energies

MDPI

Article

Mitigating the Impacts of Covert Cyber Attacks in Smart Grids Via Reconstruction of Measurement Data Utilizing Deep Denoising Autoencoders

Saeed Ahmed, YoungDoo Lee, Seung-Ho Hyun and Insoo Koo *

School of Electrical Engineering, University of Ulsan, Ulsan 44610, Korea
* Correspondence: iskoo@ulsan.ac.kr; Tel.: +82-52-259-1249

Received: 5 July 2019; Accepted: 7 August 2019; Published: 11 August 2019

Abstract: As one of the most diversified cyber-physical systems, the smart grid has become more decumbent to cyber vulnerabilities. An intelligently crafted, covert, data-integrity assault can insert biased values into the measurements collected by a sensor network, to elude the bad data detector in the state estimator, resulting in fallacious control decisions. Thus, such an attack can compromise the secure and reliable operations of smart grids, leading to power network disruptions, economic loss, or a combination of both. To this end, in this paper, we propose a novel idea for the reconstruction of sensor-collected measurement data from power networks, by removing the impacts of the covert data-integrity attack. The proposed reconstruction scheme is based on a latterly developed, unsupervised learning algorithm called a denoising autoencoder, which learns about the robust nonlinear representations from the data to root out the bias added into the sensor measurements by a smart attacker. For a robust, multivariate reconstruction of the attacked measurements from multiple sensors, the denoising autoencoder is used. The proposed scheme was evaluated utilizing standard IEEE 14-bus, 39-bus, 57-bus, and 118-bus systems. Simulation results confirm that the proposed scheme can handle labeled and non-labeled historical measurement data and results in a reasonably good reconstruction of the measurements affected by attacks.

Keywords: autoencoder; cyber-security; cyber-assaults; deep learning; self-healing smart grids; state estimation

1. Introduction

Integration of state-of-the-art computing and bi-directional communications technologies with the existing power infrastructure realizes the concept of the smart grid (SG) [1,2]. However, increased dependence on communications technologies is intensifying the SG's vulnerability to cyber-attacks. Typically, a supervisory control and data acquisition (SCADA) system is employed to periodically collect data from electric power grids. The SCADA system consists of communications networks and remote terminal units (RTUs) that include sensors and actuators. At the power control center (PCC), the collected data are applied to initiate command and control decisions by the energy management system (EMS). The fitness and health of the collected data are inordinately significant when making precise and correct control decisions. Therefore, conventionally, the consistency of the sensor-collected measurement data is checked by a data detector (BDD) before being utilized in the EMS. However, a recently discovered covert cyber-deception attack (CCDA) [3] is considered competent at deceiving the conventional BDD. Smartly designing the attack vector, a malicious user can compromise the integrity of the SG by injecting biased values into the sensor measurement data to dodge the BDD with a false, yet feasible system state [3]. Thus, initiating the CCDA through biased data may end in financial loss, partial disruption in power system operations, or a compound of economic loss and

disruptions [4,5]. Owing to the harmful impacts of such attacks on the reliable and secure operations of SGs, there is a need to investigate counter attack measures.

Generally, the defense measures reported in the literature can be organized into three layers: protection, detection, and mitigation [6].

Extensive investigations have been reported in the literature on the detection layer of the defense mechanism. Table 1 shows a summary of the research works carried out on the detection and mitigation tiers. In Table 1, it can be seen that less attention has been paid by researchers to the mitigation tier. Particularly, in the context of CCDA attack mitigation through the reconstruction of sensor collected measurement data, there is no existing work to the best of our knowledge. In the context of self-healing [7], which is a significant characteristic of an SG, there is a need to focus on mitigation layer and neutralize or minimize the impacts of a CCDA.

Table 1. A summary of research works on the detection and mitigation tier of defense mechanism in SGs.

Defense Tier	Application Area
Detection Tier (Intrusion detection system (IDS) without utilizing machine learning)	
Model-based techniques and game-theoretic methods for security [8,9]	EMS
Physical watermarking of control inputs [10]	PCC
Integration of run-time semantic analysis with efficient look-ahead PF analysis [11]	SCADA
Model-based IDS system to tackle attacks on auto generation control [12]	EMS
Integration of host-, and network-based IDS substations [13]	SCADA
Generic use of phasor measurement units (PMU) data considering white-list behavior and network topology [14]	PMU
IDS to detect PMU data assaulted by the GPS spoofing [15]	PMU
Detection of CP assaults on advanced metering network (AMI) systems based on behavior rules [16]	AMI
Distributed multi-layered IDS and early warning system [17,18]	AMI
Stealth attack identification with cumulative sum and quickest detection [19]	EMS and PCC
Abnormal energy consumption detection in smart meters using heuristic methods and contextual analysis of data [20–22]	Smart meters
Detection Tier (Intrusion detection system (IDS) utilizing machine learning)	
Utilizing machine learning methods for CCDA detection [23,24]	PCC/ EMS
Feature selection and supervised learning-based CCDA detection [25,26]	PCC
feature extraction and unsupervised learning-based CCDA detection [6]	PCC
Identification of CCDA utilizing joint transformation and Kullback–Leibler distance [27]	PCC
Deep learning-based recognition of the behavior features for CCDA [28]	PCC
Mitigation Tier (Restoration mechanism to eliminate impacts of attack)	
Game theroy-based mitigation (attacker-defender, zero-sum) game [29,30] and zero-sum Markov game-based mitigation [31]	Substation
Physical watermarking of control inputs [10]	Substation
Integration of run-time semantic analysis with efficient look-ahead PF analysis [11]	Substation

Figure 1 explains the prevailing circumstances and conceivable mitigations for CCDAs. We can see in Figure 1 that the sensor-collected data samples are scrutinized by the intrusion detection system

(IDS). The data identified as normal by the IDS are considered reliable enough to employ in the EMS to make control decisions. On the other hand, the data sample detected by IDS as attacked cannot be applied to the decision process. Due to the lack of any convincing solution, the attacked data sample is currently dropped, and the next sample is awaited, as depicted by Scheme-I in Figure 1. The same procedure is repeated, delaying control decisions until attack-free samples are received. However, due to the introduction of distributed generation and dynamically varying load conditions, taking immediate control decisions is desired for reliable and economic operations of power grids. Therefore, removing the impacts of a CCDA from corrupted measurement data is a more valuable solution than discarding samples. In this context, there is a need to establish a mechanism to induce a self-healing capability [7,32,33] in the PCC, which removes bias added by an intruder. As shown by Scheme-II in Figure 1, one possible option is to reconstruct the corrupted sample after its detection, which involves both detection and reconstruction time. An alternative is to reconstruct the data sample directly, omitting the detection stage, as shown by Scheme-III in Figure 1. Thus, the time consumed in detection can be saved; however, all measurement samples are reconstructed in this scheme, whether they were attacked or not. Table 2 shows a comparison of the above-mentioned mitigation schemes.

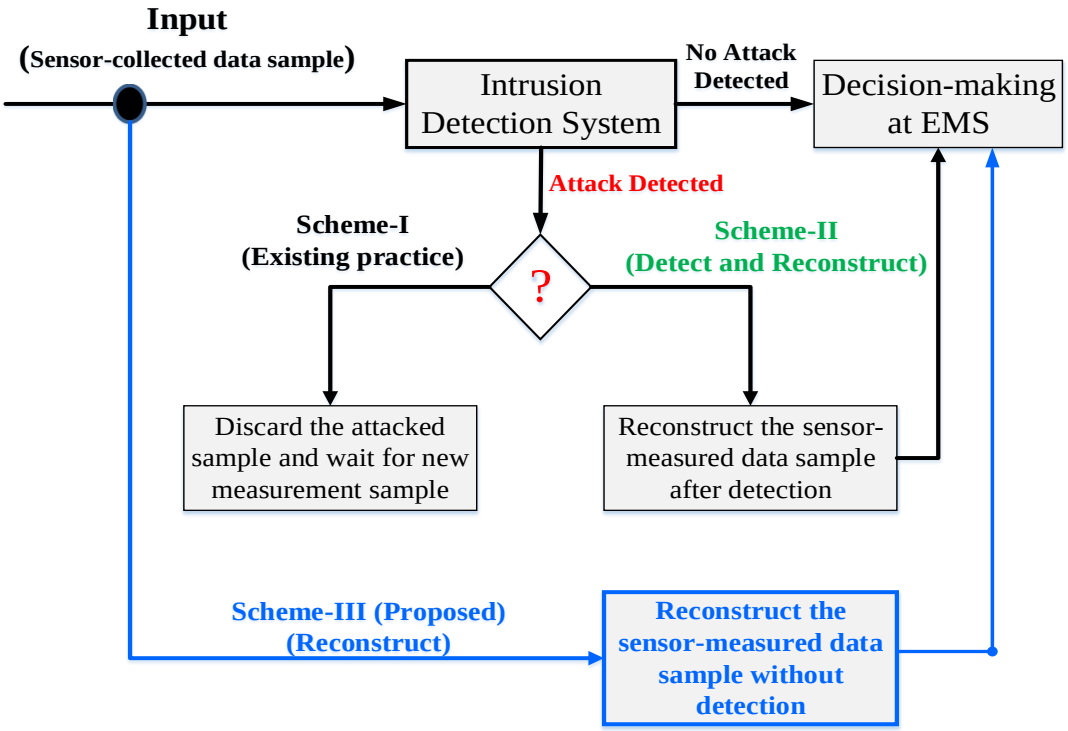

Figure 1. SG data reconstruction options against CCDA attacks.

Table 2. Possible options for mitigating attacks.

Detect and Discard (Scheme-I)	Detect and Reconstruct (Scheme-II)	Reconstruct (Scheme-III)
- Attacked samples are dropped	- Attacked samples are reconstructed	- Reconstruct without detection
- Wait for normal samples to be received	- Requires detection and reconstruction time	- Requires reconstruction time

To extend our previous work from detection [6,25,26] to mitigation, in this paper, we propose a deep neural network (DNN)-based data reconstruction scheme (Scheme-III) to mitigate the impacts of a CCDA on the SG's measurement dataset.

1.1. Motivation

Due to the vast geographical spans of power transmission networks, many sensors are deployed to collect the state data. Machine learning (ML)-based approaches may directly utilize the sensor measurements to mitigate the effects of CCDAs without requiring precise mathematical modeling. Additionally, ML-based mitigation through the reconstruction of the data does not need antecedent information of the power network. Furthermore, the state estimation (SE) measurement features (MF) data from power transmission networks via multiple sensors are multivariate and extremely correlated due to synergy and interaction between interconnected buses. Legacy multivariate-procedure monitoring techniques, such as principal component analysis (PCA) and an autoencoder (AE) with linear activation, assume linear process behavior that may not be suitable for SE-MF data. Moreover, PCA and AE are more sensitive to corruption in the data, and they have not adapted to learning robust representation from data corrupted due to attacks. In this paper, we present an SE-MF data reconstruction scheme based on a recently developed denoising autoencoder (DAE) to address the aforesaid challenges. Recent studies [34–36] have shown that a DAE can reconstruct the original signal by learning more robust representations from the attacked data. A comparison between the DAE and PCA is presented in Table 3. Inspired by its efficient application and magnificent characteristics for dealing with corrupted multivariate SE-MF data, we utilize the DAE for the robust reconstruction of attacked measurement samples while learning the nonlinear correlations embedded in an SE-MF dataset.

Table 3. Advantages of the denoising autoencoder over principal component analysis.

