COoperative Cyber prOtectiON for modern power grids

D2.1 Generic Methodology for Power Grid State Estimation

Distribution Level	PU
Responsible Partner	USE
Prepared by	Stelios C. Dimoulias (AUTH) José Maria Maza Ortega (USE) Georgios C. Kryonidis (AUTH) Esther Romero Ramos (USE) Kyriaki-Nefeli D. Malamaki (AUTH) Catalina Gómez Quiles (USE) Charis S. Demoulias (AUTH)
Checked by WP Leader	Charis S. Demoulias (AUTH)
Verified by Reviewer #1	Filip Holik (UGLA) 27/02/2025
Verified by Reviewer #2	Luna Moreno Diaz (ING) 03/03/2025
Approved by Project Coordinator	Angelos Marnerides (UCY) 14/03/2025

Co-funded by
the European Union

Disclaimer

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the Directorate General for Communications Networks, Content and Technology. Neither the European Union nor the Directorate General for Communications Networks, Content and Technology can be held responsible for them.

Deliverable Record

Planned Submission Date	17/03/2025
Actual Submission Date	14/03/2025
Status and version	FINAL (1.0)

Version (Notes)	Date	Author(s)	Notes
0.1 (Draft)	01/09/2024	José Maria Maza-Ortega (USE), Georgios C. Kryonidis (AUTH)	Table of Contents, Initial Structure, and Work allocation
0.2 (Draft)	01/11/2024	Stelios C. Dimoulias (AUTH), José Maria Maza Ortega (USE), Georgios C. Kryonidis (AUTH), Esther Romero Ramos (USE), Kyriaki-Nefeli D. Malamaki (AUTH), Catalina Gómez Quiles (USE), Charis S. Demoulias (AUTH)	Annex with detailed literature review of false data injection methods in EPES
0.3 (Draft)	01/12/2024	Stelios C. Dimoulias (AUTH), José Maria Maza Ortega (USE), Georgios C. Kryonidis (AUTH), Esther Romero Ramos (USE), Kyriaki-Nefeli D. Malamaki (AUTH), Catalina Gómez Quiles (USE), Charis S. Demoulias (AUTH)	Cyber security challenges on EPES
0.3 (Draft)	01/01/2025	Georgios C. Kryonidis (AUTH)	Executive Summary and Introduction
0.4 (Draft)	15/01/2025	Stelios C. Dimoulias (AUTH), José Maria Maza Ortega (USE), Georgios C. Kryonidis (AUTH), Esther Romero Ramos (USE), Kyriaki-Nefeli D. Malamaki (AUTH), Catalina Gómez Quiles (USE), Charis S. Demoulias (AUTH)	Literature review, statistics and proposed taxonomy
0.5 (Draft)	15/02/2025	Stelios C. Dimoulias (AUTH), José Maria Maza Ortega (USE), Georgios C. Kryonidis (AUTH), Esther Romero Ramos (USE), Kyriaki-Nefeli D. Malamaki (AUTH), Catalina Gómez Quiles (USE), Charis S. Demoulias (AUTH)	Generic FDII methodology
0.6 (Draft)	05/03/2025	Stelios C. Dimoulias (AUTH), José Maria Maza Ortega (USE), Georgios C. Kryonidis (AUTH)	Revision based on the comments received from Reviewers
1.0 (Final)	14/03/2025	Georgios C. Kryonidis (AUTH)	Revision based on the comments received from Project Coordinator

Contents

List of Figures

List of Tables

Definition of Acronyms

AD	Anomaly Diagnosis
AI	Artificial Intelligence
AS	Ancillary Services
AV	Attack Vector
CIRES	Converter-Interfaced Renewable Energy Source
COMML	Control Measurement and Monitoring Layer
CPN	COCOON Programmable Node
CSL	Cyber security Services Layer
DMS	Distribution Management System
D-FACTS	Distributed Flexible AC Transmission systems
DoS	Denial of Service
DRES	Distributed Renewable Energy Source
DS	Distribution System
DSO	Distribution System Operator
EMS	Energy Management System
ENTSO-E	European Network of Transmission System Operators for Electricity
EPES	Electric and Power Energy System
FDD	False Data Detection
FDI	False Data Injection
FDIA	False Data Injection Attack
FDII	False Data Injection Identification
FDIR	False Data Injection Response
FDL	False Data Localization
ICT	Information and Communication Technology
IED	Intelligent Electronic Devices
IOL	Instrumentation and Orchestration Layer
IT	Information Technology
LV	Low-Voltage
MbPA	Measurement-based Plausibility Analysis
MitM	Man in the Middle
MTD	Moving Target Defense
MV	Medium-Voltage
μNF	Micro Network Function
OT	Operational Technology
PI	Physics-Informed
PMU	Phasor Measurement Unit
POI	Point Of Interconnection
PPC	Power Plant Controller
PV	Photovoltaic
RES	Renewable Energy Source
RTU	Remote Terminal Unit
SCADA	Supervision Control and Data Acquisition
SE	State Estimation
SM	Smart Meter
TS	Transmission System

TSO	Transmission System Operator
WP	Work Package

Executive Summary

The decarbonization of the energy sector is one of the primary objectives set by the European Union toward a sustainable and resilient future. The cornerstone of this transition is the proliferation and efficient integration of renewable energy sources (RESs) in the electric power and energy system (EPES). As a result, conventional fossil-fuel-driven power plants are gradually being replaced by environmentally friendly generation units, such as photovoltaics parks, wind farms, etc., a significant share of which is connected in distribution systems (DSs) and is collectively referred to as distributed renewable energy sources (DRESs).

This ever-increasing RES penetration has brought to the surface a series of technical challenges that affect the secure and reliable operation of EPESs. These challenges can be classified into two main categories based on their range of impact on EPESs. The first category includes local phenomena, such as voltage violations and equipment overloading, affecting a small region of the EPES. The second category involves system-wide problems, such as frequency stability issues that affect all components connected to the EPES. As a result, the EPES is severely stressed and new solutions are required to be developed in order to guarantee the secure, reliable and cost-effective operation of the EPES.

Fortunately, the advent of digital technologies within the EPESs has brougth new possibilities to address these issues by enhancing their monitoring and control ability, thereby transforming EPESs into more complex cyber-physical systems which strongly rely on information and communication technology (ICT) infrastructures. Nevertheless, this new era is, in turn, characterized by new emerging challenges related to the cyber security of EPESs. Provided that EPESs are the backbone of modern civilization, ensuring social welfare and productive development, their protection is of utmost importance for the European Union.

In this context, the scope of the COoperative Cyber prOtectiOn for modern power grids (COCOON) is to provide a holistic and interdisciplinary solution for the cyber protection of EPESs. This is attained by supplementing state-of-the-art IT solutions with additional insights obtained from the inherent characteristics of the physical system. The cornerstone of the proposed solution is the introduction of the COCOON programmable node (CPN), a programmable networking device designed to run customized network functions, serving as the foundation for the implementation of high-level solutions, such as anomaly diagnosis (AD), false data injection identification (FDII), among others.

This deliverable, D2.1: *Generic methodology for power grid state estimation*, presents the main architecture of FDII, taking into account the physical properties of the EPES. The proposed solution is implemented at the highest abstraction layer of the CPN configuration, i.e., the cyber security services layer (CSL). The structure of the present deliverable unfolds as follows. Initially, the main challenges arising in the EPES from a cyber security perspective are analyzed and discussed. Afterward, an exhaustive literature review is presented dealing with the FDII methods applied in EPES that are further categorized using a newly proposed taxonomy. Finally, the outcomes of this analysis are used for determining the configuration of the proposed FDII method that will be adopted in frame of COCOON.

1 Introduction

False data injection (FDI) is regarded as one of the main cyber security issues that both transmission (TSOs) and distribution system operators (DSOs) shall carefully address to ensure the stable and reliable operation of the under high RES penetration. To this end, the COCOON project aims to provide a comprehensive, effective and pragmatic solution for identifying FDI attacks (FDIAs) in EPESs. This is achieved by adopting a bottom-up, system-oriented CPN configuration, consisting of three abstraction layers. The two lower layers, namely the control measurement and monitoring layer (COMML) and the instrumentation and orchestration layer (IOL), are used for translating the requirements from the top layer into micro network functions (μNFs) and executing them on a programmable data plane to perform ICT traffic operations. The top layer, namely the CSL, hosts the high-level security services that will be developed in the frame of COCOON, such as AD and FDII.

1.1 Scope of the Deliverable

This deliverable initially focuses on performing a systematic literature review of the state-of-the-art methods dealing with FDII in EPESs. In addition, a multi-facet classification scheme is proposed to categorize the FDII methods assuming various aspects, e.g., attack motivation, EPES model, type of attack, etc. This analysis paves the way for identifying the critical characteristics that will be integrated into the FDII method developed in the framework of COCOON. Special emphasis is placed on the practicality and implementation of the examined solutions under real-field conditions, as the developed FDII method will be tested in real pilot setups—specifically, the Energy Community and the photovoltaic (PV) plant—described in work packages (WP) 5 and 8, respectively.

Afterward, on the basis of the above outcomes, the outline of a generic methodology for FDII in EPESs is determined and presented. Its distinct characteristic is the consideration of the physical laws of the EPES in the cyber security analysis, which provides an additional verification layer when examining the integrity of the EPES data, thus further increasing the overall cyber security level. The core of the FDII methodology is the state estimation (SE), an optimization routine designed to estimate the operating states of the EPES, e.g., nodal voltage magnitudes and angles, using measurements acquired from various grid locations through the ICT infrastructure and leveraging the physical laws governing the EPES. FDII is supplemented with three auxiliary modules, namely measurement-based plausibility analysis (MbPA), false data detection (FDD), and FDI response (FDIR). The first module is a pre-processing tool acting on a fast timescale and aims to examine the integrity of the acquired EPES data by performing simple checks and/or more sophisticated analyses without considering the physical system of EPES. The second module is a post-processing tool that identifies false data using the outcomes of the SE as inputs. Finally, the scope of the third module is to replace the identified false data, if any, with accurate estimates of the operating states of EPES.

1.2 Relation with other Work Packages and Tasks

This is the first deliverable of WP2: *Cyber security assessment considering power system characteristics* and reflects the work carried out in Task T2.1: *Methodology of cyber security assessment considering the physical system*. The relationship of this deliverable with other tasks, deliverables, and WPs is depicted in Figure 1.1.

Specifically, this deliverable is interdependent with deliverables D1.1: *Control, Measurement and Monitoring Properties* and D1.2: *Threat Models, Vulnerability Assessment and Risk Profiling*. These deliverables are products of the work carried out within tasks T1.1-T1.3 of WP1:*Cyber Security Algorithms*. D1.1 and D1.2 are relevant to D2.1 because they determine the requirements of the lowest abstraction layer of the CPN architecture (COMML) and identify attack vectors for FDIAs, respectively.

In addition, this deliverable utilizes the outputs of deliverables D4.1: *COCOON Development Blueprint* and D4.2: *COCOON System Architecture*, which are part of the work carried out within tasks T4.1-T4.3 of WP4: *Software Prototype Development*. D4.1 and D4.2 set the foundations for the development of the CPN system architecture, and especially of the COMML and the IOL. In turn, this architecture will be leveraged by the upper layer, the CSL, where the FDII service will be established.

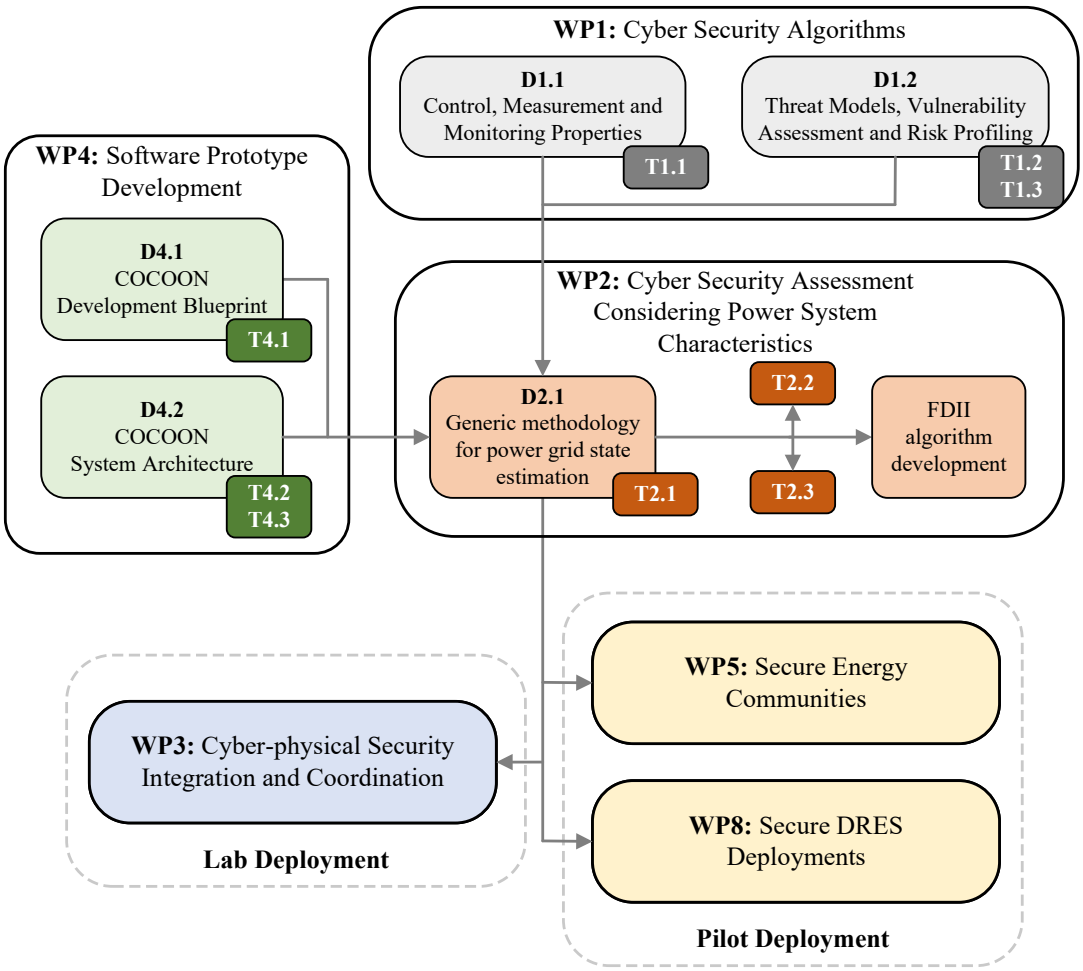

Figure 1.1: The relationship of D2.1 with other tasks, deliverables and WPs.

The output of this deliverable will be utilized in tasks T2.2: *Cyber security Assessment in PV Power Plants* and T2.3: *Cyber security Assessment in Energy Communities* to further develop the proposed generic FDII algorithm and tailor it for application to PV plants and energy communities, respectively. Finally, it will be utilized for testing the developed FDII methodology in lab and real environments, as described in WP3: *Cyber-physical Security Integration and Coordination* and WP5: *Secure Energy Communities*, and WP8: *Secure DRES Deployments*, respectively.

1.3 Outline of Deliverable

The remaining of this deliverable is organized into four Chapters and one Annex. Initially, Chapter 2 provides an overview of the cyber security challenges in EPES, emphasizing FDI. Afteward, Chapter 3 deals with the literature review of state-of-the-art methods dealing with FDIA and/or FDII. A detailed literature review is presented in Annex A summarizing the main contributions of each work, along with key insights. Moreover, a new taxonomy is presented classifying these works under a broad set of categories. In Chapter 4, the generic FDII methodology that will be adopted in the frame of COCOON is presented. Finally, Chapter 5 concludes this deliverable listing the main findings and outlining a plan for the next steps.

2 Cyber Security Threats in the EPES

2.1 Introduction

This chapter focuses on the emerging cyber security issues in modern EPESs and the countermeasures proposed in the framework of WP2. For this purpose, Section 2.2 provides a brief overview of the main components of traditional EPES composed of generation, transmission, distribution, and end users. Despite that this architecture has been maintained during the last century, the section discusses the impact that decarbonization has into the EPES, specially in the distribution and end-user levels, which makes its operation much more complex. Fortunately, the advent of digital technologies may facilitate this strained situation. In this regard, Section 2.3 states that, ultimately, the EPES can be considered a cyber-physical system composed of several control centers implemented in generation, transmission, distribution and even end-user facilities. These control centers are equipped with decision-making tools which rely on field measurements to compute proper setpoints for controllable EPES devices. The EPES requires not only conventional power assets, such as generators, lines, transformers, etc., but also more and more information and communication technologies, therefore, it can be envisioned as a cyber-physical system. This leads to the next section (Section 2.4), where it is outlined that modern EPESs are exposed to cyber attacks which may have a tremendous impact considering the electrification of the society for the sake of replacing fossil-fuel technologies. This section provides relevant information about previous cyber attacks and proposed taxonomies to characterize them. The analysis of this initial information will reveal that among denial-of service attacks, malware attacks, and many others, FDIAs can be launched from several EPES layers which makes them worthy of further analysis. In this regard, the importance of FDI attacks in the EPES and the possible categories in which this type of attack can occur, depending on whether it affects measures or setpoints, are discussed in more detail in Section 2.5. Finally, the chapter concludes with an overview of the role that state estimation techniques may have in the detection of FDI attacks by incorporating the information of the physical performance of the EPES (Section 2.6)

2.2 EPES: Current Landscape and Challenges

EPESs are fundamentally structured into four key stages: generation, transmission, distribution, and consumption by end users [1, 2]. Traditionally, the power flow in EPES was unidirectional, i.e., from bulk generation units to consumption through transmission (TSs) and DSs. In this context, the first stage involves the centralized production of electricity at large-scale power plants, mainly relying on hydro- or fuel-based primary energy sources, such as coal, natural gas, etc. These power plants are connected to EPES through synchronous generators equipped with step-up transformers. In the second stage, TSs carry electricity at high-voltage levels through an extensive network of overhead/underground/submarine lines. The high-voltage levels allow for the transmission of vast amounts of electric power over long distances, from generation centers to regional substations, with low currents and thus, under reduced thermal losses. At regional substations, step-down transformers interconnect the bulk TS with DSs, which constitute the third stage, delivering electricity through power lines and transformers directly to end users, such as residential homes, businesses, and industries (fourth stage). This EPES structure ensures that electricity is efficiently and safely delivered from the primary sources to the final consumers/end users. Although modern EPESs still comprise these four fundamental stages, the functionalities and roles of each stage have evolved, as will be analyzed in the remainder of Section 2.2.

Over the past three decades, environmental and economic factors have prompted a shift towards RESs, and particularly converter-interfaced RESs (CIRESs), such as solar PV and wind energy. Specifically, climate change and air pollution are the primary reasons for reducing greenhouse gas emissions associated with fossil fuels. In addition, rapid advancements in new materials for semiconductors, as well as improvements in the hardware and control of power electronics converters, have rendered the use of CIRESs more efficient and cost-effective. Based on the above, governments and national/international organizations have established policies and incentives for the promotion of green energy, thus leading to RES proliferation [3]. Note that RESs are installed not only at the TS level but also at the DS level as DRESs, since RES technologies are less dependent on economies

of scale, compared to traditional power plants. Therefore, DRESs enable smaller-scale energy production closer to the points of consumption, reducing transmission losses and increasing energy independence of consumers.

The main scope of TSs is to transfer electricity over large distances while maintaining high security and reliability standards. This is attained by adopting a meshed configuration, where each network node can be supplied through multiple paths [4]. As a result, in case a TS component, e.g., a line, experiences a fault or requires maintenance, the electricity supply remains unaffected, thereby avoiding the risk of widespread outages. TSOs are responsible for the the secure and reliable TS operation by applying monitoring and control procedures. The main aim is to manage the complex electricity flows within the meshed TS and balance supply and demand in real time resorting to ancillary services (ASs) that were traditionally provided by large-scale generation units. Additionally, TSOs are responsible for the procurement of these ASs that are provided both during the TS scheduling, as well as during the real-time operation [5, 6]. Nevertheless, the ever-increasing CIRES penetration along with their highly variable generation pose challenges to TSOs in maintaining the supply-demand balance and ensuring the secure EPES operation. Furthermore, the lack of inertia in CIRES can result in severe frequency deviations during sudden load or generation changes, increasing the risk of blackouts.

Until recently, DSs were operated in a passive way by the DSOs, with the aim of delivering electricity from TSs to end users. Under those conditions, the DS operations were straightforward and with minimal real-time monitoring or control requirements, justifying the limited interest shown by utilities and relevant stakeholders in their operation. This lack of visibility of the DSs [7] was further reflected in the absence of a stringent regulatory framework for their operation, as opposed to TSs. In simple terms, DSs were designed and operated according to a *fit and forget* approach. Nonetheless, DRES proliferation has lead to a series of technical challenges, such as: bidirectional power flows, voltage and thermal violations, malfunction of protection systems, power quality issues, etc. [8, 9, 10]. To overcome these issues, the concept of smart grid has been introduced incorporating advanced communication, automation, and information technologies to enhance the controllability and dynamic operation of DSs. This transformation necessitates the real-time monitoring and control of DSs, requiring substantial upgrades to the existing DS infrastructure, such as the deployment of sensors, smart meters, and advanced control systems. As a result, this digitalization leads to cyber security concerns due to the increased connectivity and data exchange within active DSs.

End users have been historically passive consumers when [11]. Their consumption patterns were largely unaffected by external factors such as real-time pricing or grid demands, without having the capability of producing their own energy. This situation has radically changed due to the advent of DRES, residential energy storage systems, and electric vehicles, introducing a more active role for consumers [12]. Furthermore, recent advances in smart grid technologies and the internet of things have equipped consumers with tools to monitor and manage their energy usage in real time [13]. They can now feed excess power back into the grid, and adjust their consumption based on external signals. Hence, this interaction has transformed passive end users into active participants, evolving them into *prosumers* (i.e., both consumers and producers), who are able to determine their corresponding generation and load profiles.

This active participation not only empowers end users with greater control over their energy expenses but can also support the RES integration towards a sustainable EPES. However, this shift further increases the complexity and importance of DSs [13]. In addition, end users are increasingly exposed to cyber security risks due to the proliferation of smart meters and local control devices that can be vulnerable to hacking. Unauthorized access to these systems can lead to privacy breaches, manipulation of energy usage data, and even disruptions in power supply, highlighting the need for robust cyber security measures at the consumer level [13].

2.3 The EPES as a Cyber-Physical System

As outlined in Section 2.2, EPESs need to be monitored and controlled in real time to guarantee a secure, reliable, and cost-effective power and energy supply to end users. The requirements and objectives of this real-time control vary with respect to the system level they address, e.g., generation, transmission, or distribution. However, irrespective of the application stage, control centers with human operators, assisted by decision-making tools, are deployed along the EPES to dispatch control signals and perform appropriate control actions. In this context, supervisory control and data acquisition (SCADA) systems are one of the main pillars of any

control center, as they are responsible for gathering information from field sensors and sending proper setpoints (i.e., commands for operation) to controllable assets.

Generation companies establish control centers to monitor and optimize the operation of their assets within the framework of liberalized power systems and energy markets. Ultimately, generation companies aim at maximizing their profits for a given portfolio of assets, which generally includes controllable thermal and hydro power plants, as well as non-dispatchable RES. The generation management system provides the required tools for bidding in the existing energy markets, monitoring the actual generation, forecasting the RES generation, and performing automatic generation control for the dispatchable generation, in coordination with the corresponding TSO. In the case of conventional generation units, the system sends operating active and reactive power setpoints to the load following control and the automatic voltage regulation mechanisms of synchronous generators. In the case of RES, the setpoints are sent to the corresponding power plant controller (PPC), which is in charge of managing several individual inverter-based generation units as a single device in the point of interconnection (POI) with EPES.

From a system-wide perspective, the TS serves as the backbone of the EPES, where a major problem could affect a large number of customers. To prevent this, TSOs are in charge of voltage and frequency regulation. On the one hand, voltage control requires the coordinated control of different assets such as synchronous generators, transformer on-load tap changers, capacitor banks, CIRESs and power electronics devices, such as static VAR compensators or static synchronous compensators. With the exception of the on-load tap changer, which directly adjusts the voltage, these devices regulate the reactive power injections at their respective connection nodes in order to control the TS voltage. Reactive power control is an effective voltage regulation method thanks to the high X/R ratio of the TS lines. Voltage regulation is coordinated across different time scales, ranging from primary to tertiary control levels, in order to prevent interactions between controllers and geographical areas, given the inherently local nature of the process.

On the other hand, frequency control is intended to achieve a perfect balance of the uncontrolled load demand and generated power. Unlike voltage control, frequency regulation is a global problem that requires the appropriate control actions by the controllable generators, i.e., synchronous generators and/or CIRESs, while also assuring adequate reserves to respond to any unexpected demand deviations. Similarly to voltage regulation, the problem is solved across different time scales to efficient exploit the available reserves.

These voltage and frequency control schemes are embedded within the so-called energy management system of the TSO control center, alongside other decision-making tools for supporting the operator actions.

As discussed above, traditional DSs are not characterized by the same degree of control and monitoring as TSs. DSs comprise medium- (MV) and low-voltage (LV) networks departing from primary and secondary substations, respectively. In this regard, traditional passive DSs typically incorporate real-time monitoring only at the head of MV feeders, while the rest of the network is not observable at all by the DSO. Note that passive DSs can be operated in this way due to their inherent simplicity, i.e., unidirectional power flows from the primary substations to secondary substations and, ultimately, to end users. The usual controllable assets are limited to transformer on-load tap changers and capacitor banks at primary substations for centralized voltage control. Nevertheless, modern active DSs experience increasing needs for real-time monitoring and control, as necessitated by the proliferation of DERs and as dictated by the smart grid paradigm. In this context, the controllability and observability of DSs is expected to be extended to LV networks as well. Therefore, this ongoing revolution in the field of active DSs requires continuous advancements in monitoring technologies and decision-support tools for operators, all incorporated within a tailored distribution management system (DMS).

The devices employed across the EPES vary with respect to the intended functionality and the application, with the following categories being the most notable:

- **Remote terminal units (RTUs):** They are used to acquire, process and transmit electrical measurements from components connected either at the TS or the DS.
- **Intelligent electronic devices (IEDs):** They are intended for automation and protection purposes in EPES substations.

- **Phasor measurement units (PMUs):** They estimate the magnitude and phase angle of voltage phasors and ensure synchronicity among distributed measurements.
- **Smart meters (SMs):** They are installed at the end-user side to provide the energy consumption reading for billing purposes.

The communication capabilities of these devices are determined by the given ICT infrastructure, consisting of physical layers (e.g. fiber optics, ethernet cables, wireless communication, power line carriers, etc.) and protocols (IEC-101, IEC-104, IEC-61850, Modbus-TCP/IP, ModBus-RTU, DNP3, DLMS/COSEM, etc.), which are adapted to the specific characteristic of the data flows.

Based on the above analysis, it is concluded that the modern EPES can be envisioned as a large, complex, and interacting cyber-physical system composed of operational (OT) and information technologies (IT).

The OT, physical part is formed by the traditional power assets: generators, transformers, lines, switchgear, capacitors, reactors, etc. The cyber, IT part consists of the modern ICT devices used to effectively control the EPES: RTUs, IEDs, PMUs, SMs, control centers, SCADA systems, decision-making software tools, etc. It is worth emphasizing that the cyber side is just as important as the physical side. More specifically, as discussed above, operations and actions critical for the reliable, secure, and cost-effective energy supply rely on field data extraction, communication infrastructure, and advanced control schemes. Furthermore, the continuous advancement of the cyber-related technologies, as well as the stringent and highly-regulated operational framework of modern EPESs further highlight the importance of the cyber part.

In summary, the highly digitalized and cyber-dependent paradigm of modern EPESs introduces a dual reality. On the one hand, these new conditions allow for the better operation of the EPES and benefit all stakeholders involved. On the other hand, modern EPESs are increasingly dependent on these cyber technologies, thereby becoming vulnerable to their failure, malfunction, or cyber attacks [14]. For this reason, cyber security is of utmost importance in this context to prevent economic losses, physical damage to EPES installations, service disruptions, and generally ensure the social welfare through the high-quality production and supply of electricity.

2.4 Overview of Cyber-Attacks against EPESs

A cyber attack is a deliberate and malicious attempt by an individual, group, or organization to breach the security of another entity's digital systems, networks, or devices. The primary goals of cyber attacks are often to steal sensitive data, disrupt operations, damage digital infrastructure, or gain unauthorized access to systems for further exploitative activities [15, 16]. These attacks can take various forms, such as deploying malware (such as viruses or ransomware), conducting phishing schemes to trick users into revealing confidential information, executing denial-of-service (DoS) attacks to overwhelm and disable services, or exploiting vulnerabilities in software and hardware to infiltrate systems. Cyber attacks pose significant risks to individuals, businesses, and governments by compromising the confidentiality, integrity, and availability of digital information and services. For a comprehensive listing of all significant cyber incidents from May 2006 to December 2024, the reader is referred to the recent report [17] of the Center for Strategic and International Studies.

Critical infrastructure sectors such as power, gas, and water utilities are prime targets for malicious cyber activities, with the energy sector reporting the highest number of vulnerabilities among all network infrastructures, as energy utilities increasingly adopt digital technologies to manage power plants, grids, and business operations [14]. Specifically, reports from the US ICS-CERT [18] and Kaspersky ICS-CERT [19] indicate that the energy sector experienced 178 (2017), 110 (2018), and 283 (2019) cyber attack incidents, outpacing other industrial control systems [20].

While the digital transformation and smart grids enhance energy security by improving supply quality, customer services, and facilitate RES integration, they also introduce new vulnerabilities [14, 20]. Each digital system, telecommunication device, and sensor becomes a potential entry point for cyber criminals, expanding the utilities' exposure to attacks. Cyber actors exploiting weaknesses in smart grid devices pose significant threats to the reliability and performance of power grids.

Publicly available information on significant cyber security incidents is limited due to under-reporting and lack

of detection [14]. However, evidence shows a rapid increase in cyber attacks on utilities since 2018, reaching critical levels in 2022, partly influenced by geopolitical events such as Russia's invasion of Ukraine. Recent incidents have disabled remote controls for wind farms, disrupted prepaid meter systems, and led to data leakage involving sensitive client information, with the average cost of a data breach in the energy sector reaching a record USD 4.72 million in 2022 [14]. These cyber incidents can lead to severe security risks, including the loss or malicious alteration of critical data necessary for control operations. Potential consequences range from incorrect customer billing [21] and price manipulation in energy markets [22, 23, 24] to large-scale power outages [25, 26] and disruptions to the social and economic order [27, 28].

Over the past 14 years, there have been several major attacks targeting the EPES and its components. Figure 2.1 illustrates the timeline of the major global cyber incidents against the power sector, as reported in [20], along with other, more recent attacks, including the following:

1. SolarWinds Supply Chain Attack (December 2020) [29]: A sophisticated cyber attack compromised the SolarWinds Orion software, affecting numerous U.S. government agencies and private companies. The breach allowed attackers to insert malicious code into software updates, leading to widespread data exposure.

2. Colonial Pipeline Ransomware Attack (May 2021) [30]: A ransomware attack targeted Colonial Pipeline, a major US fuel pipeline operator, causing a temporary shutdown of operations. This incident disrupted fuel supply along the East Coast and highlighted vulnerabilities in critical infrastructure.

Due to the diverse nature and motives of the attacks conducted so far, as well as of those theoretically described in the scientific literature, their systematic analysis through a taxonomy classification of entities and concepts is crucial. Cyber attacks can be categorized based on their techniques, targets, and the layers of systems they impact. In this regard, a comprehensive taxonomy was introduced by [31] and is visualized in Figure 2.3.

As observed, the analysis of [31] classifies cyber-physical attacks into four broader delivery categories, with respect to the layer of execution: cyber-based, physical-based, network-based, and communication-based attacks. Among these four broader categories, the only type of attack that is possible to be instantiated across all layers of the cyber-physical EPES is the FDIA. This pervasive exposure of EPESs to the threat of FDIAs explains why the COCOON project addresses this type of cyber attacks within the frame of WP2, as it is further discussed in Section 2.5. Apart from FDIAs, other important types of cyber attacks include:

- **Denial-of-Service (DoS)** attacks [32, 33], in which perpetrators aim to render a computer or network temporarily unavailable to its intended users, most commonly by overwhelming the targeted system with excessive data traffic. In the context of EPESs, DoS attacks could result in data packet dropouts, affecting the monitoring and control functionalities of smart grids. DoS attacks are typically instantiated as flooding, resource exhaustion, or reflection-amplification attacks.
- **Time delay** attacks [34, 35], in which cyber actors introduce delays in the communication channels of a smart grid with the aim of sabotaging several critical operations, e.g., timely opening of breakers.
- **Malware** attacks [36], in which harmful software such as viruses or ransomware is injected. The aim of such attacks is multifaceted, including gaining unauthorized access to operations and stealing confidential information. Notable malware software includes the Stuxnet computer worm (2010 attack in Iran) and the BlackEnergy trojan (2015 attack in Ukraine) [37, 38].

Moreover, the latest attempt for a taxonomy of cyber attacks and cyber vulnerabilities across the generation, transmission, and distribution sectors of EPESs has been presented by [39] and is illustrated in Figure 2.2. For additional insights regarding the taxonomy of cyber attacks against EPESs the reader is referred to the following reviews and surveys [25, 40, 28, 41, 42, 26, 39].

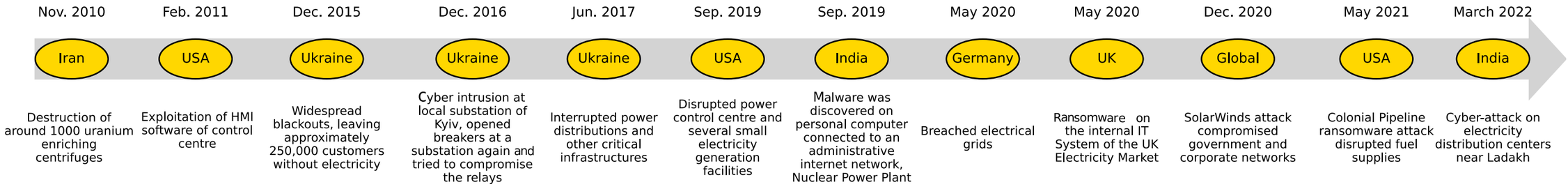

Figure 2.1: Timeline of recent major global cyber attacks on EPES based on [20] and [39], enriched with more recent events. [29, 30]

.

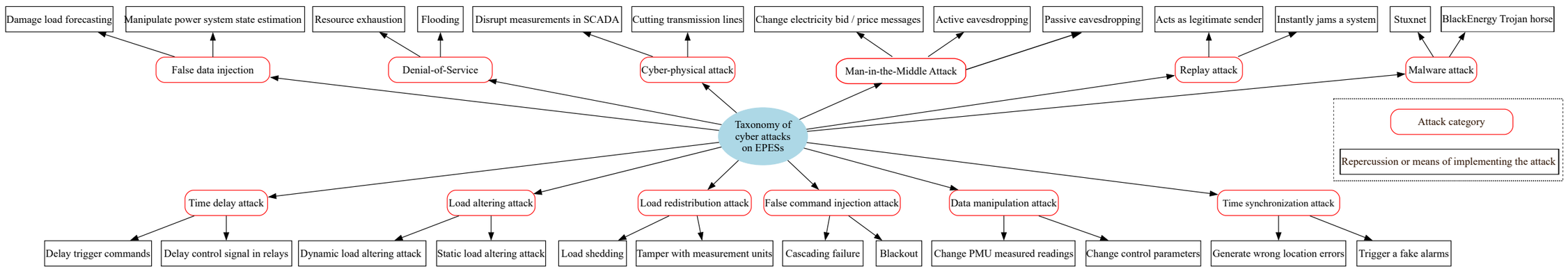

Figure 2.2: Taxonomy based on [39].

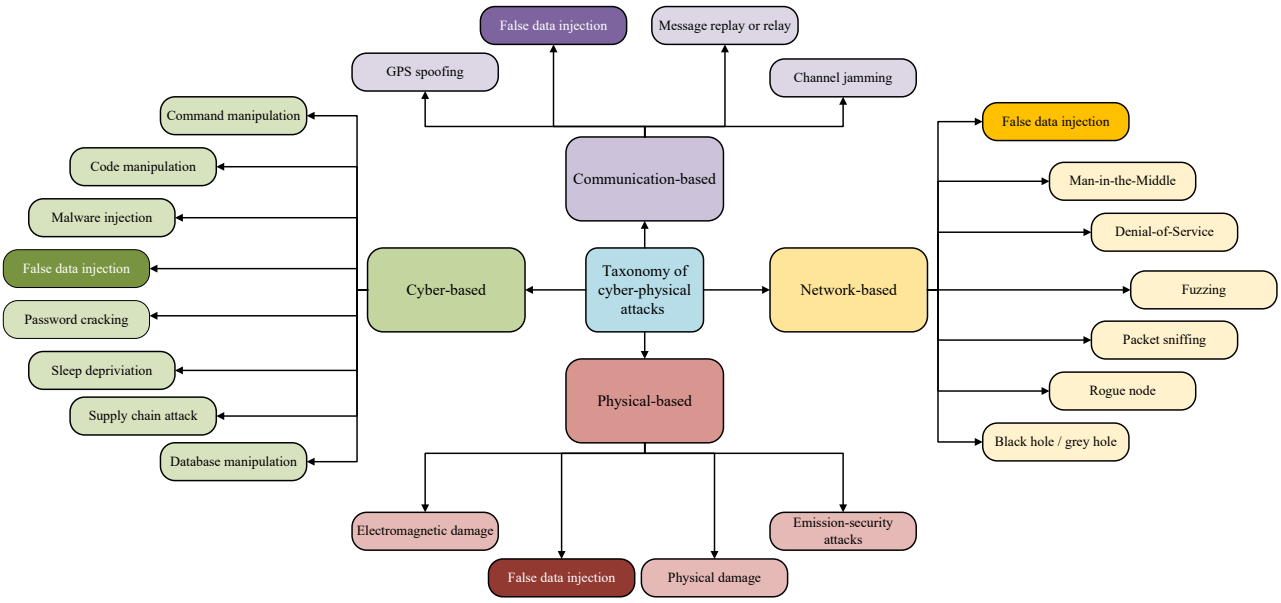

Figure 2.3: Taxonomy based on [31].

