IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received January 6, 2019, accepted January 20, 2019, date of publication February 25, 2019, date of current version March 13, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2899293

Modelling Cyber Attacks on Electricity Market Using Mathematical Programming With Equilibrium Constraints

SAEED AHMADIAN[1], (Member, IEEE), XIAO TANG[2], (Member, IEEE), HEIDAR A. MALKI[3], (Senior Member, IEEE), AND ZHU HAN[1], (Fellow, IEEE)

[1]Department of Electrical and Computer Engineering, University of Houston, Houston, TX 77004, USA
[2]Department of Communication Engineering, Northwestern Polytechnical University, Xi'an 710065, China
[3]Department of Engineering Technology, University of Houston, Houston, TX 77004, USA

Corresponding author: Saeed Ahmadian (sahmadian3@uh.edu)

This work was supported in part by the US Multidisciplinary Research Program of the University Research Initiative (MURI) AFOSR MURI under Grant 18RT0073, in part by the NSF under Grant CNS-1717454, Grant CNS-1731424, Grant CNS-1702850, Grant CNS-1646607, and in part by the Texas Research Incentive Program under Grant C108430.

ABSTRACT With the development of communication infrastructure in smart grids, cyber security reinforcement has become one of the most challenging issues for power system operators. In this paper, an attacker is considered a participant in the virtual bidding procedure in the day-ahead (DA) and real-time (RT) electricity markets to maximize its profit. The cyber attacker attempts to identify the optimal power system measurements to attack along with the false data injected into measurement devices. Towards the maximum profit, the attacker needs to specify the relation between manipulated meters, virtual power traded in the markets, and electricity prices. Meanwhile, to avoid being detected by the system operator, the attacker considers the physical power system constraints existing in the DA and RT markets. Then, a bi-level optimization model is presented which combines the real electricity market state variables with the attacker decision-variables. Using the mathematical problem with equilibrium constraints, the presented bi-level model is converted into a single level optimization problem and the optimal decision variables for the attacker are obtained. Finally, simulation results are provided to demonstrate the performance of the attacker, which also provides insights for security improvement.

INDEX TERMS Cyber-security, electricity market, false data injection, mathematical programming with equilibrium constraints.

I. INTRODUCTION

Smart grids are exposed to risk of cyber-attacks since the heavily integration of communication devices into system. Smart grids work based on different physical and cyber levels of communication. Generally, measurements devices in the field send the data to the closest data collector device (IP-based router) using radio frequency signals. Then using Supervisory Control And Data Acquisition (SCADA) system, the collected data is transferred to central energy management system. However, the SCADA system uses firewalls to secure the communication network, still there are many cyber-attacks in different communication steps from disrupting radio frequency devices to manipulating the firewall systems. Thus, the cyber-security of electrical network becomes increasingly challenging for electrical independent system operators (ISOs) [1]. In this paper, it is aimed to present a comprehensive model which indicates the pattern of attacker's behaviours.

The associate editor coordinating the review of this manuscript and approving it for publication was Giambattista Gruosso.

A. CURRENT STATE OF ART

The research works on power system hacking can be categorized into two types. The first type of cyber-attacks concentrates on any general damage to electrical system operation and the second type of attacks focus on monetary issues in electricity markets.

In the first type, the attacker desires to malfunction system operation through injecting false data into the measurement system. This may cause a wide range of possible damages to power grid from a small amount of not-supplied load to large generator control malfunction and Automatic

Generation Control (AGC) unit failures [2]–[4]. To detect such attacks, different methods have been presented. In [5], using the Kalman filter, the residual value of the AGC system is predicted for the next time step of the operation. Then this value is used to obtain the Mahalanobis norm of residual vector. Comparing the Mahalanobis norm with the predefined threshold the attack is determined. In [6], FDI attacks to AGC are detected using load frequency control system of generators.The major signal for the attack's identification process is area control error (ACE). Then different types of attacks including over/under compensation and negative compensation attacks are formulated. Adding all possible formulations for the attacker, the author minimized the attack duration possibility. In [7], the principle components analysis (PCA) and support vector machine (SVM) algorithms are used to detect false data injection (FDI) attacks. In [8], a statistic decentralize learning algorithm is presented. In [9], by minimizing the residual value of power state estimator, the authors attempt to find the optimal attack vector.

One of the most common issues in this type of cyber-security investigation methods (first type) is to concentrate on power system state estimation attacks. The FDI attack against Alternative Current (AC) state estimator is discussed in many studies [10]. In [11], cyber attacks toward wide-area monitoring, protection and control (WAMPAC) system in energy systems are investigated. The paper, presents common attacks in WAMPAC and three different stages including attack prevention, attack detection and attack mitigation common methods.

In [12], both FDI and jamming attacks are modeled via a non-Bayesian approach with unknown information about type or the time of attack. To detect the attacks, using the Kalman filters, the system states are predicted and using generalized likelihood ratio approach the attack parameters are estimated. Considering low rank matrix of the measurement and sparse Jacobian of the electricity system, three different detection methods of FDI are compared in [13]. In this reference, augmented Lagrangian multipliers (ALM) and low rank matrix factorization method as bench mark algorithms to detect noisy measurements, are compared with Go decomposition method. Basically, the authors changed singular value decomposition in ALM with bilateral random projection. The problem with this method is that the attacker is consider to have the limited number of options and can pick its desired Phase Measurement Unit (PMU) to attack.

In [14], a new dynamic state estimation model is presented to mitigate the risk level of cyber attacks and also unknown inputs to power system. The authors used a linearized state model of the system including power system states, control variables and system parameters. Then, they considered unknown inputs and cyber attacks as extra inputs that effect on the state variables of power system (specifically rotor speed and system frequency). In fact, unknown inputs, cyber-attacks and detection filters are integrated into dynamic state estimation formula. Finally, based on the presented state estimation algorithm, the authors introduced a new risk mitigation method to detect possible faulty measurements and maximize security-level.

Mousavian *et al.* [15] presented a probabilistic formulation to model the probability of compromised measurement devices (e.g. PMU). They model random variable to show whether the PMU is attacked or not. Then, the probability of a compromised PMU and also other influenced PMUs by attacking a specific PMUs are presented using exponential probabilistic models. Regarding these models, the threat level of each PMUs is formulated which shows the probability of detecting the specific compromised PMU after multiple time intervals. Finally, an optimization problem is presented to minimize the number of disabled PMUs in order to keep power system observable and detect the affected PMUs.

In the second type of cyber-attacks, the attacker injects malicious data into power system measurement devices to obtain the monetary profit [16]. The main idea in existing literature is to increase the transmission line congestion using deterministic or stochastic methods [17]. Esmalifalak *et al.* [18] target at the highest line congestion to make profit from Financial Transmission Rights (FTRs). In [19], within the zero sum game theory, the attacker maximizes transmission congestion while the system operator aims to defend selected measurements.

FTR markets are not the only financial transactions that can be influenced by the cyber-attacks. The virtual bidding markets for the players who wants to invest on electricity markets, are the other types of markets can be manipulated by the attackers. The virtual bidding problem as one of the newly-emerged challenges in electricity markets is introduced [20]. Although the virtual bidding decreases the associated risk of real-time pricing and improves the efficiency of the market clearing process, the concerns of the electricity market's manipulation are still reported in [21]. Thus, to guarantee the efficiency of virtual bidding mechanism, the FDI attack issue within the virtual bidding market is of paramount importance to be addressed. To this end, the attack through electricity markets by considering the attacker as one of the virtual bidders is not-well presented in the literature.

B. CONTRIBUTIONS

In this paper, to address FDI attacks in the virtual bidding problem, a new optimization model is proposed. Each participant in the virtual bidding process, buys (sells) electricity in the DA market at cheap (expensive) prices and sells (buys) electricity in the RT market at expensive (cheap) prices. An attacker which participates in the virtual bidding injects false data into electricity grid's measurement devices (so called remote terminal units (RTUs)) such that maximize its profit. Meanwhile, the attacker aims to remain undetected by the ISO. Better to say, the attacker replaces the real RTU's measurements with fake data so that the manipulated RTUs look like the rests.

To avoid being detected by system operators, the proposed method considers the actual physical constraints in power system including load balance and power flow constraints

within the attacker's optimization problem. To reflect the power system constraints in the attacker decision-making problem, a bi-level programming model is presented. Then, using a mathematical programming equilibrium constraint (MPEC) model, sensitive measurements along with false data values are obtained. Using the proposed method, ISOs can identify vulnerable RTUs in the smart grids and strengthen the security on those parts. In particular, to detect cyber-sensitive parts in smart grid, an attacker is confronted with two main questions: what is the optimal pair of traded power and electricity price regarding the DA and RT markets, and what is the relation between the pair of traded power-electricity price and manipulated RTUs.

To be more specific, the contribution of this paper can be summarized as below:

- A novel virtual bidding procedure for an attacker to buy/sell electricity in the DA and RT markets is presented. The presented model consists of two levels: The upper level consists of the attacker's objective function and the being undetected related constraints. The lower level contains the physical layer of the power system including power flow equations in the DA market and incremental balance between generation and demand in the RT market.
- The presented model depicts inter-correlation between the DA and RT market electricity prices. Indeed, this research marks on the relation between the DA and RT electricity prices and how the attacker can alter them by manipulating measurement devices (RTUs specifically). New equations are presented to show the effects of malicious data on the DA and RT prices and power flow directions in the electricity grid. These sets of new equations are embedded into the power flow equations.
- Using MPEC modelling, the implicit relation between the first and the second above-mentioned contributions turn into set of explicit equations. Introducing some auxiliary variables in the presented model, the exact relation between the DA and RT electricity market prices, traded power in both markets by the attacker, the optimal RTU to attack and the value of FDI are specified.

