Two-Stage Optimization Framework for Detecting and Correcting Parameter Cyber-Attacks in Power System State Estimation

Nader Aljohani
*Electrical & Computer Engineering
University of Florida
Gainesville, Florida, United States
nzjohani@taibahu.edu.sa

Arturo Bretas*†
†Distributed Systems Group
Pacific Northwest National Laboratory
Richland, Washington, United States
arturo.bretas@pnnl.gov

Newton G Bretas
Electrical and Computer Engineering
University of Sao Paulo
Sao Paulo, Brazil
ngbretas@sc.usp.br

***Abstract*—One major tool of Energy Management Systems for monitoring the status of the power grid is State Estimation (SE). Since the results of state estimation are used within the energy management system, the security of the power system state estimation tool is most important. The research in this area is targeting detection of False Data Injection attacks on measurements. Though this aspect is crucial, SE also depends on database that are used to describe the relationship between measurements and systems' states. This paper presents a two-stage optimization framework to not only detect, but also correct cyber-attacks pertaining the measurements' model parameters used by the SE routine. In the first stage, an estimate of the line parameters ratios are obtained. In the second stage, the estimated ratios from stage I are used in a Bi-Level model for obtaining a final estimate of the measurements' model parameters. Hence, the presented framework does not only unify the detection and correction in a single optimization run, but also provide a monitoring scheme for the SE database that is typically considered static. In addition, in the two stages, linear programming framework is preserved. For validation, the IEEE 118 bus system is used for implementation. The results illustrate the effectiveness of the proposed model for detecting attacks in the database used in the state estimation process.**

***Index Terms*—state estimation, two-stage optimization, cyber-physical security, false data injections, real-time monitoring**

I. Introduction

Monitoring the power grid in real-time has been achieved through the evolving developments on the Power System State Estimator (PSSE). The objectives of the PSSE is the estimation of complex voltages at each system bus, which are considered as the typical system states. The outcomes of the State Estimation (SE) process are used for other applications such as protection schemes, power flow, post-processing of gross errors etc. PSSE relies on two main inputs: a set of measurements and model parameters. The former is obtained through communicated reading of deployed meters in the field. Such measurements could be complex power flow, complex power injections and/or voltage magnitudes. The latter is a representation of the physical grid such as line impedance, transformers taps and capacitor banks. Any perturbation in the aforementioned inputs of PSSE could lead to a wrong estimate of the system states which could affect previously mentioned applications that depend on the outcomes of the SE routine. The research of power systems data trustworthiness has been invariably addressed through False Data Injection (FDI) models. However, there is limited research regarding FDI attacks on measurement model parameters. Model parameters or system database are typically stored in files and considered static. Hence, these database are considered error free while performing SE routine. In the evolving power grid system database is though inevitably exposed to cyber-threats. One form of such could be an outside user who can obtain access to the victim machine and modify/change the parameters in the database. This type of attack is called Remote-to-Local attack. The other form could be an individual user who is capable of gaining sufficient permissions to manipulate the database. This form of attack is classified as User-to-Root attack [1]–[5].

978-1-6654-8537-1/22/$31.00 © 2022 IEEE

The work on state estimation can be categorized into two main groups: DC and AC state estimation. While DC SE considers relationships between measurements and state variables are linear, the AC SE, on the other hand, uses non-linear algebraic models. In both models, research focus on FDI attacks on measurements. However, SE also uses measurement model parameters, which is static data, to compose the objective function. In the literature though, cyber-attacks in the measurement model parameters has been much less considered [6]. Current works on parameters estimation [7]–[11] use iterative solutions that could be costly in real-time applications. In addition, for the case of simultaneous attack on measurements and parameters, the order to which one to correct could affect the final estimate of the states.

