleurs en temps continu et en échelle de temps exécutent avec succès la tâche de navigation sous une attaque DoS légère durant $0.5\,\mathrm{s}$ (Figure 11.3(a) et Figure 11.3(b)). Cependant, lors d'une attaque plus prolongée de $1.5\,\mathrm{s}$, le contrôleur en temps continu (Figure 11.3(c)) échoue à s'aligner avec le premier point de passage en raison d'un dépassement persistant, tandis que le contrôleur en échelle de temps (Figure 11.3(d)) connaît seulement une dégradation partielle de la performance de la trajectoire.

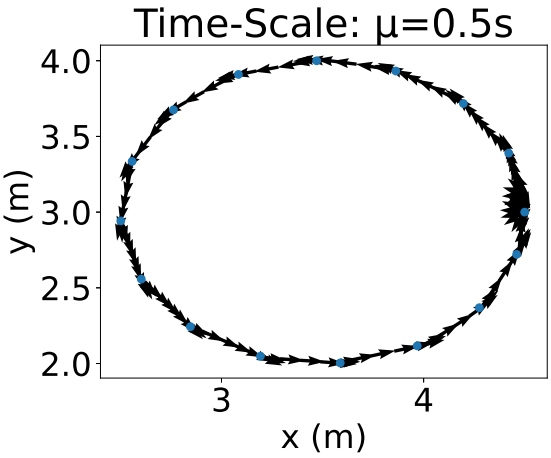
Sous une durée d'attaque de $1.9\,\mathrm{s}$, le contrôleur en échelle de temps (Figure 11.3(f)) maintient sa capacité à naviguer le robot à travers les points de passage, bien que avec une efficacité réduite. Notamment, le contrôleur

[3] `https://github.com/acristoffers/phd-turtlebot/blob/main/workspace/gains.json`

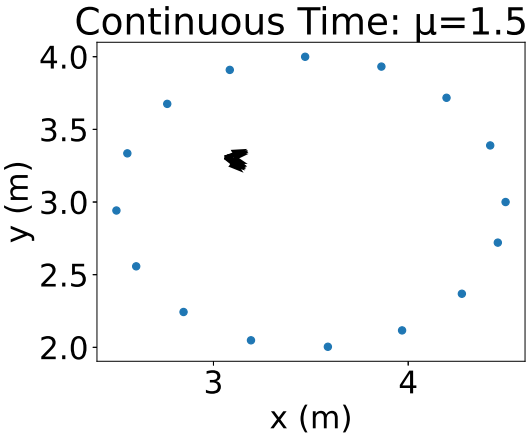
en temps continu (Figure 11.3(e)) montre une instabilité lorsque la durée de l'attaque s'étend à environ $0.8\,\mathrm{s}$, tandis que le contrôleur en échelle de temps reste stable jusqu'à environ $2.3\,\mathrm{s}$, dépassant son seuil opérationnel attendu.



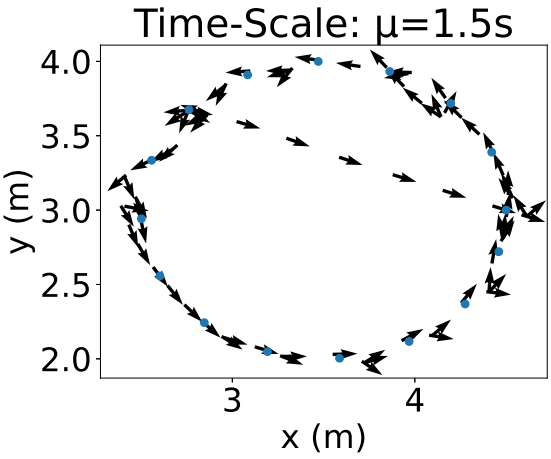
((a)) Contrôleur en temps continu avec une attaque DoS de 0.5 s



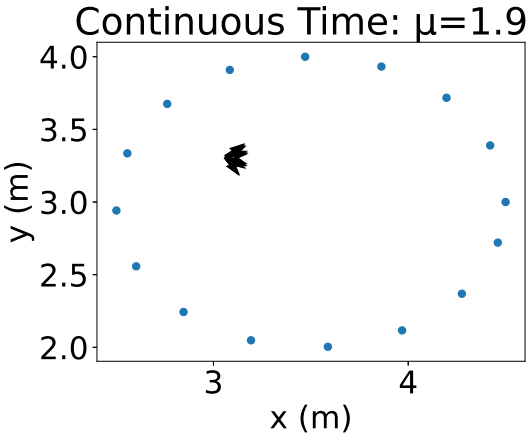
((b)) Contrôleur à échelle temporelle avec une attaque DoS de 0.5 s



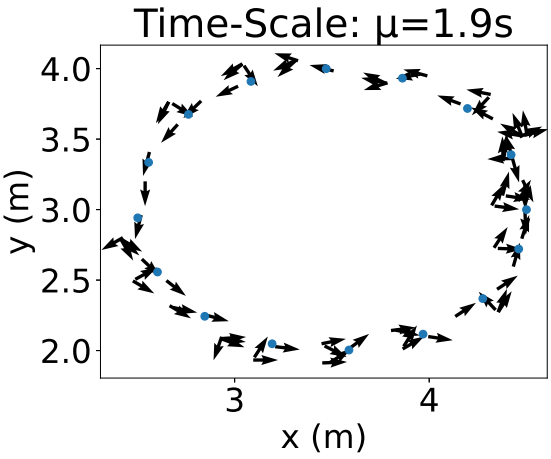
((c)) Contrôleur en temps continu avec une attaque DoS de 1.5 s



((d)) Contrôleur à échelle temporelle avec une attaque DoS de 1.5 s



((e)) Contrôleur en temps continu avec une attaque DoS de 1.9 s



((f)) Contrôleur à échelle temporelle avec une attaque DoS de 1.9 s

Figure 11.3 – Attaques DoS de différentes intensités sur les contrôleurs en Temps Continu et à Échelle Temporelle.

11.5 Conclusion

Dans ce chapitre, nous avons exploré les répercussions et les stratégies d'atténuation concernant les attaques par déni de service (DoS) au sein des systèmes cyber-physiques (CPS). Ces attaques représentent une menace significative en raison de leur capacité à perturber les réseaux de communication, un élément critique soutenant la fonctionnalité des CPS. En obstruant le flux des données de contrôle et de mesure, les attaques par DoS compromettent l'intégrité opérationnelle de ces systèmes.

L'enquête académique sur la vulnérabilité des CPS aux assauts par DoS a généré une pléthore de stratégies visant à comprendre et à neutraliser ces menaces. Cela inclut le déploiement de systèmes de détection d'intrusions (IDS), l'utilisation de ressources de file d'attente inutilisées pour la détection des attaques, et la mise en œuvre d'architectures réseau avancées conçues pour améliorer la sécurité et la résilience. De plus, des pipelines et des cadres de travail en apprentissage automatique exploitant l'informatique en brouillard et le réseautage défini par logiciel (SDN) ont été proposés pour détecter et atténuer efficacement ces attaques.

En outre, le chapitre se penche sur l'application de la Théorie des Échelles de Temps pour concevoir des contrôleurs et des observateurs capables de maintenir la stabilité du système sous conditions d'attaque par DoS. En traitant l'attaque comme une forme de délai et en ajustant le domaine temporel opérationnel du contrôleur et de l'observateur, la résilience du système est considérablement améliorée. Le fondement mathématique sous-jacent à cette approche, incluant la synthèse de systèmes μ fixes et variables, fournit un cadre solide pour assurer la stabilité du système malgré la présence d'attaques par DoS.

