IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received April 9, 2018, accepted May 7, 2018, date of publication May 11, 2018, date of current version June 5, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2835527

Feature Selection–Based Detection of Covert Cyber Deception Assaults in Smart Grid Communications Networks Using Machine Learning

SAEED AHMED, YOUNGDOO LEE, SEUNG-HO HYUN, AND INSOO KOO
School of Electrical Engineering, University of Ulsan, Ulsan 44610, South Korea

Corresponding author: Seung-Ho Hyun (takeitez@ulsan.ac.kr)

This work was supported by the Basic Science Research Program through the National Research Foundation of Korea funded by the Ministry of Education under Grant 2015R1D1A1A09057077 and the Korean Government (MSIT) under Grant 2018R1A2B6001714.

ABSTRACT The integration of computing and modern wireless communications techniques is enabling prolific intelligent monitoring and efficient control of electric power systems in the frameworks of smart grids. In parallel, an enhanced reliance on such technologies has increased the susceptibility of today's smart grids to cyber-assaults. Recently, a new type of assault, termed *covert cyber deception assault*, has been introduced to infringe upon the integrity of smart grid data. Such assaults are designed and initiated by hackers who have considerably good knowledge of the power network topology and the security measures in place, and therefore, these assaults cannot be effectively detected by the bad-data detectors in traditional state estimators. In this paper, we propose a supervised machine learning–based scheme to detect a covert cyber deception assault in the state estimation–measurement feature data that are collected through a smart-grid communications network. The distinctive characteristic of the paper is that we use a genetic algorithm–based feature selection in our scheme to improve detection accuracy and reduce computational complexity. The proposed detection scheme is evaluated using standard IEEE 14-bus, 39-bus, 57-bus, and 118-bus test systems. Through performance analysis, it is shown that the proposed scheme provides a significant improvement in *covert cyber deception assault* detection accuracy, compared with existing machine learning–based schemes.

INDEX TERMS Cyber assaults, feature selection, genetic algorithm, machine learning, smart grids, state estimation, support vector machines.

I. INTRODUCTION

Rapid growth in human population, increased consumerism, and induction of renewable energy has multiplied the challenges for the electric power industry. These challenges have given birth to the idea of transition from traditional power grids to a new paradigm of the SG. The concept of the SG as a complex cyber-physical system is to insert adequate intelligence to augment control of the traditional electric power grid and make it more autonomous, fault-tolerant, reliable, and efficient. Because bulk storage of the generated electric energy is not possible, generation and consumption should be closely equated; otherwise, there can be a deviation in the electrical quantities. Thus, a PCC needs to closely monitor the power network to make sure that the operation of the power system is safe and reliable. The SE is a well-organized method for online monitoring of states in power networks. The fundamental building blocks, i.e., generation, transmission, and consumption, of an SG (along with the communications links) are illustrated in Figure 1. Distributed sensors, actuators, and meters (referred to as RTUs) are installed in the electric power grid, mainly in substations, to collect the measurements, including power injections into the buses and power flow in the branches. These measurements are combined at the PCC via communications links and are further used to estimate the states, i.e., voltage magnitudes and angles, at buses. These state variables form the basis for correct decisions by the EMS about AGC and OPF to maintain the electric power systems in a safe operating zone.

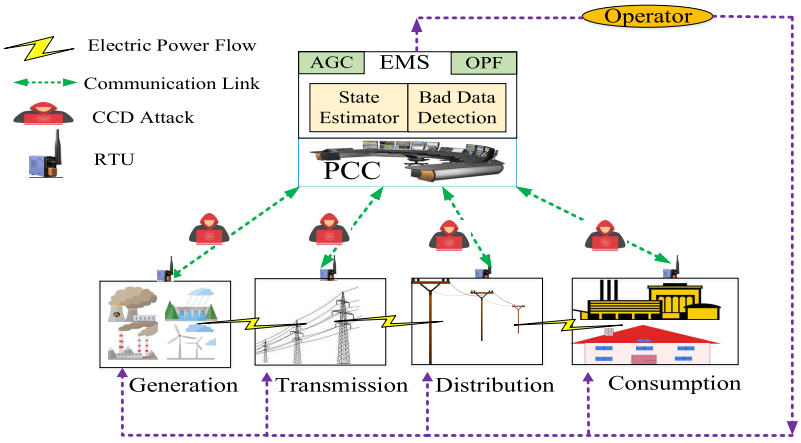


FIGURE 1. Covert cyber deception assault in a smart grid communications network.

On one hand, the existence of a communications infrastructure is compulsory for realization of efficient monitoring and intelligent control in the framework of an SG, but a communications infrastructure is prone to malicious cyber-assaults [1]–[3], owing to certain incentives for the attacker, like fiscal benefits, inserting technical faults resulting in partial or complete power blackouts, or a combination of both. A substantial amount of sensed information and control signals flow on the bi-directional communications network in SGs. Therefore, it becomes important to study a special type of malicious user behavior that attempts to violate the integrity of the measurement data by inserting a deceptive bias value into the state estimation. Such malicious behavior is mostly undetectable by the BDD present in the legacy PCC. We call this kind of attack a CCD assault. Due to its grave, negative impact on correct decisions in the PCC, detection and elimination of the susceptibilities injected by a CCD assault on the measurement data are critical for the safe and reliable operation of SGs. Normal data that are not affected by a CCD assault are consistent with electrical laws, like Kirchhoff's current and voltage laws, whereas data that are affected by a CCD assault are inconsistent with these laws. This fundamental distinction between normal and compromised data inspires ML-based algorithms for detection of CCD assaults.

Detection mechanisms against the CCD assaults have been developed along multiple directions. Numerous CCD assault detection techniques, which are not based upon ML rules, have been proposed in the literature [4]–[7]. However, ML-based schemes are gaining the attention of the researchers due to their effectiveness in classifying the data that have different underlying distributions. In the context of CCD assaults on SG, proficient ML-based detection schemes have been proposed by the researchers [8]–[12]. With the growing size of power systems, the curse of dimensionality [13] becomes challenging in CCD assault detection using ML-based algorithms. In summary, the dimensionality issue has not been addressed in above mentioned works with the feature selection prospective.

Unlike prior efforts, in this paper we use a GA-based FS technique to tackle the curse of dimensionality [13]. The optimal features selected from the SE-MF dataset are then used as input by an SVM classifier for the detection of a CCD assault. The feature selection-based method does not alter the original representation of the data [14]. Contributions of this paper can be summarized as follows:

- We study the CCD assault on SE-MF dataset, launched by a hacker who is equipped with knowledge of the topology of the power system network, and we investigate how such an attack goes undetected in legacy systems that use bad-data detectors.
- We propose an ML-based solution to detect the CCD assault. To tackle the increasing computational complexity with growing sizes of power systems, we use a GA for the selection of independent and discriminative features from the SE-MF dataset. The selection of discriminative features leads to lower computational costs, a shorter time delay, and improved accuracy. Then, the selected optimal features are used as input to a binary SVM classifier to detect the presence of compromised data.
- We use IEEE standard 14-bus, 39-bus, 57-bus, and 118-bus test systems to evaluate the efficiency of proposed FS-based ML scheme for identification of CCD assaults. Performance evaluation shows that the proposed scheme can provide better accuracy in comparison to the existing machine learning techniques for CCD assault detection.

The remainder of this paper is organized as follows. In section II, we survey some related works, and in section III, we present both the system model and the construction of covert cyber-assault vectors. In section IV, we first describe the proposed GA-based feature-selection scheme, and then the SVM-based detection scheme to detect a CCD assault. Simulation results are presented in Section IV. We conclude the paper in Section V. The abbreviations and notations used in this paper are summarized in Table 1 and 2, respectively.

TABLE 1. Nomenclature.

