Multi-Objective Evolutionary Optimization for Worst-Case Analysis of False Data Injection Attacks in the Smart Grid

Moshfeka Rahman, Yuanliang Li and Jun Yan
Concordia Institute for Information and Systems Engineering, Concordia University
Montréal, QC H3G 1M8, Canada
r_moshfe@encs.concordia.ca; l_yuanli@encs.concordia.ca; jun.yan@concordia.ca

***Abstract*—False data injection attacks (FDIA) have drawn significant interests recently after the discovery of vulnerabilities of bad data detectors (BDD) deployed in the smart grid. While most FDIA analyses focused separately on the aspects of stealthiness, knowledge, resources, or expected consequences of the attack, few have evaluated the relationship and trade-offs among these factors to identify the worst-case scenario in realistic operations. To fill the gap, this paper investigates a strictly stealthy FDIA scheme with multi-objective evolutionary optimization, which could compromise a small set of meters to inflict large impacts on the smart grid in realistic scenarios. Compared with existing attack schemes that relax the problem with the ℓ_1-norm, the paper introduced the Improved Strength Pareto Evolutionary Algorithm (SPEA2) as the solver to directly obtain the ℓ_0-sparse attack vector. Meanwhile, the unobservability is ensured by not only bypassing the BDD but also satisfying the physical and operational constraints. A three-step constraint handling technique is also proposed for the SPEA2 to ensure the stealthiness and improve the efficiency of attack vector identification in the worst-case scenario. Simulation results on the IEEE 14-bus and 30-bus systems demonstrate that the new multi-objective formulation discovers highly sparse attack vectors with significant impacts on the system without triggering immediate emergency responses. The influence of alternative objectives and constraints has also been evaluated to reveal the trade-offs among the attack's stealthiness, sparsity, and impact. The results are expected to facilitate better-informed risk assessment and mitigation with refined worst-case understandings.**

***Index Terms*—Cyber-physical security, false data injection, multi-objective optimization, smart grid, worst-case analysis.**

I. Introduction

Recent vulnerability analysis of state estimation (SE) in safety-critical systems has shed light on a high-threat attack called the false data injection attack (FDIA), which injects errors into the estimated states such that the conventional bad data detectors (BDD) can not identify the manipulated measurements [1]. This attack, if gone undetected, can mislead decisions in automatic generation control, contingency analysis, and economic dispatch [2], which have raised both concerns and interests to fully investigate the threat.

The threat introduced by FDIA on PSSE relies on attackers' knowledge of power grid topology and resources to manipulate substantial measurements. The availability of knowledge and resources will enhance the attacker's capability to launch a successful FDIA. However, in reality, such knowledge and resources are mostly constrained. It has been demonstrated that the probability of a successful stealth attack vector decreases with the knowledge of the topology and the number of attacked measurements [3], [4].

To accurately analyze the risk of FDIA, it is important to investigate the minimal set of measurements required to launch the stealthy FDIA. The original work on FDIA [1] did not guarantee the finding of a feasible stealth attack vector while requiring substantial amount of resources. Therefore, the succeeding works explored the possibility of a guaranteed stealthy attack vector leveraging minimal resources. For example, relaxed ℓ_1-norm minimization of the attack vector has been developed with [3] and without [5]–[7] the full knowledge of grid topology. In addition, studies have also tried to identify the minimal attack vector via heuristic search for minimal attack vector [8], manipulation of the null space [4], and the locally regularized fast recursive (LRFR) algorithm [5]. Although these works have been successful in proposing ways to minimize the sparsity of the attacked measurements, the impact of the attacks remains unaddressed as they assumed a random attack vector, which may render the attack vectors insignificant in practice.

Meanwhile, the impact analysis of FDIA on power system in the existing literature has some limitations. The potential physical and cyber consequences have been exploited in variants of FDIA schemes, including the malicious load redistribution [9], [10], the intentional line overflow [11], and the manipulated localized marginal price [12], among others. The resilience analysis of the grid under the FDI attack in the context of the voltage violation, line outage, and cascading blackouts have also been explored in [13]. The impact of FDIA taking into account the environmental and economical uncertainties, namely the environmental/economic dispatch (EED), was formulated in [14], utilizing a robust evolutionary optimization algorithm (REOA). In the effort to maximize the impact, however, these works did not consider how the intensified attempts could also raise the chance of being detected by the defender or the the number of resources that were being leveraged by the attacker.

From attackers' perspective, launching the FDIA with fewer resources is beneficial if attackers can inflict desired impacts on the power system and remain stealthy for a long term. However, to the authors' knowledge, maximizing the attack impact with minimal resources while remaining

978-1-7281-6929-3/20/$31.00 ©2020 IEEE

stealthy remains a challenge to be addressed and this presents a novel multi-objective optimization problem to solve.

Motivated by these gaps in the existing literature, this work proposes a novel multi-objective optimization formulation of the FDIA, which simultaneously minimizes the number of compromised measurements and maximizes the impact of the attack vector while remaining completely stealthy. The first objective is achieved via the ℓ_0-norm minimization of the attack vector. For the second objective, we investigate two different expected impacts on the system based on the assumed motivation of the attacker. In the first case, we assume the attacker aims at direct manipulation of the states to mislead control room decisions, and the objective is thereby to maximize the error of estimated states injected by the attacker. In the second case, the attacker aims to report false branch flows to gain financial benefit without overloading the line, and the objective is to maximize the error of the branch power flow. Thus, two different multi-objective optimization problems are formulated in this work.

To solve the proposed multi-objective optimization formulations, the Improved Strength Pareto Evolutionary Algorithm (SPEA2) [15] is implemented. This evolutionary multi-objective algorithm efficiently demonstrates the trade-off between sparsity and impact and can solve the problem of direct minimization of ℓ_0-norm, which poses an NP-hard problem for the conventional optimization algorithms. Moreover, To make the FDIA meet the stealthiness requirements, the paper formulates constraints and proposes a three-step constraints handling method to improve the efficiency in the multi-objective optimization problems.

The contributions of this work are summarized as follows:

1) We propose a novel multi-objective optimization problem using the Improved Strength Pareto Evolutionary Algorithm (SPEA2) to obtain the ℓ_0-sparse attack vector with maximized impacts.
2) We consider the physical and operational constraints on the measurements in generating the attack vector, which will not only bypasses the BDD but also not raise violations in the control centre. To meet these stealthiness requirements, we apply a three-step constraints handling method.
3) We provide a comparative analysis between errors injected into the state vector and branch power flows as the potential impacts of the attack. Evaluations of the impacts considering the attackers' motivation are offered.
4) We also provide analysis on how the constraints influence the impact performance of the attack model, and were able to provide insights into the possible trade-off between stealthiness and impact.