La validation sur un système réel, menée à travers une étude de cas détaillée impliquant un TurtleBot3 Waffle Pi, souligne l'applicabilité pratique des solutions proposées. Le dispositif expérimental a démontré l'efficacité des contrôleurs et des observateurs à échelles de temps dans la navigation du robot à travers des points de passage prédéfinis, même sous des intensités variables d'attaques par DoS. Ceci non seulement exemplifie la robustesse des stratégies de contrôle conçues, mais met également en lumière le potentiel de leur application dans des scénarios réels.

Selon la compréhension des auteurs, l'emploi de systèmes à échelles de temps pour contrer les attaques par DoS reste un domaine non investigué. Traditionnellement, les stratégies d'atténuation des DoS sont mises

en œuvre dans le domaine de la technologie de l'information (Information Technology (IT)), qui correspond directement au lieu de ces attaques. Néanmoins, l'intégration d'un contrôleur et d'un observateur résilients aux DoS permet aux concepteurs d'assurer une fonctionnalité continue du système, même dans les cas où les techniques d'atténuation des attaques basées sur l'IT ne sont pas totalement efficaces. Ainsi, cette approche sert de couche supplémentaire de défense, complétant efficacement les mesures de sécurité existantes.

En conclusion, ce chapitre contribue à l'enrichissement du corpus de connaissances concernant la défense contre les attaques par DoS dans les CPS. En intégrant des perspectives de la technologie de l'information et de l'automatisation des systèmes, et en exploitant des cadres théoriques avancés tels que la Théorie des Échelles de Temps, cette recherche délimite une approche holistique pour renforcer la résilience des CPS contre les attaques par DoS.

Le chapitre conclusif qui suit fournit un résumé de cette thèse, tout en offrant des aperçus et des projections pour les entreprises de recherche futures.

Chapter 12

Conclusion

Alors que nous concluons cette thèse, il est évident que la vulnérabilité des systèmes cyber-physiques aux attaques représente des risques significatifs pour leur intégrité et leur fonctionnalité. Ces systèmes, essentiels à l'infrastructure moderne, sont confrontés à des menaces par accès non autorisé et manipulation des couches réseau, pouvant entraîner la perturbation des dispositifs physiques. L'initiation d'une attaque peut compromettre l'ensemble du réseau, permettant aux adversaires de fabriquer des données et de saper les méthodologies de détection. Ainsi, le développement de mécanismes de détection robustes devient crucial pour préserver la sécurité, la fonctionnalité et la résilience de ces systèmes dans un monde axé sur la technologie.

Pour relever ces défis, les stratégies de cyberdéfense doivent englober à la fois la technologie de l'information (TI) et l'automatisation. Les défenses basées sur la TI, incluant les pare-feu, le cryptage et le contrôle d'accès, constituent la première couche de protection contre l'accès non autorisé. Cependant, pour contrer les tactiques sophistiquées employées par les attaquants, y compris l'exploitation des vulnérabilités humaines, une couche secondaire de défense émerge à travers l'automatisation. Cela implique le déploiement de contrôleurs résilients et de schémas de détection avancés conçus pour identifier et atténuer les impacts des cyberattaques. Cette approche à double couche souligne la complexité de la sécurisation des systèmes cyber-physiques et souligne la nécessité de mécanismes de défense ciblés adaptés à des types spécifiques d'attaques.

Dans le Chapitre 9, nous nous sommes concentrés sur l'attaque par Injection de Fausses Données. La contribution de ce chapitre est le développement d'une méthode basée sur les Observateurs Fonctionnels et le critère de stabilité de Lyapunov, contournant ainsi la dépendance en-

vers des méthodes susceptibles d'inexactitudes mathématiques et abordant les défis mathématiques posés par les matrices mal conditionnées fréquemment rencontrées dans les systèmes énergétiques. De plus, l'adoption d'Observateurs Fonctionnels offre une perspective alternative sur l'observabilité, ancrée dans la théorie des graphes, qui est moins contraignante que les approches conventionnelles.

Dans le Chapitre 10, nous avons changé notre focus vers l'attaque par Dynamique Nulle. La contribution de ce chapitre est l'utilisation du cadre Temps-Échelle pour développer un observateur à la fois simple et efficace en termes de calcul pour les applications en temps réel. Cette méthodologie fournit une approche novatrice pour la détection des attaques par dynamiques nulles, en utilisant une technique encore inexplorée. Le principe de fonctionnement de la méthodologie proposée est similaire à celui des méthodes basées sur le Porteur d'Ordre Zéro, rendant la nouvelle solution fondée sur des méthodes déjà établies.

Dans le Chapitre 11, nous appliquons la technique Temps-Échelle pour atténuer les attaques par Déni de Service. La contribution de ce chapitre est la conception d'un observateur et d'un contrôleur qui s'adaptent à l'attaque, permettant ainsi au système de rester opérationnel malgré celle-ci. Cette stratégie permet l'ajustement de l'intervalle d'échantillonnage de l'observateur et du contrôleur en réaction aux chutes de paquets causées par de telles attaques, permettant au système de contrôle d'adapter le signal de contrôle à l'attaque, dégradant les performances mais restant fonctionnel. Cette méthode améliore la résilience du système et ouvre une nouvelle voie pour l'atténuation des attaques Denial-of-Service (DoS).

Il existe certaines limitations au sein des méthodologies proposées. Spécifiquement, les techniques basées sur le cadre temps-échelle nécessitent que la matrice dynamique $\mathcal{A}$ du système soit auto-adjointe, une condition rarement satisfaite dans les scénarios pratiques. Heureusement, une transformation peut être appliquée aux systèmes ayant uniquement des valeurs propres réelles pour leur conférer cette propriété requise. Néanmoins, des recherches supplémentaires visant à éliminer cette condition tout en conservant la simplicité inhérente aux Inégalités Matricielles Linéaires (LMI) amélioreraient considérablement l'applicabilité de l'approche. De plus, la méthode pour déterminer le paramètre $\bar{\mu}$, qui dépend des propriétés spécifiques du système et est actuellement approximée sur la base du critère de Nyquist pour éviter l'aliasing, reste peu explorée. Établir une méthodologie définitive pour le calcul de ce paramètre pourrait affiner considérablement

les techniques proposées.

Concernant l'approche des Observateurs Fonctionnels, des opportunités d'avancement sont discernables. Une voie d'exploration est le développement d'observateurs qui surveillent des chemins variés menant au même résultat, intégrant ainsi une redondance au sein du système et améliorant sa fiabilité et sa résistance aux perturbations ou aux cyber-attaques. En outre, employer des ensembles différents de résultats pour surveiller un seul capteur pourrait présenter une méthode pour augmenter la complexité rencontrée par les attaquants employant des stratégies furtives. Ces approches exploitent la nature prévisible des modèles d'observation standard. En introduisant de la variabilité dans la dynamique d'observation, l'exposition du système aux attaques furtives pourrait être considérablement diminuée, renforçant ainsi ses mécanismes de défense.

