A Taxonomy of Cyber Defence Strategies Against False Data Attacks in Smart Grids

HAFTU TASEW REDA, La Trobe University, Australia and Deakin University, Australia
ADNAN ANWAR, Deakin University, Australia
ABDUN NASER MAHMOOD, La Trobe University, Australia
ZAHIR TARI, RMIT University, Australia

The modern electric power grid, known as the Smart Grid, has fast transformed the isolated and centrally controlled power system to a fast and massively connected cyber-physical system that benefits from the revolutions happening in communications (such as 5G/6G) and the fast adoption of Internet of Things devices (such as intelligent electronic devices and smart meters). While the synergy of a vast number of cyber-physical entities has allowed the Smart Grid to be much more effective and sustainable in meeting the growing global energy challenges, it has also brought with it a large number of vulnerabilities resulting in breaches of data integrity, confidentiality, and availability. False data injection (FDI) appears to be among the most critical cyberattacks and has been a focal point of interest for both research and industry. To this end, this article presents a comprehensive review of the recent advances in defence countermeasures of FDI attacks on the Smart Grid. Relevant existing works of literature are evaluated and compared in terms of their theoretical and practical significance to Smart Grid cybersecurity. In conclusion, a range of technical limitations of existing false data attack detection research is identified, and a number of future research directions are recommended.

CCS Concepts: • Security and privacy → Systems security; • Computing methodologies → Artificial intelligence; Machine learning; Distributed computing methodologies; • Networks → Network types; Cyber-physical networks

Additional Key Words and Phrases: Communication system, cyberattack, cyber-physical system, cybersecurity, detection, false data injection, Internet of Things, power system, smart grid

ACM Reference format:
Haftu Tasew Reda, Adnan Anwar, Abdun Naser Mahmood, and Zahir Tari. 2023. A Taxonomy of Cyber Defence Strategies Against False Data Attacks in Smart Grids. *ACM Comput. Surv.* 55, 14s, Article 331 (July 2023), 37 pages.
https://doi.org/10.1145/3592797
1 INTRODUCTION

Energy is the backbone of our economic growth and is a super-critical resource on which all other national critical infrastructure sectors rely. A significant rise in threats to critical infrastructure from nation states and malicious actors poses real challenges to the understanding of operational vulnerabilities in Smart Grids as well as the different attack vectors that may jeopardize the stability and efficiency of the power system. The 2020 Global Risks report by the World Economic Forum\footnote{World Economic Forum, The Global Risks Report 2020, 15th Edition. Available: \url{http://www3.weforum.org/docs/WEF_Global_Risk_Report_2020.pdf}.} indicates that large-scale cyberattacks against critical infrastructure and networks is the top threat and will continue to be among the most likely global threats over the next 10 years.

According to vulnerability reports from the US ICS-CERT\footnote{NCCIC, ICS-CERT Year in Review (2016). Available: \url{https://us-cert.cisa.gov/sites/default/files/Annual_Reports/Year_in_Review_FY2016_Final_SS08C.pdf}.} and Kaspersky ICS-CERT,\footnote{Kaspersky ICS CERT, Threat landscape for industrial automation systems. Available: \url{https://ics-cert.kaspersky.com/media/KASPERSKY_H22019_ICS_REPORT_FINAL_EN.pdf}.} the energy sector has reported the largest number of vulnerabilities among all critical infrastructures. For example, Figure 1 shows the number of vulnerabilities of various Industrial Control System (ICS) elements between 2010 and 2019. Accordingly, 178, 110, and 283 cyberattack incidents were recorded in the energy sector out of 322, 415, and 509 ICS cyberattack incidents, respectively, across the fiscal years 2017, 2018, and 2019. These cyber incidents may lead to myriad security risks, including the loss of critical data necessary for control operations and malicious modification/deletion of critical power system states. Possible consequences include incorrect customer billing information, price manipulation in the energy market, small- to large-scale electric power outage, and the likelihood of endangering lives by limiting power to other national critical infrastructures.

This article discusses various state-of-the-art false data injection (FDI)\footnote{FDI [1]–[3]} defence countermeasures in Smart Grids.

1.1 Purpose and Scope of the Study

FDI\footnote{1} is a critical malicious cyberattack in which power system state estimation (SE) outcomes are undermined by a deliberate injection of data into meter measurements. Following SE, bad data detection (BDD) techniques\footnote{Extensive studies on potential FDI attacks have enabled Smart Grid operators to set up a range of defence mechanisms. The primary goal of this study is to conduct a systematic literature review of FDI attacks on Smart Grids. This work discusses the different FDI attack models that target various cyber-physical components and proposes a taxonomy for cyber defence strategies against FDI attacks.} are used to detect any injected bad data by determining the residual of the original measurement vector. However, it has been proven that the BDD techniques are incapable of detecting FDI attacks [1–3].

Extensive studies on potential FDI attacks have enabled Smart Grid operators to set up a range of defence mechanisms. The primary goal of this study is to conduct a systematic literature review of FDI attacks on Smart Grids. This work discusses the different FDI attack models that target various cyber-physical components and proposes a taxonomy for cyber defence strategies against FDI attacks.

1.2 Contributions

This article has analysed related and recent publications and reference materials in the mitigation techniques of false data attacks across various domains of Smart Grid infrastructures. We systematically search for older and more recent related literature, analyse the main findings covered in each study, critically evaluate them, and compare each solution within the broader conception of the cyber-physical data integrity attacks. Major contributions of this article are summarised here.

(1) The article identifies essential cybersecurity requirements of Smart Grids (Section 4), including a theoretical analysis and requirements for stealthy FDI attacks (Section 5.1),
characteristics and comparisons of the different FDI attacks (Section 5.2), and vulnerable critical Smart Grid components to FDI attacks (Section 5.3).

(2) After a thorough review of relevant existing survey papers, this work highlights their contributions and identifies the gaps that have been addressed through this survey. Detailed comparisons have been highlighted in Table 1 and related discussions have been presented in Section 6.

(3) The article comprehensively analyses defence categories (Section 6) and compiles a list of methodologies, including statistical, signal processing, and advanced artificial intelligence (AI)–based methodologies that can be used to detect and prevent FDI attacks.

(4) In addition, this article analyses various countermeasure methods and provides statistical facts on the basis of the evaluation criteria in Section 8. The article discusses main research gaps in the existing papers in Section 9.

(5) Finally, this article provides technical recommendations for emerging advanced application areas, including Internet of Things (IoT)–based Advanced Metering Infrastructure (AMI), cognitive radio, lightweight machine learning (ML) for resource-constrained IoT devices, distributed attack detection in edge computing environment, and Blockchain-based defence for privacy preservation in the Smart Grid.

1.3 Outline of the Article

Section 2 discusses related works on defence countermeasures to false data attacks and compares them with our work. Background on Smart Grid and key cyber-physical elements is presented in Section 3. Cyber-physical attacks, Cybersecurity main objectives, and security requirements of Smart Grid are highlighted in Section 4. In Section 5, we comprehensively discuss the nature of FDI attacks, including main requirements, characteristics of different classes of FDI attacks, and vulnerable targets in the Smart Grid environment. Section 6 presents the suggested taxonomy and defence strategies against false data attacks that are critical frameworks for the power system operator and other stakeholders. Literature search methodology, selection and analysis of the surveyed literature, and evaluation criteria among the multitude of algorithms of selected surveyed papers are presented in Section 7. We compare and contrast the numerous defence strategies in Section 8. Following a critical review of the shortcomings found in the literature in Section 9, our technical recommendations that can substantiate future research in the field are provided in Section 10. Section 11 concludes this survey article.
2 RELATED WORKS

The article by Guan et al. [7] is one of the earliest works to present a comprehensive survey of FDI attacks. The authors present an overview of detection and defence schemes based on centralised and distributed SE techniques. Deng et al. [8] survey data injection attacks with respect to three major cybersecurity aspects: FDI attack construction, impacts of the attacks, and countermeasures. Unlike previous studies, Deng et al. [8] thoroughly study the impacts of data injection attacks on the electricity market. Another line of survey research is studied in [9], which summarises related literature on different attack models, economic impacts of the attack, and mitigation techniques for various Smart Grid domains, including transmission, distribution, and microgrid networks. Liang et al. [10] complement previous studies in discussing various FDI attack models, physical and economic impacts of the attacks, and countermeasures in Smart Grids. The authors of [11] and [12] also comprehensively discuss FDI attacks from the attacker’s and operator’s points of view along with consequences of the attacks.

In contrast to previous surveys, the authors of [13] reviewed two main classes of detection algorithms: model-based and data-driven, and have discussed the benefits and drawbacks of each technique. As compared with other review works, which mostly focus on the energy management system (EMS), the authors of [14] discuss FDI attacks on various entities of online power system security. They review and compare studies on FDI attacks and provide a new class of cyber-oriented countermeasure: prevention (further classified into block chain- and cryptography-based techniques).

Unlike the related works, this article presents a detailed survey of recent developments in FDI and sets out a taxonomy of the incumbent cyberattack with respect to defence strategies across every Smart Grid domain, including transmission to consumption, automatic generation control (AGC) to microgrids/distributed energy resources (DERs), and substations to wide area monitoring systems. The IoT, cognitive radios [15], and software-defined networks have recently been introduced as enablers to the Smart Grid. These communication technologies are very important to address the cybersecurity aspects of today’s Smart Grids, which were missed in most of the existing related works. This article provides an in-depth survey of the latest advances of the defence measures against cyber-physical FDI attacks within the Smart Grid infrastructure. Table 1 summarises the comparison of existing survey papers and this article.

3 BACKGROUND

The Smart Grid is the convergence of various cyber and physical components of the electrical power domain. It is an evolution of the electrical power system that is co-engineered through expertise from different fields such as (operational technology OT; physical power devices, data acquisitions, control systems, and industrial automation), (information technology IT; decision-making and human interfaces), advanced (information and communication technology ICT; infrastructure), and cybersecurity to be more effective and sustainable in meeting the growing global energy challenges. As compared with the traditional power grid, the Smart Grid provides an end-to-end system of two-way electricity flow in which customers can not only utilize energy but can also feed energy back into the grid. The Smart Grid also supports a wide variety of energy sources (including renewables, which are key to low-carbon emissions). According to conceptual model of the National Institute of Standards and Technology (NIST) [16], the Smart Grid comprises seven interconnected application domains: generation, transmission, distribution, customer, market, service provider, and operator. While applications in every Smart Grid domain are critical to the scalability, efficiency, and stability of power system operations, they also introduce vulnerabilities to the Smart Grid. The probability of a successful breach is inevitable for all cyber-physical
### Table 1. Comparison of Current Survey Articles and Our Work

| Comparison attributes                                      | Literature                  | Our paper |
|-----------------------------------------------------------|----------------------------|-----------|
| **Defence based on SE type**                              |                            |           |
| Conventional BDD                                          | ×  | ✓  | ✓  | ✓  | ✓  | ✓  | ✓  | ✓  |
| Detection based on dynamic SE                             | ×  | ×  | ✓  | ×  | ✓  | ✓  | ✓  | ✓  |
| Optimal phasor measurement unit placement                 | ×  | ✓  | ✓  | ✓  | ✓  | ✓  | ✓  | ✓  |
| Optimal measurement selection                             | ×  | ✓  | ×  | ✓  | ×  | ×  | ✓  | ✓  |
| **Protection-based defence**                              |                            |           |
| Grid topology perturbation                                 | ×  | ×  | ✓  | ✓  | ×  | ×  | ✓  | ✓  |
| Generalised likelihood ratio (GLR) test detector          | ×  | ✓  | ×  | ×  | ×  | ×  | ✓  | ✓  |
| **Statistical-based detection**                           |                            |           |
| Bayesian test detector                                    | ×  | ✓  | ×  | ×  | ×  | ×  | ✓  | ✓  |
| Quickest change detector                                  | ×  | ×  | ✓  | ×  | ✓  | ✓  | ✓  | ✓  |
| Statistical distance                                      | ×  | ✓  | ×  | ×  | ✓  | ✓  | ✓  | ✓  |
| Sparse matrix recovery                                    | ×  | ✓  | ×  | ✓  | ×  | ×  | ✓  | ✓  |
| **Data-driven detection**                                 |                            |           |
| Supervised machine learning                               | ×  | ×  | ×  | ×  | ✓  | ✓  | ✓  | ✓  |
| Semi-supervised machine learning                          | ×  | ×  | ×  | ×  | ✓  | ✓  | ✓  | ✓  |
| Deep learning                                             | ×  | ×  | ×  | ×  | ✓  | ✓  | ✓  | ✓  |
| **Prevention-based defence**                              |                            |           |
| Cryptographic-based prevention                             | ×  | ×  | ×  | ×  | ✓  | ✓  | ✓  | ✓  |
| Blockchain-based prevention                               | ×  | ×  | ×  | ×  | ×  | ×  | ✓  | ✓  |
| **Evaluation criteria**                                   |                            |           |
| Duration of surveyed papers                               | 2009 to 2013               | 2019 to 2020 |
| Future directions                                          | ✓  | ✓  | ✓  | ✓  | ✓  | ✓  | ✓  | ✓  |

✓ studied/covered, ❄ partially studied, × not studied.

systems directly or indirectly linked to the Smart Grid. Hence, it is crucial to scrutinize the relations between the physical model and the cyber system in order to provide a resilient cyber and communications infrastructure in the Smart Grid environment. Therefore, in this section, we briefly discuss the main cyber-physical elements of the Smart Grid.

### 3.1 SCADA

Supervisory control and data acquisition (SCADA) [17] is an industrial and power system control application. Usually a SCADA system consists basically of three subsystems: a data acquisition subsystem that collects measurements of the power system, a supervisory subsystem that has the ability to control remote intelligent electronic devices (IEDs) by transmitting control commands (such as to close/open a circuit breaker, change transformer tap settings, lower/raise generator output, etc.), and a communication subsystem that interconnects the data acquisition subsystem to the supervisory subsystem. A typical scenario in the integrated SCADA system is as follows: when the SCADA system gathers data from diverse IEDs in a power system through various communication methods (such as Internet protocol [IP]-based wide area networks, local area networks, and software-defined networking), and then controls/monitors the data using different visualisation tools.

### 3.2 Energy Management System

Power system operations are overseen by system operators from a control center. Within the control center lies the EMS, an automated system used to monitor, control, coordinate, and optimize energy data performance in real time. The EMS depends on a SCADA system for its data monitoring and analysis events. Distribution grids (usually from substation to consumption side) are controlled via a distribution energy management system (DEMS).

A typical EMS comprises the following functional elements: state estimation (SE), optimal power flow, contingency analysis (CA), an alarm management system, planning and operations, automatic generation control, and economic dispatch. Generally, a physical power grid can be considered as a set of buses, transmission lines, loads, generators, and shunt components. Each of the
Fig. 2. Typical Smart Grid architecture.

buses or nodes are physically interconnected by lines or branches. Figure 2 is a typical Smart Grid architecture. It illustrates how the EMS and other ICT/OT components communicate with the physical power grid. An intruder can compromise the power system measurements circumvent the EMS.

3.3 Smart Grid Communication Systems
Communication systems are essential to the efficient operation of the Smart Grid. Various communication technologies are utilised across the different domains: for example, IEC 61850 in the substation automation system (SAS), PMU in the wide area monitoring system (WAMS), AMI across the customer side, and networked control system (NCS) between sensors, actuators, and controllers.

