# A Game-Theoretic Approach for PMU Deployment Against False Data Injection Attacks

Sajjad Maleki, Subhash Lakshminarayana, E. Veronica Belmega, Carsten Maple

# A Game-Theoretic Approach for PMU Deployment Against False Data Injection Attacks

Sajjad Maleki*†, Subhash Lakshminarayana*, E. Veronica Belmega‡†, and Carsten Maple§
* School of Engineering, University of Warwick, Coventry, United Kingdom
† ETIS UMR 8051, CY Cergy Paris Université, ENSEA, CNRS, F-95000, Cergy, France
‡Univ. Gustave Eiffel, CNRS, LIGM, F-77454, Marne-la-Vallée, France
§WMG, University of Warwick, Coventry, United Kingdom
Email: sajjad.maleki@warwick.ac.uk, subhash.lakshminarayana@warwick.ac.uk,
veronica.belmega@esiee.fr, cm@warwick.ac.uk

***Abstract*—Phasor Measurement Units (PMUs) are used in the measurement, control and protection of power grids. However, deploying PMUs at every bus in a power system is prohibitively expensive, necessitating their optimal placement to ensure the system observability with minimum cost. This partial PMU placement renders the system vulnerable to False Data Injection Attacks (FDIAs). This paper proposes a zero-sum game-based approach to strategically place an additional PMU (after the initial optimal PMU deployment for full observability) to bolster robustness against FDIAs by introducing redundancy in attack-susceptible areas. To compute the Nash equilibrium (NE) solution, we leverage a reinforcement learning algorithm that mitigates the need for complete knowledge of the opponent's actions. The proposed PMU deployment algorithm increases the detection rate of FDIA by $36\%$ compared to benchmark approaches.**

***Index Terms*—Cybersecurity, False Data Injection Attacks (FDIA), Zero-sum Games, Nash Equilibrium, PMU.**

## I. Introduction

Phasor Measurement Units (PMUs) have become pivotal in power systems thanks to their increased accuracy, high sampling rate, and time-synchronised measurements. However, they must be carefully and sparingly deployed because of their high cost. As a result, optimal PMU placement that ensures full system observability has been widely investigated in the literature [1], [2].

PMU measurements play a crucial role in the state estimation of power systems, providing essential data for their accurate operation. However, the communication technologies for PMU operations (communications channels and global positioning system for synchronization) introduce cyber security risks. In particular, the communication channel is susceptible to attacks aimed at manipulating PMU measurements through the injection of false data [3].

The presence of a PMU in a bus or its adjacent buses allows an operator to obtain the necessary measurements to determine the state of the bus. If all system states can be obtained through PMU measurements, then the system is considered fully observable. Full observability of the system is necessary for state estimation of the power systems. Consequently, many works including [4], [5] proposed different methods for optimal PMU placement to guarantee the observability of the system. In [4], the authors have introduced statistical criteria to determine the optimal locations for PMUs and employ a multi-criteria decision-making approach known as the analytical hierarchical process to optimize their placement. In [5], a hierarchical process for enhancing power system observability is introduced, which proposes an iterative approach in which the operator adds a PMU at a time to guarantee the maximum observability of the system with the available number of PMUs. In [6] and [7], authors propose greedy algorithms for PMU deployment not only to achieve full observability but also to detect stealthy FDIAs on the Supervisory Control and Data Acquisition (SCADA) measurements. It's worth noting that these papers assume that the PMUs themselves are secure against FDIAs. Reference [8] utilizes a modified optimal PMU placement algorithm to develop a defense solution to attacks causing a surge in the electricity price due to the false data injected by adversaries. Authors of [9] propose that PMUs themselves could be vulnerable to FDIAs. Consequently, they advocate for a new PMU placement method to mitigate such attacks. Their research demonstrates that increasing the number of PMUs reduces the probability of undetected FDIAs. However, they do not consider the presence of a strategic attacker.

To the best of our knowledge, [10] is the first work to propose a game theoretic approach to tackle the optimal PMU placement problem in the presence of cyber threats initiated by strategic attackers. Their methodology involves implementing a bi-level game in which the attacker's objective is to deviate the state estimation from the true value in a stealthy manner. On the other hand, the operator tries to optimally place the PMUs to maintain the full observability of the system and minimize the state estimation errors.

Previous works have extensively investigated the observability of the system and enhanced the robustness of PMU measurements against manipulation in state estimation by adversaries. However, there remains a gap in the literature regarding research on mitigating the likelihood of successful FDIAs in the presence of a strategic attacker. Unlike [10], which proposes a completely new PMU placement scheme to achieve robustness against FDIAs, this paper assumes that

This work has been supported in part by the EUTOPIA PhD Cofund WALL-EE project between the University of Warwick, UK and CY Cergy Paris University, France and in part by the PETRAS National Centre of Excellence for IoT Systems Cybersecurity through the U.K. EPSRC under Grant EP/S035362/1.

optimal PMU placement has already been made to ensure system observability. Our focus is on adding an additional PMU to the system to decrease the likelihood of successful FDIA through the resulting redundancy in measurements. Thus, our approach is more practical for securing power systems that are already starting to witness PMU deployment, mainly with the observability criterion in mind.