Denoising Autoencoder	Principal Component Analysis
The DAE can learn nonlinear and linear correlations in the multivariate sensor-collected data from the power grid [34,35,37,38]	PCA requires linear and Gaussian assumption about the data [34].
The DAE does not need dictionary elements to be orthogonal, making it adaptable to fluctuations in the representation of data. Thus signal reconstruction capability is improved [34,36,39].	PCA reduces the data frame by orthogonally transforming the data into a set of principal components. This limits the performance of PCA in reconstruction of data [34].
By restricting removed variables to be rebuilt from the remaining data, the DAE learns to convolute variables that tend to be correlated. This enhances robustness against noise and local fluctuations in the primary multivariate measurement data [36,39].	Reconstruction of noisy or corrupted nonlinear data is much too lossy as compared to PCA [34,35].

1.2. Contributions

In this paper, we consider multivariate SE-MF data affected by a CCDA. To reconstruct data by removing bias added by an attacker, we employ a state-of-the-art anomaly reconstruction method: the denoising autoencoder. The major technical contributions of this paper are summarized as follows.

- We investigate the impacts of smartly crafted CCDA on SG measurements, and study how such an attack can sidestep a BDD in typical power systems.
- We introduce the DAE algorithm to capture in a more robust way nonlinear correlations in multivariate SE-MF data corrupted by CCDA attacks while setting out robust signal reconstruction. To the best of our knowledge, this is the first paper to employ a DAE for reconstruction of corrupted sensor measurements in SE-MF data from SGs.
- To train the DAE model, commonly used choices for the addition of corruption are the zero-masking DAE (ZDAE) and the additive Gaussian DAE (GDAE) [27]. In addition to these schemes, we have introduced another corruption-addition scheme termed estimated DAE (EDAE). A comparison of these schemes shows that the newly introduced EDAE trains the denoising autoencoder model to obtain robust and powerful representations from the raw attacked data and results in a low reconstruction error.

- We employ IEEE standard 14-bus, 39-bus, 57-bus, and 118-bus test systems to gauge the performance of the proposed approach. Performance evaluations show that the proposed EDAE scheme results in reasonably good reconstruction with little loss of accuracy.

1.3. Paper Organization

The rest of this paper is organized as follows. In Section 2, state estimation, conventional bad data detection, and the nature of a CCD attack on an SG network are presented. In Section 3, we explain the fundamentals of the DAE, followed in Section 4 by an explanation of the proposed scheme to reconstruct the sensor measurements corrupted by a CCDA. Simulation results are presented in Section 5. We conclude the paper in Section 6.

2. System Model

2.1. Electric Power Network

The power transmission system connects various electrical generators across a vast geographic region with a host of customers. Multiple routes and lines contribute to securing the routing of the power from any generating source to any consumer, considering the economy of the transmission route. For effective monitoring and control of the power infrastructure, a communications network linking the power system components to the PCC is employed.

2.2. State Estimation

For the purpose of efficient monitoring, bi-directional RTUs, consisting of sensors and actuators, are employed in the power network. The readings of the sensors are collected at the PCC, which estimates the states (bus voltage angles and magnitudes) of the power system variables by utilizing the sensor measurements. The problem is to estimate the state variables, $\theta=[x_1, x_2, ..., x_n]^{\mathrm{T}}$, considering the sensor-collected measurements, $Z=[m_1, m_2, ..., m_m]^{\mathrm{T}}$, of the power system, where n and m are positive integers, and $x_i, m_j \in \mathbb{R}$ for $i = 1, 2, ..., n$ and $j = 1, 2, ..., m$. More specifically, the state variables are connected with the measurements in a nonlinear or alternating current (AC) model as follows:

$$Z = h(\theta)+e \tag{1}$$

where h is a nonlinear relationship between measurement vector Z and state vector, and $e = [e_1, e_2, ..., e_m]^T$ is a Gaussian measurement noise vector with standard deviation σ. On the assumption that the voltage magnitude at each bus remains close to its rated value, the model in Equation (1) can be described utilizing direct current (DC) model as follows:

$$Z = H\theta+e \tag{2}$$

where H is the Jacobian matrix in DC power flow problems and is approximated as follows [40,41]:

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

H is composed of topology and impedance data only. To find the estimate, $\hat{\theta}$, of the θ that is the best fit of the measurements, three statistical criteria are utilized in state estimation: maximum likelihood, minimum variance and weighted least squares (WLS) [42]. On the assumption that sensor error is normally distributed with a zero mean, the above-mentioned criteria result in an identical voltage phase estimation, as follows:

$$\hat{\mathrm{x}} = (H^T \Omega H)^{-1} H^T \Omega Z = GZ, \tag{4}$$

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

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

2.3. Conventional Bad Data Detection

The sensor-collected measurement data may become corrupted for many reasons, such as sensor faults, communication medium noise, and cyber attacks. Sensor-collected measurements result in an estimate of the state variables that is close to their true values under normal conditions, whereas a CCDA attack may result in shifted state variables, introducing a contrariety between the normal and attacked measurements. Typical power systems employ a residual-based detector to identify corruptions in the sensor measurements [3]. The residual, R, is the difference between sensor collected measurements, Z, and the estimated measurements, $\hat{Z}$, at the PCC, and is described as follows:

$$R = Z - \hat{Z} = Z - H\hat{\theta}. \tag{6}$$

Then, $L2 - norm \left\| Z - H\hat{\theta} \right\|$ is compared with a deliberately selected threshold, τ [11], to detect the presence of bad measurements. Therefore, the hypothesis of not being attacked is accepted if we have

$$\max_i |R_i| < \tau, \tag{7}$$

where R_i is the element of residual vector R. Otherwise, an alarm indicating existence of bad measurements is raised.

2.4. Covert Cyber Deception Attack: Basic Principle

From a complete (or even partial) familiarity with the power network topology, a smart attacker can add biased data to meter or sensor-collected measurements Z by forming an attack vector, $a = [a_1, a_2, ..., a_m]^T$, to deceive the bad data detector [27]. Let $Z_a = Z + a$ be the measurements containing the attacked data. In attack vector a, the attacker enjoys the liberty of selecting any non-zero arbitrary element. Thus, the ith non-zero element, a_i, of attack vector a, allows that the attacker to alter the ith sensor measurement, Z_i, with a forged measurement: $Z_i + a_i$.

As discussed above, the conventional bad data detector computes the L_2-norm of measurement residual R, to determine the presence of attacked or bad measurements. However, if the attacker designs attack vector a, such that $a = Hc$, where c is a non-zero vector of length n, the measurement vector containing the attacks (Z_a) can circumvent the traditional detection as long as the measurement vector containing the normal measurements can pass.

Let $\hat{\theta}_a$ denote an estimate of state variables using attacked sensor measurements Z_a such that we have

$$\hat{\theta}_a = GZ + Ga = \hat{\theta} + GHc = \hat{\theta} + c. \tag{8}$$

Now, the L_2 norm of attacked measurements residual R_a is

$$\begin{aligned} &\|R_a\|_2 = \left\|Z_a - H\hat{\theta}_a\right\|_2 \\ &= \left\|(Z + a) - H(\hat{\theta}_a + c)\right\| \\ &= \left\|\left(Z - H\hat{\theta}\right) + (a - Hc)\right\|_2 = \left\|\left(Z - H\hat{\theta}\right)\right\|_2 \\ &= \|R\|_2 < \tau. \end{aligned} \tag{9}$$

2.5. Covert Cyber Deception Attack Model

Broadly speaking, there are two kinds of CCD attack: (1) the load redistribution attack; and (2) the load change attack [43–45]. With the ultimate objective being to dodge the conventional BDD and pass the operator at the PCC, the attacker may craft the attack aiming at altering one or more measurements. In this paper, our main objective is real-time reconstruction of corrupted sensor measurements. Therefore, our approach is to formulate the most generalized attack, and thereby to come up with robust reconstruction of the attacked measurements. For the construction of the attack, we assumed that the assailant has enough knowledge about the power network topology.

During the CCD attack, the malicious user embeds a forged value in the sensor measurements, altering the real power injection and real power flows to project the desired changed state variables for the system operator. For instance, to change state variable x_2 by adding a corruption of -12%, a $(1 \times (n-1))$ attack vector c can be formulated by considering the following equation:

$$c = [-0.12x_2, 0, ..., 0]. \tag{10}$$

Utilizing the power flow equations and state vector $x_{\mathrm{a}} = \hat{\mathrm{x}} + \mathrm{c}$, the corrupted measurements are calculated as follows:

$$Z_a = H\theta_a + e. \tag{11}$$

3. Reference Model Learning

In this section, we first briefly describe the AE and denoising autoencoder (DAE) algorithm, and then, we explain the proposed data reconstruction scheme.

3.1. Autoencoder: Basic Principle

Autoencoders (AEs) are a specific type of fully connected feed-forward neural networks where the inputs are equal to the output, and therefore, the AE is trained in an unsupervised way without any label information. Fundamentally, an AE consists of three elements: an encoder, the code, and a decoder. The encoder compresses the input into a lower-dimensional code or latent-space representation, and then, the decoder reconstructs the output from this representation. Analogous to PCA, the AE aims to encode the input data into an intermediate representation that preserves most of the information in the input data to allow reconstructing it. In this paper, to capture the hidden nonlinear correlations more robustly, and to tackle a CCDA in complex multivariate data, we employ a recently developed algorithm in the field of deep learning, the DAE [39]. The DAE is an expansion of the AE and has multiple advantages over the conventional PCA-based dimension-reduction method, as explained in Table 2.

3.2. Denoising Autoencoder: Basic Principle

The fundamental concept of the DAE is to reconstruct the primary input from corrupted or attacked input [39]. Thus, it can stop an AE from just learning identity mapping between the input and the reconstructed output, can apprehend more informational latent-space patterns, and can gather a strong and robust representation from raw, attacked data. Two primary choices for addition of corruption are additive Gaussian noise (the GDAE) and zero-masking noise (the ZDAE) [34].

Similar to AE, a DAE is composed of three parts: the encoder, code or latent space, and the decoder. Given an input, x, the encoder typically transforms its corrupted or attacked input data, $\tilde{x}$, instead of original input data x, into a hidden or latent-space representation, h, employing nonlinear mapping as follows:

$$h = f\left(W_1 x + b\right), \tag{12}$$

where $f(.)$ is a nonlinear activation function, such as the sigmoid function. $W_1 \in IR^{m \times n}$ is the weight matrix, and $b \in IR^m$ is optimized in the encoding with m nodes in latent space. Then,

the decoder unfolds the latent space into a reconstructed vector, $\hat{x}$, at the output layer, utilizing nonlinear transformation as follows:

$$\hat{x} = g\left(W_1 x + c\right), \tag{13}$$

where $g(.)$ is a nonlinear activation function, such as the sigmoid function. For better learning efficiency, we employed the tied weights as $W_1 = W_2^T$ [39]. The reconstruction error can be computed for a given input training set, $\{x_i\}_{i=1}^m$, as follows: $\sum_{i=1}^m \|x_i - \hat{x}_i\|^2$. The objective training of the DAE is to find optimal parameters, $\psi = \{W_1, b, c\}$ that minimize the reconstruction error, as given below:

$$\min_{\psi} \sum_{i=1}^{m} \|x_i - \hat{x}_i\|^2. \tag{14}$$

It is obvious from Equation (13) that the reconstruction error is the difference between the reconstructed output and the actual meter measurements instead of the attacked measurements. In other words, the DAE is trained to produce output closer to the original input x, even when employing the attacked input, $\tilde{x}$.

