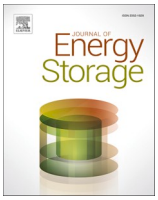
Journal of Energy Storage 92 (2024) 112225

Contents lists available at ScienceDirect

Journal of Energy Storage

journal homepage: www.elsevier.com/locate/est



Research papers

NPformer based static FDIAs detection for state-of-charge estimation of battery energy storage systems in smart distribution networks

Zhiying Liu, Yuancheng Li *, Shuang Xu, Qingle Wang, Jianbin Li

School of Control and Computer Engineering, North China Electric Power University, No. 2 Beinong Road, Changping District 102206, Beijing, China

ARTICLE INFO

Keywords:
Smart distribution networks
Battery energy storage systems
State of charge estimation
False data injection attacks

ABSTRACT

State of charge (SoC) estimation of battery energy storage systems is essential for ensuring the security, stability, and SoC estimation of battery energy storage systems (BESSs) in smart distribution networks (SDNs) is critical to the control and operation of power systems. False data injection attacks (FDIAs) can escape bad data detection, thus affecting the SoC estimation of BESSs. Existing model-based detection methods require the manual calculation of thresholds, so this paper proposes an NPformer-based method for detecting FDIAs on BESSs in SDNs. The method constructs a pyramidal attention model based on a multi-scale C-ary tree to explore the multi-resolution representation of time series. Then, the network structure of the NPformer is automatically searched with neural architecture search to obtain an effective detection model. We evaluate the performance of the proposed method using the IEEE13 and IEEE33 bus systems with BESSs. Compared to Transformer and Long Short Term Memory autoencoder, the experimental results show that the proposed method has higher detection accuracy for close-time use. The proposed method has a detection accuracy of 97 %.

1. Introduction

Battery energy storage systems (BESSs) can eliminate the volatility of distributed energy generation, improve power quality, and enhance the flexibility and reliability of smart distribution networks (SDNs) [1]. As an important energy storage element, the state of charge (SoC) of the battery directly affects the stable operation of the BESSs [2]. Traditionally, SoC estimation of BESSs has been done in a battery management system (BMS). With the proliferation of BESSs and the development of Internet of Things technologies, BMS is gradually integrating Internet of Things and cloud computing technologies to ensure that BESSs have secure, reliable, and accurate SoC estimation [3–5]. BESSs have now become complex cyber-physical systems. However, this has also led to a serious threat of malicious cyber attacks on BESSs, such as random delay attacks, man-in-the-middle attacks, false data injection attacks (FDIAs), SQL injection attacks, denial of service attacks, etc. [6]. It is worth noting that FDIAs achieve their attacks by injecting fake data into the data in question, which is different from a fault. Faults in sensors, communications, etc. are often caused by faulty hardware.

FDIAs can bypass the bad data detection (BDD) of distribution system state estimation (DSSE) and then cause BESSs to produce false SoC estimation. The incorrect SoC estimation may cause the BESSs to be overcharged or discharged, leading to battery aging, system damage, and even fire or explosion. These will seriously affect the safe, stable, and economic operation of SDNs [7]. For example, a former NSA researcher used the BMS to transfer low SoC in 2015. This allowed the BESSs to perform charge and discharge control in violation of mechanical and electrical safety requirements, thus impacting the safe and stable operation of SDNs [6]. Therefore, effective detection of FDIAs for SoC estimation of BESSs is relevant to ensure the reliable operation of SDNs.

In recent years, FDIAs have attracted the attention of domestic and international scholars. A cyber-physical model that integrates capturing ideal measurements and a generative adversarial network approach to defend against FDIAs is proposed by Y. Li et al. [8]. In [9], the SoC of each cell is estimated by using an adaptive extended Kalman filter (EKF) algorithm, and the voltage residuals are evaluated by statistical inference methods to determine the presence of FDIAs. In [10], attacks are constructed by adding data to a set of sensors, and then it analyzes the limited range of detectable attacks with a cardinality detector. Obrien V et al. propose to introduce a priori residual data (that is estimated by the EKF) into the cumulative sum algorithm to obtain the minimum detection diameter of the model, which detects small-magnitude false count data injection attacks [11]. The cumulative sum, physics-based models,

* Corresponding author.
E-mail address: ycli@ncepu.edu.cn (Y. Li).

https://doi.org/10.1016/j.est.2024.112225
Received 8 July 2023; Received in revised form 19 May 2024; Accepted 20 May 2024
Available online 25 May 2024

and EKF algorithm are used in [12] to accurately estimate the SoC of a series-connected battery, enabling the detection of random FDIAs for cell and stack voltage sensors. For the FDIAs of SOC estimation in BESSs. In [13], an improved moving target defence approach is proposed to provide active defence against static FDIAs in battery energy storage systems. Z. Liu et al. proposed a Transformer-based improved GAN approach for passive defence against FDIAs in battery energy storage systems [14]. Zhuang P et al. propose a static FDIAs attack model for SoC estimation of BESSs in SDNs that can escape the BDD in conventional DSSE, and then they proposed a residual-based detection method [6]. However, the thresholds for this method require complex manual calculation and determination.

Therefore, inspired by the FDIAs intelligent detection algorithm, this paper innovatively proposes an NAS-Pyramidal-Transformer (NPformer) detection method for FDIAs of BESSs in SDNs. The main contributions of this paper are as follows: (i) To avoid the complex computational process in existing threshold-based detection methods, we innovatively propose a machine-learning approach to detect FDIAs. (ii) Combining a C-ary tree based Pyramidal Attention Model with Neural Architecture Search to better learn the spatiotemporal characteristics of measurement data. (iii) Avoiding the complex manual tuning process of traditional machine learning methods and achieving automatic setting of optimal model parameters.

The remainder of this paper is organized as follows: Section 2 describes the principles for the construction of static FDIAs for SoC estimation of BESSs in SDNs. Section 3 proposes a static FDIAs detection method for BESSs based on NPformer. Section 4 conducts an experimental study of the IEEE13 and IEEE33 bus systems with BESSs, and Section 5 summarizes this manuscript.

2. Problem description

In SDNs, BESSs convert the electric energy of the grid into storable chemical energy through an external interface and convert it into electric energy when needed [15]. The power conversion system provides the external interface, which is the key component of BESSs in SDNs [16]. In this paper, an SDN containing a BESS is used as an example and its structure is shown in Fig. 1. The BESS is connected to a point of common coupling on the substation through the voltage source converter (VSC), transformer, and filter [17]. The BESS consists of a battery pack, battery management system (BMS), BESS controller, VSC, transformer, and filter. The BMS not only provides real-time measurements of battery pack voltage, current and temperature through sensors, but is also capable of battery SoC estimation and battery equalization management [18]. The BESS controller determines the magnitude modulation m^{abc} and phase displacement angle $\Delta\theta^{abc}$ based on the set values of active and reactive power calculated by the substation control center, and transmits the data and statuses of the BESS to the substation control center.

In this paper, the measurement data and state vectors of the system contain information about the SDNs and the BESSs. In addition, potential FDIAs can occur at smart inverter and BMS (which are shown in Fig. 1). The attack vectors can be divided into amplitude modulation, the voltage and current on the DC side of the BESSs, and the corresponding components of active and reactive power on the AC side of the BESSs. Therefore, to achieve FDIAs for SoC estimation of BESSs, the attack must be able to escape the BDD [6]. Considering that the attack can occur in smart inverters on the AC side of the BESSs and the system voltage and current are variable, analysis of physical battery properties (e.g. C-rate) cannot detect FDIAs well [19]. Coupled with the fact that existing threshold-based detection methods require complex calculations, we innovatively propose the NPformer detection method. In this paper, system steady states for discrete-time slots with equal length of Δt is considered, and the model error estimated by the DSSE and SoC is negligible [20].