2.5 FDIAs against EPESs

2.5.1 Significance

As mentioned in Chapter 1, WP2 of the COCOON project, titled *Cyber Security Assessment Considering Power System Characteristics*, focuses on fortifying the EPES against FDIAs, also known as data integrity attacks. This particular attention and interest in FDIAs emanate from their following properties:

- ***Pervasiveness:*** FDIAs are a pervasive cyber threat, i.e., they can be launched through all four basic layers of the cyber-physical EPES, as discussed in Section 2.4.
- ***Stealthiness:*** FDIAs are an inherently stealthy type of attack, since they can indirectly damage the EPES without being detected. To elaborate, in the absence of an FDII tool, as is currently the case with most DSs and RES parks, even basic "dummy" FDIAs can remain undetected and cause damage to the system, while sophisticated FDIAs can successfully bypass advanced FDII tools.

The properties of *pervasiveness* and *stealthiness* shape the philosophy adopted in WP2, and especially in tasks T2.1-T2.3. More specifically, due to the *pervasiveness* of FDIAs and the various vulnerabilities that potential intruders can exploit to launch FDIAs, the scope of WP2 and deliverable D2.1 is not to prevent the implementation of FDIAs but rather to successfully identify them, thereby introducing an additional and final level of security. In order for this detection to be successful, the *stealthiness* of FDIAs needs to be counteracted and negated, which can be achieved by incorporating the physical characteristics of the EPES, as well as historical and operational patterns, into the detection strategy.

2.5.2 Overview

As discussed in Section 2.4, FDIAs can be implemented through all four basic layers of the cyber-physical EPES, as described below:

- **Physical layer:** FDIAs can be implemented by means of physical tampering of the meters, e.g., through electromagnetic interference or hardware modification.
- **Communication layer:** In this type of FDIAs, attackers target the vulnerabilities in the communication protocols of the EPES to transmit false data [20, 43]. Communication-based FDIAs are also known as spoofing attacks [44]. A special case of spoofing attack, which regards an indirect FDIA, is time synchronization attacks. In time synchronization attacks, intruders compromise the time stamps of measurements, achieving an effect equivalent to falsifying measurement values directly [45, 46].
- **Network layer:** FDIAs through the network layer (or data layer) are possible when the adversaries manipulate data packets within the network. In this case, FDIAs are instantiated as man-in-the-middle (MitM) attacks

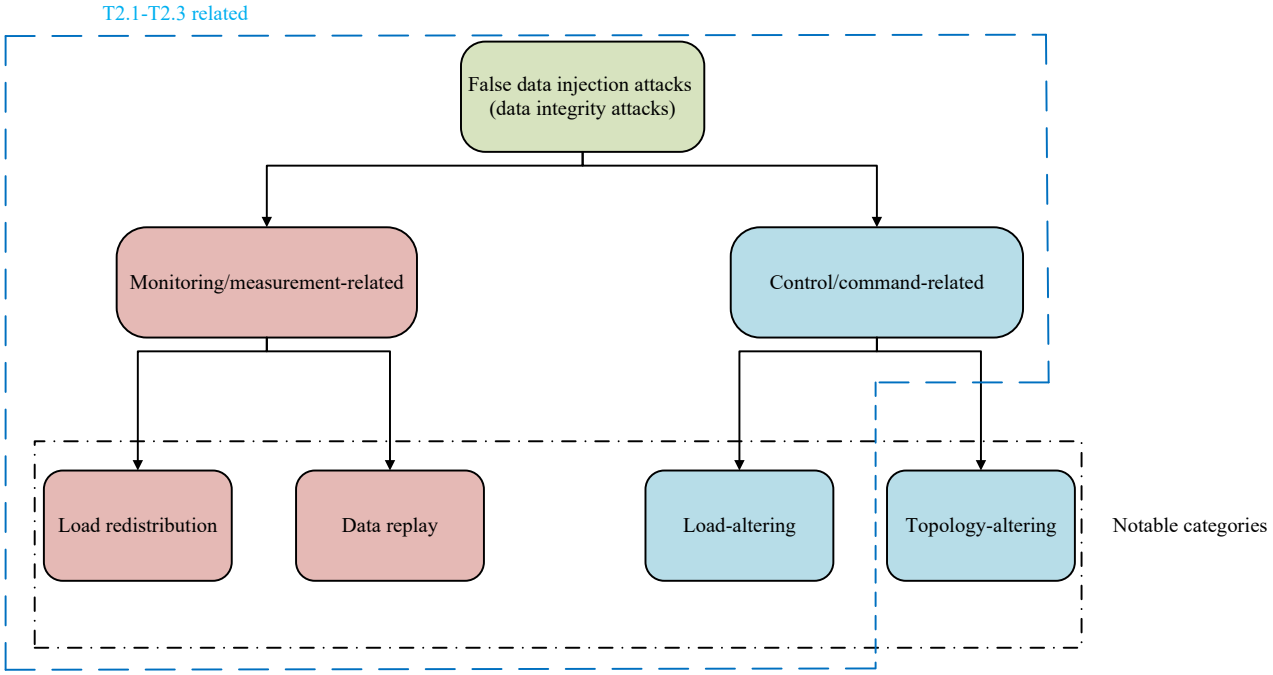

Figure 2.4: Categories of FDIAs.

[47], i.e., attacks where the intruders position themselves within a data exchange between two parties to either eavesdrop or impersonate one of the devices. In MitM FDIAs, the intruders intercept and falsify data, while the software, firmware, and/or physical communication link remain unaffected. Therefore, the two endpoints are unaware of the intrusion.

- **Cyber layer**: FDIAs can be launched in the cyber layer through software manipulation of SCADA systems and metering devices. Moreover, cyber-based attacks could target the firmware of metering devices, such as RTUs [48].

In turn, FDIAs can be further divided into monitoring/measurement-related and control/command-related attacks [49]. In the former, the falsified data correspond to measurements transmitted from the RTUs to the premises of the system operator. In the latter, the falsified data involve commands or operational setpoints that are sent from the system operator to the components and devices of the grid. This classification is illustrated in Figure 2.4, where notable subcategories of FDIAs are also highlighted. More specifically:

- **Monitoring/Measurement-related attacks**:
  - **Load redistribution attacks**: First introduced in [50], they are a specific type of measurement-related FDIA, where the power consumption measurements of system loads and/or generation measurements are falsified while their total value remains unaltered [51].
  - **Data replay attacks**: In this case, attackers re-transmit historical measurement data, corresponding to previous time instants, in place of the actual, concurrent measurement values.
- **Control/Command-related attacks**:
  - **Load-altering attacks**: This is a special case of attacks that manipulate load/generation operational setpoints, where cyber actors modify remotely the power demand of system loads. Most commonly, the ultimate aim is to disrupt the automatic generation control of the generation units within the EPES [52, 53].
  - **Topology-altering attacks**: The target of such attacks is the status of switches and breakers, which can be remotely controlled to change the topology of an EPES [54].

As depicted in Figure 2.4, the methodologies developed within the frame of tasks T2.1-T2.3 of the COCOON project will focus on all categories of FDIAs, except for the topology-altering attacks. The latter, being closely associated with the concept of digital substations, pertains to task T2.4 of the COCOON project.

2.6 State Estimation as a Tool for Detecting FDIAs

2.6.1 Motivation and applications

As mentioned in Section 1.1, WP2 focuses on exploiting the inherent physical characteristics of the EPES through model-based methodologies. By developing accurate models of physical systems, such as PV power plants or energy communities, it is possible to gain insight into the expected behavior of these systems under normal operating conditions. When real-time measurements from the field are compared against these models, inconsistencies can be detected. In this way, this model-based approach allows for the detection of discrepancies between the expected and actual data, providing a means to verify the authenticity and consistency of measurements and setpoints. Therefore, methods that leverage the inherent physical properties, behaviors, and operating patterns of the EPES can be used not only to handle inherent measurement uncertainties but also to identify and respond to FDIAs.

In this context, the most well-established model-based method for such purposes is SE. In the field of EPESs, SE was firstly introduced by Fred Schweppe in 1968, as a procedure to transform redundant, raw, real-time measurements, along with other information, into an estimation of the internal operating states of the EPES. The real-time measurements are collected through SCADA systems and typically include active and reactive power flows through branches, active and reactive power injections at nodes, bus voltages, branch currents, and sometimes even phase measurements. This set of real-time measurements is incomplete, inconsistent, and not exact, since every measurement has an associated error that depends on several factors, e.g., error of the measurement device, precision class of voltage/current transformers, noise in the communication channel, etc. Furthermore, the measurement set is redundant, meaning that the number of measurements exceeds the minimum required to define the EPES state. State estimators take advantage of this property of redundancy to compute the most accurate estimate of the EPES state.

Gradually, SE evolved into a cornerstone of control centers in modern EPES, serving as a foundational layer for various decision-making applications that support the real-time operation of the power system. In simple terms, SE increases the situational awareness regarding the EPES operation, allowing for more effective and precise control actions. Notably without application of SE and its filtering action of raw measurements, a high risk of erroneous control decisions arises, potentially leading to economic inefficiencies or even compromising the security of the EPES.

Nowadays, state estimators are not only employed in TSs, as was traditionally the case, but also in DSs, as the latter become increasingly digitalized and better monitored. Nevertheless, it should be noted that the mathematical formulation of the state estimation algorithm must be adapted to each application, taking into account the particular features and characteristics of the monitored system, such as existing redundancy, measurement latencies, the power system model, measurement types, among others.

2.6.2 Mathematical formulation and basic components

As previously mentioned, state estimators are model-based algorithms that derive the state of the system from the raw measurements. From a technical point of view, the state of the EPES usually corresponds to the values of bus voltage magnitudes and phase angles with respect to the slack bus. Note that the calculation of the system state allows for the computation of any other magnitude/quantity in the power system, e.g. branch currents, power flows, etc.

In general, the system state is estimated by leveraging the underlying physical information of the EPES to create a mathematical model that correlates the real-time, raw measurements and the state variables. Therefore, the first step is the derivation of such a model to relate the observed real-time measurements, $\mathbf{z}$ (nodal voltages, active and reactive power flows, active and reactive power injections, etc.), acquired by different RTUs dispersed across the EPES, with the states, $\mathbf{x}$ (voltage magnitudes and phase angles):

$$\mathbf{z} = \mathbf{h}(\mathbf{x}) + \boldsymbol{\epsilon} \tag{2.1}$$

Here, $\mathbf{h}$ represents the underlying, non-linear model that correlates measurements and states, whereas $\boldsymbol{\epsilon}$ denotes the error related to the measurements. The state estimator computes an estimation of the actual state, $\hat{\mathbf{x}}$, in order

to satisfy that:

$$\mathbf{h}(\hat{\mathbf{x}}) + \boldsymbol{\epsilon} \rightarrow \mathbf{z} \tag{2.2}$$

Since the measurements are inaccurate, the estimate obtained for the unknown state, $\hat{\mathbf{x}}$, is also inaccurate. This raises the problem of how to formulate the state estimator in order to derive the optimal (i.e., most accurate) estimate of the unknown system state from the available measurements. To this end, different formulations of the state estimation algorithm can be found in the literature depending on how this estimation is computed. Among these, the most notable are:

- The maximum likelihood criterion, where the objective is to maximize the probability that the estimate of the state variable is the true value of the state variable vector.
- The weighted least-squares criterion, where the objective is to minimize the sum of the weighted squared differences between the estimated states and the actual measurements. The estimator that employs this criterion is called the *Weighted Least Squares Estimator* (WLSE).
- The minimum variance criterion, where the aim is to minimize the expected sum of squared deviations between the estimated state components and their corresponding true values.

These three approaches result in identical estimators when unbiased meter error distributions are assumed.

The (WLSE) is by far the most popular and widespread estimator in practical applications. Nevertheless, the WLSE may exhibit poor robustness in cases of multiple bad data and/or outliers in the input measurements. Consequently, alternative, more robust state estimators could prove a better choice, such as *M-estimators*, a group of estimators that follow the maximum likelihood criterion. A notable M-estimator is the *Schweppe-Huber Generalized-M Estimator* (SHGME), which can be implemented in a straightforward manner based on the conventional WLSE algorithm. Both state estimators, WLSE and SHGME, are analyzed in Section 4.3.1.

Nevertheless, and irrespective of the implemented mathematical algorithm, it is worth mentioning that a complete *SE application* for EPESs comprises the following essential functions:

- Topology processor: determines the one-line diagram of the EPES by processing the status (open/closed position) of circuit breakers and switches.
- Observability analysis: identifies whether the whole system can be estimated given the type of available measurements and their distribution in the system. If the entire system is not observable, the observable islands and unobservable branches are identified.
- State estimation solution: computes the best estimate of the system state, i.e., complex voltages (voltage magnitude and phase angle) at all the buses of the power system. This module also determines the optimal estimates for any other electrical magnitude of interest: power flows in lines and transformers, power injections, transformer taps, currents, etc.
- Bad data processing: detects and identifies gross errors in the measurement set if there is enough redundancy in the measurement configuration.

2.6.3 Role within WP2

Section 2.6.1 highlighted that model-based techniques, with emphasis on SE, can robustly analyze measurement data, effectively handling errors and uncertainties. Therefore, inconsistent data that deviate from the expected system state can be detected. In the same vein, intentionally manipulated data due to FDIAs can be identified. In addition, when suspicious data are detected, the state estimator can provide the most probable values for the system state based on the system model, ensuring continuity of operations while mitigating the impact of compromised data.

Based on the above, the methodology developed in the framework of WP2 involves tailoring SE techniques to specific use cases (PV plants and energy communities), considering factors such as the datasets used, the physical systems involved, and the characteristic time constants. More specifically, by providing appropriate actions in case of FDIAs, e.g., removing and replacing compromised data, the generic methodology of WP2 strives to maintain the integrity and reliability of the EPES against FDIAs. This approach not only enhances the detection of cyber threats, but also ensures that the EPES can continue to operate effectively even in the face of such challenges.

To ensure the practicality of implementation, the computational effort has to be minimized in order to facilitate the real-time FDI detection and response. To this end, challenges such as measurement errors, model mismatches, and the need for redundancy in measurements need to be addressed.

Finally, in the framework of the WP2 solution, the main SE function will be enhanced by additional tools, as detailed in Chapter 4.

3 Literature Review on FDI

3.1 Introduction

The significance of FDIAs, highlighted in Section 2.5.1, has sparked a growing interest in the scientific topic of FDI, from both attacking and defensive perspectives. Particularly over the past few years, the popularity of this research field has risen exponentially, with numerous research articles being published and addressing diverse aspects of the FDI problem.

Apart from the significance of FDIAs, another contributing factor to this increase in scientific popularity is the limitations of conventional state estimators. More specifically, although very effective in handling bad data, traditional SE tools have proved unable to successfully detect sophisticated FDIAs. Consequently, the FDI research field has emerged as a testing ground for a wide range of advanced algorithms and techniques [44].

As a result, a systematic review and classification of the published technical literature is required to navigate this growing wave of interest in FDI. In Section 3.2, a thorough literature review is conducted, highlighting the main focal points and current trends. Based on this review, Section 3.3 introduces a new comprehensive taxonomy for the published research on FDI.

Finally, it should be noted that the common ground of this body of literature, in alignment with the scope of WP2, is the focus on either:

- **The identification of FDI (FDII)**, leveraging information regarding the physical characteristics of the monitored EPES, as well as historical data and patterns related to its operation. This has already been highlighted in the Executive Summary, Chapter 1, and Section 2.5.1.
- **The preparation of FDIAs**, assuming that adversaries can launch FDIAs, through one or more cyber-physical layers, as described in Section 2.5. Therefore, the examined attack-related literature does not emphasize the mechanisms by which cyber actors could compromise one or more cyber-physical layers of the EPES to conduct FDIAs. In contrast, it focuses on the manipulation of the data themselves, once access to them has been achieved. Based on this, for the remainder of this deliverable, the term ”attack vector” (AV) is used to describe the final manipulated dataset after the execution of the FDIA.

3.2 Examined Literature and Bibliometric Analysis

The literature review conducted covers a 15-year period, spanning from 2011 to 2025. Year 2011 was selected as the starting point of our analysis, as several pioneering articles were published that year, such as [50, 55, 56]. In total, as of January 14, 2025, 1573 scientific journals have been published by IEEE on the field of FDI since 2011, as per IEEE Xplore. Additionally, 657 scientific works were identified from other publishers, including Elsevier, IET, etc. Among those, several review and survey studies were found, including [22, 57, 31, 27, 58, 59, 20, 23, 60]. This substantial volume of scientific publications underscores the importance and the worldwide interest in this topic.

The annual distribution of the identified IEEE publications is illustrated in Figure 3.1, demonstrating a consistent increase in interest in FDI, culminating in 378 papers published in 2024 alone. Furthermore, the multidisciplinary nature of the field and its applications is noteworthy. The distribution of publications per journal is visualized in Figure 3.2. The acronyms of the journals are clarified in Table 3.1, where the corresponding IEEE society is also mentioned. The diversity of journals and societies reveals the widespread nature of the FDI problem in distinct sectors and emphasizes the necessity for multidisciplinary strategies to address it.

This multitude of disciplinary-diverse publications needs to be narrowed, in order to identify the most relevant ones for the scope of the COCOON project. Ultimately, over 100 scientific works were selected and thoroughly reviewed. Details of these selected publications are provided in Annex A of the deliverable, which includes brief descriptions of their scope, insights, and critical assumptions.

Based on the study of the publications of Annex A, a new, comprehensive taxonomy for FDI works is proposed and elaborated in Section 3.3.

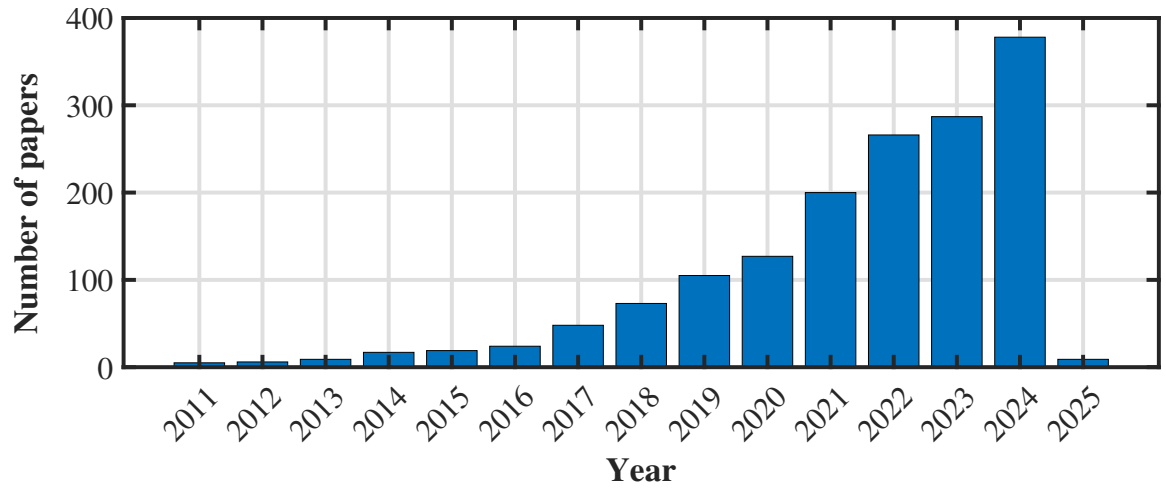

Figure 3.1: Annual distribution of IEEE Xplore publications on FDIAs.

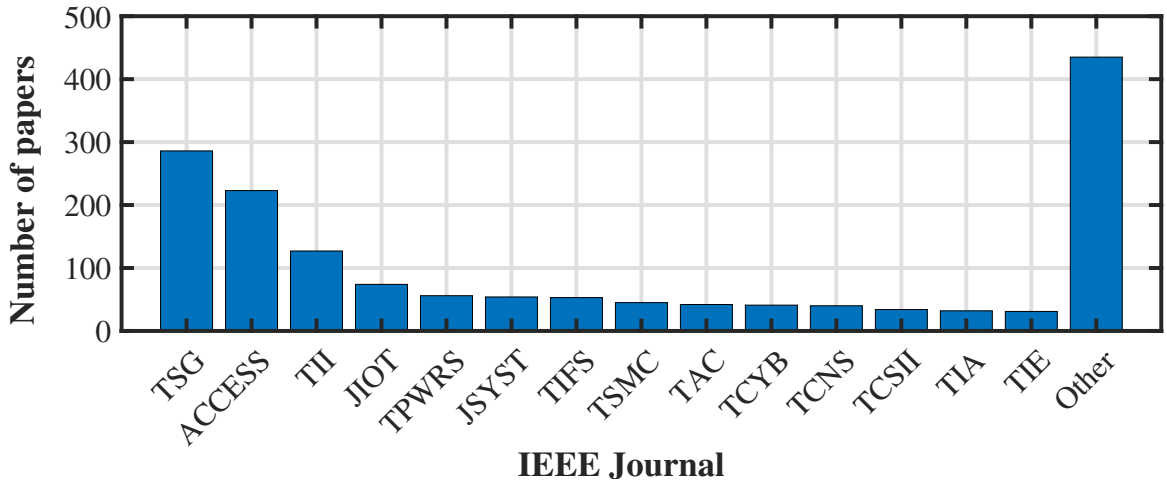

Figure 3.2: Distribution of publications per IEEE journal on FDIAs.

Table 3.1: Journal acronyms of Figure 3.2

	Full name	IEEE Society
TSG	Trans. on Smart Grid	Power and Energy
ACCESS	Access	-
TII	Trans. on Industrial Informatics	Industrial Electronics
JIOT	Internet of Things Journal	Sensors Council
TPWRS	Trans. on Power Systems	Power and Energy
JSYST	Systems Journal	Systems Council
TIFS	Trans. on Information Forensics and Security	Signal Processing
TSMC	Trans. on Systems, Man, and Cybernetics	Systems, Man, and Cybernetics
TAC	Trans. on Automatic Control	Control Systems
TCYB	Trans. on Cybernetics	Systems, Man, and Cybernetics
TCNS	Trans. on Circuits and Systems II	Circuits and Systems
TIA	Trans. on Industry Applications	Industry Applications
TIE	Trans. on Industrial Electronics	Industrial Electronics
Other	-	28 different societies

3.3 Proposed Taxonomy

In the frame of the COCOON project, a new taxonomy scheme is proposed to facilitate and systemize the study of the FDI literature, visualized in Figure 3.3. As depicted, the following six main categories are introduced for

the classification of papers:

- **Contribution:** This fundamental category divides papers based on their research scope and contribution, either proposing new attack techniques, defense mechanisms, or both.
- **Attack motivation:** Papers can be further distinguished whether the motivation behind the FDIA and its ultimate goal is technical, e.g., line overloading, or economic, e.g., financial benefit for the adversaries [24].
- **Attack**: This category organizes papers based on their assumptions and practices regarding FDIAs.
- **Defense:** This category organizes papers based on their assumptions and practices for FDII.
- **Attacked network type:** This category classifies papers with respect to the type of the power network they examine, i.e., transmission or distribution systems. In addition, distribution networks are further divided into balanced and unbalanced.
- **Measurement device:** This classification distinguishes papers that assume regular real-time measurements from the field coming from RTUs from others that take advantage of other types of measurements like those provided by PMUs.

In the subsections that follow, each category is thoroughly analyzed, while the corresponding papers are aggregated in comprehensive tables. The overall classification of the papers discussed in Annex A is presented in Table 3.16.

Finally, it is clarified that the **attack** and **defense** categories are not mutually exclusive, nor do they solely refer to papers whose contribution relates to FDIA or FDII, respectively. For papers whose contribution focuses on designing FDIAs, the **defense** category refers to the FDII assumptions under which their proposed attack schemes are tested. Similarly, for papers whose contribution focuses on designing FDII techniques, the **attack** category refers to the FDIA assumptions under which their proposed defense schemes are tested.

3.4 Main Contribution of the Reviewed Papers

The literature review of FDI reveals that contributions can be grouped into those dealing with the preparation of the attack and others tackling the defense mechanisms. Nevertheless, it is also possible to find some publications covering both issues at the same time. A summary of the papers categorized according to this feature can be found in Table 3.2. The analysis of this information indicates that most of the papers examined are devoted to defining defense tactics against FDIAs (66 papers or 61%), followed by those dealing with the matter of planning a successful attack (30 papers or 28%). Only 11 papers (10%) provide simultaneous contributions to both the attack and defense strategies.

Table 3.2: Distribution of papers with respect to their main contribution to the FDI topic.

Contribution	References
Attack	[61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 50] [73, 55, 74, 75, 76, 56, 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87]
Defense	[88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100, 101] [102, 103, 104, 105, 106, 107, 108, 109, 110, 111, 112, 113, 114] [115, 116, 117, 118, 119, 120, 121, 122, 123, 124, 125, 126, 127] [128, 129, 130, 131, 132, 133, 134, 135, 136, 137, 138, 139] [140, 141, 142, 143, 144, 145, 146, 147, 148, 149, 150, 151, 152]
Both	[153, 154, 155, 156, 157, 158, 159, 160, 161, 162, 163]

3.5 Motivation of the FDIA

The main objective of an FDIA is to introduce malicious information into the data flow between control centers and field devices. In this regard, it is possible to attack measurements coming from existing RTUs to the control center (measurement-related attacks) or setpoints sent from the control center to controllable devices

Figure 3.3: FDI Taxonomy.

(command-related attacks), as depicted in Figure 2.4. In either case, these actions may lead to an improper operation of the EPES but, beyond this, the cyber attacker has a motivation for launching such attacks. The motivation has been grouped into two broad categories: technical and economic. A technical motivation is defined as the one pursuing an impact on the EPES that may compromise the security of its operation, e.g., overvoltage/undervoltage situation, transformer or line overload, etc. On the other hand, an economically motivated attack is specifically designed to lead to a global economically inefficient operation of the EPES. Moreover, economic motivations may be driven by an attempt to improve the position of a particular stakeholder relative to its competitors. Table 3.3 shows that most of the reviewed papers have considered technical motivations (47 papers or 44%), while only 8 papers (7%) address FDIAs that aim to impact the EPES in an economic manner. A smaller number of papers (5 or 5%) identifies a simultaneous techno-economic motivation behind the cyber attack. It is also worth noting that a large number of papers (30 papers or 28%) does not provide any position on the motivations behind the FDIAs. It has to be considered, though, that a cyber attack with a technical motivation will inevitably have an associated economic impact derived from a non-optimal operation, e.g. larger

power losses, or an unsecured state that could potentially lead to partial or total outages.

Table 3.3: Distribution of papers with respect to the motivation behind the cyber attack.

Motivation	References
Technical	[61, 164, 91, 92, 93, 62, 94, 95, 96, 98, 155, 63, 64] [99, 156, 100, 102, 65, 66, 67, 103, 104, 105, 107] [82, 83, 124, 126, 73, 127, 55, 160, 74, 76, 84, 85, 132, 136] [137, 138, 150, 151, 87, 140, 80, 141, 143]
Economic	[154, 101, 69, 72, 125, 77, 136, 142]
Both	[50, 84, 144, 137, 144]
Not specified	[153, 88, 89, 90, 165, 97, 157, 122, 75, 56, 128, 129, 130, 131, 133, 134] [79, 135, 139, 149, 86, 152, 145, 81, 161, 162, 146, 147, 148, 163]

3.6 Network Type

This category refers to the type of network where the FDI attack is taking place, distinguishing between TSs and DSs. According to Table 3.4, most of the papers (86%) deal with cyber attacks in the transmission level, which is characterized by a balanced network model. Only 13 papers (12%) tackle FDIAs within the distribution level, and among those, only 6 papers consider the inherent unbalanced nature of this part of the EPES. This assumption can be justified considering that, in MV networks, the load imbalance is less severe than in LV grids. Moreover, in most cases, the assumption of a balanced network (i.e., only considering the positive-sequence component of the network) achieves sufficiently accurate results with simple models.

In any case, the results of this bibliometric analysis reveal that the study of FDIAs in the context of DSs has not yet been explored to a satisfactory extent. Nevertheless, as explained in Chapter 2, the operating conditions and the role of DSs within the EPES is rapidly evolving, with the two main changes being that:

- EPESs are shifting away from the centralized generation scheme, where a *few* large power plants connected to the transmission level are responsible for the energy production, toward a fully decentralized paradigm characterized by a *plethora* of small- and medium-size generation assets connected to DSs. This paradigm shift also affects the number of stakeholders involved in the generation activity, moving from a scenario dominated by a few generation companies with standardized technical procedures to one with diverse groups of newcomers employing uneven practices (the extreme example of this transformation corresponds to the so-called *prosumers*).
- DSs are undergoing a complete digital revolution, which is crucial for ensuring the security of supply in an increasingly stressed grid with highly volatile renewable energy resources.

The above-described phenomena highlight that DSs are becoming increasingly: (a) more important for the EPES, and thus essential and attractive targets for the malicious intentions of attackers, (b) more vulnerable cyber attacks through diverse entry points. Therefore, the cyber security of DSs requires greater attention from the scientific community in order to ensure the safe transition to a decarbonized society.

3.7 Measurement Device

Table 3.5 shows that most of the papers reviewed, particularly more than 90% of them, considered that the target of FDIAs is field data acquired with conventional measurement devices interfaced with RTUs. However, it is also possible to find some papers analyzing the role that PMUs may have in preventing FDIAs. This capability is derived from the specific communication requirements of the PMUs, particularly from their protocol and physical channel.

Table 3.4: Distribution of papers with respect to the targeted EPES part.

EPES level	References
Transmission	[61, 153, 88, 89, 90, 154, 91, 92, 93, 62, 94, 95, 97, 155, 63] [64, 99, 156, 101, 102, 67, 103, 104, 105, 106, 107, 68, 108] [109, 157, 69, 70, 110, 158, 111, 112, 113, 114, 115, 116] [117, 118, 71, 119, 78, 120, 159, 121, 122, 123, 72, 50, 82, 83] [124, 125, 126, 73, 127, 55, 160, 74, 75, 76, 56, 77, 84, 85, 129] [130, 131, 132, 79, 136, 137, 138, 150, 151, 87, 152, 140, 80] [141, 143, 144, 145, 81, 146, 147, 148, 163]
Distribution	[164, 96, 98, 100, 65, 66, 112, 78, 76, 129, 130, 142, 162]

Table 3.5: Distribution of papers with respect to the considered measurement devices.

Measurement type	References
PMUs	[153, 90, 95, 99, 156, 102, 65, 82, 74, 56, 129, 131, 137, 151, 163]
RTUs	[61, 153, 88, 164, 89, 90, 154, 91, 92, 93, 62, 94, 97, 98] [155, 63, 64, 156, 100, 102, 65, 66, 67, 103, 104, 105, 106, 107] [68, 108, 157, 69, 70, 110, 158, 111, 112, 113, 114, 115, 116, 117] [118, 71, 119, 78, 120, 159, 121, 122, 123, 72, 50, 82, 83] [124, 125, 126, 73, 127, 55, 160, 74, 76, 84, 85, 129, 130, 132, 136, 137] [138, 150, 151, 87, 152, 140, 80, 141, 81, 162, 146, 147, 163]

3.8 FDI Attack

This section performs a taxonomy of those papers focused on designing FDIAs. After the review of the papers, the following four categories have emerged: (i) accessibility of measurements, (ii) network knowledge, (iii) EPES model, and (iv) attack vector composition. The following subsection presents some details about these categories and how the reviewed papers are classified according to them.

3.8.1 Accessibility of measurements

An FDIA refers to the manipulation of measurements being transmitted from the field to the control center. Therefore, this category indicates the ability of the cyber attacker to modify the complete or just a partial set of the available measurements. Note that this ability is strongly related to the layer from which the cyber attack is carried out. To elaborate, those FDI attacks instantiated from the physical layer are usually characterized by a partial measurement manipulation, taking into account that physical devices are spread across the EPES and it is not feasible to have simultaneous access to all of them. On the contrary, if the FDI is launched from the communication, network or cyber layers, it would be possible to have a larger accessibility to the measurements. The papers considering partial accessibility to the measurements amount to 47 (44%), whereas 41 (38%) papers assume a full accessibility, as summarized in Table 3.6.

3.8.2 Network Knowledge

This is a key classification of FDIAs, since it directly influences the conditions under which the attack vectors are derived/composed, thus strongly conditioning the success of the attack. In this regard, it is convenient to resort to the mathematical formulation described in Section 2.6, which relates the states, $\mathbf{x}$, with the measurements, $\mathbf{z}$: $\mathbf{z} = \mathbf{h}(\mathbf{x}) + \boldsymbol{\epsilon}$, with $\mathbf{h}$ representing the network model. Let $\mathbf{z_a}$ denote the FDIA vector that modifies the actual field measurements $\mathbf{z}$. Therefore, with this in mind, it is possible to classify FDIAs into the following groups considering the information about the network knowledge that the attackers possess:

Table 3.6: Distribution of papers with respect to the accessibility of measurements.

Measur. access.	References
Partial	[61, 153, 88, 89, 154, 91, 93, 94, 95, 97, 98, 156, 66, 107]
	[69, 110, 111, 114, 115, 120, 50, 82, 83, 124, 55, 160, 74, 76, 56, 85]
	[130, 131, 132, 79, 137, 150, 151, 87, 152, 80, 142, 144, 145, 81, 162, 146, 147]
Full	[92, 62, 96, 155, 63, 64, 156, 100, 101, 65, 67, 103]
	[104, 68, 108, 70, 110, 158, 111, 112, 113, 114, 115, 116, 71, 119]
	[78, 159, 121, 123, 72, 125, 76, 84, 136, 140, 141, 143, 162, 147, 163]

- **Full information**: Full knowledge of the complete network model means that the cyber attacker is totally aware of $\mathbf{h}$, thus being possible for them to perform an undetectable attack. Indeed, the FDIA vector can be computed as $\mathbf{z_a} = \mathbf{h}(\mathbf{x} + \Delta\mathbf{x}) + \boldsymbol{\epsilon}$, where $\Delta\mathbf{x}$ refers to a modified value of the EPES states with respect to the actual ones $\mathbf{x}$. Note that, in this manner, the EPES states are consistent with the falsified field measurements and, therefore, the SE algorithms are not able to detect the attack. This type of FDIA is referred to as stealthy attack.
- **No information**: In this case, field measurements are simply changed without any specific criteria with respect to the network model. For this reason, $\mathbf{z_a} \neq \mathbf{h}(\mathbf{x}) + \boldsymbol{\epsilon}$, situation which can be detected by applying traditional SE techniques. This type of FDIA is referred to as dummy attack.
- **Partial information**: This is an intermediate situation, where the cyber attacker has only information about a specific region of the network. It can be demonstrated that, even in this less favorable scenario for the attack, it is possible to launch an FDIA which may compromise the EPES security [61].

The analysis of the reviewed papers, shown in Table 3.7, reveals that most of the papers consider complete knowledge of the network (about 44 papers or 41%), followed by those analyzing how to design attacks under limited information (29 papers or 27%). Finally, only 18 papers (17%) deal with attacks without network knowledge. From a practical point of view, the higher the information of the EPES to be attacked, the higher the probability of success. Obviously, the extent of network knowledge depends on where and how the FDIA is instantiated. More specifically, it is reasonable to expect that for FDIAs executed from the physical, communication, or network layers, it is unlikely for the attacker to possess full network knowledge. However, this is not the case for those FDIAs launched from the cyber layer, i.e. through SCADA systems, as in thoses cases, the cyber attacker could also gain access to the critical EPES information.

Table 3.7: Distribution of papers with respect to the considered network knowledge.

Network knowledge	References
None	[89, 91, 94, 95, 96, 97, 98, 63, 64, 108, 70, 110, 158, 72, 73, 160, 145, 81]
Partial	[61, 153, 88, 93, 156, 102, 69, 110, 111, 112, 114, 115, 120, 82]
	[83, 124, 125, 126, 127, 160, 74, 76, 85, 79, 137, 87, 152, 142, 147]
Full	[154, 92, 62, 155, 156, 100, 101, 65, 66, 67, 103, 104, 107, 68, 108, 111]
	[113, 114, 115, 116, 71, 119, 78, 159, 121, 123, 50, 55, 76, 56, 84]
	[130, 131, 132, 136, 150, 151, 140, 80, 141, 143, 144, 147, 163]

3.8.3 System Model

This category refers to the EPES model used to launch the FDIA, thus being directly related to the ”Network Knowledge” category previously analyzed. In the case of having full or partial information about the EPES, it is possible for the attacker to compute the false data to be injected by considering models of varying complex-

ity. Moreover, this complexity of the models used also affects the performance of attacks. More specifically, it should be noted that the static model of the EPES comprises a set of highly non-linear equations, which accurately represents the steady-state regime, albeit at the cost of the inherent complexity implied by a non-linear system. Therefore, the usage of an exact network model for the derivation of attack vectors increases the chances of a stealthy FDIA, as it allows for attack vectors that better align with the physical laws of the actual EPES, albeit at the expense of higher computational cost.

As an alternative, it is also possible for attackers to resort to approximate models for designing FDIAs, introducing a trade-off between the computational effort and the success of the attack. The literature review reveals that most of the approximate models used so far rely on the relaxation of the non-linear AC power flow equations to the so-called DC approximated models [83]. By performing these approximations, the problem is reduced solely to the analysis of active power flows, which are directly related to the difference of nodal phase angles. Despite its simplicity, this approximation provides satisfactory results for TSs due to the high X/R ratio. In addition to the DC approximated models, other previous works propose alternative simplified models based on the linearization of the power flow equations [141]. From a bibliometric point of view, as shown in Table 3.8, most of the papers (42%) use an approximate network model, while 33% adopt an exact model based on the non-linear power flow equations. The high percentage of papers with linearized models is justified by the fact that most studies are focused on TSs, as shown in Section 3.6, where approximated models perform adequately and thus the corresponding attack vectors can remain stealthy [87].

