Transient Stability Assessment in Power Systems: A Spatiotemporal Graph Convolutional Network Approach with Graph Simplification

Dan Zhang [1], Yuan Yang [2,*], Bingjie Shen [2], Tao Wang [2,*] and Min Cheng [1]

[1] Yunnan Electric Power Dispatching and Control Center, Kunming 650011, China; zhangdan@dltd.yn.csg.cn (D.Z.); chengmin@dltd.yn.csg.cn (M.C.)
[2] Hebei Key Laboratory of Physics and Energy Technology, Baoding 071003, China; 220232217030@ncepu.edu.cn
* Correspondence: 220222217002@ncepu.edu.cn (Y.Y.); wangtao@ncepu.edu.cn (T.W.)

Abstract: Accurate and fast transient stability assessment (TSA) of power systems is crucial for safe operation. However, deep learning-based methods require long training and fail to simultaneously extract the spatiotemporal characteristics of the transient process in power systems, limiting their performance in prediction. This paper proposes a novel TSA method based on a spatiotemporal graph convolutional network with graph simplification. First, based on the topology and node information entropy of power grids, as well as the power flow of each node, the input characteristic matrix is compressed to accelerate evaluation. Then, a high-performance TSA model combining a graph convolutional network and a Gated Convolutional Network is constructed to extract the spatial features of the power grid and the temporal features of the transient process. This model establishes a mapping relationship between spatiotemporal features and their transient stability. Finally, the focal loss function has been improved to dynamically adjust the influence of samples with different levels of difficulty on model training, adaptively addressing the challenge of sample imbalance. This improvement reduces misclassification rates and enhances overall accuracy. Case studies on the IEEE 39-bus system demonstrate that the proposed method is rapid, reliable, and generalizable.

Keywords: transient stability assessment; graph simplification; spatial features; temporal features; sample imbalance

1. Introduction

With the ever-increasing penetration of renewable energy [1,2] and the application of power electronic devices [3], the structure and operating conditions of transmission networks become increasingly complex, posing unprecedented challenges to the stability of traditional power systems. Changes in the network structure and power flow necessitate a reassessment of their impacts on power system stability [4]. Particularly, as the operational modes of traditional equipment and the distribution of power flows have been altered, various aspects such as rotor angle stability, voltage stability, and frequency stability are affected. Additionally, new stability issues, including electromechanical-like low-frequency oscillation and electromagnetic wideband oscillations, are gradually emerging [5]. Although renewable energy devices connected to the grid via power electronics improve flexibility and environmental sustainability [6], their low inertia characteristics weaken the system's resistance to frequency disturbances [7]. Furthermore, the rapid dynamic response of power electronic devices, along with their complex control strategies, complicates the analysis and control of transient stability in power systems. Consequently, there is an urgent need for an accurate and rapid TSA method to ensure the safe and stable operation of power systems under these evolving conditions.

Traditional TSA methods primarily rely on energy-based direct methods [8] and time–domain simulations [9]. However, these approaches are increasingly facing challenges in addressing the uncertainty and complexity introduced by high-penetration renewable energy systems [10]. In recent decades, the rapidly developing deep learning (DL) has been generally applied to build the TSA model for the power system [11], and compared to traditional transient stability assessment methods, DL establishes a mapping relationship between the physical characteristics of power systems and the results of transient stability assessment through offline training. Subsequently, the trained model is applied for online transient stability prediction in power systems. On the one hand, the trajectories of state and operational variables following a disturbance can directly reflect the transient stability of the power system [12,13]. In [14], a deep belief network (DBN) is presented to precisely model the mapping relationship between the input and the output, realizing the accuracy evaluation. A novel evaluation of power system transient stability assessment (TSA) based on a convolutional neural network (CNN) has been proposed, which significantly enhances the speed of batch assessment [15]. The long short-term memory (LSTM) is employed in [16,17] to extract features from voltage phasors at different time intervals, resulting in a time-adaptive TSA model. Furthermore, the integration of LSTM with parallel CNN effectively captures temporal relationships within transient processes, thereby improving the performance of the TSA model [18]. The aforementioned methods utilize electrical measurement data as inputs to construct power system stability assessment models based on temporal characteristics, achieving a certain level of effectiveness in stability assessment [19]. On the other hand, considering the influence of power grid topology on power system transient stability, some graph neural network (GNN)-based methods have been employed for TSA [20]. To mine the impact of the power grid topology, the graph convolutional network (GCN) is adopted to improve the prediction performance in TSA [21]. Additionally, the graph attention network (GAT) is utilized in [22] to effectively aggregate spatial features during transient processes, while residual structures are introduced to prevent network degradation and enhance the accuracy of TSA.

Existing DL-based TSA models typically focus on either the temporal measurement data of the power system or the spatial characteristics of grid topology [23]. However, these models rarely integrate both aspects to capture the spatiotemporal features of grid data. Incorporating both spatial and temporal features is crucial for accurately evaluating power system stability [24,25]. By analyzing and incorporating the interactions between these features, a comprehensive understanding of the dynamic behavior during transitions can be achieved, leading to more reliable assessments of power system stability. During major disturbances, such as faults and stability control actions, the transient processes of power systems exhibit significant spatiotemporal correlations. Analyzing these correlations provides a comprehensive understanding of the dynamic behavior during transitions, leading to more reliable assessments of power system stability [26,27]. Furthermore, as power systems continue to grow in scale, the challenge of prolonged model training times has become increasingly prominent [28,29]. Under limited resource conditions, the effective and rapid training of models on large-scale grid data remain a critical challenge that must be addressed [30].