En conclusion, cette thèse a démontré les avancées significatives et les voies potentielles pour des recherches futures dans le domaine de la sécurité des systèmes cyber-physiques, particulièrement à travers le prisme ciblé des attaques False Data Injection (FDI), Zero-Dynamics Attack (ZDA), et DoS. Les méthodologies proposées abordent non seulement les défis actuels posés par ces attaques mais ouvrent également la voie à des approches innovantes pour améliorer la robustesse et la résilience du système. L'intégration des Observateurs Fonctionnels, du cadre Temps-Échelle, et de l'application du critère de stabilité de Lyapunov, entre autres, représentent des contributions significatives au domaine, offrant de nouvelles perspectives et solutions aux problèmes complexes rencontrés par les systèmes cyber-physiques.

En regardant vers l'avenir, il est évident que le chemin vers la sécurisation complète de ces systèmes est en cours et nécessite un effort continu. Alors que cette recherche contribue à la base de connaissances, elle souligne également la nécessité d'approches interdisciplinaires qui englobent à la fois les perspectives Information Technology (IT) et Operational Technology (OT). La nature dynamique des menaces cybernétiques exige des stratégies de défense adaptables, intelligentes et multifacettes. Ainsi, ce travail sert non seulement de témoignage des progrès réalisés mais aussi comme un appel à l'action pour les futurs chercheurs à s'appuyer sur ces découvertes, à explorer des territoires inexplorés et à renforcer les défenses de notre monde de plus en plus interconnecté. En faisant avancer notre compréhension et nos méthodologies, nous pouvons aspirer à un avenir où les systèmes cyber-physiques fonctionnent de manière sécurisée et fiable,

protégeant les infrastructures critiques et les services qui sous-tendent la société moderne.

Bibliography

Abusorrah, Abdullah et al. (Mar. 2019). “Minimax-Regret Robust Defensive Strategy Against False Data Injection Attacks”. In: *IEEE Transactions on Smart Grid* 10.2, pp. 2068–2079. doi: 10.1109/TSG.2017.2788040.

Agarwal, Ravi et al. (2002). “Dynamic Equations on Time Scales: A Survey”. In: *Journal of Computational and Applied Mathematics* 141.1-2, pp. 1–26. doi: 10.1016/S0377-0427(01)00432-0.

Aguirre, Luis A., Leonardo L. Portes, and Christophe Letellier (2018). “Structural, Dynamical and Symbolic Observability: From Dynamical Systems to Networks”. In: *PLoS ONE* 13.10, pp. 1–21. doi: 10.1371/journal.pone.0206180.

Ahmed, Mohiuddin and Al-Sakib Khan Pathan (Dec. 2020). “False Data Injection Attack (FDIA): An Overview and New Metrics for Fair Evaluation of Its Countermeasure”. In: *Complex Adaptive Systems Modeling* 8.1, p. 4. doi: 10.1186/s40294-020-00070-w.

Al Hammadi, Ahmed Y., Chan Yeob Yeun, and Ernesto Damiani (Nov. 2020). “Novel EEG Risk Framework to Identify Insider Threats in National Critical Infrastructure Using Deep Learning Techniques”. In: *2020 IEEE International Conference on Services Computing (SCC)*. 2020 IEEE International Conference on Services Computing (SCC). Beijing, China: IEEE, pp. 469–471. doi: 10.1109/SCC49832.2020.00071.

Alsulami, Abdulaziz A. and Saleh Zein-Sabatto (Jan. 27, 2021). “Resilient Cyber-Security Approach For Aviation Cyber-Physical Systems Protection Against Sensor Spoofing Attacks”. In: *2021 IEEE 11th Annual Computing and Communication Workshop and Conference (CCWC)*. 2021 IEEE 11th Annual Computing and Communication Workshop and Conference (CCWC). NV, USA: IEEE, pp. 0565–0571. doi: 10.1109/CCWC51732.2021.9376158.

Anand, C S and Ravi Shanker (Mar. 3, 2023). “Advancing Crypto Ransomware with Multi Level Extortion: A Peril to Critical Infrastructure”. In: *2023 2nd International Conference for Innovation in Technology (INOCON)*. 2023 2nd International Conference for Innovation in Technology (INOCON). Bangalore, India: IEEE, pp. 1–5. doi: 10.1109/INOCON57975.2023.10100971.

Antezana, Jorge, Pedro Massey, and Demetrio Stojanoff (Nov. 19, 2004). *Jensen’s Inequality and Majorization*. arXiv: math/0411442. url: http://arxiv.org/abs/math/0411442 (visited on 05/16/2022). preprint.

Baniamerian, Amir, Khashayar Khorasani, and Nader Meskin (Aug. 2020). “Monitoring and Detection of Malicious Adversarial Zero Dynamics Attacks in Cyber-Physical Systems”. In: *2020 IEEE Conference on Control Technology and Applications (CCTA)*. 2020 IEEE Conference on Control Technology and Applications (CCTA). Montreal, QC, Canada: IEEE, pp. 726–731. doi: 10.1109/CCTA41146.2020.9206295.

Ben Nasser, Bacem et al. (Aug. 3, 2022). “Time Scale Observability and Constructibility of Linear Dynamic Equations”. In: *International Journal of Control* 95.8, pp. 1994–2004. doi: 10.1080/00207179.2021.1890823.

Bernstein, Dennis S. (2009). *Matrix Mathematics: Theory, Facts, and Formulas*. 2nd ed. Princeton, N.J: Princeton University Press. 1139 pp. isbn: 978-0-691-13287-7.

Bestehorn, Markus and Theodor Borsche (Oct. 2014). “Balancing Power Consumption and Production in Smart Grids”. In: *IEEE PES Innovative Smart Grid Technologies, Europe*. 2014 IEEE PES Innovative Smart Grid Technologies Conference Europe (ISGT-Europe). Istanbul, Turkey: IEEE, pp. 1–6. doi: 10.1109/ISGTEurope.2014.7028827.

Bouheroum, Ayoub et al. (Oct. 12, 2022). “A Formal Integrated Approach for Cyber Physical Systems”. In: *2022 4th International Conference on Pattern Analysis and Intelligent Systems (PAIS)*. 2022 4th International Conference on Pattern Analysis and Intelligent Systems (PAIS). Oum El Bouaghi, Algeria: IEEE, pp. 1–7. doi: 10.1109/PAIS56586.2022.9946900.

Bout, Emilie, Valeria Loscri, and Antoine Gallais (Sept. 2020). “Energy and Distance Evaluation for Jamming Attacks in Wireless Networks”. In: *2020 IEEE/ACM 24th International Symposium on Distributed Simulation and Real Time Applications (DS-RT)*. 2020 IEEE/ACM 24th International Symposium on Distributed Simulation and Real Time Applica-

tions (DS-RT). Prague, Czech Republic: IEEE, pp. 1–5. doi: 10.1109/DS-RT50469.2020.9213652.

Boyd, S. et al. (1994). *Linear Matrix Inequalities In Systems And Control Theory*. Philadelphia, PA: SIAM Studies in Applied Mathematics.

Briat, Corentin and Alexandre Seuret (Oct. 2012). "A Looped-Functional Approach for Robust Stability Analysis of Linear Impulsive Systems". In: *Systems & Control Letters* 61.10, pp. 980–988. doi: 10.1016/j.sysconle.2012.07.008.

Cameron, Calum et al. (May 2019). "Using Self-Organizing Architectures to Mitigate the Impacts of Denial-of-Service Attacks on Voltage Control Schemes". In: *IEEE Transactions on Smart Grid* 10.3, pp. 3010–3019. doi: 10.1109/TSG.2018.2817046.