Table 3.8: Distribution of papers with respect to the considered EPES model.

EPES model	References
Exact	[154, 64, 156, 101, 67, 70, 158, 115, 116, 123, 72] [50, 82, 83, 125, 73, 55, 56, 84, 130, 131, 132, 79, 150] [151, 87, 152, 143, 144, 145, 81, 146, 147]
Approximate	[61, 153, 88, 92, 93, 62, 155, 156, 100, 102, 65, 66] [67, 103, 107, 68, 108, 110, 111, 112, 113, 114, 71, 119] [78, 120, 159, 121, 124, 126, 127, 160, 74, 76, 56, 85, 136] [137, 140, 80, 141, 142, 81, 162, 163]

3.8.4 Attack Vector Composition

This field of the proposed taxonomy elaborates on the different approaches that cyber attackers might adopt to compose the attack vector, i.e., to compute the false data that will replace the actual measurement values. According to Table 3.9, three different types of strategies can be found in the papers reviewed. First, strategies based on a deterministic approach (almost a 50%), where the false data are computed considering a network model that might in turn be exact or approximated, as well as partial or complete, as per the other taxonomy fields previously analyzed. In addition, it is also possible to find papers in which false data are derived through statistical methodologies and observations of the EPES measurements, without any previous information about its model. Finally, the use of AI techniques aimed to derive a network model from observed measurements to launch attacks has also been reported but with application limited to linear systems.

3.9 FDI Defense

Six subcategories are introduced to characterize the assumptions and practices of papers with respect to FDII, as elaborated in the subsections that follow.

3.9.1 Purpose

This classification refers to the main goal of FDII, which is either to localize the injected false data, i.e., to explicitly specify to which physical quantities of the EPES the falsified data correspond (referred to as false data localization, FDL), or to simply detect the existence of false data within the dataset (No FDL).

Table 3.9: Distribution of papers with respect to the attack vector composition.

Attack vector definition	References
Deterministic	[61, 153, 88, 89, 91, 92, 93, 94, 95, 96, 97, 98] [155, 63, 99, 156, 100, 102, 66, 67, 104, 107, 68, 84] [85, 130, 131, 132, 79, 136, 137, 150, 151, 87, 152, 140] [80, 141, 142, 143, 144, 81, 162, 146, 147, 163]
Statistical	[62, 64, 65, 108, 110, 158, 72, 145, 81]
AI-based	[70]

Table 3.10: Distribution of papers with respect to the purpose of their assumed FDII schemes.

Purpose	References
FDL	[86, 89, 91, 93, 98, 102, 103, 109, 110, 158, 111, 112, 113, 115, 116, 117, 119, 120] [121, 125, 126, 73, 127, 55, 56, 129, 131, 137, 138, 151, 142, 145, 162, 146, 147, 163]
No FDL	[153, 92, 95, 97, 155, 63, 64, 99, 156, 100, 65, 103, 106, 107, 68, 108, 114, 118] [159, 124, 160, 130, 132, 136, 83, 87, 152, 140, 80, 141, 143, 144, 81]

3.9.2 Complementary Mechanisms

This category divides defense strategies into those that employ complementary defensive mechanisms alongside the basic FDII tool, and those that do not. More specifically, defensive complementary mechanisms refer to the following security schemes:

- **Installation of PMUs**: In general, PMUs are considered more secure and difficult to be tampered, in comparison with conventional meters. Consequently, their optimal strategic placement has been meticulously discussed in the literature [22, 166, 167], as an additional protective layer against FDIAs. Nevertheless, it should be noted that PMU measurements are also exposed to cyber threats, mainly through GPS spoofing practices and time synchronization attacks [45, 46, 168].
- **Moving target defense (MTD):** The concept of MTD, i.e., the deliberate modification of the system topology or parameters (admittance perturbation), has been widely addressed in the literature. By slightly modifying the topology or the parameters of the monitored grid on purpose, system operators can negate the potential knowledge that attackers might have for the grid. Devices such as Flexible AC Transmission Systems (FACTSs) or switches can be utilized to perform MTD.
- **Meter coding:** Meter coding refers to the practice of encoding data, e.g., through watermarking [169, 170]. Meter coding practices can enhance the accuracy of FDII against stealthy FDIAs.

Table 3.11: Papers that employ complementary defense mechanisms.

Complementary Mechanism	References
PMU	[153, 102, 50]
MTD	[93, 103, 105, 158, 115, 130, 65, 136]
Meter coding	[92, 98, 104, 132, 152]

3.9.3 State Estimation Architecture

This categorization divides defense strategies with respect to the type of SE they employ. More specifically, SE approaches can be classified as centralized or distributed. In the centralized approach, which is the most

common, field measurements are gathered at the premises of system operators, in order for state estimation to be centrally performed at their SCADA. In contrast, in distributed SE, the monitored grid is divided into sub-networks. For each sub-network, SE is locally and independently performed. Subsequently, based on the results of those local, independent SEs, the states of the boundary buses of the entire grid are estimated. Table 3.12 aggregates the papers where distributed SE is employed; the remaining papers employ centralized SE.

Table 3.12: Papers that employ distributed SE schemes.

SE Architecture	References
Distributed SE	[98, 159, 129, 138, 145]

3.9.4 System Model

This is a basic category and regards the type of system model considered in the FDII algorithm, differentiating between the complete, exact, and more accurate AC model, and a simplified, approximated model. It should be noted that an approximated model refers to any linearized approximation of the complete AC model, including the common linear model used for DC power flow analysis in transmission systems. Therefore, even for distribution systems, approximated models could be used, if simplicity and computational efficiency is prioritized over modeling accuracy. The bibliometric analysis reveals that the use of exact models has prevailed over the last years, while most of the initial studies employed DC models.

Table 3.13: Distribution of papers with respect to the system model assumed for FDII.

System Model	References
Approximated	[86, 89, 90, 85, 64, 101, 67, 158, 115, 123, 72, 50, 125, 73, 55, 129, 130, 131] [132, 83, 151, 87, 152, 145, 81, 146, 147]
Exact	[153, 88, 91, 92, 93, 94, 155, 63, 99, 100, 102, 65, 66, 67, 103, 104, 105, 106] [107, 68, 108, 109, 110, 111, 112, 113, 114, 116, 117, 118, 71, 119, 78, 120, 159, 121] [82, 83, 124, 126, 127, 160, 136, 138, 140, 80, 141, 71, 144, 81, 162, 163]

3.9.5 Measurement-based Plausibility Analysis

In this category, the works and their proposed/employed methods are taxonomized with respect to the type of plausibility analysis conducted, if any. The term ”measurement-based plausibility analysis” (MbPA) is used to describe the assessment of received measurements that is performed prior to applying the main SE and FDD tools. Plausibility analysis approaches can be classifed as follows:

- **Deterministic**, i.e., checking whether the received measurement values lie within their permissible technical limits.
- **Statistical**, i.e., compare the concurrent measurements with the database of historical measurements through statistical indices.
- **AI-based**: AI-based tools refer to the usage of machine learning or neural network techniques, which have been trained on the historical measurement dataset. In turn, the AI-based plausibility analysis solutions can be further classified into physics-informed (PI) or non-PI. PI techniques take into account the physical and spatial correlation between measurements of different EPES quantities, as dictated by the physical laws governing the EPES. In contrast, non-PI methods are limited to the analysis of temporal characteristics of the measurement timeseries.

3.9.6 False Data Detection

Defense strategies can be categorized with respect to the FDD type employed. Within the framework of the COCOON project and this deliverable, FDD refers to any analysis performed after the application of the main SE, to investigate the presence of suspicious data within the dataset. Moreover, FDD is differentiated from MbPA,

Table 3.14: Distribution of papers with respect to the type of measurement-based plausibility analysis performed.

Plausibility Analysis	References
Deterministic	[90, 85, 101, 102, 107, 121, 83, 151, 87]
Statistical	[153, 88, 108, 110]
Non-PI AI	[93, 95, 97, 155, 156, 109, 113, 124, 125, 126, 136, 137]
PI AI	[155, 110, 116, 118, 127]

as the former is conducted on the derived states or residuals dataset, after the initial SE execution. Therefore, the notion of FDD in the COCOON project extends beyond traditional residual-based bad data detectors.

Similarly to the measurement-based plausibility analysis, the FDD types are classified into deterministic, statistical, and AI-based.

- **Deterministic**: This is the fundamental category, referring to approaches such as the residual-based analysis of traditional FDD.
- **Statistical or AI-based**: The same taxonomy logic as for the measurement-based plausibility analysis.
- **Other:** This category includes FDD techniques that do not fit into the above classifications.

Table 3.15: Distribution of papers with respect to the FDD type.

FDD	References
Deterministic	[86, 89, 93, 63, 99, 100, 102, 65, 66, 67, 103, 105, 106, 68, 158, 111] [112, 114, 115, 71, 78, 72, 50, 82, 83, 73, 55, 130, 132, 136] [150, 152, 80, 81, 162, 146, 147]
Statistical	[153, 91, 92, 93, 94, 63, 64, 112, 117, 120, 160, 129, 131, 138, 87, 142, 143, 145, 163]
AI	[155, 106, 107, 114, 116, 119, 159, 123, 140, 141, 148]
Other	[111, 151, 144]

3.10 Proposed FDI Taxonomy and Paper Classification

Table 3.16: Papers categorization based on the proposed FDI Taxonomy.

Ref.	Contribution		Network Type				Attack										Defense															Measurement Device		Attack Motivation	
							Measurement Accessibility		Network Knowledge			System Model		Attack Vector Composition			System Model		SE Architecture		Measurement-based Plausibility Analysis				FDD				Complementary Mechanism			Purpose					
																							AI-based												
	Attack	Defense	TS	DS	Balanced	Unbalanced	Partial	Full	None	Partial	Full	Approx.	Exact	Deterministic	Statistical	AI-based	Approx.	Exact	Centralized	Distributed	Deterministic	Statistical	No PI	PI	Statistical	AI-based	Deterministic	Others	Meter Coding	MTD	PMU	No FDL	FDL	PMU	RTU	Economic	Technical
[61]	✓		✓		✓		✓			✓			✓	✓			✓									✓						✓			✓	✓
[153]	✓	✓	✓		✓		✓			✓			✓	✓				✓	✓			✓			✓						✓	✓	✓	✓		
[88]		✓	✓		✓		✓			✓			✓	✓				✓	✓			✓											✓			
[164]				✓																													✓			✓
[89]		✓	✓		✓		✓		✓					✓			✓		✓								✓					✓	✓			
[90]		✓	✓		✓												✓		✓		✓									✓			✓			
[154]	✓	✓	✓		✓		✓				✓	✓					✓		✓		✓													✓	✓	
[91]		✓	✓		✓		✓		✓					✓				✓	✓						✓								✓			✓
[92]		✓	✓		✓			✓			✓		✓	✓				✓	✓						✓				✓			✓		✓		✓
[93]		✓	✓		✓		✓			✓			✓	✓				✓	✓				✓		✓		✓			✓			✓			✓
[62]	✓		✓		✓			✓			✓		✓		✓																		✓			✓
[94]		✓	✓		✓		✓		✓					✓				✓	✓						✓								✓			✓
[95]		✓	✓		✓		✓		✓					✓					✓				✓									✓		✓		✓
[96]		✓		✓		✓		✓	✓					✓																						✓
[165]																																				
[97]		✓	✓		✓		✓		✓					✓					✓				✓									✓		✓		
[98]		✓		✓			✓		✓					✓						✓			✓				✓			✓		✓		✓		✓
[155]	✓	✓	✓		✓			✓			✓		✓	✓					✓				✓	✓		✓						✓		✓		✓
[63]	✓		✓		✓			✓	✓					✓					✓						✓		✓					✓		✓		✓
[64]	✓		✓		✓			✓	✓			✓			✓		✓		✓						✓								✓		✓		✓
[99]		✓	✓		✓									✓				✓	✓								✓						✓			✓
[156]	✓	✓	✓		✓		✓	✓		✓	✓	✓	✓	✓					✓				✓										✓		✓		✓
[100]		✓		✓	✓	✓		✓			✓		✓	✓				✓	✓								✓						✓			✓	✓
[101]		✓	✓		✓			✓			✓	✓					✓		✓		✓														✓		
[102]		✓	✓		✓					✓			✓	✓				✓			✓						✓			✓		✓	✓	✓		✓
[65]	✓			✓	✓			✓			✓		✓		✓			✓	✓								✓				✓		✓	✓		✓
[66]	✓			✓		✓	✓				✓		✓	✓				✓									✓							✓		✓
[67]	✓		✓		✓			✓			✓	✓	✓	✓			✓	✓	✓								✓							✓		✓
[103]		✓	✓		✓			✓			✓		✓					✓	✓								✓			✓		✓	✓	✓		✓
[104]		✓	✓					✓			✓			✓				✓										✓						✓			✓
[105]		✓	✓		✓													✓	✓								✓			✓				✓		✓
[106]		✓	✓		✓													✓	✓						✓		✓					✓		✓		✓
[107]		✓	✓		✓		✓				✓		✓	✓				✓	✓		✓				✓								✓		✓		✓
[68]	✓		✓		✓			✓			✓		✓	✓				✓	✓								✓					✓		✓			✓
[108]		✓	✓		✓			✓	✓		✓		✓	✓	✓			✓	✓			✓										✓		✓			✓
[109]		✓	✓		✓													✓	✓				✓										✓		✓		✓
[157]	✓	✓																																		
[69]	✓				✓		✓			✓				✓																					✓	
[70]	✓							✓	✓			✓				✓																				✓
[110]		✓	✓		✓		✓	✓	✓	✓			✓	✓	✓			✓	✓			✓		✓									✓		✓		✓
[158]	✓	✓	✓		✓			✓	✓			✓			✓		✓		✓								✓			✓			✓		✓		✓
[111]		✓	✓		✓		✓	✓		✓	✓		✓	✓				✓	✓								✓	✓					✓		✓		✓
[112]		✓		✓		✓		✓		✓			✓	✓				✓	✓						✓		✓						✓		✓		✓
[113]		✓	✓		✓			✓			✓		✓	✓				✓	✓				✓										✓		✓		✓

Table 3.16: Papers categorization based on the proposed FDI Taxonomy.

Ref.	Contribution		Network Type				Attack										Defense																Measurement Device		Attack Motivation	
							Measurement Accessibility		Network Knowledge			System Model		Attack Vector Composition			System Model		SE Architecture		Measurement-based Plausibility Analysis				FDD				Complementary Mechanism			Purpose					
																							AI-based														
	Attack	Defense	TS	DS	Balanced	Unbalanced	Partial	Full	None	Partial	Full	Approx.	Exact	Deterministic	Statistical	AI-based	Approx.	Exact	Centralized	Distributed	Deterministic	Statistical	No PI	PI	Statistical	AI-based	Deterministic	Others	Meter Coding	MTD	PMU	No FDL	FDL	PMU	RTU	Economic	Technical
[114]		✓	✓		✓		✓	✓		✓	✓		✓	✓				✓	✓						✓	✓						✓			✓		✓
[115]		✓	✓		✓		✓	✓		✓	✓	✓		✓			✓		✓							✓				✓			✓		✓		✓
[116]		✓	✓		✓			✓			✓	✓		✓				✓	✓				✓			✓							✓		✓		✓
[117]		✓	✓		✓									✓				✓	✓					✓									✓		✓		✓
[118]		✓	✓		✓													✓	✓				✓									✓		✓		✓	✓
[71]	✓		✓		✓			✓			✓		✓	✓				✓	✓							✓									✓		✓
[119]		✓	✓		✓			✓			✓		✓	✓				✓	✓						✓								✓		✓		✓
[78]	✓			✓		✓		✓			✓		✓	✓				✓	✓							✓									✓		✓
[120]		✓	✓		✓		✓			✓			✓	✓				✓	✓						✓								✓		✓		✓
[159]	✓	✓	✓		✓			✓			✓		✓	✓				✓		✓						✓						✓			✓		✓
[121]		✓	✓		✓			✓			✓		✓	✓				✓	✓		✓											✓					✓
[122]		✓																																			
[123]		✓	✓		✓			✓			✓	✓		✓			✓		✓						✓									✓			✓
[72]	✓		✓		✓			✓	✓			✓			✓		✓		✓							✓								✓		✓	
[50]	✓		✓		✓		✓				✓	✓		✓			✓		✓							✓								✓		✓	✓
[82]	✓		✓		✓		✓			✓		✓		✓				✓	✓							✓				✓			✓	✓			✓
[83]	✓		✓		✓		✓			✓		✓		✓				✓	✓							✓								✓			✓
[124]		✓	✓		✓		✓			✓			✓	✓				✓	✓				✓								✓			✓			✓
[125]		✓	✓		✓			✓		✓		✓		✓			✓		✓				✓									✓		✓	✓		
[126]		✓	✓		✓		✓			✓			✓	✓				✓	✓				✓									✓		✓			✓
[73]	✓		✓		✓		✓		✓			✓		✓			✓		✓							✓						✓		✓			✓
[127]		✓	✓		✓		✓			✓			✓	✓				✓	✓					✓								✓		✓			✓
[55]	✓		✓		✓		✓				✓	✓		✓			✓		✓							✓						✓		✓			✓
[160]	✓	✓	✓		✓		✓		✓	✓			✓	✓				✓	✓					✓							✓			✓			✓
[74]	✓		✓		✓		✓			✓			✓	✓																		✓					✓
[75]	✓																																				
[76]	✓			✓	✓		✓	✓		✓	✓		✓	✓																				✓			✓
[56]	✓		✓		✓		✓				✓	✓	✓	✓																		✓	✓				
[77]	✓		✓		✓																														✓		
[84]	✓		✓		✓			✓			✓	✓		✓																				✓		✓	✓
[128]		✓																																			
[85]	✓		✓		✓		✓			✓			✓	✓																				✓			✓
[129]		✓		✓		✓											✓			✓					✓								✓	✓	✓		
[130]		✓		✓		✓	✓				✓	✓		✓			✓		✓							✓				✓		✓			✓		
[131]		✓	✓		✓		✓				✓	✓		✓			✓		✓						✓								✓	✓			
[132]		✓	✓		✓		✓				✓	✓		✓			✓		✓							✓			✓	✓		✓			✓		✓
[133]		✓																																			
[134]		✓																																			
[79]	✓		✓		✓		✓			✓		✓		✓																							
[135]		✓																																			
[136]		✓	✓		✓			✓			✓		✓	✓				✓	✓				✓			✓				✓		✓			✓	✓	✓
[137]		✓	✓		✓		✓			✓			✓	✓					✓				✓									✓	✓	✓		✓	✓
[138]		✓	✓		✓													✓	✓						✓							✓		✓			✓
[139]		✓																																			
[149]		✓																																			

Table 3.16: Papers categorization based on the proposed FDI Taxonomy.

Ref.	Contribution		Network Type				Attack										Defense																Measurement Device		Attack Motivation	
							Measurement Accessibility		Network Knowledge			System Model		Attack Vector Composition			System Model		SE Architecture		Measurement-based Plausibility Analysis				FDD				Complementary Mechanism			Purpose					
	Attack	Defense	TS	DS	Balanced	Unbalanced	Partial	Full	None	Partial	Full	Approx.	Exact	Deterministic	Statistical	AI-based	Approx.	Exact	Centralized	Distributed	Deterministic	Statistical	AI-based		Statistical	AI-based	Deterministic	Others	Meter Coding	MTD	PMU	No FDL	FDL	PMU	RTU	Economic	Technical
																							No PI	PI													
[150]		✓	✓		✓		✓				✓	✓		✓			✓		✓		✓						✓					✓			✓		✓
[86]	✓																																				
[151]		✓	✓		✓		✓				✓	✓		✓			✓		✓		✓							✓					✓	✓	✓		✓
[87]	✓		✓		✓		✓			✓		✓		✓			✓		✓		✓				✓							✓			✓		✓
[152]		✓	✓		✓		✓			✓		✓		✓			✓		✓								✓		✓			✓			✓		
[140]		✓	✓		✓			✓			✓		✓	✓				✓	✓							✓						✓			✓		✓
[80]	✓		✓		✓		✓				✓		✓	✓				✓	✓								✓					✓			✓		✓
[141]		✓	✓		✓			✓			✓		✓	✓				✓	✓							✓						✓			✓		✓
[142]		✓		✓	✓		✓			✓			✓	✓				✓		✓					✓								✓			✓		
[143]		✓	✓		✓			✓			✓	✓		✓					✓						✓							✓					✓
[144]		✓	✓		✓		✓				✓	✓		✓				✓	✓						✓			✓				✓				✓	✓
[145]		✓	✓				✓		✓			✓			✓		✓			✓					✓								✓					
[81]	✓		✓		✓		✓		✓			✓	✓	✓	✓		✓	✓	✓								✓					✓			✓		
[161]	✓	✓																																			
[162]	✓	✓		✓	✓		✓	✓					✓	✓				✓	✓								✓					✓			✓		
[146]		✓	✓		✓		✓					✓		✓			✓		✓								✓					✓			✓		
[147]		✓	✓		✓		✓	✓		✓	✓	✓		✓			✓		✓								✓					✓			✓		
[148]		✓	✓		✓																				✓												
[163]	✓	✓	✓		✓			✓			✓		✓	✓				✓	✓						✓								✓	✓	✓		

4 COCOON Generic FDII Methodology

4.1 Introduction

This chapter describes the generic FDII methodology to be implemented in COCOON. For this purpose, it is important to understand the application field of the developed FDII methodology, i.e., the PV power parks and energy communities. Both fields are characterized by being almost totally balanced EPES applications with measurements mainly gathered from RTUs. With this application field in mind, the main assumptions and minimum requirements for the FDII methodology are outlined considering the taxonomy outcomes detailed in the previous chapter. Then, the architecture of the FDII procedure is detailed with emphasis on the main envisioned components, namely: MbPA, SE, FDD, and FDI response (FDIR). It has to be noted that some of these components are particularized for the WLS and Huber formulations of the state estimator. In addition to the mathematical description of each of these components, an algorithmic implementation in the form of a flowchart is included. Moreover, further details regarding the programming language to be used for the implementation, along with the expected required software libraries are included to provide a straightforward integration within the COCOON ecosystem. Finally, the chapter closes with a description of the tests required to evaluate the performance of the developed FDII methodology. This includes a description of how the attack vectors will be computed and the definition of some well-established metrics in order to characterized in a quantitative manner the performance of the FDII methodology.

4.2 Application Field, Assumptions and Requirements

According to Table 3.16, it can be observed that DSs have been overlooked from the FDI analysis, since the vast majority of research works focuses on TSs. In addition, to the best of authors' knowledge, FDI analysis at RES level, e.g., within a PV plant, has been not examined in the literature. The above can be regarded as important shortcomings for modern EPESs, where a large RES share is connected to the DSs. As a result, cyber security threats, especially FDIAs, may jeopardize the secure and reliable operation of DRESs and/or DSs, potentially triggering cascading effects across the EPES.

COCOON aims to fill these gaps by developing a generic FDII methodology that can be readily applied to both RES and DS levels. Special emphasis is placed on MV DSs, where the largest share of the overall RES installed capacity is connected. Contrary to LV DSs, where single-phase loads and/or RESs are connected, MV DSs are characterized by a balanced configuration and loading. Therefore, COCOON focuses on RESs connected to balanced, MV DSs, representing a typical European EPES configuration.

Regarding the monitoring infrastructure at RES and DS levels, it primarily involves the use of RTUs, which are the most common type of measurement devices employed by RES plant operators and DSOs. In addition to RTUs, PMUs will also be considered in the generic FDII methodology due to their ever-increasing use at the DS level. This way, a diverse set of measurement devices will be examined including RTUs, which are more vulnerable to cyber attacks, and PMUs, whose measurements are generally considered more trustworthy.

Finally, COCOON mainly focuses on the technical motivations behind an FDIA, e.g., line overloading, voltage violation, etc., which, in turn, may trigger cascading events within EPES. It is evident that these FDIAs not only jeopardize the secure and reliable operation of the EPES but also have important economic implications, affecting key financial indices such as market prices, operational costs, etc.

Based on the above, Table 4.1 summarizes the application field of COCOON generic FDII methodology.

Regarding the assumptions in the development of the FDII methodology, an important prerequisite for a successful FDII is to exploit additional knowledge/insights that are usually not available to the attacker, mainly due to their limited financial and/or computational resources. In addition, the attacker may have limited expertise in the field of electrical engineering and especially on EPES modelling, operation, and control. This is also reflected in Table 4.2, where the assumptions on both FDIA and FDII are presented. Specifically, unlike the DSO or the RES plant operator that has full access to the measurement devices, the attacker can tamper only

Table 4.1: Application field of COCOON FDII methodology.

Network Type	Balanced, MV DS
Measurement Devices	RTUs and PMUs
Attack Motivation	Technical

Table 4.2: Assumptions and requirements on FDI.

		FDIA	**FDII**
Assumptions	**Measurement Accessibility**	Partial	Full
	Network Knowledge	Partial	Full
	System Model	Exact	Exact
	AV Composition	Deterministic	–
Requirements	**MbPA**	–	Deterministic*
	SE Architecture	–	Centralized
	FDD	–	Deterministic*
	Complementary Mechanisms	–	No
	Purpose	–	FDL

* *Minimum requirement.*

with a portion of them, resulting in partial accessibility. This is also the most common approach adopted in relevant research works, as shown in Table 3.16. Similarly, the attacker has partial knowledge of the DS/RES configuration in terms of network topology, line parameters, etc. On the contrary, the DSO or the RES plant operator has a detailed and accurate knowledge of the DS/RES configuration, which is essential for ensuring highly accurate monitoring and control activities.

In addition, an exact system model is considered for both FDIA and FDII where the full, nonlinear, AC power flow equations are used to model the grid operation of either the DS or within the RES power plant. Note that this is a well-established approach adopted by most of the relevant works, as verified in Table 3.16. Finally, regarding attack vector composition, a deterministic approach is adopted, since it has been widely examined in the literature and can be readily applied from an attacker point of view.

Table 4.2 also contains the main functional requirements for the generic FDII methodology that will be developed in the frame of COCOON. In particular, a centralized SE architecture will be adopted in order to be aligned with the centralized control and monitoring schemes that are commonly employed by DSOs and RES plant operators. Apart from identifying the existence of false data, the generic FDII methodology aims to explicitly specify which EPES data have been falsified (FDL functionality). Moreover, concerning MbPA and FDD, the minimum requirement involves the use of deterministic approaches, as it will be further analyzed in Section 4.3. However, in the framework of COCOON, advanced methods will also be explored, e.g., statistical approaches, to further improve the performance and accuracy of the generic FDII methodology. Finally, no complementary mechanisms will be used, since they are characterized by increased investment and operational costs, hindering their applicability under real field conditions.

4.3 Architecture

This section provides the details about the architecture of the FDII procedure. With this purpose, the main components of the application are first described by providing the required mathematical details. This will unlock the software requirements to integrate the application into the COCOON ecosystem. Subsequently, the

section analyzes the interaction of these components through a description of the implementation algorithm, providing the details about the sequence of execution and the corresponding input data, as well as the expected outputs. It is worth mentioning that a preliminary analysis of the proposed generic FDII methodology has been presented in [171].

4.3.1 Main Components

Figure 4.1 shows a schematic of the four main components of the proposed methodology for identifying FDIAs, considering the underlying physics of an EPES as per its corresponding steady-state model. Each of these components is analyzed in the next subsections.

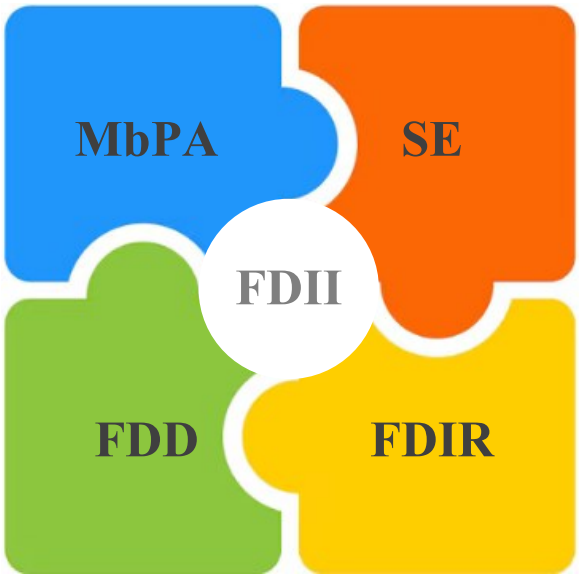

Figure 4.1: Main Components of COCOON generic FDII methodology.

Measurement-based Plausibility Analysis (MbPA)

The aim of this component is to act as a pre-filtering stage prior to the application of the SE algorithm. In this regard, it is important to recall the aim of the FDII methodology: guarantee that all the data passing through the OT communication infrastructure from the field RTUs to the control centers are free of intentionally corrupted information. Let us consider that the RTUs inform at regular time intervals about a measured set of electrical magnitudes to the control center. Basically, this first stage of the FDII methodology consists on questioning if each of these measurements considered as individuals, i.e., without considering their possible inter-relationships because of the actual physical system, is consistent. For this purpose, it is proposed to apply different techniques, summarized as follows:

Deterministic methods: This is probably the most straightforward strategy, since it simply exploits the comparison of the measured electrical magnitude to a representative reference value established, considering some properties of the measured quantity. As an example, let's consider a PV inverter with the following rated magnitudes: voltage U_n, current I_n, and apparent power S_n. The measured electrical variables, U_m, I_m, P_m, and Q_m related to this inverter must satisfy the following equations:

$$U_{min} \leq U_m \leq U_{max} \rightarrow (1-k)U_n \leq U_m \leq (1+k)U_n \tag{4.1}$$

$$I_m \leq I_n \tag{4.2}$$

$$0 \leq P_m \leq S_n \tag{4.3}$$

$$-\sqrt{S_n^2 - P_m^2} \leq Q_m \leq \sqrt{S_n^2 - P_m^2} \tag{4.4}$$

Equation (4.1) indicates that the measured voltages have to be within a range around the rated voltage, where k refers to the percentage of permissible deviation. k depends on each specific application but generally lies around $\pm 10\%$. Equation (4.2) imposes that the measured current has to be always below the rated one because, otherwise, the PV inverter would have experienced damage due to overloading. Similarly, (4.3) indicates that the

active power has to be always below the rated apparent power of the PV inverter, given the fact that usually the operation is with a power factor close to unity. Finally, (4.4) refers to the available reactive power capability of the PV inverter considering the measured active power and the rated apparent powers. This deterministic method would provide some protection against naive cyber attacks in which the attackers do not possess knowledge about the physical system and have not conducted a prior analysis on regular data of the OT traffic.
Statistical methods. In contrast to the previous deterministic approach, these methods take advantage of the statistical properties of the time series of the measured electrical magnitudes. Let's assume a time series during a representative time interval, e.g. one year, of a given electrical magnitude, $x_k, (k = 1, ..., N)$. This dataset can be characterized by a mean and a standard deviation, μ_x and σ_x, which can be computed as:

$$\mu_x = \frac{1}{N}\sum_{k=1}^{N} x_k \tag{4.5}$$

$$\sigma_x = \sqrt{\frac{1}{N}\sum_{k=1}^{N}(x_k - \mu_x)^2} \tag{4.6}$$

$$\tag{4.7}$$

After the analysis of the dataset, it is possible to define a range around the mean value, i.e., $\mu_x \pm \alpha\sigma_x$, in which most of the time series lies. Any value outside this range is suspicious of indicating an FDIA. Note that α is a user-defined parameter: the larger α is, the higher the number of the dataset samples within the defined range. Therefore, this parameter has to be carefully adjusted to balance the performance of the method (false positives versus undetected *mild* FDIAs). In addition, it is also possible to find other approaches in the specialized literature:

- Benford's law [172]. Also known as the law of anomalous numbers or the first-digit law, it is based on the observation that in many datasets, the most significant digit of the data is usually small and follows a probability distribution:

$$P_b(d) = \log_b(d+1) - \log_d(d) = \log_b(1 + \frac{1}{1+d}) \tag{4.8}$$

  where b is the base in which the data is expressed and d is the digit. Figure 4.2 shows the distribution of first digits according to Benford's law in the case of decimal numbers ($b = 10$). Therefore, if the dataset naturally follows the Benford's law and a FDIA takes place without considering this fact, it would be possible to detect it. Particularly, this occurs if the FDI is executed following a uniform distribution, i.e., if the most significant digit of the introduced false data has the same probability to assume any value.
- Hellinger distance [143]. This index measures the similarity between two probability distributions $p(x)$ and $q(x)$ as:

$$HD(p,q)^2 = 1 - \sqrt{\int p(x)q(x)dx} \tag{4.9}$$

  where $HD(p, q)$ is the Hellinger distance ($0 \leq HD(p, q) \leq 1$). Note that when $HD(p, q) = 1$, the probability distributions $p(x)$ and $q(x)$ are exactly the same, while $HD(p, q) = 0$ means just the opposite. Therefore, if an FDI cyber attack takes place, disturbing the natural probability distribution of some measurements, the Hellinger distance between unattacked datasets (i.e., historical values) and the attacked ones must decrease.

State estimator
As explained in Section 2.6, SE is a procedure in which measurements are used to estimate the value of one or more unknown parameters of a system, in the COCOON case, the EPES state. Since the measurements are inaccurate, the estimate obtained for the unknown state is also inaccurate. This raises the problem of how to formulate the optimal estimate of the unknown system state from the available measurements. As commented in Section 2.6, the development of a SE algorithm can follow several strategies, depending on the statistical criterion chosen. Next, two fundamental estimators are presented, the WLSE and the SHGME.

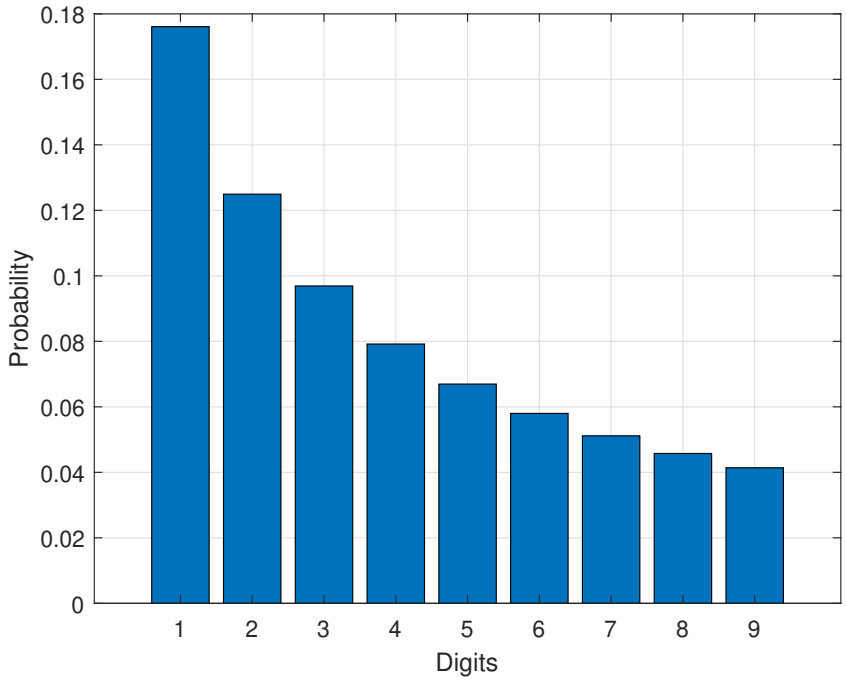

Figure 4.2: Benford's law.

Weighted-least square state estimator (WLSE): Recalling the notation used, let **x** denote the state of the system, i.e., voltage phasors at all the N system buses at a given time instant. The network topology and parameters are perfectly known and assumed to be error-free. The set of m collected measurements is specified by **z**. Each component of vector **z**, i.e. z_i, can be expressed as a nonlinear function $h_i(\mathbf{x})$, which relates the system vector **x** to the *i-th* measurement, plus an error ϵ_i because of the measurement uncertainty:

$$\mathbf{z} = \mathbf{h}(\mathbf{x}) + \boldsymbol{\epsilon} \tag{4.10}$$

where $\boldsymbol{\epsilon}$ is the vector of measurement errors. Regarding the statistical properties of each component of the measurement errors vector $\boldsymbol{\epsilon}$, a Gaussian distribution with null mean and variance σ_i^2 is usually assumed. The standard deviation σ_i of *i-th* measurement is calculated by reflecting the expected accuracy of the used meter device. In addition, measurement errors are also considered to be independent.

Under these assumptions the weighted least squares estimator (WLSE) minimizes the weighted sum of squares of the residuals r_i, defined as the difference between the measurement and its expected value $E(z_i) = \mu_i$, as:

$$min \quad J(x) = \sum_{i=1}^{m} w_{ii} r_i^2 = \sum_{i=1}^{m} w_{ii} \left(z_i - h_i(x)\right)^2 \tag{4.11}$$

Note that μ_i is equal to $h_i(x)$ by considering (4.10) and the null mean of measurement errors. In this methodology, the square of each residual r_i is weighted by $w_{ii} = 1/\sigma_i^2$, i.e., the weights are inversely proportional to the variance of the measurements.

The solution of the optimization problem (4.11) is called the *Weighted Least Squares* (WLS) estimator for x. Denoting by R the diagonal matrix constituted by the covariances of measurements σ_i^2, the optimal solution of (4.11) is reached by satisfying the first-order optimality conditions:

$$\mathbf{f}(\mathbf{x}) = \frac{\partial J}{\partial \mathbf{x}} = -\mathbf{H}^T(\mathbf{x})\mathbf{R}^{-1}(\mathbf{z} - \mathbf{h}(\mathbf{x})) = \mathbf{0} \tag{4.12}$$

where $\mathbf{H}(\mathbf{x}) = \dfrac{\partial \mathbf{h}(\mathbf{x})}{\partial \mathbf{x}}$. The Taylor's series expansion of the functions $\mathbf{f}(\mathbf{x})$ around the point $\mathbf{x}^k$ results in (4.13):

$$\mathbf{f}(\mathbf{x}) = \mathbf{f}(\mathbf{x}^k) + \mathbf{H}(\mathbf{x})(\mathbf{x} - \mathbf{x}^k) + ... = \mathbf{0} \tag{4.13}$$

Rearranging (4.13) after neglecting the second order and higher terms of the Taylor's series, the iterative solution scheme known as the Gauss-Newton method allows to obtain the system state by solving the *normal equations* at each iteration:

$$\mathbf{G}(\mathbf{x}^k)\Delta\mathbf{x}^{k+1} = \mathbf{H}^T(\mathbf{x}^k)\mathbf{R}^{-1}(\mathbf{z} - \mathbf{h}(\mathbf{x}^k)) \tag{4.14}$$

where $\mathbf{G}(\mathbf{x}^k) = \mathbf{H}^T(\mathbf{x}^k)\mathbf{R}^{-1}\mathbf{H}^T(\mathbf{x}^k)$ is called the *gain matrix*.