2.1. State estimation of SDNs with BESSs

In SDNs with BESSs, the measurement vector $z = \left[|\overline{V}|, |\overline{I}|, \overline{\theta}, \overline{P}, \overline{Q}, |\overline{V}|_{ac}, |\overline{I}|_{ac}, \overline{P}_{ac}, \overline{Q}_{ac}, \overline{m}, \Delta\overline{\theta}, \overline{V}_{dc}, \overline{I}_{dc}\right]$. Where $|\overline{V}|$ is the bus voltage magnitude. The branch's current magnitude is $|\overline{I}|$, which is measured by remote terminal units of intelligent electronic devices and Supervisory Control And Data Acquisition systems. The bus voltage phase angle $\overline{\theta}$ is measured by power management units (PMUs). $\overline{P}$ and $\overline{Q}$ are the bus active and reactive power respectively. $\overline{P}_{ac}$ and $\overline{Q}_{ac}$ are the active power and reactive power of the BESSs. The measurement vector also contains the battery terminal voltage $\overline{V}_{dc}$ and current $\overline{I}_{dc}$ on the DC-side of BESSs, the AC-side voltage magnitude $|\overline{V}|_{ac}$, current magnitude $|\overline{I}|_{ac}$ of BESSs, the magnitude modulation $\overline{m}$ and phase displacement

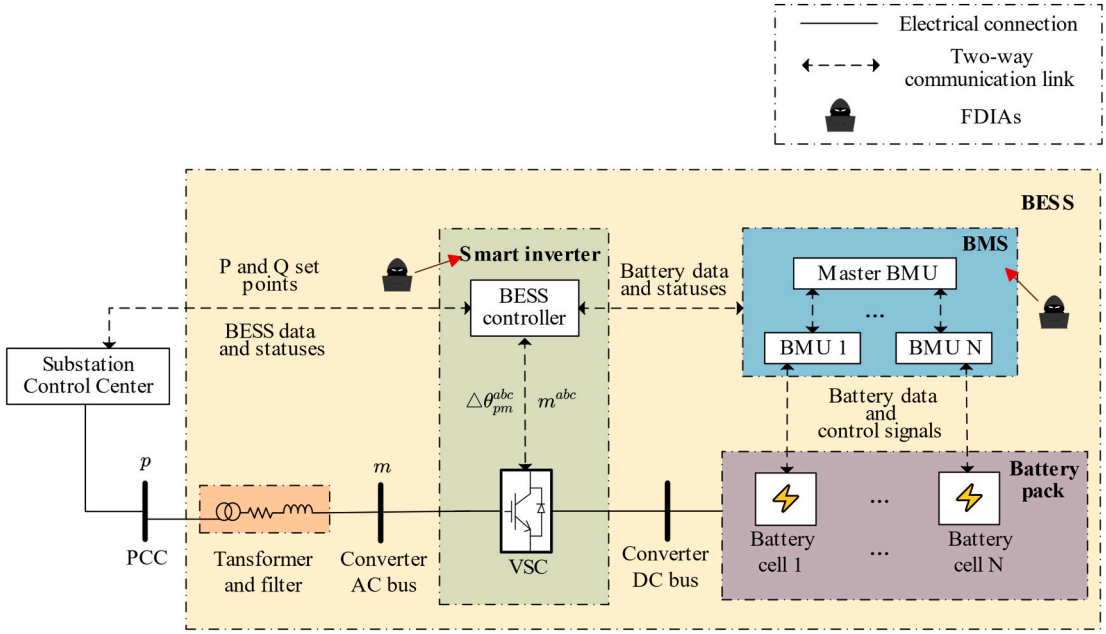


Fig. 1. Structure of a smart distribution network with a BESS.

angle $\Delta\bar{\theta}$ of VSC [6].

In addition, due to the phase imbalance in the distribution network, a three-phase model is required. According to Kirchhoff's current law, the current $I^{abc} = \left[I_1^{abc},\ I_2^{abc},\ \ldots\ , I_N^{abc}\right]^T$ injected into the bus can be calculated by Eq. (1).

$$I^{abc} = Y^{abc} V^{abc} \tag{1}$$

where $V^{abc} = \left[V_1^{abc},\ V_2^{abc},\ \ldots\ , V_N^{abc}\right]^T$ is the bus voltage, Y^{abc} is the nodal admittance matrix, which is calculated by the diagonal (to eliminate the voltage variables, we insert virtual derivatives for the missing phases) [21].

For BESSs, the coupling transformer and filter can be modeled as the admittance matrix y_{mp}^{abc}. The VSC can be modeled using the magnitude modulation m^{abc}, phase displacement angle $\Delta\theta^{abc}$, AC side resistor R_{ac}, and DC side resistor R_{dc} [22]. The magnitude modulation and phase displacement angle in phase α are calculated as:

$$m^{\alpha} = \sqrt{2}|V|_m^{\alpha} V_{dc} \tag{2}$$

$$\Delta\theta_{pm}^{\alpha} = \theta_m^{\alpha} - \theta_p^{\alpha} \tag{3}$$

where $|V|_m^{\alpha}$ is the AC-side voltage magnitude (e.g. the voltage magnitude of phase α of bus m in Fig. 1). V_{dc} is the DC-side voltage. Separately, θ_m and θ_p are the phase angles of the terminal buses of the coupling transformer (e.g. the phase angle of bus m and bus p in Fig. 1). The state vector of the system includes the magnitude and phase angle of all bus voltages, both on the AC and DC sides of the VSC. The state vector is $x = \left[|\overline{V}|_1^{abc},\ \ \theta_1^{abc}\right]^T$.

In the AC model, the measurement vector is related to the state vector [6].

$$z = h(x) + \delta \tag{4}$$

where $h(\cdot)$ is a nonlinear mapping of the state vector to the measurement vector. δ is the measurement noise.

The purpose of DSSE is to obtain the system state estimation that best fits the measurement value according to Eq. (4). The weighted least squares (WLS) method is generally used for state estimation in power systems. The optimization problem of WLS-based state estimation is shown in the following equation:

$$\hat{x} = \underset{x}{\operatorname{argmin}} (z - h(x))^T \mathrm{W} (z - h(x)) \tag{5}$$

where W is the measurement error variance array with σ_i^2 as the diagonal element. In the objective function, $h(\cdot)$ is a nonlinear function, which cannot be solved directly by linear equations for the state variables and needs to be solved iteratively by the Gauss-Newton method.

Traditionally, DSSE performs bad data detection based on measurement residuals $\|r\|_2 = \|z - h(\hat{x})\|_2$, where $\hat{x}$ is the estimated system state vector. The threshold τ is determined by a hypothesis test with a significance level λ, and $\|r\|_2 > \tau$ represents bad data with a false positive rate of λ.

2.2. SoC estimation of BESSs based on EKF

To meet the load voltage and power requirements in SDNs, the battery pack consists of multiple battery cells. Considering the high computational cost of SoC estimation for all battery cells, a battery cell model with a battery filter is selected [23]. The battery cell model adopts the first-order RC network, as shown in Fig. 2. $g(\Phi)$ is a nonlinear function that maps the SoC to the open circuit voltage, which can be linearized to a linear function with slope β. R_b is the ohmic resistance of the cell. V_{dc} and I_{dc} are the terminal voltage and current of the cell respectively. These parameters are identified by the recursive least squares method.

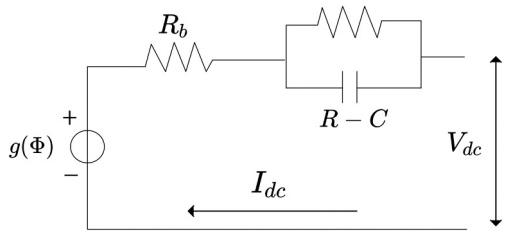


Fig. 2. Battery cell model.