3.4 Distributed Energy Resources
DERs are decentralised, versatile, and modular architecture that incorporate a number of renewable sources, including solar, wind, and geothermal. Compared with conventional approaches, in which energy is generated by centralised and big power plants, DERs now allow energy production and delivery from many areas, including millions of homes and businesses. Microgrid technology is one of the enablers of Smart Grids that provides smooth collaboration between DERs offering isolation options (also known as islanding) or access to conventional grid electricity.

4 CYBER-PHYSICAL SECURITY OF SMART GRIDS
The security issues of Smart Grids have emerged from both physical and cyber spaces: physical security [16] (security policies with respect to personnel, physical equipment protection, and contingency analysis), cybersecurity (focusing on the information security of Smart Grids pertaining to IT, OT, and network and communication systems), and cyber-physical security (incorporating strength in all physical and cybersecurity measures against inadvertent cyber-physical incidents within an integrated Smart Grid framework). In this section, Smart Grid cybersecurity goals, cybersecurity requirements, and cyber-physical attacks are highlighted.

4.1 Smart Grid Cybersecurity Goals
Quality of service and secure power supply are the primary concern of power companies and industrial sectors. The goal of Smart Grid utilization is to build a much more efficient and reliable energy source; cybersecurity threats can inevitably slow down progress toward this goal. The Smart Grid needs to ensure basic security goals, such as data integrity, availability, confidentiality,
accountability, and more for the various cyber-physical elements. While these security principles have been developed to govern policies on generic information security within organisations, the principles of Smart Grid cybersecurity have also been identified by the NIST [16].

Availability ensures timeliness of electricity in our day-to-day life, which is by far the most critical security goal within the Smart Grid environment. Availability can be quantified in terms of latency, the time required for data to be transmitted across the Smart Grid. Smart Grid cybersecurity solutions should provide acceptable latency thresholds for various applications by minimising detrimental effects on availability. Integrity is the second, yet highly critical, Smart Grid security requirement, which ensures that data should not be altered without authorized access, the source of data needs to be verified, the timestamp linked with the data must be identified/validated, and quality of service is in the acceptable range. Confidentiality seems to be the least important as compared with availability and integrity. Nevertheless, the proliferation of smart meters and AMIs across the Smart Grid implies the increasing importance of confidentiality to prevent unauthorized disclosure of information, and to preserve customer privacy or proprietary information. Another security objective within the Smart Grid ecosystem is accountability, a requirement that consumers should be responsible for the actions they take. Accountability is very important, particularly when customers obtain their billing information from the utility center. At that time, they can have sufficient evidence to prove the total power load that they have used.

4.2 Smart Grid Security Requirements
The dynamics of cyber-physical interaction in the Smart Grid poses extrinsic system dependencies. Further, the open interconnectivity of the Smart Grid with the Internet brings various security challenges. Therefore, the Smart Grid requires stringent holistic security solutions to uphold the security objectives discussed earlier and to provide salient features within the Smart Grid infrastructure. First, the security solutions need to be robust enough to counteract against increasing security breaches that can lead to loss of data availability, data integrity, and data confidentiality. The operation of a power system should continue 24/7 regardless of cyber incident, maintaining the power grid reliability (consistent to data availability and to almost 99.9% [16] of data integrity across the power system), and ensuring consumer privacy. Second, resilient cyber-physical operations are required. According to the NIST’s recommendation [16], cybersecurity in critical infrastructure such as the Smart Grid can adopt a comprehensive security framework containing five main features. These include identifying risks or cyber incidents, providing protective mechanisms against the impact of a potential cybersecurity event, providing defence mechanisms to allow prompt discovery of security breaches, appropriate response to minimise the effect of the incident, and recovery plans to restore any systems that have been disrupted due to cyber incidents. Moreover, as attacks from cyber criminals on the power grid continue to rise in complexity and frequency, it is inevitable that various parts of the Smart Grid are vulnerable to incumbent attacks. Therefore, it is required to provide strong attack defence across the EMS and to deploy secure communication protocols.

5 FALSE DATA INJECTION ATTACKS
This section discusses the main requirements for stealthy FDI attacks, the various attack models that fall in this category of cyberattack, and cyber-physical components that can be potentially targeted by the incumbent cyberattack.

5.1 Requirements for Stealthy FDI Attacks
The requirements of FDI attacks vary depending on the application domain. For instance, in wireless sensor networks (WSNs), the underlying wireless communication and broadcast channels
between nodes increase vulnerability to adversaries who can intrude on all traffic, inject bad data reports containing incorrect sensor readings, or even reduce the nodes’ already scarce energy capacity. In contrast, in the Smart Grid, it is difficult for an intruder to access the network parameters. Hence, a more intelligent approach is necessary to launch a successful attack. In general, FDI attacks create strong requirements from the perspectives of both attackers and system operators. Some of the main requirements for FDI attacks in Smart Grid security are discussed next.

5.1.1 Sparsity of FDI Attack. During an FDI attack, the adversary’s goal is to introduce an attack vector $\mathbf{a}$ (where $\mathbf{a} = [a_1, a_2, \ldots, a_m]^T = \mathbf{Hb}$ and $\mathbf{b} = [b_1, b_2, \ldots, b_n]^T$ denotes the estimated error vector injected by the adversary) into the measurements without being noticed by the operator. Usually $\mathbf{a}$ is assumed as a linear combination of the columns of $\mathbf{H}$ [1]. However, the adversary’s control can be limited to only over a few measurement devices. It could be because either the system has secure measurement devices that the attacker cannot access or the attacker has limited physical access to the devices. This results in a sparse FDI attack [1]. An FDI attack designed with only a few non-zero components is called a sparse attack and only a small number of devices (let us say $k$) are required to launch the attack. Let the attack $\mathcal{A} = (\mathbf{a}, k)$ contain the attack vector $\mathbf{a}$ and $k$ sets of compromised meters. Then, the sparse attack [1] with $||\mathbf{a}||_0 \leq k$ can be defined as an $\ell_0$-norm minimization problem and can be given as $\mathbf{a} = \begin{cases} \mathbf{Hb}^i, & \text{for } i \in \{1, \ldots, k\} \\ 0, & \text{for } i \notin \{1, \ldots, k\}. \end{cases}$ where the injected vector $\mathbf{b}^i$ is given by $\mathbf{b}^i = [0, \ldots, 0, b_0^i, 0, \ldots, 0]^T$.

5.1.2 Rendering Power System Unobservability [1]. Through the injection of false data, an adversary can hijack and compromise the power system measurements which results in the system unobservability. Typically, the attacker can remain undetectable at the control center while resulting in incorrect decisions of the state estimator. Even if the cyberattack can be detected by the SE, part of the power network may become unobservable where the SE cannot determine the system states.

5.1.3 Partial-Parameter-Information. Earlier studies on FDI attack models are based on the premise that adversaries are capable of getting complete information on system topology. The authors of [18] assert that it is also possible to construct stealthy attacks based on partial network information. However, attacks based on partial information require satisfying observability criteria. Another research direction ensures that stealthiness (i.e., undetectability) of FDI attacks can also be modeled through data-driven or other partial-parameter-information approaches.

5.1.4 Minimal Attack Vectors. For many reasons, the adversary’s control can be limited to a few measurement devices. For this reason, stealthy FDI attacks should be designed with a very small attack magnitude and with only a few non-zero components (i.e., attack sparsity) [1]. Consequently, the attacker would need to compromise only a small set of devices to cause network unobservability.

5.1.5 Attack Specificity. Depending on the motives of a cyber criminal, the strategy behind the attack may be either indiscriminate or targeted. The scope and impact of these two adversarial approaches are different. The indiscriminate attack may not require a specific knowledge of network devices and may be launched arbitrarily against random Smart Grid elements. The targeted attack would require a sophisticated approach that can be launched against targeted nodes.

5.1.6 Requirement for the Influence of an Attack. Attackers can approach in various ways to launch a successful attack and to cause a security risk on the Smart Grid. Some attackers want to exploit the data collected from sensors and networked devices across the power system. They may
intend to exploit the weaknesses of sensors and communication protocols and launch the attack vector. Some typical examples of attack scenarios can be attacks against sensor measurements: tampering with power system parameter values in remote terminal units (RTUs) and PMUs. Another example can be leveraging communication protocols, in which remote tripping injection can be performed by adversaries. In addition, attackers can infiltrate AMI-based communications networks in order to tamper with the contents of customer data, which can result in disorder of the SE and other EMS functionalities. Others may intend to directly falsify the outcome of the state estimators [1].

5.1.7 Requirement Based on Security Violations. Some FDI-based malicious attackers try to infringe data availability, some violate data integrity, and others target data confidentiality. Loss of data integrity: For example, by injecting systematically generated false data, a cyber intruder may compromise the integrity of the SE by hijacking a subset of meters and returning modified data. The modification may involve deletion of data from the original meter readings, addition of bad data to sensor readings, or alteration of values in the hijacked measurements. The majority of FDI attacks, including, but not limited to, those described in [1, 19, 20] are based on this type of security violation. Loss of data availability: FDI attacks can compromise the availability of critical information that is either intended to disrupt the power system or to stop its availability by shutting down network and communication devices [7, 8]. Attack on confidentiality: Although the effect of FDI on data confidentiality ranks among the least of all security objectives, the injection of false data could also violate the privacy of customers, especially in AMIs of the Smart Grid. This has become common these days as illustrated in [21, 22].

5.1.8 Requirements Based on Attack Impact on the Power System. Threat actors can exploit Smart Grid security vulnerabilities that may lead to malfunctions in energy systems, operational failures in communications equipment as well as physical devices, and may even trigger a cascading failure. According to a report by the NIST [16], three potential impact levels — low, moderate, and high — have been assessed for each of the Smart Grid security objectives based on the degree of adversarial effect and associated risk level. The ultimate aim of FDI adversarial strategies is to have significant impact on the Smart Grid, such as causing sequential transmission line outages, maximizing operation costs of the system by injecting falsified vectors into a subset of targeted meters, culminating in large-scale failure of the power system operation.

5.2 FDI-Based Attack Models

Adversaries employ a variety of FDI attack strategies, with the end goal of breaching the cyber-physical infrastructure of the Smart Grid. Although some adversarial approaches necessitate complete network data and topological configurations, others require limited resources only. Data-driven strategies are also employed to build stealthy FDI attacks. In this research, different FDI attacks are classified based on power flow model, network architecture, and adversarial construction method. These are discussed in this subsection. Figure 3 summarizes the different FDI attack classes.

5.2.1 Classification of Threat Models Based on Power Flow Model. Most FDI attack studies are conducted in a constrained environment on the basis that the functions from power system states to measurements are linear (DC-based power flow models) whereas most industry standard state estimators are based on a nonlinear AC power flow model. One of the pioneer FDI attacks under the DC model is proposed by Liu et al. [1]. Since then, other similar lines of research have been conducted, including what is covered in [23, 24]. In most situations, the study of AC power flow models has to be accompanied by solving complete nonlinear power flow equations, which
are involved in the nonlinear models. Consequently, the complexity of analysis must be reduced and the nonlinear constraints completely ignored while modeling the cyberattacks. While most of the FDI techniques available in the literature rely on the simplified DC state estimators, such techniques are not valid with AC-based SEs.

5.2.2 Classification of Threat Models Based on Network Architecture. In general, the operation of a Smart Grid depends upon the availability of information from hierarchically distributed cyber-physical elements and the outcomes of the central control center. It is important to investigate the FDI attacks from the viewpoint of network architecture: centralised and distributed. Centralised FDI attacks target the centralised state estimator. Once the adversary manipulates the measurement reports sent from different communication devices to the control center, the SE fails to estimate the optimal system states, which further affects other functional elements such as optimal power flow, economic dispatch, and CA that rely on the SE outcome. A great many of the FDI attack construction methodologies are introduced using centralized network architecture, including those discussed in [1, 25, 26]. However, the centralised attacks may be difficult to implement in distribution systems, which require knowledge of local states. Adversaries may also intend to forge the injection of bad data against the energy system’s supply-side, energy control commands, communication links of energy transmission, and distributed energy routing processes [27].

5.2.3 Classification of Threat Models Based on Construction Methodology. The various adversarial construction methods used in the literature are presented here.

(1) Attacks with Complete Topology Information: In these types of FDI attacks, adversaries typically require a complete knowledge of network topology, transmission system parameters, details of SE algorithm, and/or BDD methods. This case presumes that the adversary has access to several resources of the electric power system and can successfully construct the FDI attack vector. Although most FDI attack studies consider this type of strategy, it is impractical to assume that adversarial models have access to a large number of measurements. Liu et al. [1] have demonstrated the constraints faced by adversaries. Accordingly, the adversary can be restrained only to a certain set of sensor readings due to the fact that the sensors may have specific physical defences or the adversary may have a limited budget to compromise the sensors.

According to the findings of Liu et al. [1], the objective of the adversaries may be to randomly inject bad data, in which they aim to locate any attack vector so long as it can bring a wrong SE performance of state variables or to launch more targeted attack vectors, in which the adversaries aim to build bad data injections into some chosen state variables. Studies include random and targeted FDI attacks from the SE to other cyber-physical components. In [27], random bad data were injected to a distributed system to compromise the supply–demand of the energy system. In contrast, Kosut et al. [25] view the nature of stealthy FDI attacks as a matter of basic constraint
on the detectability of malicious data attacks. Unlike Liu et al. [1], Kosut et al. came up with the concept of a detectability heuristic to find the attacks that would render BDD the most vulnerable provided a specific set of meters controlled by the attacker. The authors proposed an FDI attack algorithm [25] based on minimal energy leakage by considering two forms of attacks: the strong attack and the weak attack. In the strong attack regime, the adversary compromises a sufficient number of meters such that the system state becomes unobservable by the SE utilising a graph theoretic approach, whereas in the weak attack regime, the adversary controls a limited number of meters.

However, the FDI attacks pose several stringent requirements against the intruders. For instance, the topology settings of the power system are typically only available at the operator’s EMS, whose physical access is strongly restricted and secured. Further, these settings change very often due to routine normal maintenance of electrical power grid devices and unplanned incidents such as unexpected field device failure. In general, intruders have restricted physical access to most power grid infrastructure and they barely have real-time knowledge with respect to topology configurations and physical states such as the transformer tap changes, circuit breakers, and switches. Therefore, attackers need to pursue alternative approaches, discussed next.

(2) Attacks with Partial Topology Information: As discussed earlier, the construction of a valid FDI attack is subject to certain constraints. Although it is ideally fair to implicitly presume that the topology information can be accessible to the adversary in order to build the attack vector, it is more realistic to believe that the adversary has incomplete topology knowledge for certain transmission line networks due to the adversary’s lack of real-time knowledge with respect to topology configurations and physical status such as the transformer tap changes, circuit breakers, and switches. Therefore, a realistic FDI attack can be launched with incomplete information as the adversary can have access to limited resources only [18]. Liang et al. [10] have reviewed various scenarios under which adversaries can get partial topology information necessary to launch a successful FDI of this attack category. One is a manual or online mode [18] in which before generating the FDI attack, the adversary collects grid topology information either manually or online where the adversary can use one’s own meters to access the grid. The other is through a market database (extracting the topology information from locational marginal prices). Finally, extraction of \( H \) from power flow measurements is achieved.

(3) Load Redistribution Attacks: Under restricted access to specific meters, the load redistribution (LR) attack [18] is a special type of FDI attack targeting load measurements of nodal power injections and power flows. This kind of FDI attack aims to generate biased load estimates.