In the author's previous work [12], a bi-level model for cyber-attack detection and correction was presented. However, the solution is guaranteed for a known X/R ratio. Any uncertainty in this ratio would lead to a solution that is inaccurate. Hence, the solution is sensitive to the setting of this ratio. In this work, the uncertainty in the model's constraints is addressed. Therefore, this paper highlights the following contributions:

- Estimating the ratios of the parameters of lines used in PSSE under uncertainty;
- Two-Stage optimization framework for measurements' model parameter estimation.

The organization of the paper is as follows. In Section II, a highlight on the theoretical background on relationships between measurements and system states is presented. The two-stage optimization formulation is illustrated in Section III. A case study is presented in Section IV while closing remarks are provided in Section V.

II. Background

Consider a π model of a line km that link node k and node m. The line admittance can be expressed as:

$$y_{km} = 1/z_{km} = 1/(r_{km} + jx_{km}) = g_{km} + jb_{km}. \quad (1)$$

Hence, with the line admittance y_{km}, the complex power flow in line km in its conjugate form can be written as follows [13], [14]:

$$S^*_{km} = P_{km} - jQ_{km} = E^*_k I_{km} + jb^{sh}_{km} V^2_k. \quad (2)$$

where the voltage at node k in complex form is E_k is , the current in the line km is I_{km}, Vk is the voltage magnitude of node k, and is the line shunt susceptance. Since I_{km} also equal to $(E_k - E_m)y_{km}$, the complex power could be written as:

$$P_{km} - jQ_{km} = y_{km}E^*_k(E_k - E_m) + jb^{sh}_{km}V^2_k. \quad (3)$$

By expressing E_k and E_m in term of states, i.e., voltage magnitude and angles, one can derive the following:

$$\begin{aligned} P_{km} - jQ_{km} = y_{km}V_k e^{-j\theta_k}(V_k e^{j\theta_k} - V_m e^{j\theta_m}) \\ + jb^{sh}_{km}V^2_k. \end{aligned} \quad (4)$$

where V_k and θ_k are the voltage magnitude and angle of bus k, V_m and θ_m are the voltage magnitude and angle of bus m respectively. By expanding (4) and decomposing it to real part and imaginary part, it would result in 6 and 5 respectively:

$$P_{km} = (V^2_k - V_kV_m \cos\theta_{km})g_{km} - (V_kV_m \sin\theta_{km})b_{km}. \quad (5)$$

$$\begin{aligned} Q_{km} =&(-V_kV_m sin\theta_{km})g_{km} + (-V^2_k)b^{sh}_{km} + \\ &(V_kV_m cos\theta_{km} - V^2_k)b_{km}. \end{aligned} \quad (6)$$

where g_{km} is the conductance, i.e., $\Re\{y_{km} = 1/z_{km} = 1/(r_{km} + jx_{km})\}$, and b_{km} is the susceptance, i.e., $\Im\{y_{km} = 1/z_{km} = 1/(r_{km}+jx_{km})\}$, V_k and V_m are voltage magnitudes of bus k and bus m respectively, and $\theta_{km} = \theta_k - \theta_m$ is the angle difference between bus k and bus m.

The expressions in (5) and (6) refers to the real power flow and reactive power flow in the line km. Injection measurements at bus k, i.e., real and reactive power injections P_k and Q_k respectively, can be written in terms of flow measurements as follows [14]:

$$P_k = \sum_{l\in\Omega_k} P_{kl}. \quad (7)$$

$$Q_k = -V^2_k b^{sh}_k + \sum_{l\in\Omega_k} Q_{kl}. \quad (8)$$

where Ω_k is the set of buses adjacent to the bus k, and b^{sh}_k is the bus shunt admittance at bus k. With the power flows from opposite direction, i.e., bus m to bus k: P_{mk} and Q_{mk}, the power losses of line km could be written as:

$$\begin{aligned} P^{loss}_{km} &= P_{km} + P_{mk} \\ &= g_{km}(V^2_k + V^2_m - 2V_kV_m cos\theta_{km}) \\ &= g_{km}|E_k - E_m|^2, \end{aligned} \quad (9)$$

