Multi-task Transient Contingency Screening with Temporal Graph Convolutional Network in Power Systems

Yinsheng Su1, Jiyu Huang2, Haicheng Yao1, Lin Guan2,3*, Mengxuan Guo2 and Zhi Zhong2

1 Department of Power System Operation, China Southern Grid Co., Ltd.
2 School of Electric Power, South China University of Technology, Guangzhou 510641, P.R. China
3 Guangdong Provincial Key Laboratory of Intelligent Operation and Control for New Energy Power System, Guangzhou 510663, China

*Corresponding author email: lguan@scut.edu.cn

Abstract. Rapid transient stability assessment (TSA) is an essential requirement for power system security. In real-world applications, transient contingency screening (TCS) applies TSA approaches to address the pre-defined contingency sets under online operation conditions. TSA by time domain simulation (TDS) is time-consuming, hence we propose a high-speed temporal graph convolutional network (TGCN) that achieves TSA decisions such that a large-scale contingency set can be scanned quickly with enough precision. Based on multi-graph inputs to reflect the transient process, the TGCN utilizes the graph convolutional network (GCN) to extract topology representations and temporal convolution (TC) layers to encode temporal relations. After above graph embedding, two downstream networks are designed for stability classification and critical generator prediction, respectively. Test results on IEEE 39 Bus system demonstrate its superiority over existing models under different operation topologies, fault locations and clearing modes.

Keywords: Temporal graph convolutional network (TGCN); Temporal convolution (TC); Transient contingency screening (TCS); Critical generator prediction.

1. Introduction
During daily operation in power system, the increasing expansion of interconnected grids and the large-scale penetration of low-inertia renewable energy cause more stressed operation conditions. To capture the risks owing to the changeable conditions, transient stability assessment (TSA) provides an early warning as well as guidance for pre-fault prevention control. It is triggered every 15min to screen the pre-defined contingency set in real-world power systems, which is named transient contingency screening (TCS) in this paper with the aim of implementing TSA on very large contingency sets. Conventional TSA depends on time domain simulation (TDS) [1], which is time-consuming and difficult to handle increasingly complex power system models in real time operation. Data-driven machine learning (ML) benefits from flexibility, fast speed and its ability to predict stability margins such that it serves as a supplement to the TDS. The application of shallow networks has drawn wide attention, while they need manual feature selection and suffer from poor generalization[2]. To deal with the information loss, major researchers have to feed the models with post-fault dynamics. In recent years, deep learning (DL) methods characterized by high-performance computing and big data are introduced in TSA. On one hand, stacked auto encoder (SAE) [3] and long short term memory (LSTM) [4-5] automatically reduces the overfitting risk of the model through pre-training and temporal...

Content from this work may be used under the terms of the Creative Commons Attribution 3.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

Published under licence by IOP Publishing Ltd
parameter sharing respectively. On the other hand, convolutional neural network (CNN) [6-7] avoids significant parameter increases via convolution kernels of space sharing, which allows its ability to face large-scale inputs. Their advantage of automatic feature aggregation reduces the time span covered by the dynamic information to the moment of fault clearance. However, preset TDS for the dynamics is still a bottleneck for the speed enhancement, unless they acquire more information from the operation topologies to reflect the changes of operation conditions or fault locations. Regarding that the power system can be treated as a graph when buses are denoted as nodes and transmission lines are denoted as edges[8]. A natural solution complementary to the gap is the graph learning approaches [9]. More recently, graph convolutional networks (GCNs) are reported in power system applications, such as fault location prediction, load shedding decision and missing data recovery [10-12]. Huang et al. [13] verify the effectiveness of GCN in topology learning for multiple TSA tasks. Therefore, we follow GCN and propose a temporal graph convolutional network (TGCN) to fulfill two TCS tasks, i.e. stability classification as well as critical generator prediction. The design objective is that a well-trained TGCN can cover various operation topologies, fault locations. A temporal convolution (TC) module help capture the time-evolving nature of the transient process. The main contributions are summarized as follows:

1) Our model can make TSA decisions only based on the electrical variables at steady-state (t0-) and the fault occurrence (t0+) with the fault information.
2) The multi-task TGCN model can fulfill both stability category and critical generator prediction under the scenarios with different fault locations, operation topologies and clearing modes (i.e., a fault cleared with or without line tripping).

2. Proposed Methodology

![Figure 1. The structure of TGCN.](image)

As figure 1, the whole TGCN contains graph embedding and downstream networks. Three graphs concerning the transient process are required before the representations are captured by utilizing GCN and TC iteratively. Two multilayer perceptron (MLP), the stability category predictor (SCP) and critical generator predictor (CGP) transform these expressive representations into task-specific outputs. Assume the superscript "l" as the times a graph passes GCN and TC while the subscript "m" as the graph index.