Considering that the SoC in the state space equation of the BESSs is non-linear and the EKF algorithm has been widely used [24], we use the EKF algorithm for SoC estimation. The EKF algorithm for the SoC estimation process mainly consists of two phases: prediction and correction. The forecasting phase focuses on projecting the forecast SoC ($\widetilde{\Phi}_t$) at time slot t based on the estimated SoC ($\widehat{\Phi}_{t-1}$) at time slot $t-1$ [6].

$$\widetilde{\Phi}_t = \widehat{\Phi}_{t-1} + \frac{\Delta t}{c} \vec{I}_{dc}^{t-1} \tag{6}$$

where c is the battery pack capacity and $\vec{I}_{dc}^{t-1}$ is the measured current at time slot $t-1$. The covariance matrix of prediction is $\widetilde{\mathscr{P}}_t$ (calculated as $\widetilde{\mathscr{P}}_t = \widehat{\mathscr{P}}_{t-1} + \mathbb{Q}$, where $\mathbb{Q}$ is the variance of the Gaussian process noise.). In practice, $\mathbb{Q}$ is a combination of the overall modeling error. The value of $\mathbb{Q}$ is defined concerning the measurement noise variance and is determined empirically [25].

In the correction phase, the predicted SoC is corrected iteratively based on the difference between the measured voltage $\overline{V}_{dc}^t$ and the predicted voltage $\widetilde{V}_{dc}^t$ at time slot t.

$$K_t = \frac{\partial g(\Phi)}{\partial \Phi}\bigg|_{\widetilde{\Phi}_t} \widetilde{\mathscr{P}}_t \left[\left(\frac{\partial g(\Phi)}{\partial \Phi}\bigg|_{\widetilde{\Phi}_t} \right)^2 \widetilde{\mathscr{P}}_t + \sigma^2 \right] \tag{7}$$

$$\widehat{\Phi}_t = \widetilde{\Phi}_t + K_t (\overline{V}_{dc}^t - \widetilde{V}_{dc}^t) \tag{8}$$

where K_t is the Kalman gain and σ^2 is the assumed measurement noise variance. The covariance matrix of estimation is updated to $\widehat{\mathscr{P}}_t = \left(1 - K_t \frac{\partial g(\Phi)}{\partial \Phi}\Big|_{\widetilde{\Phi}_t}\right) \widetilde{\mathscr{P}}_t$. The initial SoC estimation $\widehat{\Phi}_0$ is obtained after the battery pack has been staying in the open circuit for several hours.

It can be seen that the EKF-based SoC estimation is related to both the terminal voltage and the current of the BESSs. Therefore, the injection of false data into the measurement data will affect the SoC estimation of BESSs, threaten the safe, stable, and economical operation of the smart distribution networks with BESSs.

2.3. Construction of FDIAs

Since the distribution network has a BDD mechanism, the WLS-based DSSE algorithm must be considered to successfully inject false data into the measurement data. Moreover, in order to produce a specific attack effect on the SoC estimation of BESSs, we no longer perform random attacks as [11], but construct FDIAs on the EKF-based SoC estimation algorithm in collaboration [6].

For FDIAs in SDNs, the principle of constructing static FDIAs for a time slot in [26], adds an attack vector to the measurement data and can bypass the detection of bad data based on measurement residuals, as shown in Eq. (9).

$$\varepsilon = h(\hat{x} + v) - h(\hat{x}) \tag{9}$$

where ε is the attack vector and v is the state estimation error caused by

the attack vector.

For FDIAs of BESSs, the attacker can cause SoC estimation errors with a moderate attack and thus gain illegal profits through the deregulated electricity market. Constructing FDIAs in this way has less impact on the operation of SDNs and the attack effect lasts longer [6]. In [6], the SoC estimation error at time slot t can be expressed by the following equation.

$$\widehat{\Phi}_t = \widehat{\Phi}_{t-1} + \frac{\Delta t}{c}\overline{I}_{dc}^{t-1} + K_t\left(\overline{V}_{dc}^{t} - \widetilde{V}_{dc}^{t}\right) \tag{10}$$

By injecting the attack vector to the terminal voltage and current measurement at time slot t, the difference between the real SoC and the estimated SoC is $\Delta\Phi_t^{\varepsilon}$, which can be expressed by the following equation:

$$\begin{aligned}\Delta\Phi_t^{\varepsilon} &= \Phi_t - \widehat{\Phi}_t^{\varepsilon} \\ &= (1-\beta K_t)\left(\Delta\Phi_{t-1} - \frac{\Delta t}{c}\delta_{\overline{I}_{dc}}^{t-1}\right) \\ &\quad + K_t\left(R_b\delta_{\overline{I}_{dc}}^{t} - \delta_{\overline{V}_{dc}}^{t} + R_b\varepsilon_{\overline{I}_{dc}}^{t} - \varepsilon_{\overline{V}_{dc}}^{t}\right) \\ &= K_t\left(R_b\varepsilon_{\overline{I}_{dc}}^{t} - \varepsilon_{\overline{V}_{dc}}^{t}\right) + \Delta\Phi_t\end{aligned} \tag{11}$$

where $\widehat{\Phi}_t^{\varepsilon}$ is an incorrect SoC estimate caused by FDIAs ε under the time slot t. $\varepsilon_{\overline{I}_{dc}}^{t}$ and $\varepsilon_{\overline{I}_{dc}}^{t}$ are the values of the DC side voltage and current of the BESSs injected by the FDIAs with false data at time slot t. $\delta_{\overline{V}_{dc}}^{t}$ and $\delta_{\overline{I}_{dc}}^{t-1}$ are the measurement noise of the BESSs DC side voltage and current at time slot t and time slot $t-1$, respectively.

According to Eq. (11), the attacker deviates $\Delta\Phi_t$ to $\Delta\Phi_t^{\varepsilon}$ by injecting an attack ε_t. The attacker's objective is to maximize SoC estimation error, so the attack optimization objective of the attacker can be expressed by the following equation:

$$\begin{aligned}&\max_{\varepsilon_t} |\Delta\Phi_t^{\varepsilon}| \\ &subject\ to\ \varepsilon_t = h(\widehat{x}_t + v_t) - h(\widehat{x}_t) \\ &\quad e_{\overline{m}^a}^{T}(z_t+\varepsilon_t) \le \overline{m}, \forall a \in \{a,b,c\} \\ &\quad -\overline{I_{dc}} \le e_{\overline{I}_{dc}}^{T}(z_t+\varepsilon_t) \le \overline{I_{dc}} \\ &\quad \underline{V_{dc}} \le e_{\overline{V}_{dc}}^{T}(z_t+\varepsilon_t) \le \overline{V_{dc}} \\ &\quad e_{\overline{P}_{ac}}^{T}(z_t+\varepsilon_t) = P_{ac}^{set} \\ &\quad e_{\overline{Q}_{ac}}^{T}(z_t+\varepsilon_t) = Q_{ac}^{set}\end{aligned} \tag{12}$$

where $e_{\{\cdot\}}$ is the base vector. $\overline{m}$ is the upper limit of magnitude modulation. $\overline{I_{dc}}$ is the limitation of the battery pack current. $\underline{V_{dc}}$ and $\overline{V_{dc}}$ are the limitation of the battery pack voltage. P_{ac}^{set} and Q_{ac}^{set} are the active power and reactive power of BESSs. When the EKF is in a steady state with stable Kalman gain coefficients, the attacker can construct the maximum attack vector according to Eq. (12), which causes the maximum SoC estimation error and threatens the safe, stable, and economical operation of the distribution network.

3. NPformer-based static FDIAs detection model for BESSs

To avoid the complex computational process when using thresholding for detection in [6], this paper chooses a machine learning approach for FDIAs detection. Compared with convolutional neural network (CNN) and Long Short Term Memory (LSTM) networks, Transformer can better extract data depth features [27]. However, Transformer does not easily capture multi-scale time dependencies and is computationally very expensive. In order to solve this problem and explore a multi-resolution representation of the measurement data, this paper proposes to construct a multi-scale C-ary tree using the coarser-scale construction module (CSCM) module and build a pyramidal attention model (PAM) on top of it. In addition, the Transformer's network structure is complex in design and the parameters need to be set manually (such as learning rate, position embedding, layernorm, etc.). Therefore, inspired by [28], the Neural Architecture Search (NAS) algorithm is used in the NPformer method. The NPformer network structure is designed as shown in Fig. 3. The NPformer network structure mainly consists of the embedding module, CSCM module, PAM module, Classification module, and NAS algorithm.