4. Proposed CCDA Mitigation–Data Reconstruction Scheme

Recovering the original sensor measurements from the attacked signals is required for a self-healing SG. In this section, we propose a scheme utilizing a DAE for the reconstruction of measurements corrupted due to a CCDA in an SG communications network. Figure 2 illustrates the reference model's learning process. The bulk power-generated at the power producing plants is transported to consumers via the power transmission and distribution networks. The RTUs collect measurement data X_0 from the power network and transmit the collected data over wireless media. A smart attacker can compromise the integrity of the data by adding biased values to the collected data. At the power control center, the proposed DAE-based reconstruction scheme attempts to reconstruct the attacked data, $\tilde{X}_0$, by removing any bias added by the attacker. A healthy DAE model reconstructs the data well enough to be employed in the EMS for initiating command and control decisions. The proposed reconstruction scheme is explained in the following subsection. Figure 3 shows the flowchart of the proposed scheme.

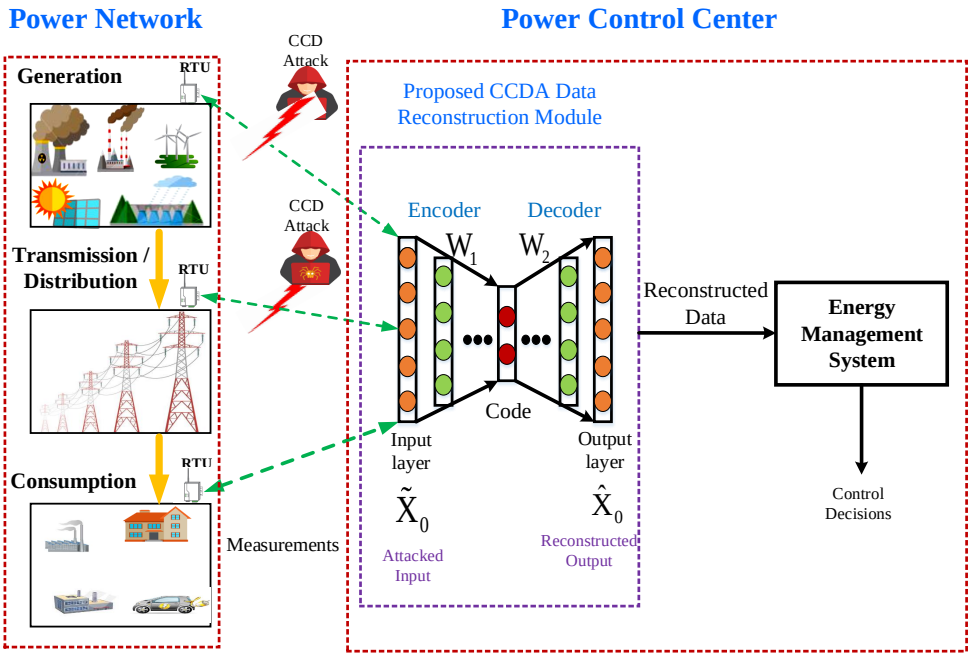

Figure 2. SG Data reconstruction options against CCDA attacks.

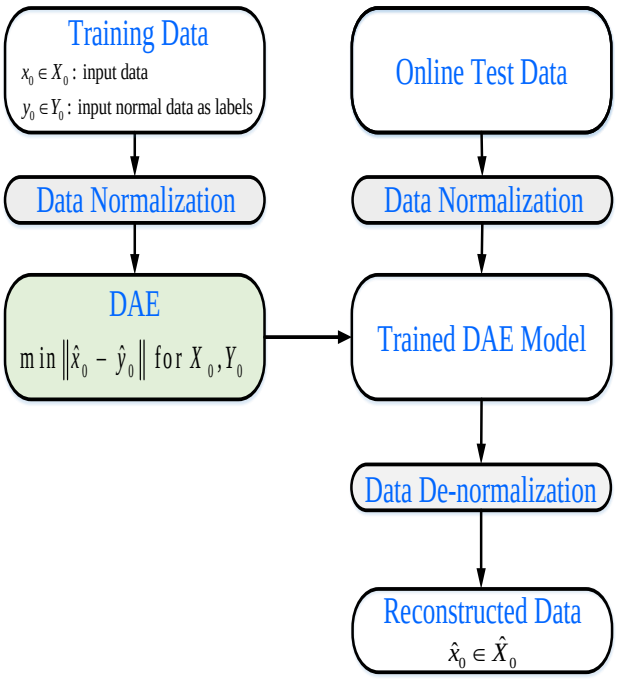

Figure 3. A flowchart of the proposed scheme for reconstruction of data affected by CCDA.

4.1. Proposed Corruption Addition Scheme: EDAE

As mentioned above, there are two common choices of adding corruption to the DAE model for training: the ZDAE and GDAE [34]. In the ZDAE scheme, some features of each sample are set to zero randomly with probability v. Practically speaking, zero-masking can be viewed as the nonexistence of sensor measurements (due to attack or noise). In addition to the above-mentioned choices, in this paper, we introduce another corruption addition scheme EDAE, in which the noise is generated based on Gaussian distribution with mean and variance obtained from analysis of the SE-MF data. The corruption addition procedure following the EDAE is explained as follows.

To train the DAE model, we insert noise or corruption to the SE-MF training data $X = \{x_1, x_2, ..., x_m\}$, where m is the number of measurement samples, and $x_i \in X$, where $i \in \{1, 2, ..., m\}$. The x_i is a sample consisting of n features, $x_i = \{f_{i1}, f_{i2}, ..., f_{in}\}$. The corruption or noise data δ follows a normal distribution $\mathcal{N}(\mu, \nu)$, where μ is a vector of mean values $\mu = \{\mu_1, \mu_2, ..., \mu_n\}$, which is calculated as $\mu_j = \frac{1}{m}\sum_{i=1}^{m} f_{ij}$. The ν is a variance vector $\nu = \{\nu_1, \nu_2, ..., \nu_n\}$ where $\nu_j = \frac{1}{m}\sum_{i=1}^{m} (\mu_j - f_{ij})^2$. Finally, the training data are obtained as $X_0 = X + \delta$, and used for training the DAE model.

4.2. DAE Model Training

Let $x_0 \in X_0$ be the historical SE-MF input data, where $X_0 \in I\!R^{m \times n}$. To train the DAE model to minimize any loss in the reconstruction process, we insert normal (attack-free) data, $y_0 \in Y_0$, as labels, where $Y_0 \in I\!R^{m \times n}$, in which m is the number of measurement samples, and n is the number of measurement features in a sample.

Before employing the input data to train the learning model (i.e., the DAE), data normalization is required as a pre-processing step; otherwise, the data are not reconstructed well enough. Data normalization means converting all variables in the data within a particular range. Normalization is essential for steady convergence of weights and biases. There are several types of normalization, such as min-max normalization, decimal scaling, and the standard deviation method [46,47]. Choosing a suitable normalization method depends on the application and the algorithm in which the normalized

data will be used. The min-max normalization approach is a simple normalization technique and is usually more efficient. To linearly scale each feature to the range $0 \sim 1$, in this paper, we employ a min-max normalization function as follows [46]:

$$\tilde{f}_i = \frac{(f_i - f_{\min})}{(f_{\max} - f_{\min})}, \tag{15}$$

where $\tilde{f}_i$ is the normalized, or scaled, value of feature i, and $f_{\min}$ and $f_{\max}$ are the minimum and maximum values, respectively, of feature i in the dataset. Normalization of training and testing data according to the same scale helps in a good reconstruction. After normalization, the data are inserted to train the DAE model. Equations (1)–(3) are applied to the input data to find the optimal parameters defined by ψ, and to acquire the hidden representation h and the rebuilt output by minimizing the reconstruction error as follows: $\min \|\hat{x}_0 - y_0\| \ \ for \ X_0, Y_0$.

4.3. DAE Model Testing

During the testing phase, online data are inserted to the trained DAE model. The test samples are also rescaled according to the minimum and maximum values of the training data, ensuring both datasets are in a similar range. A trained DAE model attempts to reconstruct the test data identical to the normal data. Each feature in the reconstructed data is denormalized using the inverse transformation, as follows:

$$f_i = \tilde{f}_i \, (f_{\max} - f_{\min}) + f_{\min}. \tag{16}$$

A well-trained, healthy DAE will generate the reconstructed data as close to the original input as possible and the reconstructed data can be employed in the EMS with high enough level of confidence to initiate control decisions.

5. Experimental Results

In this section, we gauge the performance of the proposed reconstruction scheme for CCDA-corrupted data.

5.1. Power System Data and Attack Data Generation

We utilized various power system test cases, from standard IEEE 14-, 39-, 57-, and 118-bus systems, to endorse the performance of the proposed mitigation scheme. To set up the configuration of these standard IEEE test systems, and explicitly the Jacobian matrix, we applied the Matpower 6.0 toolbox [48]. To generate the measurements, Z, operating points of the test systems provided in Matpower case files were employed. We used the DC power flow analysis to approximate the state vectors employed in the AC power flow model. The state variable vector θ, for a B-bus system, consists of $(B-1)$ bus voltage phase angles, and measurement vector Z comprises of active power flows in the lines and active power injections into the buses. To carry out a fair comparison with a real-world power system, we adopted stochastic loads with uniform load distributions identical to those in [24], i.e., ranging from $0.9 \times Q_0$ to $1.1 \times Q_0$, where Q_0 is the base load. The features employed in these simulations were the active power flows in the branches and active power injections into the buses.

As mentioned above, the CCDA can be modeled to deceive the BDD with the ultimate objective being to falsify single or multiple system states. We assumed that the attacker has complete knowledge of the topology of the power network, and the attack is initiated following the model described in [6]. The attack formulation is explained in Section 2. In the simulations, we considered various attack scenarios, described as follows.

Scenario 1: The attacker is stationary and has access to specific RTUs or meters only. Thus, the attacker can initiate a fixed attack, i.e., fixed or the same features in the measurement samples are corrupted with the attack.

Scenario 2: The attacker is moving and can randomly access different RTUs or meters. Initiating such an attack, the attacker randomly adds biased values into the features or meter measurements. This sort of attack is stronger than a fixed attack, and arduous for the recovery of the original data through reconstruction. For the above-mentioned fixed- and random-attack scenarios, the measurement data attacks were on 40% of the features and 20% of the features.

5.2. Parameter Tuning for the DAE Structure

We employed multiple power system test cases to validate the performance of the proposed SG data-reconstruction scheme. Power variable states and measurement features for standard IEEE systems increase as the size of the power system increases. Therefore, for each power system test case, the DAE structure has a different number of input nodes. The DAE reconstructs the output so it is identical to original input; therefore, the number of input and output nodes in a DAE structure is the same. We used 50% of the historical SE-MF data for training the DAE model and 50% for testing it. For each standard IEEE test bus system employed in this work, tuning parameters were chosen based on optimal reconstructed data. The simulation parameters for the DAE structure in various test bus cases are shown in Table 5. If familiar with the topology of the power system network, an attacker can adopt various strategies to dilute SG measurements. Keeping in mind an expanded attack choice for of the attacker, we employed different corruption addition schemes: the GDAE, ZDAE, and EDAE. In the following subsection, we present the simulation results for various standard IEEE test bus systems.

5.3. Simulation Results

The accuracy of the reconstructed measurements is significantly important for making correct decisions at the PCC. Therefore, the reconstruction error and error average sum (EAS) were considered performance measuring metrics for the proposed data reconstruction schemes. We simulated the system for standard IEEE 14-, 39-, 57-, and 118-bus systems, as mentioned above. However, we present simulation results only from the standard IEEE 14- and 39-bus systems due to space limitations.

