# Topology-Aware Hybrid FDIA Against Third-Party Aggregators in Smart Grids

Javad Mokhtari Koushyar*, Mina Guirguis*, George Atia†
*Department of Computer Science, Texas State University
†Department of Electrical and Computer Engineering, University of Central Florida

***Abstract*—Power utility companies are increasingly interfacing with equipment and services they neither own nor directly control. These are typically provided by third-party entities and now form critical components of the power grid's supply chain. Among them, aggregators collect and aggregate power usage data from prosumers, and report them for billing, and for monitoring usage and forecasting demand. However, many aggregators are small entities with limited security capabilities, introducing new vulnerabilities into the grid. In this work, we expose Hybrid False Data Injection Attacks (FDIA), whereby an adversary compromises an aggregator and subtly manipulates load reports in a topology-aware fashion to trigger blackouts or power losses. We develop a time-slotted model that abstracts low-level power flow dynamics while capturing interactions among key components in the power grid such as prosumers, aggregators, substations, and the control unit. To assess the impact of Hybrid FDIA, we introduce new performance metrics that quantify different dimensions of inflicted damage. Using realistic prosumer distributions in a regional smart grid, we examine how aggregator dominance can be exploited to deteriorate grid stability and demonstrate how safety margins and energy storage create tradeoffs between resiliency and operational cost.**

***Index Terms*—False Data Injection Attacks, Smart Grids, Grid Topology, Third-Party Aggregators**

## I. Introduction

**Motivation:** Power utility companies are increasingly interfacing with a wide variety of equipment and services supplied by third-parties. These include smart meters, networking equipment, solar panels, and more. Decisions and control actions taken by power companies rely heavily on the behavior and integrity of this third-party infrastructure. In particular, a compromise of such entities poses significant security and safety risks, not only to the utility company, but to the broader power grid as well. Power companies are generally less concerned with isolated attacks–such as those manipulating a single customer's meter data for financial gain–than with coordinated, large-scale attacks involving compromised third parties. The threat becomes especially serious when sophisticated adversaries target smaller aggregators or vendors with limited security postures [5].

One class of attacks that has been extensively studied is False Data Injection Attacks (FDIA), which can compromise state estimators while evading detection [14], [11]. This has spurred a large body of work on developing defense mechanisms [17]. However, most studies have focused on specific estimation or detection techniques (e.g., $\chi^2$ detectors) and have not considered the broader *topological* impact of FDIA that remain stealthy under most conventional defenses.

The main goal of this work is to develop a *system-level* approach for analyzing and quantifying the impact of a new class of FDIA, which we term Hybrid FDIA, launched through compromised third-party aggregators. By *system-level*, we refer to analyzing the collective impact of an attack across multiple interacting components of the smart grid, such as aggregators, substations, and the control unit, rather than focusing on isolated devices or detection algorithms. Our model explicitly considers both the *footprint* of the aggregator (i.e., the number of prosumers under its control) and the *physical topology* of their connections within the power grid. We evaluate our models using a smart grid simulator that offers fine-grained control over all components of the system, from the consumer/prosumer[1] level, to transmission and distribution infrastructure, power generation, and load forecasting. Our simulator operates on a regional grid structure based on the IEEE 57-Bus system, which provides sufficient complexity for modeling realistic topologies and evaluating attack propagation.

**Scope:** Due to the presence of different types of "third-parties", we focus on the specific class of aggregators [7]. These entities are responsible for collecting and aggregating power consumption/production data from prosumers and sending reports to the upper levels for further usage (billing, future demand prediction, etc.). Thus, they correspond to providers of smart meters, solar panels, or networking components that are vulnerable to cyber attacks. Furthermore, in order to capture system-level impact of compromises, our evaluation is focused on regional realistic setups composed of thousands of prosumers and distribution lines, and a small number of transmission substations and generators. We abstract many of the complex details of the power dynamics in favor of a scalable system-level assessment.

**Contributions:** In this paper, we make the following contributions:

1) We develop an abstract time-slotted power system model that allows us to study the impact of third-party compromises on the behavior of the power grid. This model provides the appropriate level of abstraction that hides the complex details of power systems while allowing various

[1]A consumer who is also involved in power generation.

components of the power system to interact over longer time scales.
2) We identify hybrid FDIA targeting a single third-party aggregator, in which the attacker manipulates load reports in a topology-aware manner to stealthily exploit the grid's inherent load redistribution mechanisms [21].
3) To assess the impact of hybrid FDIA, we develop novel metrics that capture the deviation of the system under attack from its nominal operation, as well as metrics that expose the imbalance caused by the attacks.
4) We conduct extensive simulations using realistic topologies with various prosumer-aggregator connectivity models, demonstrating how aggregator dominance and prosumer clustering significantly amplify the impact of attacks.

**Paper organization:** This paper is organized as follows. In Section II, we provide an overview of related work. In Section III, we present our methodology, our abstract time-slotted model, the adversary models, and the performance metrics. In Section IV, we present our evaluation of various Hybrid FDIA across the power grid. Finally, we conclude and summarize the key findings in Section V.