Cecil, J. (Feb. 17, 2017). "Internet of Things (Iot)-Based Cyber–Physical Frameworks for Advanced Manufacturing and Medicine". In: *Internet of Things and Data Analytics Handbook*. Ed. by Hwaiyu Geng. 1st ed. Wiley, pp. 545–561. doi: 10.1002/9781119173601.ch33.

Cetinkaya, Ahmet, Hideaki Ishii, and Tomohisa Hayakawa (Feb. 22, 2019). "An Overview on Denial-of-Service Attacks in Control Systems: Attack Models and Security Analyses". In: *Entropy* 21.2, p. 210. doi: 10.3390/e21020210.

Chen, Guan-Rong (2013). "Problems and Challenges in Control Theory under Complex Dynamical Network Environments". In: *Acta Automatica Sinica* 39.4, pp. 312–321. doi: 10.1016/s1874-1029(13)60032-4.

Cheng, Zheyuan and Mo-Yuen Chow (Nov. 2020). "Resilient Collaborative Distributed Energy Management System Framework for Cyber-Physical DC Microgrids". In: *IEEE Transactions on Smart Grid* 11.6, pp. 4637–4649. doi: 10.1109/TSG.2020.3001059.

Chlela, Martine et al. (Sept. 2018). "Fallback Control for Isochronous Energy Storage Systems in Autonomous Microgrids Under Denial-of-Service Cyber-Attacks". In: *IEEE Transactions on Smart Grid* 9.5, pp. 4702–4711. doi: 10.1109/TSG.2017.2667586.

Choeum, Daranith and Dae Hyun Choi (Jan. 2021). "Vulnerability Assessment of Conservation Voltage Reduction to Load Redistribution Attack in Unbalanced Active Distribution Networks". In: *IEEE Transactions on Industrial Informatics* 17.1, pp. 473–483. doi: 10.1109/TII.2020.2980590.

Cowan, Noah J. et al. (2012). "Nodal Dynamics, Not Degree Distributions, Determine the Structural Controllability of Complex Networks". In: *PLoS ONE* 7.6. doi: 10.1371/journal.pone.0038398.

Davis, John M. et al. (2010). “Algebraic and Dynamic Lyapunov Equations on Time Scales”. In: *Proceedings of the Annual Southeastern Symposium on System Theory*, pp. 329–334. doi: 10.1109/SSST.2010.5442815.

Devi, Boddu Rama and E. Susmitha (Nov. 2017). “Energy Distribution and Pricing Based on Load Demand Taxonomy in a Smart Grid Tree Network”. In: *2017 International Conference on Inventive Computing and Informatics (ICICI)*. 2017 International Conference on Inventive Computing and Informatics (ICICI). Coimbatore: IEEE, pp. 230–235. doi: 10.1109/ICICI.2017.8365344.

Ding, Derui et al. (2018). “A Survey on Security Control and Attack Detection for Industrial Cyber-Physical Systems”. In: *Neurocomputing* 275, pp. 1674–1683. doi: 10.1016/j.neucom.2017.10.009.

Ding, Kemi et al. (June 2019). “DoS Attacks on Remote State Estimation With Asymmetric Information”. In: *IEEE Transactions on Control of Network Systems* 6.2, pp. 653–666. doi: 10.1109/TCNS.2018.2867157.

Ding, Peili et al. (Oct. 2017). “DoS Attacks in Electrical Cyber-Physical Systems: A Case Study Using TrueTime Simulation Tool”. In: *2017 Chinese Automation Congress (CAC)*. 2017 Chinese Automation Congress (CAC). Jinan: IEEE, pp. 6392–6396. doi: 10.1109/CAC.2017.8243929.

Djenna, Amir and Diamel Eddine Saidouni (Oct. 2018). “Cyber Attacks Classification in IoT-Based-Healthcare Infrastructure”. In: *2018 2nd Cyber Security in Networking Conference (CSNet)*. 2018 2nd Cyber Security in Networking Conference (CSNet). Paris: IEEE, pp. 1–4. doi: 10.1109/CSNET.2018.8602974.

Dorfler, F., Michael Chertkov, and Francesco Bullo (Feb. 2013). “Synchronization in Complex Oscillator Networks and Smart Grids”. In: *Proceedings of the National Academy of Sciences* 110.6, pp. 2005–2010. doi: 10.1073/pnas.1212134110.

Duan, Guang-Ren and Hai-Hua Yu (2013). *LMIs in Control Systems: Analysis, Design and Applications*. Boca Raton, Fla.: CRC Press. 453 pp. isbn: 978-1-4665-8300-9 978-1-4665-8299-6 978-1-4665-8303-0.

Emami, Kianoush et al. (Dec. 2013). “A Fault Detection Technique for Dynamical Systems”. In: *2013 IEEE 8th International Conference on Industrial and Information Systems*. 2013 IEEE 8th International Conference on Industrial and Information Systems (ICIIS). Peradeniya, Sri Lanka: IEEE, pp. 201–206. doi: 10.1109/ICIInfS.2013.6731981.

E Sousa, Álan, Nadhir Messai, and Noureddine Manamanni (July 2022). “Load-Altering Attack Detection on Smart Grid Using Functional Ob-

servers". In: *International Journal of Critical Infrastructure Protection* 37, p. 100518. doi: 10.1016/j.ijcip.2022.100518.

Farraj, Abdallah, Eman Hammad, and Deepa Kundur (Mar. 2018). "A Cyber-Physical Control Framework for Transient Stability in Smart Grids". In: *IEEE Transactions on Smart Grid* 9.2, pp. 1205–1215. doi: 10.1109/TSG.2016.2581588.

Foundation, US National Science (Jan. 2022). *Cyber-Physical Systems (CPS)*. url: https://www.nsf.gov/publications/pub_summ.jsp?ods_key=nsf21551&org=NSF.

Garagad, Vishwanath G., Nalini C. Iyer, and Heera G. Wali (July 2020). "Data Integrity: A Security Threat for Internet of Things and Cyber-Physical Systems". In: *2020 International Conference on Computational Performance Evaluation (ComPE)*. 2020 International Conference on Computational Performance Evaluation (ComPE). Shillong, India: IEEE, pp. 244–249. doi: 10.1109/ComPE49325.2020.9200170.

Giraldo, Jairo et al. (2018). "A Survey of Physics-Based Attack Detection in Cyber-Physical Systems". In: *ACM Computing Surveys* 51.4. doi: 10.1145/3203245.

Gorbenko, Anna and Vladimir Popov (Sept. 4, 2022). "Zero-Day Attacks Detection Using an Analysis of Mobile Robot Motor Primitives". In: *2022 International Russian Automation Conference (RusAutoCon)*. 2022 International Russian Automation Conference (RusAutoCon). Sochi, Russian Federation: IEEE, pp. 278–283. doi: 10.1109/RusAutoCon54946.2022.9896253.

Gupta, Himanshu et al. (Dec. 2019). "Impact of Side Channel Attack in Information Security". In: *2019 International Conference on Computational Intelligence and Knowledge Economy (ICCIKE)*. 2019 International Conference on Computational Intelligence and Knowledge Economy (ICCIKE). Dubai, United Arab Emirates: IEEE, pp. 291–295. doi: 10.1109/ICCIKE47802.2019.9004435.

Haes Alhelou, Hassan, Mohamad Esmail Hamedani Golshan, and Nikos D. Hatziargyriou (Nov. 2019). "A Decentralized Functional Observer Based Optimal LFC Considering Unknown Inputs, Uncertainties, and Cyber-Attacks". In: *IEEE Transactions on Power Systems* 34.6, pp. 4408–4417. doi: 10.1109/TPWRS.2019.2916558.