To overcome these limitations, we propose a spatiotemporal graph convolutional network approach with graph simplification (GS-STGCN) for TSA, which improves prediction performance and accelerates TSA evaluation. The contributions of this paper are as follows:

1. A node influence-based graph simplification method is proposed, incorporating both the topological and electrical characteristics of the power system. This approach introduces pruning parameters to remove less influential nodes, thereby optimizing the model input and enhancing the training efficiency of the TSA model.

2. A spatiotemporal graph convolutional network (STGCN) is designed to simultaneously capture spatial and temporal features during transient processes. It combines Gated Convolutional Network (Gated CNN) with GCN to effectively extract spa-

tiotemporal features, aiming to integrate temporal sequence data with topological information and develop a TSA model with enhanced performance.

3. The focal loss function has been improved to adjust the impact of samples with varying levels of difficulty on model training. This adjustment helps prevent premature convergence caused by excessively small losses in later stages, which could otherwise degrade assessment performance. Additionally, an adaptive class weighting factor is introduced to reduce the need for manual parameter tuning.

The remainder of this paper is organized as follows: Section 2 introduces the GS-STGCN methodology; Section 3 elaborates on the proposed TSA framework; and Section 4 conducts case studies on the IEEE 39-bus system; finally, Section 5 concludes the whole paper's work.

2. GS-STGCN Methodology

2.1. Node Influence-Based Graph Simplification Method

The expansion of the power system scale will cause the growth of input data. In power systems, some nodes exhibit noteworthy importance and carry rich information, exerting a substantial impact on TSA. Conversely, other nodes contribute relatively less to model training, but they may increase training time and memory usage. Therefore, a node influence-based graph simplification method (GS) is proposed to compress large-scale power grid input data, thus improving the calculation speed of the TSA model.

- Influence Indicators

The K-Shell algorithm (KS) is a coarse-grained node importance assessment method based on the entire network structure. The algorithm stratifies the network from the periphery to the core and assigns a KS value to each node, reflecting its importance within the network. The magnitude of the KS value effectively distinguishes core nodes from non-core nodes. A larger KS value indicates that the node is closer to the core of the network and possesses greater influence.

Additionally, the node information entropy is utilized to assess the amount of information a node carries within the power system. Calculating the information entropy of a node mainly relies on the node's neighborhood degree information, which can be expressed as

$$e_i = -\sum_{j \in N(i)} I_j \cdot \ln I_j \tag{1}$$

where $N(i)$ is the set of neighbor nodes of node i; I_i is defined as

$$I_i = \frac{k_i}{\sum\limits_{j \in N(i)} k_j} \tag{2}$$

where k_i is the degree of node i.

Considering the location and information entropy of the nodes in the network, the topological entropy impact indicator is defined as

$$e_i' = -K_i \sum_{j \in N(i)} I_j \cdot \ln I_j \cdot K_j \tag{3}$$

where K_i is the KS value of node i.

In power systems, power flow is a critical indicator of a node's output capacity and load demand, reflecting the node's importance within the system. The active power and reactive power at steady-state conditions are selected to jointly assess the node's power flow. The power flow influence indicator can be defined as

$$f_i = \ln\left(\sqrt{P_i^2 + Q_i^2}\right) \tag{4}$$

where P_i and Q_i are the active power and reactive power of node i in the steady state, respectively.

Considering both the topological entropy and power flow influence of a node within the system, a comprehensive influence indicator can be expressed as

$$S_i = w_1 \cdot \gamma\left(e_i'\right) + w_2 \cdot \gamma(f_i) \tag{5}$$

where w_1 and w_2 are the weights of the two impact assessment indicators; $\gamma(\cdot)$ is the MIN-MAX normalization method. The comprehensive influence indicator of a node quantifies the importance of each node in the power grid by considering both the system's topological entropy and power flow influence.

- The steps of GS

The importance of each node in the power grid is quantified using a comprehensive influence indicator. This indicator is then used to simplify and compress the system's graph data structure. The steps are as follows:

Comprehensive Influence Index Calculation The comprehensive influence index is calculated by considering both the topological entropy and the power flow influence of each node within the system. Using Equations (1)–(5), the comprehensive influence index for each node in the power grid is computed and ranked in ascending order, resulting in the sorted node influence vector:

$$\boldsymbol{S} = [S_1, S_2, \cdots, S_m] \tag{6}$$

where m is the number of nodes.

Node Selection and Elimination Nodes and their connected transmission lines are removed sequentially according to their order in the set $\boldsymbol{S}$, with continuous checks to ensure that the power grid remains connected. If the grid becomes disconnected after a pruning operation, the removal is reversed and the process moves to the next node. This procedure continues until the proportion of removed nodes exceeds the predefined threshold ξ, at which point the process is terminated.

Reconstruction of the Input Feature Matrix After pruning, the topology and feature matrix of the remaining nodes in the power grid are reconstructed, resulting in a reduced number of nodes and corresponding feature data. The impact on the grid topology before and after pruning is illustrated in Figure 1.
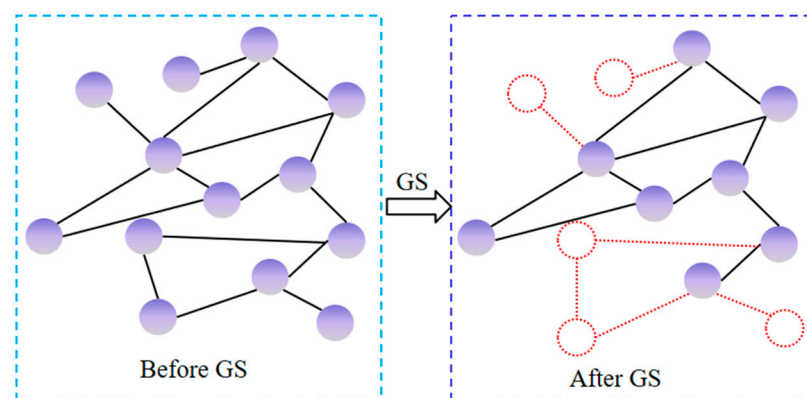

Figure 1. The effect of topology pruning on the power grid diagram before and after.