## II. Related Work

### A. False Data Injection Attacks (FDIA) on State Estimators

FDIA is a class of attacks that target state estimation by injecting false measurement data that can bypass detection methods [14], [11]. Attackers mount FDIA for various purposes, including power theft, economic gains, power disruption, and equipment damage. The early work of Liu *et al.* [14] demonstrated how FDIA can bypass the $\chi^2$ bad data detection method, even with limited access to measurement meters. This has spurred a large body of work on improving bad data detection and optimizing FDIA strategies under physical constraints and partial knowledge of system topology (e.g., [3], [13]). For example, Che *et al.* [3] optimize FDIA to cause cascading failures by targeting specific branches in the grid. Liu *et al.* [13] propose identifying feasible attack regions using Mixed Integer Linear Programming, even without full topology knowledge.

On the defense side, many studies have explored detection and mitigation techniques for FDIA (see [17] for a survey). These include methods based on measurement protection [16], Phasor Measurement Units (PMUs) [15], and machine learning approaches [6], [8]. For example, Manandhar *et al.* [16] use a Kalman filter for state estimation, combining its output with raw measurements in a $\chi^2$ detector to identify FDIA. Mahapatra *et al.* [15] apply Principal Component Analysis (PCA) to differentiate between FDIA and normal disturbances. He *et al.* [8] use deep learning to detect FDIA in real time based on historical measurement features. The recent DAMGAT framework [19] employs a multi-head graph attention network to spatially and temporally model node features across grid topology. Unsupervised approaches such as deep latent space clustering [2] have also shown promise in detecting stealthy FDIA on AC state estimators. Cui *et al.* [6] apply machine learning to detect FDIA in load forecasting.

Despite these advances, most existing works target specific state estimation algorithms (e.g., $\chi^2$) and are validated on small-scale systems (e.g., IEEE-14). As a result, it remains difficult to assess the broader impact of FDIA on grid-wide operations. In contrast, our work abstracts away from detailed physical modeling to focus on system-level impacts. Our framework can integrate various FDIA optimizations while capturing their cascading effects across the power grid hierarchy.

### B. Broader Impact of FDIA

Beyond state estimation, several studies explore the broader impacts of FDIA. Choi *et al.* [4] analyze the economic effects on locational marginal pricing due to topological misrepresentation caused by FDIA. Xie *et al.* [22] study how attackers can financially benefit from manipulating real-time market operations through virtual bidding. Aflaki *et al.* [1] introduce a variant of FDIA — the "Civil Attack" — where historical loads are swapped across zones, affecting forecasting. Sreeram *et al.* [18] use graph-based methods to rank vulnerabilities and guide optimal meter placements. Lin *et al.* [12] examine general attack strategies that imbalance supply and demand, increase distribution costs, and disrupt power delivery, albeit in a simplified model of the US grid. Xiang *et al.* [20] analyze long-term impacts of load redistribution attacks using a semi-Markov process to model manipulated measurements.

In our prior work [10], we empirically studied the role of the aggregators' topology on the impact of inflation and deflation FDIAs. In this work, we develop a novel systematic model to assess the impact of FDIAs across multiple layers of the grid hierarchy. We expose Hybrid FDIAs and evaluate their impact using newly introduced metrics.

## III. Methodology

In this section, we first present an illustrative setup, followed by our system model, adversary model, and performance metrics.

### A. An Illustrative Setup with Third Parties

Figure 1 illustrates an example setup of prosumers physically connected to substations and to aggregators (third-party companies). In this regional setup, we have two neighborhoods, each served by a substation and three aggregators. The customers of these aggregators are distributed across the two neighborhoods. In this setup, the top aggregator provides services to more prosumers than the bottom one. If an attacker compromises the top aggregator, they can inject FDIA into more reports. In this work, we are interested to capture the impact of an aggregator compromise given their footprint in terms of the number of prosumers they control and the manner in which those prosumers are physically connected to substations.

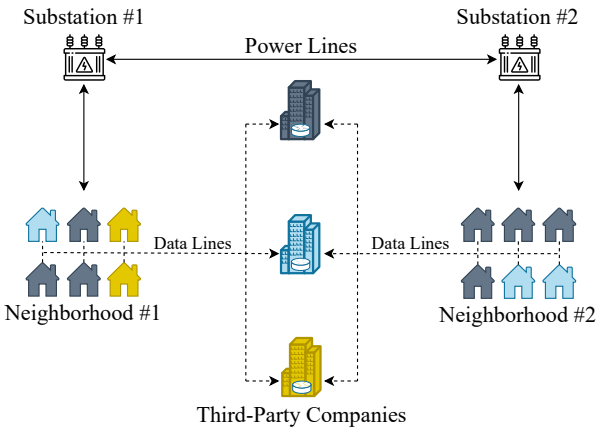



Fig. 1: Involved third parties and their customers distribution.