$$\begin{aligned} Q^{loss}_{km} &= Q_{km} + Q_{mk} \\ &= -(V^2_k + V^2_m)b^{sh}_{km} \\ &- (V^2_k + V^2_m - 2V_kV_m cos\theta_{km})b_{km} \\ &= -(V^2_k + V^2_m)b^{sh}_{km} - (|E_k - E_m|^2)b_{km}. \end{aligned} \quad (10)$$

where E_m is voltage at node m and E_k is voltage at node k in their complex form. Since state estimator routine uses a set of measurements that are composed of real and reactive power flows as well as injections, any perturbation in the line parameters, i.e., g_{km}, b_{km} and b^{sh}_{km} would result in an inaccurate state estimation. Hence, a bi-level model for detecting changes in the SE database was proposed in [12]. The main goal of the such model is to provide a monitoring scheme for validating the database used by SE. These database are considered static.

III. Two-Stage Framework

To address the uncertainty in (x_{km}/r_{km}) ratio, consider the expression in (3). Let the term $E^*_k(E_k - E_m)$ be expressed as follow:

$$E^*_k(E_k - E_m) = V^{km}_{re} + jV^{km}_{im}. \quad (11)$$

By substituting (1) and (11) in (3), we derive the following:

$$P_{km} - jQ_{km} = (g_{km} + jb_{km})(V^{km}_{re} + jV^{km}_{im}) + jb^{sh}_{km}V^2_k. \quad (12)$$

Expanding (12) into real and imaginary parts and rearranging terms, we derive:

$$\begin{aligned} P_{km} - jQ_{km} = (V^{km}_{re}g_{km} - V^{km}_{im}b_{km}) \\ + j(V^{km}_{im}g_{km} + V^{km}_{re}b_{km} + V^2_k b^{sh}_{km}). \end{aligned} \quad (13)$$

Hence, real and reactive power flows from bus k to bus m can also be expressed as follows:

$$P_{km} = V^{km}_{re}g_{km} - V^{km}_{im}b_{km}. \quad (14)$$

$$-Q_{km} = V^{km}_{im}g_{km} + V^{km}_{re}b_{km} + V^2_k b^{sh}_{km}. \quad (15)$$

Dividing (14) and (15) by g_{km}, the following expressions can be obtained:

$$\frac{P_{km}}{g_{km}} = V^{km}_{re} - V^{km}_{im}\left(\frac{b_{km}}{g_{km}}\right). \quad (16)$$

$$\frac{-Q_{km}}{g_{km}} = V^{km}_{im} + V^{km}_{re}\left(\frac{b_{km}}{g_{km}}\right) + V^2_k\left(\frac{b^{sh}_{km}}{g_{km}}\right). \quad (17)$$

Dividing (17) by (16), the following expression is obtained:

$$\frac{-Q_{km}}{P_{km}} = \frac{V^{km}_{im} + V^{km}_{re}\left(\frac{b_{km}}{g_{km}}\right) + V^2_k\left(\frac{b^{sh}_{km}}{g_{km}}\right)}{V^{km}_{re} - V^{km}_{im}\left(\frac{b_{km}}{g_{km}}\right)}. \quad (18)$$

It is important to note that the term V_k in (18) is based on the measurement scenario of Q_{km}. If Q_{mk} is the case, then V_m should be used instead, since shunt susceptance is related to the imaginary part as shown in (13). Expanding and rearranging terms in (18), one can derive the following:

$$\begin{aligned}&\left(\left(\frac{Q_{km}}{P_{km}}\right)V_{im}^{km}-V_{re}^{km}\right)\left(\frac{b_{km}}{g_{km}}\right)-V_k^2\left(\frac{b_{km}^{sh}}{g_{km}}\right)\\&=\left(\left(\frac{Q_{km}}{P_{km}}\right)V_{re}^{km}+V_{im}^{km}\right).\end{aligned} \tag{19}$$