AGC	auto generation control	ML	machine learning
ADAB	Adaboost	MLP	multi layer perceptron
BDD	bad data detection	NB	Naive Bayesian
CCD	covert cyber deception	OPF	optimal power flow
CSD	cyber stealthy deception	PCA	principle component analysis
EMS	energy management system	PCC	power control center
FDI	false data injection	ROC	receiver operating characteristic
FE	feature extraction	SE	state estimation
FS	feature selection	SG	smart grid
GA	genetic algorithm	RTU	remote terminal unit
KNN	K nearest neighbours	SVM	support vector machine
LR	load redistribution	MF	measurement features

TABLE 2. Notations in CCD assaults in SG networks.

Symbol	Quantity	Explanation
Z_{meter}	True Measurements	Measurements received from meters.
$Z_{estimated}$	Estimated Measurements	Using weighted least square (WLS).
δ	$\delta = [\delta_2, \delta_3, ..., \delta_n]^T$	DC model: voltage angles are states.
AC model	$Z_{meter} = h(\delta) + e$	$h(\delta)$:nonlinear measurement function
DC model	$Z_{meter} = H(\delta) + e$	H is Jacobian matrix .
e	$e = [e_1, e_2, ..., e_m]^T$	Gaussian error
R	$R = Z_{meter} - Z_{estimated}$	Residual$(m \times 1)$, $R = (I - A)Z_{meter}$
$Z_{assault}$	$Z_{assault} = H\delta + a + e$	Assailed measurements.
E	$E = (H^T \Omega H)^{-1} H^T \Omega$	Pseudo-Inverse of H, since $EH = I$
a	$a = Hc$	Non-zero attack vector.
Ω	$(m \times n)$ co-variance matrix	For measurement errors
m	Power system measurements	Number of real-time measurements
n	States/ buses	$(n - 1)$ number of states(buses);

II. RELATED WORK

The opportunities and challenges in employing communication techniques along with legacy power networks have been extensively surveyed [15]–[19]. Intrusion into a communications network by a malicious user who is aiming to destroy the integrity of the data can have a catastrophic impact on the secure and reliable operation of an SG [20]–[22]. Therefore, in the context of the security of SGs, understanding the nature of the assault and identification of compromised data has been the focus of research in electric power systems. The conventional state estimator in a PCC utilizes the BDD to separate bad data for state estimation. However, Liu *et al.* [23] demonstrated that a smart attacker who has information on the network topology can realize the construction of a set of falsified data that can dodge the legacy BDD. This type of attack is known as an unobservable (or covert) cyber-assault, a CSD attack, an FDI attack, a malicious data attack, an LR attack, data integrity assault, and so on [24].

Numerous schemes considering the construction of intrusion assaults against state estimation, and the subsequent defense measures against them, have been discussed in the literature [4]–[12], [19]–[26]. Xie *et al.* [27] demonstrated that a data integrity assault can methodically result in considerable economic loss in real-time market operations. Similarly, Esmalifalak *et al.* [28] studied the economic impact of a false data injection attack on electric power market operations. An encryption-based security mechanism integrated into power system devices was proposed [6] to improve the security of the power system against FDI attacks. Li *et al.* [4] proposed a decentralized conjunctive rule–based majority voting algorithm to detect compromised or assaulted phase measurement units. Huang *et al.* [5] proposed cumulative sum hypothesis test-based bad-data detection in a state estimator.

Recently, ML-based techniques have gained the attention of researchers to identify and eliminate faults, intrusions, and abnormalities in many fields. Machine learning algorithms are finding application areas in the field of SG security as well. Detection of pernicious activity at the network layer of SG communications using ML schemes is investigated in [8] and [9]. Ozay *et al.* [10] employed a variety of ML-based schemes to detect a CCD assault on SG communications at the physical layer level. The curse of dimensionality [13] becomes crucial with the growing power system. Ozay *et al.* in [10] processed the samples in small sizes by selecting a single measurement vector as a sample and do not use the FS or FE-based ML schemes to tackle the dimensionality issue. For growing size of the power system, however, they only mentioned at high level of abstract, without specifying underlying FS scheme, that the FS method may be a promising direction for handling the dimensionality issue in the context of SG security. Wang *et al.* [11] developed an ML-based algorithm to detect an attack (dissimilar to the CCD or FDI assault), termed as time synchronization attack, in which the adversaries change the time stamps of the SE measurements. Furthermore, tackling the curse of dimensionality has not been emphasized in their work. Esmalifalak *et al.* [12] proposed a machine learning-based method for the detection of a stealthy data-injection assault in SE for SGs. To cope with the problem of dimensionality using FE, Esmalifalak *et al.* [12] employed PCA to allow the transformation of the original SE-MF data to a new representation in a low-dimensional space. This new representation describes most (but not all) of the variance within the features of SE-MF data. The authors proposed a distributed SVM based algorithm for the classification of the compromised and normal data samples in SE-MF data.

Unlike the above-mentioned approaches, in this paper, we focus on feature selection to improve the classification accuracy and reduce the computational complexity and associated time-delay at PCC. Feature selection is a special method for dimensionality reduction, in which a subset of the original set of features is selected without any transformation to a low-dimensional space, i.e., the features in the original set that represent measurements of physical quantities retain their units. Only selected are those features of the SE-MF dataset that are discriminative and that can be used to accurately differentiate between compromised and uncompromised data. We employ a GA-based FS technique to select the most discriminative features, and we then employ an SVM-based ML algorithm on the selected features to detect CCD assaults. Performance evaluation shows that proposed scheme results in good accuracy in detecting the presence of bad data.

III. CONSTRUCTING THE COVERT CYBER DECEPTION ASSAULT

State estimation at the PCC is the fundamental mechanism to maintain reliable and efficient operations of SG systems [29]. As illustrated in Figure 1, the measurement data collected

from RTUs via communications networks are used by the state estimator to determine the system state over time. The problem with state estimation is how to approximate power system state variables $\delta = [\delta_1, \delta_2, \delta_3, ..., \delta_n]^T$, based on the meter measurements $Z_{meter} = [Z_1, Z_2, Z_3, ..., Z_m]^T$ where n and m are positive integers and $\delta_i, Z_j \in \mathbb{R}$ for $i = 1, 2, \ldots, n$ and $j = 1, 2, \ldots, m$. The measurement data and the state variables are related through the following AC power flow observation model:

$$Z_{meter} = h(\delta) + e, \tag{1}$$

where $h(\delta)$ is a non-linear relationship between Z_{meter} and δ; $e = [e_1, e_2, \ldots, e_m]^T$ is Gaussian measurement noise with co-variance matrix σ.

Using a linear or DC power flow model, the observation model in (1) becomes further simplified with a small sacrifice of accuracy, as follows [30], [31]:

$$Z_{meter} = H\delta + e. \tag{2}$$

In a DC power flow problem, the Jacobian matrix H can be approximated as follows:

$$H = \left.\frac{\partial h(\delta)}{\partial \delta}\right|_{\delta=0}, \tag{3}$$

where H is composed of topology and impedance data only. One objective of (2) is to determine the estimated state, $\hat{\delta}$, that is the best fit for the meter measurements. In other words, we can say that the best estimated value can minimize estimation weighted least square (WLS) error$\left(Z_{meter} - H\hat{\delta}\right)^T \Omega \left(Z_{meter} - H\hat{\delta}\right)$. By applying the weighted least square statistical estimation criteria, estimated voltage phase angle is given as follows:

$$\hat{\delta} = (H^T \Omega H)^{-1} H^T \Omega Z_{meter} = E Z_{meter}. \tag{4}$$

Here $E = (H^T \Omega H)^{-1} H^T \Omega$ and Ω is a diagonal matrix where elements are reciprocals of the variances of meter errors. That is,

$$\Omega = \begin{bmatrix} \sigma_1^{-2} & & & \\ & \cdot & & \\ & & \cdot & \\ & & & \sigma_m^{-2} \end{bmatrix}, \tag{5}$$

where σ_i^{-2} is the variance of the i-th RTU ($1 \le i \le m$).

