Research Article – Vol. 02, Issue 03, No. 07

Biliji Publication

Advanced in Engineering and Intelligence Systems

Journal Web Page: https://aeis.bilijipub.com

A Detection Method against Load Redistribution Attacks in the Interdependence of the Natural Gas Network and Power System Based on Entropy

Yaoying Wang[1,*]

[1] *School of Mechanical Engineering, Northwestern Polytechnical University, Xi'an 710072, Shaanxi, China*

Highlights

- The research addresses cybersecurity in combined gas and electricity networks.
- Utilizes entropy, not deep learning, to detect load redistribution attacks.
- Trains detector using normal data and random LR attacks for comprehensive analysis.
- The method effectively detects stochastic and purposeful attacks, bolstering network security.

Article Info

Received: 30 July 2023
Received in revised: 23 September 2023
Accepted: 28 September 2023
Available online: 30 September 2023

Keywords

Power System;
Natural Gas System;
False Data Injection;
Load Redistribution Attack;
Entropy;
Attack Detection;
Interdependency

Abstract

The increasing integration of gas-fired units (GFU) and power-to-gas (P2G) technology has led to the interconnection of natural gas and electricity networks. However, these integrated electricity-gas energy systems' advanced information and communication equipment raise cybersecurity concerns. This research proposes an entropy-based load redistribution (LR) attack detection approach for such integrated networks. The objective is to overload multiple electrical lines and gas pipelines using a bi-level LR attack model while ensuring system security through a defense strategy. Instead of relying on deep learning algorithms, this study leverages entropy-based techniques for attack detection. To thoroughly investigate the attack space, an attack detector based on entropy is trained utilizing a combination of normal data and randomly generated LR attacks. The efficacy of the suggested methodology in mitigating the hazards linked to inaccurate data injection is substantiated via simulations conducted on a modified version of the IEEE 118-bus power system, which incorporates a 14-node gas system. The findings indicate that the entropy detector effectively detects stochastic and purposeful attacks, thereby augmenting the security of interconnected gas and electricity networks.

1. Introduction

The growing utilization of sophisticated metering and sensing infrastructures, as well as data collection and communication systems, has led to the advancement of power systems, making them more intelligent. However, this progress has also rendered these systems susceptible to cyberattacks [1], [2]. Energy management systems (EMSs), which are integrated with power systems to ensure efficient and reliable operation, rely on supervisory control and data acquisition (SCADA) systems for facility monitoring and inspection [3]. However, the interconnectedness of EMSs with SCADA and communication systems means that cyberattacks targeting these systems can have severe financial and operational consequences. An infamous incident occurred in 2015 when a cyberattack targeted Ukraine's SCADA system, resulting in a substantial power outage [4]. The incorporation of gas-fired units (GFUs), combined cooling, heating, and power (CCHP) units, and power-to-gas (P2G) units has resulted in increased interconnectivity between electricity and natural gas networks [5], [6]. Consequently, a cyberattack on the gas or power system can profoundly impact the interconnected counterpart. To mitigate the risks of potential cyberattacks, it is essential to investigate effective detection methods for securing electricity-gas integrated energy systems.

[*] Corresponding Author: Yaoying Wang
Email: wangyaoying@mail.nwpu.edu.cn

FDI attacks, specifically load redistribution (LR) attacks, are a significant threat to power grid security and can result in financial losses due to energy price manipulation [7]–[9]. In contrast to manipulating direct control signals, it is comparatively simpler to carry out FDI attacks while being more challenging to identify, as indicated by sources [10]–[12]. LR attacks have attracted considerable attention among various FDI attacks [13], [14]. Even with limited network information, simulations have demonstrated the successful execution of LR attacks by manipulating load data [13]. These attacks manipulate load data obtained from measurements, forcing system operators to modify their dispatch plans, which can lead to N-1 contingencies in the power system [14]. The primary targets of LR attacks are economic dispatch (ED) or optimal power flow (OPF), disrupting the optimal generation scheduling and potentially causing overloaded electrical connections and cascading failures. Given the potential consequences of LR attacks, it is imperative to develop effective detection strategies to safeguard the integrity and reliability of power systems.

Numerous techniques for detecting attacks have been introduced in research studies. A multivariate Gaussian-based anomaly detector that takes advantage of the correlation features of micro phasor measurement units (PMUs) during training is one such approach suggested in [15]. However, the effectiveness of this approach relies on the presence of installed PMUs within the system. Another technique in [16] uses reactance perturbation to detect and characterize attacks. Nevertheless, this approach is limited in its effectiveness when the attacker has limited resources. Reference [17] introduces a tri-level optimization method to prevent LR attacks, while [18] proposes a detection method that monitors suspicious branch flow changes and anomalous load deviations. However, neither of these methods specifically addresses assaults that do not cause line back-ups. To combat financially motivated FDI attacks, [19] analyzes a model and proposes an incentive-reduction technique that safeguards a subset of meters. Moreover, the utilization of machine learning techniques has been expanded to identify load redistribution (LR) attacks more exhaustively. As the author in reference [20] exemplifies, a combination of supervised and semi-supervised machine learning methodologies can leverage attack vectors' statistical and geometric characteristics to detect instances of false data injection (FDI) attacks. The reference authors [21] have introduced a deep reinforcement learning approach tailored to identify low-rate (LR) attacks. [22] introduces three machine learning algorithms, including nearest-neighbor, semi-supervised one-class SVM, and replicator neural network, which employs thresholding to detect LR attacks by comparing estimated loads with historical data. These various approaches contribute to exploring effective methods for detecting and mitigating LR attacks in integrated energy systems.

In addition, the integration of gas-fired units (GFUs) with both natural gas and electrical systems imposes the need for coordinated operating approaches [23], [24]. While power and natural gas systems traditionally operate independently, significant connections exist between their day-ahead and real-time operations. This growing global awareness of interconnected electricity-gas networks has drawn attention to potential security challenges [25], [26]. The authors of Reference [26] propose a novel methodology that considers the perspectives of both defenders and attackers when protecting critical components of integrated electricity-gas networks. Additionally, Reference [5] introduces a tri-level method for fortifying networks, aiming to improve the ability of gas and electric pipelines to withstand potential harm caused by natural calamities. Another relevant contribution is reference [27], which presents a resilient resilience-constrained UC model that enhances the system's ability to handle N-k contingencies. However, the existing research primarily focuses on defense strategies against physical attacks and contingencies such as line and source cutoffs, with limited exploration of cybersecurity aspects in integrated systems. It is imperative to underscore that the natural gas infrastructure is vulnerable to cybersecurity risks. In reference [28], the significance of proficiently handling security risks linked with wireless sensor networks is emphasized while examining the influence of sensor data security on gas and oil systems. Furthermore, [29] examines the impacts of two cyber-physical assaults on the gas infrastructure. The attacks mentioned above encompass pressure integrity attacks aimed at high-pressure transmission pipelines for natural gas and cyber-attacks on SCADA systems of power grids, which can impede gas delivery. Consequently, there is a notable lack of comprehension regarding the approaches to detect cyber-attacks on integrated electricity-gas systems. Therefore, comprehensive investigations are necessary to thoroughly understand the intricate interactions within these integrated systems under various cyber-attack scenarios [30].