5.3.1. Reconstruction Errors for Fixed- and Random-Attack Dataset

As mentioned in Tables 4 and 5, the IEEE 14-bus system has 53 measurement features, and it is challenging to present reconstruction of all the features in the dataset due to space limitations. Therefore, to show the performance of the proposed cyber-attack-mitigation scheme, the seven best and the seven worst reconstructed features are presented for the proposed EDAE and the existing ZDAE and GDAE schemes. The fixed and random attacks targeted either 20% or 40% features in the dataset. Furthermore, fixed- and random-attack scenarios were also considered in the simulations. The seven best-reconstructed features for the measurement data, in which a fixed attack was initiated on 20% of the features, are shown in Table 6 for the EDAE model. Actual values of the features and the reconstructed values (along with the reconstruction error) are shown. The error ratio is also presented in the table. A reduced error ratio is essential for attack mitigation in the SG measurement data. We see that features have been reconstructed well, because the reconstruction error and the error ratio are very small. The seven worst-reconstructed features are depicted in Table 7 for the EDAE scheme when 20% of the features were targeted in a fixed attack. We see that the reconstruction error and the error ratio are low for the reconstructed features, except for features with values closer to zero. Tables 8 and 9, respectively, show the seven best- and worst-reconstructed features with the EDAE scheme for the data in which 40% of the features were subjected to a fixed attack. Similarly, Tables 10–13 show the reconstruction of the data affected from a fixed attack on 20% of the features and 40% of the features, for the ZDAE schemes. In addition, Tables 14–17 show GDAE scheme's reconstructed data that were subjected to fixed attack. Similarly, from random attacks, the seven best-, and worst-reconstructed features are shown in Tables 18–29 for the earlier-mentioned DAE schemes. We see that features have been reconstructed well, because the reconstruction error and the error ratio are small for the EDAE scheme, compared to the other schemes.

Table 4. Dimension growth with increasing sizes of power systems.

System	States	DAE Nodes (Input and Output)
IEEE 14-bus	13	53
IEEE 39-bus	38	130
IEEE 57-bus	56	216
IEEE 118-bus	117	489

Table 5. Simulation parameters for DAE model.

System	IEEE 14-Bus	IEEE 39-Bus	IEEE 57-Bus	IEEE 118-Bus
Input and Output Nodes	53	130	13	53
DAE Layers	3	2	2	2
Code Layer size	34	75	140	318
Total Data	200,000	200,000	200,000	200,000
Training Data	100,000	100,000	100,000	100,000
Test Data	100,000	100,000	100,000	100,000
Optimizer	MSE	MSE	MSE	MSE
Normalization	Min-Max	Min-Max	Min-Max	Min-Max
Batch Size	1	1	1	1
Epochs	30	30	30	30

Discussion: In Tables 6–29, it is observed that all features have been reconstructed well with the EDAE scheme, compared to other schemes. However, the error ratio is high for the features with values closer to zero. These results were obtained by using the mean squared error (MSE) objective function. It would be interesting future work to reduce the error ratio by investigating more objective functions.

Table 6. Seven best-reconstructed features (20% fixed-attack on 20% of the features in a standard IEEE 14-bus system) with the EDAE scheme.

Feature Number	19	47	38	46	49	24	18
Feature Value	−24.31	−0.10896	−42	−17.574	−6.8649	8.09	50.4
Reconstructed Value	-2.43×10^{1}	-1.09×10^{-1}	-4.20×10^{1}	-1.76×10^{1}	-6.86×10^{0}	8.09×10^{0}	5.04×10^{1}
Error	5.88×10^{-6}	1.39×10^{-5}	5.18×10^{-5}	6.14×10^{-5}	6.47×10^{-5}	8.01×10^{-5}	0.00010462
Error Ratio	8.50×10^{-8}	3.84×10^{-6}	6.75×10^{-7}	7.19×10^{-6}	1.12×10^{-6}	4.24×10^{-6}	6.09×10^{-6}

Table 7. Seven worst-reconstructed features(fixed attack on 20% of the features in a standard IEEE 14-bus system) with the EDAE scheme.

Feature Number	33	8	1	9	13	44	50
Feature Value	5.2503	−35.862	16.015	−10.707	−15.029	−6.7417	−11.735
Reconstructed Value	5.25731468	−35.85805672	16.01861077	−10.70396887	−15.02630777	−6.739086843	−11.73276872
Error	5.25731468	−35.85805672	16.01861077	−10.70396887	−15.02630777	−6.739086843	−11.73276872
Error Ratio	0.001336053	0.000109957	0.000225462	0.000283098	0.000179135	0.000387611	0.000190139

Table 8. Seven best-reconstructed features (fixed attack on 40% of the features in a standard IEEE 14-bus system) with the EDAE scheme.

Feature Number	16	24	18	44	36	43	38
Feature Value	71.248	6.7785	50.363	−8.1342	−71.248	−52.386	−41.969
Reconstructed Value	7.12×10^{1}	6.78×10^{0}	5.04×10^{1}	-8.13×10^{0}	-7.12×10^{1}	-5.24×10^{1}	-4.20×10^{1}
Error	6.05×10^{-6}	2.60×10^{-5}	3.40×10^{-5}	5.84×10^{-5}	7.95×10^{-5}	0.000221883	0.000255396
Error Ratio	8.50×10^{-8}	3.84×10^{-6}	6.75×10^{-7}	7.19×10^{-6}	1.12×10^{-6}	4.24×10^{-6}	6.09×10^{-6}

Table 9. Seven worst-reconstructed features (fixed attack on 40% of the features in a standard IEEE 14-bus system) with the EDAE scheme.

Feature Number	4	28	7	13	2	33	5
Feature Value	−9.3102	28.447	−0.21257	−15.038	−114.81	5.2384	−11.576
Reconstructed Value	−9.305423719	28.45157074	−0.208140482	−15.03418581	−114.806619	5.241757108	−11.57308739
Error	0.004776281	0.004570744	0.004429518	0.003814195	0.003381042	0.003357108	0.002912608
Error Ratio	0.000513016	0.000160676	0.020837926	0.000253637	2.9449×10^{-5}	0.000640865	0.000251607

Table 10. Seven best-reconstructed features (fixed attack on 20% of the features in a standard IEEE 14-bus system) with the ZDAE scheme.

Feature Number	21	49	29	41	9	35	15
Feature Value	29.115	−5.7354	−5.7354	−29.115	−10.608	−73.443	73.443
Reconstructed Value	29.11551588	−5.73057252	−5.729508174	−29.10641015	−10.59362174	−73.42701639	73.46094698
Error	0.00051588	0.00482748	0.005891826	0.008589851	0.01437826	0.015983606	0.017946978
Error Ratio	1.77187×10^{-5}	0.000841699	0.001027274	0.000295032	0.001355417	0.000217633	0.000244366

Table 11. Seven worst-reconstructed features (fixed attack on 20% of the features in a standard IEEE 14-bus system) with the ZDAE scheme.

Feature Number	22	42	4	1	5	12	17
Feature Value	16.856	−20.227	−9.3102	16.01	−11.576	−13.834	56.302
Reconstructed Value	17.3320005	−19.75872138	−9.015861877	16.29326843	−11.40118695	−13.66738703	56.46786937
Error	0.476000501	0.468278621	0.294338123	0.283268428	0.174813047	0.166612974	0.165869371
Error Ratio	0.028239232	0.023151165	0.031614587	0.017693218	0.015101334	0.012043731	0.002946065

Table 12. Seven best-reconstructed features (fixed attack on 40% of the features in a standard IEEE 14-bus system) with the ZDAE scheme.

Feature Number	39	9	35	15	32	52	19
Feature Value	24.31	−10.707	−73.47	88.164	1.7611	−1.4676	−24.31
Reconstructed Value	2.43×10^{1}	-1.07×10^{1}	-7.35×10^{1}	8.82×10^{1}	1.77×10^{0}	-1.45×10^{0}	-2.43×10^{1}
Error	9.29×10^{-5}	0.003913956	0.009265381	0.009673824	0.010849409	0.013501805	0.021025564
Error Ratio	3.82×10^{-6}	3.66×10^{-4}	1.26×10^{-4}	1.10×10^{-4}	6.16×10^{-3}	9.20×10^{-3}	8.65×10^{-4}

Table 13. Seven worst-reconstructed features (fixed attack on 40% of the features in a standard IEEE 14-bus system) with the ZDAE scheme.

Feature Number	22	42	11	17	1	14	8
Feature Value	16.855	−20.226	−6.335	56.346	16.015	153.52	−35.862
Reconstructed Value	17.2521351	−19.8600899	−5.976216655	56.6757069	16.30432973	153.773394	−35.6177516
Error	0.397135099	0.365910095	0.358783345	0.329706898	0.289329731	0.253393963	0.244248405
Error Ratio	3.82×10^{-6}	3.66×10^{-4}	1.26×10^{-4}	1.10×10^{-4}	6.16×10^{-3}	9.20×10^{-3}	8.65×10^{-4}

Table 14. Seven best-reconstructed features (fixed attack on 20% of the features in a standard IEEE 14-bus system) with the GDAE scheme.

Feature Number	23	38	47	12	27	18	25
Feature Value	6.7785	24.427	−28.447	−15.038	28.447	−24.427	17.551
Reconstructed Value	6.78×10^{0}	2.44×10^{1}	-2.84×10^{1}	-1.50×10^{1}	2.84×10^{1}	-2.44×10^{1}	1.76×10^{1}
Error	2.40×10^{-5}	0.000135149	0.000241691	0.000330428	0.000685705	0.000773076	0.001103067
Error Ratio	3.55×10^{-6}	5.53×10^{-6}	8.50×10^{-6}	2.20×10^{-5}	2.41×10^{-5}	3.16×10^{-5}	6.28×10^{-5}

Table 15. Seven worst-reconstructed features (fixed attack on 20% of the features in a standard IEEE 14-bus system) with the GDAE scheme.

Feature Number	3	10	7	28	48	31	51
Feature Value	−9.3102	−6.2287	−29.767	5.7354	−5.7354	1.5209	−1.5209
Reconstructed Value	−9.271860897	−6.190712099	−29.73700456	5.75936093	−5.711731496	1.543764792	−1.498413224
Error	0.038339103	0.037987901	0.02999544	0.02396093	0.023668504	0.022864792	0.022486776
Error Ratio	0.004117968	0.006098849	0.001007674	0.004177726	0.00412674	0.015033725	0.014785177

Table 16. Seven best-reconstructed features (fixed attack on 40% of the features in a standard IEEE 14-bus system) with the GDAE scheme.

Feature Number	48	22	37	28	17	1	6
Feature Value	−5.7354	43.655	−41.969	5.7354	41.969	−95.675	−0.21257
Reconstructed Value	−5.734781558	43.6563242	−41.966989	5.738713765	41.97239975	−95.67142491	−0.207780056
Error	0.159876943	0.145020078	0.13172484	0.131127364	0.10495939	0.086672587	0.083504682
Error Ratio	−0.000107829	3.03333×10^{-5}	-4.79163×10^{-5}	0.000577774	8.10061×10^{-5}	-3.73671×10^{-5}	−0.022533488

Table 17. Seven worst-reconstructed features (fixed attack on 40% of the features in a standard IEEE 14-bus system) with the GDAE scheme.

