A UNSUPERVISED MACHINE LEARNING PROTECTION SYSTEM FOR LARGE SCALE IOT GRIDS

AUTHOR'S NAME:

K.P.PORKODI APS/CSE
C.VASANTHAKUMAR AP/CSE
S.NARMATHA ME/CSE

ABSTRACT

The electrical grids are more dependable, secure, and efficient thanks to smart grid technology.For effective and dependable power distribution, new vulnerabilities are raised by its strong reliance on digital communication technology. The unsupervised anomaly in this paper.The idea of measurement correlation-based detection has been put forth. The objective is to create a scalable a large-scale anomaly detection engine for smart grids that can distinguish between an actual fault and a commotion and a clever cyber-attack. The suggested technique utilises feature extraction by Using symbolic dynamic filtering (SDF) to lighten the computational load and find causal relationships among the subsystems. The outcomes of the simulations on the bus systems support the performance of the suggested method under various operating circumstances. The outcomes demonstrate accuracy of 99%, a 98% true positive rate, and a less than 2% false positive rate.

INTRODUCTION

In today's power systems, distributed energy resources (DERs) and advanced metering infrastructure is organised to provide reliable energy generation and a network of sensors and generators that enable two-way communication inside the system's infrastructure (AMI). Even though energy efficiency, dependability, and manageability of this advanced communication system are all improved, the system's sensitivity to cyber-attacks is raised by the large number of devices and access points that operate outside the normal administrative domain.It is critical to look at how cyber attacks affect power systems since malfunctions in the power grid could result in catastrophic events.According to [1], the primary cause of the blackouts in North America is a lack of system awareness**.**which emphasise the significance of cyber-attack analysis in order to keep the power supply running steady and reliably. Cyberattacks have the potential to destroy equipment through overload or generate a lot of energy through erroneous demand requests [2][4]. Additionally, a malicious attack could result in fake overload conditions, or false negatives, in a power system.It's also potential for the infrastructure supporting electric vehicles to experience more difficulties. As

demonstrated in [5], [6], malicious attacks that prevent communications with a device can halt substation computers' services.

The reliable operation of essential infrastructure, including smart grids, depends on real-time cyberattack detection. It is necessary to monitor systems continuously and online in order to spot targeted cyberattacks.In a large-scale network, individual sensors are typically the focus of security breaches. Information stored in a corrupted system can simply be accessed by a compromised insider node. Applying an authentication method to sensor networks theoretically enables key revocation of any compromised node. However, due to the system's computational and storage limitations, authentication methods based on cryptography or security gateway architecture, such those described in [8], [9], are impractical. The existing research in the context of the smart power grid primarily focuses on sophisticated anomaly detection techniques [13], [14], and secure control theories based on various state estimate approaches [15], as well as the networking security of the cyber elements [10], [12], and [13], [14].

The presence of a cyberattack on a power system is thoroughly analysed and described. Despite the fact that the aforementioned technologies can protect power systems, the majority of them are too expensive mathematically, physically impracticable, and not scaleable for large-scale complex networks. Today, massive amounts of data are produced across the grids, making it easier to monitor systems in real time. The performance monitoring, diagnosis, and prognosis of anomaly in complex systems are all considerably improved by exploring these data.Anomalies and potential assaults can be found using historical data describing the system's operation. Due to the enormous amount of data created by the smart grid, standard Bad Data Detection (BDD) algorithms are not equipped to handle real-time computing and storage difficulties. Due to these difficulties, it is now possible to use data analysis methods like machine learning (ML) to handle data sets with complicated structures. AMI, states, and control actions can be used to examine different combinations of measures using ML algorithms.Through becoming familiar with their patterns [17, 18]. It has a false data detector.Attack using the non-linear, complex injection (FDI) a connection between measures. This is achievable in a similar manner to how effective strategies are used in other issues with the power system as reported in the research literature [19]. There are only a few research on the use of ML on smart grids' cybersecurity. The following ML methods are evaluated and compared in [20] for FDI attack detection.General conclusions about the effectiveness of machines were drawn.learning how to categorise FDI attacks. [21] suggested an amalgam.Using common path mining to discover intrusions a technique to identify unusual power system incidentsPMU's ML algorithm logs from the energy management system (EMS), relays, and data. A cyber-attack detection method based on the Pearson correlation coefficient between two PMU characteristics was utilised in [22]. Using the Pearson correlation coefficient, this method examined how the correlation between two PMU parameters changed over time. The attack approach for anomaly detection was modelled by the authors of [20] using the Gaussian process in conjunction with ML. A supervised ML-based method is put out in [23] to identify a cyber-deception attack during the state estimate

phase. A deep learning approach that can instantly identify key FDI attack characteristics is also put out in [24].