The iterative solution algorithm for the WLS estimator can be finally outlined as follows:

1. Set the iteration index: $k = 0$.
2. Initialize the vector x^0, usually using a flat start criterium, i.e., $U_i = 1$ and $\theta_i = 0$.
3. Calculate the gain matrix, i.e., the right hand side of (4.14).
4. Solve the sparse linear set of equations (4.14) by decomposing the gain matrix into its triangular factors and using forward/backward substitutions. A solution for $\Delta\mathbf{x}^{k+1}$ is obtained.
5. Convergence test: $max\left|\Delta\mathbf{x}^{k+1}\right| \leq \epsilon$? If no, update the state vector $\mathbf{x}^{k+1} = \mathbf{x}^k + \Delta\mathbf{x}^{k+1}$, and the iteration index $k = k + 1$, and go to step 3. If yes, stop and get the final estimation of the system state, $\hat{\mathbf{x}} = \mathbf{x}^{k+1}$.

Schweppe-Huber Generalized-M estimator: This is a robust estimator able to remain unbiased despite the existence of different types of outliers. It belongs to the *M-estimators* family, i.e., maximum likelihood estimators that minimize a function of measurement residuals subject to the constraints given by the measurement equations. M-estimators are considered robust estimators in which bad measurements are filtered out and suppressed during the iterative estimation process. Specifically, SHGME solves the following optimization problem:

$$\begin{aligned} \min \quad & \sum_{i=1}^{m} \rho\left(r_i\right) \\ \text{s.t.} \quad & \mathbf{z} = \mathbf{h}(\mathbf{x}) + \mathbf{r} \end{aligned} \tag{4.15}$$

where the function $\rho\left(r_i\right)$ is defined by:

$$\rho\left(r_i\right) = \begin{cases} \frac{1}{2}\frac{r_i^2}{\sigma_i^2} & \left|\frac{r_i}{\sigma_i}\right| \leq \lambda \\ \lambda\left|\frac{r_i}{\sigma_i}\right| - \frac{1}{2}\lambda^2 & \text{otherwise.} \end{cases} \tag{4.16}$$

Note that depending on the user-defined parameter λ, the cost function switches from a quadratic to a linear behaviour with respect to the residuals r_i, as shown in Figure 4.3. In the case of low residuals (lower than λ), the performance is similar to a WLSE since the cost function is quadratic and weighted by the corresponding standard deviations. On the contrary, for residuals larger than λ, the cost function depends on the absolute value of the residuals, turning the performance towards an absolute mean value estimator. In this latter case, and with enough redundancy, the estimator exactly satisfies a set of measurements and is able to directly discard the false ones.

The SHGME is usually solved by using an iteratively re-weighted least squares method. Indeed, writing the Karush-Kuhn-Tucker necessary conditions for a minimum of (4.15):

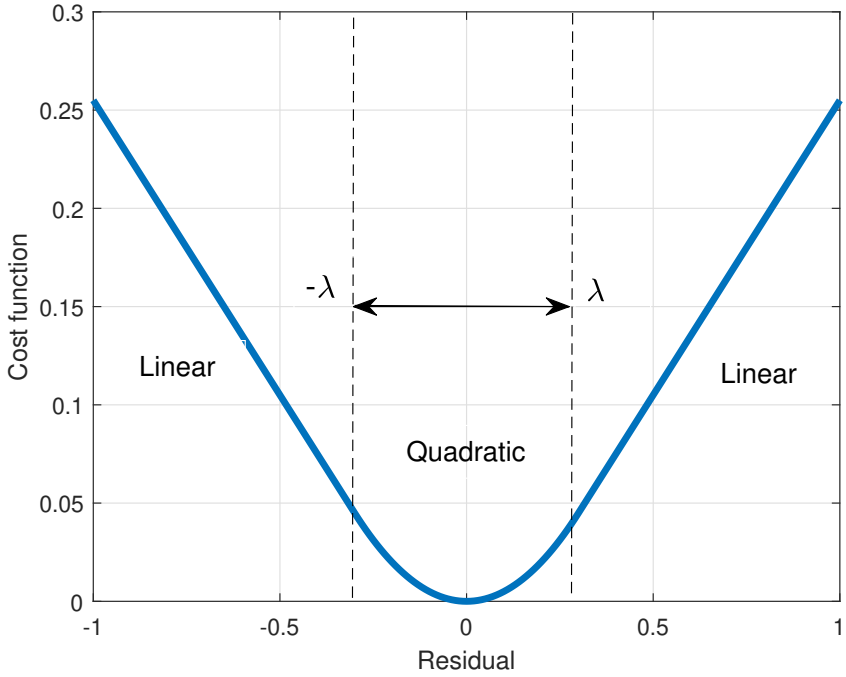

Figure 4.3: Cost function in the SHGME as a function of the λ parameter

.

$$\sum_{i=1}^{m} \frac{\partial \rho}{\partial r_i} \frac{\partial r_i}{\partial \mathbf{x}} = \mathbf{0} \rightarrow \sum_{i=1}^{m} \frac{\Upsilon\left(r_i\right)}{r_i} \cdot r_i \mathbf{H}_i = \mathbf{0} \rightarrow \sum_{i=1}^{m} \Phi\left(r_i\right) \cdot r_i \mathbf{H}_i = 0 \longrightarrow H^T \mathbf{\Phi}\left(\mathbf{z} - \mathbf{h}(\mathbf{x})\right) = \mathbf{0} \tag{4.17}$$

where $\mathbf{H}_i$ is the *i-th* row of the Jacobian matrix $\mathbf{H}$, $\Upsilon\left(r_i\right) = \dfrac{\partial \rho}{\partial r_i}$ and $\mathbf{\Phi}$ is a diagonal weight matrix whose entries $\Phi_{ii} = \dfrac{\Upsilon\left(r_i\right)}{r_i}$ are defined by:

$$\Phi_{ii} = \begin{cases} \frac{1}{\sigma_i^2} & \left|\frac{r_i}{\sigma_i}\right| \leq \lambda \\ \frac{\lambda}{r_i \sigma_i} \cdot sign\left(r_i\right) & \text{otherwise.} \end{cases} \tag{4.18}$$

When the first-order Taylor approximation for $\mathbf{h}(\mathbf{x})$ is used in (4.17), the system of equations to solve at each iteration is:

$$\mathbf{H}^T \cdot \mathbf{\Phi} \cdot \mathbf{H}\, \Delta \mathbf{x}^{k+1} = \mathbf{H}^T \cdot \mathbf{\Phi} \cdot \left(\mathbf{z} - \mathbf{h}(\mathbf{x}^k)\right) \tag{4.19}$$

Note that (4.19) are identical to the normal equations (4.14) except for the presence of the variable weight matrix $\mathbf{\Phi}$, instead of the constant weight matrix $\mathbf{R}^{-1}$. Matrix $\mathbf{\Phi}$ changes from one iteration to the next according to (4.18).

The main non-trivial issue related to SHGME lies in the selection of the λ parameter. In the Gaussian distribution, the lower the λ value, the more robust the estimator is in the presence of outliers; the higher the λ value, the more efficient the estimator is with regard to the variance of the estimates. Also, as the λ value decreases, the estimator becomes more similar to the *least absolute value estimator* [1], while as the λ value increases, the estimator becomes more similar to the WLSE.

[1] Least absolute estimator is a M-estimator where $\rho\left(r_i\right) = |r_i|$

False Data Detection
One of the most interesting uses of a state estimator is to detect, identify, and eliminate measurement errors whenever possible. This feature is highly dependent on sufficient redundancy in the measurements and on the type, location, and number of erroneous measurements. In this regard, erroneous measurements are equivalent to possible FDIAs

When using the WLSE, the detection and identification of bad data are implemented after the estimation process, by processing the measurement residuals through the *large normalized residual* (LNR) test. This test is based on using the normalized values of the residuals. It can be shown that when measurement errors ϵ_i have a Gaussian distribution defined by $N(0, R_{ii})$, the measurement residuals r_i will follow the Gaussian distribution $N(0, \Omega_{ii})$, where Ωii is the corresponding diagonal entry of the residual covariance matrix $\mathbf{\Omega}$:

$$\mathbf{\Omega} = \mathbf{R} - \mathbf{H}\mathbf{G}^{-1}\mathbf{H}^T \tag{4.20}$$

So, the normalized value of the residual for the *i-th* measurement can be obtained by:

$$r_i^N = \frac{|r_i|}{\sqrt{\Omega_{ii}}} \tag{4.21}$$

The normalized residual vector $\mathbf{r}^N$ will have a standard normal distribution, i.e., $N(0, 1)$. This implies that the largest element in $\mathbf{r}^N$ can be compared against a statistical threshold to decide on the existence of bad data or, equivalently, of an FDIA.

It can be demonstrated that if there is a single bad datum in the measurement set [2], the largest normalized residual corresponds to the erroneous measurement. This allows to define the LNR test, which is composed by the following steps:

1. Solve the WLSE and compute the elements of the measurement residual vector $\mathbf{r} = \mathbf{z} - \mathbf{h}(\hat{\mathbf{x}})$, for all the m measurements.
2. Compute the normalized residuals by using (4.20) and (4.21).
3. Locate the largest r_k^N among all the components of vector $\mathbf{r}^N$.
4. If $r_k^N > c$, the *k-th* measurement is suspected to be bad datum [3]. Otherwise, stop.
5. Eliminate the *k-th* measurement from the measurement set and go to step 1.

The main weakness of the LNR method is that it relies on the residuals, which may be strongly correlated. Therefore, in case of multiple bad data, this correlation may result in a comparable size of residuals for both good and bad measurements. In addition, it is worth mentioning that the implementation of the LNR method may require several iterations of identification and elimination if multiple measurements are attacked, which implies not only the computation of the residual covariance matrix at each cycle, but also the recomputation of the new structure of the gain matrix due to the removal of the corresponding identified bad data from the measurement set. All this implies a high computational cost, which becomes critical in real-time operation.

On the contrary, the bad data processing under an SHGME is directly incorporated into SE procedure. In fact, the variable weight matrix $\mathbf{\Phi}$ changes from one iteration to the next according to (4.18), i.e:

- If the weighting residual $\frac{|r_i|}{\sigma_i}$ of *i-th* measurement is small enough, the corresponding diagonal entry Φ_{ii} is exactly the same as the R_{ii}^{-1} entry of WLSE, which means that the *i-th* measurement is healthy.

[2] This bad datum cannot be a critical measurement or a member of a critical pair. As critical measurement is called a measurement whose elimination results in an unobservable system. A critical pair comprises two redundant measurements whose simultaneous removal also renders the system unobservable.

[3] c is a specified identification threshold. Typically, c is 3, since if the normalized residuals have a standard normal distribution, 99.7% of them should be within 3 times the standard deviations of the mean

- Conversely, if $\frac{|r_i|}{\sigma_i}$ exceeds the threshold λ (which means a high deviation of the estimated measurement with respect to the acquired one), a low value is assigned to the corresponding diagonal entry Φ_{ii}, thus reducing the influence of that *i-th* measurement on the state estimation process. The weighting residual corresponding to this measurement will continue to be calculated and checked against λ, so that they can again increase their weight in the iterative estimation process. This means that during the iterative solution of the Huber estimator, the weighting residuals of measurements can fluctuate between values corresponding to healthy or suspicious data.

Note that the computational cost of this phase of false data detection is much lower for SHGME than for WLSE, since no covariance matrix needs to be computed and the structure of the H and G matrices does not change.

False Data Injection Response
Any of the two former SE methods, WLSE or SHGME, after converging and including the application of the LNR test for the WLSE, yields the state estimate $\hat{\mathbf{x}}$. Once $\hat{\mathbf{x}}$ is known, a more accurate estimation of the measured quantities $\hat{\mathbf{z}} = h(\hat{\mathbf{x}})$ can be obtained. This set of estimates for the originally measured quantities, $\hat{\mathbf{z}}$, turns out to be more accurate than the initial set of measurements $\mathbf{z}$, regardless of whether or not there are bad data in $\mathbf{z}$. This means that the state estimation tool not only is a suitable tool for identifying and detecting FDIAs, but also provides a solution for inferring the most accurate and probable values for the attacked quantities.

The covariance matrix $\mathbf{R}_{\hat{\mathbf{z}}}$ of the estimate of the measurement vector $\hat{\mathbf{z}}$ can be computed as:

$$\mathbf{R}_{\hat{\mathbf{z}}} = \mathbf{H}\mathbf{G}^{-1}\mathbf{H}^T \tag{4.22}$$

which allows deducing the standard deviation of measurement estimates $\sigma_{\hat{z}_i} = \sqrt{R_{\hat{z}})_{ii}}$, and with it, the dispersion of the measurement estimate relative to its mean.

4.3.2 Algorithmic Implementation

This section is devoted to explain how the different components of the general FDII procedure described in the previous section interact in a given implementation according to the flowchart depicted in Figure 4.4.

First of all, the FDII method is a real-time, high-level application, executed in the CSL. Once the application is started at $t = 0$, the FDII tool will be executed in regular time intervals, $\Delta\tau$, which will depend on the latency required by the particular application and using the measurements gathered from the field devices, usually ranging from milliseconds to seconds/minutes.

The first step is the MbPA which can be conducted based on a mix of deterministic and statistical methods. Note that, for this step, input data such as device ratings and/or historical data will be required. The result of this pre-filtering stage can determine if all the measurements are normal or, on the contrary, suspicious of being intentionally corrupted. In the latter case, the set of attacked measurements is identified, with the tool moving to the FDIR step, which will be described afterwards. In the former case, it is not yet possible to guarantee that the measurements are *healthy*, since the MbPA analysis does not take into account the underlying laws of the physical EPES model. For this reason, it is required to move forward to the next step.

The second step is the SE algorithm, which can be implemented according to any of the formulations reviewed in the previous sections, either the WLSE or the SHGMSE. The state estimator will compute the most likely EPES state, e.g., complex nodal voltages, from the real-time raw measurements and considering some additional input data. The SE tool requires two inputs. First, the model of the examined EPES is required, in order to compute its mathematical representation $\mathbf{h}$. Moreover, information about the accuracy of the measurement devices is critical to adequately adjust the weights, i.e., matrix $\mathbf{R}$ in the WLSE formulation, for each one of the measurements. Once the estimated states are computed, it is possible to compute all the EPES magnitudes of interest, i.e., active and reactive power flows.

The third step refers to the FDD, the implementation of which depends on the selected state estimator. In the case of the WLSE, the false data are identified after the estimation process is completed, and specifically when the LNR test is executed. On the contrary, the Huber estimator naturally decreases the weight of the suspicious measurements, thus not introducing an additional step after the estimation process ends. In this second case,

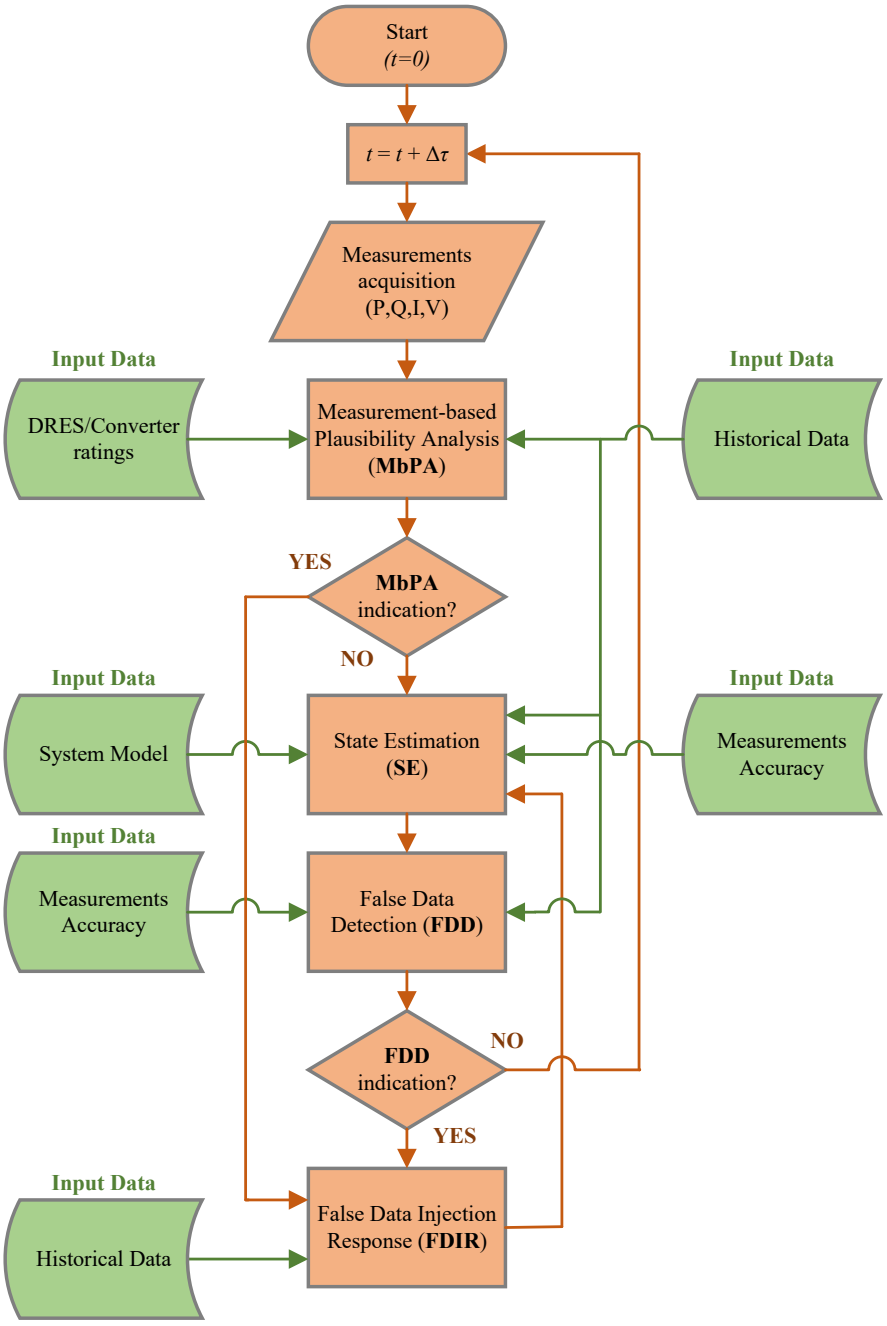

Figure 4.4: Algorithmic implementation of COCOON generic FDII methodology.

the set of suspicious measurements is defined by analyzing the final weights, particularly the lowest ones. Therefore, once the set of false measurements has been identified, it is possible to move to the fourth and final step.

Finally, the fourth step, i.e., FDI response, is activated in case an FDIA has taken place. In turn, the activation of FDIR unfolds in various steps. First, the tool informs the RES plant operator or the DSO (depending on the application, i.e., PV power parks or energy communities) that an FDIA has occurred. Moreover, the RES plant operator or the DSO is provided with a set of suspicious measurements. Those measurements are candidates of having being targeted in the frame of an FDIA. Finally, the suspicious measurements can be replaced by historical data or the estimated value derived by the state estimator, once the effect of the suspicious dataset has been removed.

In addition to the flowchart depicted in Figure 4.4, the algorithmic implementation also requires information about the software requirements of the proposed methodology. With this regard, it is expected to develop the FDI identification tool using Python (versions 3.8 - 3.11) as programming language with the use of the following libraries:

- pandas 1.4, it is a fast, powerful, flexible and easy to use open source data analysis and manipulation tool, built on top of the Python programming language (https://pandas.pydata.org/).
- numpy 1.23, it is an open source project that enables numerical computing with Python (https://numpy.org/).
- scipy 1.8, it provides algorithms for optimization, integration, interpolation, eigenvalue problems, algebraic equations, differential equations, statistics and many other classes of problems (https://scipy.org/).
- networkx 2.7, it is a Python package for the creation, manipulation, and study of the structure, dynamics, and functions of complex networks (https://networkx.org/).
- sqlalchemy 1.4, it is a SQL toolkit and Object Relational Mapper providing developers the full power and flexibility of SQL.
- seaborn 0.12, it is a data visualization library based which provides a high-level interface for drawing attractive and informative graphics.
- pyomo 6.5, it is an open-source software package that supports a diverse set of optimization capabilities for formulating, solving, and analyzing optimization models.
- GAMSpy, it acts as a bridge between the Python language and the robust GAMS system, allowing to create complex mathematical models effortlessly.

In the case of using a microservice architecture, it will be required the use of:

- nats-py 2.1, it provides a simple, open-source messaging system with a high-throughput, low-latency platform for cloud-native distributed systems.
- pangres 4.1, it updates and inserts pandas dataframes in PostgreSQL, MySQL, SQlite and other databases.
- pymongo 4.2, it is a distribution containing tools for working with MongoDB.

4.4 Integration within the COCOON Ecosystem

Two instantiations of the generic FDII methodology will be created. The first will focus on the application of FDII in Energy Communities located at balanced, MV DSs, while the second one will target at the FDII within RES power parks. Both instantiations will lie on the CSL layer which is the highest abstraction layer of the CPN architecture presented in Figure 4.5. Details regarding the CPN architecture are presented in deliverables D4.1: *COCOON Development Blueprint* and D4.2: *COCOON System Architecture*.

Using the Northbound API, both instantiations of the generic FDII methodology will interact with the two lower layers of the CPN architecture, i.e., COMML and IOL, to retrieve all the necessary information from the communication network. This mainly involves obtaining electrical measurements (voltage, current, active and powers) from predefined locations within the DS and/or RES power park. Note that IEC104 and Modbus TCP/IP will be considered the main communication protocols, as they are commonly used by the DSO and the RES park operator, respectively. The outputs of the FDII include: a) EPES states, b) an indication about the occurrence of FDIA, and c) which EPES data have been falsified. This information can be retrieved by the COCOON dashboard either regularly or on demand, thus providing the operator with a clear overview of the EPES state from a cyber security perspective.

4.5 Performance Assessment

The performance of the two instantiations of the generic FDII methodology will be thoroughly assessed by adopting a three stage procedure. The first stage is devoted to the offline evaluation through simulations on benchmark systems, e.g., CIGRE European MV grid, as well as the COCOON pilot setups for the Energy Community and PV power plant, analyzed in WP5 and WP8, respectively. By exploiting benchmark datasets and historical measurements from the pilot setups, a parametric analysis will be performed by examining different FDIA scenarios that can be classified in the following main categories: (a) measurement accessibility, (b) network knowledge, and (c) attack goal. The second stage will integrate the proposed FDII methodology into the cyber-physical laboratory setup of the University of Sevilla as part of the activities taking place in WP3.

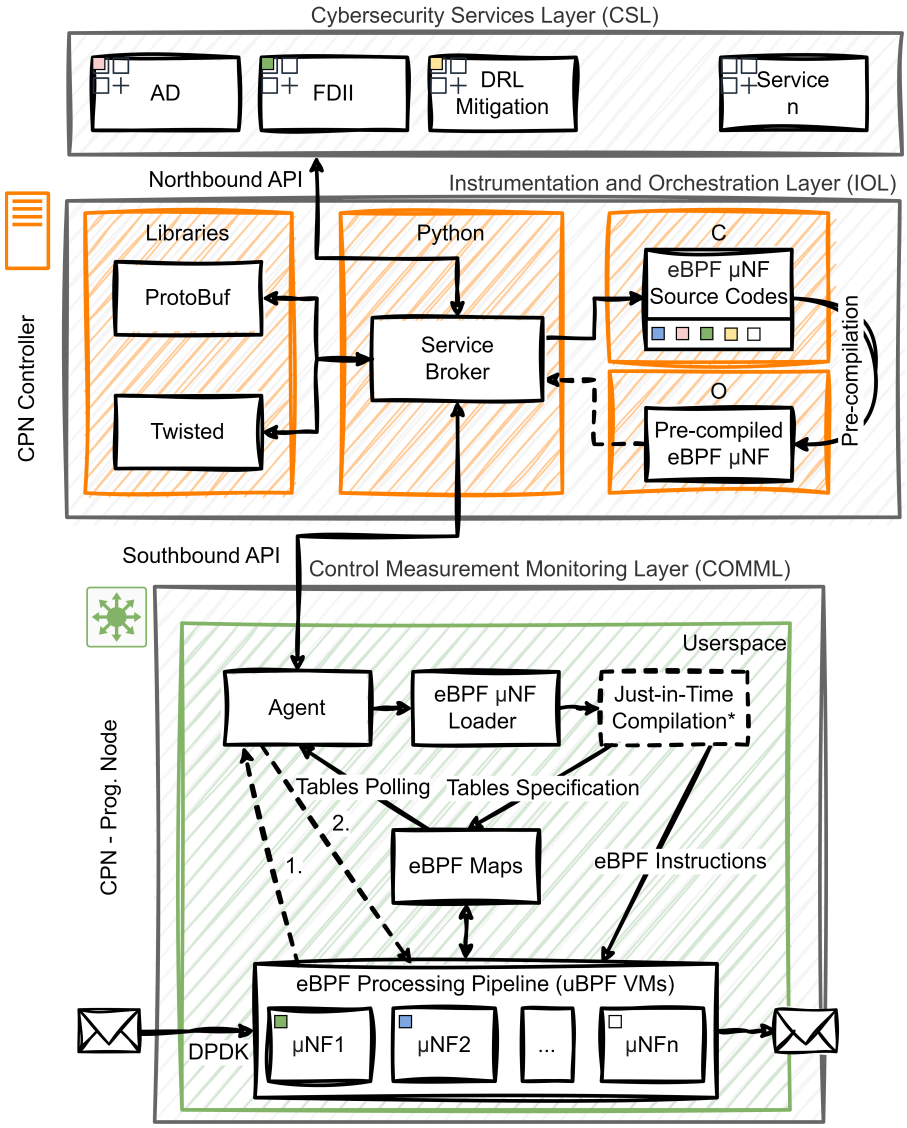

Figure 4.5: CPN Architecture.

This laboratory will mimic the OT infrastructure of the pilots mentioned previously in order to have a real-time testing as close as possible to real-world conditions. Finally, in the third stage, the performance of the generic FDII methodology will be evaluated under real field conditions in the COCOON pilot setups, using a subset of the above-examined scenarios.

4.5.1 AV Composition

Two procedures for composing AVs will be assessed. The first is a simple process where the values of the tampered measurements are determined in an uncoordinated way without considering the inherent characteristics of the physical system of EPES. The second is a more sophisticated approach, in which the physical laws of EPES are included in the determination of the values of the tampered measurements to further increase the stealthiness of the attacks. This can be achieved by solving an optimization problem that is mathematically formulated according to (4.23)-(4.26).

$$\min \sum_{i \in N_{\mathrm{m}}} (V_i^{\mathrm{a}} - V_i^{\mathrm{m}})^2 + \sum_{i \in N_{\mathrm{m}}} (P_i^{\mathrm{a}} - P_i^{\mathrm{m}})^2 + \sum_{i \in N_{\mathrm{m}}} (Q_i^{\mathrm{a}} - Q_i^{\mathrm{m}})^2 \tag{4.23}$$

s.t.

$$\bar{\mathbf{I}} = \bar{\mathbf{Y}}\bar{\mathbf{V}} \tag{4.24}$$

$$\bar{\mathbf{S}} = \bar{\mathbf{V}} \circ \bar{\mathbf{I}}^* \tag{4.25}$$

$$f\left(\bar{\mathbf{V}}, \bar{\mathbf{I}}\right) \geq \mathbf{A} \tag{4.26}$$

Here, V_i^x, P_i^x, Q_i^x denote the voltage magnitude, active and reactive power injection at node i of the grid, while superscript $x \in \{a, m\}$ indicates the tampered (a) and the actual, measured quantity (m), respectively. Furthermore, N_m is the set of attacked/tampered measurements. Scope of the objective function is to minimize the deviation between the tampered and the actual, measured quantities in order to achieve a stealthy attack. In addition, (4.24) models the power flow equations of EPES, where $\bar{\mathbf{I}}, \bar{\mathbf{V}}$ are the Nx1 vectors of nodal complex currents and voltages, respectively. Moreover, N is the set of network nodes, $\bar{\mathbf{Y}}$ is the NxN complex admittance matrix, while $\bar{\mathbf{S}}$ stands for the Nx1 vector of nodal complex power calculated based on (4.25). Finally, (4.26) is an inequality constraint that models a fictitious operating point ($\mathbf{A}$) that is properly selected to represent an emergency condition across EPES, e.g., voltage violation of a specific node, that would trigger unnecessary control actions by the DSO or the PV plant operator, which, in turn, could lead to cascading effects.

4.5.2 Evaluation Metrics

The FDII methodology will be tested through several scenarios following different procedures for the AV composition, according to subsection 4.5.1. In addition to the identification of a FDI attack, it is required to define performance metrics which consider the distinction between the actual attacked and healthy variables and those identified by the algorithm. With this regard, TP is defined as the number of attacked measurements effectively detected by the FDI algorithm while FN is the number of those attacked measurements identified as healthy ones. Similarly, TN refers to healthy measurements correctly identified and FP denotes the number of healthy ones considered as attacked. Considering these definitions, COCOON uses the most well-established performance metrics for FDII according to Annex 5:

- Accuracy. It is introduced to measure the overall correctness of the FDII by calculating the proportion of correctly predicted attacked values:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{4.27}$$

- Precision. It indicates how many of the detected attacked values are actually correct:

$$Precision = \frac{TP}{TP + FP} \tag{4.28}$$

- Recall. This metric shows how well the FDII identifies actual attacked values:

$$Recall = \frac{TP}{TP + FN} \tag{4.29}$$

- F1 score. It is a balance between precision and recall, which helps assessing the FDII performance when both false positives and false negatives are of significant importance.

$$F_1 = \frac{2 \cdot Precision \cdot Recall}{Precision + Recall} \tag{4.30}$$

5 Conclusions

This deliverable presented a comprehensive analysis of the challenges that EPES may face under FDIAs, focusing also on their effective detection and mitigation. By performing an exhaustive literature review on state-of-the-art methods, a new taxonomy was proposed to categorize FDIA and FDII techniques based on various parameters such as attack motivation, system model, type of attack, etc. By leveraging this, a generic FDII methodology is proposed that integrates physical system characteristics into the cyber security assessments, thereby enhancing the detection accuracy and EPES resilience.

Initially, the application field, assumptions and requirements of the proposed generic FDII methodology were identified. Based on the outcomes of this analysis, the final design of the FDII architecture was presented. In particular, SE is the core of the proposed generic FDII methodology, which is an iterative approach used to determine the network states of EPES, such as voltage magnitudes and angles, by combining real-time measurements with the underlying physical laws governing the system. The FDII methodology is further reinforced by three auxiliary modules: (a) MbPA, a pre-processing tool designed to validate data integrity without considering the physical system of EPES; (b) FDD, a post-processing module responsible for identifying and isolating false data; and (c) FDIR which replaces compromised/false data with accurate state estimates. The combination of these components enhances the robustness of the cyber security framework.

The interactions among the above modules have been demonstrated from an algorithmic point of view. Moreover, the integration within the COCOON ecosystem has been thoroughly analyzed and discussed. Finally, a framework for assessing the performance of the generic FDII methodology has been presented. The proposed framework involves two different scenarios for AV composition and a preliminary list of metrics for assessing the accuracy of the FDII methodology.

Future work will be carried out to develop two instantiations of the generic FDII methodology. The first will aim to address FDIAs on DSs where Energy Communities are connected, while the second will focus on PV power plants.

Annex A

X. Liu and Z. Li, ”False Data Attacks Against AC State Estimation With Incomplete Network Information,” *IEEE Transactions on Smart Grid*, vol. 8, no. 5, pp. 2239-2248, Sept. 2017, doi: 10.1109/TSG.2016.2521178

Proposal
This work focuses on showing how it is possible to build False Data Injection attacks in a specific area of a transmission network without needing all the topology and parameter information of the whole system. The main idea of the proposed model is to construct an attack vector based on AC power flow equations with only a few measurements with the boundary buses of the attacked region.

Insights

- The authors propose a methodology to construct the attack vector by knowing the topology and electrical parameters of the attacked region and the voltage and power flow measurements at the boundary of that region. Non PMU measurements are needed.
- The False data injection attack is demonstrated to be practically undetectable by AC estimators; the larger the region under attack, the more so.
- The proposed methodology allows the system operator to estimate the effort required by the attacker if he wants to be succeed, quantified in terms of the minimum information the attacker needs.

J. Zhao, L. Mili and M. Wang, ”A Generalized False Data Injection Attacks Against Power System Nonlinear State Estimator and Countermeasures,” *IEEE Transactions on Power Systems*, vol. 33, no. 5, pp. 4868-4877, Sept. 2018, doi: 10.1109/TPWRS.2018.2794468

Proposal
This work offers three contributions: 1.- derive and analyze the uncertainties for launching successful FDIAs along with their upper bounds 2.- a robust detector that checks the measurement statistical consistency using a subset of secure PMU measurements to know if a FDIA occurs 3.- If the previous step fails because of these secure PMUs are attacked, the authors state that detectability is also ensured by using alternative redundant measurements from short-term nodal synchrophasor predictions together with the robust Huber M-estimator

Insights

- The authors state that although hackers can have access to SCADA measurements of a part of the system, and even knowing the estimated state of the systems, they assume an uncertainty as a consequence of not knowing the exact topology and/or electrical parameters, or because the estimated state by the hacker is computed from an approximated point of view.
- The operator of control center can analyze how large uncertainties the hacker can assume so that a success FDIA results. This allow the system operator to quantify the level of minimal knowledge that the hacker must have about the system to be succeed
- The authors propose a robust FDIA detector that uses a set of minimal secure PMU measurements, forecasting of these PMUs, and a robust Huber M-estimator. A final binary hypothesis test on the measurement consistency allows to identify no FDIA or just the opposite, the occurrence of a FDIA.

B. Li, T. Ding, C. Huang, J. Zhao, Y. Yang and Y. Chen, ”Detecting False Data Injection Attacks Against Power System State Estimation With Fast Go-Decomposition Approach,” in *IEEE Transactions on Industrial Informatics*, vol. 15, no. 5, pp. 2892-2904, May 2019, doi: 10.1109/TII.2018.2875529

Proposal
The authors focus on detecting fault data injection attacks and recuperating the original uncorrupted set of measurements. Under a series of hypotheses, the proposed solution reformulates the FDIA detection as a matrix

separation problem. Then a new methodology to solve the matrix separation problem is presented and compared with other three well-known techniques, outperforming all the three in terms of computational efficiency and accuracy.

Insights

- The implemented solution to detect false data injection attacks assumes the attacker knows the full DC model of the system and the measurements available at each scenario, although only a limited number of the set of measurements are corrupted. It is also considered that the dynamic changes of the state of the system are slow. These two hypothesis allows to assume the data matrix composed of historical measurements has low-rank and the attack matrix is sparse, being these hypothesis essentials to apply the proposed solution.
- The attacked measurement matrix is dealt with the proposed matrix separation technique to detect if a sparse attack component has been added to the original uncorrupted measurement component. If so, both components are identified and the FDIA is cancelled.
- The proposed new methodology to solve the matrix separation problem takes into account the measurement noise, question not contemplated for the rest of the three methodologies implemented for comparative purposes.
- A thorough comparative analysis is performed proving the improvement in accuracy and saving in computational time because of the new proposed solution

---

S. K. Singh, K. Khanna, R. Bose, B. K. Panigrahi and A. Joshi, ”Joint-Transformation-Based Detection of False Data Injection Attacks in Smart Grid,” in *IEEE Transactions on Industrial Informatics*, vol. 14, no. 1, pp. 89-97, Jan. 2018, doi: 10.1109/TII.2017.2720726

Proposal

An image-processing-based technique to detect FDIA in real-time is proposed. This technique, called joint transformation, is applied on the dynamics of measurement variations, using the Kullback-Leibler distance (KLD) to find the difference between probability distributions obtained from measurements variation with and without false data, and comparing that difference with a threshold value. The methodology is tested in small transmission networks, outperforming better than equivalent solutions that work directly with probability distribution of measurement variations instead.

Insights

- The two considered image transformation techniques are the Power-Low (Gamma) and log transformation. The first one is considered to identify threshold values by working on reliable historical measurements, while the second one is used to calculate runtime distance KLD.
- The tested system is the IEEE14-bus network, studying a great variety of scenarios with different network topologies and different level of gross errors in the FDIA.
- Computational times show a fast methodology, although none large transmission network is analyzed.

K. Khanna, B. K. Panigrahi and A. Joshi, ”Priority-Based Protection Against the Malicious Data Injection Attacks on State Estimation,” in *IEEE Systems Journal*, vol. 14, no. 2, pp. 1945-1952, June 2020, doi: 10.1109/JSYST.2019.2933023

Proposal

A defense strategy against false data injection attacks is tackled in this work. The authors propose a new protection strategy by identifying sensitive measurements through the normalized measurement Jacobian of the state estimation tool. The specific protection of these sensitive measurements to any attack, ensures a secure and reliable operation of the power system. The specific case in which not only conventional measurements but also PMUs exist is also considered.

Insights

- The authors propose solving a mixed integer linear programming to select and protect the most critical measurements. A dc-model is considered.
- When the PMU are incorporated, a new optimization problem is defined and solved, resulting the minimum number of Scada measurements and PMUs to protect.
- The proposed methodology is tested in large transmission networks and compared with a previous equivalent solution based on a simplified measurement Jacobian, obtaining the same minimum number of sensitive measurements in all the cases. However, the authors show as the specific resulting solution considers the level of impact of each measurement on the state of each bus.