### B. Power System Model

We introduce a time-slotted model with $G$ generators, $S$ substations, $P$ Prosumers, $N$ Aggregators, and arbitrary transmission and distribution lines. Generators are divided into subsets connected to specific transmission substations. Aggregators report their demand directly to the *Utility Company*, which then passes this information to a central *Control Unit*. The Control Unit applies a safety margin, incorporating random noise, to predict the required power generation for future time steps.

In the following, without loss of generality, we focus on a single substation (say substation $s$). We use $P_s$ and $N_s$ to denote the number of prosumers and aggregators associated with substation $s$. The same analysis applies to prosumers connected to any other substations. Let $\mathcal{P}_j \subseteq [P_s]$ denote the set of prosumers connected to aggregator $j \in [N_s]$, where $[n] := \{1, \ldots, n\}$ and $[m, n] := \{m, m+1, \ldots, n\}$.

*1)* ***Prosumers****:* Each prosumer has a time-varying demand and may generate renewable energy (e.g., solar) in addition to receiving grid power. The surplus of prosumer $i$ at each time step is:

$$s_i(t) = r_i(t) + u_i(t) - d_i(t), \tag{1}$$

where $s_i(t)$ is the surplus power, $r_i(t)$ is renewable power, $u_i(t)$ is grid power, and $d_i(t)$ is demand. Thus, the surplus power at time $t$ is:

$$p_i(t) = (s_i(t))^+ := \max(s_i(t), 0)\,, \tag{2}$$

which represents the excess power prosumer $i$ can feed back into the grid or store.

*2)* ***Aggregator Reporting****:* Aggregators report their aggregated demand directly to the Utility Company every $T$ time slots, at times $t_k = kT$. The *actual aggregated energy demand* for aggregator $j$ over time slots $[t_k - T, \ldots, t_k - 1]$ (in kWh) is:

$$D_j(t_k) = \sum_{t=t_k-T}^{t_k-1} \sum_{i \in \mathcal{P}_j} d_i(t)\Delta t\,, \tag{3}$$

where $\Delta t$ is the duration of each time slot. The *reported aggregated demand* by aggregator $j$ for the next period is $\hat{D}_j(t_k)$, which may differ from $D_j(t_k)$ if aggregator $j$ is compromised. We introduce an *aggregator reporting error*:

$$\hat{D}_j(t_k) = D_j(t_k) + \epsilon_j(t_k), \tag{4}$$

where $\epsilon_j(t_k)$ represents the reporting error, zero for honest reporting, and non-zero for compromised reporting.

*3)* ***Control Unit and Power Generation****:* The *Control Unit* receives the aggregated demands from the Utility Company and, using an integrated Predictor, forecasts the required power generation for future time steps. The Predictor uses historical reported demands to estimate future demand. Specifically, the predicted demand for aggregator $j$ is:

$$\hat{D}_j^{\text{pred}}(t_k) = \text{Predictor}(\{\hat{D}_j(t_{k-m})\}_{m=1}^{M}), \tag{5}$$

where the Predictor function forecasts future demand based on recent reports over the past $M$ periods.

To ensure robustness, the Control Unit applies a *safety margin* $\Gamma > 1$. Therefore, the power generation required at the transmission substation $s$ is:

$$G_s(t) = \Gamma \cdot \frac{1}{T\Delta t} \sum_{j=1}^{N_s} \hat{D}_j^{\text{pred}}(t_k) \quad \text{for } t_k \le t < t_k + T, \tag{6}$$

where $G_s(t)$ is the predicted power requirement at substation $s$, recalling that $N_s$ is the number of aggregators served by the substation $s$. Hence, the total power generation required from all substations at time $t$ is:

$$G(t) = \sum_{s=1}^{S} G_s(t). \tag{7}$$

We also let $G_0(t)$ denote the total power generated if the reporting error $\epsilon_j(t_k)$ is 0 for all $j \in [N]$, meaning none of the aggregators are reporting erroneous demands in the corresponding time slot.

*4)* ***Predictor****:* The Predictor, as part of the Control Unit, is a model that forecasts demand. By adjusting the safety margins, the Predictor can approximate real-world demand fluctuations and control needs. To account for prediction inaccuracies, we generate the safety margin from a probability distribution in our experiments. For simplicity, we omitted this randomness from the current model description and represent the safety margin as $\Gamma$ in (6).

*5)* ***Power Flow Balance****:* Power balance must be maintained at each time slot taking into account the generator-substation mappings, as well as losses in transmission and distribution lines. Hence,

$$\sum_{s=1}^{S} G_s(t) + \sum_{i=1}^{P} p_i(t) = \sum_{i=1}^{P} d_i(t) + L_{\text{trans}}(t) + L_{\text{dist}}(t), \tag{8}$$

where $L_{\text{trans}}(t)$ and $L_{\text{dist}}(t)$ represent losses in transmission and distribution lines, respectively, which are affected by line load and efficiency.

One objective is to study the relationship between discrepancies in reported demand $(D_j(t_k) - \hat{D}_j^{\text{pred}}(t_k))$ and the combined losses $(L_{\text{trans}}(t) + L_{\text{dist}}(t))$.