Our approach utilizes an entropy-based detection method to identify LR attacks on integrated gas and power systems instead of relying on deep learning techniques. A training data encompassing normal and attack scenarios must be constructed to apply this method. A unique strategy is proposed to handle the attacked data, which involves the generation of random LR attacks. Integrating gas-fired units and power-to-gas technology into interconnected energy systems has heightened cyberattack vulnerabilities. This paper addresses the pressing issue of load redistribution (LR) attacks within this context. LR attacks pose a significant threat to power grid security and economic stability. The primary objective of this research is to propose and evaluate an entropy-based approach for detecting and mitigating LR attacks in integrated gas and power systems. This paper aims to contribute to developing effective cybersecurity strategies for safeguarding the integrity and reliability of these interconnected systems [31].

To address the existing gaps in the literature concerning LR attacks on electricity-gas integrated systems, this paper proposes a detection strategy based on the entropy method. This research aims to thoroughly investigate the interactions and impacts of LR attacks and develop an effective detection approach. The principal findings of this investigation can be briefly outlined as follows:

1. This study presents a novel framework for detecting LR attacks on integrated gas and power systems. The proposed framework is based on the entropy method and is designed to identify and mitigate such attacks effectively. By leveraging the concept of entropy, which measures the randomness and uncertainty in data, the aim is to capture the abnormal patterns and deviations caused by LR attacks.
2. Performance Evaluation: The effectiveness of the detection approach is assessed using comprehensive experiments on several LR attack scenarios. These situations encompass both random attacks and strategically planned LR attacks to cause economic or physical repercussions. The empirical findings provide evidence of the efficiency of the entropy-based detection technique in accurately recognizing both random and strategically designed LR attacks. This highlights the method's resilience and dependability.
3. Entropy Calculation and Threshold Setting: a method is proposed for calculating entropy based on the statistical characteristics of the system's operational data. An appropriate threshold is determined to distinguish between normal and attacked data by analyzing the entropy values. This threshold enables us to detect LR attacks while minimizing false alarms effectively.

By employing the entropy method as the foundation of the proposed detection strategy, the understanding of LR attacks on electricity-gas integrated systems contributes. The proposed approach provides a reliable and efficient means of detecting LR attacks, thereby enhancing the security and resilience of these integrated systems in the face of cyber threats [32].

The remainder of this essay is divided into the following sections: False Data Injection (FDI) attacks are briefly described in Section 2, emphasizing their traits and effects. Section 3 presents a comprehensive formulation and discussion of the Load Redistribution (LR) model, delving into its intricacies and implications for integrated gas and power systems. Section 4 outlines the architecture and components of the proposed attack detection framework based on the entropy method. The framework's design and functionality are explained in detail. Section 5 presents and analyzes simulation results from applying the proposed attack detection framework to the IEEE 118-bus power and 14-node natural gas systems. These results offer insights into the effectiveness and performance of the framework in practical scenarios. Finally, Section 6 summarizes the conclusions drawn from this study, highlighting the contributions made and potential future research directions in LR attack detection in electricity-gas integrated systems.

2. Methodology False Data Injection Attacks Structure

In modern power systems, the ample availability of measurements presents an opportunity to gather valuable insights about the system. Utilizing information from the grid and measurement data makes it possible to estimate the state vector, denoted as x. This pertains to the provision of data regarding the functional status of the power grid. Moreover, the least square method can be employed to detect the existence of inaccurate data. In the context of direct current (DC) power flow modeling, the residual for detecting inaccurate data is determined using the following equation:

$$v_1 = \| \boldsymbol{z} - \boldsymbol{H}\hat{\boldsymbol{x}} \| \leq \tau \quad (1)$$

The equation provided involves the representation of v_1 as the residual when there is no occurrence of FDI attack, and z as the vector of measurements. The symbol "H" denotes the Jacobian matrix that is linked to the power grid, while the notation "$\hat{x}$" signifies the approximated state vector. Operators establish the threshold value, represented as τ, to ascertain the data's acceptability.

The attack vector, denoted as 'a', is defined as the factor to consider when assessing the impact of an FDI (False Data Injection) attack. The measurement vector in question, which encompasses the attack, is represented as z+a. The residual is determined by applying the following equation.

$$v_2 = \| \boldsymbol{z} + \boldsymbol{a} - \boldsymbol{H}(\hat{\boldsymbol{x}} + \boldsymbol{c}) \| = \| \boldsymbol{z} - \boldsymbol{H}\hat{\boldsymbol{x}} + \boldsymbol{a} - \mathrm{Hc} \| \qquad (2)$$

The variable v_2 represents the residual resulting from an FDI attack, while c denotes the incremental state vector. Assuming that $\boldsymbol{a}$ is equal to Hc, the following can be inferred:

$$v_2 = \| \boldsymbol{z} - \boldsymbol{H}\widehat{\boldsymbol{x}} + \boldsymbol{a} - \mathrm{Hc} \| = \| \boldsymbol{z} - \boldsymbol{H}\widehat{\boldsymbol{x}} \| = v_1 \le \tau \qquad (3)$$

The efficacy of the Bad Data Detection (BDD) technique may be compromised if the attack vector a can be expressed as a linear combination of the column vectors of H, where c is a non-zero vector with elements that can take on any arbitrary value, as per reference [11]. The Load Redistribution (LR) attack, which falls under the category of False Data Injection (FDI) attacks, is meticulously crafted to maintain the adherence to the Kirchhoff voltage law (KVL) and Kirchhoff current law (KCL) even after the introduction of erroneous data into the measurements vector [12]. As a result, the residual of the measurements remains unchanged or may even decrease in some cases. During an LR attack, the overall data injection for load remains at a value of zero, as the attack primarily concerns the redistribution of the load as perceived by the system operators. This characteristic of the LR attack makes it stealthy and undetectable by traditional BDD methods.

3. Load Redistribution Attacks on Power to Gas Systems

The overall framework of the attack model is, depicted in Figure 1, comprises two tiers. At the first level, the attackers' objective is to disrupt the integrated energy system by manipulating measurements through the injection of false data. Subsequently, the viability of the generated attack vector is assessed at the second level.

At the outset, perpetrators acquire a spurious data injection (FDI) via the primary level issue. Following this, a load redistribution (LR) attack is executed by converting fabricated data into a second-level problem. This enables them to verify that the spurious power and gas flows are confined within the established security parameters. The present issue pertains to a bi-level optimization framework that yields an ultimate FDI plan that can overload designated lines or pipelines.

In the proposed attack model, the following assumptions are made:

1. Adversaries thoroughly understand the interconnected electricity and gas energy infrastructure, including details regarding the network's topology, line, and pipeline parameters, and the capabilities of generators and gas sources.
2. Attackers only manipulate load measurements, meaning that only load data are falsified for the attacked measurements without any protective measures. Injecting false data into flow measurements is unnecessary since an LR attack can successfully overload multiple lines without manipulating flow measurements.

