Adversarial Classification of the Attacks on Smart Grids Using Game Theory and Deep Learning

Kian Hamedani
ECE Department of Virginia Tech, Blacksburg, VA, USA.
Marconi-Rosenblatt AI/ML Innovation Lab,
ANDRO Computational Solutions LLC, Rome, NY, USA
hkian@vt.edu

Lingjia Liu
ECE Department of Virginia Tech
Blacksburg, VA, USA
ljliu@vt.edu

Jithin Jagannath
Marconi-Rosenblatt AI/ML Innovation Lab,
ANDRO Computational Solutions LLC, Rome, NY, USA
jjagannath@androcs.com

Yang (Cindy) Yi
ECE Department of Virginia Tech
Blacksburg, VA, USA
yangyi8@vt.edu

ABSTRACT

Smart grids are vulnerable to cyber-attacks. This paper proposes a game-theoretic approach to evaluate the variations caused by an attacker on the power measurements. Adversaries can gain financial benefits through the manipulation of the meters of smart grids. On the other hand, there is a defender that tries to maintain the accuracy of the meters. A zero-sum game is used to model the interactions between the attacker and defender. In this paper, two different defenders are used and the effectiveness of each defender in different scenarios is evaluated. Multi-layer perceptrons (MLPs) and traditional state estimators are the two defenders that are studied in this paper. The utility of the defender is also investigated in adversary-aware and adversary-unaware situations. Our simulations suggest that the utility which is gained by the adversary drops significantly when the MLP is used as the defender. It will be shown that the utility of the defender is variant in different scenarios, based on the defender that is being used. In the end, we will show that this zero-sum game does not yield a pure strategy, and the mixed strategy of the game is calculated.

CCS CONCEPTS

• **Security and Privacy** → **Intrusion/anomaly detection**; • **Deep Learning** → **Adversarial examples**.

1 INTRODUCTION

Smart grids are new infrastructure that integrates energy by many different technologies, such as telecommunication, the internet, and electronic devices in this era of Internet-of-Things (IoT). This convergence of different technologies brings up some opportunities and challenges as well. The main opportunity that smart grids provide is a bidirectional flow of electricity and information between power suppliers and customers, which will result in a more efficient distribution of power. However, due to the integration of different technologies, the smart grids are more vulnerable to cyber-attacks [11]. False data injection (FDI) attacks are known to be one of the most malicious cyber-attacks in smart grids[8]. Cyber security tries to maintain reliable and secure communication between different components of the network, including communication networks and computer systems[7]. As a result of this secure communication, supervisory control and data acquisition (SCADA) system can have a better estimation of the network state. State estimation is a critical process in a control system and due to this fact, the SCADA is usually a target for attackers. Injecting false data to the SCADA manipulates the state estimation and it can cause economic gains for the attacker[7].

Game theory has shown to be a powerful approach to model and capture the complex interactions among the different players of electricity markets, which in some cases have a conflict of interests together[14]. During the last decades, game theory has been applied in many different fields, including economics, politics, and psychology. Due to the demand of having intelligent, autonomous, and flexible networks, in which devices can make rational decisions, game theory has also been applied in wireless communication networks and smart grids[6, 13–15]. Different devices or software can be the players in the network security domain. These players make their decisions independently and can be cooperative, non-cooperative, or even malicious towards each other[9]. As a result of this interaction among the players, the rational strategic decisions that they affect the network security and can be modeled as a game[9]. However, it is essential to consider the role of bad data detectors (BDD) while analyzing the strategic decisions and interactions among the players. The game theory can only study the interactions between the attackers and defenders, while the machine learning methods are being vastly used as the BDD. Therefore,

it is necessary to incorporate the machine learning-based BDDs in the game-theoretical formulation of the smart grids.

Machine learning-based approaches have also been widely studied for several applications in the realm of IoT [5]. Similarly, machine learning has been applied for FDI detection in smart grids [11]. Artificial neural networks, support vector machines (SVM), and k-nearest neighbor (KNN) are few examples of different machine learning algorithms that have been used for this purpose [11]. Machine learning-based approaches have shown better performance than the traditional state vector estimation (SVE) in detecting the FDI in smart grids. Esmalifalak *et al* [2] have applied dimension reduction for mapping the data collected from the network and used both supervised and unsupervised machine learning algorithms to detect the stealth data injections in smart grids. In [16], the performance of SVM, and KNN in detecting the false data in an IEEE-30 bus system under balanced and imbalanced data scenarios are studied. Mohammadpourfard *et al* [10] proposed an unsupervised anomaly detection method in smart grids which considers the effect of wind power generation and topology configuration. Zhao *et al* [17] have proposed a method based on short-term state forecasting which considers a temporal correlation among measurements. Moslemi *et al* [12], proposed an approach based on maximum likelihood (ML) which is decentralized, and the near chordal sparsity of smart grids is considered to detect the FDI. Esmailfalak *et al* [3] proposed a game theory-based approach to study the utility of the attacker and defender, while the defender is using a traditional state estimator as the BDD.