### C. Adversary Model

Figure 2 illustrates an overview of a hybrid FDIA. We assume the attacker has compromised a single aggregator $j$ and is capable of launching a FDIA on the set of prosumers $\mathcal{P}_j$ connected to that aggregator. Additionally, the attacker is assumed to have knowledge of the physical locations of the prosumers and the substations to which they are connected. Such knowledge enables the attacker to find the optimal partitioning that maximizes the impact of the FDIA.

We identify a hybrid FDIA in which the attacker strategically inflates the reported demands of a subset of prosumers while simultaneously deflating those of others, such that the aggregated load measurement reported by the compromised aggregator remains close to its expected value. This manipulation preserves stealth by evading anomaly detection mechanisms that monitor aggregated reports, while introducing a hidden power imbalance that must be compensated for elsewhere in the grid.

**Algorithm 1** Find Multipliers for Compromised Aggregator

**Initialize:**

$\mathcal{P}_j$: compromised aggregator
$\mathcal{P}_j^+$, $\mathcal{P}_j^-$: partitions of $\mathcal{P}_j$ for inflation and deflation
$\Phi_i(t) \in [1-\eta, 1+\eta]$: multiplier for consumer $i$ at time $t$
$\mathcal{X}(t)$: maximum allowed deviation from nominal

**Optimize:**

**for** $t \in [t_{\text{start}}, t_{\text{end}}]$ **do**
Solve:

$$\begin{aligned} \min_{\{\Phi_i(t)\}} \quad & \epsilon_j(t) \\ \text{subject to} \quad & \epsilon_j(t) \le \mathcal{X}(t) \\ & 1 \le \Phi_i(t) \le 1+\eta, \quad \forall i \in \mathcal{P}_j^+ \\ & 1-\eta \le \Phi_i(t) \le 1, \quad \forall i \in \mathcal{P}_j^- \end{aligned}$$

**end for**
**return** $\{\Phi_i(t)\}$

To execute the attack, the adversary formulates a linear program shown in Algorithm 1 that determines the inflation and deflation weights across the prosumers in $\mathcal{P}_j$, ensuring that the total deviation in the aggregated measurement remains minimal. This approach enables the attacker to remain stealthy while strategically disrupting load distribution within the grid.

### D. Performance Metrics

To evaluate the impact of grid topology on Hybrid FDIA, we define the following performance metrics. These metrics are designed to efficiently quantify the effects of compromised power reports on the grid's operation and identify the potential costs associated with load redistribution and power outages.

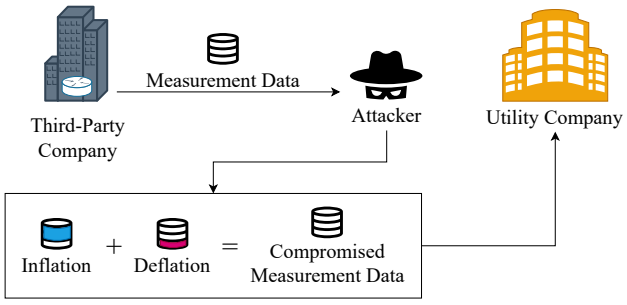



Fig. 2: Hybrid FDIA on a Third-Party Aggregator

*1) **Deviation From Nominal**:* The Deviation from Nominal (DN) metric quantifies the difference between the power generated in a compromised mode and that in a secure mode. Recalling that $G_0(t)$ and $G(t)$ represent the total power generated without compromise and the actual total power generated at time $t$, respectively, DN is defined as:

$$\text{DN} = \sum_{t=t_{\text{start}}}^{t_{\text{end}}} \frac{|G_0(t) - G(t)|}{G_0(t)} \tag{9}$$

where the summation is over the attack window. This metric is normalized by the power generated in the secure mode to ensure independence from the absolute scale of power generation.

*2) **Redistribution Cost**:* The Redistribution Cost (RC) applies where one substation receives excess power while another experiences a deficit. The load redistribution operation, which redistributes power between substations, incurs a cost that depends on the amount of power transferred, the distance between substations and the reliability of power transfer. We define the redistribution cost over the attack window by:

$$\text{RC} = \sum_{t=t_{\text{start}}}^{t_{\text{end}}} \sum_{(i,j)\in S_t} \left(\frac{\Delta G_{ij}}{\rho_{ij}}\right)^{\beta} k_{ij} \tag{10}$$

where $\Delta G_{ij}$ is the amount of power transferred from substation $s_i$ to substation $s_j$, $k_{ij}$ is the distance between $s_i$ and $s_j$, $\rho_{ij}$ is the reliability rate of the transmission line connecting them, $S_t$ is the set of substation pairs involved in redistribution at time $t$, and the exponent $\beta$ models nonlinear relationships between the power transferred and the cost. The RC metric quantifies the operational cost of load redistribution actions triggered by the hybrid attack. This cost is proportional to both the amount of power transferred and the distance between substations, and inversely proportional to the reliability of the transmission lines.