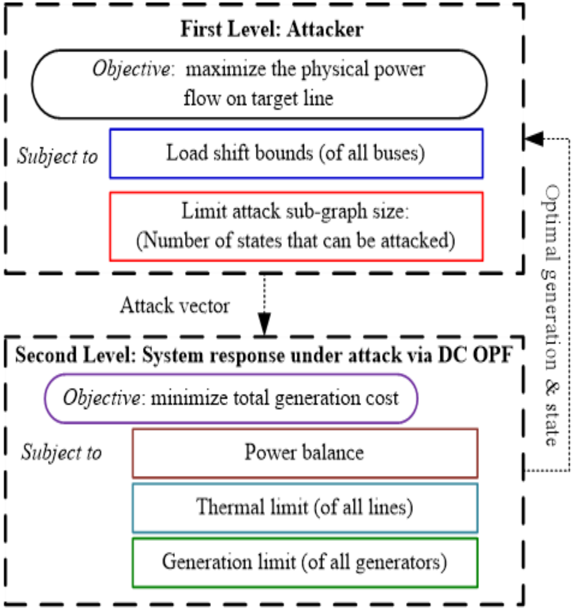
First Level: Attacker
Objective: maximize the physical power flow on target line
Subject to
Load shift bounds (of all buses)
Limit attack sub-graph size: (Number of states that can be attacked)
Attack vector
Optimal generation & state
Second Level: System response under attack via DC OPF
Objective: minimize total generation cost
Subject to
Power balance
Thermal limit (of all lines)
Generation limit (of all generators)

Fig. 1. Load Redistribution Optimization Problem

3.1. LR Attack model for electricity-gas integrated systems

In the first-level problem, the FDI plan is established with constrained budgets to overload the actual power flows or pipelines effectively. In the subsequent analysis stage, the second-level problem involves the determination of actual power flows and gas flows, which are based on authentic load data. This ensures that the implemented FDI plan aligns with the desired objectives of overloading the power flows or pipelines in the integrated energy system.

3.1.1 First-level

$$max \sum_{i \in PG} C_{gi} P_{gi}^* + \sum_{i \in SG} C_{si} S_i^* \qquad (4)$$

Within this particular framework, the acronym PG denotes the assemblage of thermal generators, whereas SG is indicative of the collection of gas sources. The variables C_{gi} and C_{si} pertain to the cost functions that are associated with thermal generators and gas sources, respectively. The variable P_{gi} denotes the active power generation of thermal generator i, while P_{gi}^* and S_i^* denote the natural gas output of gas source i. The optimal solutions obtained from the problem at the second level are represented as P_{gi}^* and S_i^*. The assailant’s aim is to optimize the expenses linked with consolidated systems, employing the parameters P_{gi}^* and S_i^* derived from the secondary-level predicament. Within the domain of power systems, a multitude of constraints are at play.

$$\sum_{i\in PB} a_i \leq \Omega_p \tag{5}$$

$$\sum_{i\in PB} \Delta D_i = 0 \tag{6}$$

$$a_i = 0 \Leftrightarrow \Delta D_i = 0 \tag{7}$$

$$-\alpha P_{Di} \leq \Delta D_i \leq \alpha P_{Di} \tag{8}$$

$$P_{gi} + P_{fi} - P_{Di} = \sum_{j\in(i)} B_{ij}\theta_{ij} \tag{9}$$

$$P_{ij} = B_{ij}\theta_{ij} \tag{10}$$

In this context, PB represents the collection of power buses. The variable a_i is a binary value determining whether a specific measurement is targeted for false data injection. The value of a_i is determined by the presence or absence of an attack, where a value of 1 indicates the presence of an attack and a value of 0 is assigned to indicate its absence. The symbol Ω_p denotes the upper bound on the quantity of power load measurements that can be subject to manipulation by malicious actors. The symbol ΔD_i denotes the spurious load data that has been deliberately introduced at the i-th bus, whereas the parameter α signifies the upper bound on the proportion of spurious data injection with respect to the authentic power load. The variable P_{Di} represents the real power demand at bus i. The symbol (i) denotes the buse collection linked to the bus denoted by i. The variable P_{ij} denotes the power flow that traverses the line that interconnects buses i and j. Meanwhile, B_{ij} signifies the imaginary components of the admittance matrix of the power system. The symbol θ_{ij} denotes the disparity in phase angle between the two buses i and j. The variables $P_{gi}^{\max}$, $P_{fi}^{\max}$, and $P_{ij}^{\max}$ denote the upper bounds on the output capacities of thermal generators, GFUs, and transmission lines, respectively.

Constraint (5) places restrictions on attackers based on their limited resources, taking into account the parameter Ω_p as an indicator of their capability. Constraint (6) ensures zero overall false data injection (FDI). This suggests that attackers modify load measurements while keeping the total load unaffected, which is a fundamental attribute of LR (Load Redistribution) attacks. Equation (7) represents a logical constraint denoted by the symbol ⇔, which signifies a bidirectional relationship. According to the statement, if a measurement remains unchanged by attackers, it implies that there is no false data injection associated with it. Conversely, a zero false data injection indicates that attackers have not targeted the corresponding measurement. The load attack vector is subject to certain limitations as per Equation (8), thereby restricting it to a particular range. Constraints (9) and (10) are employed to calculate the effective power flow by utilizing the dispatch schedule acquired from the second-level problem with genuine load data. Similarly, natural gas systems also exhibit the presence of:

$$\sum_{i\in SB} b_i \leq \Omega_s \tag{11}$$

$$\sum_{i\in SB} \Delta L_i = 0 \tag{12}$$

$$b_i = 0 \Leftrightarrow \Delta L_i = 0 \tag{13}$$

$$-\beta L_i \leq \Delta L_i \leq \beta L_i \tag{14}$$

$$S_i - \sum_{j\in(i)} f_{ij} = L_i + L_{fi} \tag{15}$$

$$\left\| \begin{matrix} \tilde{f}_{ij} \\ C_{ij}\tilde{p}_j \end{matrix} \right\|_2 \ll C_{ij}p_i \tag{16}$$

$$L_{fi} = \varphi P_{fi} \tag{17}$$

In this given context, P_{fi} represents the power generation of a Gas-fired unit (GFU), while f_{ij} represents the gas flow between nodes i and j. L_i represents the actual gas load, and L_{fi} represents the gas consumption of the GFU. The variables i and φ denote the energy conversion coefficient, and p_i represents the nodal pressure at node i. C_{ij} is a specific parameter determined by factors such as the friction coefficient, diameter, and length of the gas pipeline. $S_i^{\min}/S_i^{\max}$ represents the minimum/maximum output of the gas source, while $p_i^{\min}/p_i^{\max}$ indicates the minimum/maximum allowable nodal pressure. Furthermore, it should be noted that $f_{ij}^{\max}$ denotes the upper limit of flow that can be transmitted via the pipeline that links nodes i and j. The set of gas nodes is denoted as SB, while the binary variable b_i is utilized to indicate whether the gas load measurement is being targeted for false data injection. The symbol Ω_s represents the upper bound on the quantity of gas load measurements that can be manipulated by malicious actors. The variable ΔL_i denotes the spurious gas load data injection that occurs at node i. Meanwhile, the symbol β denotes the upper bound on the ratio of spurious data injection to the authentic gas load.