(4) Grid Topology Attacks and Line Outages: Research of attacks against power grid topology (GT) and attempts to cause outages of transmission lines are very recent developments [28]. Most of the adversarial models mentioned earlier are focused on the premise that the power grid topology stays unchanged. This implies that the adversary can only inject false data to the measurement data of the power system. As a matter of fact, topology configurations do change very often due to routine normal maintenance of electrical power grid devices and unplanned incidents such as unexpected field device failure. Therefore, the state-of-the-art literature on FDI attack strategy targeting power system states has further been extended to reflect on real-time grid topology. The purpose of such attacks is to concurrently alter network measurements and topology configurations such as physical states of transformer tap changes, circuit breakers, and switches so that the estimated topology is consistent with the received network data.

(5) Data-Driven Attack: In this type of attack, also known as the blind attack method, undetectable FDI attacks are constructed without prior power grid knowledge, typically using statistical inferences (e.g., sparse optimization [29]), heuristic methods [26], and ML [30, 31] algorithms. The adversary is expected to make inferences from the correlations of measurement
### Table 2. Categories of FDI-Based Attack Models

| Category of FDI attack models          | Description                                                                                                                                 |
|---------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------|
| Based on power flow                   | **DC power flow**                                                                                                                             |
|                                       | Attacks are constructed in a constrained environment with the assumption that the functions from power system states to measurements are linear (DC-based power flow models). |
|                                       | **AC power flow**                                                                                                                             |
|                                       | The attack modeling is accompanied by solving complete nonlinear power flow equations and nonlinear constraints found in nonlinear power system models. |
| Based on architecture                 | **Centralised**                                                                                                                             |
|                                       | In the centralised-based network architecture, FDI attacks are modeled to target centralised SE, in which the adversary may attempt to manipulate measurement reports sent from different communication devices to the control center or tries to manipulate the SE outcomes. |
|                                       | **Decentralised**                                                                                                                            |
|                                       | FDI attacks are modeled to target distributed SE in the decentralised network architecture, in which the adversary may attempt to manipulate measurement reports within distributed control centers and/or tie-line parameters shared among the distributed control centers. |
| Based on construction methodology     | **Attacks with complete topology info**                                                                                                      |
|                                       | The construction of FDI attack requires a thorough understanding of network topology, transmission system parameters, SE algorithm details, or BDD methods. |
|                                       | **Attacks with partial topology info**                                                                                                       |
|                                       | Certain constraints apply to the design of an FDI attack. Due to the adversary’s lack of real-time knowledge of topology configurations or physical device status, adversaries have incomplete topology knowledge for all or certain transmission line networks. |
|                                       | **Load redistribution attacks**                                                                                                             |
|                                       | This attack category targets Smart grid load measurements of nodal power injections and power flows in order to generate biased load estimates. |
|                                       | **Grid topology attacks**                                                                                                                   |
|                                       | This adversarial model focuses on grid topology and transmission lines, with the adversary attempting to simultaneously alter measurements of network and topology configurations, such as physical states of transformer tap changes and circuit breakers. |
|                                       | **Data-driven attacks**                                                                                                                     |
|                                       | In this case, attackers are expected to design FDI attacks without prior knowledge of the power grid, for example, by using statistical inferences, machine learning, or other data-driven approaches to infer correlations between measurement data and smart grid topology parameters. |

Data and/or topology parameters of the power system. Table 2 summarizes the different categories of FDI-based threats.

### 5.3 Smart Grid Components Vulnerable to FDI-Based Threats

Various cyber-physical elements are essential for monitoring and controlling grid operations (see Section 3). However, they also make the Smart Grid vulnerable to a variety of data breaches that may bring a greater exposure to attacks on data integrity, data confidentiality, data availability, and so forth. FDI attacks target various cyber-physical components of the Smart Grid ranging across all domains, namely generation, transmission, distribution, consumption, market, and operations. In this section, vulnerabilities of some of the principal cyber-physical elements are discussed.

**EMS.** The EMS within the control center is the most vulnerable element in the Smart Grid. This is quite important, particularly because the processes within the EMS are temporally sequential. For example, the output of the SCADA or PMU systems are critically demanded by the state estimator, and the other subsequent EMS modules require the output of the state estimator to a great
extent. As a result, the state estimator is the most important target for cyber attackers. Coordinated and sophisticated cyberattacks such as FDI can compromise measurement data, cause unbounded estimation errors, and deceive the system operator. Since the first paper [1] on FDI attacks, the majority of FDI attack methodologies described in Section 5.2 target the SE. The vulnerability issues in the SE problem can be investigated with respect to the various cyber-physical elements, including physical properties of the power system, communication systems, IEDs, and AMIs.

**Automation Generation Control.** In the power grid, data between AGC and generator units or NCSs is transmitted via communication systems such as SCADA and PMU, making them vulnerable to cyberattacks. It has been experimentally evaluated in [32] that the AGC algorithm can be manipulated by adversaries on frequency measurements, generation of load balance, and control commands between AGC and generator units.

**Contingency Analysis.** Attackers could stealthily introduce contingency of transmission lines to a normal contingency list by misleading the CA process through the injection of false data into the SE. The exploited contingency would then be embedded as security constraints in the security constrained economic dispatch (SCED), which may result in various impacts.

**Distribution Energy Management System.** The DEMS [17] has become instrumental for handling real-time networks and dynamic decisions that could not otherwise be taken by conventional EMS. Despite their popularity in the power grid, they face an unprecedented challenge from incumbent cyberattacks. The vulnerability of the DEMS to FDI was studied in [27]. It was found that the manipulated data introduced by the attacks would cause imbalanced demand and response, increase costs for electricity transmission and distribution, and affect the reliability of energy supplies in the power grid.

**Market Management System.** The market management system (MMS) is the national electricity market of the grid that dictates energy prices. The MMS is designed to facilitate standardised transactions between service providers and utility consumers in the energy industry. The MMS provides market information based on variables such as price, dispatch, and other constraints obtained from EMS/DEMS modules. The MMS has become a primary goal for adversaries to manipulate intelligence on the utility market or otherwise to make illicit financial gains.

**Communication System.** There are numerous cyber-dependent communication technologies [17] in the Smart Grid, including in the SAS (e.g., IEC 61850), WAMS (e.g., PMU), SCADA, AMI, and NCS. These components are among the most vulnerable to FDI attacks. Power system measurements are also vulnerable to FDI attacks via the SCADA. This may further affect other cyber-physical elements such as the SE or AMI. For example, if adversaries get access to the SCADA system, they can damage AMI, where the intruders can carry out falsifying customer billing information. Communication protocols (such as the DNP-3 and IEC 61850) are also vulnerable to FDI attacks [17]. Other communication systems that can potentially bring vulnerability to the Smart Grid environment include WAMS (i.e., PMUs) [33], IEEE C37.118 [17], and wide area network communication infrastructure [16].

**Intelligent Electronic Devices.** IEDs link field devices to a communication infrastructure that enables SCADA and SAS to gather critical grid information (see Section 3). FDI attacks have been found to jeopardise such critical information by breaching IEDs [34]. For example, FDI attacks can temper voltage readings over the IEDs, and they can modify IED settings that can also cause the relay to trip. This can lead to an abrupt voltage drop below the critical level, resulting in load shedding and, much worse, power outages.

**Renewable DERs.** DERs are among the most vulnerable cyber-physical components to FDI attacks. Lin et al. [27] discuss the vulnerabilities of DERs in regard to the routing process. The authors’ confirm that the forged data injected by the attackers would induce imbalanced demand and response, cause higher costs for energy transmission, distribution, and number of outage
customers. Microgrids have become increasingly popular in the Smart Grid infrastructure due to their versatility and integration with renewable energy. However, they have also become potentially susceptible to a variety of cyber threats, such as FDIs [34].

6 CLASSIFICATION OF FDI ATTACK DEFENCE STRATEGIES

The success of cyber-physical attacks in general and FDI attacks in particular depends on both the perspective of the adversary and the operator. It is highly likely that adversaries are subject to a trade-off between maximizing the probability of impact on various cyber-physical system components and minimizing the probability of detection of the launched attack. This section provides an extensive review of existing state-of-the-art research on the defence against the incumbent cyber-attacks from the point of view of the power grid operator. There have been substantial research works on mitigation strategies against the FDI attack. We believe that taxonomy of the different countermeasures will help other researchers in the cyberattack defence arena to see correlations, differences, and to foresee future perspectives of these concepts. Here, we broadly classify the countermeasures into five categories, with taxonomy depicted in Figure 4 and details of each class presented next.

6.1 Countermeasures Based on SE Type

6.1.1 Conventional Bad Data Detectors. The BDD [35] has been an integral part of power system state estimators. It is used to detect and remove faulty measurements caused by error due to device malfunctions, communication channel problems, or cyberattacks, and is still widely used in various commercial EMS software. The $\chi^2$ distribution test (or $\chi^2$-detector) [35] is the most widely used BDD in the power system SE, employing hypothesis testing based on WLS estimation to determine cyber anomalies or bad data. The detail implementation of the $\chi^2$ distribution test can be found in [4]. The largest normalised residue (LNR) [35] test is another metric for the identification of bad data. Given the residue between the observed measurement and the estimated measurement as $r = y - H\hat{x}$, the LNR test is based on the largest or maximum value of the normalised residue $\hat{r}_i$ for each measurement index $i$.

Typically, BDDs are used in centralised state estimators. A few studies [4, 6, 36] extended the use of BDDs in distributed based state estimators. In [4], the authors suggest dividing a power system into many non-overlapping subsystems according to the physical topology and applying SE and a $\chi^2$ distribution test for the detection of bad data in each subsystem. The findings reveal that the local degree of freedom is less than that obtained from the centralised SE, which results in a better...
identification of bad data. Similarly, in [6] under distributed SE, a BDD is examined by taking into account the weight of local measurement residual of subareas of the power system and the overall change of measurement residual. A substation-level BDD is also proposed in [36]. However, the BDD-based detection approaches did not address FDI attacks and, thus, are vulnerable to FDI attacks [1].

6.1.2 Detection Based on SE Partitioning. By decomposing the power grid into several subsystems, measurement redundancy can be reduced and the false data threshold in each subsystem can be lower than the original system, improving the likelihood of attack detection (as proposed in [37] using adaptive graph partitioning-based SE and $\chi^2$- detector).

6.1.3 Detection Based on Dynamic SE. The absence of real-time information in the power system operation can be attributed to its use of steady-state estimators that produce input data for many EMS modules. Dynamic SE methods, on the other hand, model the time varying behaviour of the process, making it possible to predict the state variables ahead of time. In this case, the SE proves to be a great advantage for the system operator to conduct security analysis as well as other EMS functions. The Kalman filter (KF) has been extensively utilised in dynamic SE. There are different extensions of KFs available, including extended KFs, unscented KF (UKFs), ensemble KFs, and particle filters that are designed for non-linear systems [38]. The discussion of these techniques is beyond the scope of this article; details of each dynamic SE are found in [38].

With the advent of dynamic state estimators, more effective countermeasure strategies against FDI attacks than BDDs have emerged. In [39], combined $\chi^2$-detector and cosine similarity matching techniques are employed for the detection of FDI attacks in Smart Grids in which KF estimation has been used to measure any deviation from actual measurements. In the $\chi^2$-detector, the variation in the KF-estimated measurements is used to identify malicious attacks; in the cosine-similarity metric, the cosine of the angle between the received measurements and the KF-estimated measurements is computed to detect attacks. While the $\chi^2$-detector has been confirmed as vulnerable to FDI attacks, the cosine similarity is found to have a better detection probability against FDI attacks. Yet, the cosine-similarity criterion is not efficient for sparse FDI attacks in which the cosine angle between the received measurements and the injected data becomes almost unity, which bypasses the underlying detector.

Similarly, in [40], the authors developed a detection of FDI attacks in Smart Grids based on a KF-based state estimator. A Euclidean distance detector was employed to detect the discrepancies between KF-estimated data and the received measurements. If the Euclidean distance-based difference is greater than a predetermined threshold, the detector triggers a decision on whether an attack has occurred. Although the proposed approach achieves better detection accuracy than conventional BDD, there are two drawbacks to this approach. First, it considers only time-invariant states, and the dynamic nature of the state variables is ignored. Moreover, the proposed detector cannot distinguish between an FDI attack and a failure due to physical faults. In contrast, the authors of [41] suggest the use of a combined UKF state prediction and WLS-based SE algorithm to detect inconsistencies between state vector estimates and, as a result, to detect false data attacks for non-linear measurement models. Although the combined WLS and UKF estimates has a better detection rate of FDI than BDD and KF techniques, it has drawbacks. First, UKF state predictions are highly influenced by the non-linear transition matrix and system noise, which can potentially make it difficult to distinguish between attack-free and compromised states. Second, the accuracy of detection relies on the UKF predicted outcome, whose uncertainties can result in high false positives. Third, a generalised FDI attack is considered, rather than a more stealthy and sparse FDI attack, which may pass the proposed detector. Therefore, data of various load forecasts, proper threshold selection, and threat models are among the critical points to
consider for the robustness of the proposed methodology. Other FDI detection approaches based on the dynamic SE include: spatio-temporal correlations [39] among states of the power system, a short-term state forecasting-based [42] approach for analysis of nodal state temporal correlations, and a graph signal processing–based [43] scheme to determine a graph Fourier transformation of the estimated states and to filter the high-frequency components of the graph.

6.2 Protection-Based Defence

In Smart Grid cybersecurity, protection-based defence aims to deter attacks by identifying a set of measurement devices and making them immune to incumbent cyberattacks for ensuring observability of the states. The objective of introducing protection to components of the Smart Grid is that the attacker would not get enough measurements to start FDI attacks, which otherwise will render the power system unobservable. The idea behind this defence technique is that, for a given grid topology, certain sensor readings affect more state vectors than others and should thus have a better cost–benefit ratio when protected, and certain state vectors are more reliant on sensor data than others, thus separately checking their estimation can limit the ability of hackers to exploit the sensor data without being noticed. Three major research approaches have been investigated and are discussed further next.

6.2.1 Optimal PMU Placement. It has been found that the cyberattack protection capability of a power grid can be significantly enhanced with the integration of a few secure PMUs in the grid [44]. This is because PMUs measure voltage and current phasors using a standard time source based on a global positioning system and therefore have the potential to provide precise timestamped measurements for geographically distributed nodes. As a result, they have secured measurements and are usually resilient against bad data injection attacks. For the same reason, in [45], a linear programming–based PMU placement algorithm is used to determine the number of PMU placements across a grid with $b$ number of branches and $n$ number of buses. For PMUs with $c$ current phasor measurements, they calculated $k = \binom{b}{c}$ possible combinations to assign those $c$ PMU configurations. Thus, the number of possible PMU configurations for all buses is calculated as $N = \sum_{i=1}^{n} k_i$. A semi-definite programming approach [46] has a better solution than the approach featured in [45] for the problem of optimal placement of PMUs for protecting measurements against malicious attacks. Similarly, a mixed integer programming method [19] determines the minimum number of PMUs needed to protect against unobservable data integrity attacks.

