A Data-Based Detection Method Against False Data Injection Attacks

Charalambos Konstantinou, *Member, IEEE,* and Michail Maniatakos, *Member, IEEE*

***Abstract*—State estimation is one of the fundamental functions in power grid. In this paper, we address the vulnerability of state estimators to false data injection attacks (FDIAs) by proposing a data-driven anomaly detection algorithm. The proposed technique applies dimensionality reduction on grid measurements along with a density-based Local Outlier Factor (LOF) analysis and a feature bagging framework of combining predictions from multiple LOF outlier detection outputs. The work also addresses the handling of critical measurements. Instead of removing the attacked measurements, which may cause the system to become unobservable, we replace them by forecasted measurements. Numerical tests on IEEE 14-bus system verify the effectiveness and performance of the proposed method.**

***Index Terms*—Cybersecurity, false data injection attacks, state estimation, outlier detection, dimensionality reduction.**

I. Introduction

STATE Estimation (SE) is the method that estimates the power system operation based on the real-time network model. This is achieved by filtering measurement errors and leveraging redundant measurements. The SE measurements are collected by Remote Terminal Units (RTUs) and Phasor Measurement Units (PMUs) and transmitted through the Supervisory Control And Data Acquisition (SCADA) network to the control center. Grid operators use the SE results to adjust controls, estimate unmeasured variables, and detect faulty data. SE is classified into static SE (SSE)[1] and dynamic SE (DSE). While SSE operation is based on a single set of measurements, DSE takes into account the continuous variation of system loads, i.e., it keeps track of the real-time system states and predicts the state progressively in short time-steps.

A particular class against SE is False Data Injection Attacks (FDIAs). In a FDIA, an attacker compromises measurements from grid sensors in order SE to produce erroneous results and also bypass bad data detection modules [1]. The circumstances of the Ukraine blackout in 2015 underline the plausibility of common assumptions regarding the adversary's knowledge and capabilities in order to mount a successful FDIA [2]. This real-world example demonstrates the effects that a sophisticated FDI-based attack can cause on the grid infrastructure.

Several research efforts have highlighted the vulnerability of SE to FDIAs. Existing mitigation actions either require protection of a measurement set or the placement of PMUs [3]. The protection of measurements is primarily achieved by improving the security of associated sensors or masking the effect of attacks through redundant measurements. In addition to the impractical process of enhancing the security of existing devices or the deployment of redundant sensors, this approach also involves significant deployment costs and risks. Similarly, a major limitation to large-scale deployment of PMUs is their high capital cost. Also, PMUs can be exposed to GPS spoofing attacks [4]. Other defence methods against FDIAs are based on offline methodologies that often do not consider dynamic system models and require adversaries to possess system parameters and network topology information. In addition, existing data-driven solutions involve high computational cost (e.g, nonlinear optimizations) that constitutes a major obstacle for their use in practical real-time systems [3].

In this paper, a data-based approach is proposed to detect FDIAs within SCADA measurements. The method relies on an anomaly detection scheme that applies Local Outlier Factor (LOF) analysis techniques. We leverage the correlation among measurements over multiple time instants. Anomalous data points in the current SCADA measurement snapshot are identified by measuring their local deviation with respect to their neighbours. The algorithm requires no prior knowledge of system parameters or topology. In order to alleviate the high dimensionality of system data, the algorithm uses spectral methods in the unsupervised learning process to reduce dimensionality and transform the data to a new space.

The proposed scheme detects the presence of FDIAs. If there is no FDIA, the SE results can be reliably used for monitoring and control. If FDIA is detected, the processing of the faulty measurements must allow SE to run again and produce accurate results. In this work, the data outliers are marked as the attacked measurements. Typically, the attacked data are removed from the measurement set so that will not impact the final SE solution. However, the removal of the attacked measurements will cause the system to become unobservable and result in unavailable SE. In order to handle this problem, instead of removing these measurements, we replace them by forecasted data. For the DSE, a regression analysis method is utilized to calculate the power system state transition matrix [5]. For the SSE, we implement a forecasting algorithm that is based on the economic dispatch of the system, the load forecasts, and the generation schedules [6].