The rest of the paper is organized as follows: Section II gives the introduction of the basic FDIA threat and the challenges to be addressed. Section III describes the formulation of the multi-objective FDIA scheme for the worst-case analysis and gives the details of specific methods to solve the proposed problems. Section IV presents the experiments and performance evaluation. Section V draws the conclusions.

II. The FDIA Threat

A. The System Model

The state estimator processes the topology and estimates accurate states of the system from raw measurements. In the DC state estimation, we can write the problem as, $\mathbf{Z} = \mathbf{Hs} + \mathbf{e}$, where $\mathbf{Z}$ is the measurement vector, $\mathbf{H}$ is the topological matrix, also known as the Jacobian matrix, $\mathbf{s}$ is the state vector, and $\mathbf{e}$ is the random noise often modeled as the white Gaussian noise. The weighted least square (WLS) solution can be obtained by $\hat{\mathbf{s}} = (\mathbf{H^T R^{-1} H})^{-1} \mathbf{H^T R^{-1} Z}$ [16], where $\mathbf{R}$ is the noise covariance matrix. If M and N are the numbers of measurements and states, respectively, then $\mathbf{H}$ is an $M \times N$ matrix with $M \gg N$.

In PSSE, the estimated state $\mathbf{s}$ is processed by the bad data detection (BDD), which detects and eliminates bad data that might be caused by meter noise, sensor failure or communication loss. A residual-based hypothesis test [16] is utilized to validate the normalized ℓ_2-norm of the residual between the observed and estimated states, that is $||\mathbf{Z} - \mathbf{H}\hat{\mathbf{s}}||$, via the χ^2-test. The test statistics indicates if a measurement is corrupted, which then will be eliminated for better estimation of the states. Specifically, $||\mathbf{Z} - \mathbf{H}\hat{\mathbf{s}}||$ is compared with a bad data threshold τ and $||\mathbf{Z} - \mathbf{H}\hat{\mathbf{s}}|| > \tau$ indicated the presence of bad data.

B. Stealthiness, Sparsity and Impact of FDIA Schemes

FDIA was first proposed by Liu *et al.* [1] and was demonstrated that the availability of the full knowledge of the topology $\mathbf{H}$ could allow the attacker to obtain the attack vector $\mathbf{a} = \mathbf{Hc}$, where $\mathbf{c}$ is the injected state error, to be undetectable by the residual-based BDD [1]. Mathematically,

$$\mathbf{z_a} = \mathbf{Hs} + \mathbf{a} + \mathbf{e} = \mathbf{Hs} + \mathbf{Hc} + \mathbf{e} = \mathbf{Hs_a} + \mathbf{e} \qquad (1)$$

The residual of the attacked measurements becomes $\mathbf{r_a} = \mathbf{z_a} - \mathbf{H}\hat{\mathbf{s_a}} = \mathbf{z} + \mathbf{a} - \mathbf{H}(\hat{\mathbf{s}} + \mathbf{c}) = \mathbf{z} - \mathbf{H}\hat{\mathbf{s}}$, which is the same as residual before the attack and hence the attack remains stealthy.

Stealthiness has been the primary goal of FDIA schemes [17]. From the attacker's perspective, however, the stealthiness of FDIA often requests the attack vector to be non-sparse [1], as more compromised measurements will make it easier to impair the situational awareness and bypass the BDD [3] [4]. In such case, the stealthy but non-sparse (i.e., large number of meters) attack may also inflict larger impact onto the system with more misleading information.

However, a non-sparse attack often requires over 60% of the measurements to be compromised [18]. As meters are geographically dispersed and individually protected, it could be challenging to compromise and manipulate such large number of measurements concurrently with limited resources. Nevertheless, minimizing the resources could also lead to limited impact. Therefore, a two-fold challenge of formulating an attack model, that simultaneously considers the stealthiness, sparsity and impact, should be addressed in the multi-objective FDIA scheme.

III. Multi-Objective FDIA Scheme

Multi-objective Optimization Problem (MOP) involves more than one objective function to be optimized simultaneously. Generally, a MOP can be formulated as [19]:

$$\text{Minimize} \quad \mathbf{F}(\mathbf{c}) = [f_1(\mathbf{c}), f_2(\mathbf{c}), ..., f_m(\mathbf{c})] \tag{2}$$

$$\text{subject to} \quad \begin{cases} h_i(\mathbf{c}) = 0 & i = 1, ..., p \\ g_i(\mathbf{c}) \leq 0 & i = 1, ..., q \\ c_i^{min} \leq c_i \leq c_i^{max} & i = 1, ..., d \end{cases} \tag{3}$$

where $\mathbf{c}$ is the d-dimensional decision vector, c_i is the ith decision variable; $\mathbf{F}$ is the objective vector consisting of m objectives, which are competing and conflicting with each other, $f_i(\mathbf{c})$ is the ith objective; p is the number of equality constraints, and $h_i(\mathbf{c}) = 0$ is the ith equality constraint; q is the number of inequality constraints, $g_i(\mathbf{c}) \leq 0$ is the ith inequality constraint; c_i^{min} and c_i^{max} are lower and upper bound of c_i, respectively.

The stealthiness requirements in this work are formulated as different constraints. We consider not only the attack model in section II.B [1], but also two practical power grid constraints. First, the manipulated measurements do not include generation and the load demand measurements of the zero-load buses, as any change on these measurements can be easily flagged as anomalies based on routine checks. In other words, the additional false power injection will only be added into the measurement of the load buses. Second, we only consider the case where the attacker does not trigger violations of voltage, power flow, etc. in the control room, as the violations will still activate reactions even if they are not recognized as intrusions.

Furthermore, we consider two different scenarios based on the impact that an attacker may intend to inflict on the system. One is the error injected into the state vector, which the estimated voltage angle for DC SE. The other one is the error injected into the branch power flow. The detailed descriptions of these two scenarios are discussed in the following subsections.