Compared with a polynomial time-complexity of linear programming and semi-definite programming, and an exponential time-complexity of mixed integer programming, greedy heuristics [26] can provide more optimal placement of secure PMUs to defend against bad data injection attacks. Most of the strategies mentioned focus on evaluating the optimal placement of PMUs to enhance power system observability, cost, and protection, and improvement of SE. However, considering the adversary-operator dynamics, the adversary might have partial knowledge about the operator’s corresponding defense measures, in which they could optimize their attack strategy. For instance, the PMUs and the power system can be compromised by the adversary during the device configuration process. Consequently, the aforementioned approaches are insufficient. As a solution to the drawback, the authors of [47] propose a predeployment PMU greedy algorithm against the attack in which the most vulnerable buses are first secured and then a greedy-based algorithm is used to deploy other PMUs until the entire power system is observable. The defence space against the FDI attack can also be strengthened using a hybrid protection-based and detection-based scheme as suggested in [48], in which the former is utilised to protect essential measurements from the intruder by means of physical defences and the latter is used to identify modified data. They proposed a zero-sum static game-theoretic approach for the optimal deployment of the PMUs.
However, PMUs are very costly. Thus, it is not practical to install enough PMUs to secure sensor readings. It is definitely much more expensive, especially with the emerging ubiquitous sensing infrastructure in the large-scale Smart Grid. In addition, research has shown that PMUs are vulnerable to FDI attacks via GPS spoofing [33]. Therefore, a more appealing security scheme is required to protect the power system against FDI.

6.2.2 Protection Via Selection of Optimal Measurements. This is a security technique developed to defend SE against the injection of bad data through a carefully selected subset of measurements. For instance, the authors of [23] employed a brute-force search for identifying an optimal set of measurements and state vectors to ensure that stealthy data injection attacks are detected by the grid operator. The method enables the grid operator to choose a random number \( q \) out of \( n \) state variables and to pick a random number \( p \) out of \( m \) sensors, which should fulfill \( \binom{m}{p} \cdot \binom{n}{q} \) combinations for a given choice of \( q \) and \( p \), where \( 0 \leq q \leq n \) and \( 0 \leq p \leq m \). Similar to the brute-force method, a fast greedy search algorithm [54] can find an optimal subset of measurements for protecting against stealth FDI attacks. Further, by decomposing the connected elements of the power grid into many subnetworks, approximate solutions for the minimal number of measurements can be achieved, for example, using a mixed integer linear programming [55] model.

In these three approaches, the system operator has to randomly select the number of measurements to be protected. Therefore, although the proposed method can be feasible for a small number of power systems, it is costly for a large-scale power grid. In contrast with [23], protection measures are introduced in [56], taking into account perfectly protected measurements (an ideal assumption that no stealth data injection attacks are possible) and non-perfectly protected measurements (in which the operator seeks to maximize its protection level through some metric) considering the operator’s budget as a constraint. However, determining such a subset of measurements is a large-complexity problem. To alleviate these complexities, other approaches in this research direction include graph-theoretical and game-theoretical, both discussed next.

(a) Protection Based on Graph-Theoretic: Graph-theoretic approaches are widely used for power system observability analysis [112]. They have also been used to define optimal protection problems to safeguard state variables with a minimal set of measurements. Some of the methods considered include Steiner tree–based graph theory [57] (which defend a set of priority-based critical state vectors), and optimal and suboptimal solutions for state protections by modelling the Smart Grid as a minimum Steiner tree measurement problem [58].

(b) Protection Based on Game-Theoretic: Game theories are important theoretical frameworks for the development of optimal decision-making of competing players, such as the adversary and the operator in the defending space of the Smart Grid. The optimal set of protection can be formulated as a three-level defender-attacker-operator problem [59] to deter the success of the coordinated attacks. A zero-sum Markov game-theoretic [60] models the defender-attacker relationships, in which the defender can maximise benefit by misleading the adversary to use incorrect cost functions of the grid. An adaptive Markov technique [34] dynamically computes an optimal defence scheme against malicious attackers with dynamic and unpredictable behaviors.

6.2.3 Grid Topology Perturbation. Most legacy IT systems are static, giving hackers enough time to conduct reconnaissance against the system, learn about flaws and potential attack vectors, and ultimately launch attacks against the system. Recently, moving target defence (MTD) [61–64] has emerged as a proactive defence strategy that has been studied in various areas of cybersecurity. MTD is helpful to maximise the complexity against adversaries by implementing uncertainty or
Table 3. Defence Methods against FDI Attacks in Smart Grids

| Category                                      | Subcategory                          | Approaches/Algorithms                              | References                      |
|-----------------------------------------------|--------------------------------------|---------------------------------------------------|---------------------------------|
| Based on SE                                   | Conventional BDD                     | $\chi^2$-detector                                 | [5, 35]                        |
|                                               |                                      | L NR detector                                     | [49, 50]                       |
|                                               |                                      | Detection based on SE partitioning                | [37]                           |
|                                               | Detection based on dynamic SE        | KP and extensions                                 | [40, 41, 51, 52]               |
|                                               |                                      | Spatio-temporal correlations                      | [39]                           |
|                                               |                                      | State forecasting                                 | [42]                           |
| Protection-based defence                      | Optimal PMU placement                | Integer linear programming                        | [44, 45]                       |
|                                               |                                      | Mixed integer semi-definite programming           | [19, 46]                       |
|                                               |                                      | Greedy algorithm                                  | [20, 26, 53]                   |
|                                               |                                      | Predeployment PMU greedy                           | [47]                           |
|                                               |                                      | Hybrid protection-detection                        | [48]                           |
|                                               | Optimal measurement selection        | Heuristic search (greedy algorithm and others)     | [23, 54–56]                    |
|                                               |                                      | Graph-theoretic                                   | [57, 58]                       |
|                                               |                                      | Game-theoretic                                    | [34, 59, 60]                   |
|                                               | Grid topology perturbation           | MTD                                               | [61–67]                        |
|                                               |                                      | Hidden MTD                                        | [68–70]                        |
| Statistical model/Protection-based defence    | GLR test detector                    | $\ell_1$-norm minimisation                        | [25, 71]                       |
|                                               |                                      | Auto-regressive                                   | [72]                           |
|                                               | Bayesian test detector               | Game-theoretic                                    | [73]                           |
|                                               |                                      | Joint estimation-detection                         | [74, 75]                       |
|                                               | Quickest change detection            | CUSUM and adaptive CUSUM                          | [24, 76–78]                    |
|                                               |                                      | Sequential change detector                        | [79–82]                        |
|                                               | Statistical distance                 | KL distance                                       | [83, 84]                       |
|                                               |                                      | JS distance                                       | [85]                           |
|                                               | Low-rank and sparse matrix recovery  | Sparse matrix optimisation                        | [29, 86]                       |
|                                               |                                      | Fast Go Decomposition                             | [87]                           |
| Data-driven                                   | Supervised ML                       | SVM                                               | [88–90]                        |
|                                               |                                      | ANN                                               | [91–93]                        |
|                                               |                                      | KNN                                               | [94]                           |
|                                               | Semi-supervised ML                  | Semi-supervised ANN                               | [88, 89]                       |
|                                               |                                      | Semi-supervised GMM                               | [91]                           |
|                                               | Deep learning                       | DFFNN                                             | [95]                           |
|                                               |                                      | CDBN                                              | [96]                           |
|                                               |                                      | DRNN                                              | [97, 98]                       |
|                                               |                                      | CNN                                               | [99, 100]                      |
|                                               |                                      | G AN                                              | [101]                          |
|                                               | Reinforcement learning              | Q-learning                                        | [31]                           |
|                                               |                                      | SARSA                                             | [30]                           |
|                                               |                                      | Bayesian Bandit                                    | [102]                          |
|                                               | Deep reinforcement learning         | Deep-Q-network                                    | [103]                          |
| Prevention                                    | Cryptographic schemes               | Encryption and dynamic key management              | [56, 104, 105]                 |
|                                               |                                      | Authentication                                    | [106–108]                      |
|                                               |                                      | End-to-end signature                              | [109]                          |
|                                               | Blockchain-based defence            | Data protection                                   | [110, 111]                     |
|                                               |                                      | Privacy preservation                              | [21, 22]                       |

to increase the cost of attack. MTD has gained popularity among grid operators due to its ability to proactively protect measurements from malicious attackers by introducing perturbations to network data or topology. Perturbation can be done by systematically changing system settings that adversaries might need to launch their attacks in order to nullify their prior knowledge of the
system, making it impossible for the adversaries to adapt their attack space. In this regard, as the topology perturbation patterns are hidden from the hackers, they cannot compute and generate the proper response for the measurements or topology under their control, which makes the FDI attack unable to correct to remain undetectable.

There are different kinds of perturbations for protecting key grid elements against FDI attacks. In [61], for example, the authors applied impedance changes through a key space approach to a number of selected transmission lines by leveraging D-FACTS devices in order to generate noticeable system changes that the adversary cannot foresee. The anticipated system response is predicted and compared with the observed measurements. MTD can also utilise both a randomized set of measurements considered in SE and the topology of transmission lines [62–64]. However, these MTD strategies have been implemented under a weak adversarial environment in which they overlook the likelihood that sophisticated FDI attackers may also attempt to identify MTD changes before they execute the attack. As a remedy for this limitation, the authors of [68] introduce a hidden MTD, an approach that hardens the stealthiness of the MTD.

6.3 Detection Based on Statistical Modelling
Several research efforts of statistical-based detection frameworks against falsified injection of data have been addressed by the research community, including the GLR test detector, Bayesian test framework, quickest change detection, statistical distance index, and sparse matrix recovery.

6.3.1 GLR Test Detector. The GLR test detector is a statistical model used for detecting cyberattacks in the power system by leveraging the likelihood ratio of statistical tests. While it is usually not feasible to use the GLR test detector to detect a large number of compromised samples, it can do well to detect weak FDI attacks [25], in which \( \ell_1 \)-norm minimisation is proposed to solve the detection problem. In particular, it has been noted in [25] that if multiple measurement samples are available under the same sparse FDI attack, the GLR test detector can be asymptotically optimal in the sense that it gives a very low probability of missed detection. Although the FDI detector in [25] is valid under AWGN distribution, a study [72] has shown that it does not work properly when the measurements are corrupted by non-Gaussian noise distributions. The authors of [72] used an independent component analysis along with the GLR test detector for an FDI attack on the basis that the power system measurements are subject to a colored Gaussian noise (modeled through an auto-regressive process).

6.3.2 Bayesian Test Detector. Bayesian-based statistical frameworks are essential for decision-making by leveraging prior knowledge and new evidence. For example, a strategic attacker–defender Bayesian game-theoretic detection technique [73] against FDI may be established when the Bayesian game is played on each node in the event of an attack on that node and a critical set of measurements to be defended is obtained for the particular node. Further, in [74], a Bayesian-based detector has been proposed for each monitoring node using a distributed architecture in WAMS. Once the probability of FDI attack vectors is determined by Bayesian inference, then a recursive Bayesian-based prediction is derived for attack detection using measurements obtained from real power transmission grids and simulated measurements. Another related work on the Bayes approach for the detection of FDI attacks is [75].

6.3.3 Quickest Change Detection. Quickest change detection (QCD) [24] is a mechanism to detect sudden changes as soon as possible using sequential or real-time observations, minimising the lag between the moment a change appears and the time it is observed. A variety of detection methods have been suggested under different conditions. Unlike static BDD procedures that
rely on a single measurement at a time, QCDs consider the use of dynamic change detection procedures. QCD-based detection techniques \cite{24} can be used with Bayesian models, non-Bayesian models (e.g., CUMulative SUM (CUSUM), adaptive CUSUM test) and statistical hypotheses tests. The following points summarize the works that apply the QCD technique to detect FDI attacks in Smart Grids: (1) A Markov chain–based QCD algorithm is proposed to detect and remove FDI attacks \cite{79}; (2) joint dynamic CUSUM and static $\chi^2$–detector \cite{77} in which the former leverages historical states and the latter utilises a single measurement at a time; (3) a generalized CUSUM algorithm \cite{76} is suggested for quickest detection of FDI attacks for dynamic KF-based state estimator under centralised and distributed settings; (4) adaptive CUSUM methodology for QCD using a linear unknown parameter solver \cite{24}; (5) Markov chain–based adaptive CUSUM \cite{78} for real-time detection of FDI attacks; (6) sequential detection of centralised and distributed FDI attacks based on GLR test \cite{80}; and (7) a generalized sequential likelihood ratio test for a decentralised system \cite{81}.

6.3.4 Detection Based on Statistical Distance Index. A statistical distance quantifies the consistency of two probability distributions through, for example, a variational distance between the distributions. Kullback–Leibler (KL) distance \cite{83} and Jensen-Shannon (JS) distance \cite{85} have recently been used for detecting malicious power system measurements by calculating the dissimilarity among probability distributions obtained from measurement variations. The KL distance metric has been suggested \cite{83, 84} to track measurement dynamics and to detect FDI attacks. When bad data are injected into power systems, the variations in the probability distributions of the measurements deviate from historical data, leading to a greater KL distance. Likewise, the JS distance–based detection framework \cite{85} monitors dynamics of probability distributions obtained from historical measurement variations and real-time measurement variations. Similar statistical distance-based approaches have also been used along with data-driven techniques (see Section 6.4).

6.3.5 Detection as Low-Rank and Sparse Matrix Recovery. The measurement matrix obtained at the control center has a low-dimensional structure due to the inherent temporal correlation of states of the power grid. Low-rank and sparse matrix recovery is another alternative for defending against existing cyberattacks, taking into account the low-rank structure of measurements and the low-sparsity of false data attacks. The detection problem of measurements with FDI attacks have been formulated as a low-rank and sparse matrix recovery \cite{29, 86, 87}. Liu et al. \cite{29} formulate the problem of detecting FDI attacks as low-rank matrix recovery in the form of augmented nuclear norm\(^4\) and $\ell_1$-norm minimisation. By considering the intrinsic low-dimensional structure of temporal attack-free measurements of power grid and sparse FDI malicious attacks, they extended their work in \cite{86} to a problem of sparse matrix optimisation in \cite{29}, solved using low-rank matrix factorization. On the other hand, while the results of \cite{86} and \cite{29} has a good computational efficiency, they have quite low FDI detection accuracy. Therefore, in order to obtain a better balance between detection accuracy and computational performance, the authors of \cite{87} proposed a new approach known as Fast Go Decomposition, which considers the low-rank behaviour of the measurement data and the sparse FDI attack.

6.4 Data-Driven Attack Detection

Various ML techniques have been employed for the detection of FDI attacks in smart power grids. Supervised learning classifiers \cite{88–94} are the most popular ML techniques for the detection of false data. These techniques can reflect the statistical characteristics of the power system using historical data and may allow the training model a better decision if redundant power system measurements are available. Historical training data can include class labels of normal versus

\(^4\)Nuclear norm is a convex optimisation problem that is used to search for low-rank matrices.
tampered; using such training data, a new observation is predicted as either false data injected or normal data. Ozay et al. [88] suggested supervised learning-based binary classifiers using statistical deviations between the FDI-corrupted and secured measurements. Similarly, in [89], an FDI detection is used based on principal component analysis for reducing dimensionality of measurement data and supervised learning over labeled data for classification. In the literature, various supervised ML algorithms are employed, including support vector machines (SVMs; e.g., in [88–90]), artificial neural networks (ANNs; e.g., in [91–93]), and k-nearest neighbor (KNN) (e.g., in [94]).

One of the disadvantages of supervised learning techniques, however, is that they require far more labelled data, which is often difficult to obtain. For this reason, semi-supervised learning techniques address the problem of supervised learning by using partially labelled samples. This approach seeks to label unlabeled data points using information gained from a limited number of labeled data points. The authors of [88, 89, 91] also employed semi-supervised algorithms. Attack strength and sparsity are the two main factors that should be considered in the detection frameworks. While most proposed supervised and semi-supervised approaches consider a relatively high magnitude of FDI attacks, their detection accuracy is low for very small attack magnitudes. On the other hand, deep learning (DL) techniques can extract high-dimensional temporal features of FDI attacks with historical measurement data and can use the known features to detect various magnitudes of FDI attacks in real time. More specifically, the latest advances in graphics processing unit (GPU) computation provide the basis for deep neural networks such as deep feedforward neural network (DFFNN) [95], deep belief network (DBN) as used in [96], deep recurrent neural network (DRNN) [97, 97], convolutional neural network (CNN) [99], and a semi-supervised deep learning approach using a generative adversarial network (GAN) framework [101].