Z. Liu and L. Wang, "Defense Strategy Against Load Redistribution Attacks on Power Systems Considering Insider Threats," in *IEEE Transactions on Smart Grid*, vol. 12, no. 2, pp. 1529-1540, March 2021, doi: 10.1109/TSG.2020.3023426

Proposal

This paper deals with a specific FDIA: the load redistribution (LR) attack carried out by insider threats. The authors define and solve a game theory-based methodology to determine both: 1) The best strategy of the attacker to maximize the system operation cost according to the decisions of the optimal economic dispatch, and 2) The best strategy of the defender to minimize the same system operation cost. The impacts of the insider in the LR attacks can be investigated, and the proposed defense strategy is able to reduce the expected operation cost of the system under the LR attacks with the information leakage due to the insider threats.

Insights

- It is supposed through this work that critical information in cybersecurity can be exposed by the insiders to the attackers: Information leakage. Also, the focus is on LR attacks, what implies the attacker manipulates the measurements of the load bus injection and line power flow in the grid to mislead the dispatch decisions of the system operator.
- A security resource allocation game model is developed for the LR attack. Based on the proposed model, the information leakage by the insiders is formulated. The optimization models to calculate the best response strategies of both the system operator and attacker in the game considering the information leakage by the insider are developed.
- In solving the problem, it is taken into account that the attacker can assume a maximum limit to carry out the attack, and that the defender also has a limit on the maximum investment in reinforcement actions against cyberattacks.
- It is made evident by the case studies that the defense actions of the system operator is highly important to reduce the damage of the LR attack, and also that the information leakage of the system operator's strategy on the critical measurements in the grid will provide great advantages to the attacker.

N. Ahmadi, Y. Chakhchoukh and H. Ishii, "Power Systems Decomposition for Robustifying State Estimation Under Cyber Attacks," in *IEEE Transactions on Power Systems*, vol. 36, no. 3, pp. 1922-1933, May 2021, doi: 10.1109/TPWRS.2020.3026951

Proposal

This paper focuses on large transmission systems, proposing decompose the system in smaller subsystems where robust state estimators can improve its performance in relation to the detection of outliers or FDIA. The authors use a regression estimator called *Least trimmed squares estimator (LTS)* already known, but extending its application to larger systems, and being capable to define a solution that can detect cyber-attacks online.

Insights

- Because of robust estimators as LTS are able to detect more outliers when the system is decomposed in subsystems or islands (containing cyles), the focus of this work is on developing an algorithm for automatically finding islands and updating the decomposition data in real-time in a computationally ef-

ficient manner whenever the topologies change. This allows the authors to implement the detection of cyber-attacks online for real-life SSE.
- For finding the cyclic islands, the authors employ an alternative approach based on methods for detection of faces in planar graphs, referring faces to regions bounded by edges for a graph drawn on a plane. They prove this solution performs better than other well-known as depth-first search (DFS) or the minimum spanning tree search (MST).
- The authors test the proposed solution in different transmission networks, the largest with 300 buses. They make extensive comparisons and demonstrates that the number of measurement outliers detected increases in all the cases. Different scenarios of measurement redundancy are also considered.

C. Liu, R. Deng, W. He, H. Liang and W. Du, "Optimal Coding Schemes for Detecting False Data Injection Attacks in Power System State Estimation," in *IEEE Transactions on Smart Grid*, vol. 13, no. 1, pp. 738-749, Jan. 2022, doi: 10.1109/TSG.2021.310797251

Proposal

This paper focuses on meter coding as a way of defending to FDIA, investigating the optimal design of coding schemes based on the analysis of detection conditions for stealthy FDIA in power system state estimation. The authors consider the cost of meter coding in their analysis.

Insights

- This work focuses on malicious modification of measurements in wide area network with attacks launched after the coding procedure of the area measurements, what implies that hackers do not know the matrix coding.
- The authors demonstrate that to detect any FDIA, it is a necessary condition that the set of encoded measurements includes, at least, a group of essential measurements that granted observability in state estimation. For the special coding scheme that only a set of essential measurements are encoded, it is a necessary condition that none critical measurement exists in the grid.
- An optimization problem that minimizes the cost of meter encoding determines not only the minimum number of meters to encode and their location but also the coding matrix. Because of the difficulty in solving such optimization problem, the authors propose a heuristic algorithm that reduces computational times.
- Small transmission systems are analyzed under the proposed methodology, comparing results with two extreme coding schemes philosophies: the fully coding scheme and the single meter coding scheme. It is demonstrated that the cost is reduced significantly with the proposed solution, and also that the probability of detecting a FDIA is almost of a 100

J. Ruan, G. Liang, J. Zhao, J. Qiu and Z. Y. Dong, "An Inertia-Based Data Recovery Scheme for False Data Injection Attack," in *IEEE Transactions on Industrial Informatics*, vol. 18, no. 11, pp. 7814-7823, Nov. 2022, doi: 10.1109/TII.2022.3146859

Proposal

This article presents a new methodology to recover the contaminated measurements and states, after a FDIA is identified, following two steps: a first gross error filtering followed by the solution of an optimization problem. The first step is based on a data inertia effect identified by the authors, while in the second one that identified inertia effect is incorporated to the optimization problem of recovery the original state of the system. Moreover, an error criterion is proposed to assess the entire performance of the proposed recovery scheme.

Insights

- The authors define a procedure to forecast the expected values of measurements taking into account the estimated values of velocity and acceleration of measurements over a moving data window and based on the well-known inertia of grid measurements. To do so, a K-nearest neighbors regression methodology

is applied to historical data of non-corrupted measurements. This forecasting is used as a first filter after a FDIA is identified by the state estimator.
- The optimization problem tries to recovery the set of measurements and state previous to the detected attack and after the first gross filtering. This optimization problem incorporates secure and normal limits in state variables that are previously deduced using interval state estimation, a methodology previously published by the same authors.
- The methodology is tested on the IEEE 30-bus system where imperfect FDIA are generated, i.e., the attacked has not complete knowledge of the network. The proposed methodology works quite well, and it is compared with other two recovery methodologies based on using current neural networks (RNN) for predicting time series.

K. Sun, I. Esnaola, A. M. Tulino and H. Vincent Poor, "Asymptotic Learning Requirements for Stealth Attacks on Linearized State Estimation," in *IEEE Transactions on Smart Grid*, vol. 14, no. 4, pp. 3189-3200, July 2023, doi: 10.1109/TSG.2023.3236785.

Proposal

This article focus on studying the performance loss by the attacker as a result of have imperfect knowledge of the distribution of the state variables. Starting from the fact that the knowledge of the Jacobian matrix as well as the covariance matrix of the distribution of the state variables allows to deduce an optimal stealthy attack, the authors propose a way of computing such attack by using historical data of uncorrupted state variables. The performance of the proposed methodology to build FDIA is assessed on medium-size transmission networks.

Insights

- The authors characterize the learning requirements to build stealth attacks from historical data via asymptotic analysis tools from random matrix theory.
- The theoretical construction of stealthy attacks implies having a knowledge of the whole network (topology and parameters) and access to historical data of noiseless state variables.
- The methodology is tested on the IEEE 30-bus and 118-bus systems. The proposed methodology works quite well, and it is compared with other two recovery methodologies based on using current neural networks (RNN) for predicting time series.

N. Ahmadi, Y. Chakhchoukh and H. Ishii, "Analysis of Targeted Coordinated Attacks on Decomposition-Based Robust State Estimation," in *IEEE Open Access Journal of Power and Energy*, vol. 10, pp. 116-127, 2023, doi: 10.1109/OAJPE.2022.3230905.

Proposal

The authors of this work improve the methodology already proposed previously by them to detect FDIA in real-time and for large systems, decomposing the whole network in smaller subsystems and using on these islands a robust state estimator. In this new work, the whole estimation process is improved by coordinating the estimation of the subsystems with the estimation of the whole network, which leads to a better success in detecting FDIA.

Insights

- The specific proposed procedure consists of running the robust least trimmed squares state estimator (LTS) decentrally at each island level, and then centrally at the entire system level; its robustness is enhanced by the final residual analysis carried out.
- The authors analyzes the performance of the new improvements considering adversarial coordinated FDIA against certain targeted buses, both attacking power injections and the corresponding rows in the Jacobian matrix. The proposed procedure overcomes up to three different state estimators, regardless of whether they are used in centralized or decomposition-based schemes..
- The methodology is tested on the IEEE 14-bus system. The threshold of the normalized residuals test of the state estimators, if too low, worsens the performance of the analysis by detecting false attacks.

H. Pan, X. Feng, C. Na and H. Yang, "A Model for Detecting False Data Injection Attacks in Smart Grids Based on the Method Utilized for Image Coding," in *IEEE Systems Journal*, vol. 17, no. 4, pp. 6181-6191, Dec. 2023, doi: 10.1109/JSYST.2023.3287924

Proposal
This article combines three techniques, Gramian angular field (GAF) ,Markov transition field (MTF), and recurrence plot (RP) to encode the relevant features of power system time-series measurements into two-dimensional (2-D) images, obtaining the underlying data features. Then, the resulting image data feed a parallel convolutional neural networks (PCNNs) classifier which, after feature extraction, is able to identify FDIA.

Insights
- The proposed methodology is tested at a very small 3-bus transmission network with two lines and using PMUS as measurement devices.
- The authors demonstrate as the images constructed by GAF-MTP-RP combine the static and dynamic statistical information in the original time series and the internal periodicity properties better than either of them separately.
- FDIA detection is performed using the PCNN as the classifier. Compared with ordinary CNN, the PCNN improves the misclassification problem of perceptually closer images, thus effectively reducing the number of misclassified samples and ensuring the reliability and correctness of FDIA detection. Computational cost increases notably compared with ordinary CNN.

Y. Zhao, J. Liu, X. Liu, K. Yuan and T. Ding, "Enhancing the Tolerance of Voltage Regulation to Cyber Contingencies via Graph-Based Deep Reinforcement Learning," in *IEEE Transactions on Power Systems*, vol. 39, no. 2, pp. 4661-4673, March 2024, doi: 10.1109/TPWRS.2023.3319699

Proposal
This paper focuses on mitigating the impact of cyber contingencies (CCs) on the voltage control problem of distribution networks with PV distributed control resources. A graph feature representation (GFR) algorithm is proposed to specifically mitigate the impact of hybrid CCs with the graph information of the cyber-physical distribution network. GFR is then embedded into a research learning methodology (proximal policy optimization PPO) to form an improved graph-based PPO algorithm for voltage regulation.

Insights
- The cyber contingencies considered are missing data, data noise and data delay, but only in relation to magnitude and/or angle voltages. FDIA are not considered.
- The voltage regulation problem is formulated as a Markov decision process (MDP), and a reward function with reward shaping is pertinently designed to improve the model performance considering CCs. PV curtailment and reactive power injections from PVs are the control variables. A comparative analysis with other three deep reinforcement learning methods and two traditional methodologies is performed on large distribution networks, demonstrating the good performance of the new proposed solution when CCs exist; under non contingencies the proposed methodology works quite similarly as traditional solutions of the voltage problem.

A. Takiddin, M. Ismail, R. Atat, K. R. Davis and E. Serpedin, "Robust Graph Autoencoder-Based Detection of False Data Injection Attacks Against Data Poisoning in Smart Grids," in *IEEE Transactions on Artificial Intelligence*, vol. 5, no. 3, pp. 1287-1301, March 2024, doi: 10.1109/TAI.2023.3286831

Proposal
This paper develops a generalized graph neural network-based anomaly detector that is robust against FDIAs and data poisoning, without requiring additional data filtering operations or computations. It requires only

benign datasets for training and employs an autoencoder with Chebyshev graph convolutional recurrent layers with attention mechanism to capture the spatial and temporal correlations within measurement data.

Insights

- Firstly, this paper quantifies the impact of data poisoning by injecting adversarial samples into the train sets of the detectors using multiple injection levels and assess the detection performance degradation. Data poisoning implies that training data might include measurement samples that are incorrectly labeled as benign. The authors demonstrate that most machine learning-based detector previously published are highly vulnerable to data poisoning
- The proposed new technique can detect cyberattacks within unseen topologies. it also offers an unsupervised anomaly detection that necessitates only benign data for training and offers detection against totally unseen FDIAs, hence offering robustness against zero-day.
- The authors perform an exhaustive comparative analysis with many other detectors: several benchmark generalized and topology-specific detectors with different characteristics, and applying hyperparameter optimization to all of them. Small and medium-size transmission networks are tested. The proposed detector offers better behavior in all the tested cases.

A. Takiddin, H. Ibrahim, J. Kim, P. Enjeti, P. R. Kumar and L. Xie, ”Detection of Cyber Attacks in Grid-tied PV Systems Using Dynamic Watermarking,” 2022 IEEE Green Technologies Conference (GreenTech), Houston, TX, USA, 2022, pp. 57-61, doi: 10.1109/GreenTech52845.2022.9772036.

Proposal

This paper proposes an active detection scheme based on digital watermarking technique for detecting cyber attacks on a grid-tied PV system. In the proposed approach a small random signal, the secret watermark, is superimposed on to the control input to the PV inverters. The watermark signal is then shown to propagate via the switching inverter and its components and appears in all the measurement data employed to control the system. The proposed approach employs the measured data to construct a system identification model, followed by two statistical tests that are designed to accurately flag cyber-attacks.

Insights

- The work focuses on protecting PV inverters from attacks on their control. In the proposed approach a small random signal with Gaussian distribution and zero mean is superimposed on the modulating signal of the inverter pulse width modulator. The watermark signal is designed to be smaller in magnitude so that it does not interfere with the operation of the PV inverter.
- The system identification method proposed in this work is based on the data collected from the system during normal operation and after applying an auto-regressive moving average technique. This system identification model is between the modulation index and the grid current; once obtained the watermark signal is superimposed.
- Two statistical variance tests are proposed to check the integrity of the actual sensor output of PV inverter against the system identification model with and without the watermark.
- The proposed solution is tested on a 3 kW laboratory grid tied PV system. The tested cases deal successfully with harmonic injection attacks, FDI attacks, stealthy attacks and amplitude reduction attacks. It is important to tune the detection system for each location of the inverters.

G. Hug and J. A. Giampapa, ”Vulnerability Assessment of AC State Estimation With Respect to False Data Injection Cyber-Attacks,” *IEEE Transactions on Smart Grid*, vol. 3, no. 3, pp. 1362-1370, Sept. 2012, doi: 10.1109/TSG.2012.2195338.

Proposal

Scope of this work is to present a comparative analysis of ac and dc state estimations in terms of identifying false data injection cyber-attacks. In addition, the minimum number of attacked measurements is assessed/quantified

for both implementations assuming various parameters, e.g., network topology, location of measurement, etc. This could serve as a basis for configuring the attack scenarios in frame of COCOON.

Insights

- Considering that all power flows on lines and power injections are measured, the attacker needs to attack all measurements in the subgraph that is bounded by buses with power injections, in order to hide the attack.
- The sum of power flow injection changes plus changes in power losses must be equal to zero.
- The minimum number of measurements that an attacker must change to hide the attack is heavily dependent on various parameters, e.g., network topology, measurements location, etc.
- In case ac state estimation is adopted and the attacker uses a dc model for false data injection, it has a greater chance of introducing errors in the measurements, which in turn, will trigger bad data detection.
- If the attacker is aware of the detailed network configuration, then it could be able to execute an attack which would probably pass unnoticed through ac state estimation.

C. Liu, H. Liang and T. Chen, "Network Parameter Coordinated False Data Injection Attacks Against Power System AC State Estimation," *IEEE Transactions on Smart Grid*, vol. 12, no. 2, pp. 1626-1639, March 2021, doi: 10.1109/TSG.2020.3033520.

Proposal

In this work, the measurements of network parameters, e.g., phase-shift angles and turns ratio of transformers, are considered in false data injection. This way, the number of attacked measurements can be significantly reduced. This could serve as a basis for configuring the attack scenarios in frame of COCOON.

Insights

- The proposed solution can be applied to cases with incomplete network knowledge, e.g., topology, line impedance, etc.
- Line impedance can be estimated using relevant measurements. However, the success rate of the proposed method may be highly affected in case the estimated value deviates from the real one.
- The success rate of the proposed false data injection method is evaluated using ac state estimation under various key performance indicators, e.g, χ^2-detector, largest normalized residual test, etc.

S. Wei, Z. Wu, J. Xu and Q. Hu, "Multiarea Probabilistic Forecasting-Aided Interval State Estimation for FDIA Identification in Power Distribution Networks," *IEEE Transactions on Industrial Informatics*, vol. 20, no. 3, pp. 4271-4282, March 2024, doi: 10.1109/TII.2023.3321098.

Proposal

Scope of this work is to present an interval state estimation algorithm that takes into account multiple measurement uncertainties and network parameter variations. This is of particular importance in grids with limited observability, e.g., distribution grids, which are examined in frame of COCOON.

Insights

- Pseudo measurements are used to address the scarcity of meters in distribution grids. They are created by employing a probabilistic forecasting algorithm.
- The conventional WLS-based state estimation is replaced with a boundary optimization framework that is further enhanced using an iterative Krawczyk operator to avoid over-conservatism when dealing with measurement uncertainties and line parameter variations.
- Contrary to the centralized implementation architectures proposed in the literature, a distributed approach is adopted to enhance the solution accuracy and computational efficiency.
- After the intervals of each variable have been identified, any state variable that falls outside of the interval can be identified as false/bad data.

R. K. Rajasekaran, B. Natarajan and A. Pahwa, "Modified Matrix Completion-Based Detection of Stealthy Data Manipulation Attacks in Low Observable Distribution Systems," *IEEE Transactions on Smart Grid*, vol. 14, no. 6, pp. 4851-4862, Nov. 2023, doi: 10.1109/TSG.2023. 3266834.

Proposal
This paper deals with the identification of stealthy attacks in distribution systems with limited available measurements. The main characteristics involve a bad data detection algorithm based on χ^2 squared index combined with a moving target defense scheme.

Insights
- The estimation accuracy is evaluated using the well-established mean absolute percetange error (MAPE) which is defined as follows:
$$MAPE(\%) = \frac{x - \bar{x}}{\bar{x}} 100\%$$
where x and $\bar{x}$ stand for the measured and estimated value, respectively.
- The computational burden of the estimation process is reduced by introducing a linear relationship between voltages and currents at measured/unmeasured nodes.
- Bad data are identified by examining the well-established χ^2 test, considering a detection probability equal to 0.95 and degree of freedom equal to $2N$ where N denotes the number of measured nodes.
- A moving target defense strategy is proposed to identify stealthy attacks. Specifically, the network parameters are perturbed to values known only to the operator. This is attained by using D-FACTS across the various lines in the distribution grid to change the line reactance between 80% to 120% of the nominal value.

J. Zhang and X. Wang, "Low-Complexity Quickest Change Detection in Linear Systems With Unknown Time-Varying Pre- and Post-Change Distributions," *IEEE Transactions on Information Theory*, vol. 67, no. 3, pp. 1804-1824, March 2021, doi: 10.1109/TIT.2021.3049468.

Proposal
This work presents a generic framework for detecting a time-varying change in a dynamic linear regression model. This approach can be applied to identify false data injections in cases where a dc state estimation method is adopted.

Insights
- Bad/false data are identified by investigating the statistical difference on the estimated states before and after the attack.
- The generalized cumulative sum algorithm that is used for analyzing the statistic difference introduces an increased computational complexity.
- The above issue is addressed by proposing a relaxed formulation where the computational burden is linearly increased with respect to the number of measurements.
- For any given threshold employed in the proposed method, an upper bound on the expected detection delay is also provided.

C. Liu, Y. Tang, R. Deng, M. Zhou and W. Du, "Joint Meter Coding and Moving Target Defense for Detecting Stealthy False Data Injection Attacks in Power System State Estimation," *IEEE Transactions on Industrial Informatics*, vol. 20, no. 3, pp. 3371-3381, March 2024, doi: 10.1109/TII.2023.3306937.

Proposal
This work combines two solutions, i.e., meter coding and moving target defense, for identifying stealthy attacks when dc state estimation is applied.

Insights

- In the proposed meter coding method, the outputs of the sensors of each metering device are encoded with a private coding matrix before transferring through the SCADA. After the measurements are received by the SCADA, they are firstly decoded and then are forwarded to the state estimation algorithm.
- According to the proposed moving target defense strategy, D-FACTs are used to perturb line reactances across the distribution grid, which, in turn, affects matrix H in dc state estimation. Note that this information is only available to the distribution system operator.
- A heuristic searching algorithm is applied to minimize the defending cost, considering also the potential negative impact the installation of D-FACTs may have on the grid.
- The cost effectiveness (ce) of the proposed solution is evaluated by the following index:

$$ce = \frac{\frac{r}{r_{\max}} \cdot \lambda}{J + (1 - \frac{r}{r_{\max}}) \cdot \lambda}$$

where $r/r_{\max}$ is used to indicate the possibility of attack detection, while λ stands for the financial benefit of successfully defending against cyberattacks. Finally, J is the corresponding defending cost.

D. Mukherjee, "Data-Driven False Data Injection Attack: A Low-Rank Approach," *IEEE Transactions on Smart Grid*, vol. 13, no. 3, pp. 2479-2482, May 2022, doi: 10.1109/ TSG.2022.3145633.

Proposal

This work deals with the formulation of attack vectors in cases with dc state estimation techniques are applied.

Insights

- The attack vector is defined using a low-rank subspace of the mapping matrix H used in the dc state estimation.
- Go-Dec and a reduced order singular value decomposition are used to define the optimal low-rank subspace.
- It was shown that the proposed false data injection attacks can bypass the bad data detection algorithms.

W. Xu, M. Higgins, J. Wang, I. M. Jaimoukha and F. Teng, "Blending Data and Physics Against False Data Injection Attack: An Event-Triggered Moving Target Defence Approach," *IEEE Transactions on Smart Grid*, vol. 14, no. 4, pp. 3176-3188, July 2023, doi: 10.1109/ TSG.2022.3231728.

Proposal

This work combines a physics-informed attack identification method with a moving target defense algorithm to achieve high true positive ratio, low false positive ratio, less operation cost and improved interpretability.

Insights

- The proposed method consists of three main stages. In the first stage, the acquired measurements are processed in real-time through a long short-term memory autoencoder. In case an attack is identified, the second stage is triggered and the same neural network is used to approximately extract the attack vector. Afterward, the MTD is triggered to verify that the neural network has correctly identified the attack.
- The trade-off between hiddeness and efficient of the MTD algorithm is addressed by solving an optimization problem
- The following four metrics are considered: (a) average cost increase, (b) average reactance perturbation ratio due to the trigger of moving target defense (MTD), (c) attack detection probability (ADP), and defense hiddenness probability (DHP). The latter two metrics are mathematically expressed as follows:

$$ADP = \frac{\text{Number of attacks being detected}}{\text{Total number of attacks}}$$

$$DHP = \frac{\text{Number of MTDs not being detected}}{\text{Total number of MTDs}}$$

- The proposed method presents a similar performance to the robust MTD solution in terms of accurately detecting attacks (96%), while improving the hiddenness by 50%.

W. Xue and T. Wu, "Active Learning-Based XGBoost for Cyber Physical System Against Generic AC False Data Injection Attacks," *IEEE Access*, vol. 8, pp. 144575-144584, 2020, doi: 10.1109/ACCESS.2020.3014644.

Proposal

In this work, a two stage FDII method is proposed. Its distinct characteristic is the use of machine learning techniques to increase the accuracy in terms of identifying cyber-attacks in the electrical grid.

Insights

- XGBoost is used as the base classifier for attack detection in the electrical grid.
- Active learning is used during the training process of the XGBoost classifier.
- Bayesian optimization is applied to find the optimal parameters of XGBoost classifier.
- The proposed solution is validated on the IEEE 15, 57, and 118-bus systems.

G. Cheng, Y. Lin, J. Zhao and J. Yan, "A Highly Discriminative Detector Against False Data Injection Attacks in AC State Estimation," *IEEE Transactions on Smart Grid*, vol. 13, no. 3, pp. 2318-2330, May 2022, doi: 10.1109/TSG.2022.3141803.

Proposal

This work deals with improving the identification accuracy of false data injection attacks in the electrical grid. Contrary to the state-of-the-art solutions where the checking process is performed based on the statistical consistency of the measurement values, this work moves a step forward by checking the statistical consistency of the measurement residuals calculated from the ac state estimation.

Insights

- The solutions that check the statistical consistency of the measurement values are ineffective to false data injection attacks that follow the historical measurement distributions or are gradually ramped up. In addition, physical grid events may be mistakenly detected as false data injection attacks.
- The proposed method is based on the assumption that a perfect attack, i.e., full network knowledge, full measurement access, etc., can hardly be achieved under real field conditions.
- Weighted least absolute value is used as the main estimator.
- The proposed method reduces false positive rates under normal operating conditions or physical grid events, ensuring also high true positive rates during false data injection attacks.

M. Du, G. Pierrou, X. Wang and M. Kassouf, "Targeted False Data Injection Attacks Against AC State Estimation Without Network Parameters," *IEEE Transactions on Smart Grid*, vol. 12, no. 6, pp. 5349-5361, Nov. 2021, doi: 10.1109/TSG.2021.3106246.

Proposal

In this work, the authors propose a methodology for preparing false data injection attacks. The attack focuses on a specific region of the electric grid, assuming no information is available to the attacker in terms of line parameters, network topology, etc.

Insights

- The false data injection attack exploits only measurements acquired from PMUs.
- Based on the availability of the PMUs, the attack region is identified.
- A dynamic network model of the attacking region is constructed. The corresponding parameters are determined by employing the regression theorem of Ornstein-Uhlenbeck combined with weighted least squares method.

- A thorough parametric analysis is performed to assess the performance of the proposed method in terms of different: (a) attack vectors, (b) attacking regions, (c) PMU noise, (d) PMU sample rate, etc.

R. Deng, P. Zhuang and H. Liang, "False Data Injection Attacks Against State Estimation in Power Distribution Systems," *IEEE Transactions on Smart Grid*, vol. 10, no. 3, pp. 2871-2881, May 2019, doi: 10.1109/TSG.2018.2813280.

Proposal
This work investigates the conditions under which a false data injection attack will succeed in distribution grids where the system operators apply the ac state estimation technique for identifying false/bad data.

Insights
- False data injection attacks are classified into two main categories, namely global and local attacks.
- In the first category, the attacker has access to all the installed measurement units and to an approximate system state, Different approximations of the system state based either on power flow or injection measurements are assessed.
- In the second category, the attacker performs a local attack to a specific region within a distribution grid. This way, the above requirements of having full access to measurement units and a good estimate of the system are relaxed.

H. Wang *et al.*, "Deep Learning-Based Interval State Estimation of AC Smart Grids Against Sparse Cyber Attacks," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 11, pp. 4766-4778, Nov. 2018, doi: 10.1109/TII.2018.2804669.

Proposal
In this work, the conventional ac state estimation is enhanced with machine learning techniques to further improve the success rate of identifying false/bad data. Its distinct characteristic is the ability to identify sparse cyber attacks, i.e., when the operating conditions of the grid are abruptly changed.

Insights
- The proposed solution combines the deterministic state values with probability state uncertainties to estimate the bounds of the state variables.
- The deterministic state values are obtained by solving an optimal power flow problems, while the probabilistic state uncertainties are assessed by solving a dual nonlinear programming problem.
- Deep learning techniques are used as for load forecast, which, in turn, is used as an input to above modules.
- Two types of false data injection attack models are assumed, namely with complete and incomplete network information.

Y. Yuan, Z. Li and K. Ren, "Modeling Load Redistribution Attacks in Power Systems," *IEEE Transactions on Smart Grid*, vol. 2, no. 2, pp. 382-390, June 2011, doi: 10.1109/TSG.2011.2123925.

Proposal
This is one of the first works that examined the issue of false data injection attacks in power systems. The authors introduced the concept of Load Redistribution (LR) attacks, a type of attack with both technical and economic implications against transmission systems. The paper provides a comprehensive framework for systematically launching of such attacks. The attack architecture assumes full network knowledge and partial measurement accessibility, targeting exclusively load-bus power injections and line power flows. The basic principle of LR attacks is that load-bus injections and line flows are manipulated in order for: (a) the total system load not to be appear altered, thereby remaining equal to the power dispatch of the system generators; (b) the new assumed system state to be economically suboptimal. Based on the above, the attack vector is composed thanks to the solution of a bi-level optimisation problem.

Insights

- The study is not limited to presenting a new attack strategy but also examines the likely response of the TSO to such attacks and their possible corrective actions.
- In the short term, LR attacks can cause economic damage due the nonoptimal system operation, while in the long term they can even cause line overloading and load shedding.
- The authors propose the performance of statistical analysis on SE residuals in order to determine some vulnerable measurements to be protected, e.g., through the installation of PMUs at the respective buses.

L. Xie, Y. Mo and B. Sinopoli, ”Integrity Data Attacks in Power Market Operations,” *IEEE Transactions on Smart Grid*, vol. 2, no. 4, pp. 659-666, Dec. 2011, doi: 10.1109/TSG.2011.2161892.

Proposal

This is one the first works to address the issue of false data injection attacks against electricity markets. The attack model requires full topology knowledge and partial measurement accessibility, and aims to yield economic profit for the attackers through virtual market bidding. The attack vectors are defind through heuristic optimisation. Nevertheless, the strict economic and transmission-system related orientation of the study render it unsuitable for extensive consideration within the COCOON project.

Insights

- The aim of the attack model is to manipulate measurements such that system lines appear congested and thus resulting in changes in the ex-post locational marginal prices.

G. Chaojun, P. Jirutitijaroen and M. Motani, ”Detecting False Data Injection Attacks in AC State Estimation,” *IEEE Transactions on Smart Grid*, vol. 6, no. 5, pp. 2476-2483, Sept. 2015, doi: 10.1109/TSG.2015.2388545.

Proposal

The scope of this paper is to propose a new FDI technique, which will replace the conventional WLS bad data estimator. The technique is based on two probability distribution functions, q and p. Distribution q is the distribution of measurement variation from the historical data, while p is the distribution of measurement variation between the current time step and the previous time step. The authors assert that if the Kullback-Leibner distance (KLD) between q and p exceeds a user-defined, dynamic threshold, then false data have been injected. In practice, the proposed method is a time-series anomaly detection technique. The proposed technique could be incorporated into the plausibility analysis tool developed within the frame of the COCOON solution.

Insights

- As the method is purely statistical, it can potentially detect attacks launched even under full knowledge of the power grid by the attacker.
- The authors assume that power system dynamics change ”slowly” and the KLD is small under normal operating conditions.
- As the method performs statistical analysis on the variation of measurements, and not on the measurements themselves, it can also detect attacks where historical data are injected as false data.
- The selection of the threshold affects the performance of the method.
- The performance of the method deteriorates against continuous small-scale attacks or continuous replay attacks.
- The method was found not to perform satisfactorily for certain targeted quantities.

M. Jorjani, H. Seifi and A. Y. Varjani, ”A Graph Theory-Based Approach to Detect False Data Injection Attacks in Power System AC State Estimation,” *IEEE Transactions on Industrial Informatics*, vol. 17, no. 4, pp. 2465 - 2475, April 2021, doi: 10.1109/TII.2020.2999571.

Proposal

The paper proposes a new graph-based algorithm for FDII. The proposed method is intended to be used after

the execution of the conventional bad data detector. The proposed algorithm receives as input the differences between the current estimated states and the estimated states of the previous state estimation execution. Subsequently, a graph-based analysis is performed on the variations between the states, in order to detect false data. The proposed algorithm could be incorporated within the COCOON solution, as an additional layer to be executed after the WLS or Huber estimator.

Insights

- The method requires the tuning of several parameters, based on historical data.
- The method allows the detection of attacks even when the attacker has full knowledge of the power grid and full measurement accessibility.
- The authors acknowledge that the performance of their method might deteriorate in power systems with increased penetration of renewable energy sources, due to the volatile changes of the latter.

C. Konstantinou and M. Maniatakos, "A Data-Based Detection Method Against False Data Injection Attacks," *IEEE Design and Test*, vol. 37, no. 5, pp. 67 - 74, November 2019, doi: 10.1109/MDAT.2019.2952357.

Proposal

The paper proposes a time-series anomaly detection method for the identification of false data. The method is intended to be executed before the actual state estimator, and it is purely statistical, i.e., it does not incorporate the knowledge of the power system topology and parameters. The authors employ the local outlier factor (LOF) method to determine the correlation between measurements and detect false data. It is assumed that faulty measurements present a lower correlation with respect to genuine measurements. Additionally, the paper develops a forecasting tool, for the accurate replacement of removed false data with reliable pseudo-measurements. The proposed FDII technique can be incorporated into the plausibility analysis tool of COCOON, while the forecasting methods can be used for the derivation of pseudo-measurements.

Insights

- The dimensionality of the input data needs first to be reduced, for the method to perform satisfactorily.
- The method also employs the feature-bagging (FB) technique, to detect outliers in noisy datasets.
- Two distinct forecasting tools are developed, one for dynamic and one for static state estimation.
- The average percentage of attacked values per examined scenario is relatively low, only 4.17%. For such a low percentage of falsified measurements, the success of the FDII tool seems obvious.

J. J. Q. Yu, Y. Hou and V. O. K. Li, "Online False Data Injection Attack Detection With Wavelet Transform and Deep Neural Networks," *IEEE Transactions on Industrial Informatics*, vol. 14, no. 7, pp. 3271 - 3280, April 2018, doi: 10.1109/TII.2018.2825243.

Proposal

The paper proposes a method that employs deep neural network (DNN) and Discrete Wavelet Transform (DWT) techniques, to identify FDI by analysing the spatiotemporal correlation of system states. The method is intended to be used after the execution of the conventional state estimator, receiving as input the estimated states. The spatial correlation is captured by the DWT tool, while the temporal correlations are evaluated via an existing DNN. The proposed FDII technique can be incorporated into the COCOON solution, executed after the main state estimator.

Insights

- The method can be used against attacks with full knowledge of the network.
- The proposed method exhibits very high accuracy (higher than 90%), despite of the complex topologies of the systems under study (IEEE 118-bus system and IEEE 300-bus system).
- The method appears to require very high measurement resolution, e.g., with a sampling step of 33.3 ms.
- The advantage of the method is given by the fact that attack vectors neglect the temporal correlation between consecutive system states, focusing only on the spatial relation (Kirchhoff's Laws) between system

variables. The authors refer to the correlation between consecutive system states in the transient/dynamic time domain, due to power system inertia. Obviously, such dynamics to be captured require very high measurement resolutions, as stated above.

S. Wei, J. Xu, Z. Wu, Q. Hu, and X. Yu, "A False Data Injection Attack Detection Strategy for Unbalanced Distribution Networks State Estimation," *IEEE Transactions on Smart Grid*, vol. 14, no. 5, pp. 3992 - 4006, January 2023, doi: 10.1109/TSG.2023.3235945.

Proposal

The paper develops a new forecasting-aided state estimation (FASE) technique for unbalanced systems, employing a square-root unscented Kalman filter (SR-UKF). FASE techniques are techniques where the states estimated during previous time instants are used to enhance the accuracy of estimating the current states. In the paper, the FASE technique is further enhanced by projection statistics. Moreover, the authors develop a generalised likelihood ratio test (GLRT) that is applied to the estimated states for further investigation of false data. Additionally, the paper proposes a method for composing attack vectors under limited power system knowledge, targeting both one or multiple system buses. Regarding the COCOON solution, it could benefit from using a FASE implementation, especially in cases of high measurement resolution. The proposed attack vector composition logic could also be investigated.

Insights

- The paper demonstrates the trade-off between the success and the impact of FDI attacks. Success regards the inability of the FDII tool to detect it, while impact regards the achieved deviation between the falsified and the actual measurements.
- The developed method outperforms several conventional FDII criteria, e.g., LNR, Euclidean distance, and cosine similarity. The method maintains a true positive rate of above 90% for all examined cases.

A. Bhattacharjee, A. B. Mondal, A. Verma, S. Mishra, and T. Saha, "Deep Latent Space Clustering for Detection of Stealthy False Data Injection Attacks Against AC State Estimation in Power Systems," *IEEE Transactions on Smart Grid*, vol. 14, no. 3, pp. 2338 - 2351, October 2022, doi: 10.1109/TSG.2022.3216625.

Proposal

This paper proposes a self-supervised deep latent space clustering algorithm (DLSC), for the detection of stealthy FDI attacks. The proposed method is purely statistical, i.e., it does not take into account the topology and the parameters of the power system. Instead, several sophisticated AI tools are trained and fine-tuned. The main key component is a stacked autoencoder network. The attack vectors are formed according to references [17] and [44] of the manuscript. As the solution neglects the physical parameters of the power system and the training and fine-tuning processes are extensive and time-consuming, the proposed method is precluded from being exploited in the frame of COCOON.

Insights

- The model is trained using only active and reactive power measurements.
- The typical performance metrics of *Accuracy*, *Precision*, *Recall*, and *F-1* are employed.
- The method achieves values over 99% for all performance metrics.

S. Peng, Z. Zhang, R. Deng, and P. Cheng, "Localizing False Data Injection Attacks in Smart Grid: A Spectrum-based Neural Network Approach," *IEEE Transactions on Smart Grid*, vol. 14, no. 6, pp. 4827 - 4838, March 2023, doi: 10.1109/TSG.2023.3261970.

Proposal

This paper proposes a spectrum-based neural network approach to localise FDI attacks in real time. Localisation refers to accurately identifying the exact system buses or lines where measurements have been compromised. This is stressed in contrast to alternative techniques, which only aim to successfully determine the presence of an

attack, without explicitly specifying the attacked buses. The designed approach unfolds in two parts: modelling of the spectral-temporal correlations and modelling of the spectral-spatial correlations. For the first part, a short-time Fourier transform (STFT) is employed to extract the time-frequency domain representation of the multi-dimensional measurement time series, which is subsequently forwarded to a two-channel convolutional neural network (2C-CNN) for the derivation of the temporal dependencies. No knowledge of the actual power system is leveraged in this stage. The second part applies a physical knowledge-aided graph convolution network (GCN) to project the graph-structured multi-dimensional measurements into the spectral domain and capture their spatial correlations.