In this work, we assume that a strategic attacker aims to launch such an FDIA on a PMU's measurement that is not observable by any other PMU while having the biggest possible impact on the state estimation. We exploit a two-player zero-sum non-cooperative game [11] to analyze the interaction between the attacker and defender and investigate the robust minimax defense solutions at the NE. The goal of the defender is to increase the FDIA detection rate against strategic attacks by adding redundant observing PMU to attack-prone measurements. Then, we propose a reinforcement learning algorithm drawn from multi-armed bandits, namely exponential weights for exploration and exploitation (EXP3), to show that the robust NE solution can be computed without the full knowledge of the game at the defender's end but via iterative interactions with the attacker.

We analyze the impact of incorporating an additional PMU using a game-theoretic approach applied to the IEEE 14-bus system. This analysis involves comparing the outcomes with two benchmarks obtained by: (i) not adding any additional PMUs and (ii) randomly adding an extra PMU. A higher attack detection percentage of the proposed solution encourages the expansion of this work to larger systems which may require more than one PMU to reach a certain level of detection.

The key contributions of this paper are:

- Devising a minimax NE solution for obtaining a robust and low-cost defense strategy against strategic FDIA attacks by placing an additional PMU in the system.
- Exploiting reinforcement learning to find the NE solution in an iterative manner without requiring any knowledge about the attacker's actions.
- We present extensive simulation results using the IEEE 14-bus system and compare the performance of the proposed algorithm against a benchmark technique. Results exhibit an enhanced detection rate for strategic attacks while using our proposed method.

The rest of the paper is organised as follows: Section II describes the problem formulation; Section III discusses the proposed solution while in Section IV, our numerical results are presented and discussed and, finally, Section V concludes the paper.

## II. Problem Formulation

In this section, we first introduce the power system model and optimal PMU placement problem, before delving into the issue of the increasing robustness of PMU measurements against FDIAs.

### A. Power system model

Let the graph $\mathcal{E} = \{\mathcal{N}, \mathcal{L}\}$ represent the power system, where $\mathcal{N}$ is the set of buses and $\mathcal{L}$ is the set of lines. Also,

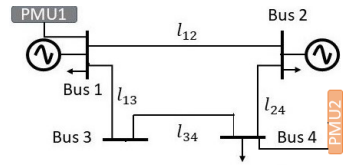


Fig. 1. An example power system with four buses and two PMUs

$\mathcal{N}_{PMU}$ and $\overline{\mathcal{N}}_{PMU}$ are the subsets of buses with and without PMUs respectively, such that $\overline{\mathcal{N}}_{PMU} \cup \mathcal{N}_{PMU} = \mathcal{N}$ and $\overline{\mathcal{N}}_{PMU} \cap \mathcal{N}_{PMU} = \emptyset$. There are $n = |\mathcal{N}|$ buses and $\ell = |\mathcal{L}|$ lines in the power system. Also, $\mathcal{A}_i$ represents the adjacent buses to the $i^{th}$ bus. Fig. 1 exhibits a sample 4-bus grid, where bus 1 is the PMU bus and buses 2, 3, and 4 are non-PMU buses (before adding the PMU 2). In this system, $\mathcal{A}_1 = \{2, 3\}$ and bus 3 is a zero injection bus (ZIB).

This work is developed based on DC state estimation, so the states of the system in this work are phase angles denoted by $\theta$. PMUs measure the state of their host buses directly, while the states of the buses in $\overline{\mathcal{N}}_{PMU}$ can be calculated as follows:

$$\theta_j = \theta_i - p_{ij} x_{ij}, \tag{1}$$

where $\theta_i$ is the phase angle of the bus with PMU and $\theta_j,\ j \in \mathcal{A}_i$, is the phase angle of the adjacent bus to the $i^{th}$ bus. Also, $p_{ij}$ represents the flowing power on the line between buses $i$ and $j$ while $x_{ij}$ is the reactance of the line. Note that all $p_{ij}$ values are measured values by PMUs which could be under attack or not. To make it possible to measure the states of all buses directly or using (1), each bus is required to either have a PMU in it or one of its adjacent buses. This requirement is the basis of optimal PMU placement.