A. ROLE OF THE CONVENTIONAL BAD-DATA DETECTOR IN STATE ESTIMATORS AT THE PCC

Environmental or medium noise in wireless communications, erroneous meters, or malicious user behavior (like CCD assaults) can be potential reasons for bad data in estimated measurements. Current power systems use a residual-based detector for BDD to protect state estimation [32]. The difference between observed meter measurements Z_{meter} and estimated measurements $\hat{Z}$ is the residual, R, and it is expressed as follows:

$$R = Z_{meter} - Z_{estimated} = (I - M)Z_{meter}. \tag{6}$$

The expected value and the co-variance of the residual are

$$\begin{aligned} E(R) &= 0, \\ cov(R) &= (I - A)R. \end{aligned} \tag{7}$$

The detection of false data or outliers is performed in BDD using the largest normalized residual (LNR) test proposed by Monticelli [32] with a predefined threshold. Therefore, the hypothesis of not being attacked is accepted if we have

$$\max_i |R_i| \le \lambda, \tag{8}$$

where R_i is the component of residual vector R and λ is the threshold.

B. THE COVERT CYBER DECEPTION ASSAULT CAN CIRCUMVENT THE CONVENTIONAL BDD

Malicious users can launch an assault if they are familiar with the topology of H. With knowledge of the H matrix, an attacker can alter the value of the meter measurement data. Let $Z_{assault} = Z_{meter} + a$ where $a \in \mathbb{R}^{m\times 1}$ denotes the malicious data injected into the meter measurement data vector. If the malicious user constructs vector a as follows:

$$a = Hc, \tag{9}$$

where $c \in \mathbb{R},^{m\times 1}$ is any arbitrary non-zero vector, the legacy BDD cannot detect such an assault. The reason is as follows: Let $\hat{\delta}_{asssault}$ denote the estimate of state variables using assaulted meter measurements $Z_{assault}$, i.e.,

$$\hat{\delta}_{asssault} = EZ_{meter} + Ea = \hat{\delta} + EHc = \hat{\delta} + c. \tag{10}$$

Now, the L_2 norm for the assaulted measurement $Z_{assault}$ residual is as follows:

$$\begin{aligned} \|R_{assault}\|_2 &= \|Z_{assault} - H\delta_{assault}\|_2 \\ &= \left\|(Z + a) - H(\hat{\delta} + c)\right\| \\ &= \left\|\left(Z - H\hat{\delta}\right) + (a - Hc)\right\|_2 = \left\|\left(Z - H\hat{\delta}\right)\right\|_2 \\ &= \|R\|_2. \end{aligned} \tag{11}$$

The assaulted measurement residual calculated here is the same as that without compromised data. Hence, $Z_{assault}$ will be able to deceive the BDD if original meter measurements Z_{meter} can pass the BDD. The compromised measurements are modeled as

$$Z_{assault} = H(\delta + \Delta\delta) + e. \tag{12}$$

Equation (12) shows that assaulted or corrupted meter measurement data results in the addition of $\Delta\delta$ to estimated state. Because the residual of the assaulted measurements is the same as the one without any assault, the BDD statistical test given in eq (8) will be futile to detect the assault which will change the system states affecting crucial operational failures. This sort of assault is termed an unobservable or covert assault [23] and [24]. Under these assumptions, the observation model in the presence of the CCD assault can be described as the following:

$$Z_{assault} = H\delta + a + e, \tag{13}$$

where a is the non-zero assault vector.

IV. MACHINE LEARNING–BASED BAD DATA DETECTION

In this section, we discusses the proposed ML-based scheme to detect CCD attacks. The proposed methodology for classification of the assaulted and unassailed SE-MF dataset is illustrated in Figure 2. The fact that normal or unassailed data follow Kirchhoff's law, and the assaulted data do not follow any physical law, suggests that both types of data will have different distributions and will therefore tend to form different clusters. These clusters would be distinguishable in a feature space of suitable dimensions. Furthermore, if the data are supplemented with class labels, then a classifier can be trained to distinguish between the two clusters (assaulted and unassailed). The curse of dimensionality [13] becomes challenging when the size of measurement features grows with an increase in the size of the power system, which results in greater computational complexity. Nonetheless, not all SE-MF dataset attributes would be equally supportive in leading to plainly distinguishable clusters in the feature space; this can have a negative impact on the classifier's performance. In this paper, we utilize GA-based FS scheme to select an optimal subset of features that would result in more tightly packed and distinctly separable clusters of vectors of chosen features in the resulting subspace. Thus, the predictive performance of the classifier is improved. Furthermore, FS reduces the measurement and storage requirements, as well as the training and prediction times [33]. In this paper, a two-level scheme for the detection of CCD assaults in the SE-MF dataset is proposed. First, the GA is used to select the optimal-feature subset from the SE-MF dataset. The selected optimal features are then employed to train an SVM-based classifier that can detect CCD assaults on the SE-MF dataset. In the following subsections, we explain the GA and support vector machine-based classification.

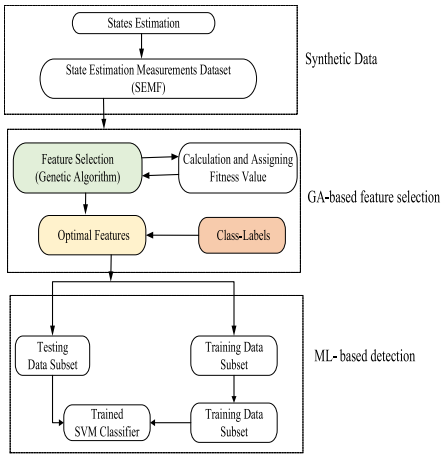


FIGURE 2. The main flowchart of the proposed machine learning-based covert cyber deception assault detection scheme.

A. DIMENSIONALITY REDUCTION USING GENETIC ALGORITHM-BASED FEATURE SELECTION

Feature size in the SE-MF dataset increases as the power system's size grows, and dimensionality reduction becomes an obvious necessity in data mining to reduce the computational costs and improve the efficiency of the classifier, which can be affected by irrelevant and redundant features. Esmalifalak *et al.* [12] used a PCA-based approach for dimensionality reduction, which transforms the original features from a high dimensional space to a low dimensional projection space. This transformation of features to a lower dimensional space may result in some loss of information. On the other hand, feature selection is the process of selecting the best subset of features from among all the features that are useful in discriminating between the two classes. In other words, the goal of FS is to choose a subset of features from a given set of features that yields minimum classification error. Many research works on dimensionality reduction reveal that FS approaches preserve data characteristics for interpretability. In addition, FS approaches also reduce the overfitting due to less redundant data, and improve modeling accuracy [34]–[38]. Janecek *et al.* [34] studied the relation between several dimensionality reduction approaches (including feature subset selection, and feature extraction with different flavors of PCA methods) and empirically tested the effects of these methods on classification accuracy on two different types of datasets email data and drug discovery data. Results revealed that feature transformation using PCA is highly sensitive to the type of data. Many methods have been proposed by researchers for FS. In general, FS methods can be divided into three categories: filters, wrappers, and embedded/ hybrid methods [14], [39]. In this paper, we use a filter-based FS mechanism that is independent of any learning algorithm or classifier. Working as a preprocessor, it selects features by considering their scores in different statistical tests for correlation with the outcome variable. We use GA to select the subset of features of the SE-MF dataset that is the best at discriminating compromised data from normal data. The GA emulates biological evolution and Darwinian selection [40]. The evolution mechanism of living beings is believed to follow natural selection, i.e., living species that are better suited to their environment thrive, whereas species that are at a disadvantage in their environment go extinct. Following the same principle, a GA improves a given solution by incrementally choosing better possible solutions, while eliminating menial solutions. The quality of each solution is calculated using a fitness value function based on the objective function. The m-dimensional set of SE-MF vector data in $\mathbb{R}^n$ is given as input to the GA as follows:

$$\left\{ X_1^{(m)}, X_2^{(m)}, \ldots, X_k^{(m)} \right\}$$

where

$$X_i^{(m)} = \begin{bmatrix} x_1 & x_2 & \ldots, & x_{m-1} & x_m \end{bmatrix}, \quad \forall i \in \{1, 2, \ldots, k\}. \tag{14}$$