The rest of this paper is organized as follows. In Section II, the detection of FDIAs as outliers is presented. The proposed strategy and forecasting algorithms are shown in Section III. The effectiveness of our scheme is evaluated in Section IV.

This work was supported by the FSU planning grant and the NYUAD global fellowship programs.

Charalambos Konstantinou is with the Department of Electrical and Computer Engineering, FAMU-FSU College of Engineering and the Center of Advanced Power Systems, Florida State University, Tallahassee, FL 32310 USA (e-mail: ckonstantinou@fsu.edu). Michail Maniatakos is with the Division of Engineering and the Center for Cyber Security, New York University Abu Dhabi, Abu Dhabi, UAE (e-mail: michail.maniatakos@nyu.edu).

[1]Tracking SE (TSE) is a type of SSE in which the estimation starts from the last calculated state variables instead of a flat point.

II. Problem Foundation

In the linear (DC) form of SE, state variables include the voltage angles while bus voltage magnitudes are set constant and equal to one. Branch resistances and shunt elements are assumed to be negligible and angle differences between buses to be small. The relationship between measurements and state variables is based on the following linear function:

$$\boldsymbol{z} = \boldsymbol{H}\boldsymbol{x} + \boldsymbol{e} \tag{1}$$

where $\boldsymbol{H} \in \mathbb{R}_{m\times n}$ denotes the linear function determined according to the physical structure of the system. $\boldsymbol{x} \in \mathbb{R}_{n\times 1}$ is the vector of the state variables of a system with cardinality $\mathcal{S}$, where $n = 2S-1$, and $\boldsymbol{z} \in \mathbb{R}_{m\times 1}$ is the vector of SCADA measurements. $\boldsymbol{e} \in \mathbb{R}_{m\times 1}$ is the vector of measurement errors.

The measurements obtained from the SCADA system are typically updated once every 4-8 seconds, and form the measurement matrix Z, that is a $m \times \tau$ matrix representing the data collected from p SCADA sensors ($p \leq m$) within τ time instants. The measurement matrix Z can be decomposed into:

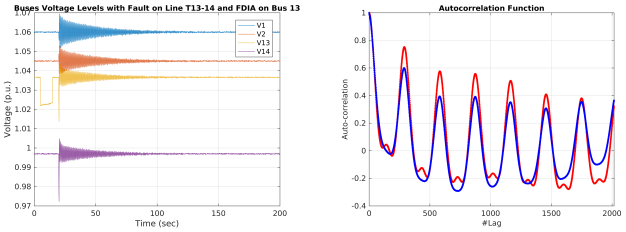
$$Z = M + A \tag{2}$$

where M is a matrix containing the true SCADA measurements without FDI data, and the non-zero elements of matrix A represent FDI measurements. A coordinated FDIA a_t at time instant t is considered to be undetectable if $\boldsymbol{z_t} = \boldsymbol{H}\boldsymbol{x_t} + \boldsymbol{r_t}$ and $\boldsymbol{\phi}(\boldsymbol{r_t}) \leq \boldsymbol{\epsilon}$, where $\boldsymbol{r_t}$ is the vector of measurement residuals at time instant t and $\boldsymbol{\phi}(\cdot)$ denotes the residual-based bad data detection criterion.

Most of the bad data detection schemes are based on the χ^2-test or the largest normalized residual method. Consequently, they rely on the network topology of the system, i.e., on the power flow equations and the SCADA measurements obtained from single time instants. Thus, a coordinated attack a_t can be successful if the relationship between x_t and z_t is satisfied without gross error. In order to overcome the above limitations, we consider the correlation of measurements among multiple sensors and multiple time instants.

Our threat model considers FDIAs affecting a limited number of measurements μ from q SCADA meters, where $\mu \leq m$ and $q \leq p$. Hence, A is a sparse matrix with only a few non-zero elements, and as a result the measurement matrix Z has a higher rank compared to matrix M. In addition, the low rank of matrix M under operating state indicates a high degree of correlation among attack-free measurements [7]. Since the rank of a matrix corresponds to the maximal number of linearly independent columns, the presence of FDIAs contributes to a higher number of linearly independent measurement vectors [8], i.e., attack-free SCADA measurements have a higher degree of correlation compared with FDI measurements.