A. Scenario I: Injected Error into State Vector

Based on the DC assumption of PSSE, suppose the original measurements consist of all branch flow measurements at the from end ($\mathbf{P_f}$) and the to end ($\mathbf{P_t}$), generated power ($\mathbf{P_g}$), load demands ($\mathbf{P_d}$), voltage angles ($\mathbf{V_a}$) for all buses. Then, the original measurement vector $\mathbf{Z}$ as the input of the state estimator could be formulated as:

$$\mathbf{Z} = [\mathbf{P_f}, \mathbf{P_t}, \mathbf{P_g} - \mathbf{P_d}, \mathbf{V_a}]^T \tag{4}$$

where $\mathbf{P_g} - \mathbf{P_d}$ is the power injection of all buses; for buses with no generator or load demands, the corresponding entries of $\mathbf{P_g}$ and $\mathbf{P_d}$ will be zero, respectively. Then, the Jacobian matrix $\mathbf{H}$ is accordingly formulated as follows:

$$\mathbf{H} = [\mathbf{B_f}, \mathbf{B_t}, \mathbf{B_{bus}}, \mathbf{I}]^T \tag{5}$$

where $\mathbf{B_f}, \mathbf{B_t}, \mathbf{B_{bus}}$ are the admittance matrices for the from end branch flow equation, the to end branch flow equation, and the power injection equation, respectively [20]; $\mathbf{I}$ is the identity matrix for the voltage angle. Based on the DC assumption, $\mathbf{B_t} = -\mathbf{B_f}$. Therefore, according to (1), the attack vector will be $\mathbf{a} = \mathbf{Hc}$. The post-attack measurement vector will be $\mathbf{Z_a} = \mathbf{Z} + \mathbf{a}$.

When optimizing the sparsity, most studies have resorted to a relaxed ℓ_1-norm minimization problem [3], [21]–[23], though the sparsity itself shall be solved through the minimization of ℓ_0-norm of $\mathbf{a}$ that is NP-hard and computationally complex. In our work, we will also tackle the problem by minimizing the ℓ_0-norm directly.

For the first scenario, the impact is considered as the maximization of the errors injected into the state vector (voltage angle), because the maximized state error can affect automatic generation control, contingency analysis and economic dispatch in the energy management system. The formulation can be written as:

$$\begin{aligned}
&\text{Minimize} && \|\mathbf{a}\|_0 && (6)\\
&\text{Maximize} && (\Sigma_{i=1}^{n_b} |c_i|)/(n_b - 1) && (7)\\
&\text{subject to} && \mathbf{a} = \mathbf{Hc} && (8)\\
& && (\mathbf{B_{bus}c})_{\{i|i \in \mathbf{ID}_0\}} = \mathbf{0} && (9)\\
& && |\mathbf{P_f} + \mathbf{B_f c}| \leq \mathbf{T_r} && (10)\\
& && \mathbf{B_{bus}c} \leq \mathbf{P_d} && (11)\\
& && |\mathbf{V_a} + \mathbf{Ic}| \leq \mathbf{30^\circ} && (12)
\end{aligned}$$

where $\mathbf{c}$ is the vector of injected state errors; n_b is the number of buses; $\mathbf{T_r}$ is the vector of branch thermal rating. $\mathbf{ID}_0$ is the set of the indices of buses with no load demands. $|\cdot|$ is used to take the absolute value of each entry of the matrix/vector. (6) is the ℓ_0-norm of the attack vector $\mathbf{a}$, representing the number of attacked meters, which should be minimized; (7) is the second objective function used to maximize the average error injected to each state. For the constraints, (8) ensures that the FDIA scheme follows the DC assumption; (9) ensures that the buses with no load demands will have no additional power injection; (10) ensures the attacked measurements of branch flow are still within the branch thermal ratings; (11) ensures that the compromised load demand is non-negative; (12) ensures that absolute compromised voltage angles should be within the threshold of 30°.

It is notable that the injected state error in this scenario will directly mislead operator's decisions in the control room through the energy management systems and could be proven costly. However, the impact on the state error might not be directly translated into specific measurements, e.g., branch power flows if the latter is the attacker's target. For instance, if the attacker aims to misinform the operator about an incoming transmission congestion and obtain financial gains from a high-demand market [12], then he/she will have to directly optimize the impact on the branch flow measurements, not the state vector. Given this consideration, we have also formulated another scenario with impacts on the measurements refined from the state vector.

B. Scenario II: Injected Error into Branch Power Flow Measurements

In this scenario, the attacker specifically aims to inject errors to the branch power flow measurements as much as possible. To achieve this impact, Scenario II aims to maximize the average relative injected error of the branch power flow while minimizing the same ℓ_0-norm of the attack vector as in Scenario I. The problem could be written as:

$$\text{Minimize} \qquad \|\mathbf{a}\|_0 \tag{13}$$

$$\text{Maximize} \qquad \frac{\Sigma_{i=1}^{n_{br}} |\overline{P}_{f_i} - P_{f_i}|/T_{r_i}}{n_{br}} \tag{14}$$

$$\text{subject to} \qquad (8)\text{-}(12) \tag{15}$$

where $\overline{P}_{f_i}$ and P_{f_i} are the manipulated and the original branch flows of the ith branch, respectively; T_{r_i} is the thermal rating of the ith branch; n_{br} is the number of branches; the constraints are the same as to Scenario I.

C. Constraint Handling and Attack Vector Optimization

To find the critical attack vector for the FDIA, the injected state error $\mathbf{c}$ is used as the decision vector for the optimization solver. From (8)-(12), if we transform the matrix expressions into a set of equalities or inequalities, we will get (n_b+n_z) equalities and $(2n_{br}+2n_b)$ inequalities, where n_z is the number of zero-load buses. For a large scale system, such number of constraints will limit the efficiency of the optimization solver to locate the candidates on the feasible region and find the optimal solutions. Therefore, we apply a three-step constraints handling method from the aspects of relaxing the number of constraints, compressing the searching space for the decision vector, and improving the searching strategy for the optimization algorithm, respectively. These methods work together with the SPEA2 optimization to solve this FDIA problem, which are described as follows.

1) Target Bus Selection: This method is used to deal with the equality constraint (9), ensuring that the attacker will not inject any error to the zero-load buses. The DC power flow of zero-load buses is described in (16) [24]:

$$P_k = \sum_{j\in k} B_{kj}(\theta_k - \theta_j) \quad k \in \mathbf{ID}_0 \tag{16}$$

where P_k is the power injection of the zero-load bus, B_{kj} is the susceptance between bus i and j, and θ_k and θ_j are angles of zero-load buses and their connected buses, respectively.