For a perturbation in the ratios in (18), we can obtain:

$$\alpha=\frac{V_{im}^{km}+V_{re}^{km}\left(\frac{b_{km}}{g_{km}}+d\right)+V_k^2\left(\frac{b_{km}^{sh}}{g_{km}}+d^{sh}\right)}{V_{re}^{km}-V_{im}^{km}\left(\frac{b_{km}}{g_{km}}+d\right)}. \tag{20}$$

where d is the deviation in the ratio b_{km}/g_{km}, d^{sh} is the deviation in the ratio b_{km}^{sh}/g_{km}, and α is the result due to the perturbation in the ratios. By expanding (20) and rearranging terms, one could derive:

$$\begin{aligned}&\left(V_{re}^{km}+\alpha V_{im}^{km}\right)\left(\frac{b_{km}}{g_{km}}\right)+V_k^2\left(\frac{b_{km}^{sh}}{g_{km}}\right)+V_k^2 d^{sh}\\&+\left(V_{re}^{km}+\alpha V_{im}^{km}\right)d=\alpha V_{re}^{km}-V_{im}^{km}.\end{aligned} \tag{21}$$

The ratios in (19) represents the (X/R) ratio of the line connecting bus k and bus m. These quantities are important and widely used in short circuit analysis. In other words,

$$\frac{b_{km}}{g_{km}}=-\frac{x_{km,L}}{r_{km}}. \tag{22}$$

$$\frac{b_{km}^{sh}}{g_{km}}=\frac{x_{km,C}}{r_{km}}. \tag{23}$$

The interpretation of (22) is the contribution to the tangent of the phase angle formed by the line impedance due to the line inductance while (23) is the contribution due to the line capacitance. Therefore, with the information of line characteristics (typically known for short circuit studies) and prior states estimate (at time t^-) from SE under no attack scenario, it will be possible to retrieve original measurement model parameters (i.e, g_{km}, b_{km} and b_{km}^{sh}) at time $t=\Delta t+t^-$. In [12], such model is developed. However, the (X/R) ratio is assumed to be correct and exact. Such information can be known with some uncertainty. In addition, it would make the model in [12] to be non-linear if it is modeled explicitly. To relax the assumption of the known (x_{km}/r_{km}), a two-stage optimization framework is presented in this paper. The obtained results in (19) and (20) can enable us to incorporate the uncertainty on the given ratio within the model presented in [12]. In particular, the model is divided into two-stages. In the first stage, an estimate of the ratio is obtained. In the second stage, the estimated ratio is used in the presented model in [12] instead of the known ratio. The two stage optimization framework would then be used as a sliding window to validate the database over time.

The overview of the framework is illustrated in Fig. 1 The State estimation process typically validate measurements only before running the state estimator routine. Database are considered as static and true. As measurements are subject

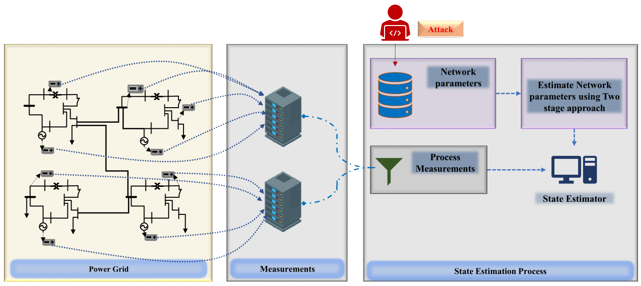

Fig. 1. Framework of Two Stage Optimization

to attack, the database of the network parameters are also subject to attack in the smart grid paradigm. For such, database validation is proposed in this paper as a pre-processing step for state estimator. The consideration in developing the model takes into account the ability to have it running in online platform. For such, the developed model process valuable information and data that are available in order to enhance the robustness of the existing state estimator software. The detailed structure of the two stage approach for validating database is illustrated in the following:

A. Optimization model for Stage I

In this stage, the primary goal is to obtain an estimate of the ratios (22)-(23). Instead of relying on an exact value that could be inaccurate or could be changed over time due to changes of line characteristic, the model would incorporate a variation in these ratios within a feasible region. Hence, a model pertaining to deviation in the ratios is developed as follows:

$$\min \quad d+d^{sh} \tag{24a}$$

$$\text{s.t.} \quad f_{1a}\left(\frac{b}{g}\right)-V_k^2\left(\frac{b^{sh}}{g}\right)=f_{1b} \tag{24b}$$

$$f_{2a}\left(\left(\frac{b}{g}\right)+d\right)+V_k^2\left(\left(\frac{b^{sh}}{g}\right)+d^{sh}\right)=f_{2b} \tag{24c}$$

$$\left(\frac{b}{g}\right)+d=\left(\frac{b}{g}\right)^t \tag{24d}$$

$$\left(\frac{b^{sh}}{g}\right)+d^{sh}=\left(\frac{b^{sh}}{g}\right)^t \tag{24e}$$

$$\left(\frac{b}{g}\right)^l\leq\left(\frac{b}{g}\right)\leq\left(\frac{b}{g}\right)^u \tag{24f}$$

$$\left(\frac{b^{sh}}{g}\right)^l\leq\left(\frac{b^{sh}}{g}\right)\leq\left(\frac{b^{sh}}{g}\right)^u \tag{24g}$$

$$d,d^{sh},\left(\frac{b}{g}\right),\left(\frac{b^{sh}}{g}\right)\in R \tag{24h}$$

where d and d^{sh} are deviations in the ratios of $\left(\frac{b}{g}\right)$ and $\left(\frac{b^{sh}}{g}\right)$ respectively. The parameters f_{1a} and f_{1b} are the evaluation of the corresponding terms in (19) given states, i.e., E_k and E_m, and measurements Q_{km} and P_{km} at time t^-. The parameters f_{2a} and f_{2b}, on the other hand, are based on the evaluation of (21) given states E_k and E_m at time t^-, and the current database to be validated i.e., g_{km}, b_{km}, and b_{km}^{sh} at time t. The ratios $\left(\frac{b}{g}\right)^t$ and $\left(\frac{b^{sh}}{g}\right)^t$ are obtained based on the database at time t. In this stage, the ratios $\left(\frac{b}{g}\right)$ and $\left(\frac{b^{sh}}{g}\right)$ are considered as decision variables to be estimated. The upper and lower limits on the ratios of $\left(\frac{b}{g}\right)$ and $\left(\frac{b^{sh}}{g}\right)$ are defined by the user, which represent the uncertainty on the ratios.

B. Optimization model for Stage II

In this stage, the presented model in [12] is performed. Since the model requires the (X/R) ratio, the estimated ratio from stage I is fed into the optimization model instead of a know ratio. The known ratio could be inaccurate. For such reason, a prior validation is needed, which is done through Stage I process. The optimization model at time $t = \Delta t + t^-$ for retrieving the original database values, which was at time t^-, is as follow [12]:

$$\min \quad dg_{km} + db_{km} + db^{sh}_{km} \tag{25a}$$

$$\text{s.t.} \quad dg_{km} = g^t_{km} - g_{km} \tag{25b}$$

$$db_{km} = b^t_{km} - b_{km} \tag{25c}$$

$$db^{sh}_{km} = b^{sh,t}_{km} - b^{sh}_{km} \tag{25d}$$

$$b_{km} = g_{km}\left(\frac{b}{g}\right)^{stageI} \tag{25e}$$

$$P_{km} = (f^g_{P_{km}})g_{km} + (f^b_{P_{km}})b_{km} \tag{25f}$$

$$Q_{km} = (f^g_{Q_{km}})g_{km} + (f^b_{Q_{km}})b_{km} + (f^{b^{sh}}_{Q_{km}})b^{sh}_{km} \tag{25g}$$