2.2. TSA Based on Spatiotemporal Graph Convolutional Networks

TSA is a classification task that relies on spatiotemporal features. The spatial structure of the power grid represents the spatial features, while the temporal variations in electrical quantities during transient processes represent the temporal features. In this study, GCN [21] and Gated CNN [31] are employed to capture these spatial and temporal features, respectively, enabling the accurate extraction of transient stability characteristics.

- Spatial Convolutional Layer

CNN and GCN are classic neural network models widely used for feature extraction. CNN is primarily designed for data in Euclidean space, excelling at extracting features from two-dimensional tensors. In contrast, GCN is specifically tailored for processing graph-structured data, making it ideal for extracting spatial features from the nodes within the power grid's graph structure.

The power grid structure can be defined using the normalized graph Laplacian matrix as follows:

$$L = E - D^{-\frac{1}{2}} A D^{-\frac{1}{2}} = U \Lambda U^{\mathrm{T}} \tag{7}$$

where E is an identity matrix, D is the diagonal degree matrix of the power grid graph structure G; A is the adjacency matrix of G; U is the matrix of eigenvectors of L; $\Lambda = [\lambda_1, \lambda_2, \ldots, \lambda_n]$ is the diagonal matrix of eigenvalues.

The graph convolution formula can be defined as

$$(X * g) = U((U^{\mathrm{T}} g) \otimes (U^{\mathrm{T}} X)) \tag{8}$$

where X is the input of the spatial convolutional layer; $*$ is graph convolution operator; g is the convolution kernel; $\otimes$ is the Hadamard product.

Using $g_\theta = U^T g$ as the learnable convolution kernel, the graph convolution formula is

$$(X * g) = U g_\theta U^T X \tag{9}$$

To accelerate the computation of Equation (9), the Chebyshev polynomial is employed to approximate the convolution kernel, which can be obtained as follows:

$$g_\theta = g_\theta(\Lambda) \approx \sum_{H=0}^{H-1} \theta_H T_H(\widetilde{\Lambda}) \tag{10}$$

where $\widetilde{\Lambda} = 2\Lambda/\lambda_{max} - E$; λ_{max} is the largest eigenvalue of L; θ_H is the vector of Chebyshev polynomial coefficients. $T_H(x)$ is the H-order Chebyshev polynomial and it can be expressed as

$$T_H(x) = 2x T_{H-1}(x) - T_{H-2}(x) \tag{11}$$

where $T_0(x) = 1$, $T_1(x) = x$.

By introducing Chebyshev polynomials, the graph convolution formula is derived as follows:

$$g * x = \sum_{H=0}^{H-1} \theta_H T_H\left(\frac{2L}{\lambda_{max}} - E\right) x \tag{12}$$

- Temporal Convolutional Layer

The temporal convolutional layer comprises two key components. First, Gated CNN is used to capture long-term dependencies across different time steps. Second, the Gated Linear Unit (GLU) serves as the nonlinear activation function, effectively integrating both linear and nonlinear transformations. In the temporal convolutional layer, the input for each node is a sequence with C_i input channels and a length of M, represented as $X \in R^{M \times C_i}$. The convolution kernel $\Gamma \in R^{k \times C_i \times 2C_0}$ is designed to perform a one-dimensional convolution operation on X, with the output can be defined as

$$X' = \Gamma * X = p \otimes \sigma(q) \in R^{(M-k+1) \times C_0} \tag{13}$$

where p and q are the input of gates in GLU respectively; $\sigma(\cdot)$ denotes the activation function; C_0 is the number of output channels; k is the size of the convolution kernel. Gated CNN is employed to extract temporal features of the transient process. Compared to Recurrent Neural Network (RNN) and their derivatives, Gated CNN offers faster training speeds and is less prone to gradient-related issues.

- Spatiotemporal graph convolutional network

To accurately and efficiently fuse spatial and temporal features, a spatiotemporal graph convolutional network is designed, as illustrated in Figure 2.
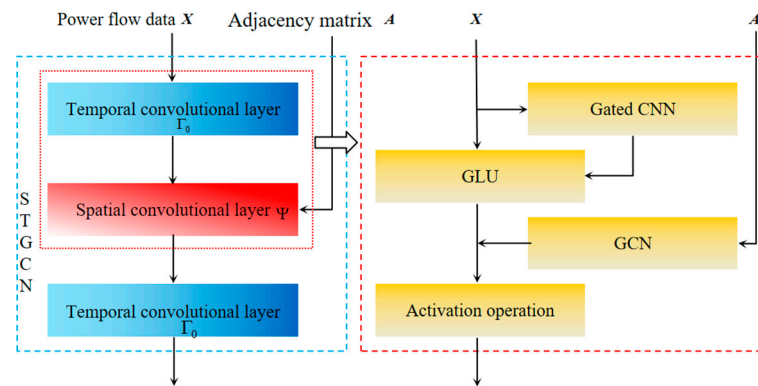

Figure 2. The spatiotemporal graph convolutional networks.