Feature Number	4	1	13	33	3	7	11
Feature Value	−5.7354	43.655	−41.969	5.7354	41.969	−95.675	−0.21257
Reconstructed Value	−5.734781558	43.6563242	−41.966989	5.738713765	41.97239975	−95.67142491	−0.207780056
Error	0.000618442	0.001324201	0.002010998	0.003313765	0.003399746	0.003575093	0.004789944
Error Ratio	0.000107829	3.03333×10^{-5}	4.79163×10^{-5}	0.000577774	8.10061×10^{-5}	3.73671×10^{-5}	0.022533488

Table 18. Seven best-reconstructed features (random attack on 20% of the features in a standard IEEE 14-bus system) with the EDAE scheme.

Feature Number	25	41	24	50	21	52	15
Feature Value	1.76×10^{1}	-1.69×10^{1}	7.73×10^{0}	2.99×10^{0}	1.69×10^{1}	-5.27×10^{0}	7.36×10^{1}
Reconstructed Value	1.76×10^{1}	-1.69×10^{1}	7.73×10^{0}	2.99×10^{0}	1.69×10^{1}	-5.27×10^{0}	7.36×10^{1}
Error	2.42×10^{-5}	3.72×10^{-5}	0.000121965	0.000145602	0.000226226	0.000232532	0.000244578
Error Ratio	1.38×10^{-6}	2.20×10^{-6}	1.58×10^{-5}	4.88×10^{-5}	1.34×10^{-5}	4.41×10^{-5}	3.33×10^{-6}

Table 19. Seven worst-reconstructed features (random attack on 20% of the features in a standard IEEE 14-bus system) with the EDAE scheme.

Feature Number	49	8	1	4	38	18	45
Feature Value	−9.9661	−8.6655	−95.556	−11.794	−42.103	−24.313	−7.734
Reconstructed Value	−9.960711568	−8.660489164	−95.55141419	−11.78952318	−42.09859985	−24.3087889	−7.730274573
Error	0.005388432	0.005010836	0.00458581	0.004476824	0.004400152	0.004211097	0.003725427
Error Ratio	0.000540676	0.000578251	4.79908×10^{-5}	0.000379585	0.000104509	0.000173203	0.000481695

Table 20. Seven best-reconstructed features (random attack on 40% of the features in a standard IEEE 14-bus system) with the EDAE scheme.

Feature Number	20	35	29	32	43	40	14
Feature Value	29.111	−73.554	5.6798	5.2682	−6.7198	−29.111	73.554
Reconstructed Value	2.91×10^{1}	-7.36×10^{1}	5.68×10^{0}	5.27×10^{0}	-6.72×10^{0}	-2.91×10^{1}	7.36×10^{1}
Error	4.64×10^{-5}	5.41×10^{-5}	6.26×10^{-5}	0.000224246	0.000406535	0.000490851	0.000534421
Error Ratio	1.59×10^{-6}	7.36×10^{-7}	1.10×10^{-5}	4.26×10^{-5}	6.05×10^{-5}	1.69×10^{-5}	7.27×10^{-6}

Table 21. Seven worst-reconstructed features (random attack on 40% of the features in a standard IEEE 14-bus system) with the EDAE scheme.

Feature Number	7	11	47	4	49	9	1
Feature Value	0.046326	−13.885	−28.613	−11.794	−9.9661	−3.734	−95.556
Reconstructed Value	0.05920957	−13.87618859	−28.60460809	−11.78692803	−9.959870581	−3.728092846	−95.55035503
Error	0.01288357	0.008811412	0.008391909	0.007071968	0.006229419	0.005907154	0.005644974
Error Ratio	0.27810667	0.000634599	0.00029329	0.000599624	0.000625061	0.001581991	5.9075×10^{-5}

Table 22. Seven best-reconstructed features (random attack on 20% of the features in a standard IEEE 14-bus system) with the ZDAE scheme.

Feature Number	41	14	21	44	42	24	49
Feature Value	−16.888	73.554	16.888	−7.734	−43.836	7.734	−9.9661
Reconstructed Value	−16.88659536	73.55574957	16.88984495	−7.732058335	−43.83405151	7.736052894	−9.963812929
Error	0.001404636	0.001749567	0.001844953	0.001941665	0.00194849	0.002052894	0.002287071
Error Ratio	8.31736×10^{-5}	2.37862×10^{-5}	0.000109246	0.000251056	4.44496×10^{-5}	0.000265438	0.000229485

Table 23. Seven worst-reconstructed features (random attack on 20% of the features in a standard IEEE 14-bus system) with the ZDAE scheme.

Feature Number	5	27	47	7	2	11	40
Feature Value	−0.5438	−0.046326	−29.855	−48.433	−6.169	−29.111	29.111
Reconstructed Value	−0.402496385	0.062047902	−29.74962789	−48.35341728	−6.09788107	−29.04398602	29.17169563
Error	0.141303615	0.108373902	0.105372107	0.079582725	0.07111893	0.067013983	0.06069563
Error Ratio	0.259844824	2.339375345	0.003529463	0.001643151	0.011528437	0.002302016	0.002084972

Table 24. Seven best-reconstructed features (random attack on 40% of the features in a standard IEEE 14-bus system) with the ZDAE scheme.

Feature Number	42	22	18	14	38	44	24
Feature Value	−43.836	16.888	−24.313	73.554	24.313	−7.734	7.734
Reconstructed Value	−43.83578598	16.88835338	−24.31091173	73.5561661	24.31673587	−7.729402089	7.739212442
Error	0.000214025	0.000353378	0.002088271	0.002166096	0.003735869	0.004597911	0.005212442
Error Ratio	4.88239×10^{-6}	2.09248×10^{-5}	8.58911×10^{-5}	2.94491×10^{-5}	0.000153657	0.000594506	0.000673965

Table 25. Seven worst-reconstructed features (random attack on 40% of the features in a standard IEEE 14-bus system) with the ZDAE scheme.

Feature Number	12	1	9	7	47	33	52
Feature Value	−15.234	16.22	−3.734	−29.855	6.7198	5.2682	−1.565
Reconstructed Value	−15.14360235	16.30842128	−3.673237667	−29.80570794	6.764957918	5.312002445	−1.523176287
Error	0.090397648	0.088421284	0.060762333	0.049292057	0.045157918	0.043802445	0.041823713
Error Ratio	0.00593394	0.005451374	0.016272719	0.001651049	0.006720128	0.008314499	0.026724418

Table 26. Seven best-reconstructed features (random attack on 20% of the features in a standard IEEE 14-bus system) with the GDAE scheme.

Feature Number	12	41	21	20	24	30	4
Feature Value	−13.885	−29.111	16.888	29.111	6.7198	9.9661	−11.794
Reconstructed Value	−13.88377492	−29.10957017	16.88965969	29.11274824	6.721632953	9.967963504	−11.79191121
Error	0.001225077	0.001429832	0.001659687	0.001748239	0.001832953	0.001863504	0.002088788
Error Ratio	8.82302×10^{-5}	4.91165×10^{-5}	9.82761×10^{-5}	6.00542×10^{-5}	0.000272769	0.000186984	0.000177106

Table 27. Seven worst-reconstructed features (random attack on 20% of the features in a standard IEEE 14-bus system) with the GDAE scheme.

Feature Number	1	7	33	13	0	2	18
Feature Value	−95.556	−95.556	−153.53	153.53	16.22	−95.556	−24.313
Reconstructed Value	−95.48854845	−95.49835622	−153.4731469	153.5859801	16.27401379	−95.50443633	−24.27054722
Error	0.067451553	0.057643775	0.056853103	0.055980148	0.054013795	0.051563669	0.042452777
Error Ratio	0.000705885	0.000603246	0.000370306	0.00036462	0.003330074	0.000539617	0.001746094

Table 28. Seven best-reconstructed features (random attack on 40% of the features in a standard IEEE 14-bus system) with the GDAE scheme.

Feature Number	14	42	22	13	33	19	38
Feature Value	73.554	-43.836	43.836	153.53	−153.53	42.103	−42.103
Reconstructed Value	73.55511276	−43.83366463	43.83835556	153.5325412	−153.5258551	42.10726753	−42.09863462
Error	0.001112764	0.002335371	0.002355563	0.00254121	0.004144926	0.004267531	0.004365384
Error Ratio	1.51285×10^{-5}	5.32752×10^{-5}	5.37358×10^{-5}	1.65519×10^{-5}	2.69975×10^{-5}	0.000101359	0.000103683

Table 29. Seven worst-reconstructed features (random attack on 40% of the features in a standard IEEE 14-bus system) with the GDAE scheme.

Feature Number	2	11	7	12	8	0	19
Feature Value	−48.433	−13.885	−95.556	−13.885	−8.6655	16.22	−62.336
Reconstructed Value	−48.30994115	−13.79823827	−95.47778432	−13.80909392	−8.591674696	16.2896147	−62.26993855
Error	0.123058854	0.086761732	0.078215681	0.075906078	0.073825304	0.069614699	0.066061451
Error Ratio	0.002540806	0.006248594	0.000818532	0.005466768	0.008519451	0.004291905	0.001059764

5.3.2. Error Average Sum (EAS)

Next, to show the reconstruction performance in the overall dataset, we present average error sum (EAS) as another performance gauge. The average error sum is given as follows: $\frac{1}{m}\sum_{i=1}^{m} e_i$, where m is the number of features and e_i is the reconstruction error for the ith feature. The EAS is measured in megawatts (MW). Figures 4–11 show the EAS for the various DAE schemes for the 20% and 40% fixed-attacked features from the standard IEEE 14-, and 39-bus systems.

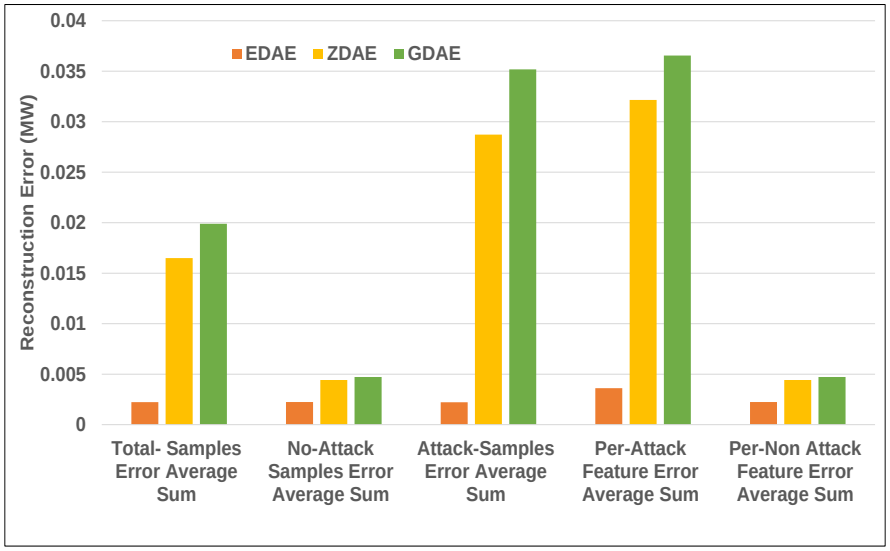

Figure 4. Error average sum for various DAE schemes in a standard IEEE 14-bus system from a 20% fixed-attack dataset.

