Exploring Physical-Based Constraints in Short-Term Load Forecasting: A Defense Mechanism Against Cyberattack

Mojtaba Dezvarei, Kevin Tomsovic, Jinyuan Stella Sun, Seddik M. Djouadi

2022 IEEE Power & Energy Society General Meeting (PESGM) | 978-1-6654-0823-3/22/$31.00 ©2022 IEEE | DOI: 10.1109/PESGM48719.2022.9917179

***Abstract*—Short term load forecasting is an essential task that supports utilities to schedule generating sufficient power for balancing supply and demand, and can become an attractive target for cyber attacks. It has been shown that the power system state estimation is vulnerable to false data injection attacks. Similarly, false data injection on input variables can result in large forecast errors. The load forecasting system should have a protective mechanism to mitigate such attacks. One approach is to model physical system constraints that would identify anomalies. This study investigates possible constraints associated with a load forecasting application. Looking at regional forecasted loads, we analyze the relation between each zone through similarity measures used in time series in order to identify constraints. Comprehensive results for historical ERCOT load data indicate variation in the measures recognizing the existence of malicious action. Still, these static measures can not be considered an efficient index across different scenarios.**

***Index Terms*—Short-term load forecasting, anomaly detection, physical constraints, time series, statistical indices, similarity measure.**

I. Introduction

Load forecasting is a critical to provide an efficient decision-making in power system operation and planning. Among the load forecasting types, short term load forecasting (STLF) plays a significant role in achieving economic, reliable, and secure operating operation. STLF supports a time horizon ranging from one hour to one week. The primary objective of the STLF function is to provide load predictions for the basic generation scheduling functions (unit commitment and hydroscheduling), for assessing the security of the power system at anytime point, and timely dispatcher information [1].

To predict the electric load, there are critical variables that accurately forecast the load. Many forecasting models with a wide range of methodologies for STLF are suggested in the literature. The models are mainly classified into two groups: statistical techniques, such as, multiple linear regression models (MLR), autoregressive and moving average (ARMA) models; and artificial intelligence (AI) techniques, such as, artificial neural network (ANN), fuzzy regression models, support vector machines (SVMs) [2]. All methods are constructed based on input data indicating the independent variables compatible with the forecasting horizon. In STLF, the primary inputs may include weather variables, calendar variables, and the load of the preceding hours [3], [4]. In [5], [6], the authors consider an additional variable to address locally increasing (or decreasing) load and the fact that electricity demand is affected by the temperatures of the preceding hours, respectively. Therefore, the quantity and quality of input data be important for an accurate estimation. Smart grid and renewable energy resources bring opportunities for improved forecasts by providing sensing and data acquisition capabilities with high resolution [7]. This work discusses big data applications and associated implementation issues in load forecasting.

With these new sources of data, the input data may face challenges, including measurement device error and cyberattacks (e.g., false data injection) that cause operators to make improper decisions. The vulnerability analysis for input data in load forecasting is analyzed in [8]. It is a vital task to check the quality of input data to guarantee accuracy. Statistical analysis and information theory [9], machine learning (ML)-based method [10], [11] are some approaches for anomaly detection schemes.

As the attackers also benefit from AI and ML methods, attackers may deploy sophisticated techniques to bypass the anomaly detection, such as, targeting state estimation and electricity market application [4], [12], [13]. Although detection scheme imposes difficulty for the attacker, this may be overcome by, for example, reinforcing the threat model [14]. To address this issue, physical-based constraints can provide a stronger obstacle for attacker. Li et al. propose adding inherent constraints in power system state estimation (SE) [15]. This idea makes a further constrained problem that needs to be solved by attackers while satisfying the physical systems' inherent constraints. In this paper, we seek to deploy a similar concept that supports building a defense mechanism for STLF against adversarial attack.