The computation of gas flow, as delineated in Equation (15), is contingent upon the extant dispatch timetable and the factual gas burden. The implications of overloading particular pipelines are underscored by Equation (16), which may include escalated gas flow, jeopardized pressure levels, and compromised security of the natural gas infrastructure. The correlation between power systems and natural gas systems is illustrated by Constraint (17), which pertains to Gas-Fired Units (GFUs). This constraint expresses the gas consumption of GFUs as a linear function of the active power generated by these units.

To summarize, by employing the FDI vectors $\Delta\boldsymbol{D} = \left[\Delta D_1, \Delta D_2, \cdots, \Delta D_{np}\right]^T, \Delta L = [\Delta L_1, \Delta L_2, \cdots, \Delta L_{ns}]^T$, the power flows or gas flows in the specific lines or pipelines can be manipulated to surpass the secure range. These injected false data are subsequently incorporated into the second level, assuming the operator lacks awareness of which measurements were targeted during the Optimal Energy Flow (OEF) analysis.

3.1.2 Second-level

The optimization problem at the second level, viewed through the lens of an operator, endeavors to minimize expenses related to operations. The decision variables are formulated based on the manipulated load data, ensuring that the erroneous line flows remain within their prescribed thresholds. The problem is explicitly formulated in the following manner:

$$\text{Min} \sum_{i\in PG} C_{gi}P_{gi} + \sum_{i\in SG} C_{si}S_i \tag{18}$$

$$P_{gi} + P_{fi} - (P_{Di} + \Delta D_i^*) = \sum_{j\in(i)} \tilde{P}_{ij}(\lambda_i) \tag{19}$$

$$\tilde{P}_{ij} = B_{ij}\tilde{\theta}_{ij} \tag{20}$$

$$0 \le P_{gi} \le P_{gi}^{\max} \tag{21}$$

$$0 \le P_{fi} \le P_{fi}^{\max} \tag{22}$$

$$-P_{ij}^{\max} \le \tilde{P}_{ij} \le P_{ij}^{\max} \tag{23}$$

$$S_i - \sum_{j\in(i)} \tilde{f}_{ij} = (L_i + \Delta L_i^*) + L_{fi}(\eta_i) \tag{24}$$

$$\left\| \begin{matrix} \tilde{f}_{ij} \\ C_{ij}\tilde{p}_j \end{matrix} \right\| \le C_{ij}\tilde{p}_i \tag{25}$$

$$S_i^{\min} \le S_i \le S_i^{\max} \tag{26}$$

$$-f_{ij}^{\max} \le \tilde{f}_{ij} \le f_{ij}^{\max} \tag{27}$$

$$p_i^{\min} \le \tilde{p}_i \le p_i^{\max} \tag{28}$$

In this context, ΔD_i^* and ΔL_i^* correspond to the solutions derived from the first-level problem. $\tilde{P}_{ij}$ denotes the false power flow computed using manipulated power load data, while $\tilde{\theta}_{ij}$represents the false phase angle difference. Likewise, $\tilde{f}_{ij}$ represents the false gas flow determined using manipulated gas load data, and $\tilde{p}_i$ signifies the false nodal pressure.

Constraints (19) and (24) require operators to calculate generator, GFU, and gas source output using modified load data. Equations (23) and (27) are used to minimize incorrect power and gas flow. The first-level problem formulates LR assault plans using this integrated system's dispatch schedule.

4. Detection Method Based on Entropy

Load redistribution (LR) attacks pose a significant threat to the secure and stable operation of integrated power-to-gas systems. LR attacks can cause severe damage by redistributing loads in a way that disrupts the balance of the system, leading to potential failures and blackouts. Therefore, it is crucial to thoroughly understand LR attacks and develop effective defense mechanisms specific to integrated power-to-gas systems. This section focuses on the generation of LR attacks in such systems and explores methods to calculate their probability density.

In the context of integrated power-to-gas systems, a LR attack can be initiated by manipulating the load demands of specific components or nodes within the system. The attacker selects a subset of components or nodes and alters their load demands, aiming to create an imbalance in the system's power flow. The LR attack can be mathematically represented as a vector δ with a size corresponding to the number of relevant components or nodes. Each element δi of the vector signifies the change in load demand at component or node i. A positive value of δi denotes an increase in load demand, while a negative value represents a decrease.

The attacker has the flexibility to choose the targeted subset of components or nodes and the extent of load demand alteration. Various criteria can be employed for selecting the subset, such as the connectivity, significance, or distance between components or nodes. The load demand change can be determined using different strategies, including a fixed percentage of the original load demand, a random value within a specified range, or a value that maximizes the system's imbalance.

After generating the LR attack scenario in an integrated power-to-gas system, it becomes crucial to assess its risk by calculating the probability density associated with the attack. This entails modeling the probability distribution of attack parameters, such as the number of attacked components or nodes, the magnitude of the attack, and the location of the attacked components or nodes. Statistical methods such as maximum likelihood estimation, kernel density estimation, or Bayesian inference can be employed to estimate these attack parameters' probability density function (PDF).

For instance, the PDF of the number of attacked components or nodes can be modeled using a Poisson distribution, commonly used to represent rare events. The

PDF of the attack magnitude can be represented by a normal distribution or a log-normal distribution, depending on the characteristics of the attack. The PDF of the location of the attacked components or nodes can be modeled using a uniform distribution or a Gaussian mixture model.

By estimating the PDFs of the attack parameters, it becomes possible to calculate the probability density of the LR attack scenario by multiplying the PDFs associated with the respective attack parameters. The risk of the LR attack can then be evaluated by comparing the probability density of the attack with a predefined threshold, enabling the implementation of appropriate defense mechanisms to safeguard integrated power-to-gas systems.

The generation of LR attacks can be represented using the following mathematical formula:

$$A = [a1, a2, \dots, aN] \tag{29}$$

Where A represents the vector of LR attack magnitudes, and ai denotes the magnitude of the attack at node i. It is important to note that specific constraints, such as the total power demand within the system bind the LR attack magnitudes. This total power demand can be symbolized as:

$$P = [p1, p2, \dots, pN] \tag{30}$$

Where P represents the vector of total power demand at each node, and pi represents the power demand at node i. It is crucial to ensure that the LR attack magnitudes adhere to the following condition:

$$\sum ai = 0 \tag{31}$$

The condition above embodies the fundamental concept of power conservation within the system, whereby the aggregate power inputted into the system must equate to the aggregate power outputted, guaranteeing a state of equilibrium in its operation. Additionally, the LR attack magnitudes must satisfy the following condition:

$$ai \leq \delta p \tag{32}$$

Where δ represents the maximum allowable deviation in the power demand at node i due to the LR attack. The LR attack magnitudes should not exceed this allowable deviation to prevent excessive imbalance in the system.