In this network, the power flow data X is the input for the temporal convolutional layer Γ_0, while the adjacency matrix A of the power grid's graph structure is the input for the spatial convolutional layer Ψ. First, X is input into Γ_0, where Gated CNN and GLU are used to capture temporal dependencies, performing convolution operations along the time dimension. The extracted temporal features are then combined with the spatial features, followed by an activation operation to perform graph convolution along the spatial dimension. Subsequently, the temporal and spatial features are further integrated by processing them through the temporal convolutional layer Γ_1, ultimately producing the fused spatiotemporal feature X^o, defined as:

$$X^o = \Gamma_1 * \sigma\big(\Psi * (\Gamma_0 * X)\big) \tag{14}$$

2.3. GS-STGCN Model

To efficiently and accurately perform TSA, this paper proposes a Graph Simplification-based spatiotemporal graph convolutional network (GS-STGCN) model, as illustrated in Figure 3. The model includes several key components: a GS module, two STGCN layers with a residual structure, and four fully connected layers followed by a SoftMax layer. The GS module preprocesses electrical measurement data by removing redundant information, enhancing computational efficiency, and allowing the model to focus on relevant features. The STGCN layers extract spatial and temporal features from the graph data, capturing complex patterns in power system behavior over time. The residual structure between the STGCN layers facilitates gradient flow during backpropagation, mitigating the vanishing gradient problem and enabling deeper architectures without sacrificing performance. The four fully connected layers process the extracted features, allowing the model to learn complex interactions. This architecture improves the model's generalization and prediction accuracy. Finally, the SoftMax layer converts the outputs into probability distributions over stability classes, yielding the final stability assessment result. This structured approach ensures that the GS-STGCN model effectively integrates graph simplification, feature extraction, and classification for reliable TSA.
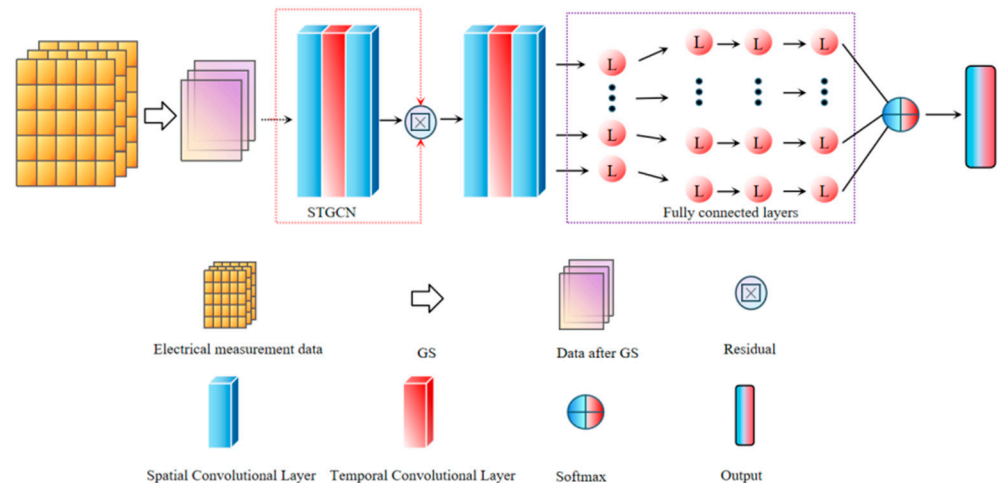
Figure 3. The GS-STGCN model.

3. GS-STGCN-Based TSA Model

3.1. Feature Input and Output

To enhance the accuracy of TSA, it is essential to comprehensively consider the spatiotemporal distribution characteristics of transient stability data. Consequently, the input features of the assessment model should be constructed from both temporal and spatial perspectives. For temporal features, variations in voltage magnitudes and phase angles can intuitively reflect the system's response characteristics, making node voltage amplitude and phase angle time series data suitable input features. For spatial features, the system's topological structure is used as an input feature to account for the impact of the power system's topological connections on energy transmission paths. And it captures the spatial evolution patterns of transient stability data by considering the dynamic interactions between nodes.

TSA is essentially a binary classification problem, where the labels 0 and 1 correspond to stable and unstable samples, respectively. The criterion for stability is determined based on the Transient Stability Index:

$$I_{\text{TSI}} = \frac{360° - |\Delta\delta_{\max}|}{360° + |\Delta\delta_{\max}|} \tag{15}$$

The I_{TSI} is calculated using the maximum rotor angle difference $\Delta\delta_{\max}$ between generators as its benchmark. A positive I_{TSI} ($I_{\text{TSI}} > 0$) corresponds to a stable system, which is labeled as (1,0), indicating that the system can maintain a synchronous operation following a disturbance. Conversely, a non-positive I_{TSI} ($I_{\text{TSI}} \leq 0$) is associated with an unstable system, labeled as (0,1), suggesting that the system has failed to sustain synchronization. Fundamentally, the I_{TSI} assesses the capacity for synchronous machine operation within an electrical power system by determining if the rotor angle difference $\Delta\delta_{\max}$ is greater than 360°. If $\Delta\delta_{\max}$ is less than this critical threshold, it indicates that the system has a sufficient stability margin to resist disturbances and allows for the resynchronization of generators. In these cases, the I_{TSI} is positive, confirming the system's stability. In contrast, if the angle difference $\Delta\delta_{\max}$ is equal to or exceeds 360°, it reflects an inability to maintain synchronous operation, resulting in a I_{TSI} that is either negative or zero, which is indicative of system instability. Moreover, I_{TSI} is intrinsically linked to the dynamics of power systems as it quantifies the system's inertial response following a significant disturbance, such as a fault or a change in load. The I_{TSI} value provides a quantitative measure of the system's resilience, which is essential for comprehending the dynamic behavior of power systems during transient stability evaluations.