In this paper, we use a zero-sum game theory approach to quantify the utility gain of the attacker. We will show that different choices of the defender can significantly affect the gain of the attacker. In one scenario, the traditional state estimator, and in the other scenario an MLP is selected as the defender. We assume that the attacker and the defender are not able to attack and defend all the meters, andcompetition exists between attacker and defender to maximize and minimize the variations caused in smart meters due to false data injection respectively. The attacker can gain higher financial utility if the false variations of the measurements read by the meters are maximized. We will demonstrate that when the defender is chosen to be based on an MLP, the utility that the attacker can gain is much less than the case the defender is a traditional state estimator. The utility of the defender is also studied in different scenarios. The utility of the defender is defined as the number of the compromised measurements that the defender can accurately detect. Adversary-aware and adversary-unaware are the two main scenarios that are used to study the utility of the defender. In the adversary-aware situation, the attacker is aware of the existence of the defender and tries to maximize its attack utility concerning the parameters of the defender. On the other hand, in an adversary-unaware scenario, the attacker is unaware that there exists a defender and it optimizes its attack regardless of the defender parameters. We will show that the utility of the defender is different in these two situations. So far in the literature, only state estimators have been used as the BDD in the game-theoretic formulations of smart grids security and to the best of our knowledge, this is the first time that MLPs have been used as the BDD in the game-theoretic approach of studying smart grids security.

The organization of this paper is as follows: Section II represents the problem formulation. In Section III, the simulation results are discussed. Section IV represents the conclusion.

2 PROBLEM FORMULATION

2.1 System Model

In direct current (DC) power flow transmission, a linear equation is used to approximate the power flow[3]:

$$P_{ij} = \frac{x_i - x_j}{R_{ij}}, \tag{1}$$

where P_{ij} is the power flowing between bus i and j, x is the bus voltage phase angle, and R_{ij} is the reactance of transmission line between buses i and j[3]. The control center observes the measurement vector $\mathbf{z}$ [4]:

$$\mathbf{z} = \mathbf{Hx} + \mathbf{e}, \tag{2}$$

where $\mathbf{z}$ is the measurements vector, $\mathbf{H}$ is the Jacobian matrix, $\mathbf{x}$ is the state vector, and $\mathbf{e}$ is the environment noise. The attacker can compromise the state estimation by injecting false data, $\mathbf{a}$. The attack formulation can be expressed as follows:

$$\mathbf{z} = \mathbf{Hx} + \mathbf{a} + \mathbf{e}. \tag{3}$$

The estimated state vector can be calculated as:

$$\hat{\mathbf{x}} = (\mathbf{H}^T \wedge^{-1} \mathbf{H})^{-1} \mathbf{H}^T \wedge^{-1} \mathbf{z} = \mathbf{Mz}, \tag{4}$$

where $\mathbf{M} = (\mathbf{H}^T \wedge^{-1} \mathbf{H})^{-1} \mathbf{H}^T \wedge^{-1}$, and $\wedge$ is the noise covariance matrix.

2.2 Attack Against State Estimation

The real time pricing can be affected if the adversaries can successfully launch attacks on the results of the real time state estimation. The main goal of the attacker is to perform the attack in a way that it will not be detected by the BDD. The first thing that the attacker needs to know, is the group of measurements that can increase or decrease the congestion in the transmission line after the false data is injected. It can be seen in Eq. 1 that any changes in the state (voltage phase angle) of any bus can change the power measurements of transmission lines. Combining Eq. 1 and 4, gives us the estimated power as follows[3]:

$$\hat{\mathbf{P}}_{ij} = \frac{\hat{\mathbf{x}}_i - \hat{\mathbf{x}}_j}{\mathbf{R}_{ij}} = \frac{(\mathbf{M}_i - \mathbf{M}_j)^T}{\mathbf{R}_{ij}} \mathbf{z}$$
$$= \mathbf{G}^T \mathbf{z} = \mathbf{G}_+^T \mathbf{z}_+ + \mathbf{G}_-^T \mathbf{z}_-, \tag{5}$$