Insights

- The typical performance metrics of *Accuracy*, *Precision*, *Recall*, and *F-1* are employed, with the method achieving values over 98% for all performance metrics.
- An ablation study is performed, to evaluate the impact of its component/tool of the proposed method. All components proved indispensable for the proposed method. Such a study could also be performed for the FDII tool developed in the frame of COCOON.
- The method is proved superior against state-of-the-art data-driven methods.

W. Xia, D. He and L. Yu, "Locational Detection of False Data Injection Attacks in Smart Grids: A Graph Convolutional Attention Network Approach," *IEEE Transactions on Smart Grid*, vol. 11, no. 6, pp. 9324 - 9337, October 2023, doi: 10.1109/JIOT.2023.3323565.

Proposal

This paper models power systems as graphs and develops an FDII technique based on the Graph Convolutional Attention Network (GCAT), for the locational detection of false data. More specifically, power systems are modelled as connected, undirected, weighted graphs. Incorporating the graph attention mechanism into the graph convolutional layer of the spatial Graph Neural Network (GNN) the adaptive learning of the unknown parameters of the graph shift operator is facilitated. The method is performed after the execution of the main state estimator, receiving the measurements or the identified states as input. As a localisation method, the employed performance metrics refer not only to the detection of the occurrence of an attack, but also to the correct identification of the compromised buses.

Insights

- The typical performance metrics of *Accuracy*, *Precision*, *Recall*, and *F-1* are employed are employed, with the method achieving values over 90% for all performance metrics.
- The focus is drawn on the *Recall* and *F-1* metrics because, in real-world scenarios, it is more important to localise all FDI attacks, even if it implies tolerating a few false alarms.
- The performance of the method is slightly improved if the system states are used as input.
- The method is proved superior against the following state-of-the-art AI tools: Decision Tree (DT), Gaussian Naive Bayes (GNB), Support Vector Machine (SVM), Multi-Label K-Nearest Neighbor (KNN), Multi-layer Perceptron (MLP), Convolutional Neural Network (CNN), and Graph Convolutional Network (GCN).

S. Nath, I. Akingeneye, J. Wu and Z. Han, "Quickest Detection of False Data Injection Attacks in Smart Grid with Dynamic Models," *IEEE Journal of Emerging and Selected Topics in Power Electronics*, vol. 10, no. 1, pp. 1292 - 1302, August 2019, doi: 10.1109/JESTPE.2019.2936587.

Proposal

This paper proposes a dynamic, forecasting-aided state estimation algorithm, for the quickest detection of FDI attacks in power systems. The main goal of the algorithm is to correctly differentiate between sudden system changes and FDI attacks. Mathematically, the algorithm aims at minimising the worst-case detection delays (WDDs) of attacks, subject to an upper bound of the false alarm rate (FAR). The detection algorithm is based

on the cumuluative sum (CUSUM) test. Provided that the COCOON solution should reach TRL7 and address several real-world issues, including detection delays, the method proposed in this paper is of significant interest.

Insights

- The state estimation model is formulated as locally linear but globally nonlinear.
- The method identifies the exact measurements that have been corrupted, not just the occurrence of an attack.
- The optimum value of the detection threshold is estimated through a Markov-chain-based analytical model.
- The performance of the method is quantified by the trade-off between the WDD minimisation and the FAR value, as well as by the correct identification of FDI attacks.
- The method cannot distinguish between system faults and FDI attacks.

A. Sahu, Z. Mao, P. Wlazlo, H. Huang, K. Davis, A. Goulart and A. S. Zonouz, "Multi-Source Multi-Domain Data Fusion for Cyberattack Detection in Power Systems," *IEEE Access*, vol. 9, pp. 119118 - 119138, August 2019, doi: 10.1109/ACCESS.2021.3106873.

Proposal

This paper proposes a framework to improve situational awareness and allow for better detection of cyber-physical intrusions, based on a data-driven hybrid information fusion algorithm that leverages real-time data from both cyber and power-based sensors. The performance of the developed framework is tested against false data and command injection (FDI and FCI)-based man-in-the-middle attacks, on a real-world power grid which is simulated through the Power World Dynamic Studio. In this context, the authors developed a new, end-to-end data fusion engine for multiple data sources. The physical data regard measurements and breaker statuses of the power system. The cyber data refer to Wireshark raw packet captures, traffic captured from Packetbeat, as well as logs and alerts stored in Snort intrusion detection system. Moreover, several unsupervised, supervised, and semi-supervised machine learning methods are comparatively evaluated within the proposed framework. The combined usage of fused cyber-physical data could be explored in the frame of the COCOON solution.

Insights

- The main performance metric used is the *F-1* score, since it serves as the harmonic mean between *Precision* and *Recall* and thus effectively balancing the importance of false positives and false negatives.
- The paper demonstrates the superior of using fused cyber-physical data, compared to the usage of pure cyber of physical data. This improvement is reflected on a 15%-20% increase on the *F-1* score.
- The performance of semi-supervised techniques was found equivalent to that of supervised techniques.

S. Gao, H. Zhang, Z. Wang, C. Huang and H. Yan, "Data-Driven Injection Attack Strategy for Linear Cyber-Physical Systems: An Input-Output Data-Based Approach," *IEEE Transactions on Network Science and Engineering*, vol. 10, no. 6, pp. 4082 - 4095, July 2023, doi: 10.1109/TNSE.2023.3292403.

Proposal

This paper proposes a general methodology for devising injection attacks against linear cyber-physical systems with repetitive movements, assuming no knowledge of the system. The attacker is only assumed to have access on input and output data of the system. These data are leveraged by iterative learning techniques to generate a virtual model of the system to be attacked. Subsequently, the FDI attack is devised via the solution of an optimisation problem, which aims to maximise the deviation of measurements and minimise the required energy consumption for launching the attack. As the power system is a nonlinear system, the framework proposed in this paper is irrelevant to the scope of the COCOON solution.

Insights

- The paper exposes the vulnerability of power systems to cyber attacks, since it highlights that intruders can launch successful attacks even with restricted data access..

G. Cheng, Y. Lin, J. Yan, J. Zhao and L. Bai, "Model-Measurement Data Integrity Attacks," *IEEE Transactions on Smart Grid*, vol. 14, no. 6, pp. 4741 - 4757, March 2023, doi: 10.1109/TSG.2023.3253781.

Proposal

This paper proposes a framework for coordinated false parameter and false data injection attacks. False parameter attacks are attacks in which intruders gain access to the database of system operators and falsify the parameter values of several power system components. Such attacks unfold in two stages: Firstly, the *pre-attack* stage determines false parameter vectors and the set of measurement channels to be manipulated offline, and the *run-time-attack* stage determines the false measurement vectors for each measurement snapshot online. Such coordinated false parameter and data injection attacks could be tested in order to investigate the limitations of the COCOON solution.

Insights

- The paper provides insights into the potential planning and reconnaissance performed by the attackers prior to the actual attack.
- The authors develop a generic attack model that covers all types of network parameters and measurements is proposed. It is observed that by strategic injection of false parameters into the model database, the attackers can drastically reduce the number of measurement channels that needs to be compromised, thus significantly facilitating the online FDI attack.
- The primary objectives are to ensure the stealthiness of the attack and to minimise the number of measurement channels to compromise, i.e., maximise *sparsity*.
- The number of falsified system parameters is also to be minimised, yet it is of secondary cruciality. The reason is that once the cyber adversaries access the system parameter database, increasing the number of falsified parameters does not cost as much as increasing the number of compromised measurement channels.

Y. Li, Y. Wang and S. Yu, "Online Generative Adversary Network Based Measurement Recovery in False Data Injection Attacks: A Cyber-Physical Approach," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 3, pp. 2031 - 2043, June 2019, doi: 10.1109/TII.2019. 2921106.

Proposal

This paper develops a new defense methodology to maintain uninterrupted state estimation when the system is under FDI attacks. Specifically, the proposed method aims at successfully recovering the actual measurement values from the tampered data, in order to restore the execution of state estimation. Therefore, the proposed method is activated after the execution of the state estimator and the indication of false data. Moreover, the method requires a perfect localisation/flagging of the falsified measurements in order to perform satisfactorily. The assumptions of the method, as well as the the computationally- and time-consuming process of the involdved neural network training render the proposed method overly theoretical for the purposes of the COCOON project.

Insights

- The main state estimation which determines the compromised measurements is designed according to reference [16].
- The unaffected measurements, as well as the number of affected measurements are forwarded into a generative adversarial network (GAN), to retrieve the original measurements.
- The authors present an enhanced version of the original GAN, which ensures improved convergence.
- The parameters of the GAN have to be updated online, in order to ensure high performance.
- The proposed technique can effectively ensure the seamless execution state estimation by generating data sufficiently close to the measurements before tampering, with an error lower than $1.5e^{-5}$ and $2e^{-2}$ p.u. for voltage amplitude and phase angle, respectively.

- The authors demonstrate that the performance of the developed technique under the proposed GAN model is superior to that under the conventional GAN model.

N. N. Tran, H. R. Pota, Q. N. Tran and J. Hu, "Designing Constraint-Based False Data-Injection Attacks Against the Unbalanced Distribution Smart Grids," *IEEE Internet of Things Journal*, vol. 8, no. 11, pp. 9422 - 9435, June 2021, doi: 10.1109/JIOT.2021.3056649.

Proposal

This paper develops an attack vector design strategy for unbalanced distribution systems. The main assumption is that the adversary successfully hijacks the SCADA system of the DSO. Thus, it gains access of the topology, the status of switches and breakers and the measurement values, while it is also capable of falsifying any targeted measurement. The attack is launched in two stages, by firstly determining the attack region and subsequently the attack vector. Optimization problems are formulated and solved for the determination of the attack region and vector. Although the assumptions regarding the data access and knowledge of the attacker are deemed unrealistic, the proposed attack strategies could be replicated in order to investigate the limitations of the COCOON solution, and especially those of the developed plausibility analysis tool.

Insights

- The size of the attack area should be minimal in order to also minimise the required resources for launching the attack and avoiding attracting the attention of the DSO.
- Regarding the formation of the attack vector, the paper abides by the principles discussed in [56].

B. Lei, G. Xiao, R. Lu, R. Deng and H. Bao, "On Feasibility and Limitations of Detecting False Data Injection Attacks on Power Grid State Estimation Using D-FACTS Devices," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 2, pp. 854 - 864, February 2020, doi: 10.1109/TII.2019.2922215.

Proposal

This paper proposes a framework for the detection of FDI attacks using the moving-target defence (MTD) approach, and specifically by deploying distributed flexible ac transmission systems (D-FACTS). More precisely, under the assumption that adversaries have a certain degree of knowledge of the power system topology and parameters, D-FACTS devices are deployed in order to partially modify the electrical parameters of the power system, thus negating the knowledge of the adversaries and allowing for the successful detection of the attacks. The proposed D-FACTS deployment framework is used in conjunction with a standard BDD and state estimator. The framework is evaluated under the following types of FDI attacks: coordinated single-bus attacks uncoordinated multiple-bus attacks, and coordinated attacks on multiple, interconnected buses. The authors adopt a dc modeling of the power system for both attack and defence purposes. Provided that the COCOON solution does not envision the usage of D-FACTS against FDI attacks, the proposed methodology is irrelevant to the scope of the project.

Insights

- The paper provides definitions regarding the *effectiveness* of the attack. Specifically, the authors consider an attack as *ineffective*, if the deviation between the actual and the falsified measurements lies within the tolerance threshold of the system state errors/faults.
- The authors prove the following theorem: The proposed approach is successful in detecting all three types of FDI attacks, but only on buses with a degree larger than 1. Elaborating, the authors define as degree of a bus the number of branches connecting it to other buses. Moreover, for the detection to be successful, D-FACTS devices must cover branches that contain at least one spanning tree of the grid graph. Consequently, the developed method is not suitable for distribution systems, as the latter are mostly operated as radial, thus containing several buses with a degree of 1.

M. Higgins, F. Teng and T. Parisini, "Stealthy MTD Against Unsupervised Learning-Based Blind FDI Attacks in Power Systems," *IEEE Transactions on Information Forensics and Security*, vol. 16, pp. 1275 - 1287,

September 2020, doi: 10.1109/TIFS.2020.3027148.

Proposal

This paper examines the issue of MTD against FDI attacks, focusing on both the adversary and defender perspective. Regarding the adversary side, the authors propose a counter-MTD framework for FDI attack launching. The attacks are devised as blind, implying that no knowledge of the power system topology and parameters is required. From the point of view of the defending system operator, the paper introduces a new MTD strategy. The assumption for launching blind attacks is that adversaries have full measurement access over a sufficiently long time period, so as to store an adequate amount of historical measurements. Although the adoption of MTD techniques is not envisioned within the frame of the COCOON solution, the insights concerning the preparation of blind attacks are valuable.

Insights

- The three main stages for the execution of a successful blind attack against MTD are as follows: *(a)* Dimensionality reduction of the utilised data set, e.g., via T-distributed stochastic neighbour embedding (T-SNE); *(b)* Application of unsupervised learning techniques on the historical data to identify the subset of measurements corresponding to the current distribution system, e.g, via density based spatial clustering of application with noise (DBSCAN); *(c)* attack vector composition using independent component analysis (ICA). The dimensionality reduction is essential for simplifying the complex measurement datasets and enabling the online execution of attacks.
- The defence strategy involves the usage of D-FACTS for MTD. However, the introduced parameter changes are strategically selected in order to to be sufficiently small and indistinguishable from the effect of AWGN on power systems. This technique is called Gaussian, physical watermarking and safeguards the power system from the above-described blind attacks.
- The proposed deployment of D-FACTS results in minimal change in the power system operational state, and thus minimal increase in cost.

F. Ye, X. Cao, Z. Cheng and M.-Y. Chow, ”CASL: A Novel Collusion Attack Against Distributed Energy Management Systems,” *IEEE Transactions on Smart Grid*, vol. 14, no. 6, pp. 4717 - 4728, November 2023, doi: 10.1109/TSG.2023.3251401.

Proposal

This paper proposes an FDI attack scheme against microgrids, based on the malicious coordination between an energy storage system (ESS) and a load. More specifically, the authors designed two secret functions, only known by the colluding ESS-load pair, to enable the transmission of secret information without the need for additional communication infrastructure. The attack is tailored for microgrids, whose assets are operated and controlled according to optimal power flow and economic dispatch schemes. The attack is economically motivated, aiming to disrupt the optimal economic operation of the microgrid and incur financial losses. Nevertheless, the attack strategy is tailored for microgrids without state estimation functionalities, where the detection of FDI is only implemented by means of cross-checking the assets setpoints. This cross-checking requires the broadcast of the economic dispatch setpoints to all distributed assets, which is not applied in distribution systems. Consequently, the proposed attack scheme cannot be exploited in the frame of the COCOON project for the protection of complex distribution systems.

Insights

- Despite of the above-mentioned inapt assumptions, the attack strategy can provide insights for the design of sophisticated attacks under ancillary-services provision schemes.
- The proposed attack scheme can achieve up to 15% economic losses.

H. Wang, J. Ruan, B. Zhou, C. Li, Q. Yu, M. Q. Raza and G.-Z. Cao, ”Dynamic Data Injection Attack Detection of Cyber Physical Power Systems With Uncertainties,” *IEEE Transactions on Smart Grid*, vol. 15, no. 10, pp. 5505 - 5518, October 2019, doi: 10.1109/TII.2019. 2902163.

Proposal

The scope of this paper is twofold: *(a)* to propose a dynamic FDI attack over multiple time snapshots; *(b)* to develop a framework for FDII against such dynamic attacks. Both the attack and the defence strategies assume that an OPF is conducted for the determination of the operating point of the system. Moreover, regarding the attack, in contrast with the majority of available literature where most developed attack models are static and refer only to a specific operational snapshot, the proposed attack framework unfolds over multiple time snapshots. The attack vectors are derived by solving an optimisation problem, whose objective function is the minimisation of the compromised meters. From the perspective of the SO, the authors leverage load consumption and renewable generation historical data and forecasts to determine the optimal operational point of the power system. Subsequently, based on the same forecasts, as well as on the assumed noise levels and network parameter discrepancies, the maximum expected deviation of the state variables from this optimal point is estimated. Any variation exceeding the maximum expected deviation threshold is deemed a potential injected false data. The statistical analysis is conducted through kernel quantile regression. Therefore, the defence strategy is purely statistical and assumes that the actual operating point is the result of an OPF. In the frame of COCOON, this only applies to the cases where ancillary services are provide. Moreover, the attack assumptions are quite unrealistic to be considered in the frame of the COCOON solution. Nevertheless, the defence mechanisms proposed in COCOON could benefit from the statistical analysis tools presented in this paper.

Insights

- The attack aims to gradually falsify the system states in order for two tie lines to eventually appear as overloaded.
- The attacking assumption is that the adversary has full knowledge and measurement accessibility for a specific attack region. This knowledge includes the operational limits of the generators within the attack area, as well as historical load profiles and renewable generation profiles. Essentially, the attacker must reconstruct the OPF problem that the system operator is supposed to solve.
- The proposed attack strategy does not introduce residual increases in the tampered measurements, thus evading detection from the standard residual-test and chi-square BDD.

A. Shefaei, M. Mohammadpourfard, B. Mohammadi-ivatloo, and Y. Weng, "Revealing a New Vulnerability of Distributed State Estimation: A Data Integrity Attack and an Unsupervised Detection Algorithm," *IEEE Transactions on Control of Network Systems*, vol. 9, no. 2, pp. 706 - 718, June 2022, doi: 10.1109/TCNS.2021.3091631.

Proposal

This paper focuses on power systems where state estimation is performed in a distributed fashion. More specifically, in distributed SE (DSE), the power system is partitioned in areas and SE is performed locally, for each area. Subsequently, the states of boundary buses are shared between neighbouring areas. In this context, the paper proposes an attack model against DSE, where only boundary buses are targeted. The attack strategy requires full system knowledge and measurement accessibility. The attack vectors are tailored to circumvent the DSE method of reference [5] and the convergence-based detector of reference [11]. The attack vectors are derived from an optimisation problem whose scope is to minimise the per area errors between neighbouring areas. Regarding FDII, the authors exploit the different statistical behaviour of estimated states, with and without the presence of false data. More specifically, the article proposes the use of kernel density estimation to calculate the statistical distribution of estimated states and infer FDI attacks. Provided that DSE is not envisioned in the frame of COCOON, the proposed attack and defence models are not directly applicable.

Insights

- The optimisation problem for the composition of attack vectors is formulated as mixed-integer second-order cone programming.
- Assuming that modern smart grids might be subject to frequent topological changes or that the actual system parameters might not be precisely known for distribution systems, unsupervised ML techniques

should be preferred for FDII.
- To infer the presence of false data by analysing the probability density function of the estimated states, the employed criteria are the skewness and the kurtosis. These criteria are chosen according to Pearson correlation-based feature selection.
- The usage of unsupervised learning allows the detection of false data even under contingencies, which imply topological changes.

S. Chakrabarty and B. Sikdar, "Unified Detection of Attacks Involving Injection of False Control Commands and Measurements in Transmission Systems of Smart Grids," *IEEE Transactions on Smart Grid*, vol. 13, no. 2, pp. 1598 - 1610, March 2022, doi: 10.1109/TSG.2021.3137835.

Proposal
This paper proposes an algorithm for unified detection of FCI and FDI attacks in transmission systems. The detection is based on the following observations on the measurement covariance matrix: *(a)* the trace of the measurement covariance matrix is greater under FCI/FDI attacks; *(b)* the absolute values of non-diagonal elements in the covariance matrix increase under FCI/FDI attack; *(c)* the covariance matrix under FCI/FDI has eigenvectors that are not standard basis vectors. The above observations/prepositions are proved. Each of the above prepositions are mathematically quantified through ad-hoc indices. The sum of those indices, denoted as β, serves as the detection metric of the proposed strategy. The presented covariance matrix analysis is straightforward and it could be easily incorporated in the plausibility analysis tool of the COCOON solution.

Insights
- The attack vectors for FDI attacks are devised according to [56].
- The performance of the method was found comparable to that of [108] and [111].
- The developed detection algorithm is non-iterative and does not require historical datasets.
- The detection of attacks requires the selection of an appropriate threshold value for the detection index β. This threshold value has to be fine-tuned.

S. Yang, K.-W. Lao. Y. Chen and H. Hui, "Resilient Distributed Control against False Data Injection Attacks for Demand Response," *IEEE Transactions on Power Systems*, vol. 13, no. 2, pp. 2837 - 2853, June 2023, doi: 10.1109/TPWRS.2023.3287205.

Proposal
This paper discusses the issue of FDI attacks against demand response systems (DRS) for building applications, which are equipped with heating, ventilation, and air conditioning (HVAC) units. More specifically, the authors showcase the potential impact of FDI attacks on such systems and develop a defence strategy. The defence strategy regards the design of an attack-resilient DRS contoller. To elaborate, the authors argue that the concepts of state estimation and BDD cannot be effectively applied to DRS due to the lack of critical parameter information, which impedes the mathematically accurate correlation of the involved physical quantities. As a result, the proposed defence strategy regards the design of a DRS controller that could seamlessly dictate the correct operational setpoints for the HVAC assets, even under the presence of attacks, and prevent discomfort for building occupants. Evidently, both the scope and the results of this paper are irrelevant to the COCOON project, as they pertain neither to distribution systems nor to FDII.

A. Khaleghi, M. S. Ghazizadeh and M. R. Aghamohammadi, "A Deep Learning-Based Attack Detection Mechanism Against Potential Cascading Failure Induced by Load Redistribution Attacks," *IEEE Transactions on Smart Grid*, vol. 14, no. 6, pp. 4772-4783, Nov. 2023, doi: 10.1109/TSG.2023.3256480.

Proposal
This paper presents a mechanism for identifying load redistribution (LR) attacks that, besides causing overflow on lines, have the potential to generate cascading failure.

Insights

- Three states can be determined: (i) potential cascading failures by LR attack; (ii) LR attacks resulting in the overflow and increasing the operating costs; and (iii) the normal state of the system.
- It is presented a mechanism to identify LR attacks based on a deep learning algorithm.
- Attacker tries to maximize the power flow of the transmission assets by falsifying the load and power flow measurements.
- The Initiating Contingency (IC) is defined as ”a combination of outages or an outage that occurs in a short period of time such that corrective action is not possible before the next case occurs.”
- It is critical to determine the IC: (i) determine the attacks in which the attacker has the ability to maximize power flow of some lines above their threshold; (ii) find the branches whose outages can cause severe overload; (iii) worst IC is the number of formed islands after line outages.

Z. Kazemi, A. Safavi, F. Naseri, L. Urbas, P. Setoodeh, “A Secure Hybrid Dynamic-State Estimation Approach for Power Systems Under False Data Injection Attacks,” *IEEE Transactions on Industrial Informatics*, vol. 16, no. 12, pp. 7275-7286, Dec. 2020, doi: 10.1109/TII.2020. 2972809.

Proposal
It is proposed an effective secure hybrid dynamic-state estimation approach that involves a dynamic model of the attack vector.

Insights
- Formulated for dynamic State Estimator.
- Initial estimation of the system states using a designed Unknown Input Observer (UIO).
- Dynamics of the attack vector were obtained using the state estimations of the first stage.
- Kalman filter is used for coestimation of the attack and the system states.
- Assumes that PMUs are installed at generator buses.
- Uses IEEE 14 and 57-bus benchmark networks for validation purposes.

H. M. Albunashee et al., “A Test Bed for Detecting False Data Injection Attacks in Systems With Distributed Energy Resources,” *IEEE Journal of Emerging and Selected Topics in Power Electronics*, vol. 10, no. 1, pp. 1303-1315, Feb. 2022, doi: 10.1109/JESTPE.2019.2948216.

Proposal
A detection algorithm based on a Kalman filter to detect a type of phase-locked loop (PLL) attack, in which the PLL reports false phase angle and frequency values to the distributed energy resources (DER) controllers.

Insights
- Test bed combination of: simulation models of utility-scale DER and distribution systems using OPAL-RT, IEDs, network devices and LabVIEW for monitoring purposes.
- Weighted Least Square State Estimator is implemented.
- A new class of attack that causes PLL reporting of false phase angle and frequency values to the controllers.
- Kalman Filter: The attacks are categorized as either
  - parametric: know the system, without violating limits,
  - nonparametric: communication.

Z.-H. Yu, W.-L. Chin, “Blind False Data Injection Attack Using PCA Approximation Method in Smart Grid,” *IEEE Transactions on Smart Grid*, vol. 6, no. 3, pp. 1219-1226, May 2015, doi: 10.1109/TSG.2014.2382714.

Proposal
It studies the general problem of blind false data injection attacks using the principal component analysis approximation (PCA) method without the knowledge of Jacobian matrix and the assumption regarding the distribution of state variables.

Insights

- Calculation of eigenvalues for the PCA
- Attack vector generated by the PCA matrix.
- Can use both DC and AC power flow models.
- The attack vector generated by the PCA matrix is proven to be approximately stealthy.
- The method is tested on the IEEE 14-Bus model.

S. Lakshminarayana, A. Kammoun, M. Debbah, H.V. Poor, "Data-Driven False Data Injection Attacks Against Power Grids: A Random Matrix Approach," *IEEE Transactions on Smart Grid*, vol. 12, no. 1, pp. 635-646, Jan. 2021, doi: 10.1109/TSG.2020.3011391.

Proposal

It is proposed an algorithm, guided by results from random matrix theory (RMT), to construct FDI attack vectors in the face of limited measurements that can nevertheless bypass the Bad Data Detection with high probability.

Insights

- The attacker is assumed to have access to only power grid measurement data traces collected over a limited period of time and no other prior knowledge about the grid.
- Trade-off between the FDI attack's BDD-bypass probability and the number of power meters in the grid.
- Note that the problem at hand is equivalent to estimating the principal eigenvalues/vectors.
- Under a limited measurement period setting, a key issue is that only a few eigenmodes can be reliably estimated from the sample covariance matrix.

H. Goyel, K.S. Swarup, "Data Integrity Attack Detection Using Ensemble-Based Learning for Cyber-Physical Power Systems," *IEEE Transactions on Smart Grid*, vol. 14, no. 2, pp. 1198-1209, March 2023, doi: 10.1109/TSG.2022.3199305.

Proposal

It is proposed a technique to generate and detect data integrity attacks in smart grids, and an optimization algorithm for generating FDIA against state estimation algorithms present at the control center. It also proposes a combining technique for the voting-based ensemble learning technique (MVCC) to detect FDIA.

Insights

- Formulation for generating AC state estimation attack with full information.
- Formulation for generating DC state estimation attacks with limited information.
- The objective function maximizes the change in meter readings or system states of the attacked bus while considering a trade-off factor (t)
- For a given training dataset S with cardinality N, it trains T classifiers independent of each other.

R. Deng, G. Xiao, R. Lu, "Defending Against False Data Injection Attacks on Power System State Estimation", *IEEE Transactions on Industrial Informatics*, 13, 1-1, 2015, doi: 10.1109/ TII.2015.2470218.

Proposal

This paper designs the least-budget defense strategy to protect power systems against FDI attacks and investigates which meters should be protected, determining how much defense budget has to be deployed on each of these meters.

Insights

- The meter selection problem is formulated as a mixed integer nonlinear programming (MINLP) problem, which can be efficiently tackled by Benders' decomposition.
- The defender-attacker interaction is investigated.
- The attacker's objective is to launch an attack at the least cost considering the following capabilities:
  - Knowledge of the power network topology and configuration of the power system, i.e., the H matrix;

  - The ability to access any set of meter measurements simultaneously.
- The attack/defense scheme is built on the measurement residual-based estimator.

Z. Wang, H. He, Z. Wan, Y. Sun, "Detection of False Data Injection Attacks in AC State Estimation Using Phasor Measurements", *IEEE Transactions on Smart Grid*, doi: 10.1109/ TSG.2020.2972781.

Proposal

It investigates the detection of False Data Injection Attacks (FDIAs) in AC state estimation where attackers cannot obtain the accurate transmission line admittances and the accurate estimated system state variables. Sufficient conditions for launching FDIAs without being detected by the traditional residual-based bad data detector are derived. A robust detection method using secure PMU measured data is proposed to detect these FDIAs effectively.

Insights

- It obtains the estimated states using PMU measurements and the Bisquare estimator.
- Then, the expected SCADA measurements are calculated based on these estimated state variables.
- Finally, the statistical consistency between the received SCADA measurements and the expected SCADA measurements is utilized to detect FDIAs.
- Simulations are performed on the IEEE 30-bus, IEEE 118-bus, and IEEE 300-bus systems.

P. Zhuang, R. Deng, H. Liang, "False Data Injection Attacks Against State Estimation in Multiphase and Unbalanced Smart Distribution Systems," *IEEE Transactions on Smart Grid*, vol. 10, no. 6, pp. 6000-6013, Nov. 2019, doi: 10.1109/TSG.2019.2895306.

Proposal

It investigates the vulnerability of DSSE to FDI attacks. Proposes a local state-based linear Distribution System State Estimator (DSSE) for multiphase and unbalanced smart distribution systems, which can facilitate the construction of FDI attacks numerically with the least information of system states. Also, the construction of three-phase coupled FDI attacks is introduced.

Insights

- For strong coupling among phases, the probabilities of successful attacks using the proposed three-phase decoupled FDI attack are derived numerically.
- The perfect three-phase decoupled FDI attacks, which consider the weak couplings among phases, are investigated.
- The FDI attacks based on the proposed linear DSSE only need the corresponding local states of the compromised power measurements to succeed against the original nonlinear DSSE.
- The FDI attacks based directly on the original nonlinear DSSE require the information of the entire system states.

Y. Liu, O. Ardakanian, I. Nikolaidis, H. Liang, "False Data Injection Attacks on Smart Grid Voltage Regulation With Stochastic Communication Model," *IEEE Transactions on Industrial Informatics*, vol. 19, no. 5, pp. 7122-7132, May 2023, doi: 10.1109/TII.2022.3209287.

Proposal

This work proposes a novel false data injection attack (FDIA) against the Voltage Regulation (VR) capacity estimation process that exploits the uncertainty in EV mobility and network conditions. It shows the attack vector with the largest expected adverse impact is the solution of a stochastic optimization problem, subject to a constraint that ensures it bypasses bad data detection. The attack vector is determined by solving a sequence of convex quadratically constrained linear programs.

Insights

- The attacker knows the distribution system model, has access to real-time and historical PMU and smart meter data, and the number of EVs reported by each EVCS.
- Stochastic Process for Characterizing the Number of EVs.
- Modified DSSE with VR variables.
- Proposes a FDIA against this VR scheme, which considers potential delays and packet losses in the communication network and the stochastic mobility pattern and charging demand of an EV fleet, to maximize its expected adverse impact on the distribution system over time.

J. Zhao, G. Zhang, Z.Y. Dong and K.P. Wong, "Forecasting-Aided Imperfect False Data Injection Attacks Against Power System Nonlinear State Estimation," *IEEE Transactions on Smart Grid*, vol. 7, no. 1, pp. 6-8, Jan. 2016, doi: 10.1109/TSG.2015.2490603.

Proposal
This work presents an imperfect FDIA model and its corresponding forecasting-aided implementation method against the nonlinear power system state estimation by introducing an attack vector relaxing error.

Insights
- Assumes that the historical measurements, the system topology and parameters for estimating the target state variables are known to the hacker, but not real-time measurements, only historical data.
- Defines an upper bound for FDIA not being detected by the residual-based bad data detection method.
- If the forecasting error is large, the attack magnitude will be greatly restricted by this bound. Otherwise, the attack magnitude can be relatively large due to a small error of state forecasting.

A.D. Syrmakesis, H.H. Alhelou and N.D. Hatziargyriou, "Novel SMO-Based Detection and Isolation of False Data Injection Attacks Against Frequency Control Systems," *IEEE Transactions on Power Systems*, vol. 39, no. 1, pp. 1434-1446, Jan. 2024, doi: 10.1109/TPWRS.2023.3242015.

Proposal
This paper proposes a novel detection and isolation method of False Data Injection Attacks (FDIAs) against Load Frequency Control (LFC). It is able to successfully distinguish the FDIAs from other system disturbances and it is robust against uncertainties in power system parameters and noisy measurements.

Insights
- FDIA detection and localization.
- The original dynamic model of the system is virtually split into two subsystems: subsystem-I is subject to disturbances but free from FDIAs, and subsystem-II is subject to FDIAs but free from disturbances.
- Attacks are identified by comparing the generated residuals with a specific adaptive threshold.

A. Kemmeugne, A. A. Jahromi, D. Kundur, "Resilience Enhancement of Pilot Protection in Power Systems," in *IEEE Transactions on Power Delivery*, vol. 37, no. 6, pp. 5255-5266, Dec. 2022, doi: 10.1109/TPWRD.2022.3175148.

Proposal
This paper investigates the resilience of pilot protection using a co-simulation platform based on OPAL-RT simulator and Riverbed Modeler. It uses software-defined networking for operational technology (OT SDN), which is a programmable architecture for communication networks that decouples the decision-making functions from packet forwarding functions in network devices and hands them to a centralized controller.

Insights
- Pilot protection is proposed for the first time and investigated for both SDN and legacy Ethernet-based communication networks.

- A software-defined networking can be integrated and implemented for pilot protection in substation environments.
- A co-simulation platform based on OPAL-RT and Riverbed.

W. Xu, I.M. Jaimoukha, F. Teng, "Robust Moving Target Defence Against False Data Injection Attacks in Power Grids," *IEEE Transactions on Information Forensics and Security*, vol. 18, pp. 29-40, 2023, doi: 10.1109/TIFS.2022.3210864.

Proposal
It proposes a Moving Target Defence (MTD) design problem in a noisy environment. The robust MTD guarantees the worst-case detection rate against all unknown attacks. The approach involves proactively triggering the distributed flexible AC transmission system (D-FACTS) devices. Theoretically proves that, for any given MTD strategy, the minimal principal angle between the Jacobian subspaces corresponds to the worst-case performance against all potential attacks.

Insights
- Proactively triggers the distributed flexible AC transmission system (D-FACTS) devices.
- Numerical simulations in IEEE case-6, 14, and 57 systems demonstrate the improved detection performance of robust MTD algorithms against the worst-case, random, and single-state attacks, under both simplified and full AC models.

J. Ruan et al., "Super-Resolution Perception Assisted Spatiotemporal Graph Deep Learning Against False Data Injection Attacks in Smart Grid," *IEEE Transactions on Smart Grid*, vol. 14, no. 5, pp. 4035-4046, Sept. 2023, doi: 10.1109/TSG.2023.3241268.

Proposal
A spatiotemporal graph deep learning (STGDL)-based scheme is proposed to detect cyberattacks without requiring attack samples. A Super-resolution perception (SRP) network is introduced, which is capable of reconstructing the high-frequency data of estimated states from low-frequency state estimation (SE) results, thereby improving the temporal learning ability in the STGDL model.

Insights
- The first step is to train the SRP networks for estimating high-frequency SE results. Input data: conventional low-frequency SE results. Output data: high-frequency SE results obtained from the offline computation.
- The second step is to train the proposed STGDL models using the high-frequency SE results obtained from the output of the well-trained SRP networks.
- The proposed FDIA detection is advantageous in reconstructing high-frequency temporal time series data from the low-frequency SE results.
- With the high-frequency temporal data fed into the system, the secure state bounds determined by the STGDL model can easily detect state anomalies, exhibiting effective FDIA detection ability.

J. Liang, L. Sankar, O. Kosut, "Vulnerability Analysis and Consequences of False Data Injection Attack on Power System State Estimation," *IEEE Transactions on Power Systems*, vol. 31, no. 5, pp. 3864-3872, Sept. 2016, doi: 10.1109/TPWRS.2015.2504950.

Proposal
This paper presents a bi-level optimization problem whose objective is to maximize the physical line flows subsequent to a False Data Injection (FDI) attack on DC state estimation (SE).

Insights
- The detectability constraint is modeled by limiting the cyber load shifts that result from the FDI attacks.

- The attacker has prior knowledge of system-wide topology, including line impedances, as well as generator cost functions.
- The attacker has real-time knowledge of measurements in a small area S bounded by buses.
- The attacker can change or replace all measurements in S.
- The attacker has arbitrary computational capability.
- The attacker maximizes the power flow on one branch while changing as few states as possible.
- The detectability constraint is modeled by limiting the cyber load shifts that result from the FDI attacks.

Y. Zheng, Z. Yan, K. Chen, J. Sun, Y. Xu, Y. Liu, "Vulnerability Assessment of Deep Reinforcement Learning Models for Power System Topology Optimization," *IEEE Transactions on Smart Grid*, vol. 12, no. 4, pp. 3613-3623, July 2021, doi: 10.1109/TSG.2021.3062700.

Proposal
This paper studies the vulnerability of deep reinforcement learning (DRL) models for power systems topology optimization under data perturbations and cyber-attacks.

Insights
- It assesses the vulnerability of a DRL model by constructing perturbations that minimize the model's performance.
- Indices based on probability and gradient criteria are proposed to identify the characteristics of perturbations that may cause malfunctions in the DRL model.
- The paper evaluates the vulnerability of DRL models specifically for power system topology optimization.
- It addresses the vulnerability of DRL-based controllers through a criticality-based adversarial perturbation model, which includes: (i) gradient-based methods; (ii) critical moments.

L. Liu, M. Esmalifalak, Q. Ding, V.A. Emesih, Z. Han, "Detecting false data injection attacks on power grid by sparse optimization", *IEEE Transactions on Smart Grid*, vol. 5, no. 2, pp. 612-621. doi: 10.1109/TSG.2013.2284438

Proposal
This method proposes to detect the FDI attack by applying techniques based on separation matrix taking advantage that the power system measurements are a time series and the FDI are sparse and of different nature. representation. The method is based on an augmented Lagrangian formulation.

Insights
- Good previous review of FDI attacks.
- Two methods for separating measurements and FDI. The most interesting is the low rank matrix factorization because its lower computational cost.
- The KPIs True Positive Rate (TPR) and False Negative Rate (FNR) appears as well as the Receiver Operating Characteristics (ROC) representation.

