

Data-driven Transient Stability Assessment Using Sparse PMU Sampling and Online Self-check Function

Guozheng Wang, *Member, IEEE*, Jianbo Guo, *Senior Member, IEEE, Fellow, CSEE*, Shicong Ma, *Member, IEEE, Senior Member, CSEE*, Xi Zhang, Qinglai Guo, *Senior Member, IEEE*, Shixiong Fan, and Haotian Xu

Abstract—Artificial intelligence technologies provide a new approach for the real-time transient stability assessment (TSA) of large-scale power systems. In this paper, we propose a data-driven transient stability assessment model (DTSA) that combines different AI algorithms. A pre-AI based on the time-delay neural network is designed to locate the dominant buses for installing the phase measurement units (PMUs) and reducing the data dimension. A post-AI is designed based on the bidirectional long-short-term memory network to generate an accurate TSA with sparse PUM sampling. An online self-check function of the online TSA's validity when the power system changes is further added by comparing the results of the pre-AI and the post-AI. The IEEE 39-bus system and the 300-bus AC/DC hybrid system established by referring to China's existing power system are adopted to verify the proposed method. Results indicate that the proposed method can effectively reduce the computation costs with ensured TSA accuracy as well as provide feedback for its applicability. The DTSA provides new insights for properly integrating varied AI algorithms to solve practical problems in modern power systems.

Index Terms—Artificial intelligence, phasor measurement units, recurrent neural networks, transient stability assessment.

NOMENCLATURE

ANN	Artificial neural network.
RRART	Relative reduction of online average response time.
b	Bias of neural network layers.
BiLSTM	Bidirectional long-short-term memory network.
C	Correlation coefficient (between η_{post} and η').

DBS	Dominant bus set.
DTSA	Data-driven TSA model.
LSTM	Long-short-term memory network.
n	Number of DBS.
PMU	Phasor measurement unit.
RNN	Recurrent neural network.
TSA-SPS	Transient stability assessment with sparse PMU sampling.
T	Length of the time series.
TDNN	Time-delay neural network.
TSA	Transient stability assessment.
$\text{Var}[X]$	Variance of X .
w	Parameter matrix of neural network.
x_t^i	Input vector of bus i at time t .
y_{Label}	Case label.
$y_{\text{pre}}/y_{\text{post}}$	Assessment of pre-AI/post-AI.
ξ_{Δ}	Relative time reduction of training.
$\Delta\delta_{ax}$	Maximum power angle difference of generators.
$\eta_{\text{pre}}/\eta_{\text{post}}$	Accuracy of pre-AI/post-AI.
η'	Correlation between pre-AI and post-AI predictions.

I. INTRODUCTION

THE transient stability is critical for ensuring the safe and stable operation of power systems, whose destruction can cause large-scale power blackouts [1], [2]. In recent years, the increasing penetration of power electronic devices and the expansion of the operating state-space creates great challenges for transient stability analysis and assessment [3].

Traditional transient stability assessment (TSA) methods can be sorted into two categories: direct methods [4] and time-domain simulations [5]. However, the speed and accuracy of the traditional methods cannot be both guaranteed for large-scale power systems. Artificial intelligence (AI) provides a new method for TSA [6], [7], which has the potential to maintain fast calculation speed and high assessment accuracy.

Some intelligent techniques, such as support vector machines (SVMs) [8], artificial neural networks (ANN) [9], and decision trees (DTs) [10], have been employed for TSA. Nevertheless, in general the inputs of the above methods are static variables, which makes it difficult to analyze the dynamic process after faults. With the widespread adoption

Manuscript received August 11, 2021; revised November 20, 2021, accepted December 28, 2021. Date of online publication January 14, 2022, date of current version March 24, 2022. This work was supported by the National Key R&D Program of China (2018AAA0101500).

G. Z. Wang (corresponding author, email: guozhengwang@foxmail.com) is with the Department of Electrical Engineering, Tsinghua University, Beijing 100084, China, and the Department of Power System, China Electric Power Research Institute, Beijing 100192, China.

J. B. Guo, S. C. Ma, S. X. Fan and H. T. Xu are with the Department of Power System, China Electric Power Research Institute, Beijing 100192, China.

X. Zhang is with the School of Automation, Beijing Institute of Technology, Beijing 100192, China.

Q. L. Guo is with the Department of Electrical Engineering, Tsinghua University, Beijing 100084, China.

DOI: 10.17775/CSEEJPES.2021.05890

of phasor measurement units (PMUs), the acquisition of the time series of the system states becomes more convenient. The recurrent neural networks (RNN) have advantages in processing time series [11], [12]. In [13], the time series of voltage phasors are directly used as the input of a long-short-term memory network (LSTM) for TSA. In [14], the ensemble learning mechanism based on LSTM is utilized to consider the PMU data missing for TSA.

The increasing complexity of modern power systems demands excessive computation that limits the application of intelligent algorithms [15], especially for LSTM networks. Considering the existing power system, PMUs monitor the operational status of the corresponding locations. However, installing a PMU in each bus of the power system is economically costly and also increases the burn for data transmission and processes [16]. Therefore, selecting the data of the key bus set in applying the intelligent model to TSA can not only reduce the installation cost of the PMU but also reduce the data dimension and improve training efficiency. In [17], a deep autoencoder network is used to build a framework for feature extraction, but the interpretability of the extraction process is unclear. To keep the meaning of features, a random forest is also used for selecting the key features [10]. Although these extraction methods can reduce the dimension of the initial data, the selection of PMU installation locations are not taken into consideration. Regarding selecting key buses, heuristic and rule-based methods have been investigated. For example, all buses are selected directly in [13] and [19], while only the generator buses are selected in [9] and [20]. In [14], the rule-based method (observability description algorithm [21]) is used to select PMU locations. However, whether these selection methods based on human experience/rule-based are suitable for data-driven methods remains to be verified, especially for black-box models, such as neural networks [9], [13], [14].

In addition, the trained intelligence models can maintain good performance on a data set with a similar distribution to the training set [22]. If the distribution of input data changes, the accuracy of the trained model will decrease. In recent years, with the increase of renewable energy, the randomness and openness of power systems have become the major features, which leads to different distribution between the input data and training set when the intelligent models are applied online [23]. Therefore, the intelligence model should have a self-check function to indicate the validity after being applied online. Nevertheless, the self-check function has rarely been discussed in the research on the artificial intelligence of power systems.

Therefore, designing an intelligent TSA method with sparse PMU samplings in the power system to reduce the data dimension as well as maintain the accuracy, and improve the applicability of the AI methods in the actual power system are urgent tasks.

In this paper, we propose a data-driven transient stability assessment model (DTSA) that combines different AI algorithms: A pre-AI is designed by time-delay neural networks to locate the dominant buses and reduce the data dimension, and a post-AI is designed based on bidirectional long-short-term

memory networks to generate an accurate TSA with sparse PUM sampling. In addition, an online self-check function of the online TSA's validity is further established by comparing the results of the pre-AI and the post-AI. DTSA provides a reference for properly combining multiple AI algorithms to solve practical problems in power systems.