The evaluation of cross-correlation seeks to identify the existence of similar behavior in the measurement patterns of different SCADA meters. This can be demonstrated in Fig. 1a in which a fault (as an eventful condition) causes the voltages at the same geographical region to have similar behavior due to their strong correlation. When a FDIA is present, the correlation of voltage tends to be weaker compared



(a) (b)

Fig. 1: IEEE 14 system data under FDIA: (a) $V_{1,2,13,14}$ after a fault ($20s$) is applied on the 13-14 line while a FDIA exists ($5s - 15s$) on V_{13}, (b) auto-correlation of Q_{51} measurements for 33.6 mins (SCADA measurements every 5 secs). If FDI measurements exist (red), the SE solution of δ_2 deviates from the true value by ≥ 0.02.

with measurements from sensors without attacks. We also consider the temporal patterns and periodicity in each sensor's measurements: we examine the auto-correlation function for each meter at different lags. Fig. 1b shows the auto-correlation function for a sensor reporting the reactive power flow between two buses. Data exhibit periodicity in their behavior. During FDIAs, the auto-correlation tends to have outlier behavior which deviates from the attack-free auto-correlation function.

The SCADA measurements in different time instants from individual meters demonstrate correlation and periodicity. Due to the degree of correlation in the scenario of FDIAs, we utilize density-based outlier detection methods to detect false data in SCADA measurements by measuring the distance metrics of outlier correlation behavior. The key idea is to track the dynamics of the current measurements (compared to historical measurements) and filter out inconsistent points due to their relative outlier correlation in both spatial and temporal space.

III. Methodology

This section explains the data-driven methodology for detecting FDIAs. The proposed scheme is independent of attacker resources and system configurations. It also covers broad attack scenarios without relying on infrastructure security mechanisms or the addition of new measurements into SE. The implementation outline is shown in Fig. 2. It makes use of the existing SCADA measurements in an anomaly detection module and utilizes existing work on predicting measurements as a way to process FDI-attacked measurements.

The measurements obtained from RTUs contain local analog and status quantities of each power system area. Typical status data are the status of breakers and disconnect switches which determine the network topology. The operating system state is constructed from analog data including *a)* voltage magnitudes, *b)* real and reactive power flows, and *c)* current magnitudes, by means of SE. The analog measurements are fed into our data-based anomaly detection module that leverages the cross- and auto-correlation among multiple time-instants, and detects local outliers using density-based techniques. Unsupervised density-based methods, however, are characterize by high computational complexity and low performance on high-dimensional spaces. Hence, our detection module first reduces the input dimensionality of data alleviating the effect of the curse of dimensionality and improving detection performance.

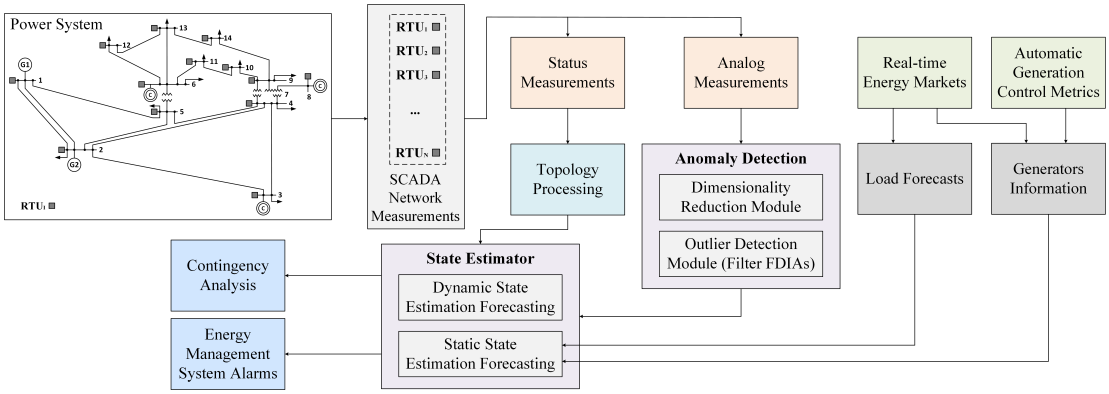


Fig. 2: Overview of the data-based approach for detecting false data injection attacks.