A. Ashok, M. Govindarasu, V. Ajjarapu, "Online Detection of Stealthy False Data Injection Attacks in Power System State Estimation", *IEEE Transactions on Smart Grid*, Vol. 9, No. 3, pp. 1636-1646, 2017, doi: 10.1109/TSG.2016.2596298

Proposal
The proposal is to compare the output of the state estimator, which can be corrupted with FDI, with other information based on historical data and PMU.

Insights
- Three major topics have been analyzed in the literature: vulnerability, impact and attack mitigation.
- Mitigation approaches:

  - Off-line mitigation: sensor placement (more data to modify by the cyberattacker); infrastructure approach (data encryption)
  - Online mitigation is based on some assumptions: therefore detection techniques may fail if these hypothesis are not valid.
- Proposal: introduce new information which is assumed to be transmitted over a different network (load forecast and synchrophasor measurements)
- Different types of attacks: FDI and topology-based attacks (the latter has no applicability in power plants).
- Normalized attack defined in Y. Liu, P. Ning, and M.K. Reiter, "False data injection attacks against state estimation in electric power grids", ACM Trans. Inf. Syst. Security, vol. 14, no. 1, pp. 1-33, May 2011, http://oi.acm.org/10.1145/1952982.1952995.
- KPIs for analyzing the algorithm performance: False Positive Rate (FPR); False Negative Rate (FNR); True Positive Rate (TPR); Receiver Operating Characteristics (ROC): classifier how well the algorithm performs considering that it is a yes/no methodology.
- Interesting methodology to set the threshold and the minimum attack magnitude.

K. Pan, A. Teixeira, M. Cvetkovic, P. Palensky, "Cyber Risk Analysis of Combined Data Attacks Against Power System State Estimation", *IEEE Transactions on Smart Grid*, Vol. 10, No. 3, pp. 3044-3056, 2019, doi: 10.1109/TSG.2018.2817387

Proposal

The proposal is focus on a risk analysis methodology of integrity and availability of combined cyber-attacks. The following definitions are provided: (i) availability attack: Dennial of Service (DoS); (ii) integrity attack: False Data Injection (FDI); (iii) stealth attack: FDI which cannot be detected by the Bad Data Detection (BDD) of the State Estimator; (iv) jamming attack: DoS related attack; (v) security index: minimum attack resources needed by the attacker to compromise the measurements while keeping stealth. Considering these definitions, stealth combined attacks launch both availability and FDI attacks without triggering the current BDD.

Insights

- Simplifications: DC power flow model.
- Defines the vulnerabilities in a risk assessment and the probability of being detected.
- FDI attack requires full (H) or partial information (topology with uncertain network parameters).
- Vulnerability assessment of integrity/availability attack: minimum number of measurements to be corrupted and not detected by the BDD. It can be formulated as an optimization problem.

C. Liu, H. Liang, T. Chen, J. Wu, C. Long, "Joint Admittance Perturbation and Meter Protection for Mitigating Stealthy FDI Attacks against Power System State Estimation", *IEEE Transactions on Power Systems*, Vol. 35, No. 2, pp. 1468-1478, 2020, doi: 10.1109/TPWRS.2019.2938223

Proposal

A methodology for detecting FDI based on changing the matrix H by using FACTS devices.

Insights

- This proposal cannot be applied to the COCOON project since it is not expected to have series FACTS in the pilots.

F. Almutairy, L. Scekic, R. Elmoudi, S. Wshah, "Accurate Detection of False Data Injection Attacks in Renewable Power Systems Using Deep Learning", IEEE Access, Vol. 9, pp. 135774-135789, 2021, doi: 10.1109/ACCESS.2021.3117230

Proposal

A methodology for detecting FDI based Deep Learning but far from the classical State Estimation technique adopted in the COCOON project.

Insights

- Different alternatives to face cybersecurity: protection versus detection.
- Protection techniques: it is required to protect the communication of a number of measurements (at least equal to the number of states of the system). PMUs are also possible but sensititve to PMU spoofing cyberattacks.
- Detection techniques: statistical, machine learning and deep learning techniques.
- KPIs to quantify the performance of the cyberatack detection algorithm:
  - True positives (TP) represents the number of correctly detected attacked samples.
  - False positives (FP) represents the number of normal samples falsely classified as attacked.
  - False negatives (FN) represents the number of attacked samples not detected.
  - Precision is computed as:
$$Precision = \frac{TP}{TP + FP}$$
  - Recall is calculated as:
$$Recall = \frac{TP}{TP + FN}$$
  - F1-score (*F1*)
$$F1 = \frac{2 \cdot Precision \cdot Recall}{Precision + Recall}$$
  - False positive rate (FPR) is the ratio between the number of normal samples falsely categorized as attacked and the total number of actual normal samples.
  - Detection rate (DR) is the number of attacks detected by the model divided by the total number of the performed attacks.

D. Mukherjee, S. Ghosh, R.K. Misra, "A Novel False Data Injection Attack Formulation Based on CUR Low-Rank Decomposition Method", *IEEE Transactions on Smart Grid*, Vol. 13, No. 6, pp. 4965-4968, 2022, doi: 10.1109/TSG.2022.3204214

Proposal

The paper describes how to attack bypassing the residue test of an state estimator taking advantage of the low-rank subspace of the topology matrix H.

Insights

- If the topology matrix H is known, it is possible to attack without being noticed in the residual tests according to equations (2)-(3) within the manuscript.
- b) Attacks can be designed by computing low-rank approximations of H, being two possibilities: (i) singular value decomposition (SVD) with a large computational cost; (ii) CUR decomposition which has the main advantage of a reduced computational cost.

C. Chen, Y. Wang, M. Cui, J. Zhao, W. Bi, Y. Chen, X. Zhang, "Data-Driven Detection of Stealthy False Data Injection Attack Against Power System State Estimation", *IEEE Transactions on Industrial Informatics*, Vol. 18, No. 12, pp. 8467-8476, 2022, doi: 10.1109/ TII.2022.3149106

Proposal

Detection of FDI based on unsupervised data-driven methods. These methods are unsupervised since since no "label" indicating if the analyzed dataset is compromised or not is available.

Insights

- Interesting discussion about the cyberattack policy: a big part of cyberattack policy is based on an optimization thinking but probably this is not the reality.
- It is designed a FDI strategy based on a System Security Index (SSI) and based on a sensitivity analysis. The idea is interesting but needs to be crafted considering a sound knowledge of power systems.

- It is proposed to classify the attacks in two groups: conservative and radical attacks depending on the electrical limit violations.
- Two unsupervised data-driven methods are explored: (i) clustering; (ii) Deep Autoencoding Gaussian Mixture Model (DAGMM).
- Some KPIs are used for analyzing the detector performance: accuracy (true positives plus true negatives), precision, recall and F1-score.

Z. Qu, J. Yang, Y. Wang, P.M. Georgievitch, "Detection of False Data Injection Attack in Power System Based on Hellinger Distance", *IEEE Transactions on Industrial Informatics*, 2023, doi: 10.1109/TII.2023.3286895

Proposal
Detection of FDI based on Hellinger distance, which is used to quantify the similarity between two probability distributions. The Hellinger distance is defined in terms of the Hellinger integral.

Insights
- The method is based on comparing real-time information, which can be attacked, with historical data, which are supposed to be correct ones. Therefore, this method does not consider the physics of the analyzed power system and its based just on statistical properties.

M. Du, X. Liu, Z. Li, H. Lin, "Robust Mitigation Strategy Against Dummy Data Attacks in Power Systems", *IEEE Transactions on Smart Grid*, Vol. 14, No. 4, pp. 3102-3113, 2023, doi: 10.1109/TSG.2022.3225469

Proposal
It is investigated the generation and defense of dummy attacks by modifying the generation scenario.

Insights
- Different attack mechanisms are described in the references [1]-[3].
- Corrective actions to avoid line overload are outlined in [10].
- Preventive actions to avoid line overload are described in [11].
- FDI attacks can be detected using anomaly detection algorithms as outliers.
- Dummy Data Attack (DDA): a new attack that cannot be seen by an anomaly detection algorithm, explained with the help of Fig. 1. This type of cyberattack can be constructed using a optimization problem posed by (7)-(13).
- Several proposed detection algorithms are based on statistical indices rather than in yes/no classification.

S. Nasiri, H. Seifi and H. Delkhosh, "A Secure Power System Distributed State Estimation via a Consensus-Based Mechanism and a Cooperative Trust Management Strategy," in IEEE Transactions on Industrial Informatics, vol. 20, no. 2, pp. 3002-3014, Feb. 2024, doi: 10.1109/TII.2023.3299385.

Proposal
This paper focuses on key challenges of future power systems, i.e., efficient decentralized decision-making and addressing cybersecurity risks, as reliable data exchanges among agents face cyber threats. Firstly, a blockchain framework for Distributed State Estimation (DSE) is developed together with a Distributed Information Kalman Filter (DIKF) that relies only on neighbors' shared state information. Consensus on predicted shared states and neighbor messaging avoiding iterative processes and P2P communications, reducing delays and costs. In addition, a trust management strategy is proposed for regions monitored by control centers (nodes) in the blockchain-based DSE, and based on this, the abnormal node behavior is detected to offer offering situational awareness by identifying corruption sources, affected nodes, unaffected nodes, and contamination levels. The trust management scheme can be divided in four distinct stages described in the Insights, which can be of potential interest to COCOON in cases of the energy communities using block-chain technology for data exchange and DSE. In addition, dynamic DSE is performed referring to the temporal relations of the system state variables, opposed to the static state estimation which only considers a single time sample. Finally, the efficiency of the proposed

DSE and Anomaly Detection are assessed via a Monte Carlo simulation (100 times) and by the calculation of metrics like the Accuracy and F1-Score.

$$\begin{aligned} \text{Accuracy} =& \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \\ F1 - \text{Score} =& 2 \times \left(\frac{\text{TP}}{2\text{TP} + \text{FN} + \text{FP}} \right) \end{aligned} \tag{30}$$

Insights

- *Self-Consistency Hypothesis*: As the data stored in blockchain are highly secure and reliable when they are finalized, their states before the attack are assumed consensual.
- *Mutual Consistency Hypothesis*: Due to practical issues in power systems, such as sudden load or renewable generation fluctuations and equipment trips causing topology changes, some states may consistently deviate more than expected. To address this, a multivariate probability distribution function is proposed, allowing nodes to assess the consistency of a target node's entire state. Mutual consistency represents the probability that all states of node i may deviate beyond their current deviations, considering the mutual covariance between the states.
- *Transmitted Information Consistency Hypothesis*: with this, the nodes can investigate the internal condition of the neighboring nodes is the transmitted information consistency index (TICI) between two nodes that share, at least, one state together. A probabilistic expression is developed to measure the trustworthiness of the information that node j sends to node i.
- *Shared States Consistency Hypothesis*: To assess the reliability of different nodes based on shared states, it is evident that achieving optimality in the proposed DSE requires neighboring nodes to estimate the shared states identically. However, in practice, there may be an unavoidable discrepancy between neighboring nodes concerning shared state variables.

H. Yang, X. He, Z. Wang, R. C. Qiu and Q. Ai, "Blind False Data Injection Attacks Against State Estimation Based on Matrix Reconstruction," in *IEEE Transactions on Smart Grid*, vol. 13, no. 4, pp. 3174-3187, July 2022, doi: 10.1109/TSG.2022.3164874.

Proposal

This paper introduces a blind FDIA strategy targeting state estimation in power grids, leveraging matrix reconstruction and subspace estimation. Unlike traditional methods, this approach requires no prior knowledge of system parameters or topology and instead relies on a limited set of measurement data. Notably, the study examines the impact of random measurement noise on subspace-based blind FDIA methods and proposes a scheme to mitigate its effects during the attack process. Compared to earlier blind FDIA techniques, this method addresses the challenge of handling measurement noise effectively, even with sparse data. Consequently, the proposed approach achieves a high success rate for FDIA and remains effective under conditions of limited measurement data. Furthermore, it demonstrates strong robustness when applied to large-scale power grids and scenarios with significant measurement noise.

Insights

- A matrix reconstruction technique is proposed to mitigate the impact of measurement noise, followed by a subspace estimation-based blind FDIA approach.
- This method operates without requiring any knowledge of the power system parameters and addresses the effects of measurement noise by adjusting the eigenvalues of the covariance matrix of the measurement data.
- The proposed approach accurately reconstructs the column space of the state estimation Jacobian matrix (SEJM), enabling the execution of more successful FDIA attacks.
- it starts with DC model for the state and then expands it to AC.

- The paper uses firstly the estimation residual to identify the false and malicious data, which is usually by the Chi2 detection with a pre-defined threshold. This BDD method determines whether there exists bad or malicious data by observing if the estimation residual is beyond a pre-defined threshold. Then, when the proposed blind FDIA is compared to other strategies, the widely-used BDD strategy based on the estimation residual is the largest normalized residual test (LNRT).

I. Zografopoulos, N. D. Hatziargyriou and C. Konstantinou, "Distributed Energy Resources Cybersecurity Outlook: Vulnerabilities, Attacks, Impacts, and Mitigations," in *IEEE Systems Journal*, vol. 17, no. 4, pp. 6695-6709, Dec. 2023, doi: 10.1109/JSYST.2023.3305757.

Proposal

This study explores the security gaps in the cyber and physical layers of DERs that threaten grid operations. It is a "classification" review paper. While prior studies highlight cyber-attacks on DER assets, they often focus on specific components, overlooks mission-critical functions, or neglects adversarial perspectives. In this work, adversarial goals and capabilities are analyzed, showing how protocol and device-level vulnerabilities can disrupt power systems. Finally, mitigation strategies are proposed and future research directions for DER cybersecurity. It could be a good review paper to serve as a basis for the COCOON threats to be tested in WP8 pilot with the PV Plants and the Power Plant Controller.

Insights

- Different attack vectors are identified and described and it is justified how they can be potential threats for DER assets.
- Definitions of Cyberattacks tailored for DER Assets (e.g., DoS, Eavesdropping, packet replay, etc.) are mapped in the relevant literature.
- The paper identifies/categorizes vulnerabilities, attacks, impacts and mitigations at DER protocol level and at DER device level.
- The paper reviews Cybersecurity Metrics for DER-Integrated Systems, e.g., the R4 resilience framework which introduces a quantitative metrics hierarchy under four main domains and EPRI's effort to quantifying security in diverse DER architectures has presented a data-driven cybersecurity metric methodology. The EPRI's metrics framework (60 metrics in total) combines real-world IT and OTdata aggregated from the system-under-test.

T. S. Sreeram and S. Krishna, "Graph-Based Assessment of Vulnerability to False Data Injection Attacks in Distribution Networks," in *IEEE Transactions on Power Systems*, vol. 39, no. 2, pp. 4510-4520, March 2024, doi: 10.1109/TPWRS.2023.3309777.

Proposal

This paper deals with the FDIA vulnerability in distribution networks. As the number of compromised measurements increases, the attacker's expenses also rise. To minimize costs, the attacker tends to select sparse attack vectors. Consequently, the attacker is likely to focus on measurements that can be compromised through a greater variety of sparse attack vectors. Since not all measurements are equally susceptible to FDIA, it is crucial to prioritize measurements by their vulnerability to such attacks. Towards this direction, in this paper, a fast graph-based algorithm is proposed in this article to rank all measurements based on FDIA vulnerability in distribution networks. While ranking meters based on their vulnerability to FDIA is an offline problem, it is inherently combinatorial, making the brute-force approach to solving it impractical. For this reason, this article uses a vulnerability index based on the number and sparsity of attack vectors. The proposed algorithm prioritizes measurements with a higher rank if they can be targeted by numerous sparse attack vectors, utilizing the radial structure of distribution networks and the topological characteristics of power flows.

Insights

- The meters are categorised into flow meters and injection meters, and the effect of having higher proportion of one of these types of meters on FDIA vulnerability is studied.

- This article uses a vulnerability index based on the number and sparsity of attack vectors
- The proposed algorithm identifies meters that should be secured to minimize vulnerability to FDIA, assessing the influence of pseudo-meters and secured meters on FDIA susceptibility, and optimizing measurement placement to reduce FDIA vulnerability.
- The article mentions that there exist two approaches to defense against FDIA: detection-based approach and protection-based approach. The method proposed in the article is related to the protection-based approach, in a way that the vulnerability assessment is identifying optimal locations for meter placement to minimize vulnerability to FDIA. This is important because both the number and location of meters significantly impact the vulnerability of measurements. Such protection-based approach could be of interest for COCOON to identify the most vulnerable inverters in a PV park or the most vulnerable RES meters in an energy community.

S. Bi and Y. J. Zhang, "Graphical Methods for Defense Against False-Data Injection Attacks on Power System State Estimation," in *IEEE Transactions on Smart Grid*, vol. 5, no. 3, pp. 1216-1227, May 2014, doi: 10.1109/TSG.2013.2294966.

Proposal

This paper focuses on using graphical methods to derive efficient strategies that defend any subset of state variables with minimum number of secure measurements. Graphical methods are employed to develop defenses against false data injection attacks on power system state estimation. By securing selected meter measurements, it ensures that no attacks can compromise any state variables. The optimal protection problem, which minimizes the number of measurements needed to safeguard state variables, is formulated as a variant of the Steiner tree problem. Both exact and approximate algorithms are proposed, with a tree-pruning-based approximation significantly reducing computational complexity while maintaining near-optimal performance.

Insights

- The conditions are established for selecting meter measurements that, when secured, prevent undetectable attacks on specified state variables. These conditions are instrumental in defining the optimal protection problem, which minimizes the cost of safeguarding state variables.
- This problem is modeled as a variant of the Steiner tree problem on a graph. Two exact solutions are proposed: a Steiner vertex enumeration algorithm and a mixed integer linear programming (MILP) model based on network flow, which reduces computational complexity by leveraging the solution's graphical structure.
- To address the problem's intractability, a polynomial-time tree-pruning heuristic (TPH) is introduced, providing near-optimal solutions with significantly lower computational demands.
- Although the approach is developed assuming the DC state estimation, a discussion on the AC state estimation is conducted. The proposed algorithms remain applicable for safeguarding AC state estimation when attackers compromise only the voltage phase angle variables, as in the DC model. Specifically, the methods continue to be both valid and optimal (for exact algorithms) in protecting state variables within AC state estimation. However, if attackers also compromise voltage amplitude state variables, the methods, while still valid, may no longer be optimal. This is because flow meter readings are now influenced by the absolute values of voltage amplitudes rather than their differences.

T. S. Sreeram and S. Krishna, "Managing False Data Injection Attacks During Contingency of Secured Meters," in *IEEE Transactions on Smart Grid*, vol. 10, no. 6, pp. 6945-6953, Nov. 2019, doi: 10.1109/TSG.2019.2914974.

Proposal

Protecting the system from FDIA by securing meters rely on the assumption that the secured meters offer 100% security all the time. For an n bus power system, a method to prevent FDIA is by securing a set of $(n-1)$ meters, called a basic measurement set (BMS). Selecting the appropriate BMS (called optimal BMS) is formulated as a multi-objective combinatorial optimization problem and the optimal BMS is found via the development of

a greedy algorithm. In this work, the case when one of the secured meters fails is studied. It is proved that that the failure in the protection system of a single meter in a BMS is sufficient to create the most vulnerable form of attacks, if the BMS is not chosen appropriately. If all the necessary meters cannot be secured, a greedy algorithm is developed for securing less than $(n-1)$ meters in an optimal BMS (subset of the optimal BMS), in such a way that the vulnerability of attack is minimum. This algorithm to solve the problem of securing a subset of meters in the optimal BMS is developed by combining an algorithm to obtain a collection of sparse attack vectors and a Binary Linear Programming problem. This could serve as a basis for COCOON to determine a BMS in in a large-scale PV plant or in an energy community or to select the subset of measuring devices that need to be secured definitely.

Insights

- The attacker has the same effort (or cost) to corrupt each meter and the attacker might not initiate an attack if the cost is high. Based on this, the number of non-zero elements in any attack vector satisfying $a = H\Delta\delta$ is called "cost of attack".
- It is assumed that an attack as in $a = H\Delta\delta$ alone is possible. It is also assumed that the rank of H is $n-1$.
- If some of the meters are secured, the entries of a corresponding to these meters are constrained to be zero. If, in spite of this constraint, there exists a nonzero a in the column space of H, FDIA is possible; otherwise FDIA is impossible.
- During the event of contingency of a single meter in a BMS, it is proved that sparse attacks are possible, the sparsity of which can be as high as that of a sparsest attack vector.
- The danger of selecting a random BMS and show that the planner can minimize this vulnerability if a BMS is selected appropriately.
- It can be concluded from the results that a significant improvement in security is possible when the optimal BMS is chosen.
- When a subset of meters in the optimal BMS is secured, the vulnerability is reduced by avoiding all attacks of cost less than a specified value.

H. Alamro, K. Mahmood, S. S. Aljameel, A. Yafoz, R. Alsini and A. Mohamed, "Modified Red Fox Optimizer With Deep Learning Enabled False Data Injection Attack Detection," in *IEEE Access*, vol. 11, pp. 79256-79264, 2023, doi: 10.1109/ACCESS.2023.3298056.

Proposal

This paper introduces a modified Red Fox Optimizer with Deep Learning-enabled FDIA detection (MRFODL-FDIAD) designed for identifying and classifying FDIAs in cyber-physical power systems (CPPS). The proposed approach operates through a three-stage process: pre-processing, detection, and parameter tuning. For FDIA detection, the MRFODL-FDIAD method leverages a multi-head attention-based long short-term memory (MBALSTM) model. To enhance the detection performance of the MBALSTM model, the MRFO technique is utilized. Experimental evaluations of the MRFODL-FDIAD approach were conducted using a standard IEEE bus system, and extensive experimentation demonstrated its good performance. This paper focuses on the detection of FDIAs but it is not relevant to COCOON objectives, since it does not use any state estimation technique nor involves any network topology information or information on measurements.

Insights

- The proposed classification of FDIAs based on the MRFODL-FDIAD model has been compared against several Machine Learning techiques, such as Deep support vector machine (DSVM), convolutional neural network (CNN), Support Vector Machines Gentle Adaboost (SVM-GAB), margin setting algorithm (MSA), and Conditional Deep Belief Network (CDBN).

Z. Cheng, H. Ren, J. Qin and R. Lu, "Security Analysis for Dynamic State Estimation of Power Systems With Measurement Delays," in *IEEE Transactions on Cybernetics*, vol. 53, no. 4, pp. 2087-2096, April 2023, doi: 10.1109/TCYB.2021.3108884.

Proposal

A modified state estimator utilizing the Kalman filter (KF) is developed to achieve optimal state estimation despite measurement delays. An FDI attack strategy is formulated to exploit measurement delays, compromising the state estimator while avoiding detection by the chi-square detector. Two measurement residual vectors are constructed by combining the attacked estimated states with uncorrupted measurement data. Using these residual vectors, a chi-square-based detection method is introduced, capable of identifying attacks even in the presence of delayed measurements. This paper can serve as a comparison basis for the COCOON pilots, in case measurement delays need to be examined and distinguished from FDIAs.

Insights

- Compared to references [10] and [17] of the manuscript, the proposed algorithm can be robust to partial measurement delays and efficient in tackling both RTU and PMU measurements, which is closer to the practical system.
- Compared to general attack vectors in references [40] and [41] of the manuscript, the proposed attack is more difficult to be identified by the detector because of measurement delays.
- Compared to the detection methods in references [23] and [32] of the manuscript, this approach can achieve effective detection even if the measurement delay occurs. This is achieved by introducing historical measurement information that has not been attacked.

Bibliography

[1] A. v. Meier. “AC Power”. In: *Electric Power Systems: A Conceptual Introduction*. 2006, pp. 49–84. DOI: 10.1002/0470036427.ch3.

[2] M. E. El-Hawary. “Introduction”. In: *Introduction to Electrical Power Systems*. 2008, pp. 1–7. DOI: 10.1002/9780470411377.ch1.

[3] F. Careri, C. Genesi, P. Marannino, M. Montagna, S. Rossi, and I. Siviero. “Generation Expansion Planning in the Age of Green Economy”. In: *IEEE Transactions on Power Systems* 26.4 (2011), pp. 2214–2223.

[4] M. E. El-Hawary. “Electric Power Transmission”. In: *Introduction to Electrical Power Systems*. 2008, pp. 129–189. DOI: 10.1002/9780470411377.ch5.

[5] C. S. Demoulias, K.-N. D. Malamaki, G. C. Kryonidis, E. O. Kontis, S. I. Gkavanoudis, K. O. Oureilidis, and J. M. Mauricio. “Ancillary services provision in terminal distribution systems”. In: *Encyclopedia of Electrical and Electronic Power Engineering*. Ed. by J. Garcia. Oxford: Elsevier, 2023, pp. 411–424. DOI: https://doi.org/10.1016/B978-0-12-821204-2.00042-8.

[6] K. Oureilidis, K.-N. Malamaki, K. Gallos, A. Tsitsimelis, C. Dikaiakos, S. Gkavanoudis, M. Cvetkovic, J. M. Mauricio, J. M. Maza Ortega, J. L. M. Ramos, G. Papaioannou, and C. Demoulias. “Ancillary Services Market Design in Distribution Networks: Review and Identification of Barriers”. In: *Energies* 13.4 (2020). DOI: 10.3390/en13040917.

[7] C. S. Demoulias, K.-N. D. Malamaki, S. Gkavanoudis, J. M. Mauricio, G. C. Kryonidis, K. O. Oureilidis, E. O. Kontis, and J. L. Martinez Ramos. “Ancillary Services Offered by Distributed Renewable Energy Sources at the Distribution Grid Level: An Attempt at Proper Definition and Quantification”. In: *Applied Sciences* 10.20 (2020). DOI: 10.3390/app10207106.

[8] G. C. Kryonidis, K.-N. D. Malamaki, S. I. Gkavanoudis, K. O. Oureilidis, E. O. Kontis, J. M. Mauricio, J. M. Maza-Ortega, and C. S. Demoulias. “Distributed Reactive Power Control Scheme for the Voltage Regulation of Unbalanced LV Grids”. In: *IEEE Transactions on Sustainable Energy* 12.2 (2021), pp. 1301–1310. DOI: 10.1109/TSTE.2020.3042855.

[9] K.-N. D. Malamaki, D.-A. Christofis, G. C. Kryonidis, and C. S. Demoulias. “Voltage Harmonic Mitigation by Distributed Renewable Energy Sources in Low-Voltage Distribution Networks: Sensitivity Analysis”. In: *IEEE Transactions on Industry Applications* 60.5 (2024), pp. 7656–7671. DOI: 10.1109/TIA.2024.3420826.

[10] S. C. Dimoulias, K.-N. D. Malamaki, and G. C. Kryonidis. “Power Smoothing as a Mitigation Action against Rapid Voltage Changes: A Comparative Study”. In: *2024 International Conference on Smart Energy Systems and Technologies (SEST)*. 2024, pp. 1–6. DOI: 10.1109/SEST61601.2024.10694205.

[11] M. Goulden, B. Bedwell, S. Rennick-Egglestone, T. Rodden, and A. Spence. “Smart grids, smart users? The role of the user in demand side management”. In: *Energy Research & Social Science* 2 (2014), pp. 21–29. DOI: https://doi.org/10.1016/j.erss.2014.04.008.

[12] P. Siano. “Demand response and smart grids—A survey”. In: *Renewable and Sustainable Energy Reviews* 30 (2014), pp. 461–478. DOI: https://doi.org/10.1016/j.rser.2013.10.022.

[13] Y. Yan, Y. Qian, H. Sharif, and D. Tipper. “A Survey on Smart Grid Communication Infrastructures: Motivations, Requirements and Challenges”. In: *IEEE Communications Surveys & Tutorials* 15.1 (2013), pp. 5–20. DOI: 10.1109/SURV.2012.021312.00034.

[14] International Energy Agency (IEA). *Cybersecurity: Is the Power System Lagging Behind?* https://www.iea.org/commentaries/cybersecurity-is-the-power-system-lagging-behind. [Accessed: December 4, 2024]. 2023.

[15] National Institute of Standards and Technology (NIST). *Enhanced Security Requirements for Protecting Controlled Unclassified Information: A Supplement to NIST Special Publication 800-171*. Tech. rep. 800-172. Accessed: 2024-12-04. National Institute of Standards and Technology, 2021.

[16] European Union Agency for Cybersecurity (ENISA). *ENISA Threat Landscape 2024*. Tech. rep. Accessed: 2024-12-04. European Union Agency for Cybersecurity (ENISA), 2024.

[17] Center for Strategic and International Studies (CSIS). *Significant Cyber Incidents*. Accessed: 2024-12-04. 2024.

[18] Cybersecurity and Infrastructure Security Agency (CISA). *Cybersecurity and Infrastructure Security Agency (CISA)*. Accessed: 2024-12-04. 2024.

[19] Kaspersky ICS-CERT. *Threat Landscape for Industrial Automation Systems: First Half of 2019*. Tech. rep. Accessed: 2024-12-04. Kaspersky ICS-CERT, 2019.

[20] H. T. Reda, A. Anwar, and A. Mahmood. "Comprehensive survey and taxonomies of false data injection attacks in smart grids: attack models, targets, and impacts". In: *Renewable and Sustainable Energy Reviews* 163 (2022), p. 112423. DOI: https://doi.org/10.1016/j.rser.2022.112423.

[21] Y. Kabalci. "A survey on smart metering and smart grid communication". In: *Renewable and Sustainable Energy Reviews* 57 (2016), pp. 302–318. DOI: https://doi.org/10.1016/j.rser.2015.12.114.

[22] G. Liang, J. Zhao, F. Luo, S. R. Weller, and Z. Y. Dong. "A Review of False Data Injection Attacks Against Modern Power Systems". In: *IEEE Transactions on Smart Grid* 8.4 (2017), pp. 1630–1638. DOI: 10.1109/TSG.2015.2495133.

[23] Q. Wang, W. Tai, Y. Tang, and M. Ni. "Review of the false data injection attack against the cyber-physical power system". In: *IET Cyber-Phys. Syst.: Theory Appl.* 4.2 (2019), pp. 101–107. DOI: https://doi.org/10.1049/iet-cps.2018.5022.

[24] Q. Zhang, F. Li, Q. Shi, K. Tomsovic, J. Sun, and L. Ren. "Profit-Oriented False Data Injection on Electricity Market: Reviews, Analyses, and Insights". In: *IEEE Transactions on Industrial Informatics* 17.9 (2021), pp. 5876–5886. DOI: 10.1109/TII.2020.3036104.

[25] H. Zhang, B. Liu, and H. Wu. "Smart Grid Cyber-Physical Attack and Defense: A Review". In: *IEEE Access* 9 (2021), pp. 29641–29659. DOI: 10.1109/ACCESS.2021.3058628.

[26] D. Du, M. Zhu, X. Li, M. Fei, S. Bu, L. Wu, and K. Li. "A Review on Cybersecurity Analysis, Attack Detection, and Attack Defense Methods in Cyber-physical Power Systems". In: *Journal of Modern Power Systems and Clean Energy* 11.3 (2023), pp. 727–743. DOI: 10.35833/MPCE.2021.000604.

[27] S. Aoufi, A. Derhab, and M. Guerroumi. "Survey of false data injection in smart power grid: Attacks, countermeasures and challenges". In: *Journal of Information Security and Applications* 54 (2020), p. 102518. DOI: https://doi.org/10.1016/j.jisa.2020.102518.

[28] N. D. Tuyen, N. S. Quan, V. B. Linh, V. V. Tuyen, and G. Fujita. "A Comprehensive Review of Cybersecurity in Inverter-Based Smart Power System Amid the Boom of Renewable Energy". In: *IEEE Access* 10 (2022), pp. 35846–35875. DOI: 10.1109/access.2022.3163551.

[29] TechTarget. *SolarWinds hack explained: Everything you need to know*. Accessed: 2024-12-04. 2023.

[30] Cybersecurity and I. S. A. (CISA). *The Attack on Colonial Pipeline: What We've Learned & What We've Done Over the Past Two Years*. https://www.cisa.gov/news-events/news/attack-colonial-pipeline-what-weve-learned-what-weve-done-over-past-two-years. Accessed: 2025-01-23. 2023.

[31] A. S. Musleh, G. Chen, and Z. Y. Dong. "A Survey on the Detection Algorithms for False Data Injection Attacks in Smart Grids". In: *IEEE Transactions on Smart Grid* 11.3 (2020), pp. 2218–2234. DOI: 10.1109/TSG.2019.2949998.

[32] A. Huseinovic, S. Mrdovic, K. Bicakci, and S. Uludag. "A Survey of Denial-of-Service Attacks and Solutions in the Smart Grid". In: *IEEE Access* 8 (2020), pp. 177447–177470.

[33] I. Ortega-Fernandez and F. Liberati. “A Review of Denial of Service Attack and Mitigation in the Smart Grid Using Reinforcement Learning”. In: *Energies* 16 (2023).

[34] X. Lou, C. Tran, R. Tan, D. K. Y. Yau, Z. T. Kalbarczyk, A. K. Banerjee, and P. Ganesh. “Assessing and Mitigating Impact of Time Delay Attack: Case Studies for Power Grid Controls”. In: *IEEE J. Sel. Areas Commun.* 38.1 (2020), pp. 141–155. DOI: 10.1109/JSAC.2019.2951982.

[35] Z. Zhang, R. Deng, P. Cheng, and Q. Wei. “On Feasibility of Coordinated Time-Delay and False Data Injection Attacks on Cyber–Physical Systems”. In: *IEEE Internet Things J.* 9.11 (2022), pp. 8720–8736. DOI: 10.1109/JIOT.2021.3118065.

[36] S. e. a. Abdelkader. “Securing modern power systems: Implementing comprehensive strategies to enhance resilience and reliability against cyber-attacks”. In: *Results Eng.* 23.102647 (2022), pp. 581–595. DOI: 10.1016/j.rineng.2024.102647.

[37] D. E. Whitehead, K. Owens, D. Gammel, and J. Smith. “Ukraine cyber-induced power outage: Analysis and practical mitigation strategies”. In: *Proc. 2017 70th Annual Con. for Protect. Relay Engin. (CPRE)*. 2017, pp. 1–8. DOI: 10.1109/CPRE.2017.8090056.

[38] G. Liang, S. R. Weller, J. Zhao, F. Luo, and Z. Y. Dong. “The 2015 Ukraine Blackout: Implications for False Data Injection Attacks”. In: *IEEE Trans. Power Syst.* 32.4 (2017), pp. 3317–3318. DOI: 10.1109/TPWRS.2016.2631891.

[39] N. Tatipatri and S. L. Arun. “A Comprehensive Review on Cyber-Attacks in Power Systems: Impact Analysis, Detection, and Cyber Security”. In: *IEEE Access* 12 (2024), pp. 18147–18167. DOI: 10.1109/ACCESS.2024.3361039.

[40] F. Liberati, E. Garone, and A. Di Giorgio. “Review of Cyber-Physical Attacks in Smart Grids: A System-Theoretic Perspective”. In: *Electronics* 10.10 (2021). DOI: 10.3390/electronics10101153.

[41] T. Yang, Y. Liu, and W. Li. “Attack and defence methods in cyber-physical power system”. In: *IET Energy Syst. Integr.* 4 (2 2022), pp. 159–170. DOI: 10.1049/esi2.12068.

[42] M. Ghiasi, T. Niknam, Z. Wang, M. Mehrandezh, M. Dehghani, and N. Ghadimi. “A comprehensive review of cyber-attacks and defense mechanisms for improving security in smart grid energy systems: Past, present and future”. In: *Electr. Power Syst. Res.* 215.108975 (2023), pp. 727–743. DOI: 10.1016/j.epsr.2022.108975.

[43] A. Amulya, K. Swarup, and R. Ramanathan. “Cyber Security of Smart-Grid Frequency Control: A Review and Vulnerability Assessment Framework”. In: *ACM Trans. Cyber-Phys. Syst.* 8.4 (2024), pp. 1–27. DOI: 10.1145/3661827.

[44] A. Sayghe, Y. Hu, I. Zografopoulos, X. Liu, R. G. Dutta, J. Yin, and K. Charalampos. “Survey of machine learning methods for detecting false data injection attacks in power systems”. In: *IET Smart Grid* 3 (5 2022), pp. 581–595. DOI: 10.1049/iet-stg.2020.0015.

[45] Z. Zhang, S. Gong, A. D. Dimitrovski, and H. Li. “Time Synchronization Attack in Smart Grid: Impact and Analysis”. In: *IEEE Transactions on Smart Grid* 4.1 (2013), pp. 87–98. DOI: 10.1109/TSG.2012.2227342.

[46] P. Risbud, N. Gatsis, and A. Taha. “Vulnerability Analysis of Smart Grids to GPS Spoofing”. In: *IEEE Transactions on Smart Grid* 10.4 (2019), pp. 3535–3548. DOI: 10.1109/TSG.2018.2830118.

[47] P. e. a. Wlazlo. “Man-in-the-middle attacks and defence in a power system cyber-physical testbed”. In: *IET Cyber-Phys. Syst.: Theory Appl.* 6 (3 2020), pp. 164–177. DOI: 10.1049/cps2.12014.

[48] C. Konstantinou and M. Maniatakos. “A Case Study on Implementing False Data Injection Attacks Against Nonlinear State Estimation”. In: *Proc. 2nd ACM Workshop on Cyber-Physical Systems Security and Privacy* (2016), pp. 81–92. DOI: 10.1145/2994487.2994491.

[49] M. Usama and M. N. Aman. “Command Injection Attacks in Smart Grids: A Survey”. In: *IEEE Open Journal of Industry Applications* 5 (2024), pp. 75–85. DOI: 10.1109/OJIA.2024.3365576.

[50] Y. Yuan, Z. Li, and K. Ren. "Modeling Load Redistribution Attacks in Power Systems". In: *IEEE Transactions on Smart Grid* 2.2 (2011), pp. 382–390. DOI: 10.1109/TSG.2011.2123925.

[51] P. Verma and C. Chakraborty. "Load Redistribution Attacks Against Smart Grids–Models, Impacts, and Defense: A Review". In: *IEEE Transactions on Industrial Informatics* 20.8 (2024), pp. 10192–10208. DOI: 10.1109/TII.2024.3393005.