According to (16), if the angle deviation $\Delta\theta$ of a zero-load bus is the same as that of its directly connected buses, the power injection to the zero-load bus will not be changed. Thus, we find all the connected buses for each zero-load bus, and impose them with the same injected state error. Moreover, if a zero-load bus shares the same connected buses with another zero-load bus, all of these buses' angle deviation should be the same. Moreover, if the zero-load bus connects to or be the slack bus, all of these buses' angle deviation should be zero. Therefore, we apply the following process to determine the new decision variables:

Step 1: For each zero-load bus, find all its connected buses and get a set $\mathbf{G}_i$ (i= 1,..., n_z), which contains the indices of its connected buses and itself. All these sets are added into a set container $\mathbf{\Omega}$.

Step 2: Compare each set in pairs, if $\mathbf{G}_i \cap \mathbf{G}_j \neq \emptyset$, then, $\mathbf{G}_i, \mathbf{G}_j = \mathbf{G}_i \cup \mathbf{G}_j$. Then, delete the duplicate sets in $\mathbf{\Omega}$.

Step 3: Find the complementary set Ψ of the union of all sets in $\mathbf{\Omega}$. Each element in Ψ is considered as a new individual set and added into $\mathbf{\Omega}$.

Step 4: Delete the set contains the slack bus in $\mathbf{\Omega}$.

At the end of **Step 4**, we can inject the same state error to buses that belong to the same set in $\mathbf{\Omega}$, and will not inject any error to buses that do not appear in $\mathbf{\Omega}$, which ensures that the zero-load buses remains unattacked. Thus, we get a new decision vector $\mathbf{x} = (x_1, x_2, ..., x_{n_{dim}})^T$, where n_{dim} is the number of sets in $\mathbf{\Omega}$, and x_i is the injected state error to the buses belong to the ith set in $\mathbf{\Omega}$. Then, constraint (9) is accordingly removed and will not affect the process of identifying $\mathbf{x}$.

2) Searching Space Compression with Linear Programming (LP): LP is a linear optimization technique used to optimize a linear function that subjects to a set of linear constraints [25]. Here, we apply it to determine the lower and upper boundaries for each decision variable in $\mathbf{x}$ by solving the linear inequalities of (10)-(12). Since the decision vector now has become $\mathbf{x}$ rather than $\mathbf{c}$, the inequalities need to be updated accordingly. The formulation of this LP problem is expressed as:

$$\text{Minimize} \qquad x_i, \qquad i = 1, .., n_{dim} \tag{17}$$

$$\text{subject to} \qquad |\mathbf{P_f} + \overline{\mathbf{B}}_{\mathbf{f}}\mathbf{x}| \leq \mathbf{T_r} \tag{18}$$

$$\overline{\mathbf{B}}_{\mathbf{bus}}\mathbf{x} \leq \mathbf{P_d} \tag{19}$$

$$|\mathbf{V_a} + \bar{\mathbf{I}}\mathbf{x}| \leq \mathbf{30^\circ} \tag{20}$$

$$\text{Maximize} \qquad x_i, \qquad i = 1, .., n_{dim} \tag{21}$$

$$\text{subject to} \qquad (18) - (20) \tag{22}$$

where the minimization and maximization of x_i are corresponding to solving the lower and upper boundaries, respectively; $\overline{\mathbf{B}}_{\mathbf{f}}$, $\overline{\mathbf{B}}_{\mathbf{bus}}$, and $\bar{\mathbf{I}}$ are updated matrices, which are expressed as follows:

$$\overline{\mathbf{B}}_{\mathbf{f}}(:, i) = \sigma(\mathbf{B_f}(:, \mathbf{\Omega}_i)) \qquad i = 1, .., n_{dim} \tag{23}$$

$$\overline{\mathbf{B}}_{\mathbf{bus}}(:, i) = \sigma(\mathbf{B_{bus}}(:, \mathbf{\Omega}_i)) \qquad i = 1, .., n_{dim} \tag{24}$$

$$\bar{\mathbf{I}}(:, i) = \sigma(\mathbf{I}(:, \mathbf{\Omega}_i)) \qquad i = 1, .., n_{dim} \tag{25}$$

where $(:, i)$ is the ith column of the matrix; $(:, \mathbf{\Omega}_i)$ is the matrix's columns coming from the ith set in $\mathbf{\Omega}$; $\sigma(\cdot)$ is an operation for the matrix, which returns a column vector containing the summation of each row.

In this work, LP problems are solvable by most commercial solvers to get the boundaries of each decision variable. It should be noted that although LP can narrow the searching space to some extent, it still cannot ensure that the candidate solutions of evolutionary algorithm meet all the constraints with high probabilities during the evolutionary process, because LP cannot accurately describe the shape of the feasible region of the decision vector.

To efficiently guide individuals to the feasible region during the optimization process, we further apply another constraint handling technique, which is the constraint domination method, and will combine it with the Improved Strength Pareto Evolutionary Algorithm (SPEA2).

3) SPEA2 with Constraint Domination: SPEA is a type of evolutionary algorithm to solve multi-objective optimization problems [26]. SPEA2 is the improved version of the SPEA, which incorporates, in contrast to its predecessor, a fine-grained fitness assignment strategy, a density estimation technique, and an enhanced archive truncation method [15]. In each iteration, SPEA2 sorts the individuals of the population, and select the next generation from the combination of the current population and off-springs created by genetic operators. The iteratively converged non-dominated solutions in the archive set, which make up the estimated Pareto front in the objective space, are the expected results of the multi-objective optimization problem. All of these solutions are feasible, which reflects the trade-off between the objectives.

The details of the SPEA2 are given as follows [15].

Definitions: $\mathbf{A}_t$: population set at iteration t; $\overline{\mathbf{A}}_t$: archive set at iteration t; N: population size; $\overline{N}$: archive size; N_g: maximum number of generations.

Step 1: Initialization: Set $t = 0$. Generate an initial population $\mathbf{A}_0$ and create the empty archive set $\overline{\mathbf{A}}_0 = \emptyset$.