Load forecasts can be represented as a time series, which offers various tools to identify characteristics. In this regards, similarity measures can facilitate us to identify physical-based constraints. The paper's main contribution is that this is a first attempt to define a protective mechanism in STLF associated with these inherent constraints.

The paper is organized as follows: The STLF formulation

This work was partially supported by NSF grant CNS-2038922. Mojtaba Dezvarei, Kevin Tomsovic, Jinyuan Stella Sun, and Seddik M. Djouadi are with the Department of Electrical Engineering and Computer Science, The University of Tennessee, Knoxville, TN 37996 USA. (e-mail: mdezvare@utk.edu; tomsovic@utk.edu; jysun@utk.edu; mdjouadi@utk.edu)

978-1-6654-0823-3/22/$31.00 ©2022 IEEE

and modeling are presented in Section II. Section III presents the problem statement in STLF capturing the constraints. Exploring constraints using similarity measures is explained in Section III. Then, simulation results and observations are discussed in Section V. Section VI provides some concluding remarks.

II. Short term load forecasting modeling

In STLF, the objective is to estimate the load for a short time horizon, typically, ranging from the next hour to one week. The main intention of STLF is an hourly prediction of aggregated load. In addition, the STLF may focus on forecasting the daily peak system load, the value of system load at a specific time of day, hourly or half-hourly of system energy, and the daily and weekly system energy [1]. In this regard, the STLF is concerned with scheduling purposes for the most economic commitment of generation sources consistent with reliability requirements, operational constraints. Furthermore, the prediction of load may be used for the assessment power system security for the off-line network analysis under which the part of systems may face stress.

The system load is the sum of all the individual demands in all nodes of power system. Many factors impact the STLF model of aggregate load. The factors are mainly categorized into weather, calendar day, economic, social events and other random factors. So, the dependent variable that needs to be estimated is aggregate load, and other variables relating to effective factor categories are considered as independent variables (explanatory). A practical report indicating independent variables for different load-serving entities (LSEs) is found in [16].

The general formulation of STLF is to define a forecasting function f as $y_t = f(\bar{X}_t)$ where y_t, and $\bar{X}_t$ are the load and the independent variables at time t. There are many ways to find f, which are commonly divided into statistical and AI methods. The former is statistical analysis such as MLR, ARMA, and exponential smoothing models. An AI approach may use an ensemble of machine learning algorithms to forecasts, such as, ANN, fuzzy regression models, SVM and so on. A high-level comparison of load forecasting techniques is presented in [2]. In this paper, the MLR is chosen to apply the proposed methods, although it should be generally applicable to other techniques. MLR is a broader class of regressions that encompasses linear and nonlinear regressions with multiple independent variables $\vec{X}$ and one dependent variable Y. The general linear regression model can be defined as

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \cdots + \beta_n X_n + \epsilon \quad (1)$$

where $\beta_0, \beta_1, \ldots$ are the parameters and ϵ is normal random error. Note that second or higher-ordered terms of a independent variables is equivalent to the first ordered one and it is a special case of Eq. 1 [5].

A. Vulnerability of independent variables

Some independent variables cannot be easily corrupted as their values are widely known. This includes calendar variables (month of the year, day of the week, hour of the day, etc.). Economic factors are not explicitly expressed in STLF models due to the longer time scale. Among other variables, weather factors significantly in the load pattern because many loads are sensitive to temperatures, such as, heating and cooling.

The temperature data is the most used features in many STLF models to capture the load behavior. In this case, the quality of temperature is crucial to STLF accuracy. The temperature used in the STLF model is an estimated value where is obtained from commercial weather services or external services/APIs. This also provides a vulnerable area for data disruptions and false attack injections. Sobhani et al. suggest temperature anomaly detection with the help of local load information collected by power companies [17]. Although the detection scheme can be helpful to find the anomaly in temperature, an attacker who also benefits from AI methods would be more challenging. The adversarial attack on temperature data exploited in [8] indicates that STLF models are quite vulnerable to malicious data in temperature from online weather services. In addition, an attacker could manipulate load forecasts in arbitrary directions and cause significant and targeted damages to system operations.