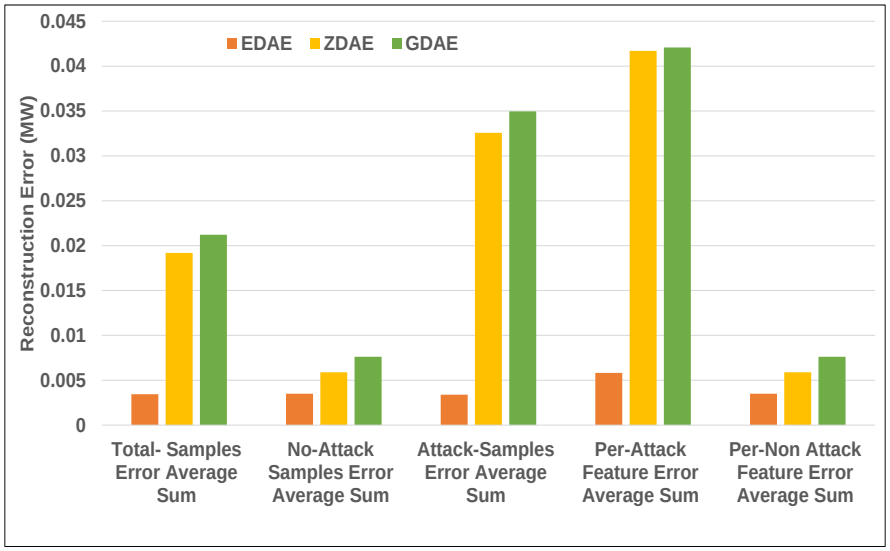

Figure 5. Error average sum for various DAE schemes in a standard IEEE 14-bus system from a 40% fixed-attack dataset.

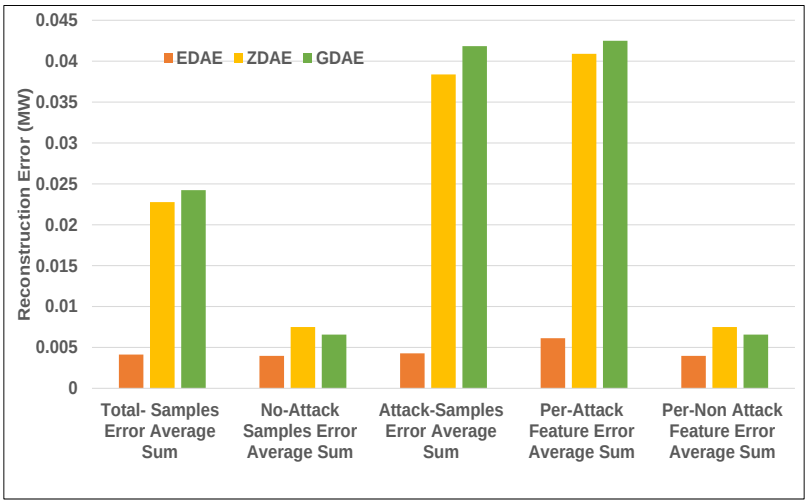

Figure 6. Error average sum for various DAE schemes in a standard IEEE 39-bus system from a 20% fixed-attack dataset.

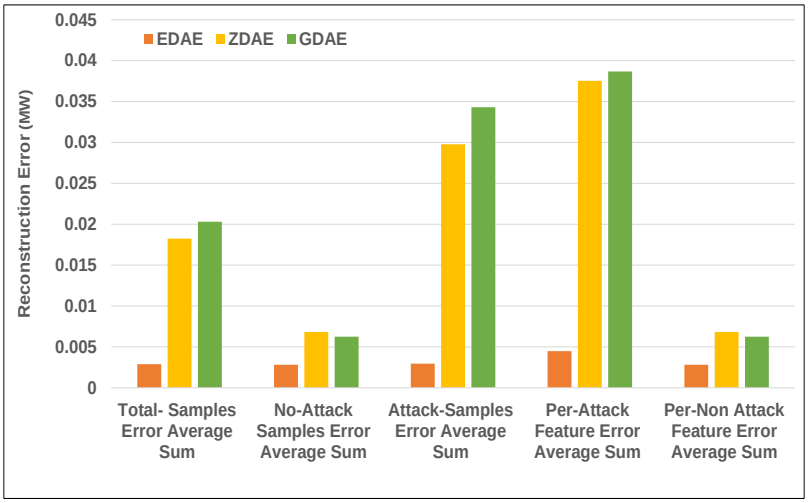

Figure 7. Error average sum for various DAE schemes in a standard IEEE 39-bus system from a 40% fixed-attack dataset.

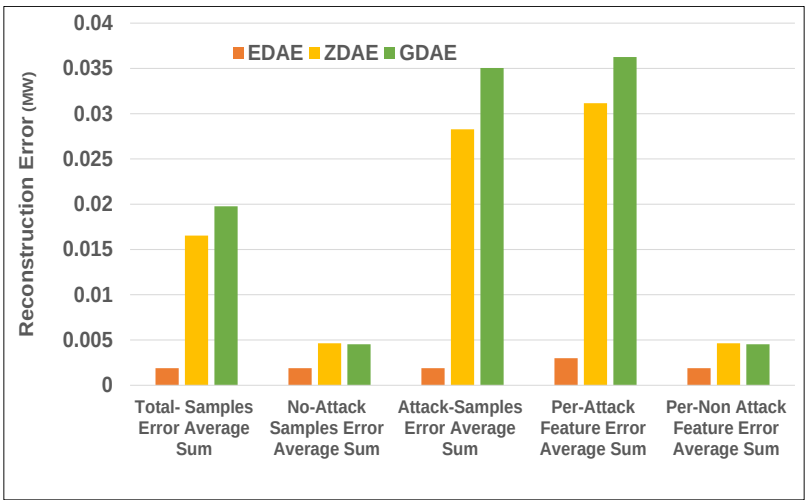

Figure 8. Error average sum for various DAE schemes in a standard IEEE 14-bus system from a 20% random-attack dataset.

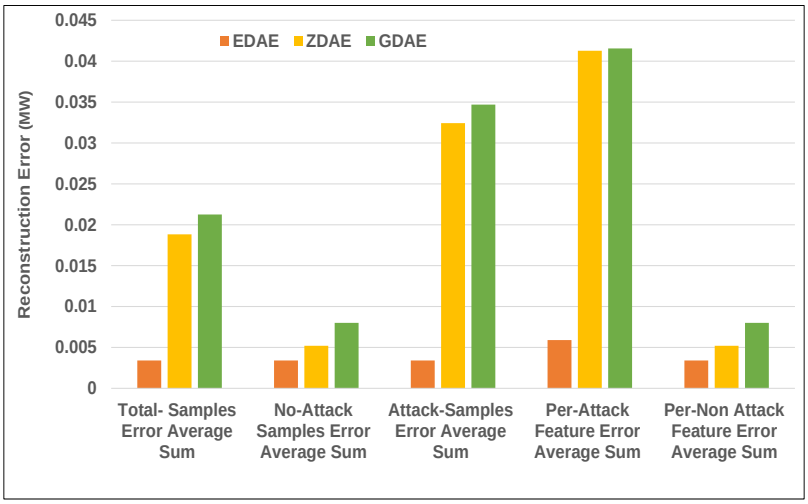

Figure 9. Error average sum for various DAE schemes in a standard IEEE 14-bus system from a 40% random-attack dataset.

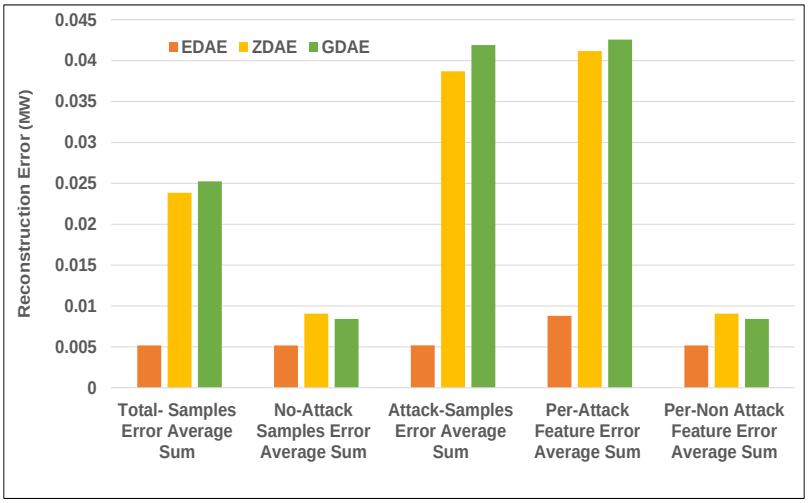

Figure 10. Error average sum for various DAE schemes in a standard IEEE 39-bus system from a 20% random-attack dataset.

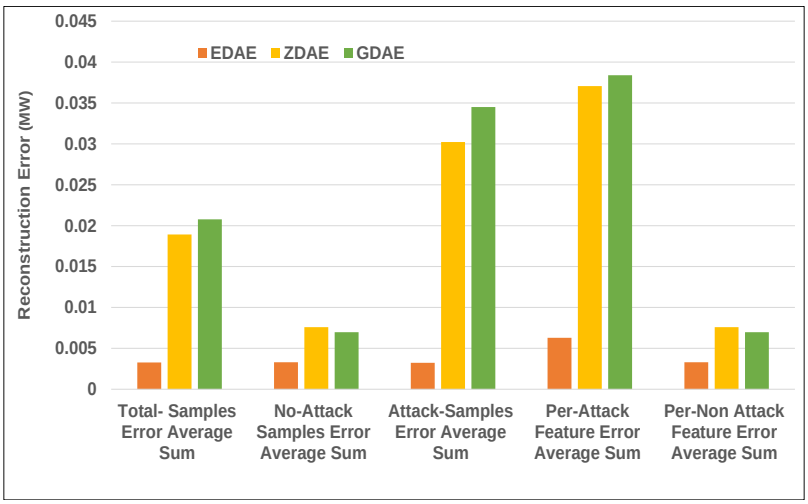

Figure 11. Error average sum for various DAE schemes in a standard IEEE 39-bus system from a 40% random-attack dataset.

Discussion: It is observed that the proposed EDAE scheme has the lowest reconstruction error, compared to the other schemes. The ZDAE outperforms the GDAE, which has the highest reconstruction error.

5.3.3. Training and Validation Costs

Figures 12–19 show the reconstruction error on the training data as a function of the number of epochs for 20% and 40% fixed- and random-attack datasets. These results are shown for standard IEEE 14- and 39-bus systems. The training and validation cost is measured in megawatts (MW).

Discussion: The results of the training and validation cost show that the proposed EDAE scheme results in much less reconstruction error, compared to the ZDAE and GDAE schemes for all test bus cases. It is also observed in the figures that the reconstruction performance of ZDAE is better as compared to the one achieved by the GDAE scheme.

From the overall results, we also observe that, during the reconstruction process, the EDAE scheme performs better than other schemes. The reconstruction error, EAS, and training and validation costs are the lowest in the case of the proposed EDAE-based reconstruction scheme.

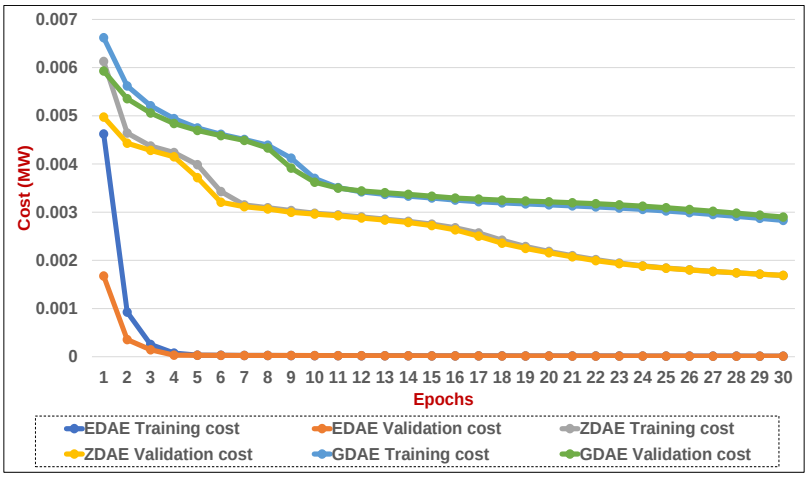

Figure 12. Training and validation costs for various DAE corruption-addition schemes (20% of the features in a fixed attack) in a standard IEEE 14-bus system.