The GA yields a set of n-dimensional vectors in subspace $\mathbb{R}^n$, described as

$$\left\{ X_1^{(n)}, X_2^{(n)}, \ldots, X_k^{(n)} \right\}$$

where

$$X_i^{(n)} = \begin{bmatrix} x_1 & x_2 & , \ldots , & x_{n-1} & x_n \end{bmatrix}, \quad \forall i \in \{1, 2, \ldots, k\}. \tag{15}$$

It is notable that the GA reduces the dimensionality of each vector in the set without affecting the cardinality of the set of vectors in Eq. (14), i.e., $n \ll m$. The selected dimensions are chosen to optimize the fitness function. Hence, $X_i^{(n)}$ denotes an instance of the feature vector in the subspace that optimizes the fitness function. Fitness function F which is adopted in the paper is given as Eq.(16):

$$F = \frac{\bar{C}}{\bar{S}}. \tag{16}$$

In (16), $\bar{C}$ is the mean compactness of classes and it is expressed as follows:

$$\bar{C} = \frac{1}{L} \sum_{i}^{L} C_i, \tag{17}$$

where the mean separability, denoted by $\bar{S}$ in (16), is the separation between any two classes in an L-class problem, given as follows:

$$\bar{S} = \frac{2}{L\,(L-1)} \sum_{i \neq j}^{L} S_{ij}. \tag{18}$$

In the paper, we are dealing with a binary classification problem such that we have $L = 2$, i.e., assaulted and unassailed SE measurement data. The GA finds a feature subspace that would minimize the ratio of the mean values of inter-class separability and intra-class compactness.

- Inter-class separability to measure how well separated two different clusters are from each other.
- Intra-class compactness to measure how well clustered the sample vectors are for a given class.

Two dimensional representation of inter-class separability and intra-class compactness is illustrated in Figure 3. To measure the compactness of a given class, the GA calculates the mean or centroid $\mu^{(i)}$ of class i as follows:

$$\mu = \frac{1}{N} (X_1 + X_2 + \ldots + X_N), \tag{19}$$

where N is the total number of samples of class i. After that, the compactness of class i is determined by finding the mean value of the Euclidean norm as follows:

$$C_i = \frac{1}{N} \sum_{j=1}^{N} \left\| X_j - \mu^{(i)} \right\|. \tag{20}$$

The Euclidean distance between the centroids of the two classes describes the separability between the two classes i and j. It is determined as follows:

$$S_{ij} = \left\| \mu^{(i)} - \mu^{(j)} \right\|. \tag{21}$$

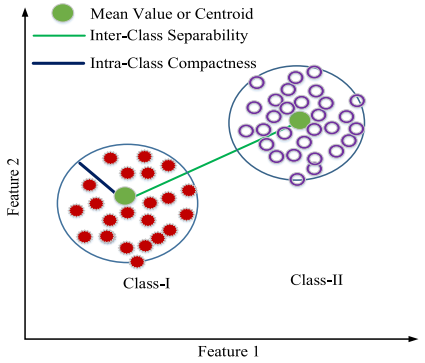


FIGURE 3. The concept of intra-class compactness and inter-class separability in a two-dimension feature space.

The GA encodes the SE-MF into chromosomes, which are going through the crossovers and mutations. Thus, new generations of chromosomes are yielded, which substitutes for their parents, provided they are healthier, i.e., their fitness or objective function value is higher. This process is iterated for many generations until there is no further improvement in the fitness function [33]. A binary encoding scheme is used to represent the features or attributes of the SEMF dataset as chromosomes. A chromosome is simply a string of binary 1's and 0's, where 1 indicates that a certain SEMF feature is selected, and 0 means it is rejected. The index of each 1 and 0 in the chromosome corresponds to a distinct SEMF attribute. In the beginning, the GA randomly selects different subsets of the SEMF. In other words, a primary population of chromosomes (a string of 1's and 0's) initiates the algorithm. A new population of chromosomes is created by subjecting the primary (parent) chromosomes to the crossover and mutations. Two parent chromosomes exchange information or swap fragments at randomly chosen crossover points during the crossover process. However, during the mutation process, the bits are flipped at randomly selected positions in a chromosome. Then, based on their respective fitness function value, chromosomes are ranked in the evaluation process. Finally, the chromosomes that minimize the proposed fitness function are selected to produce new chromosomes. This process is repeated for many generations until there is no further decrease in the value of the proposed fitness or objective function.

B. SVM-BASED CCD ASSAULT DETECTION

Originally established by Vapnik [41], the concept of SVMs is grounded in the theory of statistical learning and structural risk minimization. This machine learning method has provided accurate performance in classification and prediction problems and finds its application in the expanse of the detection field as well. In this subsection, we propose SVM-based CCD assault detection in the SE-MF dataset, which is a binary classification problem i.e., classification of assaulted and unassailed data. The binary classification problem is solved by the SVM by determining the hyperplane

with the largest margin that disconnects the two classes in the feature space of the training data. The sign of the hyperplane function is employed to determine the labels of the test samples from the assaulted and unassailed data. In this study, we use the Gaussian radial basis function as a kernel function. By means of Lagrange multipliers, the SVM algorithm can be condensed to solve the following optimization problem:

$$\arg\max \left\{ \sum \alpha_i - \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_i \alpha_j y_i y_j K(x_i, x_j) \right\},$$
$$s.t \sum_{i=1}^{N} \alpha_i y_j = 0, \quad 0 \le \alpha_i \le C \quad \forall i = 1, 2, \ldots, N. \quad (22)$$

Here x_i and y_i denote samples from the training dataset, α_i's are the Lagrangian multipliers, and C is the penalty variable for regulating the generalization performance of the SVM, and its value should be fine-tuned. $K(x_i, x_j)$ is the kernel function for the SVM derived by Mercer's Theorem [42]. The corresponding decision or classification function for the SVM is obtained as follows:

$$F(x) = sgn\{f(x)\},$$
$$where \quad f(x) = \sum_{i=1}^{N} \alpha_i^* y_i^* K(x_i^*, x) + b. \quad (23)$$

The Lagrange multiplier corresponding to the support vector, x_i^*, is denoted by α_i^* . Function value $f(x)$ has a range from $-\infty$ to $+\infty$ and represents the signed distance of unknown data sample from the decision boundary. A positive decision value for a class indicates that x is predicted to be in that class (an unassailed sample), whereas a negative value indicates otherwise [43]. The similarity between two input samples is measured using the kernel measures. As a kernel function for the SVM, the Gaussian radial basis function (RBF) kernel was used in this research. The RBF kernel is ordinarily used for linearly non-separable data and can be calculated as follows:

$$k(x_i, x_j) = \exp\left(-\omega \left\| x_i - x_j \right\|^2\right), \quad (24)$$

where $\omega = \frac{1}{2\sigma^2}$ and σ is an adjustable parameter to be carefully selected. The exponential is linear when σ is small and the higher-dimensional projection drops its non-linear potential. Conversely, the decision boundary becomes very sensitive to noise during training, when σ is large, due to the lack of regularization.