The remainder of this paper is organized as follows: Section II introduces the framework and work process of DTSA. Section III elaborates on the mathematical models of the pre-AI and post-AI. In Section IV, the IEEE 39-bus system and the 300-bus AC/DC hybrid system are adopted to verify the proposed DTSA. Finally, we conclude this study in Section V.

II. DATA-DRIVEN TRANSIENT STABILITY ASSESSMENT MODEL

A. Transient Stability Assessment

Power systems are high-dimensional nonlinear autonomous systems. The transient stability can be determined according to whether the post-fault state is in the attraction region of the stable equilibrium point $A(x_s)$. $A(x_s)$ is shown as follows:

$$A(x_s) = \{x \in \mathbb{R}^n : \lim_{t \rightarrow \infty} \phi(x, t) = x_s\} \quad (1)$$

where x is an n -dimensional state variable, x_s is a stable equilibrium point. The strict boundary of $A(x_s)$ is difficult to obtain. Generally, the transient energy at relevant unstable equilibrium points can be used to approximate the boundary, that is, the critical energy V_{cr} in transient energy function methods. If the post-fault energy is less than V_{cr} , the system can be considered stable. However, calculating the critical energy V_{cr} is also a difficult challenge in high-dimensional systems.

In practical applications, time domain simulation is generally used to calculate the transient stability of complex systems. Fig. 1 shows the dynamic process of the rotor angle of a fault in the IEEE 39-bus system, where the line 17–18 suffered a three-phase short circuit, and the fault starting time and clearing time are 0.1 s and 0.3 s respectively. If the system will lose its synchronism, we need to make timely judgments to take emergency control. However, this is difficult to do with the time domain simulation.

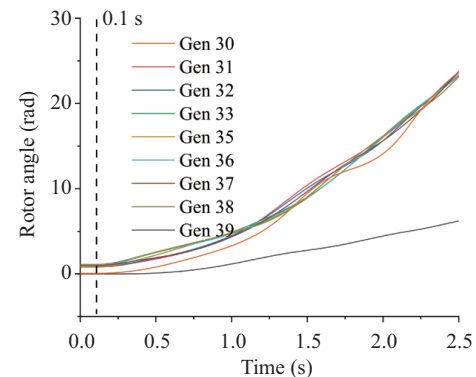


Fig. 1. Dynamic process of rotor angle after fault.

The widespread adoption of PMUs provides a data foundation for fast TSA based on data-driven methods. In this

work, according to the time series collected by the PMUs (as shown in Fig. 1), we propose a data-driven transient stability assessment method with sparse PMU sampling, which can effectively reduce the computation cost with ensured TSA accuracy as well as give feedback for its applicability.

B. Framework of DTSA

According to [24], it is a good way to combine multiple algorithms to solve complex tasks. In this part, a data-driven transient stability assessment model, as shown in Fig. 2, is presented to assess the transient stability with sparse PMU sampling and implement an online self-check of power systems.

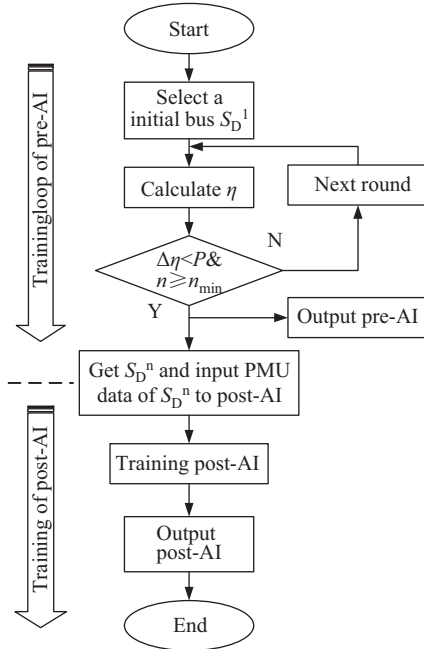


Fig. 2. The Schematic of the proposed DTSA.

In Fig. 2, the total dataset is the state variables of each bus, where M is the total number of buses in the studied power system. To make the extracted features correspond to the PMU locations, a recursive process is designed to select the dominant bus set (DBS) based on the accuracy of the trained pre-AI. Then, the state variables of the DBS are input into the post-AI to generate the final and accurate TSA

results. Since the pre-AI can determine the DBS, the transient stability assessment with sparse PMU sampling (TSA-SPS) is implemented, which can reduce the number of PMUs and relieve the burden of communication systems [16].

In the online application, the proposed DTSA can verify its validity by comparing the correlation between the predictions of pre-AI and post-AI (the detailed implementation process of this function is elaborated in Section II.C). If the correlation does not meet the requirements, it means that the trained DTSA is no longer applicable to the new environment (e.g., a major change in the operational mode or topology of the systems). The DTSA should send a signal to retrain the intelligent model, as shown in the lower part of Fig. 2.

C. TSA-SPS of DTSA

For a large power system, if the state variables of all buses are used for training DTSA, it will consume a lot of computing power, and the cost of the PMU installation will increase. Therefore, it is critically important to reduce the input data dimensionality from the perspective of bus location.

In DTSA, the pre-AI is designed to locate the DBS of the power systems, based on which post-AI generates an accurate TSA. The training flowchart of the DTSA is shown in Fig. 3.

The determination of DBS is done in the offline training loop of pre-AI. First, the pre-AI randomly selects a bus as the initial DBS S_D^1 , and the state variables of S_D^1 are used to train the pre-AI. In the training process, the accuracy of pre-AI η_{pre} is described by:

$$\eta_{pre} = n_{actual_label} / N \quad (2)$$

where N is the total number of cases in the test set, n_{actual_label} is the number of cases in which the predictions and labels are the same.

Then, the bus with maximum η_{pre} is retained as S_D^1 after all buses are traversed. In the next round of selection, the other best bus is retained until S_D^n satisfies the following:

$$\begin{cases} \Delta\eta_{pre}^{n_{min}} = \eta_{pre}^{n_{min}+3} - \eta_{pre}^{n_{min}} \leq P \\ n \geq n_{min} \end{cases} \quad (3)$$

where $\Delta\eta_{pre}$ indicates that the accuracy of pre-AI is improved less than P in the next three rounds, P is a constant with a value of 0.5% in this paper, n_{min} is the minimum number of the DBS. To ensure the convergence of DTSA, the number of DBS n is usually greater than n_{min} .

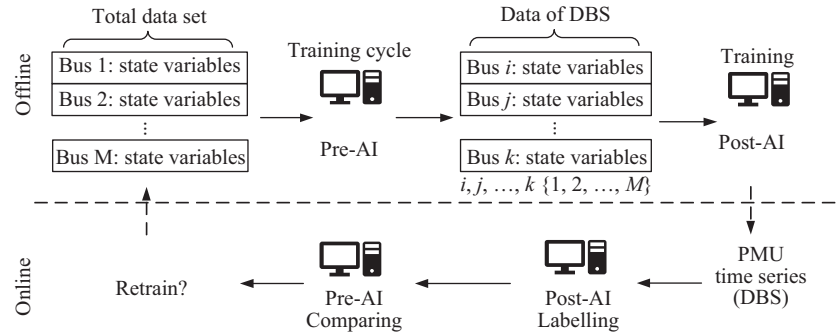


Fig. 3. The training flowchart of DTSA.