The rest of paper is organized as follows. Section II explains the structures of DA and RT markets and how an attacker can participate as a virtual bidder. Section III elaborates methodology, presents the attack optimization problem, and adopts the MPEC model towards the optimal attack. Section IV depicts the simulation results, and Section V draws the conclusion.

II. DA AND RT ELECTRICITY MARKET

Participating in the virtual bidding procedure, the attacker needs to consider the procedure of the market clearing price in order to create the optimal attack vector. Fig. 1 shows the general idea where the attacker participates in a virtual bidding procedure while compromising the selected meters to change the output of market operator calculation (electricity market state variables). Based on the Fig. 1, the attacker needs to select the optimal RTU to inject the false data instead of true measured data at first. Specifically, the proposed method considers systematic cyber-attacks against electrical system measurements. Thus, the different process of data encryption and communication infrastructure between PMUs and RTUs are not the subject of proposed method. In other word, the power system state estimator doesn't consider the difference between RTU or PMUs data. All the information gathered by the measurement devices, are sent to data center to run the state estimation. In this paper cyber-attacks to RTUs are modeled since the RTUs are easier to manipulate and are more common in smart grid measurement process.

Finding the optimal RTU or RTUs, is based on the objective function presented in Section III-C. Then, the attacker

TABLE 1. Nomenclature.

	Definitions
$C_{it}, C_{it}^{'}$	The electricity price (\$/MWh) offered by the i^{th} generator to produce electricity at the DA and RT markets at time t, respectively.
$B_{it}, B_{jt}^{'}$	The electricity price (\$/MWh) offered by the j^{th} demand to consume electricity at the DA and RT markets at time t, respectively.
p_{it}, d_{jt}	The electricity power (MWh) sold and bought by the i^{th} generator and the j^{th} demand in the DA market at time t, respectively.
$\Delta p_{it}, \Delta d_{jt}$	The incremental electricity power (MWh) sold and bought by the i^{th} generator and the j^{th} demand in the RT market at time t, respectively.
θ_{nt}	Voltage phase angel of the n^{th} bus at time t.
F_{nk}	The flow limit (MW) of the line between node n and k.
GSF_{n-k}	Generalized shift factor of the n^{th} line with respect to electricity changes in bus k.
u_{mt}	The attacker binary decision variable whether to hack the m^{th} meter at time t or not.
Δz_{mt}	The amount of false data injected to measurements at time t.
Ω_n	The set of generators and demands connected to bus n.
$\lambda_{nt}, \lambda_{nt}^{'}$	The DA and RT markets clearing price at bus n in time t.
Π_n	The set of buses (nodes) connected to bus n.
X_{nk}	The reactance of the line between bus n and k.
C_{DA}	The total DA market cost (\$).
C_{RT}	The total RT market cost (\$).
N_g	The set of generators in the smart grid.
N_d	The set of demand in the smart grid.
N_{line}	The set of lines in the smart grid.

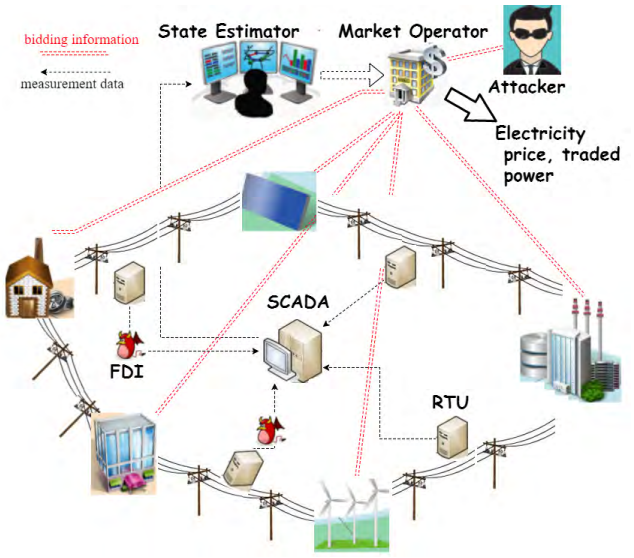

FIGURE 1. Inter-correlation of FDI and attacker's profit as virtual bidder.

needs to know what is/are the value/values that must be injected to selected RTU instead of actual measured value. This value/values are modeled by the constraint presented in Section III-D.

In brief, the attacker needs to find the optimal traded power and price, both of which are the functions of the measurement data (RTUs information). Thus, the relation between RTUs and these two parameters (power and price) must be clarified initially. To do so, the DA (section II-A) and the RT (section II-B) procedures are required and in this section, the general formulation for the DA and RT markets are presented.

A. DAY-AHEAD MARKET

The aim of the ISO is to maximize social welfare and also minimize total generation costs. Meanwhile, ISO must consider the power system constraints and limitations. In the following optimization, p_{it} and d_{jt} (MW) are the generation and demand for the i^{th} and j^{th} participants in the DA market with the C_{it} and B_{jt} prices ($/MWh) to sell and buy electricity. Bus voltage angels are shown with θ_{nt} (rad) and transmission line impedance between buses n and k as well as flow capacity are introduced with X_{nk} (Ω) and F_{nk} (MW). Therefore, the DA market clearing price can be optimized as below;

$$\min_{\lambda_{nt},p_{it},d_{jt}} C_{DA} = \left[\sum_{i=1}^{i=N_g}\sum_{t=1}^{t=T} C_{it}p_{it} - \sum_{j=1}^{j=N_D}\sum_{t=1}^{t=T} B_{jt}d_{jt}\right] \quad (1a)$$

$$s.t. \sum_{i\in\Omega_n} p_{it} - \sum_{j\in\Omega_n} d_{jt} = \sum_{k\in\Pi_n}\left(\frac{\theta_{nt}-\theta_{kt}}{X_{nk}}\right); \quad \forall n \in N_{bus}, \quad (1b)$$

$$\left|\frac{\theta_{nt}-\theta_{kt}}{X_{nk}}\right| \le F_{nk}; \quad \forall n,k \in N_{bus}, \quad (1c)$$

$$p_i^{min} \le p_{it} \le p_i^{max}; \quad \forall i \in N_{Gen}, \quad (1d)$$

$$d_{jt}^{min} \le d_{jt} \le d_{jt}^{max}; \quad \forall j \in N_d, \quad (1e)$$

$$-\pi \le \theta_{nt} \le \pi;\ \eta_{nt}^{max}, \quad \forall n \in N_{bus}. \quad (1f)$$

The problem in (1) shows the DA market mechanism. Eq. (1a) shows the total system cost that the ISO aims to minimize. In (1b), the power balance in each bus is depicted. The difference between generated electricity and consumed power is flowed into transmission lines. In (1c), the limits of transmission lines in the network are presented. In (1d) and (1e) the upper and lower bounds are represented for generated electricity and traded demand in the DA market, respectively. Finally, Eq. (1f) corresponds to the limitation of each bus voltage angle.

B. REAL-TIME MARKET

To operate the RT market, the ISO needs to collect all the data from different measurement devices in the network and minimize the incremental cost due to generation and load variations. Specifically, ISO attempts to minimize the difference between the scheduled dispatch generation-demand set in the DA market and the real-time measurements. Consequently, the following RT market optimization problem is formulated by the ISO [22].

$$\min_{\lambda'_{nt},\Delta p_{it},\Delta d_{jt}} C_{RT} = \left[\sum_{i=1}^{i=N_g}\sum_{t=1}^{t=T} C'_{it}\Delta p_{it} - \sum_{j=1}^{j=N_D}\sum_{t=1}^{t=T} B'_{jt}\Delta d_{jt}\right] \quad (2a)$$

$$\text{s.t.} \sum_{i\in\Omega_n}\Delta p_{it} - \sum_{j\in\Omega_n}\Delta d_{jt} = 0; \quad \forall n \in N_{bus}, \quad (2b)$$

$$\Delta p_{it}^{min} \le \Delta p_{it} \le \Delta p_{it}^{max}; \quad \forall i \in N_g, \quad (2c)$$

$$\Delta d_{jt}^{min} \le \Delta d_{jt} \le \Delta d_{jt}^{max}; \quad \forall j \in N_d, \quad (2d)$$

$$\sum_{n=1}^{N_{bus}} GSF_{n-k}\left(\Delta p_{nt} - \Delta d_{nt}\right) \le 0; \quad \forall k \in N_{line}. \quad (2e)$$

Eq. (2a) declares the system incremental cost which must be minimized. Eq. (2b) represents incremental balance between the generation and demand in each bus. In (2c) and (2d) the upper and lower bounds are determined for the generation and load increments for the attacker and other participants, respectively. In (2e), the line flow increment constraint is presented. the ISO tries to keep flow increment negative so that it prevents line congestion [23]. Regarding both equations in the DA and RT (1a-1e,2a-2e), if the attacker buys d_{it} MW in the DA market at λ_{nt} $/MWh, and after manipulating measurement data sells Δp_{it} MW at $(\lambda'_{nt} - \sum_m^{N_{line}} GSF_{n-m}\xi_{mt})$ $/MWh, its profit would be $\left[\lambda'_{nt} - \sum_m^{N_{line}} GSF_{n-m}\xi_{mt}\right]\Delta p_{it} - \lambda_{nt}d_{it}$ $/h. The same situation exists when the attacker sells expensive electricity at the DA market and buys cheap power at the RT market.