Step 2: Fitness assignment: Calculate fitness values of individuals in $\mathbf{A}_t$ and $\overline{\mathbf{A}}_t$. For each individual i in $\mathbf{A}_t$ and $\overline{\mathbf{A}}_t$, the strength value is calculated using the following equation:

$$S(i) = |\{j| \in \mathbf{A}_t \cup \overline{\mathbf{A}}_t \wedge i \succ j\}| \qquad (26)$$

where $|\cdot|$ denotes the cardinality of a set, $\cup$ stands for multi-set union, $\succ$ corresponds to the Pareto dominance relation, and $\wedge$ is logical AND. Then, the fitness $F(i)$ is defined as follow:

$$F(i) = R(i) + D(i) \qquad (27)$$

where the raw fitness $R(i)$ of an individual i is calculated by the following equation:

$$R(i) = \sum_{j \in \mathbf{A}_t \cup \overline{\mathbf{A}}_t, i \succ j} S(j) \qquad (28)$$

To distinguish individuals with the same raw fitness, the individual density is calculated by the K-nearest neighbor [15]:

$$D(i) = \frac{1}{\eta_{i,k} + 2} \qquad (29)$$

where $\eta_{i,k}$ is the distance between the ith and kth nearest neighbors in the objective space. As a common setting, we use $k = \sqrt{N + \overline{N}}$.

Step 3: Environmental selection: Copy all non-dominated individuals in both $\mathbf{A}_t$ and $\overline{\mathbf{A}}_t$ to $\overline{\mathbf{A}}_{t+1}$. If the size of $\overline{\mathbf{A}}_{t+1}$ exceeds $\overline{N}$, reduce $\overline{\mathbf{A}}_{t+1}$ with the truncation operator [15]. Otherwise, fill $\overline{\mathbf{A}}_{t+1}$ with dominated individuals in $\mathbf{A}_t$ and $\overline{\mathbf{A}}_t$.

Step 4: Termination: If the stopping criterion $t \geq N_g$ is satisfied, the output decision vectors are represented by the non-dominated individuals in $\overline{\mathbf{A}}_{t+1}$, then, stop. Otherwise, proceed to **Step 5:**.

Step 5: Genetic operation: Perform binary tournament selection with replacement on $\overline{\mathbf{A}}_{t+1}$ in order to fill the mating pool, in which crossover and mutation operators are applied. The new generation is set to $\mathbf{A}_{t+1}$. Set $t = t + 1$.

The constraint domination method is applied to deal with the dominance relationship between each individual, which will gradually evolve individuals to the feasible region in an efficient manner [27].

Specifically, given two individuals $\mathbf{a}$ and $\mathbf{b}$ with n_{obj} number of objective functions, we calculate the number of constraint violations $n_v(\mathbf{a})$ and $n_v(\mathbf{b})$ as well as the objective vectors $\mathbf{F}(\mathbf{a})$ and $\mathbf{F}(\mathbf{b})$, respectively. Then $\mathbf{a}$ is called dominating $\mathbf{b}$ when one of the following conditions is met:

- $\mathrm{n}_v(\mathbf{a}) = 0, n_v(\mathbf{b}) > 0.$
- $\mathrm{n}_v(\mathbf{a}) > 0, n_v(\mathbf{b}) > 0, n_v(\mathbf{a}) < n_v(\mathbf{b}).$
- $\mathrm{n}_v(\mathbf{a}) = 0, n_v(\mathbf{b}) = 0, \mathbf{F}(\mathbf{a}) \prec \mathbf{F}(\mathbf{b}).$

where $\mathbf{F}(\mathbf{a}) \prec \mathbf{F}(\mathbf{b})$ means $\forall F_i(\mathbf{a}) \leq F_i(\mathbf{b})$, and at least one $F_i(\mathbf{a}) < F_i(\mathbf{b})$, $i = 1, 2, ..., n_{obj}$ (Here we suppose all objectives are to be minimized. If there are objectives that need to be maximized, objectives' sign should be reversed when determining the dominance relation).

D. Execution of Multi-Objective FDIA

Fig. 1 shows the flow-chart of steps that we follow to obtain the sparse attack vector with the great impact based on the methods discussed above. The descriptions of each step are given as follows:

Step 1: Obtain measurements from the SCADA and formulate the measurement vector $\mathbf{Z}$ according to (4). Get the Jacobian matrix H according to (5).

Step 2: Determine the new decision vector $\mathbf{x}$ with the purpose of bypassing the zero-load buses by applying the target bus selection proposed in section III.C.

Step 3: Compress the searching space for $\mathbf{x}$ by applying the Linear Programming as the second constraint handling technique introduced in section III.C.

Step 4: Implement the SPEA2 with constraint domination to estimate the Pareto front based on the method in section III.C. From the final Pareto front, we can get a set of non-dominated solutions of $\mathbf{x}$, based on which we can further get a set of optional injected state error $\mathbf{c}$ and the corresponding attack vector $\mathbf{a}$.

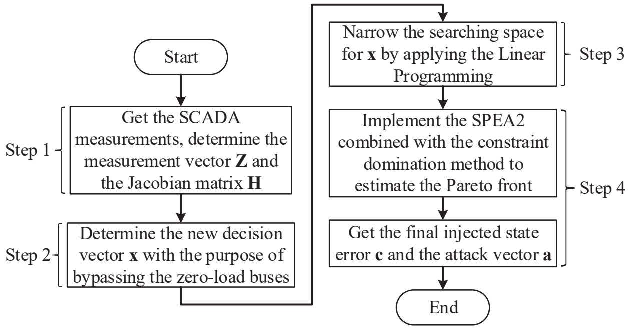
Start
Step 1
Get the SCADA measurements, determine the measurement vector Z and the Jacobian matrix H
Step 2
Determine the new decision vector x with the purpose of bypassing the zero-load buses
Narrow the searching space for x by applying the Linear Programming
Step 3
Implement the SPEA2 combined with the constraint domination method to estimate the Pareto front
Step 4
Get the final injected state error c and the attack vector a
End

Fig. 1. Flowchart of the multi-objective optimization process to find the sparse attack vector with the great impact.

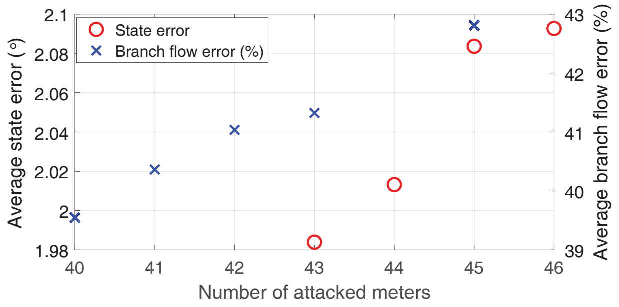

Fig. 2. The estimated Pareto fronts of two scenarios for IEEE 14-bus system.