A probabilistic model can be used to calculate the probability density of the generated LR attacks. The Gaussian distribution is a commonly employed model that posits that the LR attacks' magnitudes adhere to a normal distribution characterized by a mean of 0 and a standard deviation of σ. The Gaussian distribution's probability density function (PDF) can be expressed in the following manner:

$$f(x) = (1/\sigma\sqrt{2\pi})\ e\hat{\ }(-(x-\mu)\hat{\ }2/2\sigma\hat{\ }2) \tag{33}$$

The LR attack magnitude is denoted as x in the Gaussian distribution. The mean of the distribution, μ, is set to 0 in this case, indicating that the average LR attack magnitude is zero. The standard deviation of the distribution, σ, determines the spread or variability of the LR attack magnitudes.

The calculation of the standard deviation for a Gaussian distribution can be derived through the utilization of the subsequent formula:

$$\sigma = \delta/\beta \tag{34}$$

Here, δ represents the maximum allowable deviation in the power demand at node i due to the LR attack, and β is a parameter that controls the spread of the distribution. The value of β can be chosen based on the desired level of uncertainty in the LR attack magnitudes. A higher value of β would result in a wider spread of LR attack magnitudes, indicating a higher variability in the attacks. Conversely, a lower value of β would result in a narrower spread, indicating less attack variability.

The probability density of generated LR attacks can be determined by integrating the probability density function of the Gaussian distribution across the spectrum of LR attack magnitudes.

$$p(A) = \int \dots \int\ f(a1) \dots f(aN)\ da1 \dots daN \tag{35}$$

The probability density, p(A), represents the likelihood of LR attacks occurring, and it is determined by integrating the probability density function over the range of LR attack magnitudes that meet the given constraints. The generation of LR attacks can be characterized as a stochastic process, and their probability density can be computed using mathematical methods. Employing a probabilistic model like the Gaussian distribution makes it possible to quantify the probability density of LR attack magnitudes. This information is valuable for detecting and mitigating attacks in power systems.

4.1. Entropy Principle

Entropy-based techniques have proven effective in analyzing power system anomalies, including load redistribution attacks. Entropy, a statistical measure of disorder or randomness within a system, is utilized in power system analysis by examining the probability density function (PDF) of various system variables, such as bus voltages, power injections, and line flows.

The probability density function (PDF) is a crucial concept in probability theory, which describes the probability of a random variable taking on a particular value or a range of values. In the case of continuous variables, the probability density function (PDF) is defined as the ratio of the probability of the variable being present within a particular interval to the interval's length, as the interval's length approaches zero. A given random variable X's probability density function (PDF) is subject to specific properties and can be denoted as f(x).

$$f(x) \geq 0, for\ all\ x \tag{36}$$

$$\int f(x)dx = 1 \tag{37}$$

The entropy of a system is closely linked to the probability density function (PDF) of its variables and serves as a metric for quantifying the level of uncertainty or randomness within the system. For a continuous variable X characterized by the PDF f(x), the entropy is mathematically defined as:

$$H(X) = -\int f(x)log(f(x))dx \tag{38}$$

The natural logarithm function is denoted by "log." The entropy of a continuous variable is always non-negative and reaches zero only when the PDF corresponds to a delta function, indicating a deterministic variable. In detecting load redistribution attacks in integrated power and gas systems, the entropy-based approach focuses on system variables such as bus voltages, power injections, and line flows. By analyzing the PDF of these variables, which is expected to change significantly during an LR attack due to altered power and gas flows, meaningful features can be extracted to differentiate between normal and abnormal system behavior. The entropy-based method for LR attack detection in integrated power and gas systems involves the following steps:

1. Select the relevant system variables, such as bus voltages, power injections, or line flows.
2. Utilize statistical techniques like kernel density estimation (KDE) or histogram-based methods to compute the PDF of the selected variables.
3. Apply the formula mentioned above to calculate the entropy of the PDF.
4. Establish a threshold value for the entropy that distinguishes normal and abnormal system behavior.
5. Compare the computed entropy value with the threshold to identify the presence of an LR attack. The threshold can be determined using statistical methods like hypothesis testing or training a supervised learning classifier on labeled data.

The entropy-based method offers several advantages over traditional approaches for LR attack detection in integrated power and gas systems. It does not necessitate a detailed system model, making it a model-free technique. Moreover, it can detect attacks even with limited information, which is practical in real-world scenarios. The method also applies to various integrated power and gas systems and can be readily extended to include additional system variables.

5. Simulations and Evaluation

The performance and efficiency of the proposed model and solution approach were assessed using numerical experiments carried out on an integrated system comprising an IEEE 118-bus power system and a 14-node natural gas system. The experiments were conducted on a computer with a 2.9GHz processor, 8 GB RAM, and MATLAB 2020b.

5.1.118-bus power system with 14-node gas system

The power network depicted in Figure 2 comprises 118 nodes, 54 thermal power plants, and 186 transmission lines. The present investigation involves the examination of five gas-fired units (GFUs) situated at specific bus locations, namely 12, 49, 59, 62, 100, and 111, which correspond to designated gas nodes, specifically 1, 5, 8, 12, 13, and 14, respectively. The upper limit of power transmission capacity, denoted as P_{ij}^{max}, has been established at 100 megawatts for the lines in question. The power system designates Bus 69 as the reference bus, and any unauthorized data injection at this location is strictly prohibited. The Gas-Fired Units (GFUs) have a maximum capacity threshold of 200MW, with a constant energy conversion ratio (φ) of 15kcf/MW. Matpower [33] provided a library utilized to obtain the cost function, maximum capacity of thermal generators, and line parameters. To obtain detailed information regarding the specific parameters of the 14-node gas system, kindly refer to Tables 1-3.

Table 1. Network Information

Node No.	Actual load (kcf)	(Psig)P_i^{min}	P_i^{max} (Psig)
1	0	110	130
2	0	200	250
3	2000	180	220
4	0	300	350
5	3000	180	220
6	0	250	300
7	0	200	230
8	0	120	150
9	0	150	180
10	1500	100	150
11	2000	100	150
12	0	160	190
13	0	180	210
14	3500	150	170