where $\mathbf{G}^T = \frac{(\mathbf{M}_i - \mathbf{M}_j)^T}{\mathbf{R}_{ij}}$. The group of positive and negative measurements where increasing or decreasing their values by adding a false data, $\mathbf{z}_a > 0$, leads to increasing or decreasing the congestion in transmission line are shown with $\mathbf{z}_+$ and $\mathbf{z}_-$, respectively. In this paper we assume that $\mathbf{z}_+$ and $\mathbf{z}_-$ belong to groups K and L, respectively. It is also assumed in this paper that the attacker knows $\mathbf{H}$, and as a result, it can make a distinction between groups K and L. Adversaries try to perform their attacks in a way that cannot be detected by the BDD. In the case that BDD is chosen as the traditional state estimator, the residual value between the actual measurements and the estimated values, ρ, has to be less

than a certain threshold value, ζ, that the attack cannot be detected by the BDD. The residual value is calculated as follows:

$$\rho = \mathbf{z} - \mathbf{H}\hat{\mathbf{x}} = \mathbf{z}_0 + \mathbf{z}_a - \mathbf{H}(\mathbf{M}\mathbf{z}_0 + \mathbf{M}\mathbf{z}_a)$$
$$= \mathbf{z}_0 - \mathbf{H}\mathbf{M}\mathbf{z}_0 + \mathbf{z}_a - \mathbf{H}\mathbf{M}\mathbf{z}_a = \rho_0 + \rho_a \tag{6}$$

where $\mathbf{z}_0$ is the safe measurement vector, $\mathbf{z}_a$ is the measurement vector which is compromised by the false data injected by the adversary, $\rho_0 = \mathbf{z}_0 - \mathbf{H}\mathbf{M}\mathbf{z}_0$, and $\rho_a = \mathbf{z}_a - \mathbf{H}\mathbf{M}\mathbf{z}_a$. It can be seen in Eq. 6 that ρ_a corresponds to the residual value caused by the false data injected by the adversary[3]. In order to perform the attack such that it cannot be detected, ρ_a has to be smaller than a certain threshold value,

$$\rho_a = ||(\mathbf{I} - \mathbf{H}\mathbf{M})\mathbf{z}_a|| \leq \zeta. \tag{7}$$

This threshold value, ζ, is a design parameter that has to be specified by the attacker. The smaller values of ζ make it more challenging for the BDD to detect the attack. However, the small values of ζ will provide less gain for the attacker, because the attacker faces more strict constraints to manipulate the measurements. Adversaries have to solve an optimization problem to be able to perform the maximum manipulation on the measurements without being detected by the BDD. This optimization problem is expressed in Eq. 8[3].

$$\max_{\mathbf{z}_a} \quad \sum_{i \epsilon K} \mathbf{z}_a(i) - \sum_{j \epsilon L} \mathbf{z}_a(j)$$
$$s.t. \begin{cases} ||(\mathbf{I} - \mathbf{H}\mathbf{M})\mathbf{z}_a|| \leq \zeta \\ \mathbf{z}_a(k) = 0 \quad k \, \epsilon \, \{NK\}, \end{cases} \tag{8}$$

where $\{NK\}$ is the set of measurements that the attacker is not able to compromise. The utility of the attacker corresponds to values of $\mathbf{z}_a$ that are achieved from Eq. 8. It means that for the higher values of $\mathbf{z}_a$, the measurements are manipulated more significantly and the attacker can gain higher utility. Online pricing of the power depends on the estimated power of the smart meters. Therefore, the attacker can gain some financial utility through injecting false data and changing the actual estimations of the measurements. The utility of the attacker can be written as follows:

$$\Delta\hat{\mathbf{P}}_{ij} = \frac{(\mathbf{M}_i - \mathbf{M}_j)^T}{\mathbf{R}_{ij}} \mathbf{z}_a. \tag{9}$$

2.3 Attack Against Artificial Neural Networks

In Section II B, we assumed that the BDD is a closed- form state estimator. However, it is possible to use other tools as the BDD. In this section, we propose the idea to use deep learning (MLP) instead of a state estimator. As it was mentioned in Sections II A and B, the attacker injects the false data in a way that cannot be detected by the BDD. In the case of using MLP as the BDD, the attacker has to perform its attack in a way that cannot be detected by the MLP. It means that the attacker needs to solve the optimization problem expressed in Eq. 8. However, the constraints of that optimization problem have to be modified with respect to the parameters of the MLP. In Section II B, it was assumed that the attacker knows $\mathbf{H}$, in the case that MLP is used as the BDD we will assume that the attacker knows the parameters of the MLP. In Fig. 1, the structure of a one-layer MLP is depicted. There are two sets of weights in this network, w_{ij}, which connects the input to the first hidden layer, and