Finally, the profit from participation in these two markets (as the i^{th} participant) can be given as ($\lambda_{nt}^{RT} = \lambda'_{nt} - \sum_{m}^{N_{line}} GSF_{n-m}\xi_{mt}$);

$$R_i = \sum_{i\in\Omega_n}\sum_{t}^{T}\left[\left(\lambda_{nt}^{RT}\Delta p_{it} - \lambda_{nt}d_{it}\right) + \left(\lambda_{nt}p_{it} - \lambda_{nt}^{RT}\Delta d_{it}\right)\right]$$
$$= \sum_{i\in\Omega_n}\sum_{t}^{T}\left[\lambda_{nt}^{RT}\left(\Delta p_{it} - \Delta d_{it}\right) + \lambda_{nt}\left(p_{it} - d_{it}\right)\right]. \quad (3)$$

III. OPTIMAL CYBER ATTACK ON ELECTRICITY MARKET

To obtain the optimal attack vector, the attacker needs to know the relation between the traded power and electricity price in the DA and RT markets with measurement data. Noticing this relation, the attacker can create its attack model and choose the meters that are required to be compromised. Thus, except the problems in (1) and (2), an attacker needs to take power system state estimator calculation as well as real-time market clearing process [23] into account. For the attack model, in subsection III-A, the power system state estimation model is introduced briefly. Then, the attacker's decision variables and equations are presented in Subsection III-B. In the Subsection III-C, the attacker's optimization model is presented. Finally, Subsection III-D shows the linear relation between the market state variables and the attackers' decision-variables by embedding Subsection III-B into the Subsection III-C via the MPEC model.

A. POWER SYSTEM STATE ESTIMATOR

In the electricity market, denote $\mathbf{Z}$ as measurement vector and $\mathbf{H}$ as Jacobian matrix of linear power flow, and then the state estimator is obtained as $\mathbf{Z} = \mathbf{HX} + \mathbf{e}$, where $\mathbf{X}$ is the power system state variable vector, $\mathbf{e}$ is the measurement error. The error is considered to follow the Gaussian distribution with zero mean and co-variance matrix $\mathbf{C}$. The aim of state estimator is to estimate the state vector as $\hat{\mathbf{X}} = \arg\min_{\hat{\mathbf{x}}} \mathbf{E}\left\|\mathbf{X} - \hat{\mathbf{X}}\right\|_2^2$. Therefore, the optimal value for $\hat{\mathbf{X}}$ is given by [19]

$$\hat{\mathbf{X}} = \left(\mathbf{H}'\mathbf{C}^{-1}\mathbf{H}\right)'\mathbf{H}'\mathbf{C}^{-1}\mathbf{Z} = [\mathbf{M}]\,\mathbf{Z}. \quad (4)$$

The residual value of state estimator is given by $\|\mathbf{r}\|_2 = \left\|\mathbf{Z} - \mathbf{H}\hat{\mathbf{X}}\right\|_2 = \|\mathbf{Z} - \mathbf{HMZ}\|_2 \le threshold$. Adding $\Delta\mathbf{z}$ to $\mathbf{Z}$ changes residual value to [19]

$$\left\|\mathbf{r}^{\mathbf{new}}\right\|_2 = \|\mathbf{Z} + \Delta\mathbf{z} - \mathbf{HM}\left(\mathbf{Z} + \Delta\mathbf{z}\right)\|_2$$
$$\Rightarrow \left\|\mathbf{r}^{\mathbf{new}}\right\|_2 = \|\mathbf{r}\|_2 + \|\left(\mathbf{I} - \mathbf{HM}\right)\Delta\mathbf{z}\|_2. \quad (5)$$

Therefore, to avoid being caught, the following constraint is presented at the attacker $\|\left(\mathbf{I} - \mathbf{HM}\right)\Delta\mathbf{z}\|_2 \le \epsilon$.

B. ATTACKING ELECTRICITY MARKET

Based on (4), the line flow changes in the RT market is due to changes in loads or generations or both. So, the attacker needs to decide which elements of ΔZ in (5) must be compromised such that its profit in (3) can be maximized. Consequently, the attacker has two choices: whether let the m^{th} measurements carry the real data Δz_{mt}^1 or compromise the m^{th} meters Δz_{mt}^2. Mathematically speaking, the attacker has a binary decision variable (u_{mt})

$$u_{mt} = \begin{cases} 1, & \text{the } m^{th} \text{ meter is hacked} \quad (\Delta z_{mt} = \Delta z_{mt}^2), \\ 0, & \text{the } m^{th} \text{ meter is not-hacked} \quad (\Delta z_{mt} = \Delta z_{mt}^1). \end{cases} \quad (6)$$

Regarding (6), there can be two cases for the attacker as following.

Case 1 (The m^{th} Meter is Not Compromised): In this case, the real data values can be estimated using

$$\Delta z_{mt}^1 = \sum_{n=1}^{N_{bus}} GSF_{n-m}\left(\Delta p_{nt} - \Delta d_{nt}\right) \le 0; \quad \xi_{mt}^{(1)'}\forall m \in N_{line}. \quad (7)$$

Case 2 (The m^{th} Meter is Compromised): In this case, the injected false data must follow the physical constraints of the power system. In other word, the m^{th} line flow has to be smaller than the difference between the maximum line capacity and the scheduled line flow in the DA market, given as

$$\Delta z_{mt}^2 + \Delta z_{mt}^1 + \frac{\theta_{nt} - \theta_{kt}}{X_m} - F_m \le 0\xi_{mt}^{(2)'}, \quad \forall m \in N_{line}. \quad (8)$$

C. ATTACKING METHODOLOGY

The attacker needs a comprehensive model which maximizes its profit in (3) and embeds the exact relation between the pair of traded powers $[p_{it}, d_{it}]$, DA and RT prices $[\lambda_{nt}, \lambda_{nt}^{RT}]$ and malicious data (compromised RTUs) Δz_{mt}^2 injected into the Supervisory Control And Data Acquisition (SCADA) system. Let's define $X_{DA} = [p_{it}, d_{jt}, \theta_{nt}]$ as the DA market state variables and $X_{RT} = [\Delta p_{it}, \Delta d_{jt}]$ as the RT market state variables, and then the corresponding attack model (bi-level optimization problem) is given as

$$\max R_{Att} = \sum_{i\in\Omega_n}\sum_{t}^{T}\left[\lambda_{nt}^{RT}\left(\Delta p_{it} - \Delta d_{it}\right) + \lambda_{nt}\left(p_{it} - d_{it}\right)\right] \quad (9a)$$

$$\text{s.t.} \begin{cases} \sum_{m=1}^{M} u_{mt} \le N_{comp}, & (9b) \\ \left\|\left(I - HM\right)\Delta z_{mt}^2\right\|_2 \le \Delta\epsilon, & (9c) \\ \left(\Delta p_{it} - d_{it}\right) = 0, & (9d) \\ \left(\Delta d_{it} - p_{it}\right) = 0, & (9e) \\ (7) \text{ or } (8) & (9f) \\ \left[X_{DA}, \lambda_{nt}, X_{RT}, \lambda'_{nt}\right] \in \arg\min C_{DA} + C_{RT} & (9g) \\ \text{s.t.} \begin{cases} (1b) - (1f) \\ (2b) - (2e) \end{cases} & (9h) \end{cases}$$

In the presented attack model, Eq. (9a) shows the attacker's objective function as the i^{th} participant in the DA and RT markets (the upper level problem). Since, the attacker has limited

resources to attack toward smart grid, Eq. (9b) is presented in which the total number of compromised RTUs are restricted to a predefined number (N_{comp}). To control the injected false data and avoid being detected by the ISO, the attacker needs to keep FDI below a predefined value, corresponding to the constraint in (9b) implies this constraint. In fact, to form the attack model, the attacker should determine the number of the RTUs (N_{comp}) for injecting the malicious data. Then, the corresponding compromised RTU will send the false data (Δz^2_{mt}) to the SCADA system. Eq. (9b), helps the attacker to decide which of the RTUs must be chosen based on the specified number of attacked RTUs in (9b). Since the optimization problem in Eq. (9) is run by the attacker, and (Δz^2_{mt}) is the false data injected, it is important for attacker to stay below the thresholds to avoid being detected by the system operator. In fact, the attacker knows that there are errors in the measurement devices (RTUs errors) and wants to keep the amount of manipulated data (change in the data) within acceptable range. In (9b) $\Delta\epsilon$ emphasizes that the attacker only cares about the resulted error caused by himself. This error ($\Delta\epsilon$) will increase by increasing the number of manipulated RTUs. Eqs. (9b) and (9b) also emphasize that amount of the traded power in the DA market must be equal with the traded power in the RT market. Eq. (9b) shows whether the attacker choose to attack the m^{th} RTU or not. Finally, Eq. (9b) states the physical layer of the power system as it is discussed in Sections II and III-B (the lower level problem). In other word, the key to remain undetected by the ISO, is to consider the ISO's perspective by the attacker. To consider the ISO's perspective, the DA and the RT markets equations must be considered. Thus, the attacker's decision-making problem turns into bi-level optimization problem. Using the dual formulation of (9b) and the strong duality condition [24], the set of equations in (9b) can be turned into the MPEC problem that will be discussed in the next subsection.