$$P_{mk} = (f^g_{P_{mk}})g_{km} + (f^b_{P_{mk}})b_{km} \tag{25h}$$

$$Q_{mk} = (f^g_{Q_{mk}})g_{km} + (f^b_{Q_{mk}})b_{km} + (f^{b^{sh}}_{Q_{mk}})b^{sh}_{km} \tag{25i}$$

$$P^{pert,loss}_{km} = P_{km} + P_{mk} + dP^{loss}_{km} \tag{25j}$$

$$dP^{loss}_{km} = (f^g_{P_{km}} + f^g_{P_{mk}})dg_{km} + (f^b_{P_{km}} + f^b_{P_{mk}})db_{km} \tag{25k}$$

$$P^{pert,loss}_{km} = (|E^{t^-}_k - E^{t^-}_m|^2)g^t_{km} \tag{25l}$$

$$Q^{pert,loss}_{km} = Q_{km} + Q_{mk} + dQ^{loss}_{km} \tag{25m}$$

$$dQ^{loss}_{km} = (f^b_{Q_{km}} + f^b_{Q_{mk}})db_{km} + (f^{b^{sh}_{km}}_{Q_{km}} + f^{b^{sh}_{km}}_{Q_{mk}})db^{sh}_{km} \tag{25n}$$

$$Q^{pert,loss}_{km} = (f^g_{Q_{km}})g^t_{km} + (f^b_{Q_{km}})b^t_{km} + (f^{b^{sh}}_{Q_{km}})db^{sh}_{km} \tag{25o}$$

$$g_{km}, b^{sh}_{km} \geq 0 \tag{25p}$$

$$b_{km} \leq 0 \tag{25q}$$

where g^t_{km}, b^t_{km}, and $b^{sh,t}_{km}$ are the parameters of the database at time t. The variables g_{km}, b_{km}, and b^{sh}_{km} are the true parameters of the database (which is unknown) and we are interested to obtain. The $(x_{km}/r_{km})^{(stageI)}$ is the ratio of the line obtained from Stage I. The function $f^{param}_{meas_{km}}$ evaluates the coefficient coupled with the given parameter $param$ from node m to node k for the corresponding measurement type $meas$ as (5) and (6). The evaluation of those functions are obtained using V and θ of the nodes linking line km at time t^-, which is at the previous time step. Using the current system database at t and the states at t^-, the losses in the line are evaluated: i.e., $P^{pert,loss}_{km}$ and $Q^{pert,loss}_{km}$. In (25f)–(25i), at time t^- for each type of measurements, it is required to have one estimate of each. The remaining two variables are free .

TABLE I
PERCENTAGE OF ERROR FOR STAGE I OUTPUT UNDER NO PARAMETER ATTACK

Ratio	**Uncertainty in ratios in %**									
	1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
$\frac{g}{b}$	0.408	0.816	1.223	1.631	2.039	2.447	2.855	3.262	3.670	4.078

IV. CASE STUDY

The presented framework is implemented on IEEE 118 bus system that is obtained from MATPOWER package [15]. MATLAB environment is used for the computation and Gurobi solver [16] is selected as the optimization tool for solving the model. The measurement set includes complex power flows, complex power injections, and voltage magnitude at all buses. The set of measurements consists of 712 measurements. For the standard deviations of measurements, 1% of their absolute values are considered for the weights in SE process. The simulation setup in this section are carried out on a Apple computer with the following specifications: macOS High Sierra 32 GB RAM 1876 MHz DDR3, 4 GHz Intel Core i7.

