Smart Grids Cyber-Attack Defense: A Solution Based on an Incremental Learning Support Vector Machine

Helton do Nascimento Alves
Federal Institute of Maranhão, Electrical Engineering Dep.
São Luis, Brazil,
Email: helton@ifma.edu.br
Arturo S. Bretas
University of Florida, Electrical Engineering Department,
Gainesville, FL, USA.
Email: arturo@ece.ufl.edu
Newton G. Bretas
University of São Paulo, Electrical Engineering Dep.
São Paulo, Brazil
Email: ngbretas@sc.usp.br

***Abstract*— Recently false data injection attacks have defined a new class of gross errors in power system state estimation (SE). Considering that malicious intrusions can be launched on measurements, power system topology, transmission line parameters data, or even a simultaneous combination of them, it is very important to diagnosis the state estimator input data (SEID) to provide a proper defense. In the literature, this theme is seldom evaluated and strong assumptions are made on such approaches. In this paper the defense of the SEID against a false data injection using an incremental learning support vector machine (ILSVM) is presented. The proposed incremental learning algorithm adjusts, in real time, the SVM net with a fast re-training, aiming to evolve its network just as the cyber-attacks injection can do. A Monte Carlo simulation is used as reference considering different test scenarios in the IEEE 14-bus test system.**

***Index Terms*— False data injection, Incremental learning support vector machine, State estimation.**

I. INTRODUCTION

Data communication through the internet has lot of advantages, but inherently it increases the risk of exposure to cyber threats. . Cyber-attacks are a reality in smart grids. Most recently a cyber-attack in a power grid in Ukraine left about 225,000 customers without electricity [1]. Therefore, cyber security becomes a crucial resource to ensure the integrity and resilience of smart grid operations. In [2] is shown that a malicious attack in measurements set data is able to bypass classical bad data detection algorithms. In [3], similar result is obtained to malicious attacks in network data. Based on these considerations many studies have been made to analyze the power system vulnerabilities and countermeasures to cyber-attacks [2]-[9].

Power systems state estimation has three inputs data type, measurements, power system topology and transmission line parameters data. The false data injection (FDI) can be launched in any one or in both of them. Naturally, the more information one has about the possible cyber-attack the more efficient the countermeasure can be. A possible countermeasure against cyber-attack is to eliminate or correct the measurements where the cyber-attack occurred. This procedure is effective only if the cyber-attack is launched in measurements data, because the largest normalized residual test can be applied to identify the measurement with gross errors (GE) [2]. Therefore, to correct the proper input data under attack, it is necessary the correct and accurate detection and identification of the attack. In the literature, this theme is seldom evaluated. For that purpose, schemes based on the error [10] or the residual characteristic as well as on the Lagrange multipliers are used [11]-[14]. Strong assumptions are made on such approaches, as the need of high local redundancy and that the detection is only considered when just one specific attack type exists.

In the classical bad data detection, topological errors identification is usually based on the generalized state estimator [11]-[14]. The correction is performed modifying the status of the closest breaker/switch (on/off) from the branch identified with topological error. Therefore, the status of breaker/switch is considered as a source of the possible power system topological error. On the other hand, when a cyber-attack is launched on power system topology an exclusion/insertion/transference of a branch occurs directly on the branch line data instead on breaker/switches status. In this case, only the changing of status on breaker/switch will not correct the power system topology. Considering that the strategies of false data injection in state estimation inputs data can adapt and evolve overtime, it is very significant that the proposed defense solution has an on-line learning mechanism to face these issues.

More recently, using topological and geometrical approaches [10], [15]–[18], proposed solutions to compose the measurement error and then correct the measurement magnitudes for those measurements identified as containing a gross error. The simulations reported in these papers have shown many situations where the classical state estimation fails in the gross error detection as well as in the identification test but, using the composed measurement error proposal, besides detecting and identifying the measurements with gross errors correctly, it was able to estimate the measurement's errors and made the measurement's correction in an accurate way.

978-1-7281-0407-2/19/$31.00 ©2019 IEEE

In this paper a FDI attack detection and identification methodology, based on ILSVM, are presented. The paper main contributions are: (a) a geometrical view of state estimation is used to convert the measurement residuals to the errors, improving the classical bad data detection, even on areas with low local measurement redundancy; b) the SE data base under FDI is identified accurately by a SVM approach; c) an incremental Learning SVM for real-time learning is proposed. This approach aims to acquire new knowledge from possible new strategies of intruders to bypass the FDI detector.