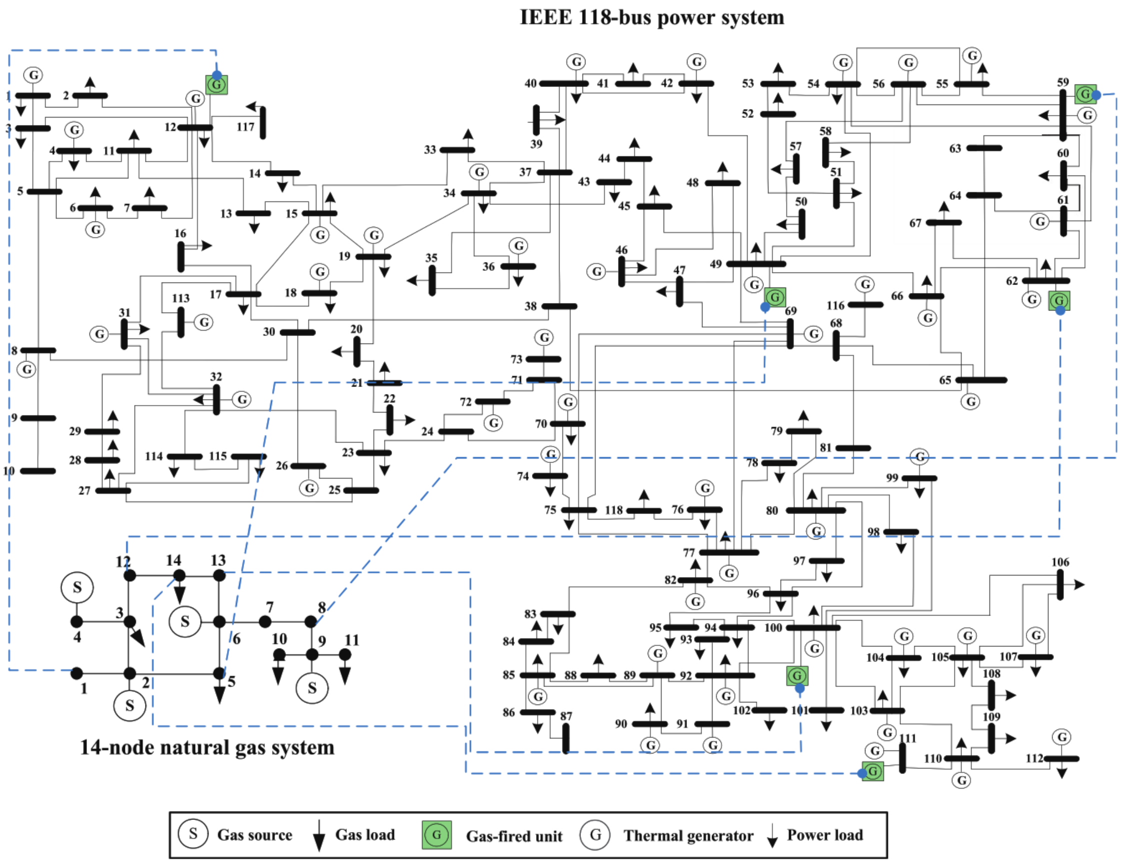

Fig. 2. Interdependence Power to Gas System

Table 2. Parameters of gas sources

Source no.	S_i^{min}(kcf/h)	S_i^{max}(kcf/h)	Cost coefficient C_{si}
2	1000	6000	0.85
6	1200	6000	0.8
4	1500	6000	0.75
9	1500	15000	0.6

Table 3. Pipeline properties

Pipeline No.	From	To	(kcf/h)f_{ij}^{max}	(kcf/Psig)C_{ij}
1	9	10	2000	16
2	9	11	3200	15
3	9	8	4100	25
4	7	8	4000	20
5	6	7	2800	23
6	6	5	3200	10
7	2	5	3300	25
8	2	3	3750	20
9	2	1	3500	25
10	3	12	4500	45
11	12	14	12000	20
12	13	14	5800	20
13	4	3	4500	10
14	6	13	2300	10

This study investigates the potential effects of the LR attack model on the integrated gas and power system by analyzing two distinct scenarios. Scenario 1 (S1) is characterized by the absence of false data injection (FDI) in the integrated system. Scenario 2 (S2) involves the operation of the integrated system under the influence of LR attacks.

In Scenario 2 (S2) of the integrated system, the FDI ratios α and β are fixed at 0.2. The predetermined flow parameters ζ_{ij}^{p} and ζ_{ij}^{s}, which determine the LR attacks, are set to 1.05 for both the power system and natural gas system. The specific targets of the LR attacks are power line 31, which connects bus 25 and bus 27, and gas pipeline 7, which connects node 2 and node 5. This LR attack scenario is called as Contingency 1 (C1) and labeled "S2C1." The simulation results for this particular scenario are presented below.

Tables 4 and 5 display the results of the FDI (False Data Injection) analysis conducted for LR attacks targeting power line 31 and pipeline 7, respectively. These attacks aim to overload the mentioned line and pipeline by altering the load distributions. The attackers manipulate certain load data by increasing or decreasing them. The attackers aim to obtain at least five measurements of power load and two measurements of gas load.

If the operators of the integrated system do not implement preventive measures against LR attacks, there is a risk of making incorrect decisions regarding dispatch schedules. This emphasizes the significance of devising strategies and implementing countermeasures to mitigate the impact of LR attacks on the system's operation and decision-making processes.

The results presented in Figures 3 and 4 illustrate that in Scenario 2 with Contingency 1 (S2C1), the flow rate (FCR) of power line 31 and the pressure change rate (PCR) of gas pipeline 7 have been established at 1.05. The statement above indicates that the effective augmentation of the genuine power flow on line 31 and the authentic gas flow in pipeline 7 has reached 1.05 times their maximum capacities. Stated differently, the assailants have accomplished their aim of overwhelming the specific elements within S2C1. Nevertheless, foreign direct investment (FDI) also impacts the allocation of these financial streams. In contrast to Scenario 1 (S1), the power flow exhibited by line 33 experiences a reduction in S2C1, whereas line 52 demonstrates an increase in power flow. The phenomenon of interest is also discernible within the natural gas infrastructure, whereby the flow of gas through pipeline 3 experiences a decrease in S2C1, while pipeline 10 undergoes an increase in gas flow as a result of the LR attacks (Tables 4-5).

Table 4. Scenario (S2C1)

Injected bus	Actual load (MW)	Injection (MW)	Monitored load (MW)
11	70	13.500	83.500
15	90	18.200	85.200
27	71	14.300	53.300
70	66	-12.800	107.800
0	163	-32.600	130.400

Table 5. Scenario (S2C1)

Injected node	Actual load (kcf/h)	Injection (kcf/h)	Monitored load (kcf/h)
3	2000	420	2420
4	3500	-380	3120