### B. Optimal PMU placement

A bus within a power system is considered "observable" under two conditions: either it contains a PMU, or a PMU is installed in any of its adjacent buses. A power system achieves "full observability" when every bus within the system meets the above criteria for observability. Minimizing the following criterion guarantees the full observability of the system while placing a minimum number of PMUs [4]:

$$\min \sum_{m=1}^{n} w_m y_m \quad s.t. \quad g(\mathbf{Y}) \geq \mathbf{B}, \tag{2}$$

where $w_m$ is the cost of installing PMU in the $m^{th}$ bus, $\mathbf{Y} = [y_1 \ \ y_2 \ldots y_n]^T$ is the binary decision variable for PMU placement with entries:

$$y_m = \begin{cases} 1, & \text{if there is a PMU in the } m^{th} \text{ bus} \\ 0, & \text{otherwise,} \end{cases} \tag{3}$$

also $\mathbf{B} = [1\ 1 \ldots 1]^T$ of dimension $n$. At last, $g(\mathbf{Y})$ is a vector of the same dimension as $\mathbf{B}$ and $\mathbf{Y}$ of entries:

$$g_m(\mathbf{Y}) = \begin{cases} 1, & \text{if } m \in \mathcal{N}_{PMU} \text{ or there is a PMU bus in } \mathcal{A}_m \\ 0, & \text{otherwise.} \end{cases} \tag{4}$$

Equation (4) is a binary function defining whether or not the bus $m$ is observable by at least one PMU.

Power systems can incorporate zero injection buses (ZIBs), which do not contain any load or generation components. According to Kirchhoff's Current Law (KCL), if the power flow is known in all lines connected to a ZIB except for one, then the current in that unknown line can be revealed. Consequently, if all buses connected to a ZIB are observable, applying KCL makes the ZIB observable as well. Furthermore, in the case where all buses connected to an observable ZIB are observable except for one, KCL can also be employed to make the previously unobservable bus observable as well [12]. Optimal PMU placement of power systems with ZIB follows the same steps with just introducing $g'_m(\mathbf{Y})$ instead of $g_m(\mathbf{Y})$. Algorithm 1 represent the computation of $g'_m(\mathbf{Y})$.

**Algorithm 1:** Observability of system with ZIB

**Data:** $ZIB$, $g_m(\mathbf{Y})$
**Result:** $g'_m(\mathbf{Y})$
**if** $g_m(\boldsymbol{Y}) = 1$ **then**
  $g'_m(\mathbf{Y}) = 1$
**else**
  **if** $(g_k(\boldsymbol{Y}) = 1,\ \forall k \in \mathcal{A}_m)$ ***and*** $(\mathcal{A}_m \cup \{m\}) \cap ZIB \neq \emptyset$
  **then**
    $g'_m(\mathbf{Y}) = 1$
  **else**
    $g'_m(\mathbf{Y}) = 0$
  **end**
**end**

In Algorithm 1, $ZIB$ represents the set of ZIB buses. If an attacker endeavors to alter the value of $\theta_i$ for a PMU bus, as per (1), all $\theta_j$ for every $j$ within the set $\mathcal{A}_i$ will also be modified. Consequently, if any bus $j \in \mathcal{A}_i$ is linked to another PMU within its vicinity, any disparity in the calculated or observed phase angles between different PMUs for the same bus will immediately raise an anomaly flag at the defender's end. In contrast, if the attacker manipulates a $p_{ij}$ value, only one phase angle, as computed by (1), will be affected. Therefore, this type of attack can go undetected if the bus at the opposite end of the target line also lacks a neighbouring PMU for monitoring purposes.

In conclusion, although certain attack scenarios have more severe consequences, they also tend to have a higher likelihood of detection.

### C. Game theoretic formulation

In this paper, a game-theoretic approach has been devised to identify the best actions for both attacker and defender and compute the FDIA detection rate based on them.

The strategic interaction between the attacker and defender can be formulated as a two-player zero-sum game [11]. This game, denoted by $\mathcal{G} = \Big(\mathcal{T} \triangleq \{D, A\};\ (\mathcal{S}_D, \mathcal{S}_A);\ (F_D, F_A)\Big)$, is composed of the set of players, the sets of possible actions of the players and the utility (or reward or payoff) functions of the players, respectively, and which will be defined next.

The set of the players of this game is $\mathcal{T} = \{D, A\}$, where $D$ denotes the defender and $A$ the attacker. The set of action profiles for the game is $\mathcal{S} = \mathcal{S}_D \times \mathcal{S}_A$, where $\mathcal{S}_D$ is the set of discrete defense actions and $\mathcal{S}_A$ is the set of discrete attacks.

The defense action $d \in \mathcal{S}_D \subseteq \overline{\mathcal{N}}_{PMU}$ is the index of nominal buses for a potential additional PMU and it is a subset of buses that do not have a PMU already. Selecting the candidates for installing additional PMU and forming $\mathcal{S}_D$ is based on the knowledge of the defender of the topology of the system. A set of defensive actions could encompass all non-PMU buses. However, to streamline the process and minimize computational costs, we can exclude buses that cause the same redundant observability as others, thereby reducing the number of actions required. The process of picking candidate buses is detailed in Section II-D.