Each of the improved Transformer selection modules on the left side of Fig. 3 corresponds to a network structure that is automatically designed by the NAS. The structure is dynamic, with the colored parts indicating the selected model parameters and network structure. The dashed parts indicate that they are not selected. On the right side, a detailed description of the improved Transformer selection module is shown. After the NAS calculates the detection accuracy of all possible options, the best network structure and model parameters will be generated by comparison.

The flowchart of NPformer is shown in Fig. 4.

3.1. Embedded modules

The embedding module separates the time data (Time) from the other data in the input data. The other data includes the EKF-based SoC estimation, bus voltage magnitude $|\overline{V}|$, branch current magnitude $|\overline{I}|$, voltage phase angle $\overline{\theta}$, bus active power $\overline{P}$ and reactive power $\overline{Q}$, BESSs AC-side voltage magnitude $|\overline{V}|_{ac}$, current magnitude $|\overline{I}|_{ac}$, active power $\overline{P}_{ac}$, reactive power $\overline{Q}_{ac}$, control variables $\overline{m}$ and $\Delta\overline{\theta}$, BESSs DC-side voltage $\overline{V}_{dc}$ and current $\overline{I}_{dc}$.

The time data is encoded separately as a timestamp, while the normal data is data encoded separately and then the corresponding location code is added to the completed data encoding. The embedding module adds up the timestamp, the data code of the common data, and the corresponding position code as the output of the embedding module.

3.2. CSCM module

The CSCM module initializes nodes on the coarser scales of the pyramidal graph so that subsequent PAM can exchange information between these nodes. Specifically, the coarser scale nodes are introduced step by step from bottom to top by convolving the sub-nodes of the original sequence. The convolutional kernel C and step size C are applied to the embedded sequence sequentially in the time dimension. Sequences of length L/C^s are generated at scale s. Sequences of different

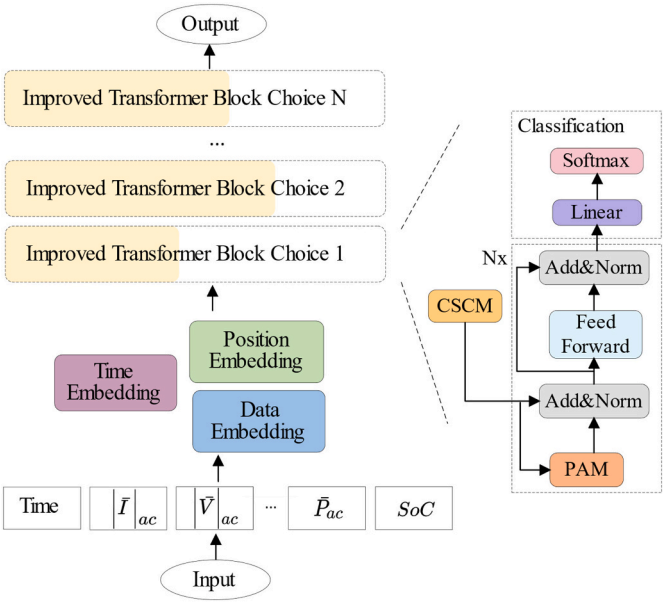


Fig. 3. NPformer overall architecture diagram.

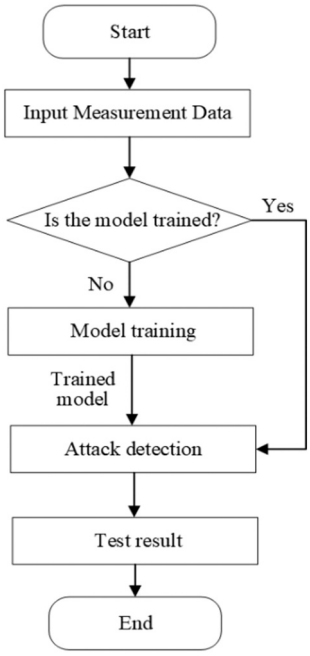


Fig. 4. NPformer flowchart.

scales form C-ary trees. In addition, to reduce the model parameters and computational effort, a fully connected layer is used to reduce the dimensionality of each node before the resulting sequences are fed into the stacked convolutional layers and recovered after all convolutions are completed. Therefore, the CSCM model reduces the number of parameters in the module and prevents over-fitting of the model.

3.3. PAM module

PAM uses pyramid diagrams to describe the time dependence of sequences in a multi-resolution manner. The pyramid diagram is divided into two main parts: finest scale connections and coarse scale connections. The connection between scales is made by associating the finest scale of the pyramid diagram with every 1-s observation of the original time series. Nodes on the coarser scales can be considered as per-minute, per-hour, or even per-day features. The resulting finest-scale connection will form a C-ary tree. The finest scale connects neighboring nodes at each level, and coarse scales are better suited to describe long-term correlations.

Before introducing the principles of PAM, we first introduce the Transformer's self-attention mechanism.

$$\begin{aligned}
&\mathbf{X}\times\mathbb{W}^{\mathscr{Q}}=\mathscr{Q}\\
&\mathbf{X}\times\mathbb{W}^{\mathscr{K}}=\mathscr{K}\\
&\mathbf{X}\times\mathbb{W}^{\mathscr{V}}=\mathscr{V}\\
&\text{attention}(\mathscr{Q},\mathscr{K},\mathscr{V})=\text{softmax}\left(\frac{\mathscr{Q}\mathscr{K}^{T}}{\sqrt{d_k}}\right)\mathscr{V}
\end{aligned}\tag{13}$$

where $\mathbf{X}$ is the input time series data. $\mathscr{K}$, $\mathscr{Q}$, and $\mathscr{V}$ are Key, Query, and Value vectors, they are all from the same input $\mathbf{X}$. d_k is the dimension of $\mathscr{Q}$ and $\mathscr{K}$. $\mathbb{W}^{\mathscr{Q}}$, $\mathbb{W}^{\mathscr{K}}$ and $\mathbb{W}^{\mathscr{V}}\in\mathbb{R}^{L\times d_k}$ are randomly initialized matrix, where L is the input sequence length. As appropriate parameters can be learned during the back-propagation of the model, the time and space complexity of the self-attention mechanism is $O(L^2)$, which greatly increases the computational effort.

In PAM, in contrast to the Transformer self-attention mechanism, each node focuses on a limited set of keys, as shown in Fig. 5. Specifically, suppose $n_l^{(s)}$ represents the l node on the scale s, where $1,2,\cdots S$

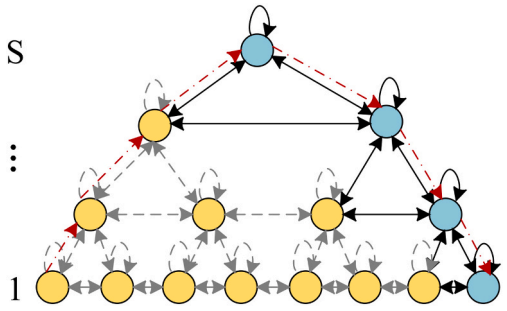


Fig. 5. PAM structure.

represent the scale number from the bottom to the top of the pyramid in turn, and S is the scales number. Each node in the figure can connect adjacent nodes of three scales which is represented as a set $\mathbb{N}_l^{(s)}$. $\mathbb{N}_l^{(s)}$ includes A adjacent nodes $\mathbb{A}_l^{(s)}$ at the same scale, C child nodes $\mathbb{C}_l^{(s)}$ and 1 parent node $\mathbb{P}_l^{(s)}$ [29].

$$\begin{cases}
\mathbb{N}_l^{(s)}=\mathbb{A}_l^{(s)}\cup\mathbb{C}_l^{(s)}\cup\mathbb{P}_l^{(s)}\\
\mathbb{A}_l^{(s)}=\left\{n_j^{(s)}:|j-l|\le\dfrac{A-1}{2},1\le j\le\dfrac{L}{c^{s-1}}\right\}\\
\mathbb{C}_l^{(s)}=\left\{n_j^{(s-1)}:l\le j\le lC\right\}\text{ if }s\ge 2\text{ else}\varnothing\\
\mathbb{P}_l^{(s)}=\left\{n_j^{(s+1)}:\left[\dfrac{l}{c}\right]\right\}\text{ if }s\le S-1\text{ else}\varnothing
\end{cases}\tag{14}$$

Therefore, for each node, the attention mechanism can be simplified to Eq. (15):

$$\text{attention}(\mathscr{Q},\mathscr{K},\mathscr{V})=\sum_{l\in\mathbb{N}_l^{(s)}}\frac{exp\left(\mathscr{Q}_i\mathscr{K}_l^{T}/\sqrt{d_k}\right)\mathscr{V}_i}{\sum\limits_{l\in\mathbb{N}_l^{(s)}}exp\left(\mathscr{Q}_i\mathscr{K}_l^{T}/\sqrt{d_k}\right)}\tag{15}$$

For a given A and L, the time and space complexity of the pyramid attention mechanism is $O(AL)$. When A is a constant, the time and space complexity are both $O(L)$, which significantly reduces computation time and memory costs.