The estimation process includes static and dynamic algorithms. Dynamic-based SE is similar to the static one except that the state can be predicted from previous network state values so that an *a priori* value for the network state is available when the measurements are incorporated. In order to deal with undetectable FDIAs able to bypass bad data detection, both DSE and SSE algorithms take as input measurements the filtered data of the anomaly detection scheme. Also, the SE module uses the output of the topology processor to update its current network model. In the presence of PMUs, the SE could potentially utilize real-time data from PMU deployments obtained from a separate communication network.

The anomaly detection algorithm filters and removes faulty data so that they will not affect the final SE. However, it is required that the system state can be computed from existing data (system observability). Therefore, SE incorporates two forecasting algorithms from literature. For SSE, real-time short-term load forecast information (at the individual bus level) and generation schedule data which are available as part of the real-time energy markets are fed as inputs. The DSE forecasting algorithm relies on regression analysis to calculate the power system state transition matrix. This matrix is used to predict the system state which is subsequently corrected through extended Kalman filter. The result of the SE routine, as a key function of the Energy Management System (EMS), is then used to perform various tasks, e.g., contingency analysis which determines the ability of the grid to tolerate failures.

A. Anomaly Detection

Let $Z(t)$ be the matrix obtained at the time t, and each column denotes the measurements acquired from each sensor. Since the number of meter measurements m must be $\geq$ than the number n of state variables x in order for x at each bus to be determined (the system becomes observable), the number of dimensions would greatly affect the performance of the anomaly detection scheme. Thus, we map each column of $Z(t)$ to a dimensionality reduced subspace and then examine its outlier behavior according to a distance function that quantifies the similarity between each sensor at multiple time instants t.

We consider the following problem: given a high dimensional dataset $\boldsymbol{Z(t)} = [Z_1(t), ..., Z_p(t)]$ where $Z_i(t) \in \mathbb{R}^m$, how can we compute p corresponding output patterns $\boldsymbol{\psi} \in \mathbb{R}^d$ that provide a low dimensional representation of the dataset with $d \ll m$. In our work we examine different spectral methods for dimensionality reduction: Principal Components Analysis (PCA), Metric Multidimensional Scaling (MDS), and Laplacian Eigen-Maps (LEM) [9]. Given a set of data on n dimensions, PCA aims to find a linear subspace of dimension d lower than m. MDS is a form of non-linear dimensionality reduction which addresses the problem of constructing a configuration of t points in Euclidean space by using information about the distances between the t patterns. LEM is a graph-based method that computes the k-nearest neighbors of the input patterns, constructs a weighted graph based on these neighborhood relations, derives a matrix and then produces an embedding from the top or bottom eigenvectors of this matrix.

The outlier behavior of the dimension-reduced dataset is calculated using the Local Outlier Factor (LOF), a density-based method that computes the degree of being an outlier for each instance based on the local density around it [10]. LOF takes as input the dimensionality reduced dataset $\boldsymbol{\Psi} = \{\boldsymbol{\psi}^{(i)}\}_{i=1}^{D}$, and iterates over all the dataset instances D and the neighbor k values between k_{min} and k_{max}. The algorithm returns the scores for each data instance and the top n outliers are reported by sorting the scores. Due to the local approach, LOF allows to identify outliers in a dataset that would not be outliers in another area of the dataset. The resulting LOF quotient-values, however, are hard to interpret: in one dataset, a value of 1.1 may already be an outlier, in another dataset and parameterization (with strong local fluctuations) a value of 2 could still be an inlier. Also, LOF is often more applicable to low-dimensional vector spaces. Thus, we also employ LOF-based Feature Bagging (FB) to detect outliers in high-dimensional and noisy datasets [11].

In our work, FB for outlier detection runs LOF on multiple projections. Multiple classifiers are trained on random feature subsets, thus high dimensional data are more easily processed. Also, the accuracy of the learner is less affected by the noisy dataset features. The outlier scores are combined to obtain the final outlier scores for each instance using two approaches: breadth-first and cumulative-sum combinations. The breadth-first combine method *a)* sorts all outlier score vectors, *b)* takes the data samples with the highest score from all outlier detection algorithms, and *c)* appends their indices at the end

of the final index vector (and so on). Cumulative-sum sums up all the score vectors and returns the result as the final outcome.