The attacker selects a target PMU to manipulate either part or all of its measurements consisting of: the phase angle of the bus and the power flows of connected lines to it. Thus, a specific attack action can be written as $a = (u, \mathcal{V}_u)$, in which $u \in \mathcal{N}_{PMU}$ is the index of the bus containing the PMU under attack and $\mathcal{V}_u \in \Pi(\mathcal{P}_u \cup \{\theta_u\})$ denotes the subset of the measurements of the target PMU at bus $u$ that are manipulated. Also, $\Pi(\mathcal{P}_u \cup \{\theta_u\})$ represents the set of partitions of $\mathcal{P}_u \cup \{\theta_u\}$ containing all measurements of the target PMU and $\mathcal{P}_u = \{p_{uk}, \forall\ k \in \mathcal{N} \mid (u,k) \in \mathcal{L}\}$ is the set of all the power flows $p_{uk}$ of the lines that are connected to bus $u$. The set of all possible attacks can be defined as follows: $\mathcal{S}_A = \{(u, \mathcal{V}_u) \in \mathcal{N}_{PMU} \times \Pi(\mathcal{P}_u \cup \{\theta_u\})\}$. The number of such attacks is $|\mathcal{S}_A| = \sum_{u \in \mathcal{N}_{PMU}} (2^{|\mathcal{L}_u|+1} - 1)$.

For the attacked line flows, there are two buses affected: (i) the attacked PMU bus, and (ii) the bus which is at the other end of the line (the line with manipulated power flow value). Also, if a phase angle is manipulated, all of the adjacent buses to the attacked PMU are affected because calculating their phase angles is dependent on the phase angle of the PMU bus. For a precise attack $a \in \mathcal{S}_A$, we form the set of affected buses and call it $\mathcal{C}_a$. For example in the 4-bus system depicted in Fig. 1, the measurements of PMU 1 that can be tampered with are: $\{p_{12}, p_{13}, \theta_1\}$. So the set $\mathcal{S}_A$ for it is thus:

$$\mathcal{S}_A = \{(1, \{p_{12}\});\ (1, \{p_{13}\});\ (1, \{\theta_1\});\ (1, \{p_{12}, p_{13}\});\ (1, \{p_{12}, \theta_1\});\ (1, \{p_{13}, \theta_1\});\ (1, \{p_{12}, p_{13}, \theta_1\})\}\,.$$

The attacker aims to maximize the deviations of phase angles measured by PMUs from their true values. The effect of the FDIA on the phase angles is:

$$\mathbf{E}(a, d) = \Theta - \Theta_{bad}, \tag{5}$$

where $\Theta = [\theta_1 \quad \theta_2 \quad \ldots \theta_n]^T$ is the vector of accurate phase angles and $\Theta_{bad} = [\theta_{bad,1} \quad \theta_{bad,2} \quad \ldots \theta_{bad,n}]^T$ is the measured phase angles vector of under attack power system. The defender's goal is to add one extra PMU for the redundant measurement of attack-prone properties in order to detect the possible FDIA. Let $O_k(d)$ represent the number of PMUs observing the bus $k$,

$$O_k(d) = |\mathcal{N}_{PMU} \cap \mathcal{A}_k|\,. \tag{6}$$

In other words, $O_k(d)$ is the total number of PMUs in the $k^{th}$ bus and its adjacency after the defensive action $d$. Consider Fig. 1 as an example. When only PMU 1 (as per the initial optimal PMU placement) is deployed within the system, the total number of PMUs observing bus 2 is equal to 1. Now, if the operator introduces another PMU in bus 4 (i.e., $d = 4$), then $O_2(4) = 2$ as there will be two PMUs observing bus 2. In this scenario, a potential anomaly is highlighted in the measurements of at least one of the PMUs if the phase angle calculations from both observing PMUs differ.

We consider that the attacker's objective is to pick an action $a$ that remains undetected and that maximizes the FDIA effect. To rigorously define the attacker's reward, we introduce $O(a, d)$ as the number of PMUs observing the bus with affected phase angle calculation after the attack and defense actions:

$$O(a,d) = \left| \mathcal{N}_{PMU} \cap \left\{ \bigcup_{i \in \mathcal{C}_a} \mathcal{A}_i \right\} \right|, \tag{7}$$

in which $\mathcal{C}_a$ is the set of buses whose phase angle measurements are affected by the attack. If $O(a, d) > 1$, then the attack is detected and the reward of the attacker is set to zero (worst case for the attacker). Otherwise, if $O(a, d) = 1$, the attack is undetected and its effect is maximized by maximizing $\|\mathbf{E}(a, d)\|$. To sum up, the reward of the attacker is:

$$F_A(a,d) = \begin{cases} \|\mathbf{E}(a,d)\|, & \text{if} \quad O(a,d) = 1 \\ 0, & \text{if} \quad O(a,d) > 1, \end{cases} \tag{8}$$

which translates that, if only one PMU observes the target bus, the attacker is undetected and can stealthily manipulate measurements, resulting in a reward of $\|\mathbf{E}(a, d)\|$. However, if multiple PMUs monitor the target bus, the attack is detected, rendering null the attacker reward.