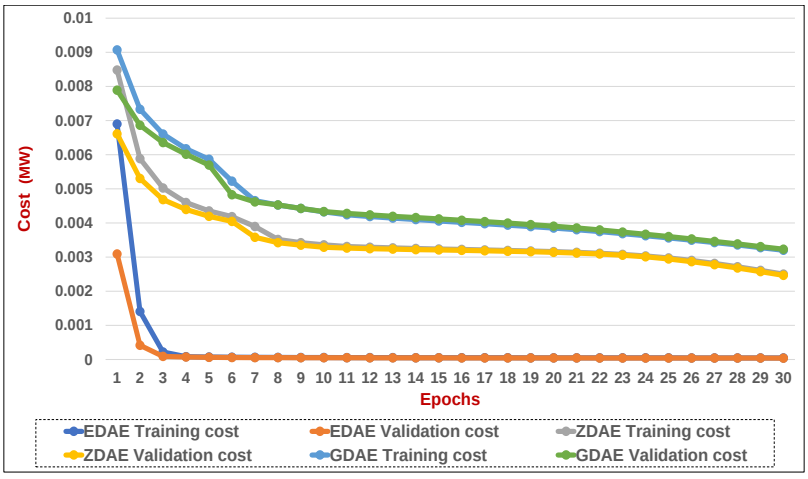

Figure 13. Training and validation costs for various DAE corruption-addition schemes (40% of the features in a fixed attack) in a standard IEEE 14-bus system.

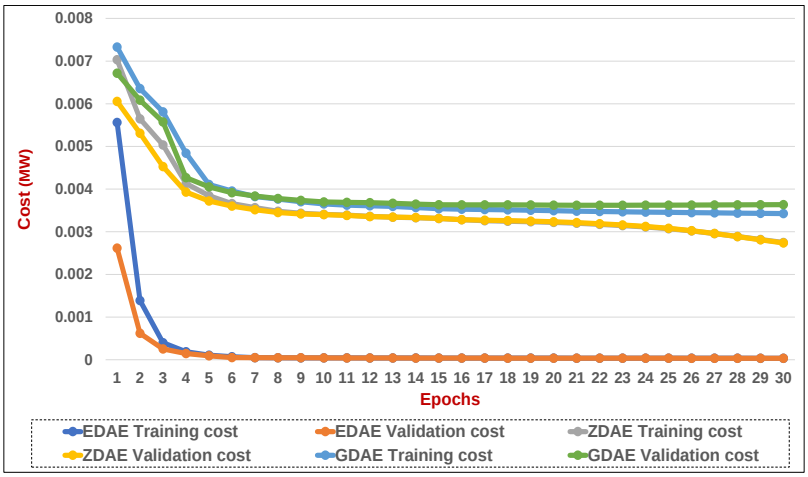

Figure 14. Training and validation costs for various DAE corruption-addition schemes (20% of the features in a fixed attack) in a standard IEEE 39-bus system.

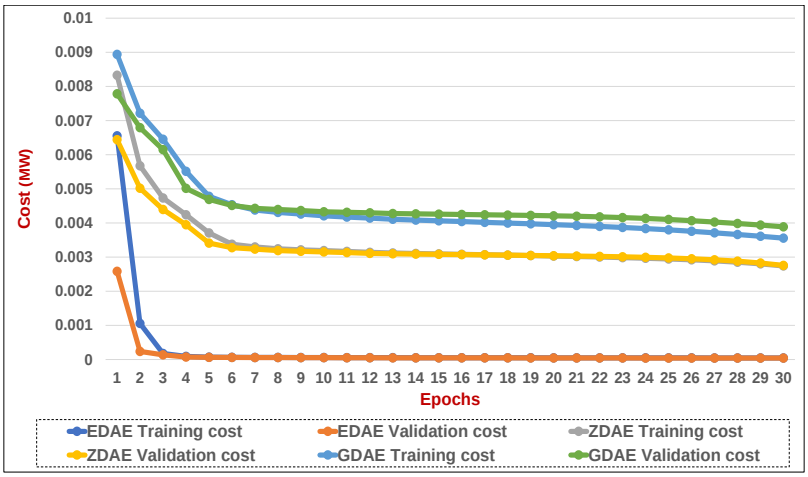

Figure 15. Training and validation costs for various DAE corruption-addition schemes (40% of the features in a fixed attack) in a standard IEEE 39-bus system.

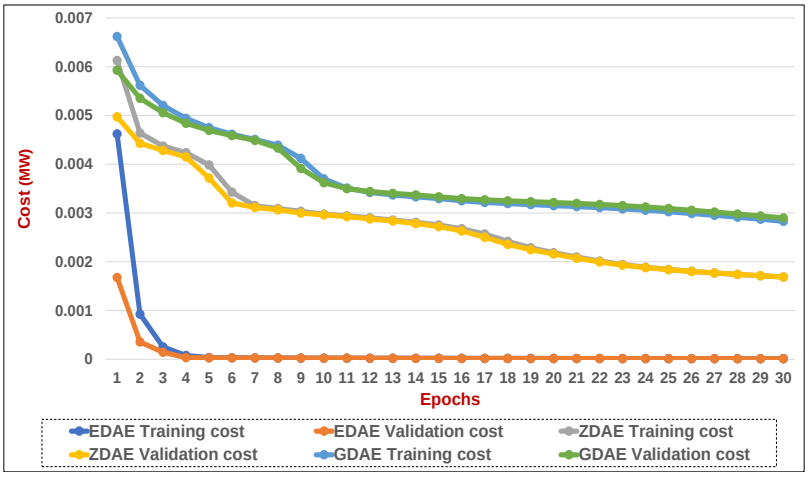

Figure 16. Training and validation costs for various DAE corruption-addition schemes (20% of the features in a random attack) in a standard IEEE 14-bus system.

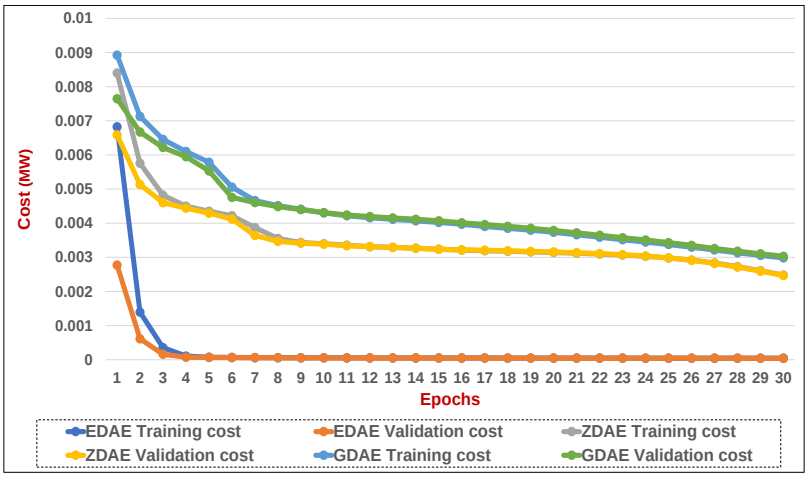

Figure 17. Training and validation costs for various DAE corruption-addition schemes (40% of the features in a random attack) in a standard IEEE 14-bus system.

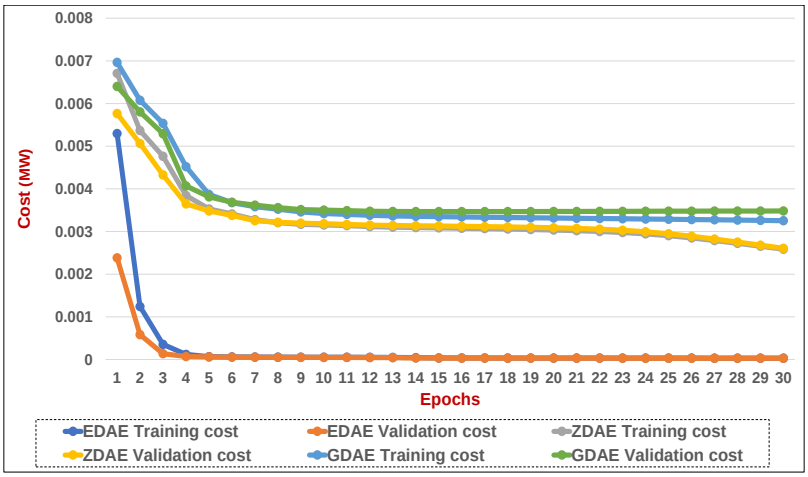

Figure 18. Training and validation costs for various DAE corruption-addition schemes (20% of the features in a random attack) in a standard IEEE 39-bus system.

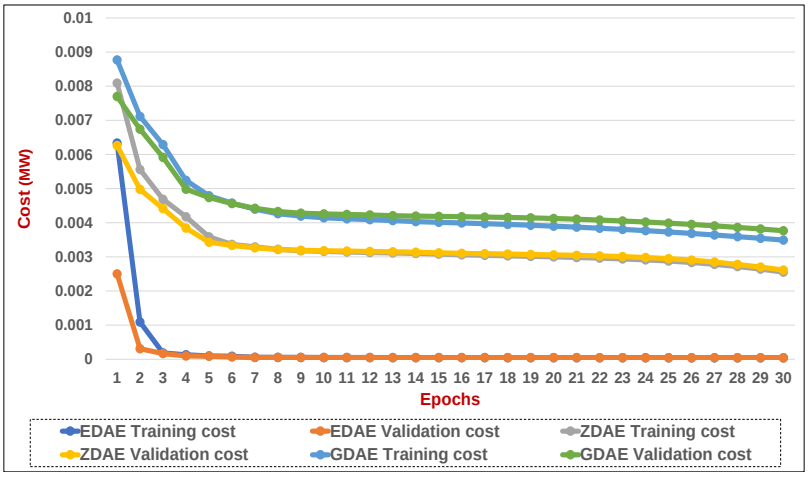

Figure 19. Training and validation costs for various DAE corruption-addition schemes (40% of the features in a random attack) in a standard IEEE 39-bus system.

6. Conclusions

In this paper, we propose a DAE-based scheme to reconstruct the measurements affected by a covert cyber-deception attack while removing the added biased values. We considered different corruption-addition schemes, such as zero-masking (the ZDAE), additive Gaussian noise (the GDAE), and an estimated corruption-addition (termed the EDEA), under diverse attack scenarios. The performance of the proposed scheme was evaluated by employing standard IEEE 14-bus, 39-bus, 57-bus, and 118-bus systems. Active power injections into the buses and active power flow measurements in the branches are the main features of the dataset. The test results show that the proposed EDAE-based reconstruction scheme results in a reasonably low reconstruction error from CCDAs on SG measurement features. Furthermore, the proposed EDAE-based reconstruction scheme results in a low error ratio, compared to the other schemes. However, the features with values closer to zero are reconstructed with a high error ratio. The results were obtained using an MSE objective function. In the future, to further reduce the error ratio for features with small values, we intend to investigate more objective functions in order to increase the reconstruction accuracy.

Author Contributions: All authors conceived and proposed the research idea. S.A. and Y.L. designed the experiments; S.A. performed the experiments; S.-H.H. and I.K. analyzed experimental results; and S.A. wrote the paper in the supervision of S.-H.H. and I.K.