Probabilistic Graphical Models (PGM) can be used to simulate complicated system behaviour and enhance the performance of the current data-driven attack detection systems. PGMs include Dynamic Bayesian Networks (DBN), which are practical representations of large systems that change over time based on the causal connections between system parts [25]. To preserve the robustness, scalability, and accuracy of the attack detection algorithms, new techniques for handling complex and high-dimensional data should be created. Feature extraction can be used to change the original features into a more meaningful representation by recreating its inputs and it entails minimising the number of resources required, which eases the computational strain in huge data sets [26], [27]. As there are abnormalities that cannot be quantified or reproduced, detection methods that do not rely on pre-classified training data are crucial. In this article, we suggest a method for extracting the patterns of changes in FDI assaults using a smart grid anomaly detection system. The attacks are immediately detected using the revealed attributes.In order to construct a computationally efficient feature extraction strategy to find causal relationships between the smart grids sub-systems through DBN, symbolic dynamic filtering (SDF) is used. Free energy is utilised as the anomaly index to detect unobservable cyber-attacks using Mutual Information (MI), DBN, and learning algorithms. By assigning a scalar energy to each variable, which acts as a gauge of compatibility, we want to capture the relationships between variables. The suggested technique's scalability is tested on a variety of IEEE test systems that were based on the PSS/E modelling programme. Under various operating situations, the results demonstrate good accuracy and little false alarm. It should be noted that the suggested method leverages the idea of free energy to distinguish between the energy level in the attacked and normal data sets in addition to relying on patterns in training data sets. As a result, even fresh and unexpected attacks can be found.These are the primary contributions of this work: a method for detecting anomalies in smart grids without labelling data sets has been developed.putting forth a strategy that is scalable by easing the strain of computation through SDF data reduction building a solid DBN-based learning model. putting forth a model-free strategy that can be used in hierarchical and topological networks for various attack scenarios. The remainder of the essay is structured as follows. Section II describes mathematical formulations. potential cyberattack Section III presents a detecting technique.

Mathematical modelling:

A. MODEL OF THE GENERATOR

This work models the smart grid as a multi-agent, cyberphysical system, where each agent has a generator, a measurement device, a distributed control agent, and an energy storage system that may add or take away actual power from the system [28]. The system's dynamic and static states are discussed. where x is the system state, which includes the static state of the network and the dynamic state of the generator (such as rotor speed and rotor angle) (voltage magnitude and phase angle). The measurements' non-linear function, h (.), and the

generators' non-linear behaviour, f (.), are both non-linear. The output and measurements vectors are denoted by u and z, respectively.It is possible to describe the generator i's 4-th order (two-axis) model. where the time derivative is denoted by (P). We discuss the generator's parameters. For synchronous generator I, the excitation system regulates the eld voltage, the related speed governor regulates the mechanical torque, and the electrical output can be derived where The conductance and susceptance between generators I and k are referred to as Gik D Gki and Bik D Bki, respectively.In order to detect anomalies and cyberattacks, this effort aims to learn and anticipate the dynamic behaviour of the smart power grid (where generators are modelled as mentioned in this section). A computationally efficient tool for identifying the interconnections between the subsystems is created using SDF, DBN, and RBM.

B. REPRESENTATION OF ATTACKS

By estimating the L-norm of the measurement residual, the BDD approach has historically been used to verify the accuracy of the state estimation process [31]. Where z 2 RN is the measurement vector, Ox 2 RD is the predicted state vector, and H 2 RND is the Jacobian matrix, the presence of incorrect data is identified.To keep the state estimation's precision, a threshold Tr is pre-dened. In addition to being circumvented by cyberattacks, the measurement redundancy needed for BDD methodologies renders them useless for smart grid technology. The objective of the adversary in clever cyber-attacks, particularly FDI attacks, is to control a portion of the measurements and arbitrarily change the state variables. You can accomplish it by inserting fake data. vector za 2 RN, which omits conventional BDD methods. Let's say the malicious assault purposefully alters the metre values provided by za. In light of this, the attack caused a measurement shift. where the measurement noise is, and Oxa the incorrectly calculated state.