y_{ij} that connects the outcome of the hidden layer to the output layer. There are also two activation functions f and $f1$. In this scenario, an MLP has to be trained with two sets of data, compromised data and safe data. For each set of data, a label is assigned to train the MLP, the label 1 is assigned for the compromised measurements, and 0 for safe data. The adversary has to perform its attack somehow that when the compromised data is tested by the MLP, the output of MLP will be closer to zero than one. In this paper, we will use an MLP with two hidden layers and 10000 total samples are used for training the MLP, half of the samples are compromised measurements and the other half are safe measurements. We assume that when the output of MLP is larger than 0.5, the measurement will be classified as the compromised data, and when it is less than 0.5, it will be considered as a safe measurement. Therefore, Eq. 8 is expressed as follows when the BDD is an MLP:

$$\max_{\mathbf{z}_a} \quad \sum_{i \epsilon K} \mathbf{z}_a(i) - \sum_{j \epsilon L} \mathbf{z}_a(j)$$
$$s.t. \begin{cases} f1(y_{11}f(z(1)w_{11} + ... + z(N)w_{N1}) \\ +y_{21}f(z(1)w_{12} + ... + z(N)w_{N2}) \\ +... \\ +y_{M1}f(z(1)w_{1N} + ... + z(N)w_{NM})) \leq 0.5 \\ z_a(k) = 0 \quad k \, \epsilon \, \{NK\}, \end{cases} \tag{10}$$

where $z = z_0 + z_a$, N is the number of measurements, and M is the number of neurons in the hidden layer. This formulation can be extended to any number of neurons and hidden layers.
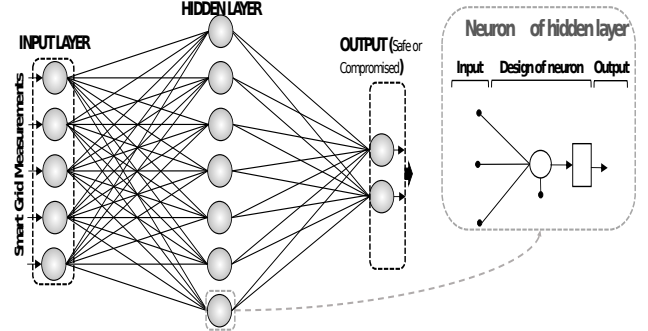

Figure 1: Structure of a MLP with one hidden layer.

2.4 Zero-Sum Game Between Two Players

In a two-person zero-sum game there are two players that each side tries to increase its gain by decreasing the other side's gain. The adversary and the defender can also be considered as two players of a game. Each player has a set of strategies, which in our case is defined as the measurements that can be attacked and defended. Based on the strategy set, for any action taken by the attacker or defender, a utility value is attributed. In this paper, $\mathbf{S} = \{s_{ij,i=1,...,r;j=1,...,c}\}$ is denoted as the game matrix. In the game matrix, the set of actions and the corresponding utilities which are taken by each player is shown. Here the utility of the attacker, U_a, corresponds to $\Delta\hat{P}_{ij}$ and the utility of defender, U_D, is equal to $-\Delta\hat{P}_{ij}$. The saddle pint or the pure strategy of this game is defined as $s_{i^*j^*}$, and it exists if and only if, $\min_{row}(\max)$ of S=$\max_{col}(\min)$ of S

where *row* and *col* correspond to the rows and columns of the **S**. In the games that a pure strategy does not exist, a mixed strategy of the game has to be calculated. The mixed strategy tells us about the probability of taking each action by the players[3]. The frequency of choosing the rows and columns of the game by the defender and the attacker will converge to the probability distribution of the defender and attacker respectively. To calculate the mixed strategy of a zero-sum game, the defender tries to minimize the average value of the outcome of the game which is defined as below:

$$min \ C(q, u) = \sum_{i=1}^{r} \sum_{j=1}^{c} q_i' s_{ij} u_j, \quad (11)$$

where q and u are the probability distributions of the strategies taken by the defender and attacker respectively. The defender's mixed strategy is calculated via solving the following optimization problem which is a linear programming problem:

$$\max_{\tilde{q}} \ \tilde{q}' \ 1_r$$
$$s.t. \begin{cases} S' \tilde{q} \leq 1_c \\ \tilde{q} \geq 0, \end{cases} \quad (12)$$

where $\tilde{q} = \dfrac{q}{h_1(q)}$, and $h_1(q)$ is defined as follows:

$$h_1(q) = \max_{U} \ q' S \ u \ \geq \ q' S \ u \quad \forall \ u \ in \ U. \quad (13)$$