B. Forecasting

In order to deal with the removal of attacked measurements while retaining the system observability, we implement two forecasting approaches from literature [5], [6]. The forecasted measurements are regarded as pseudo measurements to increase observability and enable both DSE and SSE to be performed successfully.

The implemented DSE algorithm is based on extended Kalman filtering and is comprised of three main stages: system state model identification, state forecasting, and state filtering. The "dynamic model" of the system and the forecasting of the state covariance matrix are shown below:

$$\begin{aligned} \tilde{\boldsymbol{x}}_{k+1} &= \boldsymbol{F}_k \hat{\boldsymbol{x}} + \boldsymbol{v}_k \\ \boldsymbol{R}_{\tilde{\boldsymbol{x}}_{k+1}} &= \boldsymbol{F}_k \boldsymbol{R}_{\hat{\boldsymbol{x}}} \boldsymbol{F}_k^T + \boldsymbol{R}_{\boldsymbol{v}_k} \end{aligned} \tag{3}$$

where k is the time sample, $\hat{\boldsymbol{x}} \in \mathbb{R}_{n\times 1}$ is the filtered/estimated state vector, $\boldsymbol{F}_k \in \mathbb{R}_{n\times n}$ is the state transition matrix, $\boldsymbol{v}_k$ represents the modeling uncertainties with a white gaussian noise with zero mean, $\boldsymbol{R}_{\tilde{\boldsymbol{x}}_{k+1}}$ is the forecasted state's error covariance matrix, $\boldsymbol{R}_{\hat{\boldsymbol{x}}}$ is the filtered state's error covariance matrix, and $\boldsymbol{R}_{\boldsymbol{v}_k}$ is the model error covariance matrix [5].

Our DSE forecasting scheme is based on how $\boldsymbol{F}_k$ is calculated. Similar to [5], we adopt a block diagonal form of $\boldsymbol{F}_k$ and regression analysis is utilized to identify the matrix parameters. We take into consideration that the effect of adjacent buses will result in a non-diagonal and non-constant $\boldsymbol{F}_k$ which is updated in certain time intervals. Hence, based on a number of historical terms of state variables, a sliding observation window W_{sl} is used to forecast the dynamic system state.

The SSE forecasting algorithm is based on the utilization of load forecasts and generation schedules [6]. The topology of the network as well as the real-time generation schedules and forecast information obtained from the real-time markets, are fed as an input to an economic dispatch algorithm shown in Eq. 4. The goal is to obtain the optimal generation setpoints for the next SSE interval.

$$\begin{aligned} & min \sum_{i=1}^{g} CC_i(P_{geni}) \\ & \sum_{i=1}^{g} P_{geni} - P_{loss}(P_{gen}) - P_{total\_load} = 0 \\ & P_{geni}^{min} \leq P_{geni} \leq P_{geni}^{max}, \forall i = 1, ..., g \\ & P_{geni} \geq 0, \forall i = 1, ..., g \end{aligned} \tag{4}$$

CC_i is the cost curve of each of the g generators, P_{geni} is the output of the i^{th} generator, P_{loss} are the transmission line losses, P_{total_load} is the total system load obtained from load forecasts, and P_{gen}^{min} and P_{gen}^{max} represent generators limits. The estimated generation setpoints are utilized in solving the SSE problem via the Dishonest Gauss Newton power flow iterative technique [12]. The result of the power flow solution provides the predicted state variables based on the load forecasts.

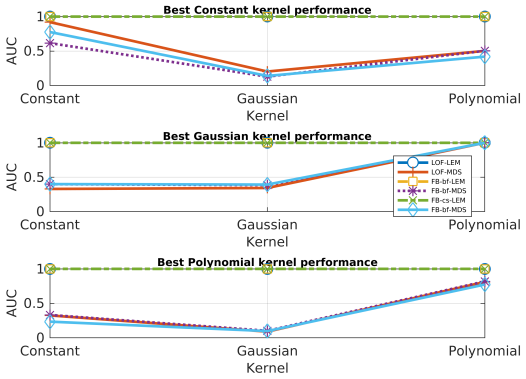


Fig. 3: AUC Vs. kernel type.