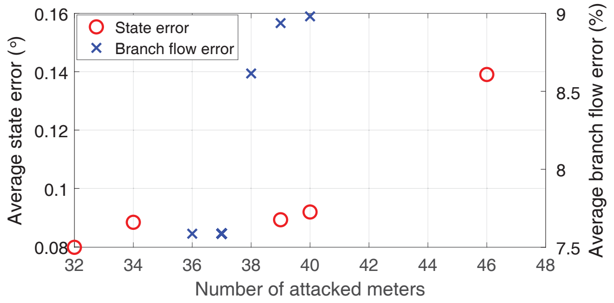

Fig. 3. The estimated Pareto fronts of two scenarios for IEEE 30-bus system.

IV. Experiment Setup and Evaluations

A. Experiment Setup

The proposed scheme is evaluated on the IEEE 14-bus and the IEEE 30-bus test systems with the MATPOWER toolbox [20]. For the IEEE 14-bus system, we get 40 active branch flow measurements (20 at the-from end and 20 at the to-end), 5 generation power outputs, 11 load demands and 13 voltage angles (except for the reference bus) summing up to a total 69 attackable meters. For the IEEE 30-bus system, we get 82 measurements of active branch flows (41 at the-from end and 41 at the to-end), 6 generation power outputs, 20 load demands, and 29 voltage angles, among which the total number of attackable measurements is 137.

For the configuration of the SPEA2 algorithm, we set the crossover parameter $\gamma = 0.3$, mutation parameter $h = 0.6$, population size $N = 500$, archives size $\overline{N} = 100$, and iteration number $N_g = 200$. γ, $\overline{N}$, and N are determined empirically, while $h = 0.6$ is determined based on a number of experiments. It was found that if we choose h smaller or greater than 0.6, it will usually have lower convergence rate compared with that of 0.6 when under the same iteration number.

B. Attack Performance Evaluation

Fig. 2 and Fig. 3 show two different Pareto fronts for two scenarios mentioned in Section III, which illustrate the impact on the system in terms of both the averaged state errors and the averaged branch power flow error for the 14-bus system and the 30-bus system, respectively. The results show that we are able to find out a group of non-dominated solutions of **c** that outline the trade-off between the number of measurements attacked and the expected impacts of FDIA for both the 14-bus and 30-bus systems.

For the Pareto fronts of Scenario I in Fig. 2 and Fig. 3, when more meters are attacked, great impact could be inflicted on the state variables. For the case of 14-bus system, the magnitude of injected state error does not increase significantly with the increasing number of attacked meters. From Fig. 2, we can see that after attacking 3 more meters from the minimum value of attacked meters (43), the magnitude of injected state error increases from $1.98°$ to approximately $2.1°$. Nevertheless, the injected error itself is a quite moderate deviation from the original state values. The minimum sparsity achieved here is approximately 58% which is lower than what is claimed in [1]. On the other hand, for the 30-bus system, from Fig. 3, it can be found that impact on the state error has a sharp increase after attacking more than 40 meters, up to which point the state error does not increase much with increased attacked meters. After this point the injected state error increases from approximately $0.09°$ to more than $0.14°$ (60%) given that 8 more meters have been attacked. The minimum sparsity achieved here is 25%, which is much more sparse than the 14-bus system. Therefore, we can conclude that for larger system, the attacker can inflict more impact in terms of the injected state error by compromising fewer meters.

The Pareto fronts for Scenario II are shown with blue crosses in Fig. 2 and Fig. 3. For the 14-bus system, the impact in terms of the branch flow is quite substantial with the minimum error being more than 39%. It could also be realized from Fig. 2, that by attacking only 5 more meters the branch error increases from approximately 39% to 45.7%. The minimum sparsity in this case is 58% which is slightly higher than in the case of Scenario I. On the other hand, for the 30-bus system, although the injected branch error is moderately high, the minimum impact being 7.6%, however, the minimum sparsity achieved here is very low (26%). Here also, from Fig. 2 and Fig. 3, we can say that for larger systems, fewer meters are needed to be compromised to inflict more impact.

C. Impact Analysis

We further compared the impacts between the two scenarios and demonstrate how the attacker may choose to optimize the impact on the system. Fig. 4 and Fig. 5 show the injected state error at each bus under two scenarios for the 14-bus and 30-bus system, respectively. It can be seen that for both systems, in Scenario II, when the branch error injection is maximized, the consequent deviation in the injected state error is greater than in Scenario I. This could be explained by the fact that injected state error in the measurements in turn generates even larger state errors that is bypassed by the BDD. The zero values of state errors indicates buses that are directly connected with a zero-load or the reference bus and thus were not attacked.

From the attacker's perspective, the goal is to inflict as much impact as possible by attacking fewer meters, which is the worst-case scenario. In this regard, we have found that the attacker may choose to maximize state error if he/she wants to mislead the control room, in which case Scenario I allows the attacker to inject moderate error into the estimated states with attacking as few as 34 out of the 137 meters

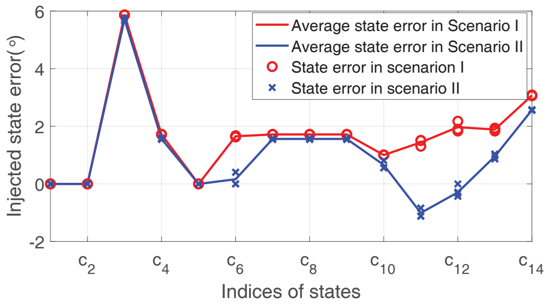

Fig. 4. The injected state error at each bus under two scenarios for the IEEE 14-bus system.

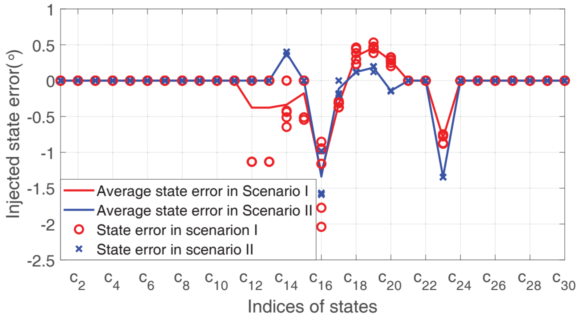

Fig. 5. The injected state error at each bus under two scenarios for the IEEE 30-bus system.