Eq. (3) provides the termination condition for the pre-AI training loop. Subsequently, the state variables of S_D^n are inputted into the post-AI for training. It should be noted that a trained pre-AI model is also obtained after the training of pre-AI, which will be used for the online self-check function (see Section II.C).

In general, the size of S_D^n is much smaller compared with the whole grid. Thus, the determined S_D^n by pre-AI can effectively reduce the input feature dimension of post-AI to improve the training efficiency. Moreover, S_D^n also simplifies the number of PMU, which can save the installation costs of PMU and relieve the burden of communication systems.

D. Online Self-check Function of DTSA

After a trained AI model is deployed online, it is difficult to evaluate the validity of the model in real-time because labeling the cases will take a lot of time. Nevertheless, the proposed DTSA can verify the validity of the online TSA (that is, the online self-check function) by comparing the correlation between the predictions of pre-AI and post-AI. This function can help operators to judge the intelligent model's validity and decide whether to retrain.

The online self-check process is shown in Fig. 4. First, we need to prepare test cases for the online self-check function. The cases can be obtained from the existing power systems or simulations. Then the predictions of post-AI are used as the labels of the cases. These labeled cases are used as a test set for pre-AI to calculate the relative accuracy η' (that is, the correlation between the predictions of pre-AI and post-AI). The calculation of the correlation is described by:

$$\eta' = n_{\text{post-AI_label}} / N_{\text{online}} \quad (4)$$

where N_{online} is the total number of the test set, $n_{\text{post-AI_label}}$ is the number of cases in which the pre-AI and post-AI have the same predictions. In general, the trained DTSA has a relatively stable value of η' when the grid environment has not significantly changed (more precisely, the real input data has a similar distribution to the training set). The value of η' will decrease as the distribution of the input data deviates from the training set. If η' is less than a predetermined value $\eta'_{\min}$, DTSA will give a warning signal to retrain the intelligence model, where $\eta'_{\min}$ can be determined according to the specific requirements of the operators and the offline training results (see Section IV for more details).

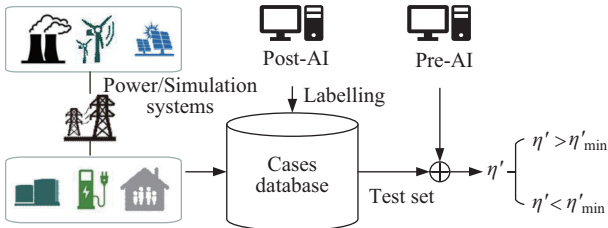


Fig. 4. Online self-check process of DTSA.

Another important thing to notice is that: If the test cases obtained from the existing power system are not enough for online self-check, we can supplement the cases by simulations.

Because the case labels are generated by post-AI predictions instead of the real transient stability label, we only need to simulate the first few cycles, which ensures real-time online self-checking.

III. MATHEMATICAL MODELS OF PRE-AI AND POST-AI

This section introduces the mathematical models of pre-AI and post-AI. The pre-AI and the post-AI are established by time-delay neural networks (TDNN) and bidirectional long-short-term memory networks (BiLSTM), respectively.

A. The Mathematical Model of Pre-AI

The following two rules should be followed in designing the pre-AI to ensure the cooperation of the pre-AI and post-AI.

- 1) The pre-AI and the post-AI should have the same input-output structure to ensure their cooperation.
- 2) The pre-AI should have a relatively simple structure of the neural networks to ensure computation efficiency.

Because TDNNs are more suitable for processing time series compared with fully connected ANNs and can share weight matrixes w in the time dimension [25], TDNN is selected to construct the pre-AI. The structure of TDNN is shown in Fig. 5, where y_{pre} is the output of TDNN at time t , the input is the time series x_{t-T}^i to x_t^i , x is a vector (that is, the state variables of buses), and the superscripts i, j, k are the bus indices of the power systems. The pre-AI constructed by TDNN can extract the time dimension features for the post-AI, which are more suitable for describing the post-fault dynamic process.

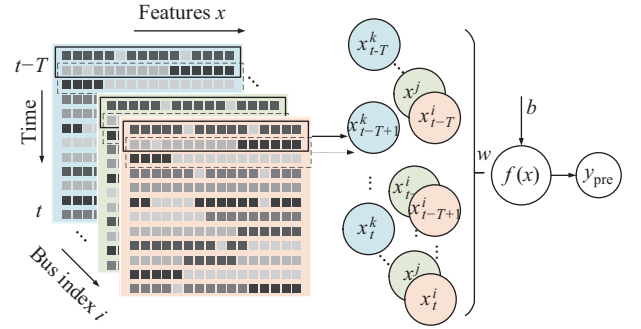


Fig. 5. The structure of the time-delay neural network.

B. The Mathematical Model of Post-AI

The post-AI should have strong feature extraction capabilities for time series to generate an accurate TSA with the sparse PUM sampling (that is, the extracted DBS by pre-AI). Therefore, the BiLSTM is adopted to design the post-AI.

A time series usually has two directions, that is, the forward and backward directions. However, the traditional RNN can only obtain one direction's information. To obtain the information in both two directions, a bidirectional RNN is adopted, as shown in Fig. 6, including the forward and backward layers. The output y_t is calculated by the following:

$$\begin{cases} h_{\rightarrow} = \sigma(w_1 x_t + w_3 h_{t-1} + b_{\rightarrow}) \\ h_{\leftarrow} = \sigma(w_2 x_t + w_5 h_{t+1} + b_{\leftarrow}) \\ y_t = w_4 h_{\rightarrow} + w_6 h_{\leftarrow} + b_y \end{cases} \quad (5)$$

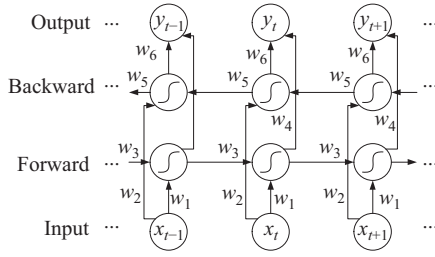


Fig. 6. The structure of bidirectional recurrent neural network.

where $h_{\rightarrow}$ and $h_{\leftarrow}$ are the output of the forward and backward layer respectively, $b_{\rightarrow}$ and $b_{\leftarrow}$ are the bias on the forward and backward layer respectively, b_y is the bias on the final output, x_t and y_t are the input and output of the bidirectional RNN at time t respectively. Compared with traditional neural cells, the LSTM cell prolongs the memory and makes up the gradient vanishing during the training process. The BiLSTM layer is established by substituting the LSTM cells for traditional cells.

As shown in Fig. 7, the post-AI network is established by using the BiLSTM layer. A dropout layer is added following the BiLSTM layer to prevent overfitting during the training process, and the probability of the dropout layer is 0.5. The BiLSTM layer contains 128 hidden units, and the output size of the fully connected layer is 2. At the end of the neural network, a softmax layer is adopted to facilitate the final classification of TSA.