We further consider that the defender's objective is the exact opposite by maximizing

$$F_D(a,d) = -F_A(a,d), \tag{9}$$

which is always negative and maximized when the FDIA attack is detected and $F_D(a, d) = F_A(a, d) = 0$.

### D. Defensive actions

The defender has full knowledge of the topology of the system and, hence, can use this information to specify more accurately the candidate defense actions by selecting only a subset of $\overline{\mathcal{N}}_{PMU}$. The reason is that the defender wants to add one extra PMU to increase the system's observability in the least observable buses, whereas some buses in $\overline{\mathcal{N}}_{PMU}$ have guaranteed double observability by the PMUs in place.

Let $\mathcal{B}_\alpha \subseteq \overline{\mathcal{N}}_{PMU}, \ \forall \ \alpha \in \overline{\mathcal{N}}_{PMU}$ denote the subset of non-PMU buses with single observability, which turn double observable after adding a PMU to $\alpha^{th}$ bus. Algorithm 2 describes the process of selecting the candidate buses.

**Algorithm 2:** Defining defender's set of actions

**Input :** $\overline{\mathcal{N}}_{PMU}$, $\mathcal{A}_i$
**Output:** $\mathcal{S}_D$

1 Specify $\mathcal{B}_\alpha$ for all of the non-PMU buses.
2 Classify the non-PMU buses $\alpha \in \overline{\mathcal{N}}_{PMU}$ as follows: buses $\alpha_1$ and $\alpha_2$ belong to the same class if: either $\mathcal{B}_{\alpha_1} \equiv \mathcal{B}_{\alpha_2}$ or $\mathcal{B}_{\alpha_1} \subset \mathcal{B}_{\alpha_2}$.
3 From each class, select the one bus with the largest set $\mathcal{B}_\alpha$.
4 The above-selected buses form $\mathcal{S}_D$.

## III. Robust minimax defense strategy

Having defined all the components of the two-player non-cooperative game $\mathcal{G}$ under study, we will now proceed to find the mixed Nash equilibrium solution of this game, which will lead to the robust minimax defense solution in terms of placing one additional PMU against a strategic attacker.

The NE is the natural outcome of a non-cooperative game and is a state, or an action profile, from which the players cannot unilaterally deviate without losing in terms of their individual rewards. The mathematical definition of the NE in pure strategies for the game $\mathcal{G}$ under study is given as follows.

**Definition 1.** *An action profile $(a^*, d^*) \in \mathcal{S}$ is a NE in pure strategy of the non-cooperative game $\mathcal{G}$, iff $F_A(a^*, d^*) \geq F_A(a, d^*), \ \forall \ a \ \in \mathcal{S}_A$ and $F_D(a^*, d^*) \geq F_D(a^*, d)$,(or equivalently, $F_A(a^*, d^*) \leq F_A(a^*, d)) \ \forall \ d \in \mathcal{S}_D$.*

Our finite and discrete game $\mathcal{G}$ might not have a such pure NE solution. Instead, the game always has at least one mixed strategy NE solution [13]. A mixed strategy NE is the solution of the extension of the game to mixed strategies, in which the players choose random actions following certain probability distributions. Therefore, our objective is to find a mixed-strategy NE solution. In our case, the attacker choose a random action $a \in \mathcal{A}$ following a discrete probability distribution $\sigma_A = (\rho_1, \rho_2, \ldots, \rho_{|\mathcal{S}_A|}) \in \Delta_A$, such that $\rho_k$ denotes the probability of selecting the $k$-th action in $\mathcal{S}_A$. Similarly, the defender choose a random action $d \in \mathcal{S}_D$ following the discrete probability $\sigma_D = (\mu_1, \mu_2, \ldots, \mu_{|\mathcal{S}_D|}) \in \Delta_D$ such that $\mu_k$ denotes the probability of selecting the $k$-th pure defense action in $\mathcal{S}_D$. We will also make use of the notations $\rho_a$ and $\mu_d$ to denote the probabilities of selecting arbitrary actions $a \in \mathcal{S}_A$ and $d \in \mathcal{S}_D$, respectively. At last, the sets $\Delta_A$ and $\Delta_D$ are the corresponding discrete probability distribution simplices:

$$\Delta_A = \left\{ \sigma_A = (\rho_1, \ldots, \rho_{|\mathcal{S}_A|}) \in [0,1]^{|\mathcal{S}_A|} \;\middle|\; \sum_{k=1}^{|\mathcal{S}_A|} \rho_k = 1 \right\}$$

$$\Delta_D = \left\{ \sigma_D = (\mu_1, \ldots, \mu_{|\mathcal{S}_D|}) \in [0,1]^{|\mathcal{S}_D|} \;\middle|\; \sum_{k=1}^{|\mathcal{S}_D|} \mu_k = 1 \right\}.$$

The modified rewards of the extended game are the mathematical expectations of the obtained rewards given the

randomly chosen actions that follow the distribution $\sigma_A$ for the attacker and $\sigma_D$ for the defender:

$$\hat{F}_A(\sigma_A, \sigma_D) = \sum_{a \in \mathcal{S}_A} \sum_{d \in \mathcal{S}_D} F_A(a, d)\ \rho_a\ \mu_d. \tag{10}$$

To sum up, the extended game to mixed strategies can be defined as $\hat{\mathcal{G}} = \Big(\mathcal{T} \triangleq \{D, A\}; (\Delta_D, \Delta_A); (\hat{F}_D, \hat{F}_A)\Big)$ and the mixed strategy NE is defined as follows.