3.2. The Adaptive Focal Loss Function

In practical power system applications, class imbalance is commonly observed, with the number of stable samples markedly exceeding that of unstable ones. During training, the abundance of stable samples results in more frequent parameter updates for these samples. Consequently, these updates dominate the overall contribution to the loss function, leading to an underfitting of the unstable samples. This imbalance increases the likelihood of the model misclassifying unknown samples as stable.

Simultaneously, the training dataset includes both easily classified samples and difficult samples that are prone to misclassification. A large number of simple samples can dominate the gradient update direction, causing the model's overall learning process to deviate and leading to ineffective learning. During iterations involving a substantial number of simple samples, the loss function may not decrease toward the desired optimization direction or may fail to reach the optimal solution.

An effective approach is to modify the model's loss function during training. The focal loss (FL) function, an improvement of the traditional cross-entropy loss function (CE), is expressed as

$$f_L = \begin{cases} -(1-w)p^b ln(1-p), & y = 0 \\ -w(1-p)^b lnp, & y = 1 \end{cases} \tag{16}$$

where p is the predicted probability; w is the class weighting factor used to adjust the balance between positive and negative samples—when w is closer to 1, the contribution of negative samples to the loss function increases. b is the difficulty weighting factor, where a larger b increases the contribution of difficult samples to the loss function.

Intuitively, the FL function addresses the class imbalance between stable and unstable samples through w, and it tackles the imbalance between easy and difficult samples via b. However, as the model's recognition ability improves during training, the distinction between easy and difficult samples evolves. Consequently, fixing the value of b may negatively impact the later stages of training. Additionally, treating w as a hyperparameter requires extensive experimentation to optimize the loss function, which can lead to considerable computational time and resource consumption.

To address the issues mentioned above, adaptive class weighting factors and dynamically changing difficulty weighting factors are introduced based on the FL function, resulting in the improved focal loss (IFL) function. The improved w and b are defined as follows:

$$\begin{cases} a = \frac{m_0}{m_0 + m_1} \\ b_c = \frac{b_s}{\ln(E - 1 + e)} \end{cases} \tag{17}$$

where m_0 and m_1 represent the number of positive and negative samples, respectively; E is the total number of training iterations; and b_s is the initially set difficulty weighting factor.

3.3. Classification Performance Metrics

The evaluation results of the TSA method are presented in Table 1. Four evaluation metrics—accuracy A_{cc}, precision P_{rec}, recall f_1, and F1-score f_1—are used to comprehensively assess the performance of the different TSA models.

$$A_{cc} = \frac{T_P + T_N}{T_P + T_N + F_P + F_N} \tag{18}$$

$$P_{rec} = \frac{T_P}{T_P + F_P} \tag{19}$$

$$R_{ec} = \frac{T_P}{F_N + T_P} \tag{20}$$

$$f_1 = 2 \times \frac{P_{rec} \times R_{ec}}{P_{rec} + R_{ec}} \tag{21}$$

where T_P is the number of samples predicted by the model to be positive and actually positive; T_N is the number of samples predicted to be negative and actually negative; F_P is the number of samples predicted by the model to be positive but actually negative, and F_N is the number of samples predicted by the model to be negative but actually positive.

Table 1. Average results of various indicators for five models.

Model	A_{cc}/%	P_{rec}/%	R_{ec}/%	f_1/%
GS-STGCN	99.51	99.98	98.92	99.46
GCN	98.62	99.77	97.05	98.49
GAT	98.87	98.91	98.65	98.78
LSTM	97.52	99.71	94.87	97.24
SVM	92.61	96.97	86.52	91.45

4. Case Study

To verify the effectiveness and applicability of the proposed method, the model is implemented using PaddlePaddle (version 2.4) within a Python (version 3.9) environment and applied to the IEEE 39-bus system for case analysis.

4.1. Database Generation

The Dynamic Simulation Program is used to simulate and generate transient stability data. The load level is increased from 70% to 110% in 10% increments. The fault type is a three-phase short circuit, with faults occurring at 20 different locations, increasing by 5% from 0% to 95%. Fault durations vary from 2 to 18 cycles. A total of 15,000 samples are generated, including 9500 stable samples and 5500 unstable samples. Simulation Data Link: https://github.com/yangyuan123654/yydata (accessed on 12 October 2024).

4.2. Performance Testing and Analysis of GS-STGCN Model Evaluation

To verify the superiority of the GS-STGCN model for TSA, it is compared against the SVM [32], GCN [21], GAT [22], and LSTM [16] models. The SVM model uses a radial basis function (RBF) kernel, with sample feature data transformed into a one-dimensional format to meet the kernel's requirements. The LSTM model comprises two LSTM layers, followed by a fully connected layer and a SoftMax output layer. The GCN model consists of two stacked GCN layers, a fully connected layer, and a SoftMax output layer. Similarly, the GAT model consists of two stacked GAT layers, a fully connected layer, and a SoftMax output layer. The first GAT layer utilizes eight attention heads, while the second GAT layer employs two attention heads, with the outputs from these heads averaged. All DL models use the IFL loss function and the Adam optimization algorithm. The pruning parameter for the GS-STGCN model is set to 0.3. The performance results of the five models are presented in Table 1.