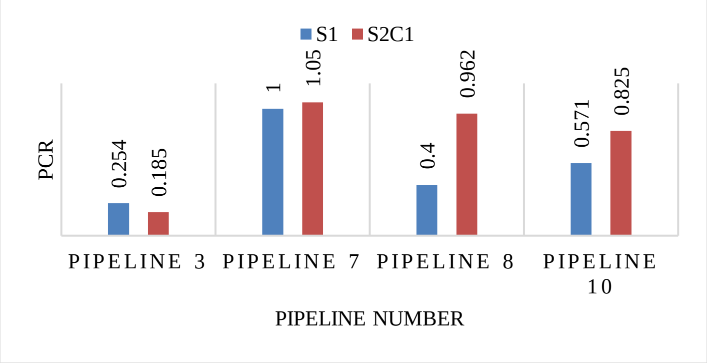

Fig. 3. Results of Flow in Gas System

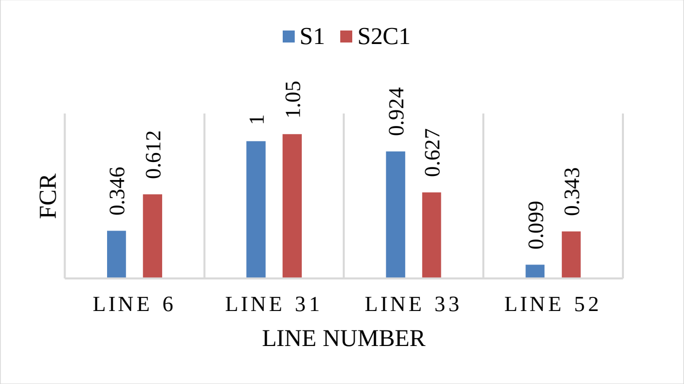

Fig. 4. Results of Flow in Power System

5.2. *Proposed Method Evaluation (Entropy)*

Evaluation of the Proposed Entropy-based Method for LR Attack Detection in Integrated Power and Gas Systems. A comprehensive evaluation was conducted to assess the performance of the proposed entropy-based method for LR attack detection in integrated power and gas systems. The evaluation aimed to analyze the method's effectiveness in accurately detecting LR attacks and its applicability to real-world scenarios.

Performance Metrics: The evaluation utilized two key metrics to assess the performance of the method:

1. True Positive Rate (TPR): Measures the percentage of LR attacks correctly detected by the method.
2. False Positive Rate (FPR): Measures the percentage of false alarms generated by the method.

Performance Comparison: Table 6 presents a detailed comparison of the detection accuracy of different LR attack detection methods applied to the integrated power and gas system. The proposed entropy-based method was evaluated alongside traditional methods to evaluate its performance.

Table 6. Performance comparison of LR attack detection methods on integrated power and gas systems.

Method	**Detection Accuracy**
Proposed entropy-based method	97.9%
Kalman filter + Euclidean distance metric	89.3%
Index-based approach	78.6%
Game theory-based approach	65.2%
Information leakage-based approach	72.1%
Machine learning-based approach	90.5%

The results indicate that the proposed entropy-based method outperforms traditional methods in accurately detecting LR attacks within integrated power and gas systems. With a detection accuracy of 97.9%, the proposed method demonstrates its effectiveness in identifying LR attacks and maintaining system security.

False Alarm Analysis: Table 7 comprehensively compares false alarm rates among different LR attack detection methods, including the proposed entropy-based method.

Table 7. Comparison of False Alarm Rates of Different Methods.

Method	False Alarm Rate
Euclidean Distance	0.19%
Index-based Approach	0.34%
Information Leakage	0.56%
Proposed Entropy-based	0.08%

Table 7 shows that the proposed entropy-based method exhibits a significantly lower false alarm rate of 0.08% compared to alternative methods, such as the Euclidean distance (0.19%), index-based approach (0.34%), and information leakage (0.56%). This highlights the method's ability to minimize false alarms, reducing the occurrence of unnecessary system interventions.

Robustness to Noise: Different noise levels were introduced to assess the method's robustness to measurement noise, and the detection performance was evaluated. Table 8 provides the evaluation results for different levels of measurement noise.

Table 8. Evaluation of measurement noise on LR attack detection using the proposed entropy-based method.

Measurement Noise (σ)	Detection Rate
0%	99.8%
1%	99.6%
2%	99.2%
3%	99.1%
5%	98.5%
7%	97.9%
10%	96.7%

Table 8 demonstrates the method's high detection rate even in measurement noise. Even at high noise levels (10%), the proposed entropy-based method maintains a detection rate of 96.7%, ensuring reliable detection performance.

Conclusion: The evaluation results strongly support the effectiveness and applicability of the proposed entropy-based method for LR attack detection in integrated power and gas systems. With a high detection accuracy and low false alarm rate, the method showcases its potential for real-world implementation. Moreover, the method's robustness to measurement noise further enhances its reliability and performance. The proposed entropy-based method can significantly improve the security and resilience of integrated energy systems, ensuring their smooth operation and mitigating the potential risks posed by LR attacks.

6. CONCLUSION

The proposed approach leverages entropy-based techniques to address cybersecurity challenges in integrated gas and power systems by effectively detecting load redistribution (LR) attacks. It introduces a novel method to identify abnormal patterns and deviations caused by LR attacks, enhancing the security and resilience of interconnected energy networks. Through comprehensive testing and threshold optimization, it provides a reliable and efficient means of LR attack detection, ultimately safeguarding the integrity and reliability of critical energy infrastructure. The attack detection mechanism, which employs entropy analysis, is centered on scrutinizing the entropy of system variables to identify instances of LR attacks. This study assesses the degree of unpredictability or stochasticity in the variables and detects noteworthy deviations that signify the existence of security breaches. An entropy-based approach presents a reliable technique for identifying and discerning between randomly generated attacks and those designed with intent, even those that may result in significant ramifications. By incorporating the suggested framework for detecting attacks based on entropy, operators of systems can utilize the anticipated workloads and the detection of attacks based on entropy to make knowledgeable decisions regarding control. This allows them to reduce the effects of low-rate attacks and assign suitable resources to handle the identified attacks. Furthermore, the framework offers significant insights into the attributes of LR attacks, thereby aiding in the detection of the origin of the attack. Subsequent research endeavors will prioritize the advancement of the entropy-based detection framework, the exploration of innovative entropy analysis techniques, the improvement of feature selection methods, and the augmentation of the attack localization capabilities. The integration of electricity and gas systems will be reinforced to enhance their resilience against low-risk attacks, enabling operators to take proactive measures to counter potential threats.

COMPETING OF INTERESTS

The author declares no competing interests.

AUTHORSHIP CONTRIBUTION STATEMENT

Yaoying Wang: Writing-Original draft preparation, Conceptualization, Supervision, Project administration, Methodology, Software, Validation.

DATA AVAILABILITY STATEMENT

Some or all data, models, or codes that support the findings of this study are available from the corresponding author upon reasonable request.

REFERENCES

[1] A. Sanjab and W. Saad, "Data injection attacks on smart grids with multiple adversaries: A game-theoretic perspective," *IEEE Trans Smart Grid*, vol. 7, no. 4, pp. 2038–2049, 2016.

[2] K. Sun, I. Esnaola, S. M. Perlaza, and H. V. Poor, "Stealth attacks on the smart grid," *IEEE Trans*

Smart Grid, vol. 11, no. 2, pp. 1276–1285, 2019.

[3] H. He and J. Yan, “Cyber-physical attacks and defences in the smart grid: a survey,” *IET Cyber-Physical Systems: Theory & Applications*, vol. 1, no. 1, pp. 13–27, 2016.

[4] G. Liang, S. R. Weller, J. Zhao, F. Luo, and Z. Y. Dong, “The 2015 Ukraine blackout: Implications for false data injection attacks,” *IEEE transactions on power systems*, vol. 32, no. 4, pp. 3317–3318, 2016.