The injected fake data (za) can be divided into two pieces, a D Hca and qa, where qa is the only observable component that is located in the complementary space where H HTH 1 HT qa D 0. Ca 2 RD is an injected vector of data that bypasses BDD tests since it is located in the column space of H.To put it another way, the stealth attack vectors (za) always exist even if the adversary only has access to a portion of the network topology and line parameters in order to create malicious attacks that entirely lay in (H), i.e., qa D 0, and therefore avoid the BDD techniques currently in use. [32]

the following assumptions are considered in the model of the attack:

The attacker in this study is thought to have few resources and be able to modify a small number of measurement readings. For a time period Ta T, this could either be power injection or power ow data.This is a reasonable assumption because it is not practical to suppose that all sensors will simultaneously report inaccurate values in the context of power networks.Furthermore, it takes a lot of time and effort for attackers to compromise all measures in practise.

It is absolutely impossible for an outsider to have a complete understanding of the system. As a result, the attacker is only partially aware of the system's topology and security features. Such information can be gathered by physically recording the security information encoded in a node or by statistically analysing the data transferred from the remote terminal units (RTUs) to the control centre.This study takes into account the least absolute shrinkage and selection operator (LASSO) with a strategic sparse FDI attack.

A row-wise method is used to decompose the Jacobian matrix (H). To represent the secure measurements, a submatrix of H called HS D is formed, where Hji;V is the ji-th row of H and HSca D 0: Submatrix HA is also built for measurements that have been attacked. The attacker's plan is then developed to find an optimal answer. where 0 is a predetermined constant. The LASSO and Regressor Selection methods are used to tackle the optimization challenges.The attack's construction is covered in more detail in [33].By breaking into the communication network, the attacker hopes to change the rotor's speed and angle via FDI attack. Since I is a constant coefficient and Ci is a constant bias in the attacked states, 8t 2 Ta indicates the impact of FDI attacks on the system state for generator i. In other words, the attacker wants to use I and Ci to change the system state. In light of this, the attacker will create za in such a manner that the attack vector is concealed from the operator and conventional BDD techniques. For the experimentation, we suppose measures, chosen at random to create a sparse attack vector, are available to the attacker.

III. PROPOSED ENERGY-BASED CYBER-ATTACK DETECTION

This section presents a methodology for cyber-attack detection that makes use of DBN modelling, feature extraction through MI, and RBM for data training. RBM is used to capture the patterns in system behaviour that are extracted by the unsupervised DBN model as it is applied to smart grid test systems with large measurements (data are not labeled).Suggested data-driven framework for anomaly identification. The system is initially divided into a number of sub-systems. Then, using SDF, causal dependence between nominal subsystem features is learned. The suggested approach is a computationally efficient tool that reduces the computing load by 1) choosing a subset of measurements by feature selection and SDF, and 2) through domain decomposition and data processing on Instead of addressing the entire system at once, several subsystems are addressed concurrently.

A. SYMBOLIC DYNAMIC FILTERING

The time series data are first converted into symbol sequences in the suggested feature extraction method based on SDF and then these sequences are dened from DBN to compress transforming the data into simple statistical patterns. The system's phase space in equation (1) is segmented into a nite amount of cells. Introducing a compact region identifies a partition B fB0;::: ; B 1g with the following elements:(Bj) mutually exclusive T 8j 6D k) and exhaustive (Bk D(S □1 cells of type jD0 Bj D). Described by the dynamic system the time-series data as O f 0;::: ; I 2,which passes through the partition B's cells [34, 35].to comprehend the ideas behind mapping and partitioning take into account the symbol alphabet Think about the

system in Think of the cell a trajectory passes through as a random variable.S has the symbol value of 2A. The symbol alphabet a collection A of various symbols designating the components in the division. Each initial state (0, 2) results in a series of symbols that can be mapped from the phase to dene the symbol with some room It is known as symbolic dynamics. Through the symbolization process, three-dimensional space is reduced to a symbol sequence, followed by a DBN.