D. CORRELATION BETWEEN MARKET AND COMPROMISED DATA USING MPEC

Based on the presented model, the electricity state variables $\left[p_{it}, d_{it}, \lambda_{nt}, \Delta p_{it}, \Delta d_{it}, \lambda^{RT}_{nt}\right]$ and the attacker's decision-variables $\left[\Delta z^2_{mt}, u_{mt}\right]$ are embedded into (9b). Replacing the dual constraints (differentiating with respect to decision variables) with the strong duality condition (at the optimal point) of (9b) leads to a MPEC problem. It must be mentioned that the strong duality for the power system physical constraints are already proven in [24], [25], and [29]. To be specific first we need to introduce the Lagrangian multipliers and Lagrangian function. In this respect, considering $\Gamma = \left[\lambda_{nt}, \xi^t_{nk}, \phi^{max}_{it}, \phi^{min}_{it}, \mu^{max}_{it}, \mu^{min}_{it}, \eta^{max}_{nt}, \eta^{min}_{nt}\right]$ as the dual variables for the DA constraints in (1b) to (1f) and $\Upsilon = \left[\lambda^{RT}_{nt}, \phi^{max\prime}_{it}, \phi^{min\prime}_{it}, \mu^{max\prime}_{jt}, \mu^{min\prime}_{jt}, \xi^{(1)\prime}_{mt}, \xi^{(2)\prime}_{mt}\right]$ as the dual variables for set of the RT constrains in (2b)-(2d), (7) and (8), the corresponding Lagrangian function is given as

$$L = C_{DA} + C_{RT} + \Gamma\mathbf{G} + \Upsilon\mathbf{G}', \tag{10}$$

where $\mathbf{G}$ and $\mathbf{G}'$ are the sets of the DA and RT constraints, respectively, as

$$\mathbf{G} = \begin{bmatrix} \sum_{i\in\Omega_n} p_{it} - \sum_{j\in\Omega_n} d_{jt} - \sum_{k\in\Pi_n}\left(\frac{\theta_{nt}-\theta_{kt}}{X_{nk}}\right) \\ F_{nk} - \frac{\theta_{nt}-\theta_{kt}}{X_{nk}} \\ p_{it} - p^{min}_i \\ p^{max}_i - p_{it} \\ d_{jt} - d^{min}_{jt} \\ d^{max}_{jt} - d_{jt} \\ \theta_{nt} + \pi \\ \pi - \theta_{nt} \end{bmatrix}, \tag{11}$$

and

$$\mathbf{G}' = \begin{bmatrix} \sum_{i\in\Omega_n} \Delta p_{it} - \sum_{j\in\Omega_n} \Delta d_{jt} \\ \Delta p_{it} - \Delta p^{min}_{it} \\ \Delta p^{max}_{it} - \Delta p_{it} \\ \Delta d_{jt} - \Delta d^{min}_{jt} \\ \Delta d^{max}_{jt} - \Delta d_{jt} \\ \sum_{n=1}^{N_{bus}} GSF_{n-m}\left(\Delta p_{nt} - \Delta d_{nt}\right) \\ \Delta z^2_{mt} + \Delta z^1_{mt} + \frac{\theta_{nt}-\theta_{kt}}{X_m} - F_m \end{bmatrix}. \tag{12}$$

Since both $\mathbf{G}$ and $\mathbf{G}'$ define a set of linear equations with respect to the state variables and parameters, they can be decomposed as fixed and variable parts as below;

$$\mathbf{G} = \mathbf{G_{var}} + \mathbf{G_{fix}} = \mathbf{G_{var}} + \begin{bmatrix} 0 & F_{nk} & -p^{min}_i & p^{max}_i & -d^{min}_{jt} & d^{max}_{jt} & \pi & \pi \end{bmatrix}^T \tag{13}$$

$$\mathbf{G}' = \mathbf{G}'_{\mathbf{var}} + \mathbf{G}'_{\mathbf{fix}} = \mathbf{G}'_{\mathbf{var}} + \begin{bmatrix} 0 & -\Delta p^{min}_{it} & \Delta p^{max}_{it} & -\Delta d^{min}_{jt} & \Delta d^{max}_{jt} & 0 & -F_m \end{bmatrix}^T \tag{14}$$

Since the sets in (9a) to (9b) induce convex regions, the corresponding MPEC problem for the set in (9) is as follows

$$\max R_{Att} \tag{15a}$$

$$\text{s.t.} \left[X_{DA}, X_{RT}, \Delta z^2_{mt}, u_{mt}\right] \in \Xi, \tag{15b}$$

$$\nabla_{X_{DA}, X_{RT}} L = 0, \tag{15c}$$

$$C_{DA} + C_{RT} = \Gamma\mathbf{G_{fix}} + \Upsilon\mathbf{G}'_{\mathbf{fix}}, \tag{15d}$$

$$\left[X_{DA}, X_{RT}, \Delta z^2_{mt}, u_{mt}\right] \in \left\{G \cup G'\right\}. \tag{15e}$$

In the presented problem, Eq. (15a) shows the attacker's objective function as the achieved profit. Eq. (15b) states the upper level constraints in (9a) to (9b). Eq. (15c) represents the dual constraints corresponding to (9b). In fact, the differential of the Lagrangian function with respect to the DA and RT variables X_{DA}, X_{RT} results in the dual constraints of (9b). Eq. (15d) illustrates the duality constraint. Since the lower level problem in (9b) is a linear problem, the strong duality must be held at the optimal point. Mathematically, the primal function ($C_{DA} + C_{RT}$) must be equal with the dual function ($\Gamma\mathbf{G_{fix}} + \Upsilon\mathbf{G}'_{\mathbf{fix}}$). Finally, Eq. (15e) depicts the primal constraints (the DA and RT markets constraints in (1) and (2))

corresponding to (9b). Furthermore, the attacker needs to add its own constraints regarding the DA and RT market clearing procedure to the MPEC problem. In fact, the attacker must define the mathematical expression to show whether the m^{th} is compromised or not (choose between (7) and (8)). In this respect, we have

$$0 \leq \xi_{mt}^{(1)'} \leq M\,(1-u_{mt}) \tag{16a}$$

$$-M\,(1-u_{mu}) \leq \sum_{n=1}^{N_{bus}} GSF_{n-m}\,(\Delta p_{nt}-\Delta d_{nt}) \leq 0, \tag{16b}$$

$$0 \leq \xi_{mt}^{(2)'} \leq M\,(u_{mt}) \tag{16c}$$

$$-M\,(u_{mt}) \leq \Delta z_{mt}^{2}+\Delta z_{mt}^{1}+\frac{\theta_{nt}-\theta_{kt}}{X_m}-F_m \leq 0. \tag{16d}$$

Eqs in (16) show the relation between the attacker's decision variables and the electricity market state variables. If the attacker chooses to manipulate the m^{th} meter's data ($u_{mt}=1$), it means that Eqs. (16a) and (16b) are zero. Thus, the attacker injects malicious data Δz_{mt}^{2} and its corresponding state variable $\xi_{mt}^{(2)'}$ into the MPEC formulation. Indeed, instead of valid Eqs. in (16a) and (16b) for the RT market, faked Eqs. in (16c) and (16d) are injected into the market clearing process.

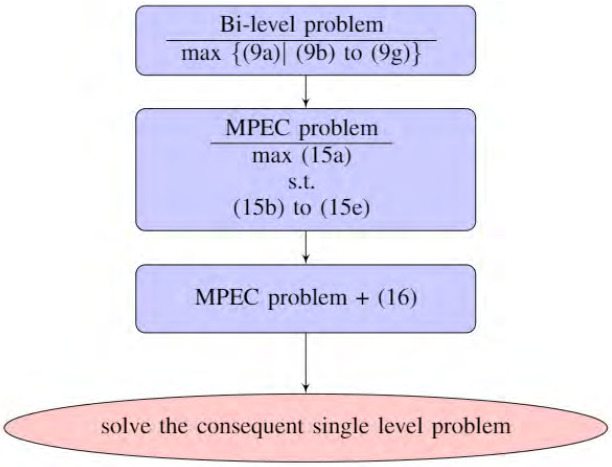

FIGURE 2. Outline of presented algorithm.

Fig. 2 shows the step by step procedure to create the final attack vector. Based on the proposed optimization model, there are three steps to create the optimal attack model as below;

1) The attacker needs to form the optimization problems in (9a) to (9b). To form this bi-level optimization problem the attacker needs to consider some assumption as follows. First, the attacker needs to specify the threshold ϵ that depends on the system operator but at most 5% error in the measurement is acceptable. Second, the attacker needs to determine the number of the RTUs that must be compromised. In this regard, the total possible RTUs (in the result section named as scenarios) can be considered. Third, the electricity network parameters and demand data can be obtained from online websites or the system operator annual reports.
2) The attacker needs to replace (9b) in the first step optimization problem with the corresponding MPEC problem (15a) to (15e). In fact, this step only replace the bi-level problem into the single level problem which is easier to solve. So far, except the attacker anyone else in the market can use this problem to maximize its profit. In other word, the MPEC problem in (15a) to (15e) is nothing but profit maximization problem for the smart grid electricity market participants.
3) The attacker decision-variables and constraints in (16a) to (16d) must be added to the second step to create the optimal attack problem. In fact, the attacker optimization problem is completed in this step. This step shows the manipulation equations that must be added to profit maximization problem. Basically, the attacker wants to see if the m^{th} RTU gets hacked with Δz_m amount of false data, how much the profit is gained. Thus, these attacks equations are added to main problem.