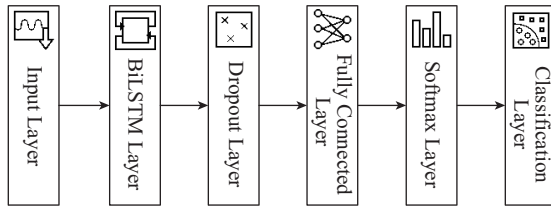


Fig. 7. The structure of post-AI neural network.

IV. CASE STUDIES

The DTSA is applied to the IEEE39-bus system in a larger power system to evaluate the computational efficiency and accuracy. The tests are conducted on a computer with an Intel Core i5 CPU (3.4 GHz), 8 G of memory. The power system department edit (PSDEdit), released by the China Electric Power Research Institute, is used for simulation.

A. Cases Generation

In the existing power system, it is difficult to collect the state variables of all buses. In addition, the actual PMU data is highly imbalanced and concentrates in a stable state. The actual operational experiences tell us that we can use the simulation system to guide the operation of the existing system (e.g., the annual calculation of the operational mode). Therefore, the simulation system is used for DTSA training. This also ensures that the DBS can be adjusted, in time, after the grid topology changes. To obtain the time series (PMU data), the sampling frequency is 50 Hz in the simulation. The voltage phasors are adopted as the state variables of the

bus [10], [26]. The format of the collected initial data is shown in (6), including voltage amplitude V and phase angle θ .

$$\begin{bmatrix} V_{1,1} & V_{1,2} & \cdots & V_{1,T} \\ V_{2,1} & V_{2,2} & \cdots & V_{2,T} \\ \vdots & \vdots & \ddots & \vdots \\ V_{M,1} & V_{M,2} & \cdots & V_{M,T} \end{bmatrix} \begin{bmatrix} \theta_{1,1} & \theta_{1,2} & \cdots & \theta_{1,T} \\ \theta_{2,1} & \theta_{2,2} & \cdots & \theta_{2,T} \\ \vdots & \vdots & \ddots & \vdots \\ \theta_{M,1} & \theta_{M,2} & \cdots & \theta_{M,T} \end{bmatrix} \quad (6)$$

where M is the total number of buses in the studied power system, and T is the length of the time series. The initial input feature dimension is $2M$ for a power system with M buses. Based on the study in [13], we take $T = 10$ to prevent the accuracy of the prediction results from being affected by the small value of T .

The case labels y_{Label} are determined by the maximum difference of the simulation rotor angle $\Delta\delta_{ax}$. The determination is described by (7) and (8).

$$TSI = \frac{360^\circ - |\Delta\delta_{\max}|}{360^\circ + |\Delta\delta_{\max}|} \quad (7)$$

$$y_{\text{Label}} = \begin{cases} 0 \text{ stable,} & TSI > 0 \\ 1 \text{ unstable,} & TSI < 0 \end{cases} \quad (8)$$

Since the TSA is a dual classification task, the cross-entropy error function L is used for training the pre-AI and post-AI, as shown in (9). During the training process, the adaptive moment estimation (ADAM) optimization algorithm is used to optimize the DTSA parameters by minimizing the loss function L . The learning rate is reduced by 0.4 times every 60 iterations, and the mini-branch size is 512.

$$L = - \sum_{n=1}^N (y_{\text{Label}}^n \log y_{\text{Pred}}^n + (1 - y_{\text{Label}}^n) \log(1 - y_{\text{Pred}}^n)) \quad (9)$$

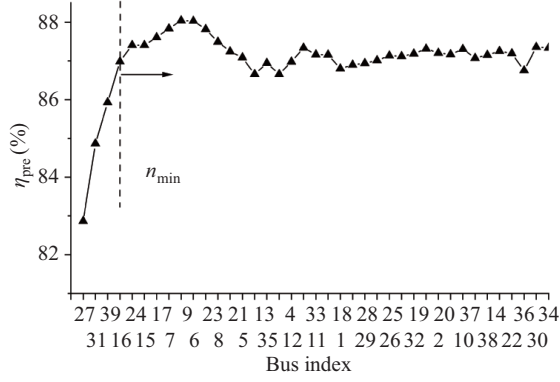
where N is the number of training cases, and y_{Pred}^n is the prediction of the pre-AI or post-AI, namely y_{pre} and y_{post} .

B. Case Studies of the IEEE 39-Bus System

In the IEEE 39-bus system, both generations and loads are scaled by the ratio that is from 80% to 120%, and the $(N - 1)$ contingencies are taken into account. The fault clearing time is set to 0–0.4 s with an increment of 0.05 s. Consequently, a total of 9,000 TSA cases are generated.

1) DBS Selection of Pre-AI

In the DBS selection of pre-AI, it is not necessary to conduct training for the whole training set, which can further accelerate the speed of DBS selection. Therefore, 1,000 cases are selected randomly to train the pre-AI, where 85% is the training set, and the remaining 15% is the testing set. Fig. 8 shows the relationship between the accuracy η_{pre} and the DBS (notice that the horizontal axis is the order in which buses are joined to the DBS). The number of DBS n should be greater than four according to (3). Fig. 8 shows that η_{pre} increases rapidly with the first four buses added to DBS. Subsequently, η_{pre} grows slowly as n increases, and η_{pre} decreases slightly when n is greater than 10 because the scale of the pre-AI limits the ability to process the high-dimensional data, this phenomenon

Fig. 8. η_{pre} of offline training in IEEE 39-bus system (test set).

is also mentioned in [14]. As n further increases, η_{pre} will remain relatively stable.

The top five buses selected by the pre-AI are listed in Table I. It is noted that the DBSs selected by the pre-AI have a high similarity, especially the first three buses, which indicates that the pre-AI have convergence and could provide a stable dataset for the post-AI training.

TABLE I
TOP FIVE BUSES SELECTED BY PRE-AI WHEN $n = 5$

No.	Bus subset
1	27, 31, 19, 39, 8
2	27, 31, 39, 16, 24
3	27, 31, 39, 17, 21
4	27, 31, 26, 39, 17
5	27, 31, 39, 15, 18

2) Computing Time and Performance

The relative computing time reduction $\xi_{\Delta t}$ is used to describe the computing demands in the training. $\xi_{\Delta t}$ is described by:

$$\xi_{\Delta t} = \frac{t_0 - (t_{pre} + t_{post})}{t_0} \quad (10)$$

where t_0 is the computing time spent on training without the pre-AI (the initial PMU data are directly inputted into post-AI), t_{pre} is the computing time spent on the DBS selection, and t_{post} is the computing time spent on the post-AI training.

The relationship between $\xi_{\Delta t}$ and n is shown as the black line in Fig. 9. It can be seen that the preprocessing of initial data by the pre-AI can effectively reduce the training time of DTSA. When the number for DBS is less than 20, more than one-fifth of the computing time can be saved. Even if 25 buses are extracted, the computing time can still be reduced by 14%. It is imperative to note that the total saved calculation time is considerable because the neural network will be trained many times to select the best option. The online response speed is another important parameter. The relative reduction of online average response time (RRART), the calculation is similar to (10)¹, is adopted to evaluate the performance of the DTSA in the online applications, see the red curve in Fig. 9. The DTSA still saves about 9% online computing time. Note that

¹In the calculation of RRART, t_0 is the online response time of the post-AI with the initial PMU data, t_{post} is the online response time of the post-AI with the extracted data by pre-AI, and note that t_{pre} is zero.