In Eq. 13, U is defined as the probability distribution of the strategies chosen by the attacker and is defined as below:

$$U = \left\{ u \in R^c : u \geq 0, \sum_{j=1}^{c} u_j = 1 \right\} \quad (14)$$

In order to calculate the mixed strategy of the attacker another linear programming problem has to be solved which is expressed as:

$$\min_{\tilde{u}} \ \tilde{u}' \ 1_c$$
$$s.t.$$
$$\begin{cases} S\tilde{u} \geq 1_r \\ \tilde{q}u \geq 0, \end{cases} \quad (15)$$

where $\tilde{u} = \dfrac{u}{h_2(u)}$, u_i' is the transpose of u, and $h_2(u)$ is defined as follows:

$$h_2(u) = \min_{Q} \ q' S \ u \ \leq \ q' S \ u \quad \forall \ q \ in \ Q. \quad (16)$$

In Eq. 16, Q is defined as the probability distribution of the strategies chosen by the defender and is expressed as below:

$$Q = \left\{ q \in R^r : q \geq 0, \sum_{i=1}^{r} q_i = 1 \right\} \quad (17)$$

3 PERFORMANCE ANALYSIS

In this paper we use MATPOWER 5.1 [18] to simulate a PJM-5 bus test system which gives us **H**, and **x**. To have a fair comparison between the MLP and the state estimator, the threshold value of the state estimator BDD, ζ, has to be chosen somehow that the false alarm rate of the MLP and the state estimator be equal. In Table 1 the ζ thresholds of the state estimator are shown when the attacker can

perform its attack on one or two meters at the same time depending on which meters are being defended and which meters are being attacked in the zero-sum game. In the PJM-5 testbed, there are total of 11 smart meters. However, as it was mentioned in Section II we assume that the adversary is not able to perform an attack on all of them. Adversary performs its attack on the meters that are more

Table 1: Threshold values of state estimator BDD in mega Watts(MW).

Compromised meter	ζ
z_1	8.54
z_4	7.5
z_5	8.32
z_{10}	8.3
$z_1 z_4$	10.89
$z_1 z_5$	12.65
$z_1 z_{10}$	12.59
$z_4 z_5$	10.48
$z_4 z_{10}$	10.33
$z_5 z_{10}$	9.48

likely to cause congestion on transmission lines. These meters are identified through DC optimal power flow (DCOPF) simulation and are shown in Table I. In Tables II and III, the utility of the attacker, in the zero-sum game between the attacker and the defender is shown. As it was mentioned in Section II D, in the zero-sum game the utility of the attacker, U_a, corresponds to $\Delta \hat{P}_{ij}$ and the utility of defender, U_D, is equal to $-\Delta \hat{P}_{ij}$. The columns of Tables II and III correspond to the meters that are attacked, and the rows correspond to the meters that are defended.

As it can be seen in Tables II and III, the utility of the attacker when the defender is using the MLP as the BDD is much lower than the case that the defender uses state estimator. The average utility of the attacker in the zero-sum game when the defender uses MLP is 0.87 MW and when it uses the state estimator, it is equal to 4.75 MW, which indicates that the MLP is a much better BDD for the defender since the attacker can gain much less utility.

**Table 2: Utility of Attacker when BDD is state estimator(MW)**

DefAtt	z1z4	z1z5	z1z10	z4z5	z4z10	z5z10
z1z4	0	6.03	5.83	6.03	5.83	11.82
z1z5	2.79	0	5.83	2.79	8.34	5.83
z1z10	2.79	6.03	0	8.6	2.79	6.03
z4z5	4.45	4.45	10.18	0	5.83	5.83
z4z10	4.45	10.44	4.45	6.03	0	6.03
z5z10	7.71	4.45	4.45	2.79	2.79	0

3.1 Adversary-aware VS Adversary-unaware classification

In adversarial attacks machine learning, the adversary tries to gain the maximum benefit through fooling the classifier. In this area, there exists a game or competition between the adversary and the

Table 3: Utility of Attacker when BDD is MLP(MW)

DefAtt	z1z4	z1z5	z1z10	z4z5	z4z10	z5z10
z1z4	0	0.77	1.5	0.77	1.5	2.33
z1z5	0.86	0	1.5	0.86	1.62	1.5
z1z10	0.86	0.77	0	1.83	0.86	0.77
z4z5	0.28	0.28	1.28	0	1.5	1.5
z4z10	0.28	0.78	0.28	0.77	0	0.77
z5z10	0.77	0.28	0.28	0.86	0.86	0
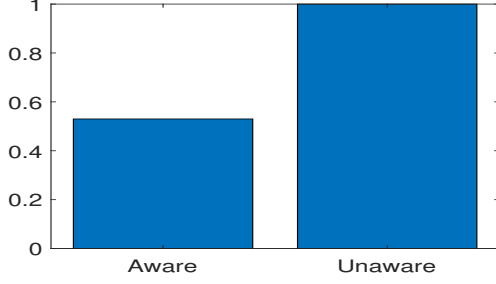
Figure 2: Detection probability of MLP.
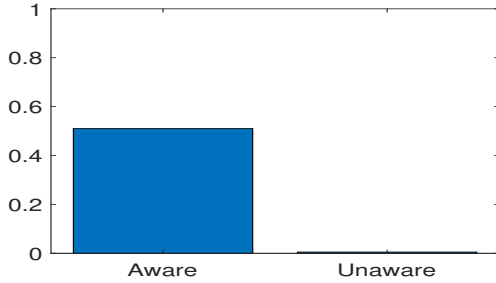
Figure 3: Detection probability of state estimator.