B. DYNAMIC BAYESIAN NETWORKS

DBNs are probabilistic graphical models that show the state of a system as a collection of variables and model the probabilistic dependencies of the variables over time. This study takes into account a high order DBN on variables xt D fx1;t ;::: ; x;t g at various time points t D 1;::: ; T. The expression of the state I at time t is represented by each xi;t. The variables that SDF sets are used to extract the symbol sequence. In order to determine the likelihood that a new symbol will appear, we assume that the DBN satisfies the L-th order Markov property. As a result, using the training data, a state transition matrix 5 that describes the L-th order Markov chain can be generated.Trial and error is used to determine the model's order. Let's use qk to represent the state at time instant k. One can define the ij-th element of five. In this work, we use a modified version of Markov chain (xL-th order Markov chain) [36] to predict the occurrence probability for a new symbol in a series A using the last L symbol for another series B because we are dealing with multiple time series. For L-th order Markov representing sub-systems A and B, 5A and 5B are denoted. In the same way, cross state transition matrices 5AB and 5BA, respectively, can be used to represent the causal dependencies of A on B and B on A.Atomic patterns (APs) are characteristics from L-th order Markov chains, and relational patterns are characteristics from xL-th order Markov chains (RPs).

One can describe the state-transition matrices 5AB and 5BA. where the state vectors for sequences A and B, respectively, are denoted by j; kQA and I lQB. A multivariate time series is given, and partitioning is used to create the symbol sequences S. The next step is to determine the subsequent states and transition probabilities between the vertices using a high order DBN. We extract significant features from an AP or RP using MI criteria. A generalised linear correlation coefficient is created by MI to calculate the correlation between two random variables. When MI has a non-zero value, the two variables are said to be independent of one another. MI can be expressed as Importance metric IAB between state sequences qA and qB. The RBM learns patterns of system behaviour once the models are prepared. Test results are used to determine the probability of the learned features. Restricted Boltzmann Machine (RBM) was utilised in this work to achieve this.

C. RESTRICTED BOLTZMANN MACHINE

A generative technique to model the unknowable distribution of data is the Boltzmann machine. Boltzmann Machine can create new data with a given joined distribution and complete patterns in the case of missing inputs, in contrast to most Machin Learning techniques that only discriminate some data vectors in favour of others. It is also thought to

be more flexible and feature-rich. The Energy-based Models (EM) subclass of stochastic models includes RBM [38]. Each system state in electromagnetics (EM) is connected to a particular energy level. A network of stochastic binary neurons (a set of visible variables v D fv1;::: ;vN g) connected to a set of hidden variables h D fh1;::: ;hK g can be used to model such a system. Based on joint configurations of the visible and hidden variables, the state of the system can be described. It has been demonstrated that estimating models in RBMasses maximises the likelihood of training data with low energy state. As a result, an anomaly will manifest as a high-energy or low-probability conjunction [39]. Given binary variables v and hidden variables h, a Boltzmann distribution function can be used to describe the joint probability of a state (Pr.v; h/) based on the energy of that state (En.v; h/). As a result, the anomaly index that ranks data instances in linear time can be free energy. In order to identify cyberattacks based on the likelihood and intensity of the event, trained RBM is used. An event with high energy or low probability serves as the analogy for an anomaly. It is presumable that cyberattacks alter how the subsystems interact, changing how DBN patterns look. IAB can be normalised into binary states for APs and RPs (0 and 1 for low and high values, respectively) to make training easier.

Finally, modifications to parameters associated with recognised patterns are used to spot cyberattacks. Based on a distance metric, a distribution of free energy is used to identify low probability events or cyberattacks. To maintain normal operations condition, the distribution of free energy will resemble that of the training data. The training data are primarily believed to have been gathered under standard operating conditions. As a result, the learned RBM can accurately capture the system's normal operation. Relative Entropy (RE) metric is used to express how different the energy distributions in training and test data are.

A measurement of the separation between two probability distributions is the relative entropy between them. RE can be defined as [35], [36], and [37] for two probability distributions P and Q on a finite set X. serves as the starting point. The supposition is that the data are within two standard deviations for 95% of the samples the median. 8 gDT is sufficient the REiVgDT REiD .In that case, where REi, DT D mingfDTg, and 0:95jfREigj; I is the training data's i-th RE. Then an anomaly is found.whenever RE.t/ DT. The following is a summary of the steps: data from a time series are transformed into a symbolic order. DBN is used to model the interactions between the subsystems. Utilizing MI, assess the information-based metric values (I ij).Utilizing I ij, create a binary vector of length L and give each I ij a state of 0 or 1.To learn the behaviour pattern, use RBM with visible nodes corresponding to APs and RPs.Using trained RBM, determine the likelihood that the current observation will result in an anomaly.The algorithm for the anomaly detection process is described

CASE STUDIES AND SIMULATION RESULTS

A distributed control agent, a measurement device, a generator as described in Section II, and an energy storage system are all included in each agent in Case 1's multi-agent cyber-physical system model, which is based on the IEEE-39 bus model. The energy that can be introduced into the system through various micro grids or renewable sources is represented by energy

storage. All case studies undergo the same analysis, but for reasons of space, only Case 1's findings are presented in this section.