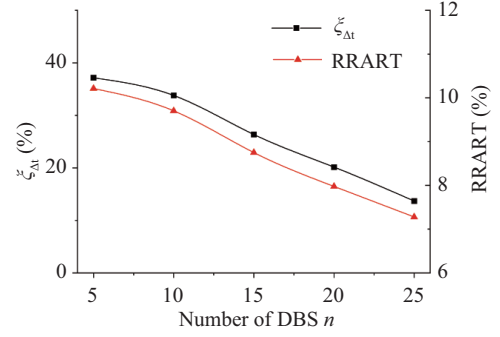
Fig. 9. $\xi_{\Delta t}$ of training and RRART of online in IEEE 39-bus system.

TABLE II
COMPUTING TIME AND ACCURACY OF DIFFERENT INTELLIGENT ALGORITHMS IN IEEE 39-BUS POWER SYSTEM

Method	Accuracy	Training Time (h)	Number of PMUs
DTSA ($n = 5$)	98.04%	0.75	5
DTSA ($n = 20$)	98.54%	0.95	20
ANN (fully connected)	92.09%	0.50	39
ANN (BiLSTM)	98.56%	1.28	39
DNN (fully connected)	96.32%	1.25	39
DNN (BiLSTM)	98.73%	1.63	39
CNN	97.01%	1.17	39
SAE	93.63%	0.72	39
SVM	93.00%	0.45	39

the DTSA will save more time in a large-scale power system, see Fig. 16 for more details.

To further verify the accuracy of DTSA, the shallow ANN, deep neural network (DNN), convolutional neural network (CNN), stacked autoencoder (SAE), and SVM for the same initial PMU data are compared in Table II. We can find that the BiLSTM can effectively improve the accuracy compared with the fully connected neural networks, SAE and SVM. This is primarily because the transient instability of power systems is a dynamic time-domain evolution process. BiLSTM layers can more effectively learn and extract the characteristics of the time series. Inversely, the fully connected networks, SAE, and SVM are better suited to handle fixed-moment features. CNN models can reuse the weights to improve the training effect due to the presence of convolutional kernels. However, for the time series, it is still inferior to BiLSTM. Therefore, the proposed DTSA and the neural networks with BiLSTM layers have higher accuracy for the time series. BiLSTM brings high accuracy but also causes the training time to be longer. By extracting the DBS, the proposed DTSA can effectively reduce the training time and ensure the accuracy of the results.

3) Different Algorithms of Pre-AI

The effect of different pre-AI algorithms is tested in this section. In the random method, the DBS is randomly selected ten times for calculating the average accuracy. Note that some recent studies use RF to extract key features [10]. Therefore, we compare the accuracy of the DTSA and RF (the RF is used to build the pre-AI). The RF uses an ensemble of unpruned decision trees to sequentially select the dominant buses (sequential forward selection, SFS).

Figure 10 shows that the average accuracy of the random method is lower than the proposed DTSA and RF, indicating

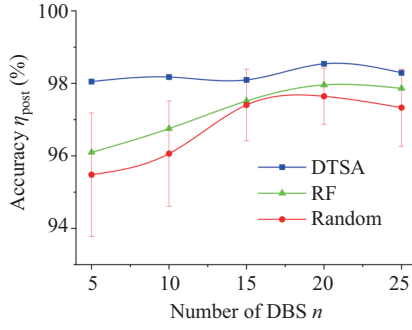


Fig. 10. The accuracy in IEEE 39-bus system with different pre-AI algorithms.

that the DBS determined by the heuristic algorithms is effective. By comparing the proposed DTSA with RF, it can be seen that the accuracy of RF is lower than DTSA because the RF is not suitable for the time series to obtain the dynamic time-domain features. That is, the pre-AI does not satisfy the first build principle in Section III.A, therefore the extracted DBS by RF cannot be effectively used by the post-AI.

4) Different Algorithms of Post-AI and the Influence of Measurement Noise

To illustrate the advantages of using BiLSTM for the post-AI, we compare BiLSTM with traditional LSTM. Fig. 11 shows that the BiLSTM has higher accuracy than LSTM and the improvement is greater when the number of DBSs are reduced, because the BiLSTM can extract features in both forward and backward directions of the time series. Therefore, the BiLSTM layer of post-AI is more effective for compressing the number of DBSs.

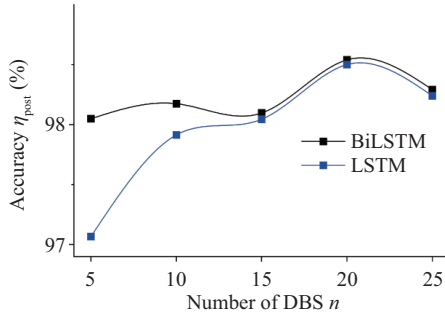


Fig. 11. The accuracy in IEEE 39-bus system with different post-AI network structures.

In the previous analyses, we assume that the acquired data are noise-free. However, in the exciting power systems, the PMUs may suffer from sampling noise. Therefore, we add different noise levels to the generated cases [27] to test the DTSA responses. The training results of DTSA are shown in Fig. 12. It can be noted that the accuracy is slightly decreased when the noise level is less than 10%. When the noise is 40%, the accuracy will drop by about 7 %. In the existing power system, we have requirements for the performance of the PMU. According to [28], the sampling error of PMU should be less than 1%. Therefore, the proposed DTSA can achieve high performance even considering the measurement noise.

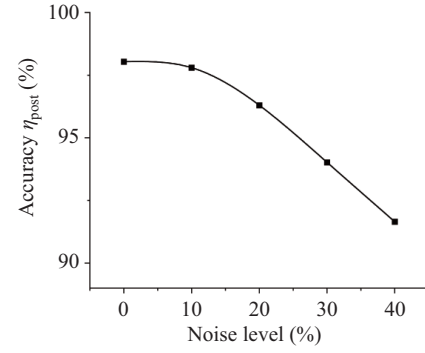


Fig. 12. Training results of DTSA in different noise levels.

5) Online Self-check Function of DTSA

Because the training process of DTSA does not consider the grid topology changes, the online self-check function is verified by changing the topology of the IEEE 39-bus system. The combinations of the cutting-off lines are listed in Table III. The lines 4–14 and 18–17 are heavy-load lines ($P(\text{active power}) > 150$ MW), while the lines 11–12, 3–4, 8–9, and 22–23 are light-load lines ($P < 80$ MW).

TABLE III
COMBINATIONS OF CUTTING LINES IN IEEE 39-BUS SYSTEM

No.	cutting off lines
1	4–14
2	4–14, 3–4
3	4–14, 18–17
4	4–14, 3–4, 18–17
5	3–4
6	3–4, 3–18
7	3–4, 8–9, 22–23
8	3–4, 8–9
9	11–12
10	11–12, 3–4
11	11–12, 3–4, 8–9
12	11–12, 3–4, 8–9, 22–23

The test sets with different grid topologies are prepared. η_{post} and η' are shown as the black and red lines in Fig. 13, respectively. The proposed DTSA has lower accuracy in combinations No.3 and No.4 and has higher accuracy in No.9, because the removed lines in No.3 and No.4 are heavily loaded, and the removed line in No.9 is lightly loaded.