II. Innovation Concept

In classical weighted least squared (WLS) estimator, considering the power system modeled as a set of non-linear equations [19]-[20], the objective is to find the best estimative for the *N*-dimensional estimated state vector $\hat{x}$, which minimizes the cost function $J(x)$:

$$z = h(x) + e \tag{1}$$

$$J(x) = [z - h(x)]^T W [z - h(x)] \tag{2}$$

where x is the state vector, $z \in \mathbb{R}^N$ is the measurement vector, N is the number of unknown state variables, $h: \mathbb{R}^N \to \mathbb{R}^m$ is a continuous nonlinear differentiable function, m is the number of measurements, e is the error vector assumed with zero mean and Gaussian probability distribution and W is the weighted matrix. The solution of the afore-mentioned minimization problem is obtained through the linearization of (2). At a certain operation point, yields,

$$\Delta z = H \Delta x + e \tag{3}$$

Where H is the matrix of first derivatives of the nonlinear functions of vector $h(x)$, known as the Jacobian, calculated at the point represented by the vector of estimated state variables.

Therefore, the solution can be obtained by:

$$\Delta x^k = [H^T W H]^{-1} H^T W \Delta z^k \tag{4}$$

The iterative process starts from an initial state vector and, at each iteration k; the corrections in the state variables Δx^k are obtained using (4). The vector of state variables update until a stopping criterion is satisfied.

This work is based on Innovation concept introduced by [10], [15]-[18]. This method is just briefly brought the concept and the formulation here. If the system described by (1) and (2) is observable, then the vector space of measurements R^m can be decomposed into a direct sum of two vector subspaces,

$$R^m = R(H) \oplus R(H)^{\perp} \tag{5}$$

in which the range space of H is an N-dimensional vector subspace into R^m and $R(H)^{\perp}$ is its orthogonal complement.

In the linear state estimation formulation, the solution of (3) can be understood as a projection of the measurement vector mismatch Δz onto $R(H)$. In [16] is defined the linear operator P that performs this projection, as follows:

$$P = H[H^T W H]^{-1} H^T W \ . \tag{6}$$

Based on (5) and (6), the linear formulation of the state estimation can be used to decompose the measurement error vector e into two parts: the detectable and the undetectable components. The detectable component is the residual measurement vector and undetectable component is orthogonal the detectable component and calculated as follow:

$$e_D = (I - P)e. \tag{7}$$

$$e_U = Pe. \tag{8}$$

$$\|e_i\|_W^2 = \|e_{Di}\|_W^2 + \|e_{Ui}\|_W^2 \tag{9}$$

In order to find the undetectable component and compose the measurement's total error for i[th] measurement, it is used *II*:

$$II_i = \frac{\|e_{Di}\|_W}{\|e_{Ui}\|_W} \tag{10}$$

A measurement with low Innovation Index (*II*) indicates that a large component of its error is not reflected in its residual as obtained by the classical WLS estimator. Consequently, even when those measurements have gross errors, their residuals will be relatively small. Based on (9) and (10) is possible to estimate the composed measurement error of the measurement i based on its standard deviation, as follow:

$$\|e_i\|_W^2 = \|e_{Di}\|_W^2 + \left\|\frac{e_{Di}}{II_i}\right\|_W^2$$

$$\|e_i\|_W = CME_i^N = \|e_{Di}\|_W \sqrt{1 + \frac{1}{II_i^2}} = \frac{r_i}{\sigma_i}\sqrt{1 + \frac{1}{II_i^2}} \tag{11}$$

Where r_i is the residue of measurement i, σ_i is the standard deviation of the measurement i.

In [10] is shown that Cyber-attack detection can be made through a Chi-square (χ^2) Hypothesis Testing (HT) applied to the composed measurement error, where bad data will be suspect if:

$$[CME^N]^T [CME^N] \geq \chi_{m,p}^2 \tag{12}$$

Where p is the detection confidence probability and m are the degrees of freedom.

In [15] is considered that in the detection stage it does not matter how reliable a measurement is because it is assumed that all them may contain errors. Therefore, it is attributed weights to the measurements as described by (13):

$$W_{ii} = 1/(0.1 z_i)^2 \tag{13}$$

III. Proposed Algorithm

The algorithm describing the cyber-attack defense proposed in this paper is shown below. More details are shown in next subsections.