III. Problem statement

Considering the temperature as a targeted variable for the attacker, we investigate constraints to act as a defense mechanism such that the attacker must satisfy these in order to have a feasible result. Being undetected from system operators is the primary task for the attacker to generate an adversarial action. Generally, the attacker's capability will be limited for a successful malicious action. That means the difference between corrupted temperature and actual temperature will be bounded [8]. Considering H is attacker simulated forecasting model parameterized by θ. The only constraint here is about bypassing the detection scheme. Adding the psychical-based constraints may cause difficulty for the attacker to find a feasible attack vector. Hence, the problem that the attacker should solve can be defined as the following optimization problem

$$\min_{\tilde{X}} \quad H_\theta(\tilde{X}) \quad (2a)$$

$$s.t. \quad \left\| X - \tilde{X} \right\|_p \leq \epsilon \quad (2b)$$

$$g(x) \leq 0 \quad (2c)$$

where $\tilde{X}, X$ are the injected and actual temperature data, ϵ is a threshold value, and p shows different norm according to detection algorithms. Here we are adding (2c) to (2a) to represent the psychical-based constraints. The minimization problem here points out attacker target as decreasing the load forecast values. The constraint $g(x)$ can be imposed either as an equality or inequality related to the physical system and topology. The right side of (2c) may be defined by an upper bound to deal with various scenarios. The way to find $g(x)$ is discussed in the next section.

978-1-6654-0823-3/22/$31.00 ©2022 IEEE

IV. Exploring Constraints for STLF

Power systems follow some physical and topological constraints. For instance, output of power metering devices should ideally capture Kirchhoff's laws. This action may present a built-in defense mechanism. Li et al. study potential vulnerabilities of ML applied in power systems by proposing constraints that adversarial examples must satisfy the intrinsic constraints of physical systems [15]. In power system SE the physical constraints have already been encoded in the mathematical model; however Liu et al. propose an approach that guarantees attacker to pass those constraints in linear case [12].

The indication of such constraints related to load forecasting applications has not been investigated in the literature. Since STLF mainly applies to aggregated load models, the dynamic of loads that follow the physics law is not easily observed. To build a set of psychical-based constraints, the spatial distribution of zones is considered. For example, ERCOT load map can be separated based on eight different weather zones. Each zone deploys its weather station to STFL, which may also represent an aggregate.

Generally speaking, we investigate the relation between the different zones to derive constraints. In this framework, we are looking for an index representing forecasted variations between zones. When an attacker intents to inject false data into some weather stations (but not all), the malicious action may be identified by these indices. Here, similarity metrics applied in clustering time series are inspected to find a relation between zones. In fact, the goal is to explore the physical-based constraints between zones $g(x)$ that may be obtained by these similarity measures.

Definition 1: The *similarity measure* $D(X, Y)$ between time series $X = \{x_0, x_1, ..., x_{N\text{-}1}\}$ and $Y = \{y_0, y_1, ..., y_{N\text{-}1}\}$ is a function taking two time series as inputs and returning the *distance* d between these series [18].

Two common approaches for similarity measure are considered to deploy as shape-based and feature-based methods.

A. Shape-based approach

Here, distances are based on directly comparing the raw values and the shape of the series in different manners. That is, distances compare the overall shape of the series. Due to the simplicity, the Euclidean distance d_{EUC} and other L_p norms are the most widely used distance measures for time series clustering. Euclidean distance is invariant when dealing with changes in the order that time features are presented. The Euclidean distance is defined as

$$d_{EUC} = \sqrt{\sum_{i=0}^{N-1} (x_i - y_i)^2} \tag{3}$$

B. Feature-based approach

The feature-Based method focuses on extracting a set of features from the data presented in time series and calculating the similarity between the features instead of using the raw values of the series.