*3) **Power Outage Count**:* Power Outage Count (POC) is relevant where the manipulated reports may cause a substation to receive less power than needed, leading to power outages. An outage occurs when the power supplied to a substation is insufficient for its prosumers.

Define a power outage at substation $s$ at time $t$ as:

$$\text{Outage}_s(t) = \begin{cases} 1 & \text{if } G_s(t) + \sum_{i \in \mathcal{P}_s} p_i(t) < \sum_{i \in \mathcal{P}_s} d_i(t), \\ 0 & \text{otherwise.} \end{cases} \tag{11}$$

The Power Outage Count over the attack window $[t_{\text{start}}, t_{\text{end}}]$ is then:

$$\text{POC} = \frac{\sum_{t=t_{\text{start}}}^{t_{\text{end}}} \sum_{s=1}^{S} \text{Outage}_s(t)}{t_{\text{end}} - t_{\text{start}} + 1} \tag{12}$$

This metric counts the total number of outages across all substations during the attack period, normalized by total attack time. It reflects the impact of the attack on power availability for prosumers.

*4) **Potential Outage Risk (POR)**:* The Potential Outage Risk (POR) metric measures the proportion of time steps during the attack window where the Net Surplus Ratio (NSR) falls below a defined threshold $\delta$. The NSR represents the relative surplus or deficit of generated power compared to consumed power. The threshold $\delta$ signifies a critical margin; if the generated power is insufficient relative to the consumed power, there is a potential risk of an outage. POR thus indicates the system's vulnerability to near-outage conditions resulting from compromised power predictions. For a given threshold ratio $\delta \geq 0$, we define the potential outage risk at time $t$ as:

$$\text{POR}_t = \begin{cases} 1 & \text{if NSR}(t) < \delta, \\ 0 & \text{otherwise.} \end{cases} \tag{13}$$

where $\text{NSR}(t)$ is the Net Surplus Ratio at time defined as

$$\text{NSR}(t) = \frac{G_{\text{net}}(t) - D(t)}{D(t)}\,. \tag{14}$$

Here, $G_{\text{net}}(t) := G(t) + \sum_{i=1}^{P} p_i(t) - L(t)$ is the net generated power available to meet demand $D(t) := \sum_{i=1}^{P} d_i(t)$ at time $t$ (in kW), with $L(t) = L_{\text{trans}}(t) + L_{\text{dist}}(t)$ representing the total system loss at time $t$. When $\text{NSR}(t) < \delta$, the net surplus is below the critical threshold, indicating a potential risk of outage due to insufficient power generation relative to demand.

The Potential Outage Risk over the attack window $[t_{\text{start}}, t_{\text{end}}]$ is then calculated as:

$$\text{POR} = \frac{\sum_{t=t_{\text{start}}}^{t_{\text{end}}} \text{POR}_t}{t_{\text{end}} - t_{\text{start}} + 1}. \tag{15}$$

This metric represents the fraction of time steps within the attack window where the net surplus ratio fell below the acceptable threshold $\delta$, highlighting the system's susceptibility to near-outage risks during compromised conditions.

## IV. Performance Evaluation

### A. Experimental Setup

We developed a time-slotted simulator to capture measurement data from the prosumer level to the control unit, enabling the study of hybrid FDIA in a regional grid context. Using the IHEPC dataset [9], which provides one-minute sampled measurements over four years, we generated data for 5,000 prosumers by adding a random *bias* between -8% and +8% at each timestamp. Our simulated grid models two substations and five third-party aggregators serving these prosumers. The IEEE 57-Bus system underpins the network topology to ensure sufficient complexity for evaluating measurement manipulation and attack propagation.

Prosumers are allocated to aggregators using either a uniform or Pareto distribution. The uniform scheme evenly distributes prosumers, while the Pareto scheme (with $\alpha$ values from 1 to 5) creates imbalance, reflecting aggregator dominance scenarios. Similarly, prosumers are assigned to substations using either a balanced (50-50) or skewed Pareto (80-20) distribution, where one substation manages the majority.

To evaluate the impact of hybrid FDIA on system performance, we instantiate Eq. (10) with $\beta = 2$, $\rho = 0.9$, and $k = 2$ Km, and Eq. (13) with $\delta = \frac{\Gamma - 1}{2}$, operating below half the additional safety margin. These metrics collectively assess the effects of hybrid attacks on efficiency, operational costs, and system resiliency.