**Definition 2.** *A mixed strategy profile $(\sigma_a^*, \sigma_d^*) \in \Delta_A \times \Delta_D$ is a mixed NE of the game $\mathcal{G}$, iff it is a NE of the extended game $\hat{\mathcal{G}}$ such that $\hat{F}_A(\sigma_A^*, \sigma_D^*) \geq \hat{F}_A(\sigma_A, \sigma_D^*),\ \forall\ \sigma_A \in \Delta_A$, and $\hat{F}_D(\sigma_A^*, \sigma_D^*) \geq \hat{F}_D(\sigma_A^*, \sigma_D),\ \forall\ \sigma_D \in \Delta_D$.*

The mixed strategy NE can be calculated via the Von-Neumann indifference principle [13] by solving a certain number of linear systems of equations and inequalities. This number grows exponentially with the number of actions of the players. When the number of actions grows large, finding the NE via the Von-Neumann indifference principle becomes intractable. The Lemke-Howson method [14] is the most efficient alternative known to date. However, in the worst cases, the complexity of the Lemke-Howson algorithm is the same as the Von-Neumann indifference principle.

Both the Von-Neumann indifference principle and the Lemke-Howson method require full knowledge of the payoffs and sets of actions of players. Lack of knowledge about the payoffs of the opponent player hinders the game from being solved by them. Inspired by [11], the NE is found using the EXP3 algorithm from the multi-armed bandits framework. The main desirable feature of EXP3 is that neither player is required to have full knowledge of the game.

Indeed, in EXP3 algorithm, the agent $A$ or $D$ draws a random action $a_t$ or $d_t$ at each iteration $t$ from the distribution $\sigma_{A,t} = (\rho_{1,t}, \rho_{2,t}, \ldots, \rho_{|\mathcal{S}_A|,t})$ or $\sigma_{D,t} = (\mu_{1,t}, \mu_{2,t}, \ldots, \mu_{|\mathcal{S}_D|,t})$ and observes its own resulting reward. The rewards of the non-chosen actions are not known and have to be estimated. For instance, at the attacker side, the estimated rewards are as follows: [1]

$$\hat{F}_{A,t}(a) = \frac{F_A(a_t, d_t)\mathbb{1}[a = a_t] + \beta_t}{\sigma_{A,t}(a)},\ \forall a \in \mathcal{S}_a \tag{11}$$

where $\beta_t > 0$ controls the estimator's variance. Additionally, $\sigma_{A,t}(a)$ signifies the probability of choosing action $a$ at iteration $t$. The next equation captures the cumulative score of these actions and determines the performance of actions in the past.

$$G_{A,t}(a) = \sum_{\tau=1}^{t} \eta_\tau \hat{F}_{A,\tau}(a) \tag{12}$$

where $\eta_\tau > 0$ is a learning parameter. Then, the updated probability distribution $\sigma_{A,t+1}$ is computed,

$$\sigma_{A,t+1}(a) = \gamma_t \frac{1}{|\mathcal{S}_A|} + (1 - \gamma_t) \frac{\exp(G_{A,t}(a))}{\sum_{r \in \mathcal{S}_A} \exp(G_{A,t}(r)}, \forall a, \tag{13}$$

[1]Similar equations can be written for the defender and are omitted here.

where $\gamma_t \in (0, 1]$ is another learning parameter. Similarly, the defender updates its own probability distribution. The process repeats until convergence. According to [15], the empirical frequency of actions of EXP3, defined below, converges to the NE.

$$\bar{\sigma}_{A,t} = \frac{1}{\sum_{\tau=1}^{t} \eta_\tau} \sum_{\tau=1}^{t} \eta_\tau \sigma_{A,\tau}. \tag{14}$$

## IV. Numerical results

In this section, we evaluate the proposed framework for the IEEE 14-bus system. The initial optimal PMU locations for satisfying the full observability condition are chosen as in [4]. In this grid, as the result of optimal PMU placement, buses 2, 6, 7, and 9 while not considering ZIB and buses 2, 6, and 9 with considering ZIB are PMU buses. Note that in this test system, $ZIB = \{7\}$ as bus 7 is the ZIB. Also, Following the Section II-D, the set of candidate buses to place the additional PMU is $\mathcal{S}_D = \{1, 3, 8, 10, 13\}$.