A)TESTING SYSTEM

Table 2 includes a list of case study specifics that were taken from Matpower [42]. It is assumed that every case study is completely observable. A level of security is added to the measurement model to ensure the accuracy of the historical data.Since large power grids have thousands of metres, protecting measurements is very expensive.Based on the best PMU placement, we identify the critical metres to protect in order to lower the cost [31].We also assume that over the course of a typical day, the system topologies will not change. Case studies are carried out in Matlab R2017a on a computer with a Core(TM) i7-7700 CPU running at 3.6 GHz and 32.00 GB of RAM. It should be noted that the data are classified as either normal or anomalous. The baseline for the normal condition, which will be used to choose the threshold for the anomaly, is obtained using training data. The distribution Q representing the dynamic behaviour of the system is computed using a moving window in a subset of the training data (with distribution P). In each subset, the RE metric is used to calculate the separation between Q and P. The testing data is set up similarly. Finally, a comparison of the two RE is done to look for anomalous conditions (cyber-attack in our case). It is clear that all of the measurements residuals caused by cyberattacks are almost the same size as the measurement residuals under normal operating conditions, indicating that the stealthy cyberattacks cannot be detected by conventional residual tests. The measurement residual will exhibit significant residual due to faults, as shown in Fig. 5. If there is a problem with the system, the operator will be alerted and can fix it. As a result, the fault won't have an impact on the system's states. The lower plot's variation is within acceptable bounds. However, the variation between the attack's 35 and 65 samples significantly increases in the top plot. This suggests that there may be a cyberattack scenario that has evaded bad data detection. As a result, the rest of the system might be fed estimated states with high error, which could cause permanent harm.

B. ACCURACY, FALSE POSITIVE AND TRUE POSITIVE

In the analysis of the smart grid, the ability to identify cyberattacks and prevent false alarms are of utmost importance. Accordingly, the True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN) listed in Table 3 are used to analyse the performance of the proposed method. The False Positive Rate (FPR), True Positive Rate (TPR), and Accuracy (Acc) values, which measure the algorithms' memorization and learning capabilities, A low FPR of 0% indicates that no secure measurements were mistakenly identified as being under attack. 100 percent claries TPR that none of the measurements under attack are incorrectly categorised as secure. When a measurement is 100% accurate, it means that a measurement that is classified as attacked, and Secure measurements are those that have received that designation.

1) illustrates the variation of FPR as a function of detection threshold for single attack (SA) and multiple attack (MA) on state variables 2 and 4. Each case's DT, which stands for

the threshold defined in Section III, was changed from 0.25 DT to 1.5 DT. As can be seen from the graph, FPR sharply drops as detection threshold rises. This shows that when the threshold is set too low, the algorithm becomes overly aggressive in detecting attacks, which results in a high rate of false alarms.

2) The impact of attack magnitude on TPR and ACC is illustrated in Fig. 8 for two attack scenarios on state variables 2 and 4. 1 (1% of the original measurement) and 10 (10% of the original measurement), respectively, denote low and high attack magnitudes. The typical type of assault on the literature is of medium magnitude (here indicated by 5). The results are plotted for two different detection thresholds to confirm the impact of detection threshold on TPR and ACC.
3) Effect of Attack Sparsity on TPR and ACC: Attacks with varying attack sparsity/N 2 [0, 1] are generated in order to examine this effect. The system's overall measurement count, N, is represented. As can be seen in Fig. 9, TPR and ACC both rise as more measurements are contaminated. Here, sparsity 1 denotes that the attacker has altered all measurements. The gauge indicates that the the measurements are attacked, which is a reasonable assumption for the attacker's attack to be successful. from a 99% TPR perspective, the algorithm is very efficient. 98% ACC, too.