classifier. Two main scenarios are considered in this situation. In the first scenario, it is assumed that the adversary is not aware of the presence of the BDD or the classifier, while in the second scenario it is assumed that the adversary is aware of that. The adversary-aware scenario requires the adversary to optimize its attack with respect to the parameters of the classifier, and the adversary-unaware requires the attacker to optimize its attack regardless of the classifier parameters[1]. In this part, we first assume that the classifier is an MLP and then we assume that the classifier or the BDD is the state estimator. For both of the cases, the adversary-aware and unaware scenarios are implemented and the normalized utility of the classifiers are shown in Figs. 2 & 3. The utility of the classifier is defined as the number of the compromised meters that it can detect.

As it can be seen in Figs. 2 & 3, when the defender is using the MLP as the BDD, in the adversary-unaware case the utility or accuracy of the attack detection is higher compared with the adversary-aware case. However, when the defender is using the state estimator as the BDD, the opposite behavior is observed. The reason for this difference in the behavior is that when the BDD is MLP and the attacker optimizes its attacks with respect to the state estimator, the magnitude of the attack will be much higher which

makes it very easy for the MLP to detect the attack. On the other hand, when the BDD is using the state estimator, and the attacker is optimizing its attack against the MLP the magnitude of the attack will be very low and the state estimator can't detect the attack. That is why in the adversary-unaware scenario, the utility of the classifier when MLP is being used is much higher than the case that the state estimator is being used. However, in the adversary-aware scenario, the utility of the classifier is almost equal in both cases, because the threshold value, ζ, of the state estimator is chosen such that both the classifiers have the same value of false alarm rate.

3.2 Mixed Strategy of the Game

As it can be seen in Tables II and III, this game does not have a pure strategy because $min(\max)$ of S $\neq max(\min)$ of S. As it can be seen in Table II, the $\underset{row}{min}(\max) = 7.71$ and $\underset{col}{max}(\min) = 0$, also in Table III $\underset{row}{min}(\max) = 0.76$ and $\underset{col}{max}(\min) = 0$. Since none of these two scenarios have a pure strategy, the mixed strategy of these two games have to be calculated. Based on the Tables II and III, and using Eq. 12, the following optimization problems have to be solved to calculate the defenders' mixed strategies for the state estimator and the MLP, respectively. By solving Eq. 18, $\tilde{q}$ can be calculated for the defender when the BDD of the defender is the state estimator. The mixed strategy of the defender when the state estimator is being used as the BDD can be calculated using Eq. 19.

$$\max_{\tilde{q}} \tilde{q}' 1_r$$

$$s.t. \begin{cases} 2.79\tilde{q}_2 + 2.79\tilde{q}_3 + 4.45\tilde{q}_4 + 4.45\tilde{q}_5 + 7.71\tilde{q}_6 \leq 1 \\ 6.03\tilde{q}_1 + 6.03\tilde{q}_3 + 4.45\tilde{q}_4 + 10.44\tilde{q}_5 + 4.45\tilde{q}_6 \leq 1 \\ 5.83\tilde{q}_1 + 5.83\tilde{q}_2 + 10.18\tilde{q}_4 + 4.45\tilde{q}_5 + 4.45\tilde{q}_6 \leq 1 \\ 6.03\tilde{q}_1 + 2.79\tilde{q}_2 + 8.6\tilde{q}_3 + 6.03\tilde{q}_5 + 2.79\tilde{q}_6 \leq 1 \\ 5.83\tilde{q}_1 + 8.34\tilde{q}_2 + 2.79\tilde{q}_3 + 5.83\tilde{q}_4 + 2.79\tilde{q}_6 \leq 1 \\ 11.82\tilde{q}_1 + 5.83\tilde{q}_2 + 6.03\tilde{q}_3 + 5.83\tilde{q}_4 + 6.03\tilde{q}_5 \leq 1 \\ \tilde{q} \geq 0. \end{cases} \quad (18)$$

$$q = \tilde{q} \, h_1(q) = \tilde{q}(\tilde{q}1_r)^{-1} \quad (19)$$

The mixed strategy of the defender when MLP is being used as the BDD can be calculated using Eq. 20. Upon solving Eq. 20, the mixed strategy of the defender when MLP is being used as the BDD can be calculated using Eq. 19. Figs. 4 & 5 demonstrate the mixed strategies of these two games. As it can be seen in Fig. 4, the attacker performs 57% of its attacks on $z4z5$, and this transmission line is defended by the defender in 21% of the times. To find the mixed strategy of the attacker, Eq. 15 has to be solved using the values in Tables II and III respectively.