Haghighi, Mohammad Sayad et al. (2023). "Cyber Attacks via Consumer Electronics: Studying the Threat of Covert Malware in Smart and Au-

tonomous Vehicles". In: *IEEE Transactions on Consumer Electronics*, pp. 1–1. doi: 10.1109/TCE.2023.3297965.

Halabi, Talal and Mohammad Zulkernine (July 2023). "The Ultimate Battle Against Zero-Day Exploits: Toward Fully Autonomous Cyber-Physical Defense". In: *2023 IEEE International Conference on Software Services Engineering (SSE)*. 2023 IEEE International Conference on Software Services Engineering (SSE). Chicago, IL, USA: IEEE, pp. 256–261. doi: 10.1109/SSE60056.2023.00041.

He, Guoqing et al. (Oct. 22, 2021). "Analysis of Cyber Attacks on Wind Power Generation Systems in Damping Inter-area Oscillations of Power Systems". In: *2021 IEEE 5th Conference on Energy Internet and Energy System Integration (EI2)*. 2021 IEEE 5th Conference on Energy Internet and Energy System Integration (EI2). Taiyuan, China: IEEE, pp. 2404–2409. doi: 10.1109/EI252483.2021.9713387.

Hossain, Md Musabbir et al. (Jan. 2022). "Bandwidth Allocation-Based Distributed Event-Triggered LFC for Smart Grids Under Hybrid Attacks". In: *IEEE Transactions on Smart Grid* 13.1, pp. 820–830. doi: 10.1109/TSG.2021.3118801.

Hu, Zhijian et al. (Apr. 2022). "Resilient Distributed Fuzzy Load Frequency Regulation for Power Systems Under Cross-Layer Random Denial-of-Service Attacks". In: *IEEE Transactions on Cybernetics* 52.4, pp. 2396–2406. doi: 10.1109/TCYB.2020.3005283.

Huang, Yu-Lun et al. (Oct. 2009). "Understanding the Physical and Economic Consequences of Attacks on Control Systems". In: *International Journal of Critical Infrastructure Protection* 2.3, pp. 73–83. doi: 10.1016/j.ijcip.2009.06.001.

Humayed, Abdulmalik and Bo Luo (Apr. 14, 2015). "Cyber-Physical Security for Smart Cars: Taxonomy of Vulnerabilities, Threats, and Attacks". In: *Proceedings of the ACM/IEEE Sixth International Conference on Cyber-Physical Systems*. ICCPS '15: ACM/IEEE 6th International Conference on Cyber-Physical Systems. Seattle Washington: ACM, pp. 252–253. doi: 10.1145/2735960.2735992.

Ibarra, Jaime et al. (Jan. 2019). "Ransomware Impact to SCADA Systems and Its Scope to Critical Infrastructure". In: *2019 IEEE 12th International Conference on Global Security, Safety and Sustainability (ICGS3)*. 2019 IEEE 12th International Conference on Global Security, Safety and Sustainability (ICGS3). London, United Kingdom: IEEE, pp. 1–12. doi: 10.1109/ICGS3.2019.8688299.

IEEE (Oct. 2013). *Smart Grid Research: Control Systems - IEEE Vision for Smart Grid Control: 2030 and Beyond Roadmap*. doi: 10.1109/IEEESTD.2013.6648362.

Ikany, Joris and Husin Jazri (Aug. 2019). "A Symptomatic Framework to Predict the Risk of Insider Threats". In: *2019 International Conference on Advances in Big Data, Computing and Data Communication Systems (icABCD)*. 2019 International Conference on Advances in Big Data, Computing and Data Communication Systems (icABCD). Winterton, South Africa: IEEE, pp. 1–5. doi: 10.1109/ICABCD.2019.8851020.

Isermann, Rolf (2006). *Fault-Diagnosis Systems: An Introduction from Fault Detection to Fault Tolerance*. Berlin ; New York: Springer. 475 pp. isbn: 978-3-540-24112-6.

Islam, Syed Imranul, Cheng-Chew Lim, and Peng Shi (Aug. 2020). "Robust Fault Detection of T-S Fuzzy Systems with Time-Delay Using Fuzzy Functional Observer". In: *Fuzzy Sets and Systems* 392, pp. 1–23. doi: 10.1016/j.fss.2019.03.020.

Jena, Satabdy, Narayana Prasad Padhy, and Josep M. Guerrero (2023). "Multi-Layered Coordinated Countermeasures for DC Microgrid Clusters Under Man in the Middle Attack". In: *IEEE Transactions on Industry Applications*, pp. 1–14. doi: 10.1109/TIA.2023.3308557.

Jennings, Les S., Tyrone Lucius Fernando, and Hieu Minh Trinh (Dec. 2011). "Existence Conditions for Functional Observability From an Eigenspace Perspective". In: *IEEE Transactions on Automatic Control* 56.12, pp. 2957–2961. doi: 10.1109/TAC.2011.2160019.

Jovanov, Ilija and Miroslav Pajic (Dec. 2019). "Relaxing Integrity Requirements for Attack-Resilient Cyber-Physical Systems". In: *IEEE Transactions on Automatic Control* 64.12, pp. 4843–4858. doi: 10.1109/TAC.2019.2898510.

Kapoor, Prateek, Ankur Vora, and Kyoung-Don Kang (Aug. 2018). "Detecting and Mitigating Spoofing Attack Against an Automotive Radar". In: *2018 IEEE 88th Vehicular Technology Conference (VTC-Fall)*. 2018 IEEE 88th Vehicular Technology Conference (VTC-Fall). Chicago, IL, USA: IEEE, pp. 1–6. doi: 10.1109/VTCFall.2018.8690734.

Kaviani, Ramin and Kory W. Hedman (Jan. 2021). "A Detection Mechanism against Load-Redistribution Attacks in Smart Grids". In: *IEEE Transactions on Smart Grid* 12.1, pp. 704–714. doi: 10.1109/TSG.2020.3017562.

Khazaei, Javad and M. Hadi Amini (Dec. 2021). "Protection of Large-Scale Smart Grids against False Data Injection Cyberattacks Leading to Black-

outs". In: *International Journal of Critical Infrastructure Protection* 35, p. 100457. doi: 10.1016/j.ijcip.2021.100457.

Kim, Daehan, Kunhee Ryu, and Juhoon Back (2020). "Security Enhancement of Sampled-Data Systems: Zero Assignment via Generalized Sampler". In: *IFAC-PapersOnLine* 53.2, pp. 3482–3487. doi: 10.1016/j.ifacol.2020.12.1696.

— (Jan. 29, 2021). "Zero-Dynamics Attack on Wind Turbines and Countermeasures Using Generalized Hold and Generalized Sampler". In: *Applied Sciences* 11.3, p. 1257. doi: 10.3390/app11031257.

Kim, Daehan, Kunhee Ryu, Jung Hoon Kim, et al. (2021). "Zero Assignment via Generalized Sampler: A Countermeasure Against Zero-Dynamics Attack". In: *IEEE Access* 9, pp. 109932–109942. doi: 10.1109/ACCESS.2021.3101637.

Kim, Jihan, Juhoon Back, et al. (Mar. 2020). "Neutralizing Zero Dynamics Attack on Sampled-Data Systems via Generalized Holds". In: *Automatica* 113, p. 108778. doi: 10.1016/j.automatica.2019.108778.