The test results indicate that, as a shallow model, the SVM is limited to capturing only surface-level features of the data and struggles to identify deeper, more complex patterns. Consequently, its performance is notably inferior to that of the GCN, GAT, and LSTM models. By effectively leveraging the topological information of the power grid, both the GCN and GAT models achieve accuracies of 98.62% and 98.87%, respectively, outperforming the LSTM model. However, the proposed GS-STGCN model surpasses the GCN, GAT, LSTM, and SVM models across all four evaluation metrics. This superior performance can be attributed to the GS-STGCN model's capability to comprehensively capture the complex topology of the power grid while simultaneously extracting temporal dynamics.

4.3. GS Performance Verification

To validate the effectiveness of the GS, it is applied independently to the STGCN, GCN, and GAT models, with the pruning parameter set to 0.3. Transient stability prediction is then conducted by comparing GS-STGCN with STGCN, GS-GCN with GCN, and GS-GAT with GAT. Each model is iterated for 100 rounds, and the results are analyzed in

terms of prediction accuracy and model training time. The comparison of GS effectiveness verification results is shown in Figure 4.
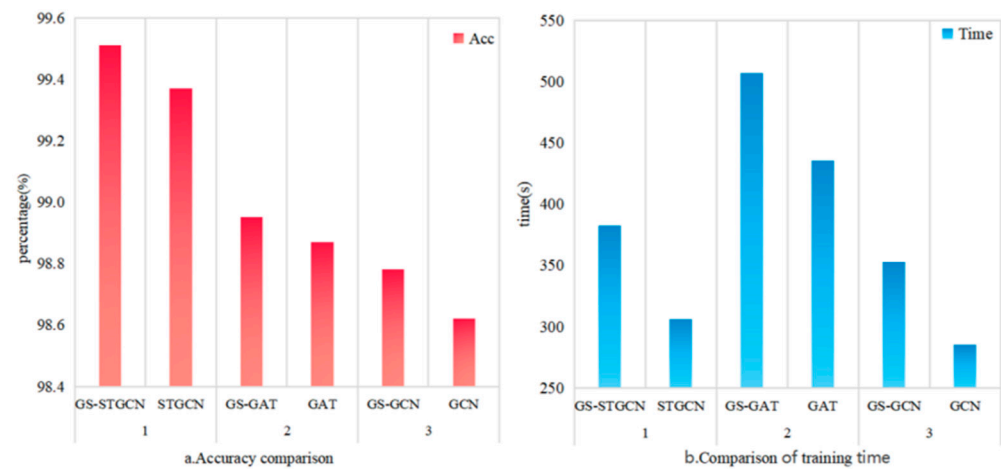

Figure 4. Comparison of GS effectiveness verification results.

Models employing the GS exhibit faster training speeds compared to their original counterparts. By reducing the complexity of topology and feature matrix of the power system, GS lowers both computational and storage requirements. Notably, the GS-GCN, GS-GAT, and GS-STGCN models improve their training speeds by 19.16%, 14.12%, and 20.08%, respectively, significantly accelerating the training process for TSA tasks. Additionally, the GS technique not only enhances task efficiency but also improves prediction accuracy. Specifically, the GS-GCN, GS-GAT, and GS-STGCN models increase their prediction accuracy by 0.16%, 0.08%, and 0.14%, respectively. This indicates that the GS technique effectively simplifies the power grid's topology and feature matrix by removing redundant information. It allows the models to focus on learning and analyzing critical information, thereby improving the precision and accuracy of the predictions.

4.4. Performance Verification of IFL Loss Function

The performance metrics of GS-STGCN on the test set, trained using the CE function, FL function, and IFL function as loss functions, are shown in Table 2. The accuracy of the model using the IFL function improves by 0.88% and 0.39% compared to the CE function and the FL function, respectively. The reason for this improvement is that the IFL function introduces an adaptive class weight factor compared to the CE function, which helps mitigate the impact of class imbalance between positive and negative samples. Compared to the FL function, the IFL function incorporates dynamic parameters to adjust the model's focus on difficult samples, leading to more effective utilization of the data during training.

Table 2. Comparison of loss function performance.

Function	A_{cc}/%	P_{rec}/%	R_{ec}/%	f_1/%
CE function	98.63	97.87	99.19	98.53
FL function	99.12	99.72	98.38	99.05
IFL function	99.51	99.98	98.92	99.46

4.5. Validation of the Robustness of the Model to Noise

In practical online applications, the test dataset is derived from real-time PMU data. Given the inherent variability in PMU measurements of dynamic data, some level of error is inevitable. To assess the robustness of the model, Gaussian white noise—characterized by instantaneous values following a Gaussian distribution and uniform power spectral density—is introduced to the test set to simulate the measurement errors typically encountered in real-world scenarios. The trained models are then evaluated using the test

set augmented with Gaussian white noise. The accuracy of various models under noisy conditions is displayed as Figure 5.
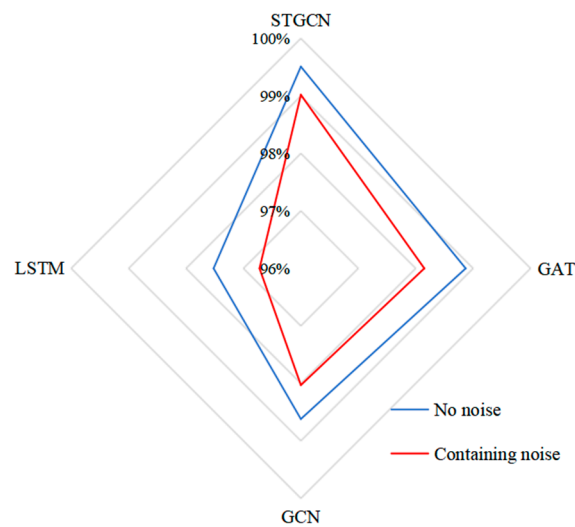

Figure 5. Accuracy of various models under noisy conditions.