$$\max_{\tilde{q}} \tilde{q}' 1_r$$

$$s.t. \begin{cases} 0.86\tilde{q}_2 + 0.86\tilde{q}_3 + 0.28\tilde{q}_4 + 0.28\tilde{q}_5 + 0.77\tilde{q}_6 \leq 1 \\ 0.77\tilde{q}_1 + 0.77\tilde{q}_3 + 0.28\tilde{q}_4 + 0.78\tilde{q}_5 + 0.28\tilde{q}_6 \leq 1 \\ 1.5\tilde{q}_1 + 1.5\tilde{q}_2 + 1.28\tilde{q}_4 + 0.28\tilde{q}_5 + 0.28\tilde{q}_6 \leq 1 \\ 0.77\tilde{q}_1 + 0.86\tilde{q}_2 + 1.83\tilde{q}_3 + 0.77\tilde{q}_5 + 0.86\tilde{q}_6 \leq 1 \\ 1.5\tilde{q}_1 + 1.62\tilde{q}_2 + 0.86\tilde{q}_3 + 1.5\tilde{q}_4 + 0.86\tilde{q}_6 \leq 1 \\ 2.33\tilde{q}_1 + 1.5\tilde{q}_2 + 0.77\tilde{q}_3 + 1.5\tilde{q}_4 + 0.77\tilde{q}_5 \leq 1 \\ \tilde{q} \geq 0. \end{cases} \quad (20)$$
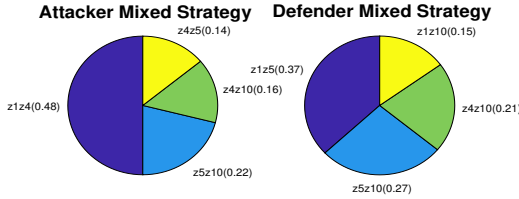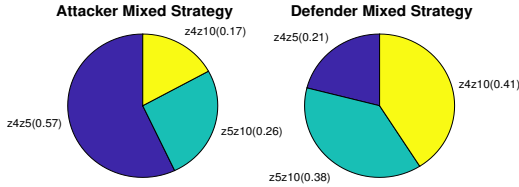
Figure 4: Mixed Strategy of Table II.

Figure 5: Mixed Strategy of Table III.

The main intuition that can be observed from Figs. 4 & 5 is that the probability of the attack is higher on the measurements that are less likely to be protected by the defender. For each set of measurement,s there is a corresponding value in G. The higher values of G correspond to the measurements that have a greater effect on $\Delta \hat{\mathbf{P}}_{ij}$. The adversary finds the measurements with greater value in G and less probability of detection to compromise. The other observation from Figs. 4 & 5 is that, while the state estimator is being used as the BDD there is a larger number of measurements that can be compromised. As it can be seen, while the BDD is an MLP there are 3 sets of measurements that have a high probability to be compromised. However, as we use the state estimator as the BDD there are 4 sets of measurements which are likely to be compromised. This intuition means that using the MLP as the BDD not only decreases the gain that the adversary can achieve but also limits the set of possible actions that the adversary can choose to perform attacks.

This behavior can be explained using Tables II & III. As it can be seen in Table II, in the scenarios that the adversary can compromise two measurements at the same without being defended it always gains more utility. As an example, while the adversary compromises $z1z4$ and the defender protects $z5z10$ the utility of the attacker is equal to 7.71MW. This behavior can be observed in other columns of Table II as well. However, this behavior cannot be seen in Table III. In the scenario where the adversary attacks $z1z4$ and the defender protects $z5z10$, the utility that the adversary gains are equal to 0.77MW which is not greater than all of the utilities that the adversary can gain in the first column of Table I. This means that while we use the MLP as the BDD there is no guarantee for the adversary to gain higher utilities if it can successfully compromise two meters simultaneously.