6.5 Prevention-Based Defence

Intelligent and integrated cyber-physical resources intended to improve the stability and reliability of the Smart Grid could be used as weapons against the grid itself. Without proper cyberattack prevention schemes, the Smart Grid can be more vulnerable, especially when it is connected to the Internet via less secure wireless communication systems such as ZigBee and Wi-Fi. Most of the defence countermeasures are based on identification of the false data attack normally after the threat compromised data integrity at the control center, during transmission, or at measurement devices. To this end, lack of adequate preventive security measures against coordinated false data attacks could be disastrous. Hence, as part of Smart Grid cybersecurity, preventive security measures are essential in the battle against attacks such as FDI. By providing prevention schemes across key cyber-physical resources, we can deter the malicious users against unauthorised access of EMS/DEMS/MMS critical OT database systems, exploitation of the communication protocols (e.g., IEC 61850), compromising user privacy or data integrity via smart meters, and tampering IEDs or interception of data transmission in WSN, IoT, and cognitive radio.

Implementing a preventive technique or a combination of preventive and detective systems can result in an effective security response across the Smart Grid. The most prominent prevention systems in Smart Grids, such as cryptographic schemes and privacy preservation using Blockchain, are discussed next.

6.5.1 Cryptographic Schemes. It is highly likely that adversaries can exploit communication channels when measurements are sent from sensors to control centers or when customer data are transmitted from smart meters to the control centers over unencrypted communication channels. For example, the unencrypted communication channel of plain text transmission over the SCADA network or the IEC 61850-SAS compliant communication protocol could be hacked by malicious actors. Although cryptographic techniques are well matured and used in a variety of domains,
they are difficult to implement in the Smart Grid due to the limited computational capabilities and the deployment of measurement sensors or related devices in hostile environments. Therefore, fast and efficient cryptographic operations are required for implementation in the Smart Grid to guarantee the accuracy and integrity of measurements against FDI attacks.

In [56], it was suggested that encrypting a sufficient number of IEDs could improve measurement protection against stealthy FDI attacks and could increase overall system security in the utility control center. Dynamic key management-based cryptographic protocols can also deter cyberattacks against privacy in Smart Grid wireless communication networks [104]. Similarly, a dynamic and periodic secret-key generation scheme over Smart Grid communication networks against various cyberthreats, including FDI, is proposed in [105], which enables a resilience so that no adversaries can exploit the network over a longer period of time even if they know a secret key.

The Smart Grid infrastructure involves millions of electronic devices that link customers to different cyber-physical entities. This calls for a strict authentication process, which is vital for the verification of the customers and devices. For example, strict authentication schemes can be implemented in the IEC 62351 EMS-compliant security standard. For Smart Grid distribution systems, an FDI prevention protocol is proposed in [106] that focuses on data integrity by preventing packet injection, replication, modification, and access to rogue nodes for the IEC 61850-90-1 SAS communication security standard. Three stages accompany the operation of their proposed protocol: node authentication (authentication techniques across the distribution network, including routers, gateways, inter-substation devices), peer authentication (authentication of a routing protocol when using a cloud platform for the distribution system), and data transmission. Lightweight hash-based message authentication protocols [106, 107] are also critical for thwarting false data attacks in IP-based data transmission in the Smart Grid environment. Further, a lightweight authentication scheme [108] with reduced energy, communication, and computational overheads can establish a secure communication between two communicating parties, such as smart meters and wireless base stations, and can provide energy efficiency in a resource-constrained environment. Other prevention methods include end-to-end signature schemes [109], which can protect data during an end-to-end communication in Smart Grids (e.g., to protect legitimate commands transmitted from the control center to IEDs against malicious commands).

6.5.2 Blockchain-Based Defence. Data protection capabilities of Smart Grids against FDI attacks can be harnessed by introducing distributed blockchain-based reconfigurable SCADA network features for geographically distributed sensors [110]. Here, during transmission or reception, each piece of information in the distributed blockchain network is cryptographically connected block by block, and includes signatures for verification. Further, blockchain can be used to preserve the privacy of the user’s energy data against coordinated data integrity attacks. A distributed blockchain network–based data management on mobile nodes for microgrid trading is proposed in [111] that aims to prevent false data attacks.

Blockchain-based privacy preservation mechanisms are used to protect network nodes or data transactions in the form of a peer-to-peer crypto connectivity [110]. In [21], a blockchain-based bi-level privacy module and anomaly-detecting module is designed to verify data integrity and mitigate false data attacksA variational autoencoder and anomaly detector is proposed in which the former is applied for transforming data into an encoded format to prevent cyberattacks and the latter is used to detect any interference attack.

7 LITERATURE REVIEW METHOD
The method of literature review represents the foundational first step that makes up the skeleton of the knowledge base and largely dictates its reconstruction in the successive analysis of the
literature. Therefore, the process of a systematic search, selection, analysis, and critical evaluation of the literature is described in this section.

7.1 Literature Search Methodology

It seems that the literature search process plays an important role in crafting a comprehensive analysis of a topic. The literature survey of this article is based on the search methodology adopted by Webster and Watson [113]. The systematic identification of high-quality publications (review articles, journals, conferences, and books), technical reports, and dissertations are reflections of the correct selection of databases, keywords, the time covered, the papers considered in the literature search, and performing backward and forward searches [114].

Figure 5 is a description of the methodology used for this article’s literature search. The following academic research databases were considered: IEEE Xplore (IEEE/IET) digital library, Elsevier ScienceDirect, Association for Computing Machinery (ACM) digital library, SpringerLink, and others. To find relevant papers, the flowchart of Figure 5 is applied for each of these academic research databases. In the first step, keywords using Google Scholar and Microsoft Academic were identified with respect to defence countermeasures. “Smart Grid”, “power system”, “false data injection”, and “cyber security” are common keywords used along with “detection”, “defence”, “mitigation”, and “countermeasure”.

7.2 Literature Selection and Analysis

Primarily, we reflect entirely on the defence of FDI threats with respect to Smart Grid cybersecurity, as there are also FDI articles related to other areas such as WSN, healthcare, software-defined networks, and so on. Another consideration is that while all of the scholarly research sources considered are prestigious and are assumed to publish quality works, further evaluations were
Table 4. Summary of Relevant Publications

| Database source | No. of relevant papers | Survey articles | Original research articles | Conferences | Books/Thesis |
|-----------------|------------------------|-----------------|---------------------------|-------------|--------------|
| IEEE Xplore     | 84                     | 4               | 57                        | 23          | 0            |
| Elsevier SD     | 7                      | 2               | 5                         | 0           | 0            |
| ACM             | 3                      | 1               | 1                         | 1           | 0            |
| Springer        | 5                      | 1               | 2                         | 1           | 1            |
| Others          | 5                      | 0               | 3                         | 2           | 0            |
| **Total**       | **104**                | **8**           | **68**                    | **27**      | **1**        |

Table 5. Evaluation Criteria for the Defence Strategies against FDI Attacks in Smart Grids

| Criterion                  | Description                                                                 |
|----------------------------|-----------------------------------------------------------------------------|
| Attack model               | Review the countermeasures from the point of view of the considered attack models. |
| Power flow model           | Adversaries use different approaches with different power flow models; countermeasures are reviewed and compared accordingly. |
| Defence algorithm          | Review various defence techniques studied in the literature.                |
| Network architecture       | Relevant articles are reviewed from network-centric point of view.          |
| Attack target              | Articles are compared on the basis of the attack target.                    |
| Performance metric         | Show the main claim of the research exemplifying the performance.            |
| Experimental platform      | Show the theoretical proofs or hardware testbeds utilized to justify the method. |

made using scientific journal ranking platforms to assess quality of the journals and the CORE\(^5\) was used for the conferences. Based on this search method, a systematic literature selection and analysis were used, which are described here. First, aggressive search was conducted using the keywords mentioned earlier and step 2 of Figure 5 that resulted in an abundant number of papers. Then, after a systematic refinement across the subcategories of the taxonomy of the FDI attack mitigation techniques, relevant works were selected. In addition to the keywords, titles and abstracts were considered for correctly categorising the selected papers. Next, important concepts were assembled for each of the chosen articles, accompanied by an overview of research results, and a thorough analysis. After an in-depth analysis of the literature, approximately 111 papers were found that, to varying degrees, dealt with the topic of defence against FDI attacks in Smart Grid cybersecurity. Note that the study of FDI attacks in Smart Grid started in late 2009. Therefore, the search for the most relevant literature of our survey starts from 2009 up to October 30, 2020 although related literature, such as that covering BDD, goes back before 2009. Table 4 is a summary of the number and source of the relevant publications considered in this survey article.

### 7.3 Evaluation Criteria

In order to quantify the efficacy and associated challenges of the various defence strategies, several key evaluation criteria for the proposed algorithms in relation to Smart Grid cybersecurity requirements are proposed. Table 5 summarizes the assessment criteria used to compare the selected defence algorithms against FDI attacks.

\(^5\)CORE: Computing Research and Education Association of Australasia ([https://www.core.edu.au/](https://www.core.edu.au/)).
One of the main evaluation criteria is the defence algorithm, a criterion that reflects the reviewed defence techniques. The attack construction methodologies discussed in Section 5.2 have been considered for the attack model criterion. These are attack with complete information, attack with partial information, LR attacks, GT attacks, and data-driven attacks. The AC and DC models are considered for the power flow model. The reviewed articles are also evaluated from a network-centric point of view (considering centralised and decentralised architecture). Additionally, the defence strategies are examined in light of the numerous cyber-physical entities that are vulnerable to FDI attack. In this regard, seven major Smart Grid components are identified for attack target evaluation criteria, including EMS, AGC, DEMS, MMS, network and communications, intelligent devices, and renewable resources. Note that the discussion in Section 3 shows the various components of the Smart Grid, and Section 5.3 discusses the vulnerability issue. Finally, two evaluation criteria, performance metrics and experimental platform, have been inspected. The evaluation criteria are used to compare and contrast among the various defence strategies as detailed in Section 8 and summarised in Table 6.

8 COMPARISON AND STATISTICS AMONG DEFENCE STRATEGIES

In this article, 104 publications are considered for the defence class. Here, the various countermeasure strategies are compared and some statistical facts based on the evaluation criteria are presented.

8.1 The Defence Strategies

The conventional BDDs, $\chi^2$ and LNR, and detection based on SE partitioning are merely used for bad data processing (see Section 6.1 for details). Consequently, the literature considered in this class did not take into account FDI attacks. Nevertheless, they have been blended with a variety of other approaches for detecting FDI attacks and serve as the basis for most countermeasure techniques. For example, the $\chi^2$ detector and the LNR detector have been employed in detection based on the dynamic SE subcategory.

Data-driven techniques and detection based on statistical models are the two most popular defence categories, comprising just under half of the total, with the former standing at approximately 25% and the latter at 23% of the total. Because of the complexity of Smart Grid infrastructure, the sheer volume of data, and the fact that high-performance computing devices are becoming available, data-driven techniques are increasingly powering various applications of the smart power system. As a result, plenty of data-driven defence techniques, especially DL and RL, are currently being pursued as a means of developing more effective detection against FDI attacks (as demonstrated in Section 6.4). The optimal placement of PMU, optimal selection of measurement quantities, and MTD, all under the protection-based category, are the other prominent defence strategies against false data attacks in Smart Grid cybersecurity (standing at 21%). Prevention-based defences (cryptographic functions and Blockchain technologies) are among the emerging security control mechanisms against incumbent cyberattacks. These techniques are especially popular across the demand side management (i.e., consumption side) of the Smart Grid.

8.2 Defence against Attack Mapping

The relationship between each defence algorithm and corresponding attack models is shown in Table 6. Although it is very challenging to establish a mapping relationship between the defence categories and attack models because each defence method can take a distinct adversarial approach, the following are the main findings for a mapping between the defence subcategories and attack models, as shown in Table 6.
Table 6. Comparison of Defence Strategies against FDI Attacks in Smart Grid Cybersecurity