[52] M. Mahrukh and M. S. Thomas. "Load Altering Attacks- a Review of Impact and Mitigation Strategies". In: *Proc. 2023 International Conference on Recent Advances in Electrical, Electronics & Digital Healthcare Technologies (REEDCON)*. 2023, pp. 397–402. DOI: 10.1109/REEDCON57544.2023.10150456.

[53] S. Yankson and M. Ghamkhari. "Transactive Energy to Thwart Load Altering Attacks on Power Distribution Systems". In: *Future Internet* 12.1 (2020). DOI: 10.3390/fi12010004.

[54] G. Liang, S. R. Weller, J. Zhao, F. Luo, and Z. Y. Dong. "A Framework for Cyber-Topology Attacks: Line-Switching and New Attack Scenarios". In: *IEEE Transactions on Smart Grid* 10.2 (2019), pp. 1704–1712. DOI: 10.1109/TSG.2017.2776325.

[55] L. Xie, Y. Mo, and B. Sinopoli. "Integrity Data Attacks in Power Market Operations". In: *IEEE Transactions on Smart Grid* 2.4 (2011), pp. 659–666. DOI: 10.1109/TSG.2011.2161892.

[56] G. Hug and J. A. Giampapa. "Vulnerability Assessment of AC State Estimation with Respect to False Data Injection Cyber-Attacks". In: *IEEE Transactions on Smart Grid* 3.3 (2012), pp. 1362–1370. DOI: 10.1109/tsg.2012.2195338.

[57] R. Deng, G. Xiao, R. Lu, H. Liang, and A. V. Vasilakos. "False Data Injection on State Estimation in Power SystemsAttacks, Impacts, and Defense: A Survey". In: *IEEE Transactions on Industrial Informatics* 13.2 (2017), pp. 411–423. DOI: 10.1109/tii.2016.2614396.

[58] P. L. Bhattar, N. M. Pindoriya, and A. Sharma. "A combined survey on distribution system state estimation and false data injection in cyber□physical power distribution networks". In: *IET Cyber-Phys. Syst.: Theory Appl.* 6 (2 2020), pp. 41–62. DOI: 10.1049/cps2.12000.

[59] L. Cui, Y. Qu, L. Gao, G. Xie, and S. Yu. "Detecting false data attacks using machine learning techniques in smart grid: A survey". In: *J. Netw. Comput. Appl.* 170 (2020), p. 102808. DOI: 10.1016/j.jnca.2020.102808.

[60] M. Irfan, A. Sadighian, A. Tanveer, S. J. Al□Naimi, and G. Oligeri. "A survey on detection and localisation of false data injection attacks in smart grids". In: *IET Cyber-Phys. Syst.: Theory Appl.* 9 (4 2024), pp. 313–333. DOI: 10.1049/cps2.12093.

[61] X. Liu and Z. Li. "False Data Attacks Against AC State Estimation With Incomplete Network Information". In: *IEEE Transactions on Smart Grid* 8.5 (2017), pp. 2239–2248. DOI: 10.1109/tsg.2016.2521178.

[62] K. Sun, I. Esnaola, A. M. Tulino, and H. V. Poor. "Asymptotic Learning Requirements for Stealth Attacks on Linearized State Estimation". In: *IEEE Transactions on Smart Grid* 14.4 (2023), pp. 3189–3200. DOI: 10.1109/tsg.2023.3236785.

[63] Z.-H. Yu and W.-L. Chin. "Blind False Data Injection Attack Using PCA Approximation Method in Smart Grid". In: *IEEE Transactions on Smart Grid* 6.3 (2015), pp. 1219–1226. DOI: 10.1109/tsg.2014.2382714.

[64] S. Lakshminarayana, A. Kammoun, M. Debbah, and H. V. Poor. "Data-Driven False Data Injection Attacks Against Power Grids: A Random Matrix Approach". In: *IEEE Transactions on Smart Grid* 12.1 (2021), pp. 635–646. DOI: 10.1109/tsg.2020.3011391.

[65] Y. Liu, O. Ardakanian, I. Nikolaidis, and H. Liang. "False Data Injection Attacks on Smart Grid Voltage Regulation With Stochastic Communication Model". In: *IEEE Transactions on Industrial Informatics* 19.5 (2023), pp. 7122–7132. DOI: 10.1109/tii.2022.3209287.

[66] P. Zhuang, R. Deng, and H. Liang. "False Data Injection Attacks Against State Estimation in Multiphase and Unbalanced Smart Distribution Systems". In: *IEEE Transactions on Smart Grid* 10.6 (2019), pp. 6000–6013. DOI: 10.1109/tsg.2019.2895306.

[67] J. Zhao, G. Zhang, Z. Y. Dong, and K. P. Wong. "Forecasting-Aided Imperfect False Data Injection Attacks Against Power System Nonlinear State Estimation". In: *IEEE Transactions on Smart Grid* 7.1 (2016), pp. 6–8. DOI: 10.1109/tsg.2015.2490603.

[68] J. Liang, L. Sankar, and O. Kosut. "Vulnerability Analysis and Consequences of False Data Injection Attack on Power System State Estimation". In: *IEEE Transactions on Power Systems* 31.5 (2015), pp. 3864–3872. DOI: 10.1109/tpwrs.2015.2504950. eprint: 1506.03774.

[69] F. Ye, X. Cao, Z. Cheng, and M.-Y. Chow. "CASL: A Novel Collusion Attack Against Distributed Energy Management Systems". In: *IEEE Transactions on Smart Grid* 14.6 (2023), pp. 4717–4728. DOI: 10.1109/tsg.2023.3251401.

[70] S. Gao, H. Zhang, Z. Wang, C. Huang, and H. Yan. "Data-Driven Injection Attack Strategy for Linear Cyber-Physical Systems: An Input-Output Data-Based Approach". In: *IEEE Transactions on Network Science and Engineering* 10.6 (2023), pp. 4082–4095. DOI: 10.1109/tnse.2023.3292403.

[71] G. Cheng, Y. Lin, J. Yan, J. Zhao, and L. Bai. "Model-Measurement Data Integrity Attacks". In: *IEEE Transactions on Smart Grid* 14.6 (2023), pp. 4741–4757. DOI: 10.1109/tsg.2023.3253781.

[72] M. Esmalifalak, H. Nguyen, R. Zheng, L. Xie, L. Song, and Z. Han. "A Stealthy Attack Against Electricity Market Using Independent Component Analysis". In: *IEEE Systems Journal* 12.1 (2018), pp. 297–307. DOI: 10.1109/JSYST.2015.2483742.

[73] H. Yang and Z. Wang. "A False Data Injection Attack Approach Without Knowledge of System Parameters Considering Measurement Noise". In: *IEEE Internet of Things Journal* 11.1 (2024), pp. 1452–1464. DOI: 10.1109/JIOT.2023.3288983.

[74] M. Du, G. Pierrou, X. Wang, and M. Kassouf. "Targeted False Data Injection Attacks Against AC State Estimation Without Network Parameters". In: *IEEE Transactions on Smart Grid* 12.6 (2021), pp. 5349–5361. DOI: 10.1109/tsg.2021.3106246.

[75] T. S. Sreeram and S. Krishna. "Protection Against False Data Injection Attacks Considering Degrees of Freedom in Attack Vectors". In: *IEEE Transactions on Smart Grid* 12.6 (2021), pp. 5258–5267. DOI: 10.1109/tsg.2021.3093498.

[76] R. Deng, P. Zhuang, and H. Liang. "False Data Injection Attacks Against State Estimation in Power Distribution Systems". In: *IEEE Transactions on Smart Grid* 10.3 (2018), pp. 2871–2881. DOI: 10.1109/tsg.2018.2813280.

[77] L. Che, X. Liu, Z. Shuai, and J. Zhao. "The Impact of Ramp-Induced Data Attacks on Power System Operational Security". In: *IEEE Transactions on Industrial Informatics* 15 (2019), pp. 5064–5075. DOI: 10.1109/tii.2019.2895058.

[78] N. N. Tran, H. R. Pota, Q. N. Tran, and J. Hu. "Designing Constraint-Based False Data-Injection Attacks against the Unbalanced Distribution Smart Grids". In: *IEEE Internet of Things Journal* 8 (2021), pp. 9422–9435. DOI: 10.1109/jiot.2021.3056649.

[79] D. Mukherjee. "Data-Driven False Data Injection Attack: A Low-Rank Approach". In: *IEEE Transactions on Smart Grid* 13.3 (2022), pp. 2479–2482. DOI: 10.1109/tsg.2022.3145633.

[80] D. Mukherjee, S. Ghosh, and R. K. Misra. "A Novel False Data Injection Attack Formulation Based on CUR Low-Rank Decomposition Method". In: *IEEE Transactions on Smart Grid* 13.6 (2022), pp. 4965–4968. DOI: 10.1109/tsg.2022.3204214.

[81] H. Yang, X. He, Z. Wang, R. C. Qiu, and Q. Ai. "Blind False Data Injection Attacks Against State Estimation Based on Matrix Reconstruction". In: *IEEE Transactions on Smart Grid* 13.4 (2022), pp. 3174–3187. DOI: 10.1109/tsg.2022.3164874.

[82] X. Liu, Z. Bao, D. Lu, and Z. Li. "Modeling of Local False Data Injection Attacks With Reduced Network Information". In: *IEEE Transactions on Smart Grid* 6.4 (2015), pp. 1686–1696. DOI: 10.1109/TSG.2015.2394358.

[83] X. Liu and Z. Li. "Local Load Redistribution Attacks in Power Systems With Incomplete Network Information". In: *IEEE Transactions on Smart Grid* 5.4 (2014), pp. 1665–1676. DOI: 10.1109/TSG.2013.2291661.

[84] J. Hou, J. Wang, Y. Song, W. Sun, and Y. Hou. "Small-Signal Angle Stability-Oriented False Data Injection Cyber-Attacks on Power Systems". In: *IEEE Transactions on Smart Grid* 14 (2023), pp. 635–648. DOI: 10.1109/tsg.2022.3199366.

[85] C. Liu, H. Liang, and T. Chen. "Network Parameter Coordinated False Data Injection Attacks Against Power System AC State Estimation". In: *IEEE Transactions on Smart Grid* 12.2 (2021), pp. 1626–1639. DOI: 10.1109/tsg.2020.3033520.

[86] X. Liu and Z. Li. "Local topology attacks in smart grids". In: *IEEE Transactions on Smart Grid* 8 (2017), pp. 2617–2626. DOI: 10.1109/tsg.2016.2532347.

[87] K. Pan, A. Teixeira, M. Cvetkovic, and P. Palensky. "Cyber Risk Analysis of Combined Data Attacks Against Power System State Estimation". In: *IEEE Transactions on Smart Grid* 10.3 (2018), pp. 3044–3056. DOI: 10.1109/tsg.2018.2817387.

[88] S. K. Singh, K. Khanna, R. Bose, B. K. Panigrahi, and A. Joshi. "Joint-Transformation-Based Detection of False Data Injection Attacks in Smart Grid". In: *IEEE Transactions on Industrial Informatics* 14.1 (2018), pp. 89–97. DOI: 10.1109/tii.2017.2720726.

[89] B. Li, T. Ding, C. Huang, J. Zhao, Y. Yang, and Y. Chen. "Detecting False Data Injection Attacks Against Power System State Estimation With Fast Go-Decomposition Approach". In: *IEEE Transactions on Industrial Informatics* 15.5 (2019), pp. 2892–2904. DOI: 10.1109/tii.2018.2875529.

[90] K. Khanna, B. K. Panigrahi, and A. Joshi. "Priority-Based Protection against the Malicious Data Injection Attacks on State Estimation". In: *IEEE Systems Journal* 14 (2020), pp. 1945–1952. DOI: 10.1109/jsyst.2019.2933023.

[91] N. Ahmadi, Y. Chakhchoukh, and H. Ishii. "Power Systems Decomposition for Robustifying State Estimation Under Cyber Attacks". In: *IEEE Transactions on Power Systems* 36.3 (2021), pp. 1922–1933. DOI: 10.1109/tpwrs.2020.3026951.

[92] C. Liu, R. Deng, W. He, H. Liang, and W. Du. "Optimal Coding Schemes for Detecting False Data Injection Attacks in Power System State Estimation". In: *IEEE Transactions on Smart Grid* 13.1 (2022), pp. 738–749. DOI: 10.1109/tsg.2021.3107972.

[93] J. Ruan, G. Liang, J. Zhao, J. Qiu, and Z. Y. Dong. "An Inertia-Based Data Recovery Scheme for False Data Injection Attack". In: *IEEE Transactions on Industrial Informatics* 18 (2022), pp. 7814–7823. DOI: 10.1109/tii.2022.3146859.

[94] N. Ahmadi, Y. Chakhchoukh, and H. Ishii. "Analysis of Targeted Coordinated Attacks on Decomposition-Based Robust State Estimation". In: *IEEE Open Access Journal of Power and Energy* 10 (2023), pp. 116–127. DOI: 10.1109/oajpe.2022.3230905.

[95] H. Pan, X. Feng, C. Na, and H. Yang. "A Model for Detecting False Data Injection Attacks in Smart Grids Based on the Method Utilized for Image Coding". In: *IEEE Systems Journal* 17.4 (2023), pp. 6181–6191. DOI: 10.1109/jsyst.2023.3287924.

[96] Y. Zhao, J. Liu, X. Liu, K. Yuan, and T. Ding. "Enhancing the Tolerance of Voltage Regulation to Cyber Contingencies via Graph-Based Deep Reinforcement Learning". In: *IEEE Transactions on Power Systems* 39.2 (2024), pp. 4661–4673. DOI: 10.1109/tpwrs.2023.3319699.

[97] A. Takiddin, M. Ismail, R. Atat, K. R. Davis, and E. Serpedin. "Robust Graph Autoencoder-Based Detection of False Data Injection Attacks Against Data Poisoning in Smart Grids". In: *IEEE Transactions on Artificial Intelligence* (2023). DOI: 10.1109/tai.2023.3286831.

[98] H. Ibrahim, J. Kim, J. Ramos-Ruiz, W. H. Ko, T. Huang, P. Enjeti, P. R. Kumar, and L. Xie. "Detection of Cyber Attacks in Grid-tied PV Systems Using Dynamic Watermarking". In: *IEEE Transactions on Industry Applications* (2023), pp. 1–10. DOI: 10.1109/tia.2023.3321588.

[99] Z. Kazemi, A. A. Safavi, F. Naseri, L. Urbas, and P. Setoodeh. "A Secure Hybrid Dynamic-State Estimation Approach for Power Systems Under False Data Injection Attacks". In: *IEEE Transactions on Industrial Informatics* 16.12 (2020), pp. 7275–7286. DOI: 10.1109/tii.2020.2972809.

[100] H. M. Albunashee, C. Farnell, A. Suchanek, K. Haulmark, R. A. McCann, J. Di, and A. Mantooth. "A Test Bed for Detecting False Data Injection Attacks in Systems With Distributed Energy Resources". In: *IEEE Journal of Emerging and Selected Topics in Power Electronics* 10 (2022), pp. 1303–1315. DOI: 10.1109/jestpe.2019.2948216.

[101] R. Deng, G. Xiao, and R. Lu. "Defending Against False Data Injection Attacks on Power System State Estimation". In: *IEEE Transactions on Industrial Informatics* 13 (2017), pp. 198–207. DOI: 10.1109/tii.2015.2470218.

[102] Z. Wang, H. He, Z. Wan, and Y. Sun. "Detection of False Data Injection Attacks in AC State Estimation Using Phasor Measurements". In: *IEEE Transactions on Smart Grid* PP.99 (2020), pp. 1–1. DOI: 10.1109/tsg.2020.2972781.

[103] A. D. Syrmakesis, H. H. Alhelou, and N. D. Hatziargyriou. "Novel SMO-Based Detection and Isolation of False Data Injection Attacks Against Frequency Control Systems". In: *IEEE Transactions on Power Systems* 39.1 (2023), pp. 1434–1446. DOI: 10.1109/tpwrs.2023.3242015.

[104] A. Kemmeugne, A. A. Jahromi, and D. Kundur. "Resilience Enhancement of Pilot Protection in Power Systems". In: *IEEE Transactions on Power Delivery* 37 (2022), pp. 5255–5266. DOI: 10.1109/tpwrd.2022.3175148.

[105] W. Xu, I. M. Jaimoukha, and F. Teng. "Robust Moving Target Defence Against False Data Injection Attacks in Power Grids". In: *IEEE Transactions on Information Forensics and Security* 18 (2023), pp. 29–40. DOI: 10.1109/tifs.2022.3210864.

[106] J. Ruan, G. Fan, Y. Zhu, G. Liang, J. Zhao, F. Wen, and Z. Y. Dong. "Super-Resolution Perception Assisted Spatiotemporal Graph Deep Learning Against False Data Injection Attacks in Smart Grid". In: *IEEE Transactions on Smart Grid* 14.5 (2023), pp. 4035–4046. DOI: 10.1109/tsg.2023.3241268.

[107] Y. Zheng, Z. Yan, K. Chen, J. Sun, Y. Xu, and Y. Liu. "Vulnerability Assessment of Deep Reinforcement Learning Models for Power System Topology Optimization". In: *IEEE Transactions on Smart Grid* 12 (2021), pp. 3613–3623. DOI: 10.1109/tsg.2021.3062700.

[108] G. Chaojun, P. Jirutitijaroen, and M. Motani. "Detecting False Data Injection Attacks in AC State Estimation". In: *IEEE Transactions on Smart Grid* 6.5 (2015), pp. 2476–2483. DOI: 10.1109/tsg.2015.2388545.

[109] C. Konstantinou and M. Maniatakos. "A data-based detection method against false data injection attacks". In: *IEEE Design and Test* 37 (2020), pp. 67–74. DOI: 10.1109/mdat.2019.2952357.

[110] S. Peng, Z. Zhang, R. Deng, and P. Cheng. "Localizing False Data Injection Attacks in Smart Grid: A Spectrum-Based Neural Network Approach". In: *IEEE Transactions on Smart Grid* 14.6 (2023), pp. 4827–4838. DOI: 10.1109/tsg.2023.3261970.

[111] M. Jorjani, H. Seifi, and A. Y. Varjani. "A Graph Theory-Based Approach to Detect False Data Injection Attacks in Power System AC State Estimation". In: *IEEE Transactions on Industrial Informatics* 17.4 (2020), pp. 2465–2475. DOI: 10.1109/tii.2020.2999571.

[112] S. Wei, J. Xu, Z. Wu, Q. Hu, and X. Yu. "A False Data Injection Attack Detection Strategy for Unbalanced Distribution Networks State Estimation". In: *IEEE Transactions on Smart Grid* 14.5 (2023), pp. 3992–4006. DOI: 10.1109/tsg.2023.3235945.

[113] A. Bhattacharjee, A. K. Mondal, A. Verma, S. Mishra, and T. K. Saha. "Deep Latent Space Clustering for Detection of Stealthy False Data Injection Attacks Against AC State Estimation in Power Systems". In: *IEEE Transactions on Smart Grid* 14.3 (2023), pp. 2338–2351. DOI: 10.1109/tsg.2022.3216625.

[114] J. J. Q. Yu, Y. Hou, and V. O. K. Li. "Online False Data Injection Attack Detection With Wavelet Transform and Deep Neural Networks". In: *IEEE Transactions on Industrial Informatics* 14.7 (2018), pp. 3271–3280. DOI: 10.1109/tii.2018.2825243.

[115] B. Li, G. Xiao, R. Lu, R. Deng, and H. Bao. "On Feasibility and Limitations of Detecting False Data Injection Attacks on Power Grid State Estimation Using D-FACTS Devices". In: *IEEE Transactions on Industrial Informatics* 16.2 (2020), pp. 854–864. DOI: 10.1109/tii.2019.2922215.

[116] W. Xia, D. He, and L. Yu. "Locational Detection of False Data Injection Attacks in Smart Grids: A Graph Convolutional Attention Network Approach". In: *IEEE Internet of Things Journal* (2023), pp. 1–1. DOI: 10.1109/jiot.2023.3323565.

[117] S. Nath, I. Akingeneye, J. Wu, and Z. Han. "Quickest Detection of False Data Injection Attacks in Smart Grid with Dynamic Models". In: *IEEE Journal of Emerging and Selected Topics in Power Electronics* 10 (2022), pp. 1292–1302. DOI: 10.1109/jestpe.2019.2936587.

[118] A. Sahu, Z. Mao, P. Wlazlo, H. Huang, K. Davis, A. Goulart, and S. Zonouz. "Multi-Source Multi-Domain Data Fusion for Cyberattack Detection in Power Systems". In: *IEEE Access* 9 (2021), pp. 119118–119138. DOI: 10.1109/access.2021.3106873. eprint: 2101.06897.

[119] Y. Li, Y. Wang, and S. Hu. "Online Generative Adversary Network Based Measurement Recovery in False Data Injection Attacks: A Cyber-Physical Approach". In: *IEEE Transactions on Industrial Informatics* 16.3 (2020), pp. 2031–2043. DOI: 10.1109/tii.2019.2921106.

[120] H. Wang, J. Ruan, B. Zhou, C. Li, Q. Wu, M. Q. Raza, and G. Z. Cao. "Dynamic Data Injection Attack Detection of Cyber Physical Power Systems with Uncertainties". In: *IEEE Transactions on Industrial Informatics* 15 (2019), pp. 5505–5518. DOI: 10.1109/tii.2019.2902163.

[121] S. Chakrabarty and B. Sikdar. "Unified Detection of Attacks Involving Injection of False Control Commands and Measurements in Transmission Systems of Smart Grids". In: *IEEE Transactions on Smart Grid* 13 (2022), pp. 1598–1610. DOI: 10.1109/tsg.2021.3137835.

[122] S. Yang, K.-W. Lao, Y. Chen, and H. Hui. "Resilient Distributed Control Against False Data Injection Attacks for Demand Response". In: *IEEE Transactions on Power Systems* 39.2 (2024), pp. 2837–2853. DOI: 10.1109/tpwrs.2023.3287205.

[123] M. Esmalifalak, L. Liu, N. Nguyen, R. Zheng, and Z. Han. "Detecting Stealthy False Data Injection Using Machine Learning in Smart Grid". In: *IEEE Systems Journal* 11.3 (2017), pp. 1644–1652. DOI: 10.1109/JSYST.2014.2341597.

[124] O. Boyaci, A. Umunnakwe, A. Sahu, M. R. Narimani, M. Ismail, K. R. Davis, and E. Serpedin. "Graph Neural Networks Based Detection of Stealth False Data Injection Attacks in Smart Grids". In: *IEEE Systems Journal* 16.2 (2022), pp. 2946–2957. DOI: 10.1109/JSYST.2021.3109082.

[125] Y. He, G. J. Mendis, and J. Wei. "Real-Time Detection of False Data Injection Attacks in Smart Grid: A Deep Learning-Based Intelligent Mechanism". In: *IEEE Transactions on Smart Grid* 8.5 (2017), pp. 2505–2516. DOI: 10.1109/TSG.2017.2703842.

[126] O. Boyaci, M. R. Narimani, K. R. Davis, M. Ismail, T. J. Overbye, and E. Serpedin. "Joint Detection and Localization of Stealth False Data Injection Attacks in Smart Grids Using Graph Neural Networks". In: *IEEE Transactions on Smart Grid* 13.1 (2022), pp. 807–819. DOI: 10.1109/TSG.2021.3117977.

[127] Y. Yu, C. Liu, L. Xiong, Y. Tang, and F. Qian. "Localization of False Data Injection Attacks in Smart Grids With Renewable Energy Integration via Spatiotemporal Network". In: *IEEE Internet of Things Journal* 11.23 (2024), pp. 37571–37581. DOI: 10.1109/JIOT.2024.3436520.

[128] A. Saad, S. Faddel, T. Youssef, and O. A. Mohammed. "On the Implementation of IoT-Based Digital Twin for Networked Microgrids Resiliency against Cyber Attacks". In: *IEEE Transactions on Smart Grid* 11 (2020), pp. 5138–5150. DOI: 10.1109/tsg.2020.3000958.

[129] S. Wei, Z. Wu, J. Xu, and Q. Hu. "Multiarea Probabilistic Forecasting-Aided Interval State Estimation for FDIA Identification in Power Distribution Networks". In: *IEEE Transactions on Industrial Informatics* (2023), pp. 1–12. DOI: 10.1109/tii.2023.3321098.

[130] J. R. K. Rajasekaran, B. Natarajan, and A. Pahwa. "Modified Matrix Completion-Based Detection of Stealthy Data Manipulation Attacks in Low Observable Distribution Systems". In: *IEEE Transactions on Smart Grid* 14.6 (2023), pp. 4851–4862. DOI: 10.1109/tsg.2023.3266834.

[131] J. Zhang and X. Wang. "Low-Complexity Quickest Change Detection in Linear Systems with Unknown Time-Varying Pre- And Post-Change Distributions". In: *IEEE Transactions on Information Theory* 67 (2021), pp. 1804–1824. DOI: 10.1109/tit.2021.3049468.

[132] C. Liu, Y. Tang, R. Deng, M. Zhou, and W. Du. "Joint Meter Coding and Moving Target Defense for Detecting Stealthy False Data Injection Attacks in Power System State Estimation". In: *IEEE Transactions on Industrial Informatics* 20.3 (2024), pp. 3371–3381. DOI: 10.1109/tii.2023.3306937.

[133] Y. Hu, X. Xian, Y. Jin, and S. Wang. "Fairness-Guaranteed DER Coordination Under False Data Injection Attacks". In: *IEEE Internet of Things Journal* 10.21 (2023), pp. 19043–19053. DOI: 10.1109/jiot.2023.3281582.

[134] H. Pang, K. He, Y. Fu, J.-N. Liu, X. Liu, and W. Tan. "Enabling Efficient and Malicious Secure Data Aggregation in Smart Grid With False Data Detection". In: *IEEE Transactions on Smart Grid* 15.2 (2024), pp. 2203–2213. DOI: 10.1109/tsg.2023.3316730.

[135] H. Li, L. Lai, and W. Zhang. "Communication Requirement for Reliable and Secure State Estimation and Control in Smart Grid". In: *IEEE Transactions on Smart Grid* 2.3 (2011), pp. 476–486. DOI: 10.1109/tsg.2011.2159817.

[136] W. Xu, M. Higgins, J. Wang, I. M. Jaimoukha, and F. Teng. "Blending Data and Physics Against False Data Injection Attack: An Event-Triggered Moving Target Defence Approach". In: *IEEE Transactions on Smart Grid* 14.4 (2023), pp. 3176–3188. DOI: 10.1109/tsg.2022.3231728.

[137] W. Xue and T. Wu. "Active Learning-Based XGBoost for Cyber Physical System Against Generic AC False Data Injection Attacks". In: *IEEE Access* 8 (2020), pp. 144575–144584. DOI: 10.1109/access.2020.3014644.

[138] G. Cheng, Y. Lin, J. Zhao, and J. Yan. "A Highly Discriminative Detector Against False Data Injection Attacks in AC State Estimation". In: *IEEE Transactions on Smart Grid* 13.3 (2022), pp. 2318–2330. DOI: 10.1109/tsg.2022.3141803.

[139] R. Zeng, Y. Cao, Y. Li, S. Hu, X. Shao, L. Xie, L. Hou, L. Zhao, and M. Shahidehpour. "A General Real-Time Cyberattack Risk Assessment Method for Distribution Network Involving the Influence of Feeder Automation System". In: *IEEE Transactions on Smart Grid* 15.2 (2024), pp. 2102–2115. DOI: 10.1109/tsg.2023.3302287.

[140] F. Almutairy, L. Scekic, R. Elmoudi, and S. Wshah. "Accurate Detection of False Data Injection Attacks in Renewable Power Systems Using Deep Learning". In: *IEEE Access* 9 (2021), pp. 135774–135789. DOI: 10.1109/access.2021.3117230.

[141] C. Chen, Y. Wang, M. Cui, J. Zhao, W. Bi, Y. Chen, and X. Zhang. "Data-Driven Detection of Stealthy False Data Injection Attack Against Power System State Estimation". In: *IEEE Transactions on Industrial Informatics* 18.12 (2022), pp. 8467–8476. DOI: 10.1109/tii.2022.3149106.

[142] Z. Cheng and M. Y. Chow. "Resilient Collaborative Distributed AC Optimal Power Flow Against False Data Injection Attacks: A Theoretical Framework". In: *IEEE Transactions on Smart Grid* 13 (2022), pp. 795–806. DOI: 10.1109/tsg.2021.3113287.

[143] Z. Qu, J. Yang, Y. Wang, and P. M. Georgievitch. "Detection of False Data Injection Attack in Power System Based on Hellinger Distance". In: *IEEE Transactions on Industrial Informatics* 20.2 (2024), pp. 2119–2128. DOI: 10.1109/tii.2023.3286895.

[144] M. Du, X. Liu, Z. Li, and H. Lin. "Robust Mitigation Strategy Against Dummy Data Attacks in Power Systems". In: *IEEE Transactions on Smart Grid* 14.4 (2023), pp. 3102–3113. DOI: 10.1109/tsg.2022.3225469.

[145] S. Nasiri, H. Seifi, and H. Delkhosh. "A Secure Power System Distributed State Estimation via a Consensus-Based Mechanism and a Cooperative Trust Management Strategy". In: *IEEE Transactions on Industrial Informatics* 20.2 (2024), pp. 3002–3014. DOI: 10.1109/tii.2023.3299385.

[146] S. Bi and Y. J. Zhang. "Graphical Methods for Defense Against False-Data Injection Attacks on Power System State Estimation". In: *IEEE Transactions on Smart Grid* 5.3 (2014), pp. 1216–1227. DOI: 10.1109/tsg.2013.2294966.

[147] T. S. Sreeram and S. Krishna. "Managing False Data Injection Attacks during Contingency of Secured Meters". In: *IEEE Transactions on Smart Grid* 10 (2019), pp. 6945–6953. DOI: 10.1109/tsg.2019.2914974.

[148] H. Alamro, K. Mahmood, S. S. Aljameel, A. Yafoz, R. Alsini, and A. Mohamed. "Modified Red Fox Optimizer With Deep Learning Enabled False Data Injection Attack Detection". In: *IEEE Access* 11 (2023), pp. 79256–79264. DOI: 10.1109/access.2023.3298056.

[149] E. Naderi and A. Asrari. "A Deep Learning Framework to Identify Remedial Action Schemes Against False Data Injection Cyberattacks Targeting Smart Power Systems". In: *IEEE Transactions on Industrial Informatics* 20.2 (2024), pp. 1208–1219. DOI: 10.1109/tii.2023.3272625.

[150] L. Liu, M. Esmalifalak, Q. Ding, V. A. Emesih, and Z. Han. "Detecting False Data Injection Attacks on Power Grid by Sparse Optimization". In: *IEEE Transactions on Smart Grid* 5.2 (2014), pp. 612–621. DOI: 10.1109/tsg.2013.2284438.

[151] A. Ashok, M. Govindarasu, and V. Ajjarapu. "Online Detection of Stealthy False Data Injection Attacks in Power System State Estimation". In: *IEEE Transactions on Smart Grid* 9.3 (2018), pp. 1636–1646. DOI: 10.1109/tsg.2016.2596298.

[152] C. Liu, H. Liang, T. Chen, J. Wu, and C. Long. "Joint Admittance Perturbation and Meter Protection for Mitigating Stealthy FDI Attacks Against Power System State Estimation". In: *IEEE Transactions on Power Systems* 35.2 (2020), pp. 1468–1478. DOI: 10.1109/tpwrs.2019.2938223.

[153] J. Zhao, L. Mili, and M. Wang. "A Generalized False Data Injection Attacks Against Power System Nonlinear State Estimator and Countermeasures". In: *IEEE Transactions on Power Systems* 33.5 (2018), pp. 4868–4877. DOI: 10.1109/tpwrs.2018.2794468.

[154] Z. Liu and L. Wang. "Defense Strategy against Load Redistribution Attacks on Power Systems Considering Insider Threats". In: *IEEE Transactions on Smart Grid* 12 (2021), pp. 1529–1540. DOI: 10.1109/tsg.2020.3023426.

[155] A. Khaleghi, M. S. Ghazizadeh, and M. R. Aghamohammadi. "A Deep Learning-Based Attack Detection Mechanism Against Potential Cascading Failure Induced by Load Redistribution Attacks". In: *IEEE Transactions on Smart Grid* 14.6 (2023), pp. 4772–4783. DOI: 10.1109/tsg.2023.3256480.

[156] H. Goyel and K. S. Swarup. "Data Integrity Attack Detection Using Ensemble-Based Learning for Cyber-Physical Power Systems". In: *IEEE Transactions on Smart Grid* 14.2 (2023), pp. 1198–1209. DOI: 10.1109/tsg.2022.3199305.

[157] M. B. D. C. Filho, J. S. d. Souza, J. D. Glover, V. B. Flr, A. Nishio, R. Lima, J. V. Daibes, L. Quintanilha, A. Coimbra, W. Soares, and H. Carneiro. "Educational Tool for Power System State Estimation Teaching and Learning". In: *IEEE Transactions on Power Systems* 38.6 (2023), pp. 5885–5895. DOI: 10.1109/tpwrs.2022.3229908.

[158] M. Higgins, F. Teng, and T. Parisini. “Stealthy MTD against Unsupervised Learning-Based Blind FDI Attacks in Power Systems”. In: *IEEE Transactions on Information Forensics and Security* 16 (2021), pp. 1275–1287. DOI: 10.1109/tifs.2020.3027148.

[159] A. Shefaei, M. Mohammadpourfard, B. Mohammadi-Ivatloo, and Y. Weng. “Revealing a New Vulnerability of Distributed State Estimation: A Data Integrity Attack and an Unsupervised Detection Algorithm”. In: *IEEE Transactions on Control of Network Systems* 9 (2022), pp. 706–718. DOI: 10.1109/tcns.2021.3091631.

[160] H. Wang, J. Ruan, G. Wang, B. Zhou, Y. Liu, X. Fu, and J. Peng. “Deep Learning-Based Interval State Estimation of AC Smart Grids Against Sparse Cyber Attacks”. In: *IEEE Transactions on Industrial Informatics* 14.11 (2018), pp. 4766–4778. DOI: 10.1109/tii.2018.2804669.

[161] I. Zografopoulos, N. D. Hatziargyriou, and C. Konstantinou. “Distributed Energy Resources Cybersecurity Outlook: Vulnerabilities, Attacks, Impacts, and Mitigations”. In: *IEEE Systems Journal* 17.4 (2023), pp. 6695–6709. DOI: 10.1109/jsyst.2023.3305757. eprint: 2205.11171.

[162] T. S. Sreeram and S. Krishna. “Graph-Based Assessment of Vulnerability to False Data Injection Attacks in Distribution Networks”. In: *IEEE Transactions on Power Systems* 39.2 (2024), pp. 4510–4520. DOI: 10.1109/tpwrs.2023.3309777.

[163] Z. Cheng, H. Ren, J. Qin, and R. Lu. “Security Analysis for Dynamic State Estimation of Power Systems with Measurement Delays”. In: *IEEE Transactions on Cybernetics* 53 (2023), pp. 2087–2096. DOI: 10.1109/tcyb.2021.3108884.

[164] C. Cameron, C. Patsios, P. C. Taylor, and Z. Pourmirza. “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 (2018), pp. 3010–3019. DOI: 10.1109/tsg.2018.2817046.

[165] X. Kong, Z. Lu, X. Guo, J. Zhang, and H. Li. “Resilience Evaluation of Cyber-Physical Power System Considering Cyber Attacks”. In: *IEEE Transactions on Reliability* 73.1 (2024), pp. 245–256. DOI: 10.1109/tr.2023.3294264.

[166] C. Pei, Y. Xiao, W. Liang, and X. Han. “PMU Placement Protection Against Coordinated False Data Injection Attacks in Smart Grid”. In: *IEEE Trans. Ind. Appl.* 56.4 (2020), pp. 4381–4393. DOI: 10.1109/TIA.2020.2979793.

[167] Q. Yang, L. Jiang, W. Hao, B. Zhou, P. Yang, and Z. Lv. “PMU Placement in Electric Transmission Networks for Reliable State Estimation Against False Data Injection Attacks”. In: *IEEE Internet Things J.* 4.6 (2017), pp. 1978–1986. DOI: 10.1109/JIOT.2017.2769134.

[168] X. Jiang, J. Zhang, B. J. Harding, J. J. Makela, and A. D. Domı´nguez-Garcı´a. “Spoofing GPS Receiver Clock Offset of Phasor Measurement Units”. In: *IEEE Transactions on Power Systems* 28.3 (2013), pp. 3253–3262. DOI: 10.1109/TPWRS.2013.2240706.

[169] C. Liu, R. Deng, W. He, H. Liang, and W. Du. “Optimal Coding Schemes for Detecting False Data Injection Attacks in Power System State Estimation”. In: *IEEE Transactions on Smart Grid* 13.1 (2022), pp. 738–749. DOI: 10.1109/TSG.2021.3107972.

[170] D. Wang, J. Huang, Y. Tang, and F. Li. “A Watermarking Strategy Against Linear Deception Attacks on Remote State Estimation Under K–L Divergence”. In: *IEEE Trans. Ind. Inform.* 17.5 (2021), pp. 3273–3281. DOI: 10.1109/TII.2020.3009874.

[171] S. C. Dimoulias, G. C. Kryonidis, K.-N. D. Malamaki, E. O. Kontis, F. P. Fotellis, A. N. Milioudis, and E. Romero-Ramos. “Droop-Control-Aided State Estimation in Active Distribution Systems”. In: *2024 3rd International Conference on Energy Transition in the Mediterranean Area (SyNERGY MED)*. 2024, pp. 1–5. DOI: 10.1109/SyNERGYMED62435.2024.10799301.

[172] F. Milano and A. Gómez-Expósito. “Detection of Cyber-Attacks of Power Systems Through Benford’s Law”. In: *IEEE Transactions on Smart Grid* 12.3 (2021), pp. 2741–2744. DOI: 10.1109/TSG.2020.3042897.