(1) Collect input data;
(2) Perform WLS and calculate *II* and CME^N;
(3) Compute mean and standard deviation of CME^N and *II*;
(4) Define X=[CME^N_{mean} II_{mean} CME^N_{STD} II_{STD}];
(5) Classify X using ILSVM in: 0. no cyber-attack detected; 1. cyber-attack detected in real-time measurements data; 2. cyber-attack detected in line parameters data; 3. cyber-attack detected in power system topology data and 4. cyber-attack detected in simultaneous SEID.
(6) If the incoming set is suspicious of being a data that will acquire new knowledge for current SVM model, re-trains the SVM model and go to step 5. In this case, the incoming set produces an upgrade of the FDI detector training data while retaining the previous knowledge learned;

A. Data model

In order to show that the results of ILSVM do not depend on meter locations, the measurement plan (MP) is generated considering two groups of measurements: default set of measurements (DSM) and probabilistic set of measurements (PSM). DSM is the set of measurements provided by meters installed in substation control house. PSM is the set of measurements obtained by Monte Carlo Simulation (MCS) between all possible meters in order to complete the measurement plan. The MCS sampling is built to determine randomly various measurement plans and, per case, 400 MCS samples are being generated. A MP is only considered eligible if the system is observable, otherwise, a new measurement plan is generated.

The measurement values used in the tests from a load flow solution (z^{lf}) were obtained. The standard deviation σ is given by $\sigma = (pr*z^{fl})/3$ where *pr* is the meter precision, equal to 4%. Every MP has an associated random noise, so that they vary from $\pm 2\sigma$ of its original values.

Eight groups of samples were generated to test ILSVM, named as: G0 - samples without false data injection; G1- samples with an associated random multiple measurement cyber-attack that vary from 4σ to 9σ applied in a group of one to five measurements chosen randomly; G2 - samples with an associated random parameter cyber-attack that vary the magnitudes of the transmission line series parameters from 20% to 80% applied in a group from 1 to 3 transmission lines (chosen randomly); G3 - samples considering an associated random exclusion/transference topological cyber-attack, where a line was excluded or transferred; G4 - samples considering simultaneous cyber-attack described in G1/G2; G5 - samples considering simultaneous cyber-attack described in G1/G3; G6 - samples considering simultaneous cyber-attack described in G2/G3; and G7 - samples considering simultaneous cyber-attack described in G1/G2/G3.

B. Train data

In [15] is presented the proof that measurements errors are independent random variables having a Gaussian distribution with zero mean and known variance. Due to specific characteristic of *II* and CME^N for each SEID when under cyber-attack, it is possible use them to distinguish one from the other. Based on this observation, the descriptive statistics mean and standard deviation of *II* and CME^N are used as training data for the algorithm. For illustration, the mean and standard deviation of *II* and CME^N was evaluated considering 1000 samples of each group described in section 3A. Fig. 1 and Fig. 2 summarize the mean and standard deviation of *II* and CME^N obtained for those groups. In these tests, the measurement plan considers 95 measurements for IEEE 14 bus system. These results show that the mean and standard deviation of CME^N and *II* present different trends for each input data under cyber-attack, indicating they have potential as training data to identify the FDI point.

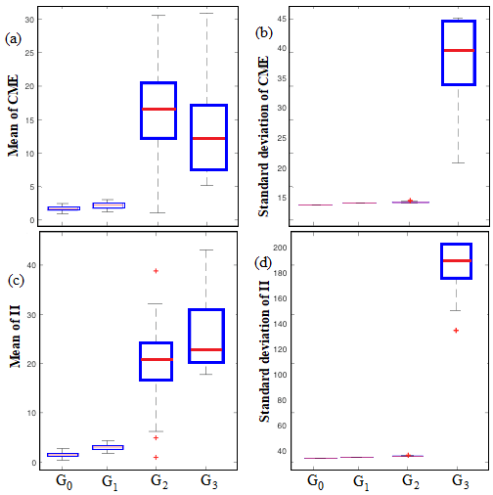

Figure 1. Mean and standard deviation of CME and II for G0, G1, G2 and G3.