Kim, Jihan and Hyungbo Shim (Sept. 2019a). "A Countermeasure against Zero-dynamics Sensor Attack via Generalized Hold Feedback". In: *2019 58th Annual Conference of the Society of Instrument and Control Engineers of Japan (SICE)*. 2019 58th Annual Conference of the Society of Instrument and Control Engineers of Japan (SICE). Hiroshima, Japan: IEEE, pp. 663–668. doi: 10.23919/SICE.2019.8859930.

— (2019b). "A Countermeasure against Zero-dynamics Sensor Attack via Generalized Hold Feedback". In: *2019 58th Annual Conference of the Society of Instrument and Control Engineers of Japan, SICE 2019*, pp. 663–668. doi: 10.23919/SICE.2019.8859930.

Kimani, Kenneth, Vitalice Oduol, and Kibet Langat (June 2019). "Cyber Security Challenges for IoT-based Smart Grid Networks". In: *International Journal of Critical Infrastructure Protection* 25, pp. 36–49. doi: 10.1016/j.ijcip.2019.01.001.

Knowles, William et al. (June 2015). "A Survey of Cyber Security Management in Industrial Control Systems". In: *International Journal of Critical Infrastructure Protection* 9, pp. 52–80. doi: 10.1016/j.ijcip.2015.02.002.

Kraus, Fritz (Dec. 1936). "Über konvexe Matrixfunktionen". In: *Mathematische Zeitschrift* 41.1, pp. 18–42. doi: 10.1007/BF01180403.

Krishnamurthy, Rajesh and J. Cecil (Apr. 2018). "Next Generation Cyber Physical Frameworks for Electronics Manufacturing". In: *2018 Annual*

IEEE International Systems Conference (SysCon). 2018 Annual IEEE International Systems Conference (SysCon). Vancouver, BC: IEEE, pp. 1–6. doi: 10.1109/SYSCON.2018.8369590.

Lakshminarayana, Subhash et al. (Jan. 2021). "Data-Driven False Data Injection Attacks Against Power Grids: A Random Matrix Approach". In: *IEEE Transactions on Smart Grid* 12.1, pp. 635–646. doi: 10.1109/TSG.2020.3011391.

Lee, Edward A. and Sanjit A. Seshia (2017). *Introduction to Embedded Systems: A Cyber-Physical Systems Approach*. Second edition. Cambridge, Massachuetts: MIT Press. 537 pp. isbn: 978-0-262-53381-2.

Lee, Joowon, Junsoo Kim, and Hyungbo Shim (Oct. 13, 2020). "Zero-Dynamics Attack on Homomorphically Encrypted Control System". In: *2020 20th International Conference on Control, Automation and Systems (ICCAS)*. 2020 20th International Conference on Control, Automation and Systems (ICCAS). Busan, Korea (South): IEEE, pp. 385–390. doi: 10.23919/ICCAS50221.2020.9268374.

Liberati, Francesco, Emanuele Garone, and Alessandro Di Giorgio (2021). "Review of Cyber-Physical Attacks in Smart Grids: A System-Theoretic Perspective". In: *Electronics (Switzerland)* 10.10, pp. 1–39. doi: 10.3390/electronics10101153.

Liu, Jianzhe, Xiaonan Lu, and Jianhui Wang (July 2019). "Resilience Analysis of DC Microgrids Under Denial of Service Threats". In: *IEEE Transactions on Power Systems* 34.4, pp. 3199–3208. doi: 10.1109/TPWRS.2019.2897499.

Liu, Zhaoxi and Lingfeng Wang (Mar. 2021). "Defense Strategy against Load Redistribution Attacks on Power Systems Considering Insider Threats". In: *IEEE Transactions on Smart Grid* 12.2, pp. 1529–1540. doi: 10.1109/TSG.2020.3023426.

Luenberger, D. (Apr. 1966). "Observers for Multivariable Systems". In: *IEEE Transactions on Automatic Control* 11.2, pp. 190–197. doi: 10.1109/TAC.1966.1098323.

Makar, A. B. et al. (June 1975). "Formate Assay in Body Fluids: Application in Methanol Poisoning". In: *Biochemical Medicine* 13.2, pp. 117–126. doi: 10.1016/0006-2944(75)90147-7.

Malamud, S.M. (Jan. 2001). "A Converse to the Jensen Inequality, Its Matrix Extensions and Inequalities for Minors and Eigenvalues". In: *Linear Algebra and its Applications* 322.1-3, pp. 19–41. doi: 10.1016/S0024-3795(00)00207-X.

Mao, Yanbing and Emrah Akyol (Oct. 2018). “Detectability of Cooperative Zero-Dynamics Attack”. In: *2018 56th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*. 2018 56th Annual Allerton Conference on Communication, Control, and Computing (Allerton). Monticello, IL, USA: IEEE, pp. 227–234. doi: 10.1109/ALLERTON.2018.8636053.

Mao, Yanbing, Hamidreza Jafarnejadsani, et al. (Sept. 2020). “Novel Stealthy Attack and Defense Strategies for Networked Control Systems”. In: *IEEE Transactions on Automatic Control* 65.9, pp. 3847–3862. doi: 10.1109/TAC.2020.2997363.

Miller, Thomas et al. (Dec. 2021). “Looking Back to Look Forward: Lessons Learnt from Cyber-Attacks on Industrial Control Systems”. In: *International Journal of Critical Infrastructure Protection* 35, p. 100464. doi: 10.1016/j.ijcip.2021.100464.

Montanari, Arthur N. and Luis A. Aguirre (Dec. 2020). “Observability of Network Systems: A Critical Review of Recent Results”. In: *Journal of Control, Automation and Electrical Systems* 31.6, pp. 1348–1374. doi: 10.1007/s40313-020-00633-5.

Montanari, Arthur Noronha (Dec. 2021). “Observability of Dynamical Networks”. PhD thesis. Universidade Federal de Minas Gerais (UFMG). 150 pp.

Mrabet, Zakaria El et al. (2018). “Cyber-Security in Smart Grid: Survey and Challenges”. In: *Computers and Electrical Engineering* 67, pp. 469–482. doi: 10.1016/j.compeleceng.2018.01.015.

Murdoch, P. (Oct. 1973). “Pole and Zero Assignment by Proportional Feedback”. In: *IEEE Transactions on Automatic Control* 18.5, pp. 542–542. doi: 10.1109/TAC.1973.1100381.

Nishikawa, Takashi and Adilson E. Motter (Jan. 2015). “Comparative Analysis of Existing Models for Power-Grid Synchronization”. In: *New Journal of Physics* 17.1, p. 015012. doi: 10.1088/1367-2630/17/1/015012.

Office of the National Coordinator for Smart Grid Interoperability (Feb. 2012). *NIST Framework and Roadmap for Smart Grid Interoperability Standards, Release 2.0*. Gaithersburg, MD: National Institute of Standards and Technology, p. 227. doi: 10.6028/NIST.SP.1108r2.

Oliva, Gabriele, Sebastian Cioaba, and Christoforos N. Hadjicostis (Dec. 2018). “Distributed Calculation of Edge-Disjoint Spanning Trees for Robustifying Distributed Algorithms Against Man-in-the-Middle Attacks”.

In: *IEEE Transactions on Control of Network Systems* 5.4, pp. 1646–1656. doi: 10.1109/TCNS.2017.2746344.