IV. Case Study

The proposed technique is tested on the IEEE 14 bus system using MATLAB. For the system data, we utilize the load data of New York (NY) state for 13 months provided by the NYISO [12]. We assume that the attacker aims to compromise SE for the voltage phase angle on bus 2, δ_2. Based on the power flow equations of the system the attacker requires to modify the measurements $\boldsymbol{z}^\star \in \mathcal{S}_2$, where $\mathcal{S}_2 = \{1, 3, 4, 5\}$. The effectiveness of the developed model is examined in the scenario which the adversary falsify the SCADA measurements in order to deviate the SE solution δ_2 from the true value by a range of $[0, 10]\%$. We test our technique for measurements collected within individual days where an average of 4.17% of the total measurements are falsified.

In rare-class classification problems, the classification accuracy is typically a good performance metric. However, since our class priors are heavily unbalanced, we use the area under the receiver operating characteristics (AUC) curve. AUC is a function of sensitivity and specificity, and thus the curve is insensitive to disparities in the class proportions. As for the training/validation set, we use $2/3$ of each dataset and the rest as the test dataset. For each dataset, we obtain 10 AUC values per outlier method by applying 5×2 cross validation on training/validation sets and report the performance on the test dataset with the optimum parameters (specific to each method).

For the anomaly detection algorithm, we apply LOF, FB-breadth-first (fbBF), and FB-cumulative-sum (fbCS) after the dimensionality reduction of datasets with PCA, MDS, and LEM. In our experiments, we use three kernels: constant (c), gaussian (g), and quadratic polynomial (p). The variance for the gaussian kernel, $Var(K)$, is set to σ which is half of the average of distances to k^{th} neighbors. We vary $Var(K)$ to $\sigma/2$, 2σ and 4σ. The value of neighbor count, $k \in \mathbb{Z}$, is varied in the interval $[3, 128]$. For the number of dimensions in the transformed space for PCA, we consider the values that explain 80%, 90%, 95% and 99% of the variance. As for LEM and MDS, we select 4 values linearly spaced in the interval $[2, d]$ where d is the input dimensionality. Based on the above, a set of experiments is performed to analyze the effect of the dimensionality reduction transformation and the density-based outlier algorithm by tuning k, d, and the kernel type.