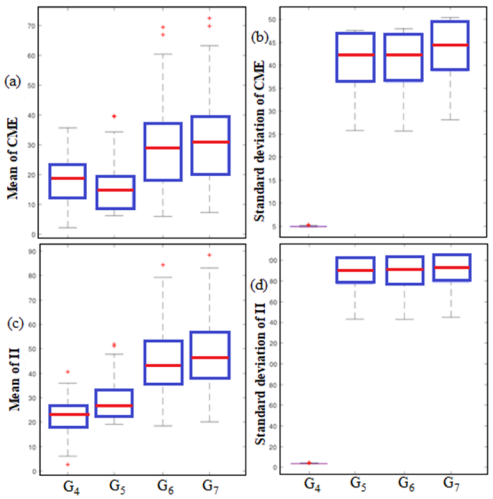

Figure 2. Mean and standard deviation of CME and II for G4, G5 G6 and G7.

C. Incremental learning SVM

SVM are a set of supervised learning methods used for analyzing patterns and classifying data [21]. SVMs seek to determine the optimization position of a linear hyperplane separator between binary data classes. The closest data points to the optimal hyperplane of each class are called Support Vectors (SVs). In situations where the data is not easily separable, SVM can use a soft margin, meaning a hyperplane that separates many, but not all data points. This process involves adding a non-negative slack vector variable ξ and a tunable penalty parameter C. ξ is upper bound on number of training errors and C is a parameter that controls the trade-off between margin and training error. However, in many real-world problems, the boundary between categories is nonlinear, not being well separated by a linear separating hyperplane even using a soft margin. SVMs try to solve this problem applying kernel functions to map the dataset into a higher dimensional feature space and to become linear separable in that new space.

Incremental learning algorithms SVM [22] search a continuous accommodation of new knowledge whenever new data becomes available (plasticity) but without forgetting the previously learned one (stability). A new model is generated when the incoming data set is considered a new knowledge for the current model. Several studies have focused on practical applications of online learning [23]-[25]. In essence, ILSVM takes advantage of an important property of SVM, i.e., the SVs form a succinct and sufficient set to ensure the same result as training on the complete example set that generate them. The algorithm proposed in [22] re-trains the SVM model using the current SVs and the new incoming instances considered as a new knowledge for the current model. Therefore, this re-training can be effective in accommodating the plasticity-stability in the new model and it is appropriate for on-line applications considering its low CPU time. The ILSVM is divided in two processes: off-line and on-line.

In the off-line process, large-scale data (thousands of samples) is used to train a multiclass SVM and generates the support vectors SV. The multiclass SVM approach is based on: one-versus-one coding design that defines $k(k-1)/2$ SVM, where k is the number of unique class labels. In this work they are used five identification error classes: 0 - cyber-attack no detected, 1 - cyber-attack launched in measurement input data, 2 - cyber-attack launched in parameter input data and 3 - cyber-attack launched in topology input data and 4 - cyber-attack launched in simultaneous input data; Error-correcting output codes (ECOC) model that reduces the problem of classification with three or more classes to a set of binary classifiers [22]. In this case is defined a coding matrix $M \in \{-1, 0, 1\}$ where the zero value indicates that a particular class is not considered in the training phase of a particular classifier. This fact provides a higher number of possible dichotomies that create different decision boundaries, allowing more accurate results. This characteristic of ECOC is very important in cyber- attack detection because if a new class of attack arrives, then the algorithm is able to adjust its SVM model to classify it; and binary loss function decoding scheme that determines how the predictions of the binary classifiers are aggregated. When a new data is to be classified, that function aggregates the binary losses for all $k(k-1)/2$ binary SVM, producing an average binary loss (ABL) per class. Therefore, an average binary loss vector ($ABLV$) is defined. The class with minimum average binary loss ($MABL$) is assigned to the new data. Using a result of ILSVM as an example, one sample with a cyber-attack launched in parameter input data (class error 2) produced $ABLV$=[class 0: 1.5, class 1: 0.83, class 2: 0.0, class 3: 1.0]. As expected, ILSVM assigned correctly the class error 2 (value zero).

The on-line process tracks if the incoming set is suspicious of being a New Knowledge (NK) for current SVM model. This tracking is made based on $ABLV$. As mentioned above, $MABL$ defines the class of the incoming set. If another value of $ABLV$ is too close of $MABL$, implies in a high uncertain about which class the incoming set actually pertains. In this case, this incoming set is considered as suspicious of being a NK. To evaluate this issue is defined three variables:

$$\begin{aligned} dABL_x &= |ABL_x - MABL| \\ dABL_y &= |ABL_y - MABL| \\ dABL_z &= |ABL_z - MABL| \end{aligned} \tag{14}$$

where $MABL$ is the smallest average binary loss, and ABL_x, ABL_y, and ABL_z are the values of $ABLV$ for the other classes.