C) PERFORMANCE ANALYSIS UNDER DIFFERENT OPERATION CONDITION

Four distinct scenarios are taken into account to validate the efficacy of the proposed method: Normal circumstances without an attack, random attack, single FDI attack on line 631, multiple, simultaneous FDI attacks on lines 6-31, and normal conditions with no attack. LNR test and Chi-Square test, the two most widely used BDD approaches, are contrasted with the proposed method. To reduce false positives caused by noise, the threshold is set to 3 while is the standard deviation [44]. As a result, the FPR caused by noise is less than 1%. The threshold is normalised for all detectors in order to allow for accurate and thorough comparison. The LNR test threshold is determined using the same criterion. Refer to [20] for more details on LNR and the Chi-Square test. When everything is operating normally, none of the detectors' outputs cross the line that indicates that there was a cyberattack or bad data in the system.

Figure 10(b) demonstrates that all techniques can identify the random attack. The operator will be alerted to the presence of an attack because the attack is not intelligent and will leave its mark in the data sets. The measurement residual vector significantly changes as a result of the random bad data that was injected into the measurement set, increasing the cost function. We assess the cost function using the measurement residual in an optimal state estimation.

Without corrupt data in the system, the cost function operates normally. adheres to a normal distribution with a mean of 0. The cost function will surpass the cutoff for accurate state estimation under a random attack.

Any FDI attack on a line or system topology typically causes similar changes in the network with minor variations. As a result, the suggested method can effectively identify a variety of FDI attacks from various sources. Furthermore, the proposed

scheme's success rate is independent of attack scenarios because it examines patterns between compromised and uncompromised data.

CONCLUSION

The solutions suggested in the literature are primarily online approaches with limitations to deal with dynamically evolving online threats in the context of smart grid anomaly detection.

In order to find causal interactions between the subsystems, this paper proposes a real-time and computationally efficient tool for anomaly detection. It uses a feature extraction scheme and time series partitioning. Free energy is used as the anomaly index in the DBN concept and learning algorithms based on the Boltzmann machine to detect unobservable attacks. Performance of the proposed algorithm was assessed for a number of measures using various IEEE test systems and operating conditions (TPR, FPR, and ACC). The outcomes showed that the system achieves 99% accuracy, 98% TPR, and less than 2% FPR.

REFERENCES

[1] J. E. Dagle, ``Postmortem analysis of power grid blackouts_The role of measurement systems," *IEEE Power Energy Mag.*, vol. 4, no. 5, pp. 30_35, Sep./Oct. 2006.
[2] Z. Huang, C. Wang, T. Zhu, and A. Nayak, ``Cascading failures in smart grid: Joint effect of load propagation and interdependence," *IEEE Access*, vol. 3, pp. 2520_2530, 2015.
[3] Y. Cai, Y. Li, Y. Cao, W. Li, and X. Zeng, ``Modeling and impact analysis of interdependent characteristics on cascading failures in smart grids," *Int. J. Elect. Power Energy Syst.*, vol. 89, pp. 106_114, Jul. 2017.
[4] H. Karimipour and V. Dinavahi, ``On false data injection attack against dynamic state estimation on smart power grids," in *Proc. IEEE Int. Conf. Smart Energy Grid Eng. (SEGE)*, Aug. 2017, pp. 388_393.
[5] G. Dondossola, J. Szanto, M. Masera, and I. N. Fovino, ``Effects of intentional threats to power substation control systems," *Int. J. Crit. Infrastruct.*, vol. 4, nos. 1_2, pp. 129_143, 2008.
[6] T. Morris, S. Pan, J. Lewis, J. Moorhead, N. Younan, R. King, M. Freund, and V. Madani, ``Cybersecurity risk testing of substation phasor measurement units and phasor data concentrators," in *Proc. CSIIRW*, Oct. 2011, Art. no. 24.
[7] A. Ameli, A. Hooshyar, E. El-Saadany, and A. M. Youssef, ``Attack detection and identi_cation for automatic generation control systems," *IEEE Trans. Power Syst.*, vol. 33, no. 5, pp. 4760_4774, Sep. 2018.
[8] H. H. Pajouh, R. Javidan, R. Khayami, D. Ali, and K.-K. R. Choo, ``A twolayer dimension reduction and two-tier classi_cation model for anomalybased intrusion detection in IoT backbone networks," *IEEE Trans. Emerg. Topics Comput.*, to be published.
[9] R. Khan, K. Mclaughlin, D. Laverty, and S. Sezer, ``Design and implementation of security gateway for synchrophasor based real-time control and monitoring in smart grid," *IEEE Access*, vol. 5, pp. 11626_11644, Jun. 2017.
[10] R. Xu, R. Wang, Z. Guan, L. Wu, J. Wu, and X. Du, ``Achieving ef_cient