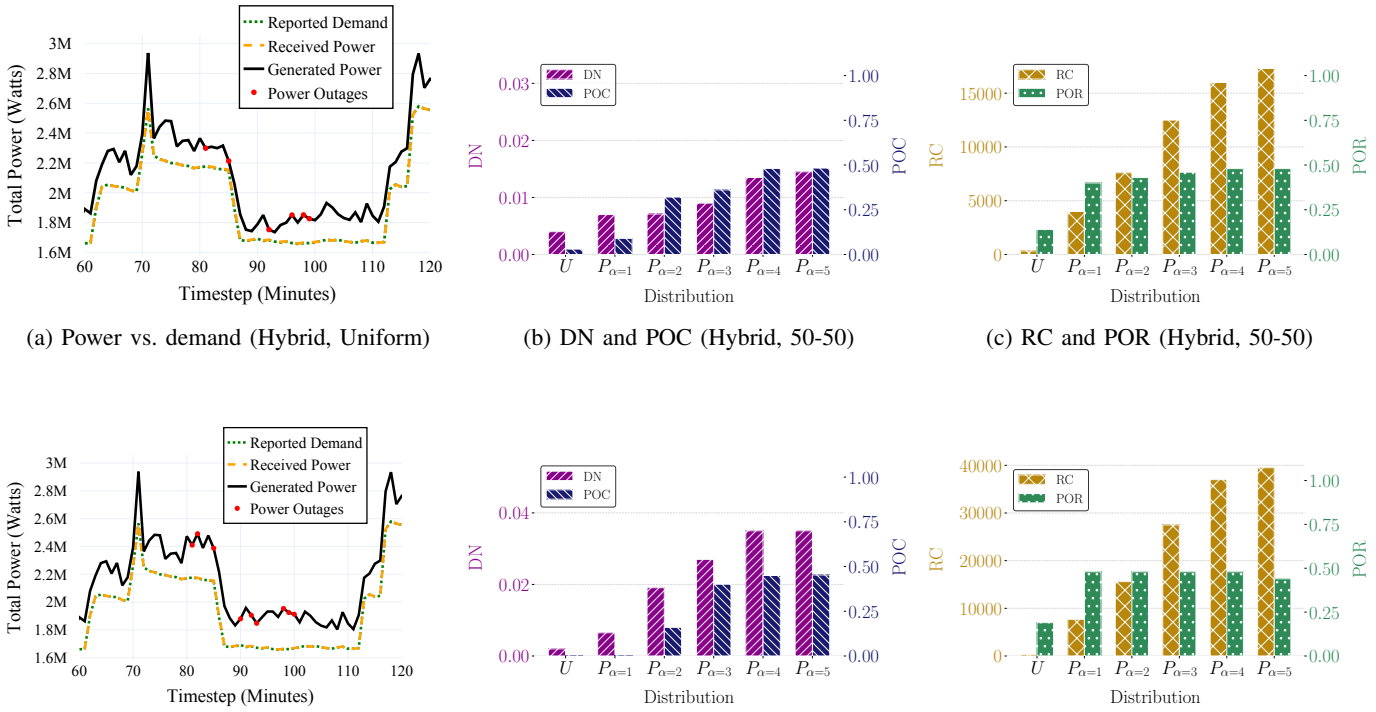
We assume the control unit uses an idealized predictor to estimate future demand within a bounded error of $\pm 5\%$. Generation requirements follow (6), with a safety margin factor $\Gamma = 1.1$, unless stated otherwise.

Experiments are conducted in the discrete-time simulation environment described in Section III. Each simulation spans 4 hours, with the attack scenario introduced for 20 minutes between time steps 80 and 100. Results are presented for a representative one-hour segment from each experiment.

### B. Impact of Hybrid FDIA on Smart Grid Stability

This section demonstrates the impact of Hybrid FDIA on smart grid stability. The attacker uses a linear program shown in Algorithm 1 to tamper with the reported demand of each prosumer connected to a compromised aggregator $\mathcal{P}_j$ while minimizing the total deviation, ensuring the aggregate values appear normal.

Figures 3a and 3d show the grid's response under two scenarios: a uniform 50-50 distribution of prosumers across substations, and an 80-20 Pareto distribution, where most prosumers are assigned to a single substation. In both cases, the Hybrid FDIA has minimal effect on the control unit's monitoring of overall generation, reported demand, and received power. However, the Pareto scenario exhibits greater fluctuations, highlighting the impact of prosumer concentration. These plots demonstrate the Control Unit perspective from Hybrid FDIA assuming that bad data detection systems are bypassed. Figures 3b and 3e capture scenarios where power redistribution is not possible, and excess generation is grounded, resulting in resource wastage. The DN metric reflects moderate deviation in generation due to offsetting inflation and deflation effects, while POC steadily rises with increasing $\alpha$, indicating that uneven prosumer distributions exacerbate outage risks despite stable overall generation. This underscores how skewed distributions lead to localized shortages, forcing excess power to be grounded.




(a) Power vs. demand (Hybrid, Uniform)

(b) DN and POC (Hybrid, 50-50)

(c) RC and POR (Hybrid, 50-50)

(d) Power vs. demand (Hybrid, Pareto, $\alpha = 3$)

(e) DN and POC (Hybrid, 80-20)

(f) RC and POR (Hybrid, 80-20)

Fig. 3: Results from hybrid attack scenarios with varying prosumer-to-aggregator distributions. (a)-(c) use a 50-50 prosumer-to-substation distribution; (d)-(f) use an 80-20 distribution.