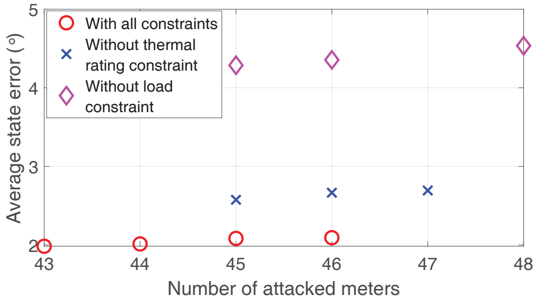

Fig. 6. Comparison of Pareto fronts with and without constraints for IEEE 14-bus system under scenario I

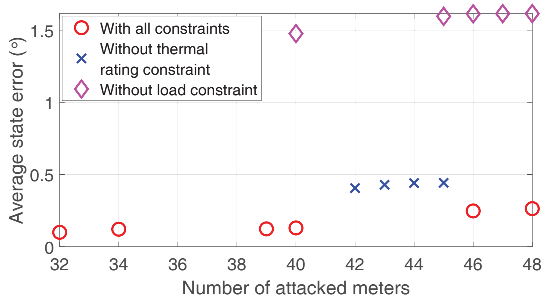

Fig. 7. Comparison of Pareto fronts with and without constraints for IEEE 30-bus system under scenario I.

for the 30-bus and 43 out of 69 for the 14-bus system. Alternatively, the attacker may choose to report false branch flow measurements and Scenario II will enable the attacker to inflict a comparatively larger error on the branch flow measurements and with compromising fewer meters. In both scenarios, by formulating the objective function accordingly, the attacker can optimize the impact with minimal efforts and remaining stealthy to the PSSE.

D. Influence of the Constraints

To make our attack model strictly stealthy, we considered a number of constraints on the attacked measurements as introduced in Section III. However, the attacker may relax some constraints with the cost of higher detection possibility but which will inflict far greater impact. We have examined the influence of these constraints to better analyze the risks and possible countermeasures from the defender's perspective. To show this we removed one constraint at a time and investigated the performance of the attack model correspondingly. Fig. 6 and Fig. 7 shows the comparison of the Pareto fronts with and without the constraints, for the 14-bus and the 30-bus system respectively under Scenario I.

Both the Figures for Scenario I show that the load demand constraint has the greatest influence on the performance of the attack model whose removal increases the injected state error impact by more than two times for the 14-bus system and more then 6 time for the 30-bus system, thereby affecting the larger systems more severely. Similarly for Scenario II, Fig. 8 and Fig. 9 also demonstrates the load demand constraint influences most, removal of which will increase the branch flow error 4 times for the 14-bus system and 7 times for the 30-bus system.

These results indicated that, if the attacker can compromise the load buses such that the load demand becomes negative, this could lead to dire consequences and hence, the defender could protect the load buses with such a detector that can instantly detect the presence of a negative load demand.

Substantial studies have been devoted to investigate ways to protect against FDIA, among which, data driven techniques are more popular recently [28]–[30]. There are also prevention based mechanisms [3], [9] where the critical meters are found and protected. However, the constraint influence analysis enables us to find the most vulnerable buses in terms of the impact and in the event of an attack that would alter the normal behavior, what should be the immediate response.

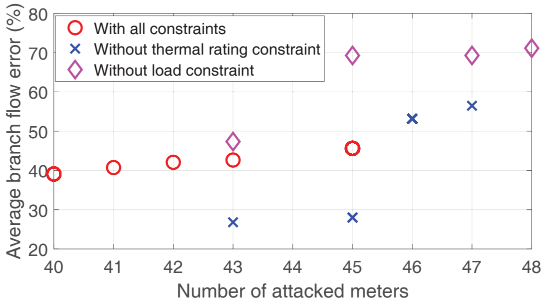

Fig. 8. Comparison of Pareto fronts with and without constraints for IEEE 14-bus system under scenario II.

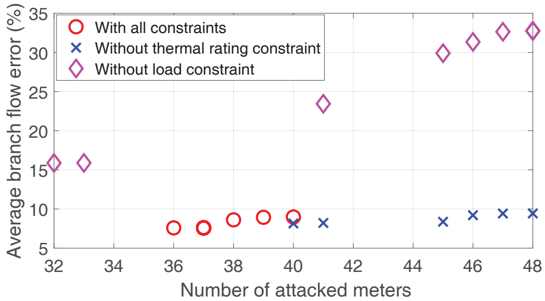

Fig. 9. Comparison of Pareto fronts with and without constraints for IEEE 30-bus system under scenario II.

V. Conclusions

This paper proposed a novel multi-objective optimization approach to analyze the false data injection attacks in power system state estimation. Using the SPEA2, the proposed approach efficiently minimizes the number of attacked meters while maximizing the impact. The results have shown that it is possible to implement a stealthy and sparse FDIA, while the attackers can optimize the expected impact in terms of injected errors on both the state vectors and the branch power flows. More compromised meters will inflict greater impact. In addition, the impact analysis also shows that, comparatively fewer compromised meters will administer greater impact to larger systems. Further, by considering alternative constraints, we demonstrated the trade-off between stealthiness and impact.

In future, we aim to investigate our attack formulation considering the AC state estimation for enhanced threat analyses. It has been shown that the attack formulated in DC environment launched upon AC state estimator have higher probability of detection [31], therefore it will be judicious to consider AC FDIA for better risk analysis. Furthermore, we aim to devise improved detection mechanism for this attack formulation that takes into account the worst-case scenario (i.e., no topology knowledge, minimum resources and maximum impact).

Acknowledgement

The research is supported by the Natural Sciences and Engineering Research Council of Canada under grant RGPIN-2018-06724 and the Fonds de Recherche du Québec-Nature et Technologies under grant 2019-NC-254971.

References

[1] Y. Liu, P. Ning, and M. K. Reiter, “False data injection attacks against state estimation in electric power grids,” *ACM Trans. Inf. Syst. Secur.*, vol. 14, no. 1, pp. 13:1–13:33, Jun. 2011.