detection against false data injection attacks in smart grid," *IEEE Access*, vol. 5, pp. 13787_13798, Jul. 2017.
[11] C.-C. Sun, A. Hahn, and C.-C. Liu, ``Cyber security of a power grid: Stateof-the-art," *Int. J. Elect. Power Energy Syst.*, vol. 99, pp. 45_56, Jul. 2018.
[12] X. Liu, N. Xiong, N. Zhang, A. Liu, H. Shen, and C. Huang, ``A trust with abstract information veri_ed routing scheme for cyber-physical network," *IEEE Access*, vol. 6, pp. 3882_3898, 2018.
[13] C. Alcaraz, C. Fernandez-Gago, and J. Lopez, ``An early warning system based on reputation for energy control systems," *IEEE Trans. Smart Grid*, vol. 2, no. 4, pp. 827_834, Dec. 2011.
[14] X. He, L. Chu, R. C. Qiu, Q. Ai, and Z. Ling, ``Anovel data-driven situation awareness approach for future grids_Using large random matrices for big data modeling," *IEEE Access*, vol. 6, pp. 13855_13865, 2018.
[15] A. Farraj, E. Hammad, and D. Kundur, ``A distributed control paradigm for smart grid to address attacks on data integrity and availability," *IEEE Trans. Signal Inf. Process. Over Netw.*, vol. 4, no. 1, pp. 70_81, Mar. 2018.
[16] I. Friedberg, X. Hong, K. Mclaughlin, P. Smith, and P. C. Miller, ``Evidential network modeling for cyber-physical system state inference," *IEEE Access*, vol. 5, pp. 17149_17164, 2017.
[17] A. Azmoodeh, A. Dehghantanha, M. Conti, and K.-K. R. Choo, ``Detecting crypto-ransomware in IoT networks based on energy consumption footprint," *J. Ambient Intell. Humanized Comput.*, vol. 9, no. 4, pp. 1141_1152, 2018.
[18] N. Milosevic, A. Dehghantanha, and K.-K. R. Choo, ``Machine learning aided Android malware classi_cation," *Comput. Elect. Eng.*, vol. 61, pp. 266_274, Jul. 2017.
[19] H. Karimipour and V. Dinavahi, ``Robust massively parallel dynamic state estimation of power systems against cyber-attack," *IEEE Access*, vol. 6, pp. 2984_2995, Dec. 2017.
[20] M. Ozay, I. Esnaola, F. T. Y. Vural, S. R. Kulkarni, and H. V. Poor, ``Machine learning methods for attack detection in the smart grid," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 27, no. 8, pp. 1773_1786, Aug. 2016.
[21] S. Pan, T. Morris, and U. Adhikari, ``Developing a hybrid intrusion detection system using data mining for power systems," *IEEE Trans. Smart Grid*, vol. 6, no. 6, pp. 3104_3113, Nov. 2015.
[22] J. Landford, R. Meier, R. Barella, X. Zhao, E. Cotilla-Sanchez, R. B. Bass, and S.Wallace, ``Fast sequence component analysis for attack detection in synchrophasor networks," in *Proc. 5th Int. Conf. Smart Cities Green ICT Syst. (SmartGreens)*, Rome, Italy, 2016.
[23] S. Ahmed, Y. Lee, S.-H. Hyun, and I. Koo, ``Feature selection_based detection of covert cyber deception assaults in smart grid communications networks using machine learning," *IEEE Access*, vol. 6, pp. 27518_27529, 2018.
[24] Y. He, G. J. Mendis, and J.Wei, ``Real-time detection of false data injection attacks in smart grid: A deep learning-based intelligent mechanism," *IEEE Trans. Smart Grid.*, vol. 8, no. 5, pp. 2505_2516, Sep. 2017.
[25] D. Codetta-Raiteri and L. Portinale, ``Dynamic Bayesian networks for fault detection, identi_cation, and recovery in autonomous spacecraft," *IEEE Trans. Syst., Man, Cybern. Syst.*, vol. 45, no. 1, pp. 13_24, Jan. 2015.
[26] S. Mohammadi, H. Mirvaziri, M. Ghazizadeh-Ahsaee, and H. Karimipour,