In addition, the Add&Norm layer is added behind the PAM model, where Add stands for residual connection to prevent module degradation and Norm stands for standardization to normalize activation values for each layer.

The feedforward neural network layer is primarily a solution to the problem that the attention mechanism does not fit complex processes well, leading to model degradation.

3.4. Classification module

The role of the Classification module is to map the data features extracted by the previous module to the sample marker space. The detection classification is performed by softmax classifier, if the result is 0, then the time slot data is free from attack, if 1, then the time slot data suffers from FDIAs.

3.5. NAS algorithm

Since BESSs can be applied to distribution networks with different node systems, the depth characteristics of the measurement data will vary. However, it is difficult for the NPformer with a single network structure to learn the features of different data accurately, which will affect the detection results and efficiency of FDIAs for SoC estimation of BESSs in SDNs. In addition, the traditional manual design of the network structure is very complicated. Therefore, in the process of NPformer implementation, the NAS algorithm is introduced in this paper to

automatically design the network structure. The NAS algorithm will select the optimal one among all possible network structures and model parameters by comparison as the final detection model in the current data environment.

3.6. Practical application of NPformer in SDN with a BESS

The application of the NPformer detection model in an SDN with a BESS is shown in Fig. 6. Power management units send the collected data to the distribution network control center through phase data concentrators. The control center performs bad data detection and integrity testing on the returned data. Then FDIAs are performed on the data that pass the test using the NPformer model. If an attack is detected, the system issues a vulnerability discovery warning, and conversely performs distribution grid-side state estimation, SoC estimation of a BESS, and stores the data (the stored historical data is used to train and improve the NPformer detection model). Based on the state estimation results, the distribution grid system is responded to by the controller and the control commands are fed back to the remote terminal, so that the generator set and the BESS can be adjusted accordingly. Meanwhile, the state estimation and behavior of the BMS and controller are recorded in the historical operation of the system for auditing and log analysis. When FDIAs occur, warnings are sent to the controller to facilitate inference of current commands based on historical operations to maintain power system stability.

4. Case studies

4.1. Static FDIAs simulation experiment

This section performs FDIAs simulation experiments on SoC estimation of BESSs in SDNs according to the improved IEEE13 bus system [26] and the IEEE33 bus system [30]. Eq. (12) is solved by using Linprog in Matlab to derive the optimal attack. The topology of the IEEE13 and IEEE33 bus systems with BESSs are shown in Figs. 7 and 8. To achieve a fully observable system state, the number of PMUs needs to be between 1/5 and 1/3 of the total nodes [31]. Therefore, we add a certain amount of PMUs to each bus system. In Fig. 7, node 650 is selected as the relaxation bus, the BESS is connected to the 632 node of the IEEE13 bus system [24,30]. The number of PMUs of the entire bus system is 3 (which are located in nodes 632, 646, and 671, respectively). In Fig. 8, the BESS is connected to the 14 node of the IEEE33 bus system, and the number of PMUs of the entire bus system is 9, which are located in nodes

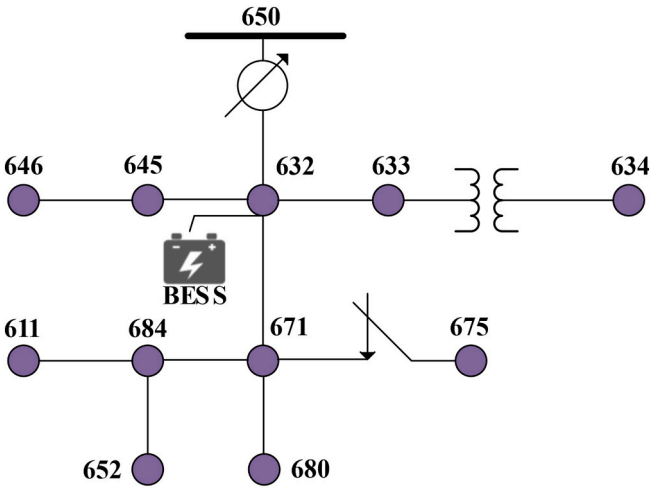


Fig. 7. IEEE13 bus system with a BESS.

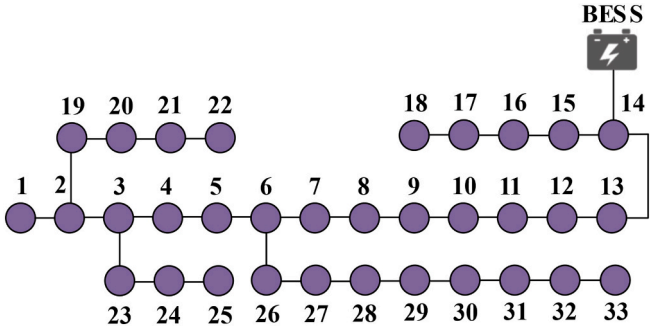


Fig. 8. IEEE33 bus system with a BESS.

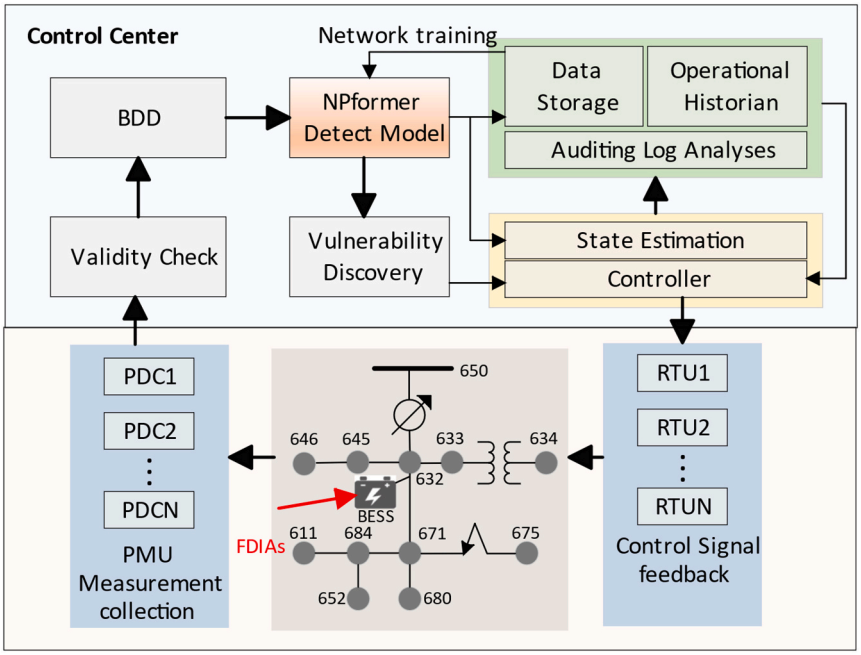


Fig. 6. Practical application of the NPformer detection model in SDN with a BESS.