Funding: This research was funded by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (2016R1A6A3A11932461).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Ahmed, S.; Lee, Y.D.; Hyun, S.H.; Koo, I. A Cognitive Radio-Based Energy-Efficient System for Power Transmission Line Monitoring in Smart Grids. *J. Sens.* **2017**, *2017*, 3862375. [CrossRef]
2. Ahmed, S.; Usman, M.; Koo, I. Sensor node selection-based lifetime maximization in sensor network assisted cognitive radio networks. *Adv. Sci. Lett.* **2016**, *22*, 2432–2437. [CrossRef]
3. Liu, Y.; Ning, P.; Reiter, M.K. False data injection attacks against state estimation in electric power grids. *ACM Trans. Inf. Syst. Secur. (TISSEC)* **2011**, *14*, 13. [CrossRef]
4. Khanna, K.; Panigrahi, B.K.; Joshi, A. Data integrity attack in smart grid: Optimised attack to gain momentary economic profit. *IET Gener. Transm. Distrib.* **2016**, *10*, 4032–4039. [CrossRef]
5. Liu, X.; Li, Z. Local load redistribution attacks in power systems with incomplete network information. *IEEE Trans. Smart Grid* **2014**, *5*, 1665–1676. [CrossRef]
6. Ahmed, S.; Lee, Y.; Seung-Ho, H.; Koo, I. Unsupervised Machine Learning—Based Detection of Covert Data Integrity Assault in Smart Grid Networks Utilizing Isolation Forest. *IEEE Trans. Inf. Forensics Secur.* **2019**, *14*, 2765–2777. [CrossRef]
7. Cinar, M.; Kaygusuz, A. Self-Healing in Smart Grid: A Review. *Bitlis Eren Üniversitesi Fen Bilimleri Dergisi* **2018**, *7*, 492–503. [CrossRef]
8. Mo, Y.; Chabukswar, R.; Sinopoli, B. Detecting integrity attacks on SCADA systems. *IEEE Trans. Control Syst. Technol.* **2014**, *22*, 1396–1407.
9. Vamvoudakis, K.G.; Hespanha, J.P.; Sinopoli, B.; Mo, Y. Detection in adversarial environments. *IEEE Trans. Autom. Control* **2014**, *59*, 3209–3223. [CrossRef]
10. Mo, Y.; Weerakkody, S.; Sinopoli, B. Physical authentication of control systems: Designing watermarked control inputs to detect counterfeit sensor outputs. *IEEE Control Syst.* **2015**, *35*, 93–109.
11. Lin, H.; Slagell, A.; Kalbarczyk, Z.; Sauer, P.; Iyer, R. Runtime semantic security analysis to detect and mitigate control-related attacks in power grids. *IEEE Trans. Smart Grid* **2016**, *9*, 163–178. [CrossRef]
12. Sridhar, S.; Govindarasu, M. Model-based attack detection and mitigation for automatic generation control. *IEEE Trans. Smart Grid* **2014**, *5*, 580–591. [CrossRef]
13. Hong, J.; Liu, C.C.; Govindarasu, M. Integrated anomaly detection for cyber security of the substations. *IEEE Trans. Smart Grid* **2014**, *5*, 1643–1653. [CrossRef]

14. Yang, Y.; McLaughlin, K.; Sezer, S.; Littler, T.; Pranggono, B.; Brogan, P.; Wang, H. Intrusion detection system for network security in synchrophasor systems. In Proceedings of the IET International Conference on Information and Communications Technologies (IETICT 2013), Beijing, China, 27–29 April 2013.
15. Fan, Y.; Zhang, Z.; Trinkle, M.; Dimitrovski, A.D.; Song, J.B.; Li, H. A cross-layer defense mechanism against GPS spoofing attacks on PMUs in smart grids. *IEEE Trans. Smart Grid* **2015**, *6*, 2659–2668. [CrossRef]
16. Mitchell, R.; Chen, R. Behavior-rule based intrusion detection systems for safety critical smart grid applications. *IEEE Trans. Smart Grid* **2013**, *4*, 1254–1263. [CrossRef]
17. Zhang, Y.; Wang, L.; Sun, W.; Green, R.C., II; Alam, M. Distributed intrusion detection system in a multi-layer network architecture of smart grids. *IEEE Trans. Smart Grid* **2011**, *2*, 796–808. [CrossRef]
18. Fadlullah, Z.M.; Fouda, M.M.; Kato, N.; Shen, X.; Nozaki, Y. An early warning system against malicious activities for smart grid communications. *IEEE Netw.* **2011**, *25*, 50–55. [CrossRef]
19. Huang, Y.; Tang, J.; Cheng, Y.; Li, H.; Campbell, K.A.; Han, Z. Real-time detection of false data injection in smart grid networks: An adaptive CUSUM method and analysis. *IEEE Syst. J.* **2014**, *10*, 532–543. [CrossRef]
20. Sial, A.; Singh, A.; Mahanti, A.; Gong, M. Heuristics-Based Detection of Abnormal Energy Consumption. In Proceedings of the International Conference on Smart Grid Inspired Future Technologies, Auckland, New Zealand, 23–24 April 2018; pp. 21–31.
21. Sial, A.; Singh, A.; Mahanti, A. Detecting anomalous energy consumption using contextual analysis of smart meter data. *Wirel. Netw.* **2019**. [CrossRef]
22. Saini, S.; Arjunan, P.; Singh, A.; Nambiar, U. E-adivino: A novel framework for electricity consumption prediction based on historical trends. In Proceedings of the 2015 ACM Sixth International Conference on Future Energy Systems, Bangalore, India, 14–17 July 2015; pp. 213–214.
23. Ozay, M.; Esnaola, I.; Vural, F.T.Y.; Kulkarni, S.R.; Poor, H.V. Machine learning methods for attack detection in the smart grid. *IEEE Trans. Neural Netw. Learn. Syst.* **2016**, *27*, 1773–1786. [CrossRef]
24. Esmalifalak, M.; Liu, L.; Nguyen, N.; Zheng, R.; Han, Z. Detecting stealthy false data injection using machine learning in smart grid. *IEEE Syst. J.* **2014**, *11*, 1644–1652. [CrossRef]
25. Ahmed, S.; Lee, Y.; Hyun, S.H.; Koo, I. Covert Cyber Assault Detection in Smart Grid Networks Utilizing Feature Selection and Euclidean Distance-Based Machine Learning. *Appl. Sci.* **2018**, *8*, 772. [CrossRef]
26. Ahmed, S.; Lee, Y.; Hyun, S.H.; Koo, I. Feature Selection—Based Detection of Covert Cyber Deception Assaults in Smart Grid Communications Networks using Machine Learning. *IEEE Access* **2018**, *6*, 27518–27529. [CrossRef]
27. Singh, S.K.; Khanna, K.; Bose, R.; Panigrahi, B.K.; Joshi, A. Joint-Transformation-Based Detection of False Data Injection Attacks in Smart Grid. *IEEE Trans. Ind. Inform.* **2018**, *14*, 89–97. [CrossRef]
28. He, Y.; Mendis, G.J.; Wei, J. Real-time detection of false data injection attacks in smart grid: A deep learning-based intelligent mechanism. *IEEE Trans. Smart Grid* **2017**, *8*, 2505–2516. [CrossRef]
29. Chen, G.; Dong, Z.Y.; Hill, D.J.; Xue, Y.S. Exploring reliable strategies for defending power systems against targeted attacks. *IEEE Trans. Power Syst.* **2010**, *26*, 1000–1009. [CrossRef]
30. Chen, P.Y.; Cheng, S.M.; Chen, K.C. Smart attacks in smart grid communication networks. *IEEE Commun. Mag.* **2012**, *50*, 24–29. [CrossRef]
31. Ma, C.Y.; Yau, D.K.; Lou, X.; Rao, N.S. Markov game analysis for attack-defense of power networks under possible misinformation. *IEEE Trans. Power Syst.* **2012**, *28*, 1676–1686. [CrossRef]
32. Framework, N. *Roadmap for Smart Grid Interoperability Standards*; National Institute of Standards and Technology: Gaithersburg, MD, USA, 2010.
33. El-Hawary, M.E. The smart grid—State-of-the-art and future trends. *Electr. Power Compon. Syst.* **2014**, *42*, 239–250. [CrossRef]
34. Jiang, G.; Xie, P.; He, H.; Yan, J. Wind turbine fault detection using a denoising autoencoder with temporal information. *IEEE/ASME Trans. Mechatron.* **2018**, *23*, 89–100. [CrossRef]
35. Ohki, T.; Otsuka, A. A Study on Autoencoder-based Reconstruction Method for Wi-Fi Location Data with Erasures. In Proceedings of the 2017 on Multimedia Privacy and Security, Dallas, TX, USA, 30 October–3 November 2017; pp. 13–18.
36. Tschannen, M.; Bachem, O.; Lucic, M. Recent Advances in Autoencoder-Based Representation Learning. *arXiv* **2018**, arXiv:1812.05069.
37. Huang, R.; Liu, C.; Li, G.; Zhou, J. Adaptive deep supervised autoencoder based image reconstruction for face recognition. *Math. Probl. Eng.* **2016**, *2016*, 6795352. [CrossRef]

38. Jiang, G.; He, H.; Xie, P.; Tang, Y. Stacked multilevel-denoising autoencoders: A new representation learning approach for wind turbine gearbox fault diagnosis. *IEEE Trans. Instrum. Meas.* **2017**, *66*, 2391–2402. [CrossRef]
39. Vincent, P.; Larochelle, H.; Bengio, Y.; Manzagol, P.A. Extracting and composing robust features with denoising autoencoders. In Proceedings of the 25th international conference on Machine Learning, Helsinki, Finland, 5–9 July 2008; pp. 1096–1103.
40. Gomez-Exposito, A.; Abur, A. *Power System State Estimation: Theory and Implementation*; CRC Press: Boca Raton, FL, USA, 2004.
41. Casazza, J.; Casazza, J.; Delea, F. *Understanding Electric Power Systems: An Overview of the Technology and the Marketplace*; John Wiley & Sons: Hoboken, NJ, USA, 2003; Volume 13.
42. Abdallah, A.; Shen, X. *Security and Privacy in Smart Grid*; Springer: Cham, Switzerland, 2018.
43. Yuan, Y.; Li, Z.; Ren, K. Modeling load redistribution attacks in power systems. *IEEE Trans. Smart Grid* **2011**, *2*, 382–390. [CrossRef]
44. Bi, S.; Zhang, Y.J. Using covert topological information for defense against malicious attacks on DC state estimation. *IEEE J. Sel. Areas Commun.* **2014**, *32*, 1471–1485. [CrossRef]
45. Deng, R.; Xiao, G.; Lu, R. Defending against false data injection attacks on power system state estimation. *IEEE Trans. Ind. Inform.* **2017**, *13*, 198–207. [CrossRef]
46. Patro, S.; Sahu, K.K. Normalization: A preprocessing stage. *arXiv* **2015**, arXiv:1503.06462.
47. Sze, V.; Chen, Y.H.; Yang, T.J.; Emer, J.S. Efficient processing of deep neural networks: A tutorial and survey. *Proc. IEEE* **2017**, *105*, 2295–2329. [CrossRef]
48. Zimmerman, R.D.; Murillo-Sánchez, C.E.; Thomas, R.J. MATPOWER: Steady-state operations, planning, and analysis tools for power systems research and education. *IEEE Trans. Power Syst.* **2011**, *26*, 12–19. [CrossRef]