Finally, the presented algorithm will show the best possible solutions for the attacker to inject the false data into the RTUs. The counter-measurement process in this paper is to reinforce the weak points in the network (possible RTUs that can be manipulated by the attacker) using physical guarding or hashing methods to code and decode the signals from RTUs.

E. NASH EQUILIBRIUM (NE)

Before proceedings to the next section, in this part, the proof of NE existence is presented.

Lemma 1: (Debreu - Fan - Glicksberg) In a non-cooperative game with N player, S_i strategy set for the i^{th} player and $U_i(S_i, S_{-i})$ as the i^{th} player's utility function ($\forall i \in N$), if $S_i, \forall i \in N$ is compact and convex and $U_i(S_i, S_{-i})$ is continuous and quasi-concave then the game has at least on pure NE [26].

Lemma 2: If $f(x)$ is a linear function, then it is quasi-concave as well as quasi-convex. Using this theorem, we need to prove that the objective function in (15a) is a linear function or equals to a linear function.

Using the KKT conditions of (15c), we have

$$\nabla_{X_{DA},X_{RT}} L \times \begin{bmatrix} p_{it} & d_{jt} & \Delta p_{it} & \Delta d_{jt} \end{bmatrix} = 0 \times \begin{bmatrix} p_{it} & d_{jt} & \Delta p_{it} & \Delta d_{jt} \end{bmatrix}, \tag{17}$$

Further, Eq. (17) is rewritten as follows

$$\begin{bmatrix} p_{it} \times \lambda_{nt} \\ d_{jt} \times (-\lambda_{nt}) \\ \Delta p_{it} \times \left[\lambda'_{nt} - \sum_{m}^{N_{line}} GSF_{n-m}\xi_{mt}^{(1)}\right] \\ \Delta d_{jt} \times \left[-\lambda'_{nt} + \sum_{m}^{N_{line}} GSF_{n-m}\xi_{mt}^{(1)}\right] \end{bmatrix} = \begin{bmatrix} p_{it} \times \left[C_{it} + \phi_{it}^{max} - \phi_{it}^{min}\right] \\ d_{jt} \times \left[-B_{jt} + \mu_{jt}^{max} - \mu_{jt}^{min}\right] \\ \Delta p_{it} \times \left[C_{it}{}' + \phi_{it}^{max\prime} - \phi_{it}^{min\prime}\right] \\ \Delta d_{jt} \times \left[-B_{jt}{}' + \mu_{jt}^{max\prime} - \mu_{jt}^{min\prime}\right] \end{bmatrix}. \tag{18}$$

The right-hand side in (18) are nonlinear terms of objective function in (15a). If it can be proven that the left-hand side terms are linear, Eq. (15a) meets the condition in Lemma 2. To this end, since $[C_{it}, -B_{jt}, C_{it}', -B_{jt}']$ are given, and thus $[C_{it}, -B_{jt}, C_{it}', -B_{jt}'] \times [p_{it}, d_{jt}, \Delta p_{it}, \Delta d_{jt}]^T$ is linear term. Using complementary conditions, terms like $[(\phi_{it}^{max} - \phi_{it}^{min}) \times p_{it}]$ are equal with $[(\phi_{it}^{max} p_i^{max} - \phi_{it}^{min} p_i^{min})]$ which are linear terms. Thus, it can be concluded that the nonlinear terms in (18) are equal with the terms in (19). Thus, the terms in (18) are linear. Therefore, the objective function is quasi-concave. Multiplying (19) by [1, 1, 1, 1] results in the objective function.

$$\begin{bmatrix} p_{it}C_{it} + \phi_{it}^{max} p_i^{max} - \phi_{it}^{min} p_i^{min} \\ -B_{jt}d_{jt} + \mu_{jt}^{max} d_{jt}^{max} - \mu_{jt}^{min} d_{jt}^{min} \\ C_{it}'\Delta p_{it} + \phi_{it}^{max\prime} \Delta p_{it}^{max} - \phi_{it}^{min\prime} \Delta p_{it}^{min} \\ -B_{jt}'\Delta d_{jt} + \mu_{jt}^{max\prime} \Delta d_{jt}^{max} - \mu_{jt}^{min\prime} \Delta d_{jt}^{min} \end{bmatrix} \quad (19)$$

Definition 1: The set X is compact if and only if it is closed and bounded. Based on the definition, the set X is bounded when there exists a scalar C such that $\|X\| \leq C for all x \in X$. Also the set X is closed if it contains all of its upper and lower limit points.

Since the MPEC problem constraints in the (15) meet the required conditions in Definition 1, the MPEC set is compact.

Definition 2: A set $\mathbb{C}$ is convex if the line segment between any two points in $\mathbb{C}$ lies in $\mathbb{C}$, i.e., $\forall x_1, x_2 \in \mathbb{C}, \forall m \in [0, 1]$; $mx_1 + (1 - m)x_2 \in \mathbb{C}$.

Using Definition 2, lines, hyper-planes, half-spaces and Euclidian balls ($\|x - x_0\|_2 \leq \epsilon$) are convex sets. So, the MPEC problem in (15) fits the required conditions.

IV. SIMULATION RESULTS

In this section, the proposed method is applied on the 5 bus Pennsylvania-Jersey-Maryland (PJM) system (see Fig. 3). The proposed method can be extended to the bigger electrical systems as far as mixed integer programming solvers can handle the optimization problem in set of Eqs. in (15) (in this paper CPLEX solver is used). The fact is that the attacker does not have access to all measurement systems (including RTUs and PMUs in the electrical system) to inject the false data. Moreover, the power systems consist of different operation zones and different zones have different communication systems to send data over system operator. Therefore, it is almost impossible to attack an entire system. In other word, the attacker's possible targets are not wide and flexible and are limited within each zone ([27]). Meanwhile to run the optimization problem in (15), the attacker needs to estimate some network parameters including the transmission lines parameters and thermal limits. However, this information can be found publicly using IEEE standards or system operator's website. Increasing the network size to attack will lead to estimating too many parameters which consequently cause inaccurate solutions for the attacker. In fact, if the attacker estimates too many parameters as inputs to run the optimization problem in (15), the obtained FDI solutions might have lots of errors. Therefore, considering all the aforementioned problems, the FDI attacks usually are investigated within each zone not entire system. The results for five different scenarios are illustrated. In the first scenario the attacker just manipulates the measurement devices on one transmission line. Similarly, in the N_c^{th} ($N_c^{th} = 2, 3, 4, 5$) scenario, there are N_c lines under attack. Table 2 shows the load and generation information. It includes the maximum and minimum amounts of power generated and consumed in the DA and RT markets as well as bidding data. The information is common knowledge that can be easily obtained using online information provided by the ISO. Therefore, it is assumed that the attacker has the access to these information. Meanwhile, the network parameters including the transmission lines impedance and thermal capacity are estimated based on the voltage level of the network.

The network data including transmission lines capacity and generation shift factors (GSF) are listed in Table 3.

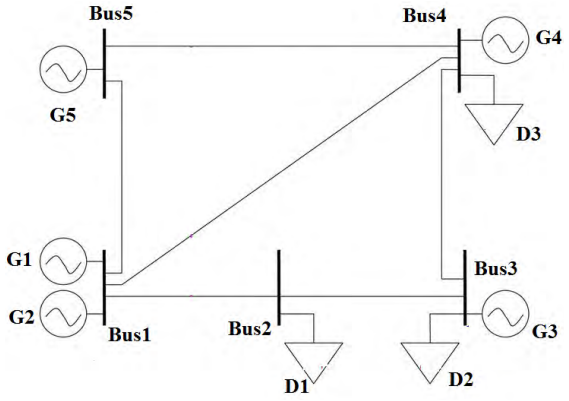

FIGURE 3. Outline of presented algorithm.

TABLE 2. Generation and load information for the DA and RT markets.

	Marginal Cost ($/MWh)	Maximum generation power (MW)	Minimum generation power (MW)	Maximum Load power (MW)	Minimum Load power (MW)
DA					
G1	14	40	0	0	0
G2	15	170	0	0	0
G3	30	520	0	0	0
G4	40	200	0	0	0
G5	20	600	0	0	0
D1	12	0	0	350	250
D2	26	0	0	350	250
D3	32	0	0	500	300
RT					
ΔG1	14	0.4	-0.4	0	0
ΔG2	15	1.7	-1.7	0	0
ΔG3	30	5.2	-5.2	0	0
ΔG4	40	2	-2	0	0
ΔG5	20	6	-6	0	0
ΔD1	12	0	0	3.5	-2.5
ΔD2	26	0	0	3.5	-2.5
ΔD3	32	0	0	5	-3

TABLE 3. Transmission line information.

	L_1	L_2	L_3	L_4	L_5	L_6
Bus1	0.1939	0.4376	0.3685	0.1939	0.1939	0.3685
Bus2	−.476	0.258	0.2176	0.5241	0.5241	0.2176
Bus3	−.349	0.1895	0.1595	−.349	0.6510	0.1595
Bus4	0	0	0	0	0	0
Bus5	0.1595	0.36	−.519	0.1595	0.1595	0.4805
Limit (MW)	320	192	192	192	192	192

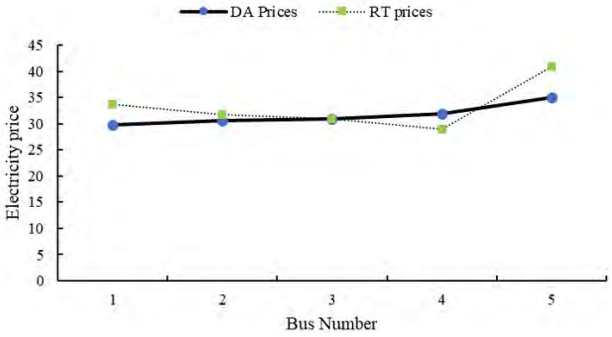

FIGURE 4. Electricity prices for DA and RT markets at different buses.