2, 3, 4, 6, 14, 19, 23, 24, and 29, respectively [24,30]. In addition, to improve the accuracy of data containing distribution network status estimates, pseudo-measured values of reactive power injections and active power injections at all nodes are required.

Considering the measurement error of the actual system, the randomly generated measurement noise of the IEEE13 bus system is normally distributed with a mean value of zero and a standard deviation of 1 %. At a 95 % confidence level, the threshold τ is 4.5862. While the mean value of the randomly generated measurement noise of the IEEE33 bus system is determined to be zero with a standard deviation of 1 % using a normal distribution. At a 95 % confidence level, the threshold τ is 4.7521.

For BESSs connected to IEEE13 and IEEE33 bus systems, we build a lithium battery pack consisting of several batteries with a rated voltage of 4.1 connected in series, referring to [32]. The battery parameter settings are shown in Table 1. The active power setpoints of IEEE13 and IEEE33 bus systems with BESSs are shown in Figs. 9 and 10, and the values of reactive power setpoints are both zero. In addition, the initial SoC value of a BESS is set to 80 %, the lithium battery temperature is 25 °C, SoC operating range is 20 %–80 %, the battery resistance is 0.02404, and the slope of the OCV function that is used in the experiment is $\beta \approx 3.015\text{V}/1\%SOC$. Finally, we build an EKF-based SoC estimation model with Simulink and set $Q = 2.5 \times 10^{-7}, R = 1 \times 10^{-4}$.

Based on Eq. (12), we construct static FDIAs for SoC estimation at 142th minute and 133th minute for IEEE13 and IEEE33 bus systems with BESSs, respectively. The effects of such static FDIAs on the error in the SoC estimation of the BESSs are shown in Figs. 11 and 12 respectively. It can be seen that static FDIAs cause a large deviation in the state of charge estimation of the BESSs, with a difference of approximately 10 % between the estimated and actual SoC values at 142th minute in Fig. 11 and at 133th minute in Fig. 12. This large error will lead to excessive charge and discharge of BESSs, which will affect the safety, stability, and economical operation of the power grid. In addition, SoC estimation errors are still large for a while after the attack point, and SoC estimation error decreases gradually with the time affected by the Kalman gain of the EKF algorithm. The measurement residuals estimated using the WLS state in the case of static FDIAs on the IEEE13 and IEEE33 buses connected to the BESSs are shown in Figs. 13 and 14 respectively. We can see that the measurement residuals for the bus under attack are essentially the same as those for the bus before the point of attack and are always below the threshold for BDD. Therefore, bad data detection in the distribution network will have a hard time detecting FDIAs.

4.2. Static FDIAs detection based on NPformer

In order to validate the effectiveness of the proposed static FDIAs detection method, this section provides experimental validation of the NPformer detection method based on the quantitative data from the previous section. Since the proposed method is for the detection of FDIAs, the selection of experimental data should include both attacked and normal data. In this paper, we use the data obtained from the simulations of the IEEE13 bus from 109 to 193th minute and IEEE33 bus from 100 to 184th minute as the sample data (the training and test sets are divided 8:2), and the batch size is 60.

Each piece of data in the dataset has a corresponding label (1 for attack, 0 for normal). T_p indicates the number of attacked samples identified as having an attack. T_n indicates the number of normal samples identified as having no attack. F_n indicates the number of attacked samples identified as non-attacking. F_p indicates the number of normal samples identified as having an attack. y_i denotes the true label of sample i. p_i is the probability that the detection algorithm will predict sample i as having an attack. Therefore, the detection accuracy is shown in Eq. (16) and cross-entropy loss function is shown in Eq. (17).

Table 1
Battery pack ratings.

Parameter	Numerical
Nominal power and capacity	240 kW and 571.2kWh/816 Ah
Nominal DC current and voltage	350 A and 700 V
DC voltage range	582–780 V
DC current range	−390 – 390 A

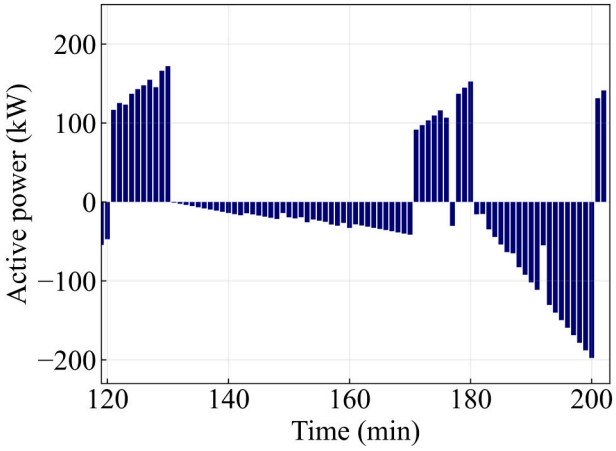


Fig. 9. Active power setpoints of IEEE13 bus systems with a BESS.

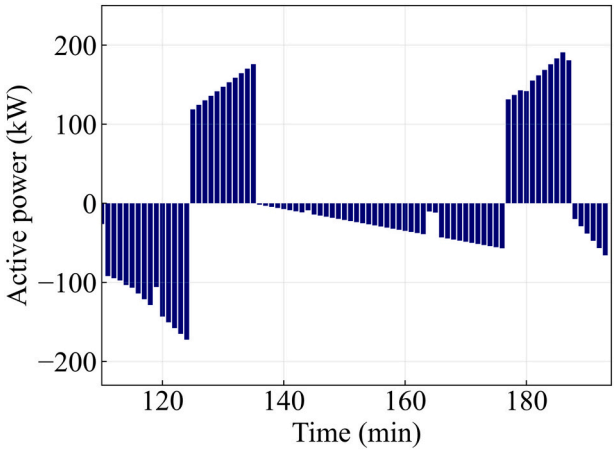


Fig. 10. Active power setpoints of IEEE33 bus systems with a BESS.

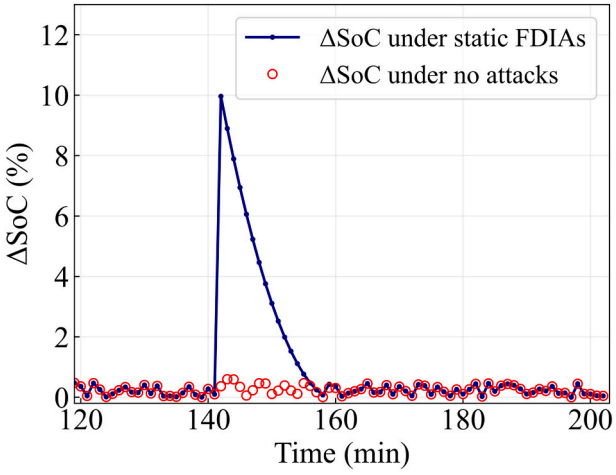


Fig. 11. FDIAs against SoC estimation of IEEE13 bus system.

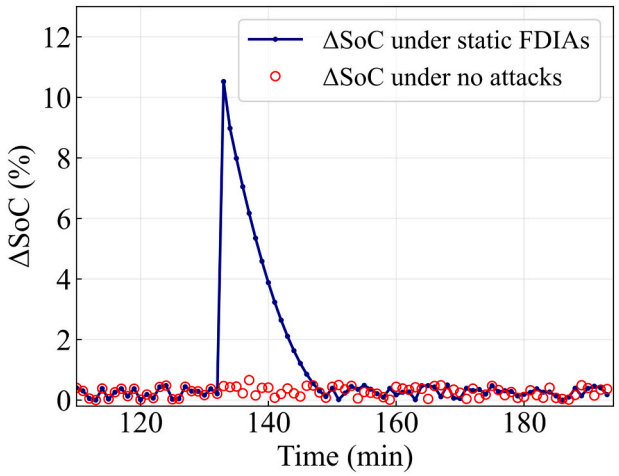


Fig. 12. FDIAs against SoC estimation of IEEE33 bus system.