1) Correlation-based distance: One index is to use the Pearson's correlation factor for two time series. The time series are similar if they are highly correlated, even though the observed values may be far apart in terms of Euclidean distance. Based on Pearson's correlation, Golay et al. suggest two correlation-based distances for time series X and Y as follows [19]

$$d_{COR\cdot 1}(X, Y) = \sqrt{2\left(1 - COR(X, Y)\right)} \tag{4}$$

$$d_{COR.2}(X, Y) = \sqrt{\left(\frac{1 - \mathrm{COR}(X, Y)}{1 + \mathrm{COR}(X, Y)}\right)^{\beta}} \tag{5}$$

where $\beta > 0$ is a distance decreasing parameter, and COR is the correlation between X and Y. It can be observed that a strong positive correlation results in small-scale correlation-based distances, which implies high similarity.

2) Autocorrelation-based distance: The other feature-based method is based on estimated autocorrelation functions. The autocorrelation function is the normalized autocovariance function indicating the estimated auto correlation vectors of two times series as $\hat{\rho}_{X_T}$ and $\hat{\rho}_{Y_T}$, the autocorrelation-based distances between X and Y is defined as follows [20]

$$d_{ACF}(X, Y) = \sqrt{\left(\hat{\rho}_{X_T} - \hat{\rho}_{Y_T}\right)^{\top} \Omega \left(\hat{\rho}_{X_T} - \hat{\rho}_{Y_T}\right)} \tag{6}$$

where Ω represents weight matrix, and T is the sample size.

3) Periodogram-based distance approach: This method computes the distance between periodograms of two time series in the frequency domain. A periodogram is an estimate of the spectral density of a signal. For a stationary time series process with finite length, the periodogram at the Fourier frequency $\omega_j = (2\pi j)/N$ for $j = 1, 2, \ldots, [N/2]$ (where $[N/2]$ is the largest integer less or equal to $n/2$) is defined as

$$\boldsymbol{P}_x(\omega) = \frac{1}{N} \left| \sum_{t=1}^{N} x_t e^{-it\omega_j} \right|^2 \tag{7}$$

The scaled periodogram distance can be defined as

$$d_{\mathrm{P}}(X, Y) = \sqrt{\sum_{j=1}^{[N/2]} \left[\boldsymbol{P}_x(w_j) - \boldsymbol{P}_y(w_j)\right]^2} \tag{8}$$

The normalized version gives more weight to the shape of the curve, while the non-normalized contributes to the scale. Measures based on the autocorrelations and measures based on the periodogram are related. However, this application could provide different useful estimation results.

4) Symbolic representation: Transformation of time series into sequences of discretized symbols can be efficiently processed to extract information about the underlying processes. The Symbolic Aggregate approXimation (SAX) method proposed in [21] allows to investigate the dissimilarity based on the symbolic representation of time series to find the true dissimilarity between the original time series. More details about the procedure are addressed in [21].

978-1-6654-0823-3/22/$31.00 ©2022 IEEE

V. Illustration and Simulation

A. Dataset

The ERCOT area can be categorized based on eight weather stations, where two West and Far West areas are selected to apply the proposed methods. The load data is for summer 2020, and it is accessible from the ERCOT website. This data set includes the overall and individual area load. As the temperature is an essential variable in STLF modeling, the temperature data was collected from Automated Surface Observing Systems (ASOSs) corresponding to the West and Far West weather stations. After data cleaning steps such as handling missing values, removing unwanted observations, particularly for the temperature data set, the recorded load verse temperature is shown in Fig. 1. As it is clear, high temperature causes increasing use of demand-side devices such as air conditioners.