The correlation coefficient C is calculated as follows:

$$C(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}[X]\text{Var}[Y]}} \quad (11)$$

where $\text{Cov}(X, Y)$ is the covariance of X and Y , $\text{Var}[X]$ and $\text{Var}[Y]$ are the variance of X and Y , respectively.

By computing the coefficient C between η_{post} and η' , we find that C is greater than 0.5 even if $n = 5$. Generally, when $C \geq 0.3$, X and Y are considered to be correlated. That is, η' can be used to reflect the actual accuracy of DTSA.

For an online self-check, two situations are considered. The first situation is that the DTSA works completely independently and should maintain high accuracy. That is, η' should be sensitive to the topology changes to ensure the accuracy of DTSA. As shown in Fig. 13, even if only one line is cut off, η' will drop below about 80%. Note that before the topology

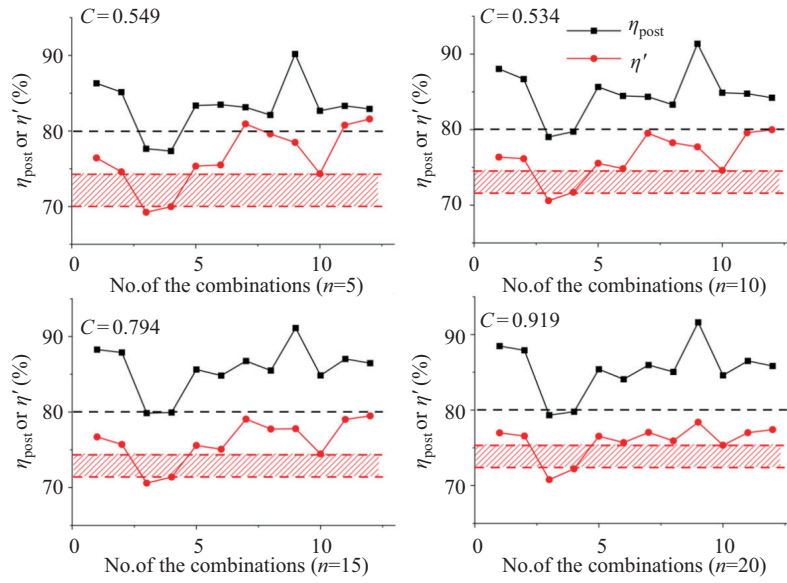


Fig. 13. Correlation between η_{post} and η' in IEEE 39-bus system.

changes, η' is approximately equal to η_{pre} (87%). If η'_{min} is set to 87%, all combinations of cutting off lines meet $\eta' < \eta'_{\text{min}}$. Thus, the online self-check can provide a warning and request to exit the run.

Since the training set cannot cover the whole operating state-space of power systems, it is difficult for current intelligent methods to guarantee the absolute reliability of TSA. The second situation is that the proposed DTSA is utilized as an auxiliary tool for operators to determine the stability of the power system. In this situation, the requirements for DTSA accuracy are not strict. For example, if the accuracy is greater than 80% after the topology changes, the DTSA still has a reference for operators to make a decision. As shown in Fig. 13, as long as combinations No.3 and No.4 are extracted, the actual DTSA accuracy can be guaranteed above 80% (the black dashed line). With the help of the correlation between η_{post} and η' , if η'_{min} is set in the red shaded area as shown in Fig. 13, the DTSA could meet the requirements. Even if $n = 5$, the online self-check can still be implemented.

The calculation of η_{post} requires real transient stability labels, which leads to a longer simulation time so that it cannot be applied online. The proposed online self-check function does not require real labels because the labels are determined by the post-AI, which greatly accelerates the generation of the test set. In the IEEE 39-bus system, it takes only 0.05 s to generate a usable case of DTSA, while it takes 10 s to generate a case with a real transient stability label. Parallel computing can further reduce the generation time of the test set. Therefore, the proposed DTSA can ensure real-time online self-checking.

C. Case Studies of the 300-Bus System

This section verifies the effectiveness of the DTSA in a large power system.

1) Introduction to the 300-bus System and Cases Generation

According to the current power supply situation in China and combined with relevant operating data, the 300-bus re-

newable energy system with an AC/DC hybrid power grid² is established as shown in Fig. 14. The 300-bus system includes 93 generators (generation capacity 26,857 MW) and 168 loads (total load 26,584 MW). This system is divided into three areas: The power supply in area-1 is primarily hydropower, area-2 includes a renewable energy power generation subarea and a hydroelectric-thermal power generation subarea, area-3 includes a distributed distribution network and large load subarea and area-3 has less installed capacity. The 300-bus system has seven cross-area DC lines, including three lines from area-1 to area-3, three lines from area-2 to area-3, and one line from area-2 to area-1.

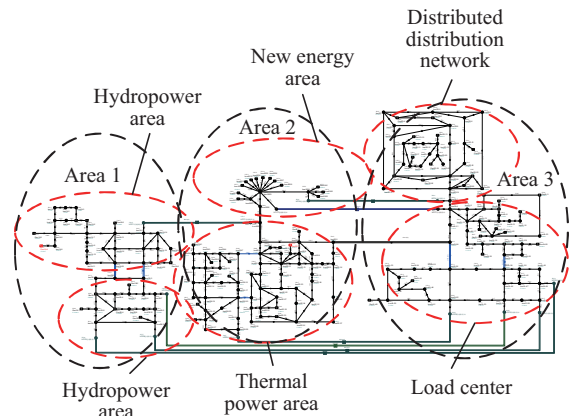


Fig. 14. Diagram of the 300-bus system.

PSDEdit software is used to simulate the test system to generate the training and test set. In the simulation, the power generation and load are between 90% and 110%, and the system is scanned for $(N - 1)$ contingency sets. The fault clearing time is set between 0 and 0.4 s, and 20,000 cases are generated.

²The 300-bus test system is developed by the China Electric Power Research Institute.

2) Automatic Selection of DBS Based on Pre-AI

One thousand cases are selected from the total cases for the pre-AI to extract the DBS, where 85% of the selected cases are the training set, and the remaining 15% are the testing set. Fig. 15 shows one calculation result of the relationship between η_{pre} and the number of DBS n . According to (3), n should be greater than 7. As n further increases, η_{pre} will remain relatively stable.

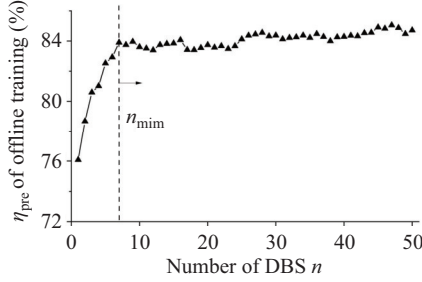


Fig. 15. Accuracy of pre-AI in the 300-bus system (test set).

In subsequent calculations, to ensure the computing speed of pre-AI, we evenly divide the 300 buses into ten groups and only need to select the optimal buses from the mini-groups. For every traversal of the 300 buses, ten buses can be extracted, which can improve computing efficiency in a large power system to define the DBS.