1) No parameter attack at time t: In this scenario, we want to validate the line parameters at time $t = t^- + \Delta t$. Both data at t and t^- are correct. The difference between the two samples is the loading condition change from t^- to t. Five lines are selected randomly. The selected lines are: $86-87$, $44-45$, $35-36$, $52-53$, and $31-32$. For this case, the uncertainty in the ratios is varied from $\pm1\%$ to $\pm10\%$ of the actual ratios. Hence, the optimization in Stage I is performed 10 times for each selected line. In other words, the constraints 24f and 24g are tested under different percentages of the actual values. The results are shown in TABLE I. In this table, the percentage of the error of the output of Stage I, which is the deviation of the ratio from its true value, is recorded and presented for each different scenario. The average value for the five lines in each scenario is calculated and reported in the table. As the relaxation increases, the observed percent of the error reached to 4% for $\left(\frac{b}{g}\right)$ when the ratios are accepted to be within $\pm10\%$ of their true values. It is worth mentioning that, under parameter attack, the attack to b and b^{sh} could be with same percentage, i.e., balance, or different percentages, i.e., unbalance [17]. Hence, the relaxed model enable us to use ratio in Stage II that is smaller than the uncertain one.

2) Parameter attack at time t: Similar to the previous scenario, the selected lines $86-87$, $44-45$, $35-36$, $52-53$, and $31-32$, which were random initially, are maintained. However, in this case, the parameters of those selected lines are attacked. The attack was chosen randomly to be between

TABLE II
PERCENTAGE OF ERROR FOR ESTIMATED PARAMETER USING [12]

Variable	Uncertainty in ratios in %									
	1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
g	21.91	24.00	41.68	42.22	42.74	43.26	43.77	44.27	44.75	45.23
b	21.66	23.55	40.32	40.43	40.53	40.63	40.73	40.83	40.92	41.02
b^{sh}	28.60	38.18	44.77	46.30	47.80	49.27	50.71	52.13	53.52	54.88

TABLE III
PERCENTAGE OF ERROR FOR ESTIMATED PARAMETER USING DECISION FROM STAGE I FOR ATTACK CASE

Variable	Uncertainty in ratios in %									
	1	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
g	0.274	0.551	0.830	1.111	1.395	1.681	1.970	2.261	2.555	2.852
b	0.118	0.2359	0.355	0.2753	0.596	0.718	0.841	0.965	1.090	1.216
b^{sh}	0.974	1.953	2.938	3.928	4.923	5.924	6.931	7.944	8.962	9.99

11% to 30% of the parameters' true values. Results of a comparative study with the state-of-the-art solution [12] are presented in TABLE III and TABLE II. The results in TABLE III are for the presented model in this paper, i.e., using two-stage approach, while the results in TABLE II are for the state-of-the-art solution [12]. The considered attack is unbalance. In other words, line parameters g_{km}, b_{km}, and b^{sh}_{km} are changed with different percentages. In the meantime, for evaluating the model in [12], the known ratio assumed to be the percentage associated with the scenario to be tested. As shown in TABLE III and TABLE II, under unbalanced attack of the parameters, which is considered the difficult case, and the uncertainty in the ratios of the line parameters, the proposed framework outperformed the model in [12].

For experimental purposes of the model, the range of the ratios are considered to up to 10%. This could be very high. However, the model is still able to enhance the database validation for the state estimator with minimizing errors in the database. In the meantime, the framework is still support parallel computation environment. In other words, the optimization framework can be computed for each line in parallel.

V. CONCLUSION

This paper presents a Two-Stage optimization framework for not only detecting FDI attacks into parameters of the SE process, but also correcting them. In the first stage, the ratios of line parameters are estimated. The estimated ratios are then applied to the model in the second stage. The model in Stage II requires X/R ratio of the line under consideration. Modeling X/R ratio explicitly turns the optimization problem to be non-linear. Hence, the presented framework relaxes the assumption of known ratios. In addition, the linear programming formulation of the original problem is preserved. Simulation results show that under uncertain ratios, the proposed two-stage model is able to reduce the error in correcting line parameters. The existing software of power system state estimation can be modified to incorporate the framework with simple modifications, enabling the two-stage approach to be adapted by utilities. Moreover, a solver without sophisticated features is sufficient to obtain a solution.