``Cyber intrusion detection by combined feature selection algorithm,"
J. Inf. Secur. Appl., vol. 44, pp. 80_88, Feb. 2019.
[27] C. A. Murthy, ``Bridging feature selection and extraction: Compound
feature generation," *IEEE Trans. Knowl. Data Eng.*, vol. 29, no. 4,
pp. 757_770, Apr. 2017.
[28] H. Karimipour and V. Dinavahi, ``Extended Kalman _lter-based parallel
dynamic state estimation," *IEEE Trans. Smart Grid*, vol. 6, no. 3,
pp. 1539_1549, May 2015.
[29] J. D. Glover, M. Sarma, and T. Overbye, *Power System Analysis and
Design*, 5th ed. Boston, MA, USA: Cengage, 2011.
[30] A. R. Bergen and V. Vittal, *Power Systems Analysis*, 2nd ed.
Upper Saddle River, NJ, USA: Prentice-Hall, 2000.
[31] A. Abur and A. Gómez-Expósito, *Power System State Estimation: Theory
and Implementation*. New York, NY, USA: Marcel Dekker, 2004.
[32] Y. Liu, P. Ning, and M. K. Reiter, ``False data injection attacks against state
estimation in electric power grids," *ACM Trans. Inf. Syst. Secur.*, vol. 14,
no. 13, pp. 1_33, May 2011.
[33] M. Ozay, I. Esnaola, F. T. Yarman Vural, S. R. Kulkarni, and H. V. Poor,
``Sparse attack construction and state estimation in the smart grid: Centralized
and distributed models," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 7,
pp. 1306_1318, Jul. 2013.
[34] A. Ray, ``Symbolic dynamic analysis of complex systems for anomaly
detection," *Signal Process.*, vol. 84, no. 7, pp. 1115_1130, 2004.
[35] C. Rao, A. Ray, S. Sarkar, and M. Yasar, ``Review and comparative evaluation
of symbolic dynamic _ltering for detection of anomaly patterns,"
Signal, Image Video Process., vol. 3, no. 2, pp. 101_114, 2009.
[36] S. Sarkar, S. Sarkar, K. Mukherjee, A. Ray, and A. Srivastav, ``Multi-sensor
information fusion for fault detection in aircraft gas turbine engines,"
J. Aerosp. Eng., vol. 227, no. 12, pp. 1988_2001, Dec. 2013.
[37] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed.
Hoboken, NJ, USA: Wiley, 2006.
[38] C. Liu, A. Akintayo, Z. Jiang, G. P. Henze, S. Sarkar, ``Multivariate
exploration of non-intrusive load monitoring via spatiotemporal pattern
network," *Appl. Energy*, vol. 211, pp. 1106_1122, Feb. 2018.

[39] C. Liu, S. Ghosal, Z. Jiang, and S. Sarkar, ``An unsupervised anomaly
detection approach using energy-based spatiotemporal graphical modeling,"
Cyber-Phys. Syst., vol. 3, nos. 1_4, pp. 66_102, 2017.
[40] B. J. Frey and N. Jojic, ``A comparison of algorithms for inference and
learning in probabilistic graphical models," *IEEE Trans. Pattern Anal.
Mach. Intell.*, vol. 27, no. 9, pp. 1392_1416, Sep. 2005.
.
[41] Computational engine performance and emission analysis using Ceiba Pentandra biodiesel(Elservier).Panneerselvam,NMurugesan,A,Porkodi,KP,Jima,Terefe ,Vijayakumar,C Subramanian,Biofuels,Volume.7,Issue.3,2015,PP.201-206.Efficient Classification of Heart Disease using machine learning Algorthim(Scopus).,Dr K.P.Porkodi., Journal of Xi'an Shiyou University,Natural scienceEdition.,Volume 17,Issue 07,PP-120-122.

[42] K.P.PORKODI, A.M.J.MD.ZUBAIRRAHMAN "A Survey of Underwater Wireless Sensor Networks and its Challenges." *Asian Journal of Research in Social Sciences* Vol.6 (2016).,pp.594-604.

[43]"H. Anandakumar, R. Arulmurugan, and C. C. Onn, "Big Data Analytics for Sustainable Computing," *Mobile Networks and Applications*, vol. 24, no. 6, pp. 1751–1754, Oct. 2019.

[44] A. Haldorai, A. Ramu, and C.-O. Chow, " Big Data Innovation for Sustainable Cognitive Computing," *Mobile Networks and Applications,* vol. 24, no. 1, pp. 221–223, Jan. 2019.