It is observed that, compared to the noise-free scenario, each method's accuracy is somewhat impacted by the presence of noise. Nevertheless, the proposed GS-STGCN achieves a transient stability prediction accuracy of 99.02%, outperforming the other methods. Furthermore, the accuracy of the GS-STGCN decreases by only 0.49% in the presence of noise, which is a relatively smaller reduction compared to the accuracy losses seen in the other methods. Therefore, the proposed TSA method based on GS-STGCN exhibits strong robustness to noise.

5. Conclusions

In this paper, we develop a TSA method based on a spatiotemporal graph convolutional network with graph simplification. First, the topology and electrical characteristics of the power grid are integrated, and the graph data of the power grid is compressed and input into a deep spatiotemporal graph convolutional network. By combining the spatial features of the power grid with the temporal features of the transient process, a high-performance TSA model is constructed. Second, the improved focal loss function is employed as the loss function of the model. The method proposed in this paper was validated through simulation studies conducted on the IEEE 39-bus system. The experimental conclusions are as follows:

(1) The GS method optimizes the model input by removing low-influence nodes, thereby compressing the large-scale power grid input data. This approach effectively enhances the training speed of the TSA model.

(2) The STGCN model can fully capture the spatiotemporal features of transient processes, effectively reflecting the changing patterns of the system's transient behavior and ensuring the accuracy of the model's evaluation.

(3) The FL function can adaptively adjust the weight factors of the loss function, allowing the model to handle sample imbalance during training. This reduces the misclassification of unstable samples while also enhancing overall accuracy.

Currently, the TSA model proposed in this paper primarily targets offline and online analysis scenarios based on simulation analysis. It effectively shortens model training time and improves evaluation accuracy. This advancement is particularly beneficial for utility companies and power system operators, as it allows them to conduct more rapid and reliable stability assessments without significant computational overhead. In future research, greater emphasis will be placed on integrating the mechanistic knowledge of

power systems into the model to further enhance its interpretability and convergence. This integration will not only improve the model's predictive capabilities but also provide deeper insights into the underlying dynamics of power systems, which can be crucial for operational decision making and strategic planning.

Author Contributions: Conceptualization, D.Z. and B.S.; methodology, D.Z.; software, D.Z. and Y.Y.; validation, D.Z., B.S. and T.W.; formal analysis, Y.Y.; investigation, M.C.; resources, M.C.; data curation, B.S.; writing—original draft preparation, D.Z.; writing—review and editing, B.S. and Y.Y.; visualization, Y.Y.; supervision, T.W.; project administration, T.W. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement: The data presented in this study are openly available in yydata at https://github.com/yangyuan123654/yydata (accessed on 12 October 2024).