If any $dABL < \gamma$, the incoming set is considered suspicious of being a NK and it participates of the re-training with its identification output exchange. For instance, if an incoming set produces an output identification in class 1 ($MABL$) and $dABL_3 < \gamma$, the incoming set participates of the SVM re-training with its class exchanged to 3. The new SVM model and new SVs obtained after re-training replaces the current SVM model and SVs respectively. Considering that ILSVM re-trains the SVM model using only the current SVs and NK, it is possible that one class presents more NK candidates than other one. This situation can increment more one class than other one, conducting the SVM model for a catastrophic forgetting [22]. In order to avoid this issue, a minimum training set is generated in the off-line process to participate in re-training process. The large-scale data used in the off-line process is divided in n data sets and every data set is trained in the SVM. The union of SVs obtained in every training form the minimum training set. The size of minimum training set range of 3 a 5% of the size of the large-scale data, depending of the number n of data sets chosen. The smaller the value of γ the rarer is the incoming set event. According to the experiments developing in this work, γ was defined as 0.02.

IV. NUMERIC TESTS

1000 samples of each cyber-attack group (G0-G7) are used as train data. All results reported are based on the expected value (mean) of each variable aggregated from MCS. The proposed algorithm is tested for 160000 cyber-attack samples with 20000 samples for each cyber-attack group. A Multilayer perceptron (MLP) is used as a benchmark. The IEEE-14 bus system is trained considering measurement plans with 95 measurements (DSM: active and reactive power injection and voltage magnitude in bus 1; PSM: 92 measurements defined by MCS considering a normal distribution function of active and reactive power injection and power flow). In Table I is reported the accuracy of the cyber-attack detection proposed in this work compared with classical largest normalized residual test detection (LNRT) and the detection proposed in [10]. Table I also reports the number of false positives and false negatives in the set of incorrectly classified cases The results reported in Table I show that the proposed cyber-attack detection identifies the features of the false data injection in a more appropriate way than the other methods.

TABLE I. ACCURACY OF THE CYBER-ATTACK DETECTION.

Cyber-attack detection	Accuracy of detection	False positive	False Negative
Proposed algorithm	97.4%	14,7%	85,7%
Proposed in [10]	89.7 %	11,5%	88,5%
Classical LNRT	77.3%	6,7%	92,3%

The accuracy of the proposed algorithm even with loss of at least 10% of meters along the power system is verified in Table II. These meters are chosen randomly among ones that do not undermine the observability of the system. The accuracy of proposed algorithm considering a power system reconfiguration is verified in Table III. In this case, the trained network is based on the former configuration. The results show that the proposed algorithm continues accurately detecting and identifying the false data injection. In relation to machine learning algorithm applied, the results show that SVM algorithm has higher generalization and better nonlinear modelling capability than MLP algorithm. Besides that, the MLP depends strongly on its initial weight and bias values. Therefore, to ensure that a neural network of good accuracy has been found, it is necessary to retrain several times.

TABLE II. ILSVM X MLP CONSIDERING POSSIBLE LOSS OF METERS.

Meters available	ILSVM accuracy			MLP accuracy		
	M (%)	P (%)	T (%)	M (%)	P (%)	T (%)
95	100.00	99.75	100.00	100.00	99.242	99.948
93	100.00	99.73	100.00	100.00	99.272	99.762
91	100.00	99.67	100.00	100.00	99.372	99.938
89	100.00	99.62	100.00	100.00	99.45	99.832
87	100.00	99.72	100.00	100.00	99.24	99.498
85	100.00	99.67	100.00	100.00	99.326	99.228

M- Measurements data; P- Parameter data; T- Topology data

In [10] an analytical method to identify the cyber-attack injection on state estimator input data is presented. This method assumes that at least one power injection of all buses and at least one power flow of all branches of the power system compose the measurement plan, but such a scenario cannot be granted. ILSVM algorithm identification does not depend of the location of the meters, but the measurement errors. Comparing the analytical method proposed in [10] for cyber-attack identification with the proposed algorithm, on average, the analytical method identified accurately just 22.3% the input data corrupted by a cyber-attack, against 99.6 % obtained by the proposed algorithm. The analytical method has a low accuracy because all measurements plans used in this work are generated randomly by MCS, therefore, they do not guarantee a favorable scenario for it.