Ozarar, Mert, Asuman Akansu, and Burkay Hasbay (Dec. 2, 2021). "Impact of Cyber Maturity Level on Health Sector". In: *2021 International Conference on Information Security and Cryptology (ISCTURKEY)*. 2021 International Conference on Information Security and Cryptology (ISCTURKEY). Ankara, Turkey: IEEE, pp. 127–131. doi: 10.1109/ISCTURKEY53027.2021.9654395.

Pang, Zhong-Hua et al. (2019). "Secure Networked Predictive Control Under DoS Attacks". In: *Networked Predictive Control of Systems with Communication Constraints and Cyber Attacks*. Singapore: Springer Singapore, pp. 205–219. doi: 10.1007/978-981-13-0520-7_12.

Park, Gyunghoon et al. (Dec. 2019). "Stealthy Adversaries Against Uncertain Cyber-Physical Systems: Threat of Robust Zero-Dynamics Attack". In: *IEEE Transactions on Automatic Control* 64.12, pp. 4907–4919. doi: 10.1109/TAC.2019.2903429.

Pasha, Syed Ahmed and Ayesha Ayub (Sept. 2021). "Zero-Dynamics Attacks on Networked Control Systems". In: *Journal of Process Control* 105, pp. 99–107. doi: 10.1016/j.jprocont.2021.07.010.

Pedramnia, Kiyana and Shayan Shojaei (Dec. 2020). "Detection of False Data Injection Attack in Smart Grid Using Decomposed Nearest Neighbor Techniques". In: *2020 10th Smart Grid Conference (SGC)*. 2020 10th Smart Grid Conference (SGC), pp. 1–6. doi: 10.1109/SGC52076.2020.9335732.

Pham, Thanh Ngoc, Amanullah Maung Than Oo, and Hieu Trinh (Feb. 2021). "Detecting and Isolating False Data Injection Attacks on Electric Vehicles of Smart Grids Using Distributed Functional Observers". In: *IET Generation, Transmission & Distribution* 15.4, pp. 762–779. doi: 10.1049/gtd2.12057.

Rios-Ruiz, Carlos et al. (July 2019). "Finite Time Functional Observers for Descriptor Systems. Application to Fault Tolerant Control". In: *2019 27th Mediterranean Conference on Control and Automation (MED)*. 2019 27th Mediterranean Conference on Control and Automation (MED). Akko, Israel: IEEE, pp. 165–170. doi: 10.1109/MED.2019.8798552.

Rushanan, Michael et al. (May 2014). "SoK: Security and Privacy in Implantable Medical Devices and Body Area Networks". In: *2014 IEEE Symposium on Security and Privacy*. 2014 IEEE Symposium on Secu-

rity and Privacy (SP). San Jose, CA: IEEE, pp. 524–539. doi: 10.1109/SP.2014.40.

Saad, Ahmed et al. (Nov. 2020). “On the Implementation of IoT-Based Digital Twin for Networked Microgrids Resiliency Against Cyber Attacks”. In: *IEEE Transactions on Smart Grid* 11.6, pp. 5138–5150. doi: 10.1109/TSG.2020.3000958.

Shi, Hanzhang, Linbo Xie, and Li Peng (May 2021). “Detection of False Data Injection Attacks in Smart Grid Based on a New Dimensionality-Reduction Method”. In: *Computers & Electrical Engineering* 91, p. 107058. doi: 10.1016/j.compeleceng.2021.107058.

Shi, Mengxuan et al. (May 2021). “Observer-Based Resilient Integrated Distributed Control Against Cyberattacks on Sensors and Actuators in Islanded AC Microgrids”. In: *IEEE Transactions on Smart Grid* 12.3, pp. 1953–1963. doi: 10.1109/TSG.2021.3050203.

Sreeram, T. S. and S. Krishna (Nov. 2019). “Managing False Data Injection Attacks During Contingency of Secured Meters”. In: *IEEE Transactions on Smart Grid* 10.6, pp. 6945–6953. doi: 10.1109/TSG.2019.2914974.

Tanaka, Kazuo and Hua O. Wang (Sept. 12, 2001). *Fuzzy Control Systems Design and Analysis*. New York, USA: John Wiley & Sons, Inc. doi: 10.1002/0471224596.

Tang, Wan, Zhengdao Zhang, and Linbo Xie (May 20, 2023). “Short Interval Replay Attack Detection for Industrial Cyber-Physical Systems Based on Wavelet Coherence”. In: *2023 35th Chinese Control and Decision Conference (CCDC)*. 2023 35th Chinese Control and Decision Conference (CCDC). Yichang, China: IEEE, pp. 3082–3087. doi: 10.1109/CCDC58219.2023.10326871.

Teixeira, Andre et al. (Oct. 2012). “Revealing Stealthy Attacks in Control Systems”. In: *2012 50th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*. 2012 50th Annual Allerton Conference on Communication, Control, and Computing (Allerton). Monticello, IL, USA: IEEE, pp. 1806–1813. doi: 10.1109/Allerton.2012.6483441.

Teixeira, André et al. (Jan. 2015). “A Secure Control Framework for Resource-Limited Adversaries”. In: *Automatica* 51, pp. 135–148. doi: 10.1016/j.automatica.2014.10.067.

Tran, H. M. and H. Trinh (2016). “Minimal-Order Functional Observer-Based Residual Generators for Fault Detection and Isolation of Dynamical Systems”. In: *Mathematical Problems in Engineering* 2016, pp. 1–17. doi: 10.1155/2016/2740645.

Tran, H. M., H. Trinh, and P. T. Nam (2015). “Functional Observer-Based Fault Detection of Time-Delay Systems via an LMI Approach”. In: *2015 Australian Control Conference, AUCC 2015*, pp. 194–199.

Tran, Hieu Manh and Hieu Trinh (Mar. 2019). “Distributed Functional Observer Based Fault Detection for Interconnected Time-Delay Systems”. In: *IEEE Systems Journal* 13.1, pp. 940–951. doi: 10.1109/JSYST.2017.2759257.

Trinh, H. et al. (Dec. 2013). “Fault Detection of Dynamical Systems Using First-Order Functional Observers”. In: *2013 IEEE 8th International Conference on Industrial and Information Systems*. 2013 IEEE 8th International Conference on Industrial and Information Systems (ICIIS). Peradeniya, Sri Lanka: IEEE, pp. 197–200. doi: 10.1109/ICIInfS.2013.6731980.

Umsonst, David et al. (May 31, 2022). *A Bayesian Nash Equilibrium-Based Moving Target Defense against Stealthy Sensor Attacks*. doi: 10.48550/arXiv.2111.06682. preprint.

Wang, Pengyuan and Manimaran Govindarasu (July 2020). “Multi-Agent Based Attack-Resilient System Integrity Protection for Smart Grid”. In: *IEEE Transactions on Smart Grid* 11.4, pp. 3447–3456. doi: 10.1109/TSG.2020.2970755.

Wang, Zhenhua, Haibo He, et al. (Feb. 2021). “Coordinated Topology Attacks in Smart Grid Using Deep Reinforcement Learning”. In: *IEEE Transactions on Industrial Informatics* 17.2, pp. 1407–1415. doi: 10.1109/TII.2020.2994977.

Wang, Zhenhua, Wei Xie, et al. (Apr. 21, 2021). “A Survey on Recent Advanced Research of CPS Security”. In: *Applied Sciences* 11.9, p. 3751. doi: 10.3390/app11093751.