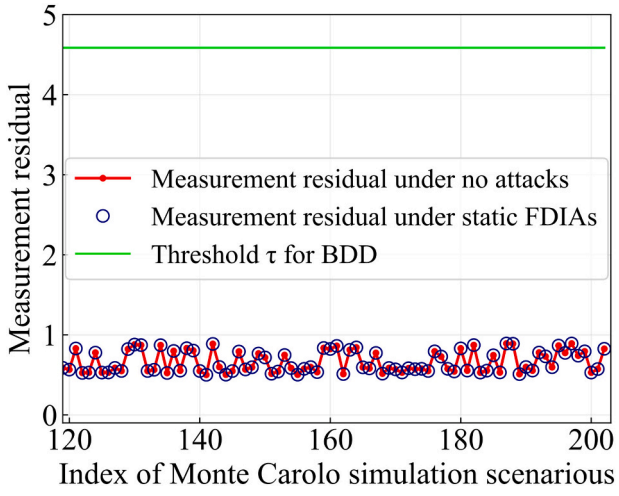


Fig. 13. Measurement residuals for IEEE13 bus system at static FDIAs.

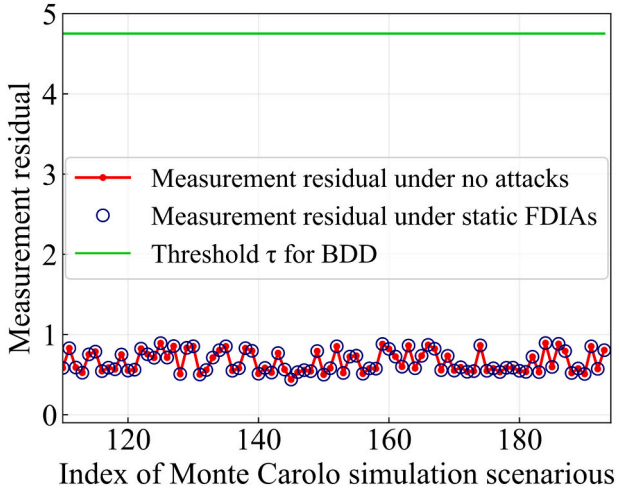


Fig. 14. Measurement residuals for IEEE33 bus system at static FDIAs.

$$Accuracy = \frac{T_n + T_p}{T_n + F_n + T_p + F_p} \tag{16}$$

$$Loss = \frac{1}{N}\sum_{i=1}^{N} -[y_i log(p_i) + (1-y_i) log(1-p_i)] \tag{17}$$

1) Comparison experiments for accuracy

Fig. 15 shows the variation of detection accuracy in the IEEE13 and IEEE33 bus systems respectively. It can be seen that the accuracy of the algorithm gradually improves and eventually stabilizes throughout the iterations. The accuracy of algorithm training under static FDIAs in the IEEE13 bus system is about 97.50 %, and the test accuracy is about 96.66 %. The accuracy of algorithm training under static FDIAs in the IEEE33 bus system is about 98.01 %, and the accuracy of testing is about 97.10 %. Experimental results show that the NPformer method can detect static FDIAs in different bus systems and has high accuracy. In addition, it can be noticed that the test results for the IEEE13 bus system are slightly lower than those for the IEEE33 bus system. This is because the SoC errors generated by an attack on the IEEE33 bus system will be greater compared to the IEEE13 bus system, allowing the NPformer method to better learn and extract multi-scale features of the data. As a result, the FDIAs of the IEEE33 bus system are detected with relatively high accuracy.

2) Comparison experiments for Loss

Fig. 16 shows the changes in NPformer method losses in IEEE13 and IEEE33 bus systems. It can be seen that the loss of the algorithm gradually decreases and eventually stabilizes throughout the iterations. In the IEEE13 bus system, the algorithm training loss under static FDIAs is about 0.0377, and the test loss is about 0.03867. In the IEEE33 bus system, the algorithm training loss under static FDIAs is about 0.0366, and the test loss is about 0.0371. Experimental results show that the NPformer method can accurately identify the static FDIAs of different bus systems.

4.3. Static FDIAs detection performance comparison experiment of different algorithms

In order to further verify the validity and reliability of the proposed method, this section compares the NPformer method with the Transformer and LSTM autoencoder algorithms using measured data from static FDIAs simulation experiments as a dataset. Figs. 17 and 18 show

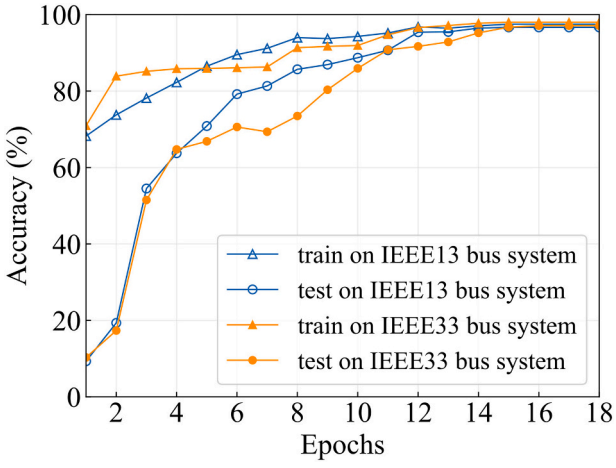


Fig. 15. NPformer detection accuracy under static FDIAs of different bus systems.

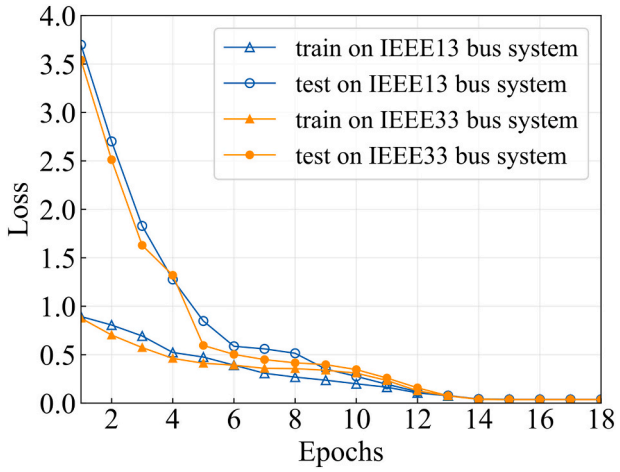


Fig. 16. NPformer detection loss under static FDIAs of different bus systems.

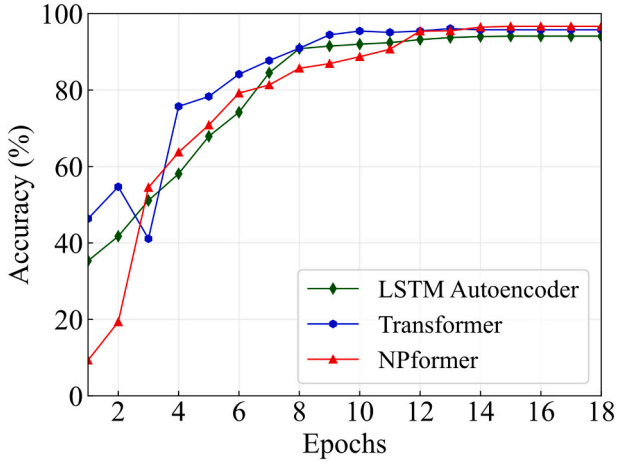


Fig. 17. Detection accuracy of different algorithms under static FDIAs of IEEE13 bus system.

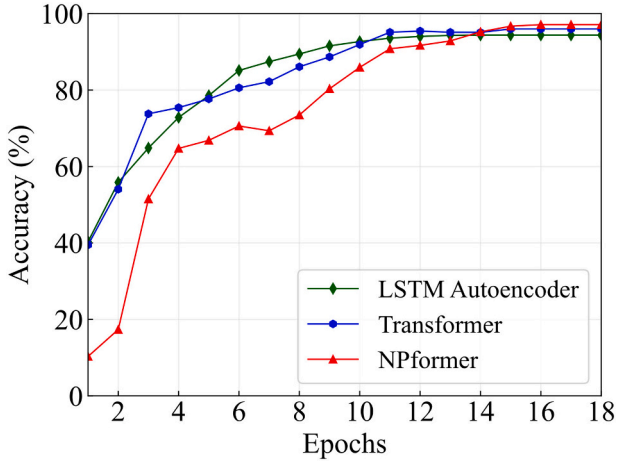


Fig. 18. Detection accuracy of different algorithms under static FDIAs of IEEE33 bus system.

the accuracy of each algorithm tested under static FDIAs in the IEEE13 and IEEE33 bus systems. It can be seen that the accuracy of the NPformer method is higher in all the different bus systems. This is mainly since to the fact that the NPformer method can learn multi-scale information of the battery sequence data and extract more attack features, which results in better detection performance of the NPformer method.