V. EXPERIMENTAL RESULTS

In this section, we evaluate the performance of the proposed FS-based CCD assault detection scheme. We have performed the simulations using MATLAB 2017b. The proposed scheme is evaluated through experiments using the standard 14-bus, 39-bus, 57-bus, and 118-bus IEEE test systems. To compare the proposed scheme with existing ML-learning algorithms, we have employed IEEE 118-bus system. Figure 4 illustrates the IEEE 39-bus system, also known as the new England

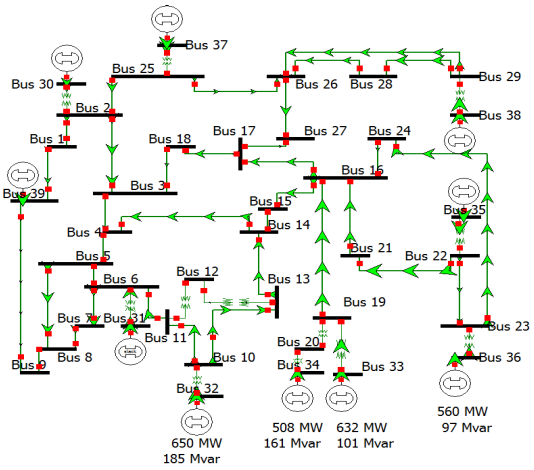


FIGURE 4. Standard IEEE 39-bus (New England 10-Machine) System [45].

10-machine system [45]. Because of space limitations, figures for other IEEE bus systems employed for testing in this work are not included. To simulate the operation of the power network, we have used the MATPOWER 6.0 toolbox [44] to generate the configuration of these test systems especially, the Jacobian matrix H. We employed the AC power flow model and used DC power flow analysis to approximate the state vectors and measurement vectors. In a B-bus system, state variable vector $\delta \in \mathbb{R}^n$ is composed of $(B-1)$ bus voltage phase angles, and the meter measurement vector consists of active power injections into the buses and branch active power flows. To perform a fair comparison with a real-world power network scenario, we have used stochastic loads with uniform load distributions, as employed in [12]. In these simulations, the active power measurement features, including the active power injections into the buses and active power flows on the branches, are input to the GA for feature selection.

A. GA-BASED FEATURE SELECTION

In this paper, the number of chromosomes in each population is 100 and the maximum number of generations is set to 80. Because the GA randomly selects different subsets of the SE-MF dataset to create a primary population of chromosomes, we iterate the GA for 30 times and choose only those features which are selected for more than 70% in these iterations. Table 3. illustrates an average number of selected features with the application of GA with 30 iterations to SE-MF dataset for various IEEE standard systems.

TABLE 3. Average number of features selected by GA.

System	States	Features	Selected Features
14-bus	13	53	23
39-bus	38	130	61
57-bus	56	216	111
118-bus	117	489	233

B. THE OPTIMAL CHOICE OF PENALTY VARIABLE (C) AND ADJUSTABLE PARAMETER (σ)

The selected optimal features by the GA are then inserted as input to the SVM for detection of compromised data in the SE-MF dataset. In this paper, using four-fold cross-validation, we employ 75% of the historical SE-MF dataset as a learning dataset and tested the accuracy of the fitted decision boundary on the remaining 25 % of the dataset. In addition, we use the radial basis function, which is given in Eq. (24). The selection of the optimal values for penalty parameter C in (22) and σ will help to improve the efficiency of the SVM in detecting compromised data for the cross-validation set. The parameter determines the smoothness of the decision boundary between the two classes, whereas the parameter σ determines the influence of a single training example. Both parameters affect the generalization performance of the SVM. Using Bayesian optimization and choosing from a wide range of values, i.e., between 10^{-5} and 10^{5}, we search for optimal values for the parameters C and σ . The optimal values for C and σ yield the minimum detection error. Figures 5(a), 5(b), 5(c), and 5(d) show the detection error with respect to different values of C and σ for standard IEEE 14-bus, 39-bus, 57-bus, and 118-bus systems, respectively. Basic performance metrices used in this work i.e., accuracy, F_1-score, and ROC curves are shown in the following subsections.

C. ACCURACY

Calculating accuracy is a standard way to evaluate learning algorithms. It is a single-number summary of the performance of the proposed algorithm and can be calculated as follows:

$$Accuracy = \left(\frac{\sum TP + \sum TN}{TotalPopulation}\right), \tag{25}$$

where true positive (TP) corresponds to the samples that the proposed algorithm detects as positive samples and that are, in fact, positive. Similarly, true negatives (TNs) are the points that the proposed algorithm detects as negative samples and that are, in fact, negative. Figure 6 illustrates the accuracy of the proposed FS-based CCD assault detection scheme in various standard IEEE bus systems according to the number of training samples. The proposed scheme was also compared with the numerous ML schemes such as ADAB, MLP [10] and FE-based SVM [12]. As a result, it can be seen that the proposed FS-based SVM with optimal C, σ, and RBF kernel outperforms other schemes and requires few number of training samples to achieve a higher accuracy. In addition, the ADAB and MLP [10] exhibit good CCD assault detection accuracy, and corresponding performance improves gradually with increasing number of training samples. However, both ADAB and MLP have slow training speed and hard to tune. The FE-based SVM [12], has lower detection accuracy compared to the proposed scheme. It is also clear that a large number of training samples are required to train the model to achieve good accuracy. The figure also shows that KNN is more sensitive to feature size and has a low detection performance for the growing size of the power system. Finally, the NB scheme shows poor detection performance.

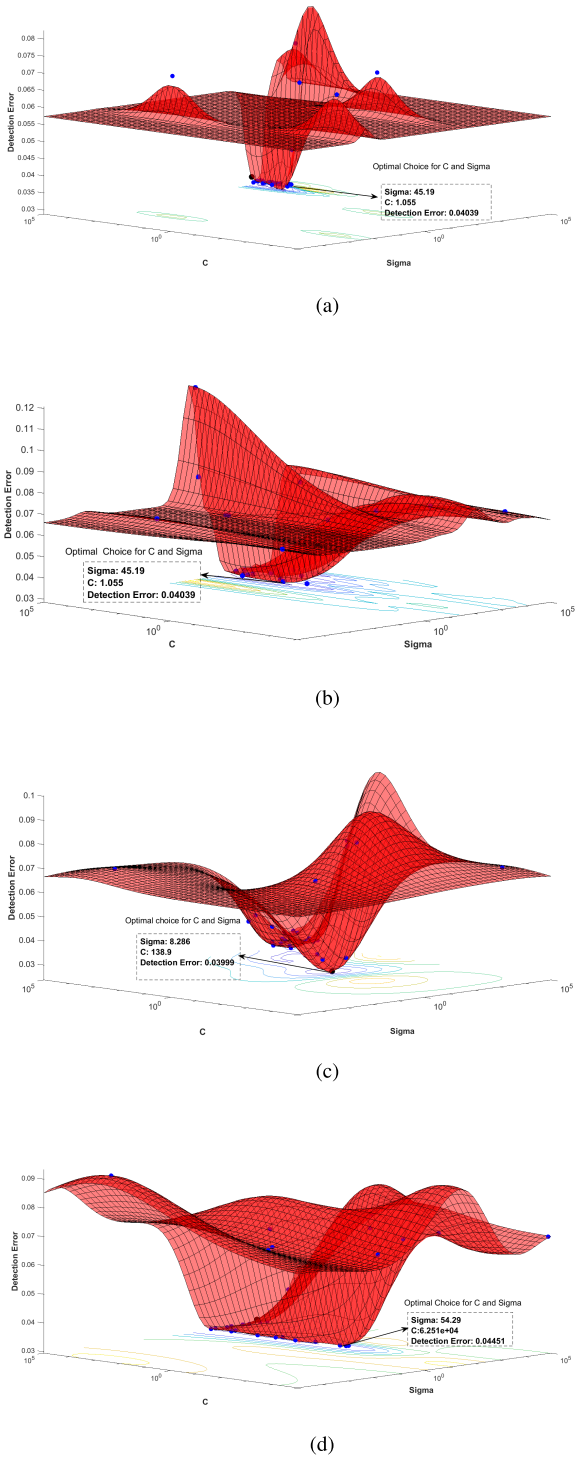


FIGURE 5. Optimal choice of C and σ for the employed IEEE systems. (a) Optimal C and σ for IEEE 14-bus system. (b) Optimal C and σ for IEEE 39-bus system. (c) Optimal C and σ for IEEE 57-bus system. (d) Optimal C and σ for IEEE 118-bus system.