| Category | Sub-category | Algorithm | Reference | Complete info | Partial info | LR attack | GT attack | Data-driven | EMS | AGC | DEHS | MASS | Core | Intelligent device | Renewable HRE | Exp. platform | Base system | Simulation | Test bed |
|----------|--------------|-----------|-----------|---------------|--------------|-----------|-----------|-------------|-----|-----|------|------|------|-------------------|--------------|-------------|-----------|-----------|---------|
|          |              | x²-detector | [35] A,c | ✓ | ✓ | SCADA | Estimated error vs r | 30 | ✓ |
|          |              |           | [51] D,c | ✓ | ✓ | SCADA | Estimated error vs r | 265 | ✓ |
|          |              | LNR detector | [49] A,c | ✓ | ✓ | PMU | Normalised residue | 30 | ✓ |
|          |              |           | [50] A,c | ✓ | ✓ | PMU | Normalised residue vs r, PE | 14 | ✓ |
|          |              | SE partitioning | [37] A,c | ✓ | ✓ | SCADA | DR vs τ | 39 | ✓ |
|          |              |           | [40] A,c | ✓ | ✓ | AMI | DR, FAR | 9 | ✓ |
|          |              | KF & extensions | [41] A,c | ✓ | ✓ | SCADA, PMU | DR, FDI | 14, 300 | ✓ |
|          |              |           | [42] A,c | ✓ | ✓ | SCADA | DR, FAR | 118 | ✓ |
|          |              |           | [52] A,c | ✓ | ✓ | SCADA | DR, τ | 14 | ✓ |
|          |              | Spatio-temporal correlations | [39] D, ed, RL | ✓ | ✓ | ✓ | AMI | ✓ | ✓ |
|          |              | State forecasting | [44] A,c | ✓ | ✓ | SCADA | DR, FAR | 118 | ✓ |
|          |              |           | [45] A,c | ✓ | ✓ | SCADA | MSE | 8 | ✓ |
|          |              | Integer LP | [19] A,c | ✓ | ✓ | SCADA | Cost of undetected attack vs number of PMUs | multiple | ✓ |
|          |              | MISDP | [19] A,c | ✓ | ✓ | SCADA | Cost of undetected attack vs number of PMUs | multiple | ✓ |
|          |              | Greedy algorithm | [26] D,c | ✓ | ✓ | SCADA | Subset of meters protection | multiple | ✓ |
|          |              |           | [20] D,c | ✓ | ✓ | SCADA | Attack cost vs PMU placement | multiple | ✓ |
|          |              | Greedy algorithm | [53] D,c | ✓ | ✓ | SCADA | SE error deviation vs placement | multiple | ✓ |
|          |              |           | [47] A,c | ✓ | ✓ | SCADA | PMU placement vs attack cost, time overhead | 9, 14, 30 | ✓ |
|          |              | Hybrid | [48] A,c | ✓ | ✓ | ✓ | SCADA | Defence probability vs nodes | 14 | ✓ |
|          |              | Greedy algorithm | [54] D,c | ✓ | ✓ | SCADA | DR vs SR, DR vs FAR | 9, 14, 57 | ✓ |
|          |              |           | [23] D,c | ✓ | ✓ | SCADA | DR vs SR, protected sensors | multiple | ✓ |
|          |              | Greedy algorithm | [55] D,d | ✓ | ✓ | SCADA | # of protected meters vs attack cost | multiple | ✓ |
|          |              |           | [56] D,c | ✓ | ✓ | SCADA | Attack cost vs # of protected IEDs | 14, 118 | ✓ |
|          |              | Game-theoretic | [59] D,c | ✓ | ✓ | ✓ | SCADA | Optimal meter protections | 14, 30 | ✓ |
|          |              |           | [60] D,c | ✓ | ✓ | ✓ | SCADA | Load shedding cost | 5, 9, 14 | ✓ |
|          |              |           | [34] D,c | ✓ | ✓ | ✓ | SCADA | Load shedding cost | 9, 14 | ✓ |
| Protection-based defence | Grid topology perturbation | MTD | Graph-theoretic |  |
|-------------------------|--------------------------|-----|----------------|---|
| [57]^{D,c}              | ✓ ✓ ✓ SCADA ✓            | ✓ ✓  | Optimal meter protections | 14, 57, ✓ |
| [58]^{D,c}              | ✓ ✓ ✓ SCADA ✓            | ✓ ✓  | Optimal meter protections | 30, 57, ✓ |
| [61]^{P,c}              | ✓ ✓ ✓ SCADA ✓            | ✓ ✓  | Power loss        | multiple ✓ |
| [62]^{P,c}              | ✓ ✓ ✓ SCADA ✓            | ✓ ✓  | DR, FDI           | 14 ✓         |
| [63]^{D,c}              | ✓ ✓ ✓ SCADA ✓            | ✓ ✓  | OPF cost, DR vs FAR | 14 ✓         |
| [64]^{P,c}              | ✓ ✓ ✓ SCADA ✓            | ✓ ✓  | AR               | 14 ✓         |
| [65]^{A,c}              | ✓ ✓ ✓ SCADA ✓            | ✓ ✓  | DR vs attacked states | 6, 57 ✓         |
| [66]^{A,c}              | ✓ ✓ ✓ SCADA ✓            | ✓ ✓  | Meter protection cost, PE vs FDI | 6, 14, 57 ✓         |
| [67]^c                  | ✓ ✓ ✓ SCADA ✓            | ✓ ✓  | DR vs FDI, TPR vs FAR | 39 ✓         |
| [68]^{A,c}              | ✓ ✓ ✓ SCADA ✓            | ✓ ✓  | DR vs FDI, TPR vs FAR | 14 ✓         |
| [69]^{D,c,RL}           | ✓ ✓ ✓ SCADA ✓            | ✓ ✓  | DR vs SR, DR vs perturbation ratio | 57, 118 ✓         |
| [70]^{A,d}              | ✓ ✓ ✓ SCADA ✓            | ✓ ✓  | Reactance rate, Power loss | 66 ✓         |
| [71]^{T_1,norm min.}    | ✓ ✓ ✓ ✓ ✓ ✓              | ✓ ✓  | DR, FAR           | 14 ✓         |
| [72]^{D,c}              | ✓ ✓ ✓ ✓ ✓ ✓              | ✓ ✓  | DR, FAR           | 30 ✓         |
| [73]^{A,c}              | ✓ ✓ ✓ ✓ ✓ ✓              | ✓ ✓  | payoffs           | 14 ✓         |
| [74]                    | ✓ ✓ ✓ ✓ ✓ ✓              | ✓ ✓  | WSN               | DR, FAR - ✓ |
| [75]                    | ✓ ✓ ✓ ✓ ✓ ✓              | ✓ ✓  | MSE, FPR          | - ✓         |
| [76]^{P,c,d}            | ✓ ✓ ✓ ✓ ✓ ✓              | ✓ ✓  | DD, FDI           | 14 ✓         |
| [77]^{P,c}              | ✓ ✓ ✓ ✓ ✓ ✓              | ✓ ✓  | DR, DD, FAR       | 4 ✓         |
| [78]^{P,d}              | ✓ ✓ ✓ ✓ ✓ ✓              | ✓ ✓  | DD, FAR           | multiple ✓ |
| [79]^{D,d}              | ✓ ✓ ✓ ✓ ✓ ✓              | ✓ ✓  | DD, FAR           | 13 ✓         |
| [80]^{D,d}              | ✓ ✓ ✓ ✓ ✓ ✓              | ✓ ✓  | DR, meters and DD, FPR | 14 ✓         |
| [81]^{D,d}              | ✓ ✓ ✓ ✓ ✓ ✓              | ✓ ✓  | SCADA             | DD, FAR 14 ✓ |
| [82]^{P,c,RL}           | ✓ ✓ ✓ ✓ ✓ ✓              | ✓ ✓  | SCADA             | DD, FDI 14 ✓ |
| [83]^{P,c,RL}           | ✓ ✓ ✓ ✓ ✓ ✓              | ✓ ✓  | SCADA             | DR, FDI 14 ✓ |
| [84]^{P,c,RL}           | ✓ ✓ ✓ ✓ ✓ ✓              | ✓ ✓  | SCADA             | DR, FDI 14 ✓ |
| [85]^{P,c,RL}           | ✓ ✓ ✓ ✓ ✓ ✓              | ✓ ✓  | SCADA             | DR, FDI 14 ✓ |
| [86]^{D,c,RL,SR}        | ✓ ✓ ✓ ✓ ✓ ✓              | ✓ ✓  | SCADA             | TPR, FPR, SR, SNR 57& 118 ✓ |
| [87]^{D,c}              | ✓ ✓ ✓ ✓ ✓ ✓              | ✓ ✓  | SCADA             | TPR, FPR, SR, SNR 57& 118 ✓ |
| [88]^{D,d}              | ✓ ✓ ✓ ✓ ✓ ✓              | ✓ ✓  | PMU               | DA, TPR, FPR 118 ✓ |
| [89]^{D,c}              | ✓ ✓ ✓ ✓ ✓ ✓              | ✓ ✓  | SCADA             | DA, P, R 9, 57, 118 ✓ |
| [90]                    | ✓ ✓ ✓ ✓ ✓ ✓              | ✓ ✓  | SCADA             | P, R, F1 118 ✓ |
| [91]^{D,c}              | ✓ ✓ ✓ ✓ ✓ ✓              | ✓ ✓  | SCADA             | F1, DR, FAR, 118 ✓ |
| [92]^{D,c}              | ✓ ✓ ✓ ✓ ✓ ✓              | ✓ ✓  | SCADA             | P, MSE 14 ✓ |
| [93]^{P,c,RL}           | ✓ ✓ ✓ ✓ ✓ ✓              | ✓ ✓  | SCADA             | AUC 14 ✓ |
| [94]^{P,c,RL}           | ✓ ✓ ✓ ✓ ✓ ✓              | ✓ ✓  | SCADA             | F1, DA 30 ✓ |

ACM Computing Surveys, Vol. 55, No. 14s, Article 331. Publication date: July 2023.
| Data-driven | Semi-supervised ANN | [88]\(^{D,d}\) | ✓ | ✓ | ✓ | PMU | DA, P, R | 9, 57, 118 ✓ |
| Semi-supervised GMM | [91]\(^{D,c}\) | ✓ | ✓ | ✓ | SCADA | Fi, DR, FAR, 118 ✓ |
| DFFNN | [95]\(^{A,c}\) | ✓ | ✓ | ✓ | SCADA | DA, P, R, TPR vs FPR, 14 ✓ |
| CDBN | [96]\(^{D,c,RL}\) | ✓ | ✓ | ✓ | ✓ | SCADA | DA, TPR vs FPR, 118, 300 ✓ |
| DRNN | [97]\(^{D,c}\) | ✓ | ✓ | ✓ | PMU | DA, TPR, FPR, 118, 300 ✓ |
| CNN | [99]\(^{A,c}\) | ✓ | ✓ | ✓ | SCADA | Location and attack DA, P, R, TPR vs FPR, 14 ✓ |
| GAN | [100]\(^{A,c}\) | ✓ | ✓ | ✓ | PMU | DA, P, R, 13, 123 ✓ |
| RL | Q-learning | [31]\(^{A,c,RL}\) | ✓ | ✓ | ✓ | ✓ | SCADA | ✓ | Voltage sag, DA, 39 ✓ |
| SARSA | [30]\(^{A,c}\) | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | DD, FAR, P, R, 14 ✓ |
| Bayesian Bandit | [102]\(^{D,c}\) | ✓ | ✓ | ✓ | ✓ | SCADA | ✓ | DR, MSE, 14 ✓ |
| DRL | Deep-Q-network | [103]\(^{A,c}\) | ✓ | ✓ | ✓ | ✓ | ✓ | DD, FAR, 9, 14, 30 ✓ |

| Prevention | Cryptographic schemes | Encryption/Decryp. | [56]\(^{D,c}\) | ✓ | ✓ | ✓ | SCADA | ✓ | Attack cost vs # of protected IEDs, 14, 118 ✓ |
| | | | [104]\(^{D,c}\) | ✓ | ✓ | ✓ | ✓ | ✓ | Modbus, AMI | Packet loss, Computational cost - ✓ |
| | | | [105]\(^{D,c}\) | ✓ | ✓ | ✓ | ✓ | ✓ | DNP3 | Packet loss, Communication overhead, Computational cost - ✓ |
| | | | [106]\(^{D,c}\) | ✓ | ✓ | ✓ | ✓ | ✓ | SCADA, IEC 61850 | Delay, Packet loss, Comm. overhead, Comp. cost - ✓ |
| | | | [107]\(^{D,c}\) | ✓ | ✓ | ✓ | ✓ | ✓ | AMI | Latency, Comm. overhead, Comp. cost - ✓ |
| | | | [108]\(^{D,c}\) | ✓ | ✓ | ✓ | ✓ | ✓ | AMI | Comm. overhead, Energy overhead, Comp. cost - ✓ |
| | | | [109]\(^{D,c}\) | ✓ | ✓ | ✓ | ✓ | ✓ | Wireless comm., C37.118 | Signature overhead 42 ✓ ✓ |
| | Blockchain-based defence | Data prot. | [110]\(^{D,c}\) | ✓ | ✓ | ✓ | ✓ | ✓ | AMI | AR vs manipulated meters 118 ✓ |
| | | | [111]\(^{D,c}\) | ✓ | ✓ | ✓ | ✓ | ✓ | AMI | Transaction verification of energy supply–demand - ✓ |
| | Privacy preserving | | [21]\(^{D,c}\) | ✓ | ✓ | ✓ | ✓ | ✓ | SCADA | DA, DR vs FAR - ✓ |
| | | | [22]\(^{D,c}\) | ✓ | ✓ | ✓ | ✓ | ✓ | Smart meter | Transaction delay, Computational cost - ✓ |

[Ref\(^{D,A}\): DC/AC model; Ref\(^{D,c}\): Centralised/decentralised architecture; Ref\(^{D,c,d}\): Centralised and decentralised architectures; Ref\(^{RL}\): Real load data considered, DR: Detection rate; DD: Detection delay; FPDR: False positive DR, DA: Detection accuracy; FPR: False-positive rate; TPR: True-positive rate; FDI: Injected magnitude of FDI attack; payoffs: Game metric of attacker–defender cost in payoffs; SR: FDI attack sparsity ratio; SNR: Signal-to-noise ratio; MAPE: Mean absolute percentage error; PE: Percentage error between true and estimated states; AR: attacking rate (attackability, or successful attacking probabilities); MSE: Mean square error.
A Taxonomy of Cyber Defence Strategies Against False Data Attacks in Smart Grids

Fig. 6. Defence against attack mapping considering FDI.

(a) The FDI attack class defined by the conventional BDD techniques is the one with complete topology information. (b) Detection techniques based on dynamic SE consider FDI with complete topology information and LR attacks. (c) Under the protection-based defence category, the optimal PMU placement methods consider FDI attacks with partial topology information. (d) Under the protection-based defence category, the optimal measurement selection methods consider FDI attacks with complete topology information, partial topology information, and LR attacks. (e) Most of the grid topology perturbation of the protection-based defence category considers FDI attacks with complete topology information, followed by partial topology information. (f) With the exception of a few methods that consider data-driven attacks and attacks with partial topology information, most statistical modeling-based detection methods take into account FDI attacks with complete topology information. (g) Data-driven attacks and attacks with partial topology information are used by the majority of data-driven detection approaches. Figure 6 summarizes the defence against attack mapping, considering the defence subcategories and various FDI attack models.

8.3 Performance Metric

The defence strategies vary, among other things, in terms of algorithmic design, adversarial method, attack target, and network architecture. For this reason, instead of providing a distinct performance metrics for all attack countermeasures, we present comprehensive qualitative metrics. Numerous performance metrics are presented for each of the countermeasure subcategories (see the 17th column of Table 6). For example, across the protection-based defence category, optimal subset of meter, optimal IED protection, and attack cost are the main metrics considered. Packet loss, computational cost, communication cost, and end-to-end delay are the main evaluation metrics adopted among the prevention schemes. In most of the detection based on dynamic SE, statistical-based models, and data-driven defence categories, detection rates in terms of probability of detection and True-Positive Rate (TPR) are compared against False-Positive Rates (FPR) or False Alarm Rates (FAR). For the adversarial models, the metrics used in the literature include magnitude of attack injection, attack sparsity, payoffs, game metric of attacker-defender cost in payoffs, attacking rate, and attacking probability, as shown in Table 6.

8.4 Experimental Platform

The vast majority of studies yielded numerical results based on simulations of IEEE standard or modified electric grid test cases. The 18th column of Table 6 contains information about the various
benchmark systems. A large number of test cases (that include small, medium, and very large sized bus systems) have been considered. Our review result indicates that the IEEE 14 bus system is the most widely referred test case. Although the vast majority of works use only a single test case to conform their numerical results, some considered multiple test cases. To further verify the efficacy of their proposal, some scholars used real-time load data, most of which used a dataset from the New York Independent System Operator. Almost all of the studies are based on simulations using MATPOWER, a MATLAB power system toolbox, and a few others utilise PowerFactory and the TOMLAB optimization toolbox. Finally, a few incorporated co-simulations and hardware testbeds.

9 MAIN GAPS OF EXISTING DEFENCE STRATEGIES AGAINST FDI ATTACKS

While detailed research reviews of each defence category have been addressed in Section 6, in what follows, we describe the key gaps of existing defence studies.

Some Emerging Smart Grid Areas Are Not Well Studied: The plethora of literature examined in this review article tried to cover a multitude of Smart Grid infrastructures. However, there are some open issues with respect to the scope (network architecture, DERs, and communication systems). The majority of existing countermeasure studies have focused on the traditional centralised EMS. While decentralised energy generation and distribution systems (such as DERs) have become very popular, they have been among the most vulnerable cyber-physical components to FDI attacks. However, only a few studies have been undertaken with respect to defence strategies of DERs. This can be seen from the 16th column of Table 6. Further, only a few papers have discussed the SAS-, AMI-, and WAMS-based communication systems.

Numerous works assert that preventive security measures are essential in the fight against FDI attacks in the Smart Grid (see Section 6.5). Lightweight cryptography and blockchain-based security systems are the least studied areas.

Throughout this report, it has been mentioned that the power system measurement data can reveal anomalies in the face of cyberattacks. It is also highly likely that physical faults contribute to the abnormal functioning of power systems. Therefore, the research on FDI attacks can be extended with respect to the identification between cyber attacks and power physical faults. Differentiating between cyber threats and physical faults can be beneficial for operators as it helps them to prevent unnecessary losses. Only two studies \[117, 118\] have been done in this respect. A real-time detection scheme is required that considers the sparsity of FDI attacks and low-dimensional property of the measurement data received at the control center.