Figures 3c and 3f examine scenarios where power redistribution mitigates imbalances at the cost of operational overhead. While grounding excess power avoids redistribution expenses but leaves shortages unresolved, redistribution offers resilience at increased cost. As $\alpha$ rises, RC increases, reflecting the growing cost of reallocating power to address disparities caused by asymmetric demand manipulation. The POR metric stabilizes at high values, indicating persistent near-outage risks in highly skewed settings.

In the 80-20 distribution case, the impact of hybrid attacks becomes more severe. DN grows faster than in the 50-50 scenario due to the difficulty of forecasting demand under a dominant aggregator. POC reaches its maximum, indicating that the deflation effect on the smaller substation severely affects power availability. RC also escalates substantially with increasing $\alpha$, reflecting the significant cost of redistributing power in such concentrated environments. The POR metric remains high, emphasizing the grid's vulnerability to shortages when aggregators control a large share of prosumers.

### C. Safety Margin Impact in Hybrid FDIAs Mitigation

In this set of experiments, we study the effect of safety margins on the impact of Hybrid FDIA, focusing on the trade-off between grid resiliency and operational costs. We consider an 80-20 distribution of prosumers to substations under a Pareto distribution with $\alpha = 3$ of prosumers to aggregators. Four safety margin levels are evaluated: $\Gamma_1 = 1.05$, $\Gamma_2 = 1.1$, $\Gamma_3 = 1.15$, and $\Gamma_4 = 1.2$.

| Safety Margin | DN | POC | RC | POR |
|---|---|---|---|---|
| $\Gamma_1$ | 0.025 | 0.63 | $3.08 \times 10^4$ | 0.76 |
| $\Gamma_2$ | 0.026 | 0.42 | $2.57 \times 10^4$ | 0.47 |
| $\Gamma_3$ | 0.026 | 0.18 | $1.77 \times 10^4$ | 0.18 |
| $\Gamma_4$ | 0.026 | 0.12 | $1.39 \times 10^4$ | 0.13 |

TABLE I: Impact of Safety Margin in Hybrid FDIA.

Table I presents the impact of different safety margins on grid performance and risk. While the DN metric remains relatively constant across all settings — reflecting the inherent nominal behavior of each setup — the POC metric significantly decreases as the safety margin increases. This highlights the role of higher safety margins in reducing unserved demand during attacks. However, increasing the safety margin incurs higher operational costs due to overcompensation in power generation. Thus, selecting an appropriate safety margin requires balancing reliability with cost-efficiency in mitigating Hybrid FDIA impacts. The RC and POR metrics further illustrate this trade-off. With lower safety margins, such as $\Gamma_1 = 1.05$, RC reaches its highest value, indicating frequent redistribution actions to balance power. As the safety margin increases, RC declines, suggesting that system over-preparedness reduces the need for reactive adjustments, thereby lowering operational costs. Likewise, POR decreases with larger safety margins, reflecting reduced risk of near-outage conditions. Maintaining a larger buffer enhances grid

stability and reduces vulnerability to critical shortages.

### D. Power Storage Impact on Grid Resilience in Hybrid FDIAs

This experiment evaluates how equipping 10% of uniformly selected prosumers with household battery systems enhances grid resilience against Hybrid FDIA. These batteries, sized to support residential loads for 2–4 hours, serve as local buffers during demand fluctuations, especially when aggregator reports are compromised. Battery capacities are determined based on the average prosumer demand over a 2–4 hour window, using historical data from our dataset.

| Prosumers Distribution | DN | POC | RC | POR |
|---|---|---|---|---|
| $U$ | 0.003 | 0.00 | $0.03 \times 10^4$ | 0.12 |
| $P_{\alpha=1}$ | 0.006 | 0.00 | $0.51 \times 10^4$ | 0.30 |
| $P_{\alpha=2}$ | 0.013 | 0.12 | $1.26 \times 10^4$ | 0.33 |
| $P_{\alpha=3}$ | 0.020 | 0.28 | $2.15 \times 10^4$ | 0.35 |
| $P_{\alpha=4}$ | 0.027 | 0.30 | $3.05 \times 10^4$ | 0.35 |
| $P_{\alpha=5}$ | 0.027 | 0.29 | $3.75 \times 10^4$ | 0.35 |

TABLE II: Impact of power storage under Hybrid FDIA.

Table II shows that adding power storage reduces DN and POC by approximately 25% and 30%, respectively, compared to non-storage scenarios. Batteries help stabilize the grid by locally compensating for demand spikes, maintaining closer alignment with predicted demand, and accelerating recovery from FDIA. The RC metric also shows an average reduction of 20%, as local storage alleviates grid stress by meeting part of the demand without triggering costly inter-substation redistributions. This effect becomes especially significant when a single aggregator controls a large prosumer base (i.e., $\alpha \geq 3$).

Similarly, POR decreases by about 20% with power storage, reflecting improved net surplus stability and a safer buffer above critical demand thresholds during attacks. Overall, integrating battery storage into prosumer systems enhances grid resilience and stability under Hybrid FDIAs.