To analyze the effectiveness of ILSVM, the experiment was designed for two situations: (i) the SVM model without incremental learning and (ii) the SVM model with incremental learning (ILSVM). The SVM model is the same in both situations. The same incoming sets with and without SEID under cyber-attacks was used in both cases and the results are shown in Fig. 3. The rate of prediction accuracy in ILSVM increased at least in 10% per class. These results show that the incremental update ensure more robust classification performance.

TABLE III. ILSVM X MLP CONSIDERING RECONFIGURATION.

Reconfiguration	ILSVM accuracy (%)			MLP accuracy (%)		
	M	P	T	M	P	T
No one	100	99.8	100	100	99.3	99.9
L Trans 13–14 → 11-14	100	99.8	100	100	99.5	99.4
L Elim 10-14	100	99.8	100	100	99.5	99.2
L Trans 4–5 → 3-5	100	99.7	100	100	99.5	99.1
L Elim 2-5	100	99.8	100	100	99.3	99.3

M- Measurements data; P- Parameter data; T- Topology data

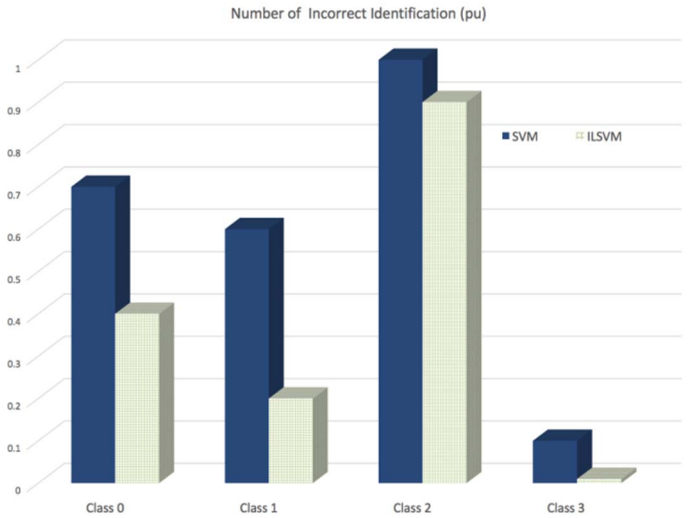

Figure 3. Number of incorrect identification obtained by SVM and ILSVM.

V. Conclusions

This research presents a cyber-attack FDI diagnostic system based on an innovation index and an incremental learning support vector machine. Descriptive statistics mean and standard deviation from variables derived of the innovation index were used as incoming sets to train the ILSVM. The method is able to detect, identify and correct malicious data attacks in smart grids. ILSVM considers potential cyber-attacks on measurements, parameters, topology or combination. In order to illustrate the performance of the algorithm, several experiments using IEEE-14-bus system were conducted. Results showed that the innovation index is a better parameter to evaluate the gross error than residual measurements. The paper simulations have shown the robustness of ILSVM even when there is loss of meters or system reconfiguration. Comparative tests demonstrated the increased accuracy of the proposed method in cases where the established method had failed. The incremental learning algorithm proposed demonstrated to be effective in accommodating the plasticity, stability and low CPU time in SVM model being appropriate for on-line applications. The results confirmed that the incremental update ensure more robust identification performance. Monte Carlo simulation application ensured the robustness of the method.

References

[1] A. H. Mohsenian-Rad and A. Leon-Garcia, "Distributed internet based load altering attacks against smart power grids," *IEEE Trans. on Smart Grid, 2(4),* pp. 667-674, 2011.

[2] E. Perez, "U.S. Investigators Find Proof of Cyberattack on Ukraine Power Grid", CNN, 2016, Available on: http://www.cnn.com/2016/02/03/politics/cyberattack-ukraine-power-grid/.

[3] S. Morgan, "Major Cyber Attack On U.S. Power Grid Is Likely", Forbes, 2016, Available on: http://www.forbes.com/sites/stevemorgan/2016/02/07/campaign-2016-major-cyber-attack-on-u-s-power-grid-is-likely/#131b743a610f.

[4] Y. Liu, M. K. Reiter and P. Ning, "False data injection attacks against state estimation in electric power grids," *in Proc. 2009 16th ACM Conference on Computer and Communications Security*.