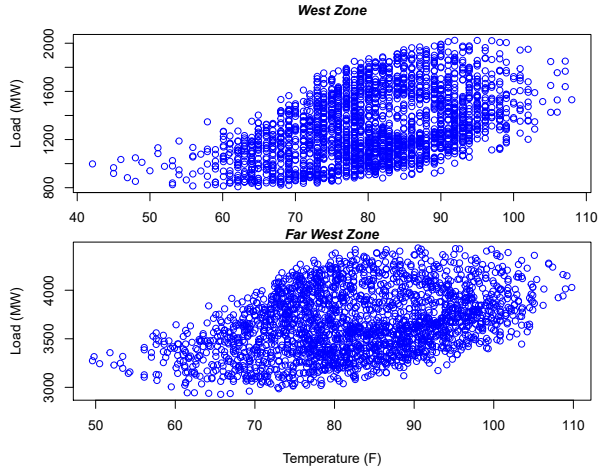

Fig. 1. Load vs. temperature

B. Correlation Analysis

To explore possible relations between loads in the west and far west in ERCOT, first, the correlation analysis between load data is calculated, and results are shown in Fig. 2. Although the two zones are separated geographically, there is a strong correlation between these loads. The reason could be the similarity of weather patterns in each zone that demonstrates hidden physical associations.

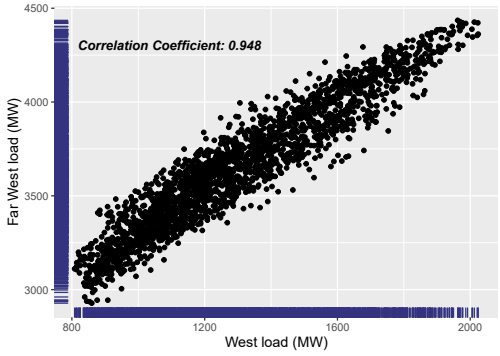

Fig. 2. West and Far West zone load correlation

C. STLF modeling and results

Two MLR models are applied for STLF modeling since the impact of temperature data as an independent variable may be seen explicitly in the formulation. Note that other techniques, either statistical or AI methods, are applicable in this case. Two models including interaction effect of independent variables based on [5] are assumed as

$$f_1 = \beta_0 + \beta_1 \times T + \beta_1 \times H + \beta_2 \times D + \beta_3 \times LL_{1w} + \beta_4 \times LL_{2w} \quad (9)$$

$$f_2 = \beta_0 + \beta_1 \times D \times H + \beta_2 \times M \times T + \beta_3 \times M \times T^2 + \beta_3 \times M \times T^3 + \beta_4 \times H \times T + \beta_5 \times H \times T^2 + \beta_6 \times H \times T^3 \quad (10)$$

where f_1, and f_2 are the foretasted load models. The model parameters $\vec{\beta}$ are obtained based on the least-squares (LS) method. The variables D, M, and H denote the day of the week, excluding the weekend, month of the year, and hour of the day, respectively. T is the temperature value, and LL_{1w} and LL_{2w} express one and two-week lagged load data, respectively.

Both models consist of continuous variables and categorical variables. The categorical variables need to be coded into a series of variables in order to be used in regression modeling. So, variables D, M, and H are encoded as a constant factor in the R programming language. The data set is randomly divided into 70% and 30% for train and test set, respectively. The models result are shown in Fig. 3. The outcomes indicate that either f_1 or f_2 model could efficiently represent the load for both zones. A Gaussian noise with zero mean and unit variance is injected into West zone temperature data to mimic malicious actions. Then, similarity measure methods for the forecasted load are applied.

Table I indicates the variation of similarity measure indices in the presence of the attack. Note that comparing the results between indices is misleading since each measure carries a different scale. The similarity measures change when there is false injection data. Thus, this may be a suitable candidate for constraints in STLF. Among the methods, the variant in d_{SAX} is the largest. The reason would be distance measure defined on the symbolic approach creates lower bound corresponding distance measures defined on the original time series. The variation in measures is not a constant value and will vary under different scenarios, such as different Gaussian signals. That is finding an explicit formula or the upper bound capturing different scenarios is an essential. In this case, to inject data with a destructive impact on the STLF, attackers must build an attack vector to bypass the measures satisfying the STLF similarity measures .