Additionally, the NE has been calculated by the Lemke-Hawson method with full game knowledge and compared to the EXP3 algorithm, which is expected to converge to the NE.

### A. Evaluation of the NE solution

We first compute the NE solution via Lemke-Howson and EXP3 algorithms for comparison purposes. Tables I and II contain the probability distributions of NE for defender and attacker. Additionally, they contain the results calculated in the last iteration of the EXP3. The provided values highlight the fact that EXP3 enables us to compute the mixed strategy NE without requiring full knowledge about the set of actions and payoffs of the opponent. indeed, for two-player zero-sum games, the EXP3 algorithm converges to the NE solution [16]. Now, to assess the performance of the NE solution in terms of FDIA detection rate, we introduce the probability of detecting and not detecting an attack given a randomly chosen defense action $d \in \mathcal{S}_D$ and attack action $a \in \mathcal{S}_A$. Since the actions of the attacker and the defender are independent of one another, we can calculate the probability of detecting and not detecting FDIA as follows:

$$\Pr[O(a, d) > 1] = \sum_{a \in \mathcal{S}_A,\ d \in \mathcal{S}_D} \Pr(a)\ \Pr(d)\ \mathbb{1}_{[O(a,d)>1]}, \tag{15}$$

and $\Pr[O(a, d) = 1] = 1 - \Pr[O(a, d) > 1]$ given that $O(a, d) \geq 1$ always, where $\mathbb{1}_{[t]}$ is the indicator function that equals one if the condition $t$ is true and zero otherwise.

At the NE $(\sigma_A^*, \sigma_D^*)$, the attack detection rate is given by:

$$\Pr[O(a, d) > 1] = \sum_{a \in \mathcal{S}_A,\ d \in \mathcal{S}_D} \rho_a^*\ \mu_d^*\ \mathbb{1}_{[O(a,d)>1]}. \tag{16}$$

The above will be compared with a naive defense strategy where the probability of choosing an action $d \in \mathcal{S}_D$ is uniformly distributed: $\Pr(d) = 1/|\mathcal{S}_D|, \forall\ d \in \mathcal{S}_D$. For this naive defense strategy, the attack detection rate is given by:

$$\Pr[O(a, d) > 1] = \frac{1}{|\mathcal{S}_D|} \sum_{a \in \mathcal{S}_A,\ d \in \mathcal{S}_D} \rho_a^*\ \ \mathbb{1}_{[O(a,d)>1]}, \tag{17}$$

TABLE I
NE SOLUTION FOR THE IEEE 14-BUS SYSTEM WITHOUT ZIB CALCULATED VIA THE LEMKE-HOWSON AND EXP3 METHODS

| Action (Attacker) | NE probability distributions | |
|---|---|---|
| | Lemke-Howson | EXP3 |
| Line 1-2 | 0.3113 | 0.3220 |
| Line 2-3 | 0.4923 | 0.4726 |
| Lines 6-11, 6-12, 6-13 | 0.1965 | 0.2044 |
| Other Actions | 0 | $\approx 0$ |
| Action (Defender) | | |
| Bus 1 | 0.3774 | 0.3758 |
| Bus 3 | 0.6071 | 0.6062 |
| Bus 10 | 0.0155 | $< 0.01$ |
| Other Actions | 0 | $\approx 0$ |

TABLE II
NE SOLUTION FOR IEEE 14-BUS SYSTEM WITH ZIB CALCULATED VIA LEMKE-HOWSON AND EXP3 METHODS

| Action (Attacker) | NE probability distributions | |
|---|---|---|
| | Lemke-Howson | EXP3 |
| Line 2-3 | 0.7685 | 0.7401 |
| Lines 6-11, 6-12, 6-13 | 0.2315 | 0.2254 |
| Other Actions | 0 | $< 0.01$ |
| Action (Defender) | | |
| Bus 3 | 0.7685 | 0.7769 |
| Bus 10 | 0.2032 | 0.1584 |
| Bus 13 | 0.0283 | 0.0630 |
| Other Actions | 0 | $< 0.01$ |

assuming the attacker follows its NE via the $\sigma_A^*$ probability distribution.

The assessment of FDIA detection rate should involve evaluating their performance against intelligent attacks. Note that none of the conducted FDIAs can be detected without the implementation of a defensive measure, such as the addition of an extra PMU, as outlined in this paper.

Using the obtained NE solution (in Table I and II), the probability of detecting FDIA has been calculated and presented in Table III in comparison with the naive defense described above. Adding the PMU based on the proposed method in this paper results in $40.75\%$ and $62.50\%$ detection rates in systems without and with ZIB, respectively. However, adding the PMU with the naive process with the uniformly distributed probability for candidate buses as defender's action results in $25.90\%$ and $23.81\%$ detection rates. So, the benefit of following the proposed algorithm is $14.85\%$ and $36.69\%$ improvements in detection rates with the same number (one) of additional PMU.