[5] J. Kim and L. Tong, "On Topology Attack of a Smart Grid: Undetectable Attacks and Countermeasures," *IEEE Journal on Selected Areas in Communications, 31(7)*, pp. 1294 – 1305, July 2013.

[6] A. Ashok and M. Govindarasu, "Cyber-attacks on power system state estimation through topology errors", *in Proc. 2012 IEEE Power and Energy Society General Meeting,* pp. 1 –8.

[7] L. Liu, M. Esmalifalak, Q. Ding, V. Emesih and Z. Han, "Detecting false data injection attacks on power grid by sparse optimization", *IEEE Trans. Smart Grid 5(2)*, pp 612–621, 2014.

[8] Y. Chakhchoukh and H. Ishii, "Coordinated cyber-attacks on the measurement function in hybrid state estimation", *IEEE Trans. Power Syst. 30(5)*, pp 2487–2497, 2015.

[9] A. Ashok, M. Govindarasu and V. Ajjarapu, "Online detection of stealthy false data injection attacks in power system state estimation", IEEE Trans. Smart Grid PP(99), pp 1–11, 2016.

[10] A.S. Bretas, N.G. Bretas, B. Carvalho, E. Baeyens and P. P. Khargonekar,"Smart grids cyber-physical security as a malicious data attack: An innovation approach", *IEEE Trans. Power Syst. 149*, pp 210-219, August 2017.

[11] O. Alsaç, N. Vempati and A. Monticelli, "Generalized state estimation," *IEEE Trans. Power Syst., 13(3)*, pp. 1069–1075, 1998.

[12] E. M. Lourenço, K. A. Clements and A. S. Costa, "Bayesian-based hypothesis testing for topology error identification in generalized state estimation," *IEEE Trans. Power Syst., 19(2)*, pp. 1206–1215, 2004.

[13] E. Lourenco, E. Coelho and B. Pal, "Topology error and bad data processing in generalized state estimation," *IEEE Trans. Power Syst., 30(6)*, pp. 3190–3200, 2015.

[14] Y. Lin and A. Abur, "Highly efficient implementation for parameter error identification method exploiting sparsity", IEEE Trans. Power Syst., vol. 32, no. 1, pp. 734-742, Jan. 2017.

[15] N.G. Bretas, S.A. Piereti, A.S. Bretas, A.C.P. Martins, A geometrical view for multiple gross errors detection, identification, and correction in power system state estimation, IEEE Transactions on Power Systems 28 (3), 2128-2135, 2013.

[16] N.G. Bretas, A.S. Bretas and S.A. Piereti, "Innovation concept for measurement gross error detection and identification in power system state estimation", *IET Gener. Transm. Distrib. 5(6)*, pp 603–608, 2011.

[17] NG Bretas, AS Bretas, ACP Martins, "Convergence property of the measurement gross error correction in power system state estimation, using geometrical background", *IEEE Trans Power Syst 28*,3729–36, 2013.

[18] N.G. Bretas and A.S. Bretas, "A two steps procedure in state estimation gross error detection, identification, and correction", *Int. J. Electr. Power Energy Syst. 73 (December)* pp 484–490, 2015.

[19] A. Monticelli, State Estimation in Electric Power Systems: A Generalized Approach Massachusetts, USA, Kluwer Academic Publishers (1999)

[20] A. Abur and A. G. Exposito, Power System State Estimation: Theory and Implementation. New York: CRC Press, 2004.

[21] V. Vapnik, Statistical Learning Theory. John Wiley & Sons, 1998.

[22] N. A. Syed, H. Liu, and K. K. Sung, "Handling concept drifts in incremental learning with support vector machines", Proc. of the 5th ACM SIGKDD International Conference on Knowledge Discovery and Data mining (KDD), pp 317–321, 1999.

[23] Y. Lv, T. Yang and J. Liu, "An adaptive least squares support vector machine model with a novel update for NOx emission prediction", Chemometrics and Intelligent Laboratory Systems, Vol. 145, pp 103-113, 2015.

[24] L. Guo, J. H. Hao, and M. Liu, "An incremental extreme learning machine for online sequential learning problems," Neurocomputing, vol. 128, no. 27, pp. 50–58, 2014.

[25] X. Song, L. Jian and Y. Song, "A chunk updating LS-SVMs based on block Gaussian elimination method", Applied Soft Computing, Vol. 51, pp 96-104, 2017.