VI. Conclusion

STLF plays an important role in power system operation. Collecting temperature, as a critical input from online services, increases the possibility of a malicious action. This paper examines some physical constraints to address deficiency of

978-1-6654-0823-3/22/$31.00 ©2022 IEEE

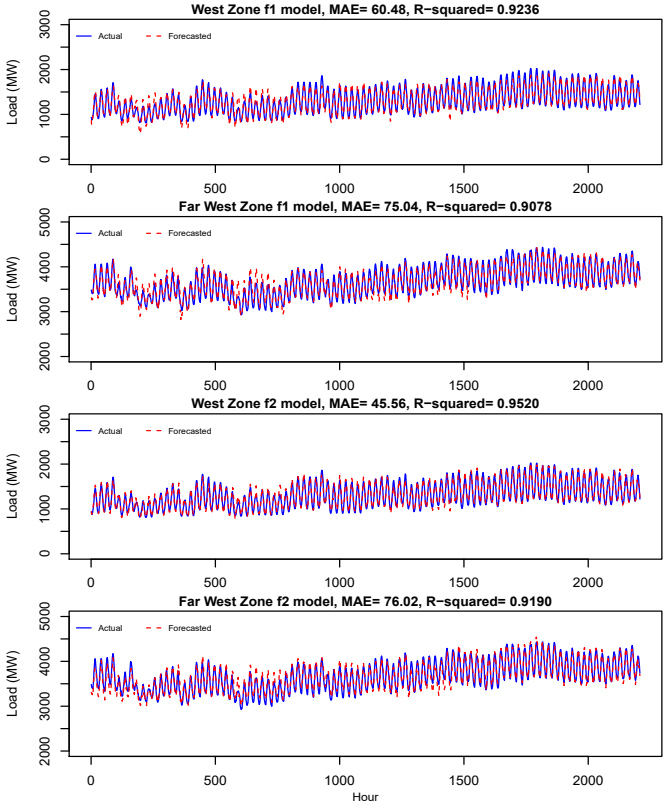

Fig. 3. Models performance: Mean absolute error (MAE) and adjusted R-squared values

TABLE I
SIMILARITY MEASURES RESULTS

Method	No False Injection			False Injection	
	No Model	f_1	f_2	f_1	f_2
d_{ECU}	110644.5	110587.1	110675.4	110652.9	110602.3
d_{COR}	0.3204611	0.3183464	0.2687631	0.326226	0.2826415
d_{ACF}	1.247954	1.146977	1.015968	1.149789	1.040784
d_p	0.1336534	0.1309215	0.1109201	0.1273667	0.1091463
d_{SAX}	2.004495	1.735943	1.417392	2.454994	1.002247

detection schemes. A new formulation of STLF based on different spatial coordination is proposed for imposing these constraints. Similarity measure techniques are applied as constraints in the proposed approach. Simulation results indicate that the applied measures can be applicable for encoding constraints when temperature data is corrupted. Distance defined in symbolic space identifies variation that can represent a strict constraint following the proposed STLF topology. Since the exploited measures vary based on scenarios, an attempt to define the explicit equation representing the constrains or to realize the upper bound will be investigated for the feature work. As part of our ongoing work, we are investigating incorporating such constraints directly into the learning algorithms for load forecasting to mitigate cyber attacks.

REFERENCES

[1] G. Gross and F. D. Galiana, "Short-term load forecasting," *Proceedings of the IEEE*, vol. 75, no. 12, pp. 1558–1573, 1987.

[2] T. Hong and S. Fan, "Probabilistic electric load forecasting: A tutorial review," *International Journal of Forecasting*, vol. 32, no. 3, pp. 914–938, 2016.

[3] I. Drezga and S. Rahman, "Input variable selection for ann-based short-term load forecasting," *IEEE Transactions on Power Systems*, vol. 13, no. 4, pp. 1238–1244, 1998.