TABLE III
ROBUST NE DEFENSE STRATEGY AND A NAIVE DEFENSE AGAINST A STRATEGIC ATTACKER.

| | Defense type | Without ZIB | With ZIB |
|---|---|---|---|
| FDIA detection rate (%) | Naive | 25.90 | 23.81 |
| | Robust NE | 40.75 | 62.50 |

## V. CONCLUSIONS

In this paper, after the first optimal PMU placement stage, a two-player zero-sum non-cooperative game is introduced to find a robust defense solution against FDIA by including a single additional PMU. The two players (attacker and defender) have opposite objectives, and neither side has complete information about the game (e.g., the opponent's actions). A reinforcement learning approach called "EXP3" is exploited to compute the robust Nash equilibrium solution. Our results show that the proposed method increases the rate of FDIA detection while being cost-efficient and robust to strategic attacks which encourages the expansion of this work to larger systems, which may require more than one PMU to reach a certain level of detection.

## REFERENCES


[1] M. Jamei, R. Ramakrishna, T. Tesfay, R. Gentz, C. Roberts, A. Scaglione, and S. Peisert, "Phasor measurement units optimal placement and performance limits for fault localization," *IEEE J. Sel. Areas Commun.*, vol. 38, no. 1, pp. 180–192, 2019.

[2] M. Elimam, Y. J. Isbeih, M. S. El Moursi, K. Elbassioni, and K. H. Al Hosani, "Novel optimal PMU placement approach based on the network parameters for enhanced system observability and wide area damping control capability," *IEEE Trans. Power Syst.*, vol. 36, no. 6, pp. 5345–5358, 2021.

[3] S. Mousavian, J. Valenzuela, and J. Wang, "A probabilistic risk mitigation model for cyber-attacks to PMU networks," *IEEE Trans. Power Syst.*, vol. 30, no. 1, pp. 156–165, 2014.

[4] P. K. Ghosh, S. Chatterjee, and B. K. Saha Roy, "Optimal PMU placement solution: graph theory and MCDM-based approach," *IET Generation, Transmission & Distribution*, vol. 11, no. 13, pp. 3371–3380, 2017.

[5] M. Zhang, Z. Wu, J. Yan, R. Lu, and X. Guan, "Attack-resilient optimal PMU placement via reinforcement learning guided tree search in smart grids," *IEEE Trans. Inf. Forensics Security*, vol. 17, pp. 1919–1929, 2022.

[6] C. Pei, Y. Xiao, W. Liang, and X. Han, "PMU placement protection against coordinated false data injection attacks in smart grid," *IEEE Trans. Ind. Appl.*, vol. 56, no. 4, pp. 4381–4393, 2020.

[7] Q. Yang, D. An, R. Min, W. Yu, X. Yang, and W. Zhao, "On optimal PMU placement-based defense against data integrity attacks in smart grid," *IEEE Trans. Inf. Forensics Security*, vol. 12, no. 7, pp. 1735–1750, 2017.

[8] H. Badrsimaei, R.-A. Hooshmand, and S. Nobakhtian, "Observable placement of phasor measurement units for defense against data integrity attacks in real time power markets," *Reliability Engineering & System Safety*, vol. 230, p. 108957, 2023.

[9] Q. Yang, L. Jiang, W. Hao, B. Zhou, P. Yang, and Z. Lv, "PMU placement in electric transmission networks for reliable state estimation against false data injection attacks," *IEEE Internet Things J.*, vol. 4, no. 6, pp. 1978–1986, 2017.

[10] S. Ghosh, M. S. Venkatraman, S. Ahmed, and C. Konstantinou, "A bi-level stochastic game model for PMU placement in power grid with cybersecurity risks," in *2023 IEEE Belgrade PowerTech*. IEEE, 2023, pp. 1–6.

[11] S. Lakshminarayana, E. V. Belmega, and H. V. Poor, "Moving-target defense against cyber-physical attacks in power grids via game theory," *IEEE Trans. Smart Grid*, vol. 12, no. 6, pp. 5244–5257, 2021.

[12] X. Chen, F. Wei, S. Cao, C. B. Soh, and K. J. Tseng, "PMU placement for measurement redundancy distribution considering zero injection bus and contingencies," *IEEE Syst. J.*, vol. 14, no. 4, pp. 5396–5406, 2020.

[13] D. Fudenberg and J. Tirole, *Game theory*. MIT press, 1991.

[14] T. Roughgarden, "Algorithmic game theory," *Communications of the ACM*, vol. 53, no. 7, pp. 78–86, 2010.

[15] S. Hart and A. Mas-Colell, "A simple adaptive procedure leading to correlated equilibrium," *Econometrica*, vol. 68, no. 5, pp. 1127–1150, 2000.

[16] A. Lazaric. (2017) Learning in zero-sum games. [Online]. Available: http://chercheurs.lille.inria.fr/∼lazaric/Webpage/MVA-RL_Course17_files/regret_games.pdf