5. Conclusion

FDIAs on the SoC estimation of BESSs in SDNs can threaten the security, stability, and economical operation of the distribution network. This paper describes the principles and construction methods of static FDIAs in smart distribution networks with battery energy storage systems, and conducts static FDIAs simulation experiments using simulation software. Next, we propose the use of NAS to implement an improved automatic design of Transformer network structure and parameters, and generate an NPformer detection model. Then, comparison experiments are conducted under different bus systems and different algorithms. The experimental results show that the NPformer detection method proposed in this paper has good detection effectiveness, high detection accuracy (nearly 97 %), and avoids the complicated calculation process when setting the threshold value. Future work will focus on the dynamic inspection of FDIAs for battery energy storage systems.

CRediT authorship contribution statement

Zhiying Liu: Writing – review & editing, Writing – original draft, Validation, Software, Methodology, Formal analysis, Conceptualization. **Yuancheng Li:** Writing – original draft, Project administration, Data curation, Conceptualization. **Shuang Xu:** Writing – original draft, Validation, Resources, Formal analysis, Data curation. **Qingle Wang:** Writing – review & editing, Visualization, Project administration, Conceptualization. **Jianbin Li:** Supervision, Project administration, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The authors are unable or have chosen not to specify which data has been used.

Acknowledgments

This work was supported by the State Grid Corporation Science and Technology Project “Research on Key Technologies to Support Network Operation of Distributed Energy Storage” (Grant No. 5100-202199544A-0-5-ZN).

References

[1] S.G. Varzaneh, et al., Optimal energy management for PV integrated residential systems including energy storage system, IET renewable power generation 15 (1) (2021) 17–29.
[2] P. He, et al., An adaptive VSG control strategy of battery energy storage system for power system frequency stability enhancement, International Journal of Electrical Power Energy Systems 294 (2021) 117022.
[3] T. Kim, et al., Cloud-based battery condition monitoring and fault diagnosis platform for large-scale Lithium-ion battery energy storage systems, Energies 11 (1) (2018) 125.
[4] K. Taesic, et al., An overview of cyber-physical security of battery management systems and adoption of Blockchain technology, IEEE J EM SEL TOP P 10 (2020) 1270–1281.
[5] N. Farshid, et al., Cyber-physical cloud battery management systems: review of security aspects, Batteries 9 (7) (2023) 382.
[6] P. Zhuang, H. Liang, False data injection attacks against state-of-charge estimation of battery energy storage Systems in Smart Distribution Networks, IEEE Trans. Smart Grid 12 (3) (2021) 2566–2577.
[7] N. Mhaisen, et al., Secure smart contract-enabled control of battery energy storage systems against cyber-attacks, Alex. Eng. J. 58 (4) (2019) 1291–1300.

[8] Y. Li, et al., Online generative adversary network based measurement recovery in false data injection attacks: a cyber-physical approach, IEEE Trans. Ind. Informat. 16 (3) (2020) 2031–2043.
[9] Z. Liu, et al., Sensor fault detection and isolation for a lithiumion battery pack in electric vehicles using adaptive extended Kalman filter, Appl. Energy 185 (2017) 2033–2044.
[10] Y. Mo, et al., On the performance degradation of cyber-physical systems under stealthy integrity attacks, IEEE Trans Automat Contr 61 (9) (2016) 2618–2624.
[11] V. Obrien et al., "Detecting False Data Injection Attacks to Battery State Estimation Using Cumulative Sum Algorithm," in NAPS, Texas, USA, 2021, pp. 1–6.
[12] V. Obrien et al., "Detection of False Data Injection Attacks in Battery Stacks Using Physics-Based Modeling and Cumulative Sum Algorithm," in PECI, Illinois, USA, 2022, pp. 1–8.
[13] Z. Liu, et al., Moving target defense of FDIAs for battery energy storage systems in smart distribution networks, J. Energy Storage 72, Part D (2023).
[14] Z. Liu, et al., TSCW-GAN based FDIAs defense for state-of-charge estimation of battery energy storage Systems in Smart Distribution Networks, IEEE Trans. Ind. Informat. 20 (4) (2024) 5048–5059.
[15] T. Kousksou, et al., Energy storage: applications and challenges, Sol. Energy Mater. Sol. Cells 120 (2014) 59–80.
[16] J. Baker, et al., New technology and possible advances in energy storage, Energy Policy 36 (12) (2008) 4368–4373.
[17] O. P. Jaga et al., "Seamless Transition between Grid-Connected and Islanded Operation Modes for Hybrid PV-BESS Combination used in Single-Phase, Critical Load Applications," in SEFET, Hyderabad, India, 2021, pp. 1–6.
[18] A.T. Elsayed, et al., Advanced battery management and diagnostic system for smart grid infrastructure, IEEE Trans. Smart Grid 7 (2) (2016) 897–905.
[19] C. Lyu, et al., An Electrochemical Thermal Coupling Model for High C-rate Conditions in Lithium-ion Batteries, in: ICIEA, Kristiansand, Norway, 2020, pp. 1733–1737.
[20] D. Rosewater, et al., Battery energy storage state-of-charge forecasting: models, optimization, and accuracy, IEEE Trans. Smart Grid 10 (3) (2019) 2453–2462.
[21] A. de la Villa Jaen, et al., Voltage source converter modeling for power system state estimation: STATCOM and VSC-HVDC, IEEE Trans. Power Syst. 23 (4) (2008) 1552–1559.
[22] S. Chen, et al., Multi-area distributed three-phase state estimation for unbalanced active distribution networks, J. Mod. Power Syst. Clean Energy 5 (5) (2017) 767–776.
[23] R. Xiong, et al., Adaptive state of charge estimator for lithium-ion cells series battery pack in electric vehicles, J. Power Sources 242 (2013) 699–713.
[24] M. Cacciato, et al., Real-time model-based estimation of SOC and SOH for energy storage systems, IEEE Trans. Power Electron. 32 (1) (2017) 794–803.
[25] S. Lee, et al., State-of-charge and capacity estimation of lithium-ion battery using a new open-circuit voltage versus state-of-charge, J. Power Sources 185 (2) (2008) 1367–1373.
[26] P. Zhuang, et al., False data injection attacks against state estimation in multiphase and unbalanced smart distribution systems, IEEE Trans. Smart Grid 10 (6) (2019) 6000–6013.
[27] S. Zhao and F. You, "A Topical Keywords Fusion Based on Transformer For Text Summarization," in ICICTA, Chongqing, China, 2020, pp. 294–297.
[28] M. Chen et al., "AutoFormer: Searching Transformers for Visual Recognition," in ICCV, Montreal, Quebec, Canada, 2021, pp. 12250–12260.
[29] C. Li, et al., DS-net++: dynamic weight slicing for efficient inference in CNNs and vision transformers, IEEE Trans. Pattern Anal. Machine Intell. (2022), https://doi.org/10.1109/TPAMI.2022.3194044.
[30] S. Song, et al., A holistic state estimation framework for active distribution network with battery energy storage system, Journal of Modern Power Systems and Clean Energy 10 (3) (2021) 627–636.
[31] T.L. Baldwin, et al., Power system observability with minimal phasor measurement placement, IEEE Trans. Power Syst. 8 (2) (1993) 707–715.
[32] G. Rancilio, et al., Modeling a large-scale battery energy storage system for power grid application analysis, Energies 12 (17) (2019) 3312.