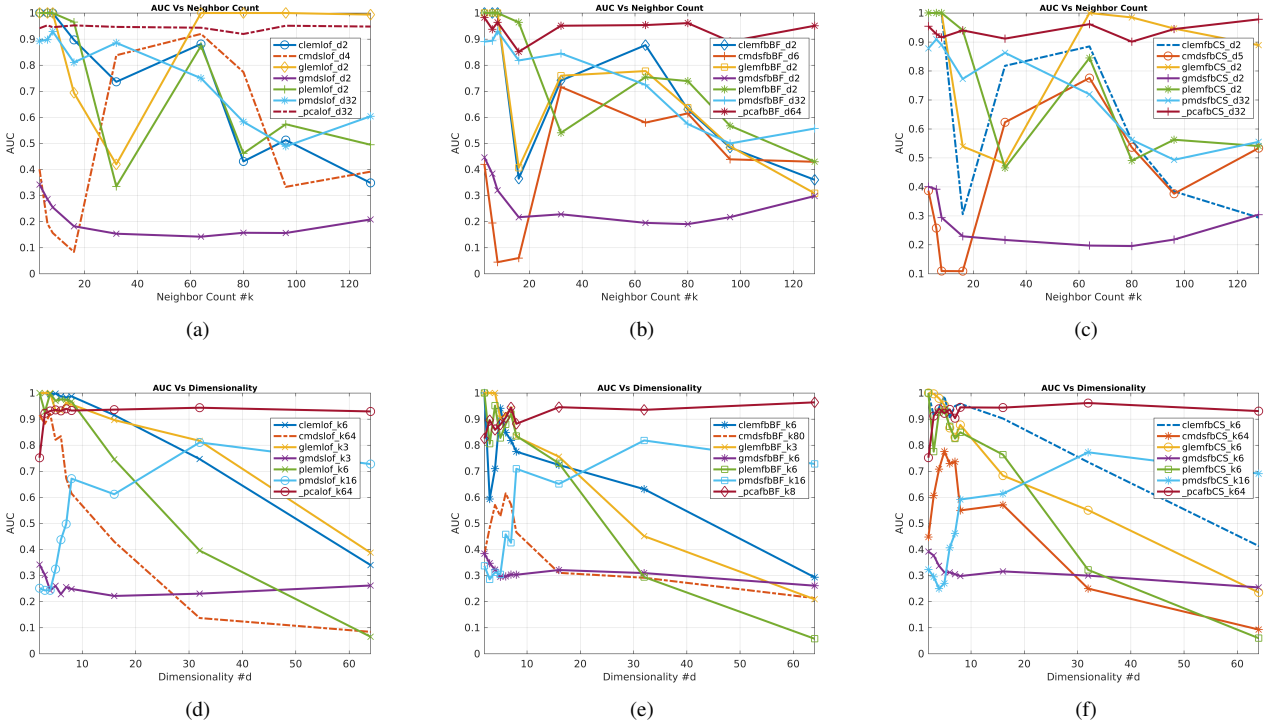


Fig. 4: (a), (b), (c): AUC vs. neighbor count, (d), (e), (f): AUC vs. input dimensionality.

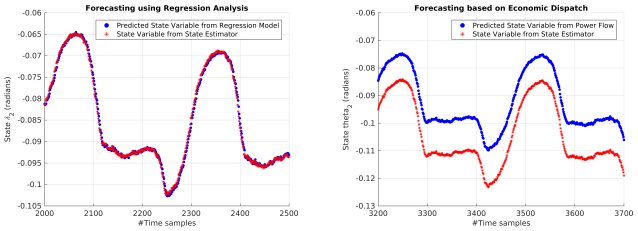


Fig. 5: δ_2 for 500 time samples using (a) DSE (mean square error $\Delta\delta = 0.0000193$) and (b) SSE ($\Delta\delta = 0.0108$) forecasting.

Fig. 3 shows the AUC values for each kernel on the dataset. The parameters of evaluation of other kernels in each graph are based on the k and d values that give the best performance of the evaluated kernel. The LEM-based transformation is less affected by the choice of kernel and exhibits an AUC close to 1. On the other hand, the performance of MDS varies depending on the kernel. For instance, with the quadratic kernel has better performance compared to the gaussian kernel.

The effects of neighbor count and dimensionality are presented in Fig. 4. First for each outlier detection method and kernel type, we find the best input dimensionality and calculate the AUC values for different neighbor counts (Fig. 4 (a)-(c)). It can be observed that the dimensionality reduction schemes exhibit unstable behavior as the performance relies heavily on the selection of a good neighbor count. Still, the overall performance of LEM is better compared to MDS. Similarly, we examine how the performance of the outlier detection methods is changed as the number of dimensions d varies (Fig. 4 (d)-(f)). For each method we present the AUC values obtained by the best performing kernels and neighbor count with different number of dimensions. The datasets are represented well in low dimensionality; adding more dimensions d allow to reach higher AUC values if $d < 10$, otherwise the performance decreases dramatically with higher dimensions.

Fig. 4 also reveals that the dimensionality reduction with PCA leads to robust AUC values in regards to the choice of k and d. This stems from the DC model utilized in this study as well as the linear interdependencies between various system measurements. Also, LEM provides better results compared to MDS: LEM maps similar instances to closer points in the new space, thus increasing the differences in the local densities and enable the detection of outliers. Of the three outlier detection techniques, no single one showed an overall performance superiority. The AUC values rely mostly on the dimensionality tranformation. In several cases, however, FB demonstrates performance improvement in higher dimensions (such datasets are more easily processed compared to LOF).

In order to demonstrate the effects of the implemented forecasting algorithms, we remove the falsified measurements and replace them with the forecasted values such that DSE and SSE to be performed successfully. For the dynamic regression model the sliding window observation interval is selected $W_{sl} = 1000$. Similar to [5], an increase in the sliding window length does not necessarily increase the forecasting accuracy drastically. For the forecasting algorithm based on the economic dispatch, the performance relies heavily on the availability of accurate real-time, short-term load forecasts. Fig. 5 presents the predicted and actual values of δ_2 for both

algorithms in random 500 time slot windows. The results validate the modeling of the system state utilizing previous historical values, load forecasts and generation schedules.