## V. Conclusion

This paper presents a novel approach for analyzing the impact of hybrid false data injection attacks (Hybrid FDIA) targeting third-party aggregators in smart grids. Our model captures the interaction of compromised aggregators with grid control mechanisms and quantifies their impact through new performance metrics. Our results highlight key insights for grid security: (1) aggregators with a large share of prosumers pose a heightened security risk and should be prioritized in risk assessments; (2) hybrid FDIA can stealthily exploit load redistribution mechanisms to cause power loss and outages without triggering anomaly detection at the control level; and (3) mitigation strategies such as higher safety margins and distributed energy storage offer a trade-off between operational cost and resiliency.

## Acknowledgments

This work was supported in part by NSF Award CCF-2106339.

## References

[1] A. Aflaki, M. Gitizadeh, and B. Kantarci. Accuracy improvement of electrical load forecasting against new cyber-attack architectures. *Sustainable Cities and Society*, 77:103523, 2 2022.

[2] A. Bhattacharjee, A. K. Mondal, A. Verma, S. Mishra, and T. K. Saha. Deep latent space clustering for detection of stealthy false data injection attacks against ac state estimation in power systems. *IEEE Transactions on Smart Grid*, 14(3):2338–2351, May 2023.

[3] L. Che, X. Liu, Z. Li, and Y. Wen. False data injection attacks induced sequential outages in power systems. *IEEE Transactions on Power Systems*, 34(2):1513–1523, 2018.

[4] D.-H. Choi and L. Xie. Impact analysis of locational marginal price subject to power system topology errors. In *IEEE International Conference on Smart Grid Communications (SmartGridComm)*, pages 55–60, 2013.

[5] CISA Cybersecurity Advisory. PRC State-Sponsored Actors Compromise and Maintain Persistent Access to U.S. Critical Infrastructure. https://www.cisa.gov/news-events/cybersecurity-advisories/aa24-038a.

[6] M. Cui, J. Wang, and M. Yue. Machine learning-based anomaly detection for load forecasting under cyberattacks. *IEEE Transactions on Smart Grid*, 10:5724–5734, 9 2019.

[7] L. Gkatzikis, I. Koutsopoulos, and T. Salonidis. The role of aggregators in smart grid demand response markets. *IEEE Journal on Selected Areas in Communications*, 31:1247–1257, 2013.

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

[9] G. Hebrail and A. Berard. Individual Household Electric Power Consumption. UCI Machine Learning Repository, 2006. DOI: https://doi.org/10.24432/C58K54.

[10] J. M. Koushyar, M. Guirguis, and G. Atia. The role of aggregator topology on the impact of false data injection attacks. In *2025 IEEE 11th International Conference on Intelligent Data and Security (IDS)*, pages 1–4, 2025.

[11] G. Liang, J. Zhao, F. Luo, S. R. Weller, and Z. Y. Dong. A review of false data injection attacks against modern power systems. *IEEE Transactions on Smart Grid*, 8(4):1630–1638, 2016.

[12] J. Lin, W. Yu, X. Yang, G. Xu, and W. Zhao. On false data injection attacks against distributed energy routing in smart grid. In *IEEE/ACM Third International Conference on Cyber-Physical Systems*, pages 183–192, 2012.

[13] X. Liu, Z. Bao, D. Lu, and Z. Li. Modeling of local false data injection attacks with reduced network information. *IEEE Transactions on Smart Grid*, 6(4):1686–1696, 2015.

[14] 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)*, 14(1):1–33, 2011.

[15] K. Mahapatra, N. R. Chaudhuri, and R. Kavasseri. Bad data detection in pmu measurements using principal component analysis. In *2016 North American Power Symposium (NAPS)*, pages 1–6. IEEE, 2016.

[16] K. Manandhar, X. Cao, F. Hu, and Y. Liu. 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):370–379, 2014.

[17] A. S. Musleh, G. Chen, and Z. Y. Dong. A survey on the detection algorithms for false data injection attacks in smart grids. *IEEE Transactions on Smart Grid*, 11(3):2218–2234, 2019.

[18] T. S. Sreeram and S. Krishna. Graph-based assessment of vulnerability to false data injection attacks in distribution networks. *IEEE Transactions on Power Systems*, 39(2):4510–4520, March 2024.

[19] X. Su, C. Deng, J. Yang, F. Li, C. Li, Y. Fu, and Z. Y. Dong. Damgat-based interpretable detection of false data injection attacks in smart grids. *IEEE Transactions on Smart Grid*, 15(4):4182–4195, July 2024.

[20] Y. Xiang, Z. Ding, Y. Zhang, and L. Wang. Power system reliability evaluation considering load redistribution attacks. *IEEE Transactions on Smart Grid*, 8(2):889–901, 2016.

[21] Y. Xiang, Z. Ding, Y. Zhang, and L. Wang. Power system reliability evaluation considering load redistribution attacks. *IEEE Transactions on Smart Grid*, 8(2):889–901, March 2017.

[22] L. Xie, Y. Mo, and B. Sinopoli. Integrity data attacks in power market operations. *IEEE Transactions on Smart Grid*, 2(4):659–666, 2011.