General Shortcomings of the Countermeasures: Performance, Computational Cost, and Feasibility of Deployment: The conventional BDD-based detection methods have not been able to handle stealthy and sparse FDI attacks and are thus vulnerable to FDI attacks. The numerous defence algorithms analysed in the literature have achieved much stronger security controls against incumbent cyberattacks. However, there are certain limitations that are worth mentioning here. In spite of their potential to defend key grid components against bad data injection attacks, protection-based defence schemes have certain drawbacks. First, deployment of PMUs in the large-scale Smart Grid is much more expensive, especially with the emerging ubiquitous sensing infrastructure. It has also been shown \[33\] that PMUs are susceptible to the injection of false data attacks via GPS spoofing, which requires a more appealing security scheme. Additionally, determining the subset of measurements is a large-complexity problem. MTD can allow grid operators to proactively protect measurements from malicious attackers by introducing perturbations to network data or topology that can inevitably lead to uncertainties and costs against adversaries. Yet, MTD protection approaches can be compromised by intelligent FDI attackers if the attackers can identify MTD changes before they perform the attack (as has been demonstrated in \[68\]).
It has been seen that detection based on dynamic SEs are more powerful countermeasure techniques than BDD techniques. However, the use of WLS and KF-based signal processors incurs an immense computational burden. Physics-aware data-driven defence approaches, on the other hand, are much more robust for power system security, especially for dynamically changing power system variables. In addition, the prevalence of GPUs makes it practical to satisfy the computational requirements of advanced ML models. Currently, deep neural networks, reinforcement learning (RL), and the convergence of the two are the most favoured ML models for detecting FDI cyberattacks.

Although currently missing in the literature using commercial-level datasets of stealthy FDI attacks is a practical way to verify the efficacy of data-driven countermeasure techniques.

**Need for Corroboration of Experimental Results Via Testbed Platform:** Although the works surveyed here have proven their cybersecurity solutions via numerical simulations benchmarked against standardised test cases, it is vital to validate experimental results via cyber-physical testbeds, which is missing in the literature except for a few articles ([31, 90, 109]). This downside can be seen from the perspectives of data- and system-oriented approaches. Most of the FDI attack schemes surveyed did not consider commercial-level datasets, which otherwise can practically validate the vulnerability of the state estimators to the stealthy FDI attacks. Even more significant, the countermeasure techniques can also incorporate real-world datasets.

Testbeds [119] are essential tools for testing the performance evaluation of algorithms and protocols in the Smart Grid. The highly complex and multidisciplinary essence of the Smart Grid requires the implementation of cyber-physical testbeds with different characteristics for comprehensive experimental validation. There is a considerable need to analyse new Smart Grid security concepts, architectures, and vulnerabilities via cyber-physical system test platforms. Recently, there has been growing attention to the study of cyber-physical Smart Grid testbeds [119]. Most notably, hardware-in-the-loop test platforms have become much more popular for the development, analysis, and testing of cyber-physical components of the electrical power system. For example, some Smart Grid stakeholders, such as ABB, Siemens Power Technologies, and OPAL RT, foster hardware-in-the-loop testing using real-time digital simulators across various Smart Grid realms, including microgrids and SAS- and WAMS-based protection environments. Therefore, we suggest that assessing the effects of FDI attacks on the Smart Grid using the hardware-in-the-loop testbed platform is critical in crafting stringent cybersecurity requirements.

**10 EMERGING ADVANCED APPLICATIONS: FUTURE RESEARCH DIRECTIONS**

Securing the electricity grid is one of the highest priorities of many countries around the world. Academic studies and industries are expected to tackle a range of issues for future research on cyber defence in the Smart Grid infrastructure. Particularly, the reliance of reliable and secure power system operation on the communication infrastructure, along with potential cyber threats, are increasingly growing. In the following, emerging advanced applications are discussed as potential future research prospects.

**Cybersecurity for Emerging Smart Grid Communication Systems:** Despite the fact that the communication infrastructure is the most critical target of FDI attacks, the countermeasures have to be studied well, especially across the SAS-compliant IEC 61850 and the WAMS-compliant IEEE C37.118. The FDI attack can be studied well with the incorporation of cyber-physical testbed platforms [119]. Moreover, although AMI is one of the most vulnerable communication systems to FDI attacks, little has been done on the defence against this attack. Given the increasing adoption of the IoT in the Smart Grid, it will be interesting to address cybersecurity issues of IoT-based AMI with regard to FDI attacks. Software-defined networking is one of the emerging networking applications. The coupling of software-defined networking with Smart Grid applications can yield efficient network monitoring. However, the security issue of this technology is worth
investigating, especially with respect to FDI attacks. Further, it has been indicated that cognitive radio can help the implementation of a control-sensing mechanism to identify and account for the detection of FDI attacks in the Smart Grid \cite{102}. In addition, countermeasures against FDI attacks on heterogeneous cognitive radio \cite{15}, WSNs, and the IoT are potential cybersecurity studies that are worth investigating. Specifically, the application of data-driven models along with the countermeasure strategies across the more intelligent communication arena of the Smart Grid seems to be a promising solution in tackling orchestrated cyberattacks.

**Security Framework Based on Lightweight ML:** Countless memory and computational-restricted wireless sensor nodes are connected to IoT applications in the Smart Grid. Several reports have shown that such limitations raise obstacles to the usage of conventional security measures over IoT systems. Security frameworks using lightweight ML can be proposed for resource-constrained IoT devices. For example, lightweight ML can be proposed for prevention schemes such as encryption, message authentication, and dynamic key management against false data attacks in an end-to-end Smart Grid communication system.

**FDI Attack Detection in Edge Computing:** The growing popularity of distributed renewable energy generation requires reduced processing costs and communication overhead. In a distributed computing environment, edge computing improves communication overhead and system bandwidth by bringing the processing and data storage near to the origin of the data source. Further, the emergence of Industry 4.0 across a number of industries, including the Smart Grid, brings ubiquitous networked elements, and intelligent edge computing. Although intelligent edge computing is expected to be able to meet the needs of the ever-increasing number of IoT users in the Smart Grid, there are inherent security threats. For example, bringing more of such IoT devices to the edge network can introduce various cybersecurity threats. FDI attacks can be challenging in the edge computing environment. Distributed detection using DL or deep RL against incumbent attacks can be a potential research direction in edge computing-based Smart Grids.

**Blockchain Technology:** A blockchain-based defence for privacy preservation and anomaly detection in Smart Grids is a very new research area, which requires further investigation.

## 11 Conclusion

The Smart Grid faces a growing threat from an emerging cyber-physical attack called FDI. By injecting stealthy falsified attack vectors, adversaries can compromise critical Smart Grid information, render the power system unobservable, and may culminate in large-scale failure of the power system operation. This survey article analysed extensive reviews of existing state-of-the-art studies on cyber defence against incumbent cyberattacks in Smart Grids. A taxonomy of five major categories and subcategories of the different countermeasures was proposed. Furthermore, in order to quantify the efficacy and associated challenges of the various proposed algorithms in the literature surveyed, key evaluation criteria were used in relation to the requirements of the power systems and Smart Grid cybersecurity. Future research directions for mitigation techniques to combat FDI attacks were proposed as a way of advancing the Smart Grid cybersecurity framework.

## References

1. Yao Liu, Peng Ning, and Michael K. Reiter. 2011. False data injection attacks against state estimation in electric power grids. *ACM Transactions on Information and System Security (TISSEC)* 14, 1 (2011), 1–33.
2. Adnan Anwar. Nov. 2017. *Data-Driven Stealthy Injection Attacks on Smart Grid*. Ph.D. Dissertation. School of Engineering and Information Technology, University of New South Wales.
3. Haftu Tasew Reda, Adnan Anwar, Abdun Mahmood, and Naveen Chilamkurti. 2023. Data-driven approach for state prediction and detection of false data injection attacks in smart grid. *Journal of Modern Power Systems and Clean Energy* 11, 2 (2023), 455–467.
A Taxonomy of Cyber Defence Strategies Against False Data Attacks in Smart Grids

[4] Y. Gu, T. Liu, D. Wang, X. Guan, and Z. Xu. 2013. Bad data detection method for smart grids based on distributed state estimation. In 2013 IEEE International Conference on Communications (ICC). IEEE, Budapest, Hungary, 4483–4487.

[5] M. Gönl and A. Abur. 2015. A modified Chi-Squares test for improved bad data detection. In 2015 IEEE Eindhoven PowerTech. IEEE, Eindhoven, Netherlands, 1–5.

[6] Y. Zhou and L. Xie. 2017. Detection of bad data in multi-area state estimation. In 2017 IEEE Texas Power and Energy Conference (TPEC). IEEE, College Station, TX, USA, 1–6.

[7] Zhihao Guan, Nan Sun, Yue Xu, and Tingting Yang. 2015. A comprehensive survey of false data injection in smart grid. International Journal of Wireless and Mobile Computing 8, 1 (2015), 27–33.

[8] R. Deng, G. Xiao, R. Lu, H. Liang, and A. V. Vasilakos. 2017. False data injection on state estimation in power systems—attacks, impacts, and defense: A survey. IEEE Transactions on Industrial Informatics 13, 2 (2017), 411–423.

[9] Xuan Liu and Zuyi Li. 2017. False data attack models, impact analyses and defense strategies in the electricity grid. The Electricity Journal 30, 4 (2017), 35–42.

[10] G. Liang, J. Zhao, F. Luo, S. R. Weller, and Z. Y. Dong. 2017. A review of false data injection attacks against modern power systems. IEEE Transactions on Smart Grid 8, 4 (2017), 1630–1638.

[11] Q. Wang, W. Tai, Y. Tang, and M. Ni. 2019. Review of the false data injection attack against the cyber-physical power system. IET Cyber-Physical Systems: Theory Applications 4, 2 (2019), 101–107.

[12] Meng Zhang, Chao Shen, Ning He, SiCong Han, Qi Li, Qian Wang, and XiaoHong Guan. 2019. False data injection attacks against smart grid state estimation: Construction, detection and defense. Science China Technological Sciences 62, 12 (2019), 2077–2087.

[13] Ahmed S. Musleh, Guo Chen, and Zhao Yang Dong. 2019. A survey on the detection algorithms for false data injection attacks in smart grids. IEEE Transactions on Smart Grid 11, 3 (2019), 2218–2234.

[14] Souhila Aoufi, Abdelouahid Derhab, and Mohamed Guerroumi. 2020. Survey of false data injection in smart power grid: Attacks, countermeasures and challenges. Journal of Information Security and Applications 54 (2020), 102518.

[15] Haftu Tasew Reda, Abdun Mahmood, Abebe Diro, Naveen Chilamkurti, and Suresh Kallam. 2020. Firefly-inspired stochastic resonance for spectrum sensing in CR-based IoT communications. Neural Computing and Applications 32, 20 (2020), 16011–16023.

[16] Patrick D. Gallagher. Volume 1 - Smart Grid Cybersecurity Strategy, Architecture, and High-Level Requirements. National Institute of Standards and Technology. Retrieved April 5, 2019 from https://nvlpubs.nist.gov/nistpubs/ir/2014/NIST.IR.7628r1.pdf.

[17] V. Cagri Gungor, Dilan Sahin, Taskin Kocak, Salih Ergut, Concettina Buccella, Carlo Cecati, and Gerhard P. Hancke. 2012. A survey on smart grid potential applications and communication requirements. IEEE Transactions on Industrial Informatics 9, 1 (2012), 28–42.

[18] M. A. Rahman and H. Mohsenian-Rad. 2012. False data injection attacks with incomplete information against smart power grids. In 2012 IEEE Global Communications Conference (GLOBECOM’12). IEEE, Anaheim, CA, 3153–3158.

[19] Annarita Giani, Russell Bent, and Feng Pan. 2014. Phasor measurement unit selection for unobservable electric power data integrity attack detection. International Journal of Critical Infrastructure Protection 7, 3 (2014), 155–164.

[20] Q. Yang, D. An, R. Min, W. Yu, X. Yang, and W. Zhao. 2017. On optimal PMU placement-based defense against data integrity attacks in smart grid. IEEE Transactions on Information Forensics and Security 12, 7 (2017), 1735–1750.

[21] M. Keshk, B. Turnbull, N. Moustafa, D. Vatsalan, and K. R. Choo. 2020. A privacy-preserving-framework-based blockchain and deep learning for protecting smart power networks. IEEE Transactions on Industrial Informatics 16, 8 (2020), 5110–5118.

[22] Xin Chen, Jiachen Shen, Zhenfu Cao, and Xiaolei Dong. 2020. A blockchain-based privacy-preserving scheme for smart grids. In Proceedings of the 2020 2nd International Conference on Blockchain Technology. ACM, Hilo, USA, 120–124.

[23] Rakesh B. Bobba, Katherine M. Rogers, Qiyan Wang, Himanshu Khurana, Klara Nahrstedt, and Thomas J. Overbye. 2010. Detecting false data injection attacks on DC state estimation. In Preprints of the First Workshop on Secure Control Systems—attacks, impacts, and defense: A survey. IEEE Transactions on Industrial Informatics 13, 2 (2017), 411–423.

[24] Yi Huang, H. Li, K. A. Campbell, and Zhu Han. 2011. Defending false data injection attack on smart grid network using adaptive CUSUM test. In 2011 45th Annual Conference on Information Sciences and Systems. IEEE, Baltimore, MD, 1–6.

[25] O. Kosut, L. Jia, R. J. Thomas, and L. Tong. 2011. Malicious data attacks on the smart grid. IEEE Transactions on Smart Grid 2, 4 (2011), 645–658.

[26] T. T. Kim and H. V. Poor. 2011. Strategic protection against data injection attacks on power grids. IEEE Transactions on Smart Grid 2, 2 (2011), 326–333.

[27] J. Lin, W. Yu, X. Yang, G. Xu, and W. Zhao. 2012. On false data injection attacks against distributed energy routing in smart grid. In 2012 IEEE/ACM 3rd International Conference on Cyber-Physical Systems. IEEE, Beijing, China, 183–192.
[33] Daniel P. Shepard, Todd E. Humphreys, and Aaron A. Fansler. 2012. Evaluation of the vulnerability of phasor measurement units to GPS spoofing attacks. *International Journal of Critical Infrastructure Protection* 5, 3–4 (2012), 146–153.

[34] J. Hao, E. Kang, J. Sun, Z. Wang, Z. Meng, X. Li, and Z. Ming. 2018. An adaptive Markov strategy for defending smart grid false data injection from malicious attackers. *IEEE Transactions on Smart Grid* 9, 4 (2018), 2398–2408.

[35] A. Abur and A. G. Exposito. 2004. *Power System State Estimation: Theory and Implementation*. CRC Press, Boca Raton: CRC, FL.

[36] Y. Wu, Y. Xiao, F. Hohn, L. Nordström, J. Wang, and W. Zhao. 2018. Bad data detection using linear WLS and sampled values in digital substations. *IEEE Transactions on Power Delivery* 33, 1 (2018), 150–157.

[37] T. Liu, Y. Gu, D. Wang, Y. Gui, and X. Guan. 2013. A novel method to detect bad data injection attack in smart grid. In *2013 Proceedings IEEE INFOCOM*, IEEE, Turin, Italy, 3423–3428.

[38] J. Zhao, A. Gómez-Expósito, M. Netto, L. Mili, A. Abur, V. Terzića, I. Kamwa, B. Pal, A. K. Singh, J. Qi, Z. Huang, and A. P. S. Meliopoulos. 2019. Power system dynamic state estimation: Motivations, definitions, methodologies, and future work. *IEEE Transactions on Power Systems* 34, 4 (2019), 3188–3198.