[2] R. Deng, G. Xiao, R. Lu, H. Liang, and A. V. Vasilakos, “False data injection on state estimation in power systems—attacks, impacts, and defense: A survey,” *IEEE Transactions on Industrial Informatics*, vol. 13, no. 2, pp. 411–423, 2016.

[3] M. Ozay, I. Esnaola, F. Y. Vural, S. Kulkarni, and V. Poor, “Sparse attack construction and state estimation in the smart grid: Centralized and distributed models,” *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 7, pp. 1306–1318, 2013.

[4] J. Hao, R. Piechocki, D. Kaleshi, W.-H. Chin, and Z. Fan, “Sparse malicious false data injection attacks and defense mechanisms in smart grids,” *IEEE Transactions on Industrial Informatics*, vol. 11, no. 5, pp. 1–12, 2015.

[5] H. Zhong, D. Du, C. Li, and X. Li, “A novel sparse false data injection attack method in smart grids with incomplete power network information,” *Complexity*, vol. 2018, 2018.

[6] S. Xie, J. Yang, K. Xie, Y. Liu, and Z. He, “Low-sparsity unobservable attacks against smart grid: Attack exposure analysis and a data-driven attack scheme,” *IEEE Access*, vol. 5, pp. 8183–8193, 2017.

[7] J. Tian, B. Wang, and X. Li, “Data-driven and low-sparsity false data injection attacks in smart grid,” *Security and Communication Networks*, vol. 2018, 2018.

[8] Q. Yang, J. Yang, W. Yu, D. An, N. Zhang, and W. Zhao, “On false data-injection attacks against power system state estimation: Modeling and countermeasures,” *IEEE Transactions on Parallel and Distributed Systems*, vol. 25, no. 3, pp. 717–729, March 2014.

[9] Y. Yuan, Z. Li, and K. Ren, “Modeling load redistribution attacks in power systems,” *IEEE Transactions on Smart Grid*, vol. 2, no. 2, pp. 382–390, 2011.

[10] Y. Yuan, Z. Li, and K. Ren, “Quantitative analysis of load redistribution attacks in power systems,” *IEEE Transactions on Parallel and Distributed Systems*, vol. 23, no. 9, pp. 1731–1738, September 2012.

[11] J. Zhang, Zhigang Chu, L. Sankar, and O. Kosut, “False data injection attacks on power system state estimation with limited information,” in *2016 IEEE Power and Energy Society General Meeting (PESGM)*, July 2016, pp. 1–5.

[12] L. Xie, Y. Mo, and B. Sinopoli, “Integrity data attacks in power market operations,” *IEEE Transactions on Smart Grid*, vol. 2, no. 4, pp. 659–666, December 2011.

[13] J. Yan, Y. Tang, Bo Tang, H. He, and Y. Sun, “Power grid resilience against false data injection attacks,” in *2016 IEEE Power and Energy Society General Meeting (PESGM)*, July 2016, pp. 1–5.

[14] J. N. Rodrigues de Assis, T. M. Machado-Coelho, G. Luís Soares, and M. H. Soares Mendes, “Robust evolutionary optimization algorithm for multi-objective environmental/economic dispatch problem with uncertainties,” in *2018 IEEE Congress on Evolutionary Computation (CEC)*, July 2018, pp. 1–6.

[15] E. Zitzler, M. Laumanns, and L. Thiele, “SPEA2: Improving the strength pareto evolutionary algorithm,” *TIK-report*, vol. 103, 2001.

[16] A. Abur and A. Gómez-Expósito, “Power system state estimation: Theory and implementation. New York: Marcel Deccer,” 2004.

[17] 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*, vol. 8, no. 4, pp. 1630–1638, 2016.

[18] B. Tang, Jun Yan, S. Kay, and H. He, “Detection of false data injection attacks in smart grid under colored gaussian noise,” in *2016 IEEE Conference on Communications and Network Security (CNS)*, Oct 2016, pp. 172–179.

[19] Z. Ma and Y. Wang, “Evolutionary constrained multiobjective optimization: Test suite construction and performance comparisons,” *IEEE Transactions on Evolutionary Computation*, vol. 23, no. 6, pp. 972–986, Dec 2019.

[20] R. D. Zimmerman and C. E. Murillo-Sánchez, “Matpower 4.1 user's manual,” *Power Systems Engineering Research Center (PSERC)*, vol. 20, 2011.

[21] W. Wang and Z. Lu, “Cyber security in the smart grid: Survey and challenges,” *Computer networks*, vol. 57, no. 5, pp. 1344–1371, 2013.

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

[23] K. C. Sou, H. Sandberg, and K. H. Johansson, “Computing critical k-tuples in power networks,” *IEEE Transactions on Power Systems*, vol. 27, no. 3, pp. 1511–1520, 2012.

[24] K. Van den Bergh, E. Delarue, and W. D'haeseleer, “Dc power flow in unit commitment models,” *no. May*, 2014.

[25] S. Chand, H. K. Singh, and T. Ray, “Team selection using multi-/many-objective optimization with integer linear programming,” in *2018 IEEE Congress on Evolutionary Computation (CEC)*, July 2018, pp. 1–8.

[26] R. Gharari, N. Poursalehi, M. Abbasi, and M. Aghaie, “Implementation of strength pareto evolutionary algorithm ii in the multiobjective burnable poison placement optimization of kwu pressurized water reactor,” *Nuclear Engineering and Technology*, vol. 48, no. 5, pp. 1126–1139, 2016.

[27] J. Li, Y. Wang, S. Yang, and Z. Cai, “A comparative study of constraint-handling techniques in evolutionary constrained multiobjective optimization,” in *2016 IEEE Congress on Evolutionary Computation (CEC)*, July 2016, pp. 4175–4182.

[28] M. Ozay, I. Esnaola, F. T. Yarman Vural, S. R. Kulkarni, and H. V. Poor, “Machine learning methods for attack detection in the smart grid,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 27, no. 8, pp. 1773–1786, Aug 2016.

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

[30] M. Esmalifalak, L. Liu, N. Nguyen, R. Zheng, and Z. Han, “Detecting stealthy false data injection using machine learning in smart grid,” *IEEE Systems Journal*, vol. 11, no. 3, pp. 1644–1652, September 2017.

[31] M. A. Rahman and H. Mohsenian-Rad, “False data injection attacks against nonlinear state estimation in smart power grids,” in *2013 IEEE Power & Energy Society General Meeting*. IEEE, 2013, pp. 1–5.