3) Computing Time and Performance

In the DTSA training, 85% of the total cases are the training set, and the remaining 15% are the testing set. The DTSA accuracy is shown as the blue curve in IV-C4. Even if only ten buses are extracted from the 300 buses, the DTSA still maintains high accuracy ($> 97.8\%$). The accuracy will slightly decrease when $n = 30$ or 40 , which is primarily due to the limitation of post-AI size. In addition, the value ξ_{Δ} is calculated by (10), as the black curve in IV-C4. When the number of the DBS is less than 50, DTSA can save more than 85% of computing power during the training process. Even for online applications, the RRART is about 65%. It is indicated that, in a large power system, the proposed DTSA can still generate an accurate TSA with sparse PUM sampling by combining the pre-AI and post-AI. Compared with the results of the IEEE 39-bus system, the proposed DTSA can save more computing power, that is, DTSA is more suitable in a large power system.

4) Online Self-check Function

To verify the online self-check function in the 300-bus system, the combinations of the cutting-off lines are shown in Table IV. The lines 3L-4-12, 3L-53-3L-54 and 3L-37-3L-38 are heavy-load lines ($P > 500$ MW), 3L-36-3L-37, 3L-26-3L-27 and 3L-12-3L-13 are light-load lines ($P < 100$ MW).

As shown in Fig. 16, the actual DTSA accuracy η_{post} is significantly reduced after heavy-loaded lines (e.g., No.3, No.5, and No.6) are cut off compared with light-loaded lines. When $n = 10$, the correlation coefficient between η_{post} and η' is 0.89. That is, η' still can be used to reflect the actual accuracy of DTSA. As long as η'_{min} is set in the red shaded

TABLE IV
COMBINATIONS OF CUTTING LINES IN 300-BUS SYSTEM

No.	cutting off lines
1	3L-36-3L-37
2	3L-36-3L-37, 3L-26-3L-27
3	3L-36-3L-37, 3L-26-3L-27, 3L-12-3L-13
4	3L-4-3L-12
5	3L-4-3L-12, 3L-53-3L-54
6	3L-4-12, 3L-53-3L-54, 3L-37-3L-38

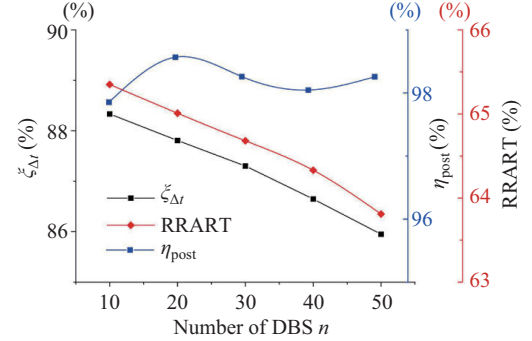


Fig. 16. $\xi_{\Delta t}$ of training and RRART of online in 300-bus system.

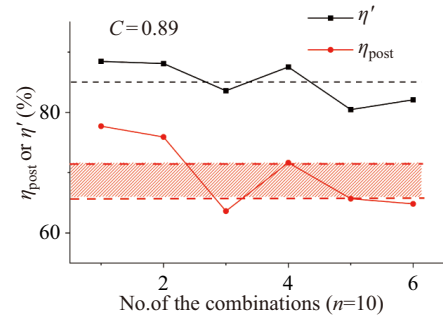


Fig. 17. Correlation between η_{post} and η' in 300-bus system.

area in Fig. 17, the value η_{post} can be guaranteed to be above 85%. Moreover, it should be noted that the decrease of η_{post} in a large-scale system is smaller than the IEEE 39-bus system because the stability of a large power system is more robust. Therefore, the proposed DTSA has more application value in a large power system.

V. CONCLUSION

The DTSA is proposed with sparse PMU sampling and online self-checking by integrating different AI algorithms. In DTSA, the pre-AI locates the DBS, which can significantly reduce the computation costs and ensure TSA accuracy for large power systems. The extracted DBS can reduce the number of PMUs and relieve the burden of the communication systems. The online self-check function of DTSA is achieved by the correlation between the pre-AI and post-AI predictions, which can ensure the online TSA's validity when the power system changes.

Although the combination of multiple AI algorithms is used for TSA in this study, it could be readily extended to other applications in power systems. The proposed method provides a reference for properly combining multiple AI algorithms to solve practical problems in power systems.