[39] D. B. Rawat and C. Bajracharya. 2015. Detection of false data injection attacks in smart grid communication systems. *IEEE Signal Processing Letters* 22, 10 (2015), 1652–1656.

[40] K. Manandhar, X. Cao, F. Hu, and Y. Liu. 2014. Detection of faults and attacks including false data injection attack in smart grid using Kalman filter. *IEEE Transactions on Control of Network Systems* 1, 4 (2014), 370–379.

[41] Nemanja Živković and Andrija T. Sarić. 2018. Detection of false data injection attacks using unscented Kalman filter. *Journal of Modern Power Systems and Clean Energy* 6, 5 (2018), 847–859.

[42] J. Zhao, G. Zhang, M. La Scala, Z. Y. Dong, C. Chen, and J. Wang. 2017. Short-term state forecasting-aided method for detection of smart grid general false data injection attacks. *IEEE Transactions on Smart Grid* 8, 4 (2017), 1580–1590.

[43] E. Drayer and T. Routtenberg. 2020. Detection of false data injection attacks in smart grids based on graph signal processing. *IEEE Systems Journal* 14, 2 (2020), 1886–1896.

[44] J. Chen and A. Abur. 2006. Placement of PMUs to enable bad data detection in state estimation. *IEEE Transactions on Power Systems* 21, 4 (2006), 1608–1615.

[45] M. Göl and A. Abur. 2013. PMU placement for robust state estimation. In *2013 North American Power Symposium (NAPS)*, IEEE, Manhattan, KS, 1–5.

[46] N. M. Manousakis and G. N. Korres. 2016. Optimal PMU placement for numerical observability considering fixed channel capacity—a semidefinite programming approach. *IEEE Transactions on Power Systems* 31, 4 (2016), 3328–3339.

[47] C. Pei, Y. Xiao, W. Liang, and X. Han. 2020. PMU placement protection against coordinated false data injection attacks in smart grid. *IEEE Transactions on Industry Applications* 56, 4 (2020), 4381–4393.

[48] Qi Wang, Wei Tai, Yi Tang, Ming Ni, and Shi You. 2019. A two-layer game theoretical attack-defense model for a false data injection attack against power systems. *International Journal of Electrical Power & Energy Systems* 104 (2019), 169–177.

[49] H. Khazraj, F. Faria da Silva, C. L. Bak, and U. Annakkage. 2017. Addressing single and multiple bad data in the modern PMU-based power system state estimation. In *52nd International Universities Power Engineering Conference (UPEC’17)*, IEEE, Heraklion, Greece, 1–6.

[50] George N. Korres and Nikolaos M. Manousakis. 2011. State estimation and bad data processing for systems including PMU and SCADA measurements. *Electric Power Systems Research* 81, 7 (2011), 1514–1524.

[51] Hadis Karimpour and Venkata Dinavahi. 2017. Robust massively parallel dynamic state estimation of power systems against cyber-attack. *IEEE Access* 6 (2017), 2984–2995.

[52] Rui Chen, Xue Li, Huixin Zhong, and Minrui Fei. 2019. A novel online detection method of data injection attack against dynamic state estimation in smart grid. *Neurocomputing* 344 (2019), 73–81.

[53] Q. Yang, L. Jiang, W. Hao, B. Zhou, P. Yang, and Z. Lv. 2017. PMU placement in electric transmission networks for reliable state estimation against false data injection attacks. *IEEE Internet of Things Journal* 4, 6 (2017), 1978–1986.
[54] J. Hao, R. J. Piechocki, D. Kaleshi, W. H. Chin, and Z. Fan. 2015. Sparse malicious false data injection attacks and defense mechanisms in smart grids. *IEEE Transactions on Industrial Informatics* 11, 5 (2015), 1–12.

[55] X. Liu, Z. Li, and Z. Li. 2017. Optimal protection strategy against false data injection attacks in power systems. *IEEE Transactions on Smart Grid* 8, 4 (2017), 1802–1810.

[56] G. Dán and H. Sandberg. 2010. Stealth attacks and protection schemes for state estimators in power systems. In *2010 1st IEEE International Conference on Smart Grid Communications*. IEEE, Gaithersburg, MD, 214–219.

[57] S. Bi and Y. J. Zhang. 2014. Graphical methods for defense against false-data injection attacks on power system state estimation. *IEEE Transactions on Smart Grid* 5, 3 (2014), 1216–1227.

[58] M. H. Ansari, V. T. Vakili, B. Bahrak, and P. Tavassoli. 2018. Graph theoretical defense mechanisms against false data injection attacks in smart grids. *Journal of Modern Power Systems and Clean Energy* 6, 5 (2018), 860–871.

[59] C. Liu, M. Zhou, J. Wu, C. Long, and D. Kundur. 2019. Financially motivated FDI on SCED in real-time electricity markets: Attacks and mitigation. *IEEE Transactions on Smart Grid* 10, 2 (2019), 1949–1959.

[60] C. Y. T. Ma, D. K. Y. Yau, X. Lou, and N. S. V. Rao. 2013. Markov game analysis for attack-defense of power networks under possible misinformation. *IEEE Transactions on Power Systems* 28, 2 (2013), 1676–1686.

[61] K. L. Morrow, E. Heine, K. M. Rogers, R. B. Bobba, and T. J. Overbye. 2012. Topology perturbation for detecting malicious data injection. In *2012 45th Hawaii International Conference on System Sciences*. IEEE, Maui, HI, 2104–2113.

[62] W. Niemira, R. B. Bobba, P. Sauer, and W. H. Sanders. 2013. Malicious data detection in state estimation leveraging system losses estimation of perturbed parameters. In *2013 IEEE International Conference on Smart Grid Communications (SmartGridComm’13)*. IEEE, Vancouver, BC, Canada, 402–407.

[63] Subhash Lakshminarayana and David K. Y. Yau. 2020. Cost-benefit analysis of moving-target defense in power grids. *IEEE Transactions on Power Systems* 36, 2 (2020), 1152–1163.

[64] Mohammad Ashiqur Rahman, Ehab Al-Shaer, and Rakesh B. Bobba. 2014. Moving target defense for hardening the security of the power system state estimation. In *Proceedings of the 1st ACM Workshop on Moving Target Defense*. ACM, Scottsdale, AZ, 59–68.

[65] C. Liu, J. Wu, C. Long, and D. Kundur. 2018. Reactance perturbation for detecting and identifying FDI attacks in power system state estimation. *IEEE Journal of Selected Topics in Signal Processing* 12, 4 (2018), 763–776.

[66] C. Liu, H. Liang, T. Chen, J. Wu, and C. Long. 2020. Joint admittance perturbation and meter protection for mitigating stealthy FDI attacks against power system state estimation. *IEEE Transactions on Power Systems* 35, 2 (2020), 1468–1478.

[67] J. Tian, R. Tan, X. Guan, Z. Xu, and T. Liu. 2020. Moving target defense approach to detecting Stuxnet-like attacks. *IEEE Transactions on Smart Grid* 11, 1 (2020), 291–300.

[68] J. Tian, R. Tan, X. Guan, and T. Liu. 2019. Enhanced hidden moving target defense in smart grids. *IEEE Transactions on Smart Grid* 10, 2 (2019), 2208–2223.

[69] Z. Zhang, R. Deng, D. K. Y. Yau, P. Cheng, and J. Chen. 2020. On hiddenness of moving target defense against false data injection attacks on power grid. *ACM Transactions on Cyber-Physical Systems* 4, 3 (2020), 1–29.

[70] B. Liu, H. Wu, A. Pahwa, F. Ding, E. Ibrahim, and T. Liu. 2018. Hidden moving target defense against false data injection in distribution network reconfiguration. In *2018 IEEE Power Energy Society General Meeting (PESGM’18)*. IEEE, Portland, OR, 1–5.

[71] Oliver Kosut, Liyan Jia, Robert J. Thomas, and Lang Tong. 2010. Malicious data attacks on smart grid state estimation: Attack strategies and countermeasures. In *2010 1st IEEE International Conference on Smart Grid Communications*. IEEE, Gaithersburg, MD, 220–225.

[72] Bo Tang, Jun Yan, Steven Kay, and Haibo He. 2016. Detection of false data injection attacks in smart grid under colored Gaussian noise. In *2016 IEEE Conference on Communications and Network Security (CNS’16)*. IEEE, Philadelphia, PA, 172–179.

[73] Sindhuja Mangalvedekar, Prashant Bansode, Faruk Kazi, and Navdeep Singh. 2017. A Bayesian game-theoretic defense strategy for false data injection attacks in smart grid. In *2017 14th IEEE India Council International Conference (INDICON’17)*. IEEE, Roorkee, India, 1–6.

[74] X. Liu, Y. Guan, and S. W. Kim. 2018. Bayesian test for detecting false data injection in wireless relay networks. *IEEE Communications Letters* 22, 2 (2018), 380–383.

[75] A. Gaber, K. G. Seddik, and A. Y. Elezabi. 2015. Joint estimation-detection of cyber attacks in smart grids: Bayesian and non-Bayesian formulations. In *2015 IEEE Wireless Communications and Networking Conference (WCNC’15)*. IEEE, New Orleans, LA, 2245–2250.

[76] M. N. Kurt, Y. Yilmaz, and X. Wang. 2018. Distributed quickest detection of cyber-attacks in smart grid. *IEEE Transactions on Information Forensics and Security* 13, 8 (2018), 2015–2030.

[77] C. Murguia and J. Ruths. 2016. CUSUM and Chi-squared attack detection of compromised sensors. In *2016 IEEE Conference on Control Applications (CCA’16)*. IEEE, Buenos Aires, Argentina, 474–480.
[78] Y. Huang, J. Tang, Y. Cheng, H. Li, K. A. Campbell, and Z. Han. 2016. Real-time detection of false data injection in smart grid networks: An adaptive CUSUM method and analysis. *IEEE Systems Journal* 10, 2 (2016), 532–543.

[79] Samrat Nath, Israel Akingenyeye, Jingxian Wu, and Zhihan Han. 2019. Quickest detection of false data injection attacks in smart grid with dynamic models. *IEEE Journal of Emerging and Selected Topics in Power Electronics* 10, 1 (2019), 1292–1302.

[80] S. Li, Y. Yilmaz, and X. Wang. 2015. Quickest detection of false data injection attack in wide-area smart grids. *IEEE Transactions on Smart Grid* 6, 6 (2015), 2725–2735.

[81] S. Li, X. Li, X. Wang, and J. Liu. 2017. Decentralized sequential composite hypothesis test based on one-bit communication. *IEEE Transactions on Information Theory* 63, 6 (2017), 3405–3424.

[82] I. Akingenyeye and J. Wu. 2018. Low latency detection of sparse false data injections in smart grids. *IEEE Access* 6 (2018), 58564–58573.

[83] G. Chaojun, P. Jiruttijaroen, and M. Motani. 2015. Detecting false data injection attacks in AC state estimation. *IEEE Transactions on Smart Grid* 6, 5 (2015), 2476–2483.

[84] S. K. Singh, K. Khanna, R. Bose, B. K. Panigrahi, and A. Joshi. 2018. Joint-transformation-based detection of false data injection attacks in smart grid. *IEEE Transactions on Industrial Informatics* 14, 1 (2018), 89–97.

[85] H. Manyun, N. Ming, L. Manli, W. Zhinong, S. Guoqiang, Z. Haixiang, and L. Zhongxi. 2018. Detecting false data injection attacks on modern power systems based on Jensen-Shannon distance. In 2018 IEEE 8th Annual International Conference on CYBER Technology in Automation, Control, and Intelligent Systems (CYBER’18). IEEE, Tianjin, China, 1154–1159.

[86] Lanchao Liu, Mohammad Esmalifalak, and Zhu Han. 2013. Detection of false data injection in power grid exploiting low rank and sparsity. In 2013 IEEE International Conference on Communications (ICC’13). IEEE, Budapest, Hungary, 4461–4465.

[87] B. Li, T. Ding, C. Huang, J. Zhao, Y. Yang, and Y. Chen. 2019. Detecting false data injection attacks against power system state estimation with fast go-decomposition approach. *IEEE Transactions on Industrial Informatics* 15, 5 (2019), 2892–2904.

[88] M. Ozay, I. Esnaola, F. T. Yarman Vural, S. R. Kulkarni, and H. V. Poor. 2016. Machine learning methods for attack detection in the smart grid. *IEEE Transactions on Neural Networks and Learning Systems* 27, 8 (2016), 1773–1786.

[89] Mohammad Esmalifalak, Lanchao Liu, Nam Nguyen, Rong Zheng, and Zhu Han. 2014. Detecting stealthy false data injection using machine learning in smart grid. *IEEE Systems Journal* 11, 3 (2014), 1644–1652.

[90] Z. Zhang, Y. Wang, and L. Xie. 2018. A novel data integrity attack detection algorithm based on improved grey relational analysis. *IEEE Access* 6 (2018), 73423–73433.

[91] S. A. Forouatan and F. R. Salmasi. 2017. Detection of false data injection attacks against state estimation in smart grids based on a mixture Gaussian distribution learning method. *IET Cyber-Physical Systems: Theory Applications* 2, 4 (2017), 161–171.

[92] Mehdi Ganjkhani, Seyedeh Narjes Fallah, Sobhan Badakhshan, Shahaboddin Shamshirband, and Kwok-wing Chau. 2019. A novel detection algorithm to identify false data injection attacks on power system state estimation. *Energies* 12, 11 (2019), 2209.

[93] D. Xue, X. Jing, and H. Liu. 2019. Detection of false data injection attacks in smart grid utilizing ELM-based OCON framework. *IEEE Access* 7 (2019), 31762–31773.

[94] J. Yan, B. Tang, and H. He. 2016. Detection of false data attacks in smart grid with supervised learning. In 2016 International Joint Conference on Neural Networks (IJCNN’16). IEEE, Vancouver, BC, Canada, 1395–1402.

[95] M. Ashrafuzzaman, Y. Chakhchoukh, A. A. Jillepalli, P. T. Tosic, D. C. de Leon, F. T. Sheldon, and B. K. Johnson. 2018. Detecting stealthy false data injection attacks in power grids using deep learning. In 2018 14th International Wireless Communications Mobile Computing Conference (IWCMC’18). IEEE, Limassol, Cyprus, 219–225.

[96] Y. He, G. J. Mendis, and J. Wei. 2017. Real-time detection of false data injection attacks in smart grid: A deep learning-based intelligent mechanism. *IEEE Transactions on Smart Grid* 8, 5 (2017), 2505–2516.

[97] J. J. Q. Yu, Y. Hou, and V. O. K. Li. 2018. Online false data injection attack detection with wavelet transform and deep neural networks. *IEEE Transactions on Industrial Informatics* 14, 7 (2018), 3271–3280.

[98] Moslem Dehghani, Abdollah Kavousi-Fard, Morteza Dabbaghjamaesh, and Omid Avatfepour. 2020. Deep learning based method for false data injection attack detection in AC smart islands. *IET Generation, Transmission & Distribution* 14, 24 (2020), 5756–5765.

[99] X. Niu, J. Li, J. Sun, and K. Tomsovic. 2019. Dynamic detection of false data injection attack in smart grid using deep learning. In 2019 IEEE Power Energy Society Innovative Smart Grid Technologies Conference (ISGT’19). IEEE, Washington, DC, 1–6.

[100] S. Wang, S. Bi, and Y. J. A. Zhang. 2020. Locational detection of the false data injection attack in a smart grid: A multilabel classification approach. *IEEE Internet of Things Journal* 7, 9 (2020), 8218–8227.