4 CONCLUSION

In this paper, an approach has been proposed which for the first time combines the MLPs with the game theory to study the false data injection attacks in the smart grids. This paper proposes to use the MLPs instead of conventional state estimators as the BDD of the defender in the game-theoretic formulation of the smart grid security. We have shown that if the MLPs are used then the adversaries can gain much less utility than when the state estimators are used. We have also studied the adversary informed and adversary uninformed scenarios in the game-theoretic formulation of this problem, and we have shown that depending on the type of the defender's BDD in each scenario different behavior can be observed. We also showed that this game does not have a pure strategy, and the mixed strategy has been calculated.

REFERENCES

[1] N. Dalvi, P. Domingos, S. Sanghai, and D. Verma. 2004. Adversarial classification. *in Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining* 1 (Aug 2004), 99–108.

[2] M. Esmalifalak, L. Liu, N. Nguyen, R. Zheng, and Z. Han. 2017. Detecting stealthy false data injection using machine learning in smart grid. *IEEE Systs Journal* 11, 3 (Sep 2017), 1644–1652.

[3] Mohammad Esmalifalak, Ge Shi, Zhu Han, and Lingyang Song. 2013. Bad data injection attack and defense in electricity market using game theory study. *IEEE Transactions on Smart Grid* 4, 1 (2013), 160–169.

[4] K. Hamedani, L. Liu, R. Atat, J. Wu, and Y. Yi. 2018. Reservoir computing meets smart grids: attack detection using delayed feedback networks. *IEEE Transaction On Ind Informatics* 14, 2 (Feb 2018), 734–743.

[5] Jithin Jagannath, Nicholas Polosky, Anu Jagannath, Francesco Restuccia, and Tommaso Melodia. 2019. Machine Learning for Wireless Communications in the Internet of Things: A Comprehensive Survey. *Ad Hoc Networks (Elsevier)* 93 (2019), 101913.

[6] Qimeng Li, Pengpeng Lv, Mengqiang Wang, Zhizhi Zhang, Shaoyin Wang, Pengbo Fang, and Lifang Gao. 2020. A Risk Assessment Method of Smart Grid in Cloud Computing Environment Based on Game Theory. In *2020 IEEE 5th International Conference on Cloud Computing and Big Data Analytics (ICCCBDA)*. IEEE, 67–72.

[7] G. Liang, L. Zhao, F. Luo, S Weller, and Z.Dong. 2017. A review of false data injection attacks against modern power systems. *IEEE Transactions on Smart Grids* 8, 4 (Jul. 2017), 1630–1638.

[8] Y. Liu, P. Ning, and M. Reiter. 2011. False data injection attacks against state estimation in electric power grids. *ACM Transactions on Information and System Security (TISSEC)* 14, 1 (2011), 21–32.

[9] M. Manshaei, Q. Zhu, T. Alpcan, T. Bacşar, and J. Hubaux. 2013. Coalitional game theory for communication networks: A tutorial. *ACM Computing Surveys (CSUR)* 45, 3 (2013), 25–39.

[10] M. Mohammadpourfard, A. Sami, and Y. Weng. 2018. Identification of False Data Injection Attacks with Considering the Impact of Wind Generation and Topology Reconfigurations. *IEEE Transactions on Sustainable Energy* 9, 3 (Jul 2018), 1349–1364.

[11] M. Ozay, I. Esnaola, F. Vural, S. Kulkarni, and H. Poor. 2016. Machine learning methods for attack detection in the smart grid. *IEEE Transactions on Neural Nets and Learning Syts* 27, 48 (Aug. 2016), 1773–1786.

[12] A. Mesbahi R. Moslemi and J. Velni. 2018. A fast, decentralized covariance selection-based approach to detect cyber attacks in smart grids. *IEEE Transactions on Smart Grids* 6, 5 (Sep 2018), 14930–4941.

[13] Walid Saad, Zhu Han, Mérouane Debbah, Are Hjorungnes, and Tamer Basar. 2009. Coalitional game theory for communication networks. *Ieee signal processing magazine* 26, 5 (2009), 77–97.

[14] D. Srinivasan, S. Rajgarhia, B. Radhakrishnan, A. Sharma, and H. Khincha. 2017. Game-Theory based dynamic pricing strategies for demand side management in smart grids. *Energy* 8, 4 (July 2017), 21–32.

[15] Kuan Wang, Jun Wu, Xi Zheng, Alireza Jolfaei, Jianhua Li, and Dongjin Yu. 2020. Leveraging energy function virtualization with game theory for fault-tolerant smart grid. *IEEE Transactions on Industrial Informatics* 17, 1 (2020), 678–687.

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

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

[18] R. Zimmerman, C. Murillo-Sanchez, and R.Thomas. 2011. Matpower:Steady-state operations, planning, and analysis tools for power systems research and education. *IEEE Transactions on Power Systems* 26, 1 (Feb 2011), 12–19.