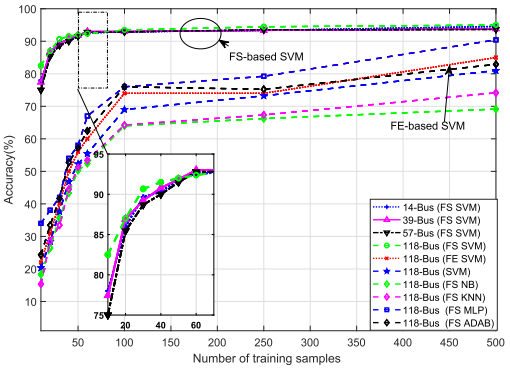


FIGURE 6. Accuracy of the feature selection-based support vector machine (SVM) in comparison with existing machine learning schemes.

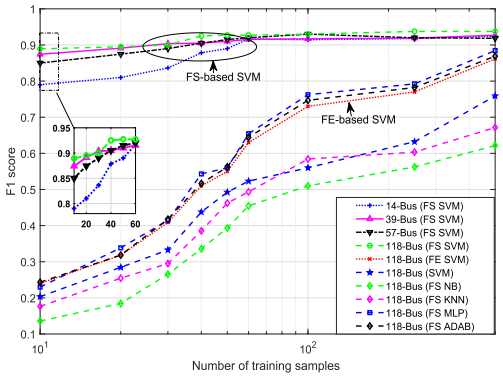


FIGURE 7. F_1 score of the proposed FS-based support vector machine(SVM) in comparison with existing machine learning schemes.

D. F_1 Score

Next, we consider the F_1 score as another metric of detection accuracy. The F_1 score is considered a measure of the precise detection or classification of the subject dataset. The F_1 score is obtained as follows:

$$F_1 = 2\left(\frac{P_r \times R_e}{P_r + R}\right), \tag{26}$$

where P_r is termed precision and is calculated as follows:

$$P_r = \left(\frac{TruePositive}{ActualPositive}\right). \tag{27}$$

The true positive corresponds to the samples that the proposed algorithm detects as positive samples, and that are, in fact, positive. Predicted positives are sample points that may include both compromised and normal ones, but the algorithm detects them all as positive. R_e is termed recall and is calculated as follows:

$$R_e = \left(\frac{TruePositive}{ActualPositive}\right). \tag{28}$$

The actual positives in (27) are all the positive points in the dataset. Generally, the F_1 score can be up to 1. In general, the closer the value is to 1, the more accurate the classifier is considered. Figure 7 shows that the proposed FS-based SVM with optimal C, σ, and RBF kernal performs well in comparison to other schemes and requires few number of training samples to achieve a higher F_1 score. It can be seen from the Figure 7 that the FE-based SVM scheme [12] requires many historical samples from the SE-MF dataset for learning in order to achieve a higher F_1 score. The ADAB and MLP (FS-based) have good CCD assault detection accuracy which improves gradually with increasing number of training samples.

However, both ADAB and MLP have slow training speed and hard to tune. Additionally, they need a large amount of data to train the model, which may require additional storage space at PCC. The figure also shows that the KNN is more sensitive to feature size and has a low detection performance for the growing size of the power system. It can be seen from the figure that Naive Bayes scheme also has poor detection performance. Next, to investigate the impact of several compromised load profiles on the F_1 score, we considered different numbers of compromised load profiles, i.e., 24, 30, 36, 40, 45, and 60. The SE-MF dataset load profiles consist of 360 samples that are collected through sensors or RTUs at regular intervals of four minutes for 24 hours. We utilize 75% of the data for training and the rest of the samples for testing. Figure 8 shows the F_1 as a measure of the accuracy of the proposed FS-based proposed detection for various standard IEEE 14-, 39-, 57- and 118-bus systems. The proposed FS-based scheme has above 90% performance for all the employed test systems. Furthermore, the average performance of our detection scheme in comparison to existing machine learning based schemes is presented using the confusion matrix in Table 4. It is evident from Table 4, that average accuracy for detecting a CCD assault is more than 90% for all the test cases. Further, the accuracy of our detection scheme improves as the system size grows.

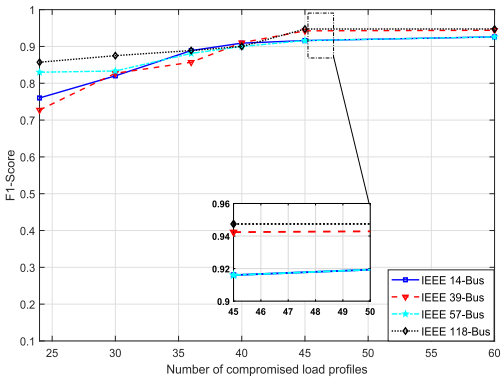


FIGURE 8. F_1 score of the proposed FS-based CCD assault detection scheme when the number of compromised load profiles changes from 24 to 60.

E. RECEIVER OPERATING CHARACTERISTIC (ROC) CURVE

Figure 9 illustrates the receiver operating characteristic curves of the proposed FS-based detection scheme for

TABLE 4. CCD assault detection accuracy comparison between proposed FS-based SVM and existing ML schemes.

Detection Scheme	Standard IEEE System	True Positive Rate (%)	True Negative Rate (%)	Accuracy (%)	F_1 score
	14-bus	76.981	94.916	92.012	0.900
	39-bus	77.324	96.545	93.912	0.912
FS-based SVM	57-bus	76.434	95.732	93.723	0.933
	118-bus	78.982	96.934	93.954	0.939
FS-based KNN	118-bus	68.378	83.165	77.234	0.691
FS-based NB	118-bus	65.435	69.776	67.321	0.631
FS-based ADAB	118-bus	74.365	93.756	85.985	0.852
FS-based MLP	118-bus	76.235	94.354	86.469	0.863
FE-based SVM	118-bus	75.234	93.653	85.714	0.847