Conflicts of Interest: Authors Dan Zhang and Min Cheng were employed by the company Yunnan Electric Power Dispatching and Control Center. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Kang, C.; Yao, L. Key Scientific Issues and Theoretical Research Framework for Power Systems with High Proportion of Renewable Energy. *Autom. Electr. Power Syst.* **2017**, *41*, 2–11.
2. Su, T.; Liu, Y.; Shen, X. Evolutionary Algorithm for Power System Transient Stability Prevention Control Driven by Deep Learning. *Proc. CSEE* **2020**, *40*, 3813–3824.
3. Ma, R.; Zhan, M.; Jiang, K.; Liu, D.; Hu, P.; Cheng, S. Transient Stability Assessment of Renewable Dominated Power Systems by the Method of Hyperplanes. *IEEE Trans. Power Deliv.* **2023**, *38*, 3455–3468. [CrossRef]
4. Musca, R.; Gonzalez-Longatt, F.; Gallego Sánchez, C.A. Power System Oscillations with Different Prevalence of Grid-Following and Grid-Forming Converters. *Energies* **2022**, *15*, 4273. [CrossRef]
5. Shair, J.; Li, H.; Hu, J.; Xie, X. Power system stability issues, classifications and research prospects in the context of high-penetration of renewables and power electronics. *Renew. Sustain. Energy Rev.* **2021**, *145*, 111111. [CrossRef]
6. Peng, Q.; Jiang, Q.; Yang, Y.; Liu, T.; Wang, H.; Blaabjerg, F. On the stability of power electronics-dominated systems: Challenges and potential solutions. *IEEE Trans.* **2019**, *55*, 7657–7670. [CrossRef]
7. Tielens, P.; Van Hertem, D. The relevance of inertia in power systems. *Renew. Sustain. Energy Rev.* **2016**, *55*, 999–1009. [CrossRef]
8. Al Araifi, S.M.; Moursi, M.S.E.; Djouadi, S.M. Individual Functions Method for Power System Transient Stability Assessment. *IEEE Trans. Power Syst.* **2020**, *35*, 1264–1273. [CrossRef]
9. Wang, N.; Duan, K.; Sun, A. Time–Power Series-Based Semi-Analytical Approach for Power System Simulation. *IEEE Trans. Power Syst.* **2019**, *34*, 841–851. [CrossRef]
10. Wang, T.; Wang, X.; Liu, G.; Wang, Z.; Xing, Q. Neural Networks Based Lyapunov Functions for Transient Stability Analysis and Assessment of Power Systems. *IEEE Trans. Ind. Appl.* **2023**, *59*, 2626–2638. [CrossRef]
11. Wu, H.; Li, J.; Yang, H. Research Methods for Transient Stability Analysis of Power Systems under Large Disturbances. *Energies* **2024**, *17*, 4330. [CrossRef]
12. Yu, J.J.Q.; Hill, D.J.; Lam, A.Y.S.; Gu, J.; Li, V.O.K. Intelligent Time-Adaptive Transient Stability Assessment System. *IEEE Trans. Power Syst.* **2018**, *33*, 1049–1058. [CrossRef]
13. Chen, Q.; Lin, N.; Wang, H. Transient stability assessment model with parallel structure and data augmentation. *Int. Trans. Electr. Energy Syst.* **2021**, *31*, 12872. [CrossRef]
14. Su, T.; Liu, Y.; Zhao, J.; Liu, J. Deep Belief Network Enabled Surrogate Modeling for Fast Preventive Control of Power System Transient Stability. *IEEE Trans. Ind. Inform.* **2022**, *18*, 315–326. [CrossRef]
15. Zhao, K.; Shi, L. Transient stability assessment of power systems based on improved one-dimensional convolutional neural networks. *Power Syst. Technol.* **2021**, *45*, 2945–2957.
16. Zhu, L.; Hill, D.J.; Lu, C. Intelligent Short-Term Voltage Stability Assessment via Spatial Attention Rectified RNN Learning. *IEEE Trans. Ind. Inform.* **2021**, *17*, 7005–7016. [CrossRef]
17. Sun, L.; Bai, J.; Zhou, Z. Transient Stability Assessment of Power Systems Based on Bidirectional Long Short Term Memory Network. *Autom. Electr. Power Syst.* **2020**, *44*, 64–72.
18. Li, X.; Liu, S.; Gao, K. Power system transient stability assessment based on bidirectional long short term memory network and convolutional neural network. *Sci. Technol. Eng.* **2020**, *20*, 2733–2739.
19. Zhang, R.; Yao, W.; Shi, Z.; Ai, X.; Tang, Y.; Wen, J. Towards Multi-Scenario Power System Stability Analysis: An Unsupervised Transfer Learning Method Combining DGAT and Data Augmentation. *IEEE Trans. Power Syst.* **2023**, *38*, 5367–5380. [CrossRef]

20. Huang, J.; Guan, L.; Guo, M. Application and Prospects of Graph Deep Learning Technology in Intelligent Transient Stability Assessment. *Power Syst. Technol.* **2023**, *47*, 1500–1511.
21. Huang, J.; Guan, L.; Su, Y. A topology adaptive high-speed transient stability assessment scheme based on multi-graph attention network with residual structure. *Int. J. Electr. Power Energy Syst.* **2021**, *130*, 106948. [CrossRef]
22. Zhong, Z.; Guan, L.; Su, Y. Transient Stability Assessment of Power Systems Based on Graph Attention Deep Network. *Power Syst. Technol.* **2021**, *45*, 2122–2130.
23. Hu, W.; Zheng, L.; Min, Y. Research on Transient Stability Assessment of Power Systems after Faults Based on Deep Learning. *Power Syst. Technol.* **2017**, *41*, 3140–3146.
24. Li, X.; Li, W.; Zhao, Z. Transient Stability Assessment of Power System Based on Spatio-temporal Information fusion. *Proc. CSU-EPSA* 2024, *in press*.
25. Jiang, T.; Dong, Y.; Wang, C.; Chen, H.; Li, G. Transient Voltage Stability Assessment of Receiving-end Power System Based on Graph Convolution and Bidirectional Long/Short-term Memory Networks. *Power Syst. Technol.* **2023**, *47*, 4937–4951.
26. Li, X.; Liu, C.; Xin, S. Analysis of the coupling mechanism and evaluation index of coupling strength between transient power angle stability and transient voltage stability. *Proc. CSEE* **2021**, *41*, 5091–5107.
27. Zhuang, Y.; Xiao, T.; Cheng, L.; Chen, Y.; Guan, H. Transient Stability Assessment of Power System Based on Spatio-Temporal Graph Convolutional Network. *Autom. Electr. Power Syst.* **2022**, *46*, 11–18.
28. Zhou, J.; Hang, X.; Hao, L. Challenges and Prospects for New Power System Network Security Protection. *Autom. Electr. Power Syst.* **2023**, *47*, 15–24.
29. Li, X.; Yang, Z.; Guo, P.; Cheng, J. An Intelligent Transient Stability Assessment Framework with Continual Learning Ability. *IEEE Trans. Ind. Inform.* **2021**, *17*, 8131–8141. [CrossRef]
30. Gu, X.; Shao, Y. A graph neural network based on influence pruning for fast computation of graph simplification. *Comput. Sci.* **2023**, *50*, 52–58.
31. Dauphin, Y. Language Modeling with Gated Convolutional Networks. In Proceedings of the 34th International Conference on Machine Learning (ICML), Sydney, Australia, 6–11 August 2017; Volume 70, pp. 933–941.
32. Gomez, F.R.; Rajapakse, A.D.; Annakkage, U.D.; Fernando, I.T. Support Vector Machine-Based Algorithm for Post-Fault Transient Stability Status Prediction Using Synchronized Measurements. *IEEE Trans. Power Syst.* **2011**, *3*, 1474–1483. [CrossRef]