REFERENCES

- [1] B. J. Li, J. B. Li, H. J. Li, W. S. He, X. F. Zeng, and C. Y. Yu, "Analysis of turkish blackout on March 31, 2015 and lessons on China power grid," *Proceedings of the CSEE*, vol. 36, no. 21, pp. 5788–5795, Nov. 2016.
- [2] Y. Liu, "Analysis on and inspiration of the "9.13" islanding and outage of Brazilian remote northwest power grid," *Proceedings of the CSEE*, vol. 38, no. 11, pp. 3204–3213, Jun. 2018.
- [3] I. Konstantelos, G. Jamgotchian, S. H. Tindemans, P. Duchesne, S. Cole, C. Merckx, G. Strbac, and P. Panciatici, "Implementation of a massively parallel dynamic security assessment platform for large-scale grids," *IEEE Transactions on Smart Grid*, vol. 8, no. 3, pp. 1417–1426, May 2017.
- [4] Y. X. Yu, Y. L. Liu, C. Qin, and T. K. Yang, "Theory and method of power system integrated security region irrelevant to operation states: an introduction," *Engineering*, vol. 6, no. 7, pp. 754–777, Jun. 2020.
- [5] P. Kundur, *Power System Stability and Control*, New York: McGraw-Hill, 1994.
- [6] L. Xi, L. P. Zhou, L. Liu, D. L. Duan, Y. C. Xu, L. Q. Yang, and S. X. Wang, "A deep reinforcement learning algorithm for the power order optimization allocation of AGC in interconnected power grids," *CSEE Journal of Power and Energy Systems*, vol. 6, no. 3, pp. 712–723, Sep. 2020.
- [7] Y. Tang, Y. H. Huang, H. Z. Wang, C. Wang, Q. Guo, and W. Yao, "Framework for artificial intelligence analysis in large-scale power grids based on digital simulation," *CSEE Journal of Power and Energy Systems*, vol. 4, no. 4, pp. 459–468, Dec. 2018.
- [8] Q. Z. Liu, Y. B. Shen, L. Wu, J. Li, L. R. Zhuang, and S. F. Wang, "A hybrid FCW-EMD and KF-BA-SVM based model for short-term load forecasting," *CSEE Journal of Power and Energy Systems*, vol. 4, no. 2, pp. 226–237, Jun. 2018.
- [9] E. A. Frimpong, P. Y. Okyere, and J. Asumadu, "On-line determination of transient stability status using MLPNN," in *2017 IEEE PES Power-Africa*, 2017, pp. 23–27.
- [10] C. Zhang, Y. S. Li, Z. H. Yu, and F. Tian, "Feature selection of power system transient stability assessment based on random forest and recursive feature elimination," in *2016 IEEE PES Asia-Pacific Power and Energy Engineering Conference (APPEEC)*, Xi'an, 2016, pp. 1264–1268.
- [11] F. Karim, S. Majumdar, H. Darabi, and S. Chen, "LSTM fully convolutional networks for time series classification," *IEEE Access*, vol. 6, pp. 1662–1669, Feb. 2018.
- [12] F. Karim, S. Majumdar, H. Darabi, and S. Harford, "Multivariate LSTM-FCNs for time series classification," *Neural Networks*, vol. 116, pp. 237–245, Aug. 2019.
- [13] J. J. Q. Yu, D. J. Hill, A. Y. S. Lam, J. T. Gu, and V. O. K. Li, "Intelligent time-adaptive transient stability assessment system," *IEEE Transactions on Power Systems*, vol. 33, no. 1, pp. 1049–1058, Jan. 2018.
- [14] B. D. Tan, J. Yang, T. Zhou, X. P. Zhang, Y. Liu, S. B. Jiang, and C. Luo, "Spatial-temporal adaptive transient stability assessment for power system under missing data," *International Journal of Electrical Power & Energy Systems*, vol. 123, pp. 106237, Dec. 2020.
- [15] N. C. Thompson, K. Greenewald, K. Lee, and G. F. Manso, "The computational limits of deep learning," *arXiv preprint arXiv: 2007.05558*, 2020.
- [16] S. Y. Liu, Y. X. Zhao, Z. Z. Lin, Y. L. Liu, Y. Ding, L. Yang, and S. M. Yi, "Data-driven event detection of power systems based on unequal-interval reduction of PMU data and local outlier factor," *IEEE Transactions on Smart Grid*, vol. 11, no. 2, pp. 1630–1643, Mar. 2020.
- [17] M. Y. Sun, I. Konstantelos, and G. Strbac, "A deep learning-based feature extraction framework for system security assessment," *IEEE Transactions on Smart Grid*, vol. 10, no. 5, pp. 5007–5020, Sep. 2019.
- [18] Y. H. Huang, Z. H. Yu, D. Y. Shi, and X. X. Zhou, "Strategy of huge electric power system stability quick judgment based on massive historical online data," *Proceedings of the CSEE*, vol. 36, no. 3, pp. 596–603, Feb. 2016.
- [19] W. Hu, Z. X. Lu, S. Wu, W. L. Zhang, Y. Dong, R. Yu, and B. S. Liu, "Real-time transient stability assessment in power system based on improved SVM," *Journal of Modern Power Systems and Clean Energy*, vol. 7, no. 1, pp. 26–37, Jan. 2019.
- [20] R. Y. Zhang, J. Y. Wu, M. Y. Shao, B. Q. Li, and Y. Z. Lu, "Transient stability prediction of power systems based on deep belief networks," in *2018 2nd IEEE Conference on Energy Internet and Energy System Integration (EI2)*, Beijing, 2018, pp. 1–6.
- [21] F. Aminifar, A. Khodaei, M. Fotuhi-Firuzabad, and M. Shahidehpour, "Contingency-constrained PMU placement in power networks," *IEEE Transactions on Power Systems*, vol. 25, no. 1, pp. 516–523, Feb. 2010.
- [22] X. Y. Yu, J. Cao, Z. Fan, M. M. Xu, and L. Y. Xiao, "Faulty feeder identification in resonant grounding distribution networks based on deep learning and transfer learning," *CSEE Journal of Power and Energy Systems*, pp. 1–10, Apr. 2021.
- [23] G. Z. Wang, J. B. Guo, S. C. Ma, Y. W. Shang, Q. L. Guo, G. Q. Bu, Y. H. Huang, S. C. Zeng, and Z. H. Zhou, "Preliminary study of power system enhanced intelligence analysis," *Proceedings of the CSEE*, vol. 40, no. 16, pp. 5079–5088, Aug. 2020.
- [24] M. Wooldridge, *An Introduction to Multiagent Systems*, Hoboken: John Wiley & Sons, 2009.
- [25] W. Ji and K. C. Chee, "Prediction of hourly solar radiation using a novel hybrid model of ARMA and TDNN," *Solar Energy*, vol. 85, no. 5, pp. 808–817, May 2011.
- [26] A. R. Messina, *Wide Area Monitoring of Interconnected Power Systems*, London: IET, 2015.
- [27] M. He, V. Vittal, and J. S. Zhang, "Online dynamic security assessment with missing PMU measurements: A data mining approach," *IEEE Transactions on Power Systems*, vol. 28, no. 2, pp. 1969–1977, May 2013.
- [28] *IEEE Standard for Synchrophasor Data Transfer for Power Systems*, IEEE Standard C37.118.2–2011, 2011.



Guozheng Wang received B.S. (2015) and M.S. (2018) degrees from Central South University, Changsha, China, and Shandong University, Jinan, China, respectively. He is currently pursuing a Ph.D. degree at Tsinghua University, Beijing, China. His research interests include power system analysis and control, and intelligent control applications.



Jianbo Guo (Senior Member, IEEE) received B.S. (1982) and M.S. (1984) degrees from Huazhong University of Science and Technology, Wuhan, China, and China Electric Power Research Institute, Beijing, China, respectively. Since 2013, he has been an elected Academician of the Chinese Academy of Engineering, Beijing, China. He has been engaged, for a significant time, in power system analysis & control.



Shicong Ma (M'12) received B.S. (2002) and Ph.D. (2008) degrees from Shandong University. He is currently employed by China Electric Power Research Institute. His research interests include power system stability and security, data mining for power system operation and control, and autonomous power systems.



Xi Zhang received B.Eng. degree from Beijing Jiaotong University, Beijing, China, in 2013 and Ph.D. degree from Hong Kong Polytechnic University, Hong Kong, in 2017, both in Electrical Engineering. He worked as a postdoc in China Electric Power Research Institute, Beijing, China from 2018 to 2019. Currently, he is an assistant professor in Beijing Institute of Technology, Beijing, China since 2020. His research interests include complexity, resilience and stability analysis of the power system with high penetrations of power electronic devices.



Shixiong Fan received a Ph.D. degree from the Department of Electrical and Electronic Engineering, Strathclyde University, Glasgow, U.K., in 2012. Currently he is working for China Electric Power Research Institute as professorate senior engineer. His current research interests include power network dispatching automation and application of artificial intelligence and big data analysis in power grid.



Qinglai Guo (SM'14) graduated from the Department of Electrical Engineering, Tsinghua University, Beijing, China, in 2000 with a B.S. degree. He received a Ph.D. degree from Tsinghua University in 2005 where he is now a professor. He is an IET Fellow, IEEE senior member and a CIGRE member. Currently he is the TCPC of the Energy Internet Coordinating Committee of IEEE PES, the co-chair of IEEE PES Work Group on "Energy Internet," IEEE PES Task Force on "Cyber-Physical Interdependence for Power System Operation and Control," and IEEE PES Task Force on "Voltage Control for Smart Grid." His research interests include smart grids, cyber-physical systems and electrical power control center applications.



Haotian Xu graduated from the Department of Mathematics and Statistics, Lanzhou University, Gansu, China, in 2011 with a B.S. degree. He received a Ph.D. degree from the Department of Electrical Engineering and Electronics, Liverpool University, UK, in 2020. Since 2020, he has been an engineer with the China Electric Power Research Institute. His research interests include stability analysis of time delay system and power system analysis and control.