Wang, Zhiwen, Jiqiang Hu, and Hongtao Sun (Dec. 2020). “False Data Injection Attacks in Smart Grid Using Gaussian Mixture Model”. In: *2020 16th International Conference on Control, Automation, Robotics and Vision (ICARCV)*. 2020 16th International Conference on Control, Automation, Robotics and Vision (ICARCV), pp. 830–837. doi: 10.1109/ICARCV50220.2020.9305398.

Wang, Ziqiang and Jie Wang (Nov. 2019). “A Novel Finite-Time Control Scheme for Enhancing Smart Grid Frequency Stability and Resilience”. In: *IEEE Transactions on Smart Grid* 10.6, pp. 6538–6551. doi: 10.1109/TSG.2019.2907144.

Xiong, Xiaoping et al. (Oct. 2020). “Detection of False Data Injection Attack Based on Improved Distortion Index Method”. In: *2020 IEEE 20th International Conference on Communication Technology (ICCT)*. 2020 IEEE 20th International Conference on Communication Technology (ICCT), pp. 1161–1168. doi: 10.1109/ICCT50939.2020.9295794.

Xu, Li Da, Eric L. Xu, and Ling Li (Apr. 18, 2018). “Industry 4.0: State of the Art and Future Trends”. In: *International Journal of Production Research* 56.8, pp. 2941–2962. doi: 10.1080/00207543.2018.1444806.

Yajun Lu and J. Cecil (Sept. 10, 2015). “An Internet of Things (IoT)-Based Collaborative Framework for Advanced Manufacturing”. In: *The International Journal of Advanced Manufacturing Technology*. doi: 10.1007/s00170-015-7772-0.

Yohanandhan, Rajaa Vikhram et al. (2020). “Cyber-Physical Power System (CPPS): A Review on Modeling, Simulation, and Analysis With Cyber Security Applications”. In: *IEEE Access* 8, pp. 151019–151064. doi: 10.1109/ACCESS.2020.3016826.

Yu, Zhiyuan et al. (2021). “Security and Privacy in the Emerging Cyber-Physical World: A Survey”. In: *IEEE Communications Surveys & Tutorials* 23.3, pp. 1879–1919. doi: 10.1109/COMST.2021.3081450.

Yuz, Juan I. and Graham C. Goodwin (2014). *Sampled-Data Models for Linear and Nonlinear Systems*. Communications and Control Engineering. 289 pp. isbn: 978-1-4471-5561-4.

Zhang, Jiazi et al. (Sept. 2018). “Can Attackers With Limited Information Exploit Historical Data to Mount Successful False Data Injection Attacks on Power Systems?” In: *IEEE Transactions on Power Systems* 33.5, pp. 4775–4786. doi: 10.1109/TPWRS.2018.2818746.

Zhao, Junbo, Lamine Mili, and Meng Wang (Sept. 2018). “A Generalized False Data Injection Attacks Against Power System Nonlinear State Estimator and Countermeasures”. In: *IEEE Transactions on Power Systems* 33.5, pp. 4868–4877. doi: 10.1109/TPWRS.2018.2794468.

Zheng, Yan et al. (July 2021). “Vulnerability Assessment of Deep Reinforcement Learning Models for Power System Topology Optimization”. In: *IEEE Transactions on Smart Grid* 12.4, pp. 3613–3623. doi: 10.1109/TSG.2021.3062700.

Amélioration de la sécurité des systèmes cyber-physiques : stratégies de détection et d'atténuation

Les systèmes cyber-physiques (SCP) se situent à l'intersection des opérations physiques, des technologies informatiques et des communications réseau. Ces systèmes sont fondamentaux pour l'infrastructure des environnements de fabrication intelligente, fournissant les moyens d'un contrôle amélioré, d'une optimisation et d'une adaptabilité au sein de divers processus de production. Cette thèse se penche sur les défis associés à la protection des SCP contre une large gamme de menaces cybernétiques qui posent des risques pour la sécurité de ces systèmes. À travers l'application de stratégies de contrôle de pointe telles que l'Observabilité Fonctionnelle et la Théorie des Échelles de Temps, ce travail introduit des méthodes innovantes visant la détection et l'atténuation de l'Injection de Données Fausses (IDF), des Attaques de Dynamique Nulle (ADN), et du Déni de Service (DoS). Ces menaces cybernétiques sont identifiées comme prévalentes et dommageables dans le contexte des SCP industriels. L'incorporation stratégique des *observateurs fonctionnels*, l'application du cadre analytique des Échelles de Temps, et l'utilisation du critère de stabilité de Lyapunov se démarquent comme contributions notables au domaine de la sécurité des SCP. Ces approches offrent de nouvelles perspectives et des solutions robustes aux défis multifacettes que rencontrent les systèmes cyber-physiques. De plus, l'efficacité des solutions proposées est testée et validée à travers une combinaison d'études expérimentales et de simulation. Ces efforts de validation soulignent la capacité des améliorations de sécurité proposées à élever les mesures protectrices au sein des environnements de fabrication intelligente. En conséquence, cette travail de thèse repousse non seulement les limites des connaissances actuelles dans le domaine de la sécurité des SCP, mais établit également une base solide pour les recherches en cours et futures visant à renforcer l'infrastructure numérique essentielle au fonctionnement des industries intelligentes. À travers cette étude, la thèse aspire à contribuer de manière significative au développement de systèmes cyber-physiques plus sûrs, fiables et résilients, soutenant ainsi l'évolution continue des technologies de fabrication intelligente.

Systèmes Cyber-Physiques, Cybersécurité, Injection de Données Fausses, Attaque de Dynamique Nulle, Déni de Service

Enhancing Cyber-Physical System Security: Detection and Mitigation Strategies

Cyber-physical systems (CPS) stand at the intersection of physical operations, computing technologies, and network communications. These sophisticated systems are fundamental to the infrastructure of smart manufacturing environments, providing the means for improved control, optimization, and adaptability within various production processes. This thesis delves into addressing the significant challenges associated with safeguarding CPS from a wide range of cyber threats that pose risks to the operational integrity and overall safety of these systems. Through the application of cutting-edge control strategies like Functional Observability and Time-Scale Theory, this work introduces innovative methods aimed at the early detection and effective mitigation of False Data Injection (FDI), Zero-Dynamics Attacks (ZDA), and Denial of Service (DoS). These particular cyber threats are identified as especially prevalent and damaging within the context of industrial CPS. The strategic incorporation of *functional observers*, the application of the Time-Scale analytical framework, and the use of the Lyapunov stability criterion stand as notable contributions to the domain of CPS security. These approaches provide fresh insights and robust solutions to the multifaceted challenges that cyber-physical systems encounter. Moreover, the efficacy of the proposed solutions is tested and validated through a combination of simulation and hands-on experimental studies. These validation efforts underscore the capability of the proposed security enhancements to elevate the protective measures within smart manufacturing environments. Consequently, this thesis not only pushes the boundaries of current knowledge in the field of CPS security but also establishes a solid foundation for ongoing and future researches aimed at fortifying the digital infrastructure essential to the operation of smart industries. Through this study, the thesis aspires to contribute meaningfully to the development of more secure, reliable, and resilient cyber-physical systems, thereby supporting the continued evolution of smart manufacturing technologies.

Cyber-Physical Systems, Cybersecurity, False Data Injection, Zero-Dynamics Attack, Denial of Service

Spécialité: Automatique et Traitement du Signal

Université de Reims Champagne-Ardenne
CReSTIC - EA 3804
CReSTIC - Bâtiment 12 - URCA