ACKNOWLEDGMENT

This material is based upon work partially supported by the National Science Foundation under Grant No. 1809739.

REFERENCES

[1] J. Raiyn *et al.*, "A survey of cyber attack detection strategies," *International Journal of Security and Its Applications*, vol. 8, no. 1, pp. 247–256, 2014.

[2] R. Gupta, S. Tanwar, S. Tyagi, and N. Kumar, "Machine learning models for secure data analytics: A taxonomy and threat model," *Computer Communications*, vol. 153, pp. 406–440, 2020.

[3] P. Sornsuwit and S. Jaiyen, "Intrusion detection model based on ensemble learning for u2r and r2l attacks," in *2015 7th international conference on information technology and electrical engineering (ICITEE)*. IEEE, 2015, pp. 354–359.

[4] P. G. Jeya, M. Ravichandran, and C. Ravichandran, "Efficient classifier for r2l and u2r attacks," *International Journal of Computer Applications*, vol. 45, no. 21, pp. 28–32, 2012.

[5] K. KR and A. Indra, "Intrusion detection tools and techniques–a survey'," *International Journal of Computer Theory and Engineering*, vol. 2, no. 6, p. 901, 2010.

[6] A. S. Bretas, N. G. Bretas, and B. E. Carvalho, "Further contributions to smart grids cyber-physical security as a malicious data attack: Proof and properties of the parameter error spreading out to the measurements and a relaxed correction model," *International Journal of Electrical Power & Energy Systems*, vol. 104, pp. 43–51, 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0142061518303946

[7] N. G. Bretas and A. S. Bretas, "A two steps procedure in state estimation gross error detection, identification, and correction," *International Journal of Electrical Power & Energy Systems*, vol. 73, pp. 484 – 490, 2015. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0142061515002495

[8] Y. Lin and A. Abur, "Fast correction of network parameter errors," *IEEE Transactions on Power Systems*, vol. 33, no. 1, pp. 1095–1096, 2018.

[9] A. Abur and Jun Zhu, "Identification of parameter errors," in *IEEE PES General Meeting*, 2010, pp. 1–4.

[10] Y. Lin and A. Abur, "A new framework for detection and identification of network parameter errors," *IEEE Transactions on Smart Grid*, vol. 9, no. 3, pp. 1698–1706, 2016.

[11] L. Zhang and A. Abur, "Identifying parameter errors via multiple measurement scans," *IEEE Transactions on Power Systems*, vol. 28, no. 4, pp. 3916–3923, 2013.

[12] N. Aljohani and A. Bretas, "A bi-level model for detecting and correcting parameter cyber-attacks in power system state estimation," *Applied Sciences*, vol. 11, no. 14, p. 6540, 2021.

[13] A. Monticelli, *State estimation in electric power systems: a generalized approach*. Springer Science & Business Media, 1999, vol. 507.

[14] A. Bretas, N. Bretas, J. B. London Jr, and B. Carvalho, *Cyber-Physical Power Systems State Estimation*. Elsevier, 2021, vol. 1.

[15] R. D. Zimmerman, C. E. Murillo-Sanchez, and R. J. Thomas, "Matpower: Steady-state operations, planning, and analysis tools for power systems research and education," *IEEE Trans. Power Syst.*, vol. 26, no. 1, pp. 12–19, Feb 2011.

[16] L. Gurobi Optimization, "Gurobi optimizer reference manual," 2021. [Online]. Available: http://www.gurobi.com

[17] T. Zou, A. S. Bretas, C. Ruben, S. C. Dhulipala, and N. Bretas, "Smart grids cyber-physical security: Parameter correction model against unbalanced false data injection attacks," *Electric Power Systems Research*, vol. 187, p. 106490, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0378779620302935