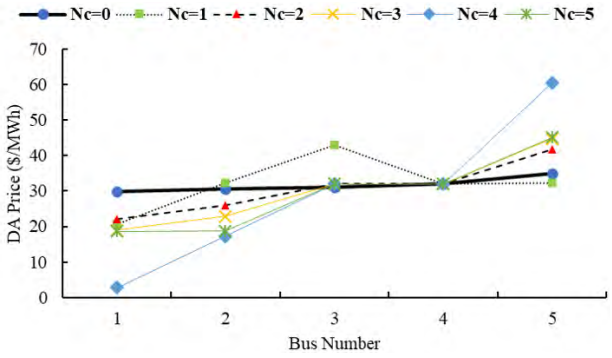

FIGURE 5. Price of electricity at DA market at the individual bus under attack.

To compute these factors, the attacker needs to estimate the transmission line's impedance based on the line capacity and voltage level which are public information. The procedure of obtaining GSFs from the lines' impedance is explained in [28].

A. BASELINE CASE-NO ATTACK TOWARD POWER SYSTEM

As a regular virtual bidder, if the attacker does not manipulate any measurement system ($N_{comp} = 0$), the electricity prices for the DA and RT markets will be as Fig. 4. If there is no attack in power system, line flows for the DA and RT markets are as Table 4. In the RT market, line 6 is congested and since line 6 is between buses 4 and 5 the electricity price difference between these two buses increased in the RT market (as it is shown in Fig. 4). The generation and demand powers for all

TABLE 4. Power Flow for the DA and RT markets (MW).

	L_1	L_2	L_3	L_4	L_5	L_6
DA	112.8	132.6	−14.3	−137.2	78.8	138.7
RT	146	177.1	−50	−104	80.9	192

TABLE 5. Generation and demand without attack (MW).

		DA				
		G_1	G_2	G_3	G_4	G_5
Generation	others	40	170	301.5	0	312.5
	attacker	0	0	0	0	0
		D_1	D_2	D_3		
Demand	others	250	250	300		
	attacker	24	0	0		
		RT				
		G_1	G_2	G_3	G_4	G_5
Generation	others	4	17	0	0	0
	attacker	4	17	3	0	0
		D_1	D_2	D_3		
Demand	others	0	0	0		
	attacker	0	0	0		

participants (named as others in table 5) in the market and also the virtual bidder (attacker) is depicted in Table 5. The amount of required power for the RT market is assumed to be 45 MW. It means that in the RT market, generation and loads must cover total amount of 45 MW.

Using Table 5 and Fig. 4, it indicates that when the attacker participates in the virtual bidding process, as a regular participant, it buys 24 MW in the DA market at $ 28.18 /MWh and sells it in the RT market at two different prices respectively $ 46.46 /MWh and $ 39.94 /MWh. Thus, the total profit for attacker (virtual bidder) is $ 404.28 without attacking the power system.

B. ATTACKING POWER SYSTEM MEASUREMENTS (RTUS)

The optimization problem in (15) to find FDI solutions, is a mixed integer linear programming (MILP). The problem solved in GAMS using CPLEX solver. Since the real-time markets are run with 15 minutes intervals, thus the solver must return the results in less than 15 minutes. Using CPLEX, even for big MILPs, the results would be produced within seconds. Specifically, for the proposed network, the solver returns the results in less than one second (the system CPU configuration is Intel Xeon E5-1603 v3 2.8 GHz, with 4 GB RAM).

In this section, different scenarios to attack power system measurements are considered. Since the PJM network consists of 6 lines, the number of potential lines with measurement to be attacked is 5. Figs. 5 and 6 show the

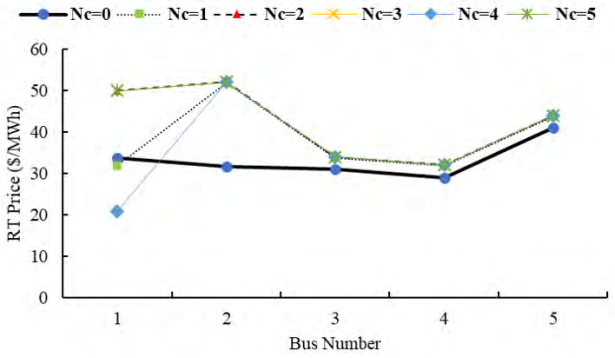

FIGURE 6. Electricity prices for RT market at different buses for different scenarios.

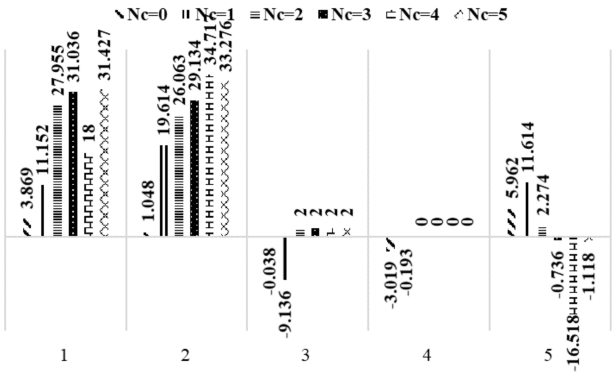

FIGURE 7. Electricity price differences for the DA and RT markets at different buses for different scenarios.

electricity prices for DA and RT under different assumptions ($N_{Comp} = 0$ for no attack to $N_{Comp} = 5$ for 5 measurements manipulation).

Generally, in the DA market, compromising more measurements decreases prices at buses 1 and 2, and increases price at bus 5. In other word, the path of network flow would change from buses 1 and 2 to bus 5. In contrast to the DA market, the attacker in the RT market tries to change the flow path from bus 5 to buses 1 and 2. Thus, the electricity prices at buses 1 and 2 are higher than buses 3, 4 and 5 in the RT market. In other world, if the attacker wants to participates in the DA market, he needs to buy electricity from buses 1 and 2 and sells the electricity to bus 5. In the RT market, the attacker needs to sell more electricity in the buses 1 and 2 rather than buses 3 to 5 and also buys electricity from bus 5 since the flow direction is changed in this bus.

Fig. (6) shows manipulating more than two measurements in the PJM system will not affect the RT prices a lot, however it will rather influence on the attacker's profit.

The electricity price differences between the DA and RT markets for different scenarios are illustrated in Fig. 7. As it is shown, buses 1, 2 and 3 have the highest variations when the attacker injects false data into measurement systems. On the other hand, smaller variation is observed under FDI attacks on buses 3 and 4. Hence, it can be concluded that buses 1, 2 and 3

TABLE 6. Power flow for the DA and RT markets (MW) for different scenarios.

N_{Comp}		L_1	L_2	L_3	L_4	L_5	L_6
1	DA	98.7	157.7	−25.4	−183.6	135.7	166.8
	RT	100.1	185.2	−12.1	192	183.1	192
2	DA	93	108.7	−161.6	−192	93	146
	RT	137.6	138.1	192.0	192	90.1	192
3	DA	93	108.7	−161.6	−192	93	146
	RT	137.6	192	192	192	90.1	192
4	DA	195.4	156.4	−120.7	−89.7	7.7	186
	RT	−320	192	192	192	20.9	−192
5	DA	99.3	132.8	−192	−185.7	109.3	177.2
	RT	143.1	192	192	192	192	−192

TABLE 7. The attacker's priority to compromise power flow data ($[u_{1t}, \ldots, u_{6t}]$).

N_{Comp}	L_1	L_2	L_3	L_4	L_5	L_6
1	0	0	0	1	0	0
2	0	0	1	1	0	0
3	0	1	1	1	0	0
4	1	0	1	1	0	1
5	0	1	1	1	1	1

need more protection compared with buses 3 and 4. The line flows for the DA and RT markets under different scenarios are depicted in Table 6.

For the first scenario (one measurement compromised), lines 4 and 6 are congested in the RT market. For $N_{Comp} = 2$, lines 3 and 4 are congested in the reverse direction with respect to the DA line flows. This path reversion in flow causes the largest difference between the DA and RT electricity prices which increases the attacker's profit. When the attacker decides to attack 3 measurement systems, lines 2, 3 and 4 are the best options to be congested. Compromising more measurements, it is still lines 3 and 4 that are among the congested lines. Thus based on table 6, it can be mentioned that lines 3 and 4 are relatively ideal to congest so as to increase the attacker's profit. Table 7 shows the attackers' priority ($[u_{1t}, \ldots, u_{6t}]$) in (6) considering different scenarios. As implied in 6, the results in Table 7 confirms that lines 3 and 4 are of the highest priorities for the attacker. If the attacker intends to compromise more than 2 measurements, lines 2 and 6 are of the second priorities for false data injection. Thus, based on information provided by Table 7, when the attacker knowledge and resources about network are limited the best choice is to manipulate the measurements on line 4 (first row in the table). In other word, by increasing the power flow measurement (based on Table 8) to 4.349, line 4 gets congested which cause profit for the attacker in the RT market.