[4] J. Xie and T. Hong, "Variable selection methods for probabilistic load forecasting: Empirical evidence from seven states of the united states," *IEEE Transactions on Smart Grid*, vol. 9, no. 6, pp. 6039–6046, 2017.

[5] T. Hong, P. Wang, and H. L. Willis, "A naïve multiple linear regression benchmark for short term load forecasting," in *2011 IEEE Power and Energy Society General Meeting*. IEEE, 2011, pp. 1–6.

[6] P. Wang, B. Liu, and T. Hong, "Electric load forecasting with recency effect: A big data approach," *International Journal of Forecasting*, vol. 32, no. 3, pp. 585–597, 2016.

[7] M. Kezunovic, P. Pinson, Z. Obradovic, S. Grijalva, T. Hong, and R. Bessa, "Big data analytics for future electricity grids," *Electric Power Systems Research*, vol. 189, p. 106788, 2020.

[8] Y. Chen, Y. Tan, and B. Zhang, "Exploiting vulnerabilities of load forecasting through adversarial attacks," in *Proceedings of the Tenth ACM International Conference on Future Energy Systems*, 2019, pp. 1–11.

[9] V. Chandola, A. Banerjee, and V. Kumar, "Anomaly detection: A survey," *ACM computing surveys (CSUR)*, vol. 41, no. 3, pp. 1–58, 2009.

[10] M. Cui, J. Wang, and M. Yue, "Machine learning-based anomaly detection for load forecasting under cyberattacks," *IEEE Transactions on Smart Grid*, vol. 10, no. 5, pp. 5724–5734, 2019.

[11] M. Esmalifalak, L. Liu, N. Nguyen, R. Zheng, and Z. Han, "Detecting stealthy false data injection using machine learning in smart grid," *IEEE Systems Journal*, vol. 11, no. 3, pp. 1644–1652, 2014.

[12] 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, pp. 1–33, 2011.

[13] Q. Zhang and F. Li, "Cyber-vulnerability analysis for real-time power market operation," *IEEE Transactions on Smart Grid*, 2021.

[14] A. Chakraborty, M. Alam, V. Dey, A. Chattopadhyay, and D. Mukhopadhyay, "Adversarial attacks and defences: A survey," *arXiv preprint arXiv:1810.00069*, 2018.

[15] J. Li, Y. Yang, J. S. Sun, K. Tomsovic, and H. Qi, "Conaml: Constrained adversarial machine learning for cyber-physical systems," in *Proceedings of the 2021 ACM Asia Conference on Computer and Communications Security*, 2021, pp. 52–66.

[16] J. P. Carvallo, P. H. Larsen, A. H. Sanstad, and C. A. Goldman, "Load forecasting in electric utility integrated resource planning," 2016.

[17] M. Sobhani, T. Hong, and C. Martin, "Temperature anomaly detection for electric load forecasting," *International Journal of Forecasting*, vol. 36, no. 2, pp. 324–333, 2020.

[18] P. Esling and C. Agon, "Time-series data mining," *ACM Computing Surveys (CSUR)*, vol. 45, no. 1, pp. 1–34, 2012.

[19] X. Golay, S. Kollias, G. Stoll, D. Meier, A. Valavanis, and P. Boesiger, "A new correlation-based fuzzy logic clustering algorithm for fmri," *Magnetic resonance in medicine*, vol. 40, no. 2, pp. 249–260, 1998.

[20] P. Galeano and D. Peña, "Multivariate analysis in vector time series," 2001.

[21] J. Lin, E. Keogh, S. Lonardi, and B. Chiu, "A symbolic representation of time series, with implications for streaming algorithms," in *Proceedings of the 8th ACM SIGMOD workshop on Research issues in data mining and knowledge discovery*, 2003, pp. 2–11.

978-1-6654-0823-3/22/$31.00 ©2022 IEEE