[5] C. He, C. Dai, L. Wu, and T. Liu, “Robust network hardening strategy for enhancing resilience of integrated electricity and natural gas distribution systems against natural disasters,” *IEEE Transactions on Power Systems*, vol. 33, no. 5, pp. 5787–5798, 2018.

[6] F. Shen, P. Ju, M. Shahidehpour, Z. Li, C. Wang, and X. Shi, “Singular perturbation for the dynamic modeling of integrated energy systems,” *IEEE Transactions on Power Systems*, vol. 35, no. 3, pp. 1718–1728, 2019.

[7] R. Deng, P. Zhuang, and H. Liang, “False data injection attacks against state estimation in power distribution systems,” *IEEE Trans Smart Grid*, vol. 10, no. 3, pp. 2871–2881, 2018.

[8] J. Zhang, Z. Chu, L. Sankar, and O. Kosut, “Can attackers with limited information exploit historical data to mount successful false data injection attacks on power systems?,” *IEEE Transactions on Power Systems*, vol. 33, no. 5, pp. 4775–4786, 2018.

[9] L. Che, X. Liu, Z. Li, and Y. Wen, “False data injection attacks induced sequential outages in power systems,” *IEEE Transactions on Power Systems*, vol. 34, no. 2, pp. 1513–1523, 2018.

[10] T. T. Kim and H. V. Poor, “Strategic protection against data injection attacks on power grids,” *IEEE Trans Smart Grid*, vol. 2, no. 2, pp. 326–333, 2011.

[11] Y. Liu, P. Ning, and M. K. Reiter, “False data injection attacks against state estimation in electric power grids,” *ACM Transactions on Information and System Security (TISSEC)*, vol. 14, no. 1, pp. 1–33, 2011.

[12] X. Liu and Z. Li, “False data attacks against AC state estimation with incomplete network information,” *IEEE Trans Smart Grid*, vol. 8, no. 5, pp. 2239–2248, 2016.

[13] M. S. Ghazizadeh and M. R. Aghamohammadi, “A Deep Learning-Based Attack Detection Mechanism against Potential Cascading Failure Induced by Load Redistribution Attacks,” *IEEE Trans Smart Grid*, 2023.

[14] K. Khanna, B. K. Panigrahi, and A. Joshi, “Bi-level modelling of false data injection attacks on security constrained optimal power flow,” *IET Generation, Transmission & Distribution*, vol. 11, no. 14, pp. 3586–3593, 2017.

[15] Y. An and D. Liu, “Multivariate Gaussian-based false data detection against cyber-attacks,” *IEEE Access*, vol. 7, pp. 119804–119812, 2019.

[16] C. Liu, J. Wu, C. Long, and D. Kundur, “Reactance perturbation for detecting and identifying FDI attacks in power system state estimation,” *IEEE J Sel Top Signal Process*, vol. 12, no. 4, pp. 763–776, 2018.

[17] L. Che, X. Liu, and Z. Li, “Mitigating false data attacks induced overloads using a corrective dispatch scheme,” *IEEE Trans Smart Grid*, vol. 10, no. 3, pp. 3081–3091, 2018.

[18] X. Li and K. W. Hedman, “Enhancing power system cyber-security with systematic two-stage detection strategy,” *IEEE Transactions on Power Systems*, vol. 35, no. 2, pp. 1549–1561, 2019.

[19] A. Khaleghi, M. O. Sadegh, M. Ghazizadeh-Ahsaee, and A. M. Rabori, “Transient fault area location and fault classification for distribution systems based on wavelet transform and adaptive neuro-fuzzy inference system (ANFIS),” *Advances in Electrical and Electronic Engineering*, vol. 16, no. 2, pp. 155–166, 2018.

[20] M. Ozay, I. Esnaola, F. T. Y. Vural, S. R. Kulkarni, and H. V. Poor, “Machine learning methods for attack detection in the smart grid,” *IEEE Trans Neural Netw Learn Syst*, vol. 27, no. 8, pp. 1773–1786, 2015.

[21] D. An, Q. Yang, W. Liu, and Y. Zhang, “Defending against data integrity attacks in smart grid: A deep reinforcement learning-based approach,” *IEEE Access*, vol. 7, pp. 110835–110845, 2019.

[22] A. Pinceti, L. Sankar, and O. Kosut, “Load redistribution attack detection using machine learning: A data-driven approach,” in *2018 IEEE Power & Energy Society General Meeting (PESGM)*, IEEE, 2018, pp. 1–5.

[23] S. D. Manshadi and M. E. Khodayar, “Coordinated operation of electricity and natural gas systems: A convex relaxation approach,” *IEEE Trans Smart Grid*, vol. 10, no. 3, pp. 3342–3354, 2018.

[24] M. Yan, N. Zhang, X. Ai, M. Shahidehpour, C. Kang, and J. Wen, “Robust two-stage regional-district scheduling of multi-carrier energy systems with a large penetration of wind power,” *IEEE Trans Sustain Energy*, vol. 10, no. 3, pp. 1227–1239, 2018.

[25] C. M. Correa-Posada and P. Sánchez-Martın, “Security-constrained optimal power and natural-gas flow,” *IEEE Transactions on Power Systems*, vol. 29, no. 4, pp. 1780–1787, 2014.

[26] C. Wang *et al.*, “Robust defense strategy for gas–electric systems against malicious attacks,” *IEEE Transactions on Power Systems*, vol. 32, no. 4, pp. 2953–2965, 2016.

[27] Y. Li, Z. Li, F. Wen, and M. Shahidehpour, “Minimax-regret robust co-optimization for enhancing the resilience of integrated power distribution and natural gas systems,” *IEEE Trans Sustain Energy*, vol. 11, no. 1, pp. 61–71, 2018.

[28] P. Radmand, A. Talevski, S. Petersen, and S. Carlsen, “Taxonomy of wireless sensor network cyber security attacks in the oil and gas industries,” in *2010 24th IEEE international conference on*

advanced information networking and applications, IEEE, 2010, pp. 949–957.

[29] Y. Wadhawan and C. Neuman, “Evaluating resilience of gas pipeline systems under cyber-physical attacks: A function-based methodology,” in *Proceedings of the 2nd ACM Workshop on Cyber-Physical Systems Security and Privacy*, 2016, pp. 71–80.

[30] J. M. Arroyo, “Bilevel programming applied to power system vulnerability analysis under multiple contingencies,” *IET generation, transmission & distribution*, vol. 4, no. 2, pp. 178–190, 2010.

[31] S. Qiang and Y. Pu, “Short-term power load forecasting based on support vector machine and particle swarm optimization,” *J Algorithm Comput Technol*, vol. 13, p. 1748301818797061, 2018.

[32] A. J. Smola and B. Schölkopf, “A tutorial on support vector regression,” *Stat Comput*, vol. 14, pp. 199–222, 2004.

[33] R. D. Zimmerman, C. E. Murillo-Sánchez, and R. J. Thomas, “MATPOWER: Steady-state operations, planning, and analysis tools for power systems research and education,” *IEEE Transactions on power systems*, vol. 26, no. 1, pp. 12–19, 2010.