The false data injected ($[\Delta z_{1t}, \ldots, \Delta z_{6t}]$ in (6)) into measurement system for different scenarios are listed in Table 8. The output generation power for the attacker in the DA market and required demand in the RT market, considering different

TABLE 8. False data injected into measurement system for different scenarios.

N_{Comp}	L_1	L_2	L_3	L_4	L_5	L_6
1	0	0	0	4.349	0	0
2	0	0	4.276	4.094	0	0
3	0	.539	4.276	4.094	0	0
4	−5.526	0	3.814	3.145	0	−4.292
5	0	0.293	4.576	4.04	0.84	−4.156

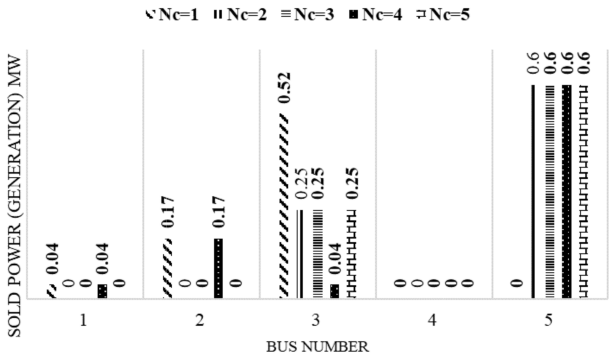

FIGURE 8. Electricity power sold by the attacker as virtual bidder in the DA market for different scenarios.

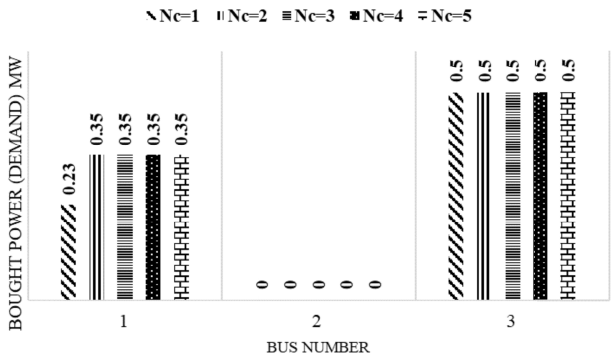

FIGURE 9. Electricity demand bought by the attacker as virtual bidder in the RT markets for different scenarios.

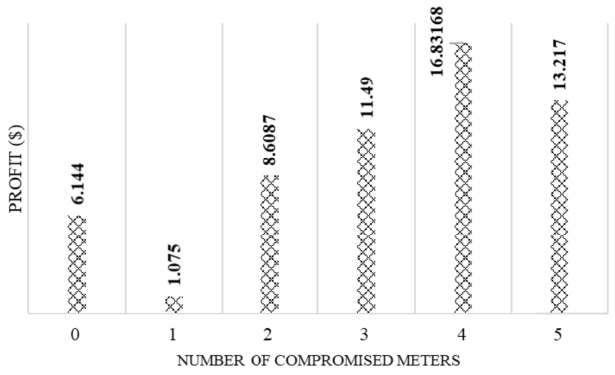

FIGURE 10. The attacker profits for different scenarios.

scenarios, are depicted in Figs. 8 and 9, respectively. Fig. 8 shows the electricity power sold by the attacker in the DA market for different scenarios. First scenario is for FDI toward one RTU and the last scenario is an attack toward five RTUs. Generally, it shows that the attacker tends to sell the electricity in the DA market at buses 3 and 5 for almost all scenarios. For the first scenario, the attacker prefers to sell the most power at bus 3 rather than 5. The reason is that based on Fig. 7 for the first scenario, the RT price will be less than the DA price. And the attacker would prefer to sell at higher price in the DA and buy at lower price in the RT market. Due to the same reason, in the fourth scenario, the most power sold by the attacker is at buses 1 and 2 (0.21 MW together) rather than bus 3 (0.04 MW). Also, for the first scenario, the attacker tends to sell power at buses 1 and 2 most of the time (respectively 0.04 and 0.17 MW). Furthermore, Fig. 9 depicts that in the RT market the attacker (virtual bidder) tends to buy the electricity from buses 2 and 4. This happens because the lines between buses 2 to 3 and bus 4 to 5 are always congested due to FDI by the attacker. In other words, the attacker practically divides the network into two zones (cheap demands and expensive demands). The cheap zone always is fed with local generation units and the attacker will buy them (electricity demands 1 and 3 in the figure). The expensive bus is fed through congested transmission lines (which is because of the attack). In Fig. 9 the electricity demand number 2 is the expensive demand to buy. It must be emphasized that both the generations and demands have reached their maximum limits (Eqs. (2c) and (2d)).

C. COUNTER-MEASUREMENT ACTIONS

In order to secure the power system measurements, in this section, the attacker profit for different scenarios is discussed. In other word, the sensitive measurement parts as well as the important state variables are introduced by analyzing the attacker's profit. Fig 10 shows the attacker profit compromising different measurement devices. The maximum profit occurs when the attacker compromises 4 measurement systems. In fact, based on network topology, the maximum total difference between the DA and RT market prices happens when 4 measurement devices are manipulated by the malicious data. However, it is to hard to manipulate all 4 measurements. It must be noted that the results are shown for the rewritten equation in (9b) as the equality equation as below

$$\sum_{m=1}^{M} u_{mt} = N_{comp}. \tag{20}$$

The attacker's profit is resulted from the difference between the DA and the RT prices by manipulating the RTUs data. In fact, by changing the number of manipulated RTUs, the inter-correlation between these two variables can be altered. However, increasing the number of attacked measurements (RTUs) generally causes more profit (based on the Fig. 10), but the inner correlation between the DA and the RT prices may affect this issue. To maximize profit, the attacker needs to divide the network into two zones: cheap zone and

expensive zone. Manipulating only one device, only leads to congestion of transmission line between bus 4 and 5 and doesn't result in the full congestion. Fig. 10 shows the results for the case of considering Eq. 20 as binding constraint. It means if we force the attacker to only manipulate one device, the question remains what is that RTU device? Therefore, in this condition the profit becomes the second priority for solver in order to meet the binding constraint (Eq. 20). Moreover, when the attacker considers only one device to attack, the amount of $\Delta\epsilon$ must be less than attacking toward two or three compromised RTUs. Moreover, by manipulating 5 RTUs the profit would be less than 4 compromised RTUs and that is because of the correlation between the DA and RT prices for buses 1, 2 and 5. Comparing Figs. 4 and 5 by attacking to 5 RTUs the difference between the DA and the RT prices has decreased (for buses 1 and 5 significantly and bus 2 a small value). This reason caused almost $3.5 less profit regarding 4 compromised RTUs. Mathematically speaking, by forcing the attacker to inject false data into all measurement devices in the sample network, the optimization problem leads to local minimum.

From Fig. 10 and Table 7, it can be seen that the attacker can increase its profit by compromising 2 measurements. Therefore, the best counter-measurement action is to increase lines 3 and 4 cyber-security. Also considering Fig. 7, electricity prices at buses 1, 2 and 5 are the most vulnerable state variables.

V. CONCLUSION

In this paper, a novel MPEC model to determine the optimal cyber attacks toward power system state estimation is proposed. The attacker participates in the DA and RT market as a virtual bidder. Meanwhile, the aim of the attacker is to inject false data into the measurement system so as to maximize its profit by trading in the DA and RT markets. The presented model considers the inter-correlation between the DA and RT state variables and the injected false data. The five bus PJM network is used to demonstrate the effectiveness of the proposed model. The results reveal the vulnerable lines and buses that are required to be more secure. Using a sensitivity analysis over the possible RTUs, the best choices for the attacker which exposes weak points of the power system are introduced. The presented optimization problem in this paper is deterministic and for the future studies, the stochastic model of the proposed method may be investigated. Due to uncertainty in the attacker's estimation for the network parameters and also, the DA and the RT electricity demands, the stochastic aspect of the model can be either two-stage stochastic optimization problem or robust methods (like chance constraint and Conditional Value at Risk modelling). Moreover, the nonlinear state estimation effects on the presented MPEC problem may be addressed. In fact, if the attacker doesn't consider the DC state estimation, then the proposed optimization problem, in this paper, needs to be modified to consider non-linearities in the decision-making process to solve the FDI problem.

REFERENCES

[1] R. Deng, G. Xiao, R. Lu, H. Liang, and A. V. Vasilakos, "False data injection on state estimation in power systemsâĂŤattacks, impacts, and defense: A survey," *IEEE Trans. Ind. Inform.*, vol. 13, no. 2, pp. 411–423, Apr. 2017.

[2] J. Hao, R. J. Piechocki, D. Kaleshi, W. H. Chin, and Z. Fan, "Sparse malicious false data injection attacks and defense mechanisms in smart grids," *IEEE Trans. Ind. Inform.*, vol. 11, no. 5, pp. 1–12, Oct. 2015.

[3] M. M. Pour, A. Anzalchi, and A. Sarwat, "A review on cyber security issues and mitigation methods in smart grid systems," in *Proc. SoutheastCon*, Charlotte, NC, USA, Mar. 2017, pp. 1–4.

[4] R. Tan *et al.*, "Modeling and mitigating impact of false data injection attacks on automatic generation control," *IEEE Trans. Inf. Forensics Security*, vol. 12, no. 7, pp. 1609–1624, Jul. 2017.