V. Conclusion

In this paper, we propose a data-based anomaly detection algorithm to detect FDIAs towards SCADA measurements by applying density-based analysis techniques after dimensionality reduction. The approach requires no prior knowledge on system parameters or topology. The proposed strategy also addresses the handling of critical measurements by implementing both dynamic and static forecasting algorithms.

References

[1] Y. Liu, P. Ning, and M. K. Reiter, "False data injection attacks against state estimation in electric power grids," *ACM Transactions on Information and System Security (TISSEC)*, vol. 14, no. 1, p. 13, 2011.

[2] G. Liang, S. R. Weller *et al.*, "The 2015 ukraine blackout: Implications for false data injection attacks," *IEEE Transactions on Power Systems*, vol. 32, no. 4, pp. 3317–3318, 2017.

[3] G. Liang, J. Zhao, F. Luo, S. R. Weller, and Z. Y. Dong, "A review of false data injection attacks against modern power systems," *IEEE Transactions on Smart Grid*, vol. 8, no. 4, pp. 1630–1638, 2017.

[4] C. Konstantinou, M. Sazos *et al.*, "GPS spoofing effect on phase angle monitoring and control in a real-time digital simulator-based hardware-in-the-loop environment," *IET Cyber-Physical Systems: Theory Applications*, vol. 2, no. 4, pp. 180–187, 2017.

[5] M. Hassanzadeh and C. Y. Evrenosoglu, "A regression analysis based state transition model for power system dynamic state estimation," in *North American Power Symposium (NAPS)*. IEEE, 2011, pp. 1–5.

[6] A. Ashok, M. Govindarasu, and V. Ajjarapu, "Online detection of stealthy false data injection attacks in power system state estimation," *IEEE Transactions on Smart Grid*, vol. PP, no. 99, 2017.

[7] L. Liu, M. Esmalifalak *et al.*, "Detecting false data injection attacks on power grid by sparse optimization," *IEEE Transactions on Smart Grid*, vol. 5, no. 2, pp. 612–621, 2014.

[8] M. Wu and L. Xie, "Online detection of low-quality synchrophasor measurements: A data-driven approach," *IEEE Transactions on Power Systems*, vol. 32, no. 4, pp. 2817–2827, 2017.

[9] A. Ghodsi, "Dimensionality reduction a short tutorial," *Department of Statistics and Actuarial Science, Univ. of Waterloo, Ontario, Canada*, vol. 37, p. 38, 2006.

[10] M. M. Breunig, H.-P. Kriegel *et al.*, "Lof: identifying density-based local outliers," in *ACM sigmod record*, vol. 29, no. 2. ACM, 2000, pp. 93–104.

[11] A. Lazarevic and V. Kumar, "Feature bagging for outlier detection," in *Proceedings of the eleventh ACM SIGKDD international conference on Knowledge discovery in data mining*. ACM, 2005, pp. 157–166.

[12] C. Konstantinou and M. Maniatakos, "A case study on implementing false data injection attacks against nonlinear state estimation," in *Proceedings of the 2nd ACM Workshop on Cyber-Physical Systems Security and Privacy*. ACM, 2016, pp. 81–92.

Charalambos Konstantinou (S'11-M'18) received his Ph.D. in Electrical Engineering from New York University (NYU), NY and the Dipl.-Ing. (M.Eng.) degree in Electrical and Computer Engineering from National Technical University of Athens (NTUA), Greece. He is currently an Assistant Professor of Electrical and Computer Engineering with Florida A&M University and Florida State University (FAMU-FSU) College of Engineering and an affiliated faculty with the Center for Advanced Power Systems (CAPS) at FSU. His research interests focus on cyber-physical, industrial control, and embedded systems security.

Michail Maniatakos (S'08-M'12-SM'17) is an Associate Professor of Electrical and Computer Engineering at New York University (NYU) Abu Dhabi, Abu Dhabi, U.A.E. He received his Ph.D. in Electrical Engineering from Yale University, New Haven, CT, USA, in 2012. He is the Director of the Modern Microprocessor Architectures Laboratory (nyuad.nyu.edu/momalab) and his research interests include privacy-preserving computation and industrial control systems security.