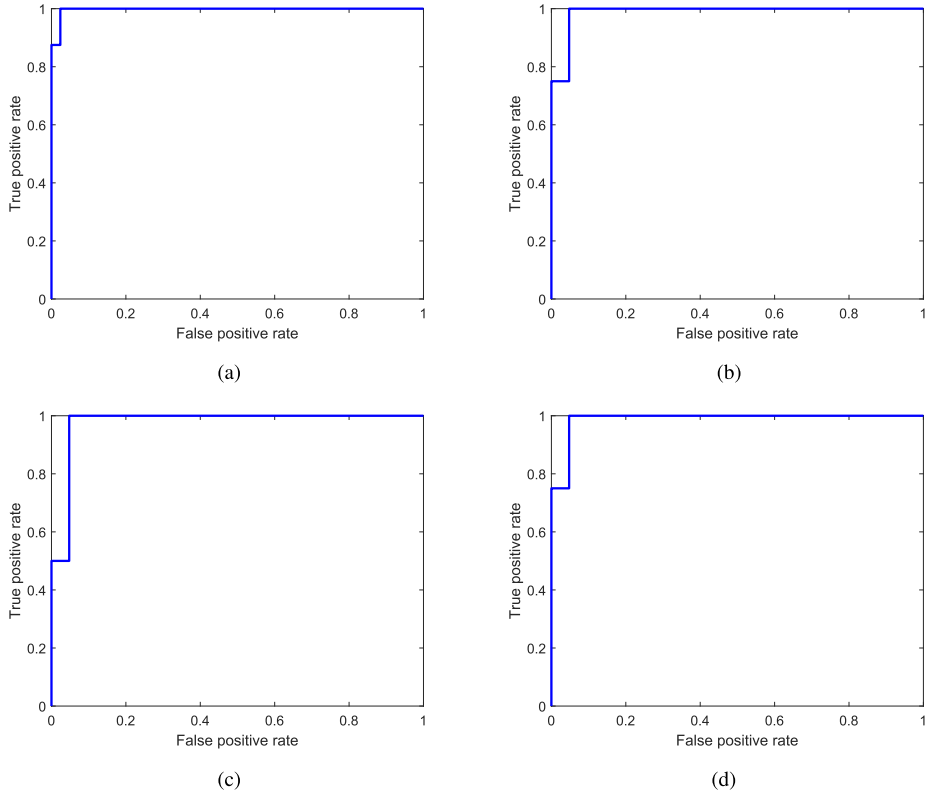


FIGURE 9. ROC curve for the proposed FS-based CCD assault detection scheme for IEEE 14, 39, 57, and 118-bus test systems. (a) IEEE 14-bus system. (b) IEEE 39-bus system. (c) IEEE 57-bus system. (d) IEEE 118-bus system.

standard IEEE 14-bus, 39-bus, 57-bus,and 118-bus systems. The ROC curve is obtained by plotting the false positive rate (FPR) versus the true positive rate (TPR). FPR is defined as the probability that normal data are identified as compromised. It is used as a measure of specificity in our detection scheme. The sensitivity of our scheme is defined as the probability that compromised data are identified as assaulted. TPR is used as a measure of sensitivity. From the figure 9, we can see that the area under the curve is approximately equal to 1 in all cases. This means that the detection accuracy of our proposed scheme is nearly equal to 1, which validates its good performance.

VI. CONCLUSION

In this work, we propose an FS-based ML mechanism for the detection of CCD assaults in SG communications networks. We employ a GA for the selection of discriminative and distinctive features. The selected optimal features are used as input to an SVM for the detection of bad data, which are inserted into the SE-MF dataset by hackers who have knowledge about the power network topology. The SVM automatically learns the decision boundary that achieves the maximum geometric deviation between unassailed and compromised data points by observing the SE-MF dataset under normal and assaulted circumstances

and then classifies test data as either compromised or uncompromised. To validate the performance of the proposed scheme, we use standard IEEE 14-bus, 39-bus, 57-bus, and 118-bus systems. The simulation results show that the proposed FS-based SVM scheme (with optimal C, σ, and RBF kernel) has reasonably good detection accuracy in comparison with the existing schemes under the occasional operational environment. AdaBoost and MLP perform well with growing power system size while requiring huge historical SE-MF dataset samples for achieving a good accuracy. We also observe that the KNN has low detection efficiency and is more sensitive to system size. Subsequently, the FS-based SVM is preferred for detection of CCD assault in SG communications network with growing size of power systems.

Finally, it would be one of the further works to consider diverse attack scenarios and to identify the compromised meters utilizing the detected compromised measurements.

REFERENCES

[1] J. Liu, Y. Xiao, S. Li, W. Liang, and C. L. P. Chen, "Cyber security and privacy issues in smart grids," *IEEE Commun. Surveys Tuts.*, vol. 14, no. 4, pp. 981–997, 4th Quart., 2012.

[2] Y. Mo *et al.*, "Cyber–physical security of a smart grid infrastructure," *Proc. IEEE*, vol. 100, no. 1, pp. 195–209, Jan. 2012.

[3] Y. Yan, Y. Qian, H. Sharif, and D. Tipper, "A survey on cyber security for smart grid communications," *IEEE Commun. Surveys Tuts.*, vol. 14, no. 4, pp. 998–1010, Oct. 2012.

[4] B. Li, R. Lu, W. Wang, and K.-K. R. Choo, "Distributed host-based collaborative detection for false data injection attacks in smart grid cyber-physical system," *J. Parallel Distrib. Comput.*, vol. 103, pp. 32–41, May 2017.

[5] Y. Huang, J. Tang, Y. Cheng, H. Li, K. A. Campbell, and Z. Han, "Real-time detection of false data injection in smart grid networks: An adaptive CUSUM method and analysis," *IEEE Syst. J.*, vol. 10, no. 2, pp. 532–543, Jun. 2016.

[6] G. Dan and H. Sandberg, "Stealth attacks and protection schemes for state estimators in power systems," in *Proc. 1st IEEE Int. Conf. Smart Grid Commun.*, Oct. 2010, pp. 214–219.

[7] H. He and J. Yan, "Cyber-physical attacks and defences in the smart grid: A survey," *IET Cyber-Phys. Syst., Theory Appl.*, vol. 1, no. 1, pp. 13–27, 2016.

[8] Z. M. Fadlullah, M. M. Fouda, N. Kato, X. Shen, and Y. Nozaki, "An early warning system against malicious activities for smart grid communications," *IEEE Netw.*, vol. 25, no. 5, pp. 50–55, Sep./Oct. 2011.

[9] Y. Zhang, L. Wang, W. Sun, R. C. Green, and M. Alam, "Distributed intrusion detection system in a multi-layer network architecture of smart grids," *IEEE Trans. Smart Grid*, vol. 2, no. 4, pp. 796–808, Dec. 2011.

[10] 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.

[11] J. Wang, W. Tu, L. C. K. Hui, S. M. Yiu, and E. K. Wang, "Detecting time synchronization attacks in cyber-physical systems with machine learning techniques," in *Proc. 37th IEEE Int. Conf. Distrib. Comput. Syst. (ICDCS)*, Jun. 2017, pp. 2246–2251.

[12] M. Esmalifalak, L. Liu, N. Nguyen, R. Zheng, and Z. Han, "Detecting stealthy false data injection using machine learning in smart grid," *IEEE Syst. J.*, vol. 11, no. 3, pp. 1644–1652, Sep. 2017.

[13] C. M. Bishop, *Neural Networks for Pattern Recognition*. London, U.K.: Oxford Univ. Press, 1995.

[14] Y. Saeys, I. Inza, and P. Larrañaga, "A review of feature selection techniques in bioinformatics," *Bioinformatics*, vol. 23, no. 19, pp. 2507–2517, 2007.

[15] F. Khan, A. U. Rehman, M. Arif, M. Aftab, and B. K. Jadoon, "A survey of communication technologies for smart grid connectivity," in *Proc. Int. Conf. Comput., Electron. Elect. Eng. (ICE Cube)*, Apr. 2016, pp. 256–261.

[16] V. C. Gungor and F. C. Lambert, "A survey on communication networks for electric system automation," *Comput. Netw.*, vol. 50, no. 7, pp. 877–897, May 2006.

[17] W. Wang, Y. Xu, and M. Khanna, "A survey on the communication architectures in smart grid," *Comput. Netw.*, vol. 55, no. 15, pp. 3604–3629, 2011.

[18] M. Emmanuel and R. Rayudu, "Communication technologies for smart grid applications: A survey," *J. Netw. Comput. Appl.*, vol. 74, pp. 138–148, Oct. 2016.

[19] W. Wang and Z. Lu, "Cyber security in the smart grid: Survey and challenges," *Comput. Netw.*, vol. 57, no. 5, pp. 1344–1371, 2013.