[5] M. Khalaf, A. Youssef, and E. El-Saadany, "Joint detection and mitigation of false data injection attacks in AGC systems," *IEEE Trans. Smart Grid*, to be published.

[6] A. Ameli, A. Hooshyar, E. F. El-Saadany, and A. M. Youssef, "Attack detection and identification for automatic generation control systems," *IEEE Trans. Power Syst.*, vol. 33, no. 5, pp. 4760–4774, Sep. 2018.

[7] M. Esmalifalak, L. Liu, N. Nguyen, R. Zheng, and Z. Han, "Detecting stealthy false data injection using machine learning in smart grid," *IEEE Syst. J.*, vol. 11, no. 3, pp. 1644–1652, Sep. 2017.

[8] H. Sedghi and E. Jonckheere, "Statistical structure learning to ensure data integrity in smart grid," *IEEE Trans. Smart Grid*, vol. 6, no. 4, pp. 1924–1933, Jul. 2015.

[9] 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 Trans. Parallel Distrib. Syst.*, vol. 25, no. 3, pp. 717–729, Mar. 2014.

[10] J. Yan, Y. Tang, B. Tang, H. He, and Y. Sun, "Power grid resilience against false data injection attacks," in *Proc. IEEE Power Energy Soc. Gen. Meeting (PESGM)*, Boston, MA, USA, Nov. 2016, pp. 1–5.

[11] A. Ashok, M. Govindarasu, and J. Wang, "Cyber-physical attack-resilient wide-area monitoring, protection, and control for the power grid," *Proc. IEEE*, vol. 105, no. 7, pp. 1389–1407, Jul. 2017.

[12] M. N. Kurt, Y. Yilmaz, and X. Wang, "Real-time detection of hybrid and stealthy cyber-attacks in smart grid," *IEEE Trans. Inf. Forensics Security*, vol. 14, no. 2, pp. 498–513, Feb. 2019.

[13] B. Li, T. Ding, C. Huang, J. Zhao, Y. Yang, and Y. Chen, "Detecting false data injection attacks against power system state estimation with fast go–decomposition (GoDec) approach," *IEEE Trans. Ind. Inform.*, to be published.

[14] A. F. Taha, J. Qi, J. Wang, and J. H. Panchal, "Risk mitigation for dynamic state estimation against cyber attacks and unknown inputs," *IEEE Trans. Smart Grid*, vol. 9, no. 2, pp. 886–899, Mar. 2018.

[15] 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, Jan. 2015.

[16] G. Liang, S. R. Weller, F. Luo, J. Zhao, and Z. Y. Dong, "Generalized FDIA-based cyber topology attack with application to the australian electricity market trading mechanism," *IEEE Trans. Smart Grid*, vol. 9, no. 4, pp. 3820–3829, Jul. 2017.

[17] L. Xie, Y. Mo, and B. Sinopoli, "Integrity data attacks in power market operations," *IEEE Trans. Smart Grid*, vol. 2, no. 4, pp. 659–666, Dec. 2011.

[18] M. Esmalifalak, H. Nguyen, R. Zheng, L. Xie, L. Song, and Z. Han, "A stealthy attack against electricity market using independent component analysis," *IEEE Syst. J.*, vol. 12, no. 1, pp. 297–307, Mar. 2018.

[19] M. Esmalifalak, G. Shi, Z. Han, and L. Song, "Bad data injection attack and defense in electricity market using game theory study," *IEEE Trans. Smart Grid*, vol. 4, no. 1, pp. 160–169, Mar. 2013.

[20] W. W. Hogan, "Virtual bidding and electricity market design," *Electr. J.*, vol. 29, no. 5, pp. 33–47, Jun. 2016.

[21] PJM Interconnection. (2015). *Virtual Transactions in the PJM Energy Markets*. [Online]. Available: http://www.pjm.com/-/media/committees-groups/committees/mc/20151019-webinar/20151019-item-02-virtual-transactions-in-the-pjm-energy-markets-whitepaper.ashx,

[22] F. Li, Y. Wei, and S. Adhikari, "Improving an unjustified common practice in ex post LMP calculation," *IEEE Trans. Power Syst.*, vol. 25, no. 2, pp. 1195–1197, May 2010.

[23] A. L. Ott, "Experience with PJM market operation, system design, and implementation," *IEEE Trans. Power Syst.*, vol. 18, no. 2, pp. 528–534, May 2003.

[24] S. Ahmadian, B. Vahidi, J. Jahanipour, S. H. Hoseinian, and H. Rastegar, "Price restricted optimal bidding model using derated sensitivity factors by considering risk concept," *IET Gener., Transmiss. Distrib.*, vol. 10, no. 2, pp. 310–324, May 2016.
[25] C. Ruiz and A. J. Conejo, "Pool strategy of a producer with endogenous formation of locational marginal prices," *IEEE Trans. Power Syst.*, vol. 24, no. 4, pp. 1855–1866, Nov. 2009.
[26] H. Lu, "On the existence of pure-strategy Nash equilibrium," *Econ. Lett.*, vol. 94, no. 3, pp. 459–462, Mar. 2007.
[27] X. Liu and Z. Li, "False data attacks against AC state estimation with incomplete network information," *IEEE Trans. Smart Grid.*, vol. 8, no. 5, pp. 2239–2248, Sep. 2017.
[28] W. Y. Ng, "Generalized generation distribution factors for power system security evaluations," *IEEE Trans. Power App. Syst.*, vols. PAS-100, no. 3, pp. 1001–1005, Mar. 1981.
[29] S. Ahmadian, H. Malki, and A. R. Sadat, "Modeling time of use pricing for load aggregators using new mathematical programming with equality constraints" in *Proc. 5th Int. Conf. Control, Decis. Inf. Technol. (CoDIT)*, Apr. 2018, pp. 38–44.

SAEED AHMADIAN received the B.S. degree in electrical and power system engineering and the M.S. degree in electrical and computer engineering and minor in computer science from the Amirkabir University of Technology, Tehran, Iran, in 2013. He is currently pursuing the Ph.D. degree with the Department of Electrical and Computer Engineering, University of Houston. From 2013 to 2016, he was a Project Manager with the Niroo Research Institute (NRI) and conducted projects on the design of first extra high voltage (765 Kv) transmission and substation in Iran. He was a Principal Investigator (PI) or a Co-PI of more than $1 million at NRI. He published a best-selling book on the power system engineering in Iran, in 2014. The book is on power system design, analysis and protection.

His current research interests include smart grids cyber-security, optimization and analysis of electricity markets, the applications of machine learning and data science in smart grids, statistical, and mathematical methods for big data mining.

XIAO TANG received the B.S. degree in information engineering (Elite Class Named after Tsien Hsue-shen) and the Ph.D. degree in information and communication engineering from Xi'an Jiaotong University in 2011 and 2018, respectively. From 2015 to 2016, he was a Visiting Student with the Department of Electrical and Computer Engineering, University of Houston. He is currently with the Department of Communication Engineering, Northwestern Polytechnical University. His research interests include wireless communications and networking, resource management, game theory, and physical layer security.

HEIDAR A. MALKI received the B.S., M.S., and Ph.D. degrees in electrical engineering from the University of Wisconsin–Milwaukee. He joined the College of Technology, University of Houston, in 1991. He is currently an Associate Dean for academic affairs with the College of Technology and a Professor of the Engineering Technology Department. He teaches undergraduate and graduate courses in engineering technology. He spent two summers at the NASA Ames Research Center as part of the ASEE-NASA Fellowship program to work on vibration control of UH-60 helicopters, in 2002 and 2003. He has published more than 75 refereed journal and conference papers and secured more than $4.6 million of external grants as PI or co-PI. His research interests include smart grids, applications of neural networks, fuzzy logic controllers, and design of fuzzy logic controllers for industrial applications. He has been either Principal Investigator (PI) or Co-PI on several NSF grants and two Texas Workforce Commission projects. He received Fulbright International Education Administrative Program–France, in 2016. He was the General Chair for the 1997 ASEE/GSW Conference and one of co-chairs of 1997 ICNN-IEEE International Conference on Neural Networks. He was an Associate Editor for the IEEE TRANSACTIONS ON FUZZY SYSTEMS.

ZHU HAN (S'01–M'04–SM'09–F'14) received the B.S. degree in electronic engineering from Tsinghua University, in 1997, and the M.S. and Ph.D. degrees in electrical and computer engineering from the University of Maryland at College Park, College Park, in 1999 and 2003, respectively.

From 2000 to 2002, he was an R&D Engineer of JDSU, Germantown, MD, USA. From 2003 to 2006, he was a Research Associate with the University of Maryland. From 2006 to 2008, he was an Assistant Professor at Boise State University, ID, USA. He is currently a John and Rebecca Moores Professor with the Electrical and Computer Engineering Department and with the Computer Science Department, University of Houston, TX, USA. His research interests include wireless resource allocation and management, wireless communications and networking, game theory, big data analysis, security, and smart grid. He received an NSF Career Award in 2010, the Fred W. Ellersick Prize of the IEEE Communication Society in 2011, the EURASIP Best Paper Award for the *Journal on Advances in Signal Processing* in 2015, IEEE Leonard G. Abraham Prize in communications systems (Best Paper Award in IEEE JSAC) in 2016, and several best paper awards in IEEE conferences. He is currently an IEEE Communications Society Distinguished Lecturer, from 2015 to 2018. He has been 1% Highly Cited Researcher, since 2017 according to Web of Science.