[20] H. Bao, R. Lu, B. Li, and R. Deng, "BLITHE: Behavior rule-based insider threat detection for smart grid," *IEEE Internet Things J.*, vol. 3, no. 2, pp. 190–205, Apr. 2016.

[21] S. Iqbal *et al.*, "On cloud security attacks: A taxonomy and intrusion detection and prevention as a service," *J. Netw. Comput. Appl.*, vol. 74, no. 2, pp. 98–120, Oct. 2016.

[22] B. Li, R. Lu, W. Wang, and K.-K. R. Choo, "DDOA: A dirichlet-based detection scheme for opportunistic attacks in smart grid cyber-physical system," *IEEE Trans. Inf. Forensics Secur.*, vol. 11, no. 11, pp. 2415–2425, Nov. 2016.

[23] Y. Liu, P. Ning, and M. K. Reiter, "False data injection attacks against state estimation in electric power grids," *ACM Trans. Inform. Syst. Secur.*, vol. 14, no. 1, p. 13, 2011.

[24] R. Deng, G. Xiao, R. Lu, H. Liang, and A. V. Vasilakos, "False data injection on state estimation in power systems—Attacks, impacts, and defense: A survey," *IEEE Trans. Ind. Informat.*, vol. 13, no. 2, pp. 411–423, Apr. 2017.

[25] M. A. Ferrag, L. A. Maglaras, H. Janicke, and J. Jiang. (Nov. 2016). "A survey on privacy-preserving schemes for smart grid communications." [Online]. Available: https://arxiv.org/abs/1611.07722

[26] T. Baumeister, "Literature review on smart grid cyber security," Collab. Softw. Develop. Lab., Univ. Hawaii, Honolulu, HI, USA, Tech. Rep. CSDL-10-11, 2010.

[27] L. Xie, Y. Mo, and B. Sinopoli, "Integrity data attacks in power market operations," *IEEE Trans. Smart Grid*, vol. 2, no. 4, pp. 659–666, Dec. 2011.

[28] M. Esmalifalak, Z. Han, and L. Song, "Effect of stealthy bad data injection on network congestion in market based power system," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Apr. 2012, pp. 2468–2472.

[29] P. Kundur, N. J. Balu, and M. G. Lauby, *Effect of Stealthy Bad Data Injection on Network Congestion in Market Based Power System*, vol. 7. New York, NY, USA: McGraw-Hill, 1994.

[30] A. Abur and A. G. Expósito, *Power System State Estimation: Theory and Implementation*. Boca Raton, FL, USA: CRC Press, 2004.

[31] F. Delea and J. Casazza, *Understanding Electric Power Systems: An Overview of the Technology, the Marketplace, and Government Regulations*. Hoboken, NJ, USA: Wiley, 2011.

[32] A. Monticelli, *State Estimation in Electric Power Systems: A Generalized Approach*. Norwell, MA, USA: Springer, 1999.

[33] S.-H. Min, J. Lee, and I. Han, "Hybrid genetic algorithms and support vector machines for bankruptcy prediction," *Expert Syst. Appl.*, vol. 31, no. 3, pp. 652–660, Oct. 2006.

[34] A. Janecek, W. N. Gansterer, M. A. Demel, and G. F. Ecker, "on the relationship between feature selection and classification accuracy," in *Proc. Int. Conf. New Challenges Feature Selection Data Mining Knowl. Discovery*, vol. 4. 2008, pp. 90–105.

[35] S. Ma, X. Song, and J. Huang, "Supervised group Lasso with applications to microarray data analysis," *BMC Bioinf.*, vol. 8, no. 1, p. 60, Feb. 2007.

[36] C. Ambroise and G. J. McLachlan, "Selection bias in gene extraction on the basis of microarray gene-expression data," in *Proc. Nat. Acad. Sci.*, vol. 99, no. 10, pp. 6562–6566, 2002.

[37] P. Jafari and F. Azuaje, "An assessment of recently published gene expression data analyses: Reporting experimental design and statistical factors," *BMC Med. Inform. Decision Making*, vol. 6, p. 27, Jun. 2006.

[38] E. R. Hruschka, Jr., E. R. Hruschka, and N. F. F. Ebecken, "Feature selection by Bayesian networks," in *Advances in Artificial Intelligence*. Berlin, Germany: Springer, 2004, pp. 370–379.

[39] A. L. Blum and P. Langley, "Selection of relevant features and examples in machine learning," *Artif. Intell.*, vol. 97, pp. 245–271, Dec. 1997.

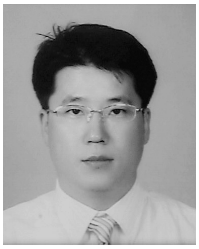
[40] M. Mitchell, *An Introduction to Genetic Algorithms*. Cambridge, MA, USA: MIT Press, 1998.

[41] V. Vapnik, *The Nature of Statistical Learning Theory*. Norwell, MA, USA: Springer, 2013.
[42] I. Steinwart and C. Scovel, "Mercer's theorem on general domains: On the interaction between measures, kernels, and RKHSs," *Construct. Approx.*, vol. 35, no. 3, pp. 363–417, Jun. 2012.
[43] S. Avidan, "Support vector tracking," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 8, pp. 1064–1072, Aug. 2004.
[44] R. D. Zimmerman, C. E. Murillo-Sánchez, and R. J. Thomas, "MATPOWER: Steady-state operations, planning, and analysis tools for power systems research and education," *IEEE Trans. Power Syst.*, vol. 26, no. 1, pp. 12–19, Feb. 2011.
[45] *Illinois Center for a Smarter Electric Grid (ICSEG)*. Accessed: Apr. 1, 2018. [Online]. Available: http://icseg.iti.illinois.edu/ieee-39-bus-system/

SAEED AHMED received the B.Sc. and M.Sc. degrees in electrical engineering from the University of AJ&K, Pakistan, in 2005 and 2010, respectively. He is currently pursuing the Ph.D. degree with the University of Ulsan, South Korea. He served as a Transmission Planning Engineer with Telecom Industry from 2005 to 2012, and has an experience in planning, surveying, deploying, and troubleshooting the microwave and optical fiber-based core and access PDH/SDH/SONET/DWDM networks. He joined the Mirpur University of Science and Technology, Mirpur, Pakistan, in 2012, as an Assistant Professor. His research area includes energy-efficient resource allocation in cognitive radios, smart grid (SG) communication technologies, SG security, and Internet of Things.

YOUNGDOO LEE received the B.E., M.E., and Ph.D. degrees from the School of Electrical Engineering, University of Ulsan, South Korea, in 2007, 2009, and 2013, respectively. Since 2013, he has been a Research Fellow with the University of Ulsan. His current research interests include artificial intelligent-based networks, cognitive radio networks, underwater sensor networks, beacon-based service networks, RFID, IoT-based service system, and next generation communication systems.

SEUNG-HO HYUN received the B.E., M.S., and Ph.D. degrees in electrical engineering from Seoul National University, South Korea, in 1991, 1993, and 1996, respectively. He was with the Korea Railroad Research Institute and Myongji University. He has been an Associate Professor with the University of Ulsan, South Korea, since 2004. His major research field is power system control, protection, and renewable energy.



INSOO KOO received the B.E. degree from Konkuk University, Seoul, South Korea, in 1996, and the M.S. and Ph.D. degrees from the Gwangju Institute of Science and Technology, Gwangju, South Korea, in 1998 and 2002, respectively. From 2002 to 2004, he was with the Ultrafast Fiber-Optic Networks Research Center, GIST, as a Research Professor. In 2003, he was a Visiting Scholar with the Royal Institute of Science and Technology, Sweden. In 2005, he joined the University of Ulsan, where he is currently a Full Professor. His research interests include next-generation wireless communication systems and wireless sensor networks.