2.1. Graph Convolutional network(GCN)

As figure 2(a), CNN performs neighborhood aggregating on inputs that typically denote images or signals in Euclidean space, which are difficult in addressing graph data attributed to the irregularity of node connections. By contrast, graph convolutional filters in figure 2(b) concentrate more on the correlation over graph edges provided by the adjacency matrix. Let $\mathcal{G}_m = (\mathcal{V}_m, \mathcal{E}_m), m = 0-, 0+, s$ denotes an undirected graph where $\mathcal{V}_m \in \mathbb{R}^N$ denotes the node set and $\mathcal{E}_m \in \mathbb{R}^L$ denotes the edge set. A graph $\mathcal{G}_m, m = 0-, 0+, s$ is characterized by its adjacency matrix $A_m$ and node feature matrix $X_m$. The degree matrix $D_m$ is derived from $A_m$ with its nonzero elements $d_m(i,j) = \sum j a_m(i,j)$. Given input node feature matrix $H_m^{(i)} \in \mathbb{R}^{N \times C^{(i)}}$ and the output one $H_m^{(l)} \in \mathbb{R}^{N \times C^{(l)}}$ with $C^{(i)}$, $C^{(l)}$ as the input and output size, the feed-forward propagation of GCN [9] with a skip connection is expressed as
\[ H^{(t)}_m = \sigma(D^{(t)-1}A^{(t)}H^{(t)}_mW^{(t)}_s) + H^{(t)}_mW^{(t)}_s \]  

(1)

where \( W^{(t)}_s \in \mathbb{R}^{C_t \times C_t} \) denotes the parameter matrix, while the scaling parameters \( W^{(t)}_s \) guarantee the consistency of the inputs and outputs.

2.2. Temporal Convolution (TC)

To encode the temporal correlation, we introduce a temporal convolution that works on the same nodes in different graphs. For the \( i^{th} \) nodes, the inter-graph propagation is as

\[ h^{(t)}_m = \sigma(\sum_{j(\in m)} h^{(t)}_j W^{(t)}_s) \]

(2)

where the convolved feature \( h^{(t)}_m = H^{(t)}_mW^{(t)}_s \in \mathbb{R}^{e(t \times c(t)} \) is calculated with temporal convolutional parameters \( W^{(t)}_s \in \mathbb{R}^{e(t \times c(t)} \).

2.3. Downstream networks

Denote the outputs of SCP as \( z^{SCP} \). It is addressed by the "softmax" operation as

\[ c^{SCP}_i = \exp(z_i^{SCP}) / \sum_i \exp(z_i^{SCP}) \]

(3)

where \( c^{SCP}_i \) is the confidence of the \( i^{th} \) category. The system is predicted as stable if \( c^{SCP}_i < 0.5 \) or otherwise unstable.

Denote the outputs of CGP as \( z^{CGP} \). It is activated by the sigmoid function as

\[ c^{CGP}_i = (1 + \exp(-z^{CGP}_i))^{-1} \]

(4)

where \( c^{CGP}_i \) is the stability confidence of the \( i^{th} \) generator. CGP considers the generator stable if \( c^{CGP}_i < 0.5 \) or otherwise unstable.

3. The Training and Evaluation of TGCN

3.1. Multi-graph Inputs and Their Labels

The element \( a_{m(i,j)} \) in \( A_m \in \mathbb{R}^{N \times N} \) is defined as

\[ a_{m(i,j)} = \begin{cases} 0 & (V_{m,i}, V_{m,j}) \notin E_m \\ \left| y_{m(i,j)} \right| / \left| y_{m(i,j)} \right|_{\text{max}} & (V_{m,i}, V_{m,j}) \in E_m \end{cases} \]

(5)

where \( \left| y_{m(i,j)} \right| \) and \( \left| y_{m(i,j)} \right|_{\text{max}} \) denote the module and the maximum module of the element in the power system admittance matrix at the \( m^{th} \) snapshot.

Each row in \( x_m(i,j) = X_m \in \mathbb{R}^{N \times 5} \) refers to a vector about system states at node \( i \), including the voltage amplitude, active and reactive power flow to loads, the active and reactive injected by generators.

The stability status of the system is practically obtained by the transient stability index \( \eta \):

\[ \eta = (180^\circ - |\Delta \delta|_{\text{max}})/180^\circ + |\Delta \delta|_{\text{max}} \]

(6)
where \( \Delta \delta_{\text{max}} \) is the absolute value of the maximum rotor angle of separation between any two generators during the simulation time. When \( \Delta \delta_{\text{max}} > 180^\circ \), i.e., \( \eta < 0 \), we label this unstable sample as the vector \( c = [1, 0]^T \). Otherwise, it is considered as a stable one and labeled with \( c = [0, 1]^T \).

Let \( N_G \) denote the number of generators, then we define \( G \) as the set of generators and \( c \in \mathbb{R}^{N_G} \) is a binary vector that denotes the stability of all the generators. The status of a generator \( i \) is expressed as:

\[
c(i) = \begin{cases} 
0 & |\Delta \delta| > 180^\circ \\
1 & |\Delta \delta| \leq 180^\circ 
\end{cases}
\]

where \( |\Delta \delta| \) is the absolute value of separation between generator \( i \) and the reference generator during the simulation time. The set of critical generators \( G_c = \{ i \in G | c(i) = 0 \} \) is a subset of \( G \) and in particular, \( G_c \) is an empty set if and only if the system is transient stable.

### 3.2. Loss Functions

The loss function \( \mathcal{L} \) contains the functions for SCP, CGP and a regularization item \( \mathcal{L}_r = \| \Theta \|^2 \):

\[
\mathcal{L} = \beta_c \mathcal{L}_{SCP} + \beta_c \mathcal{L}_{CGP} + 0.0005 \mathcal{L}_r
\]

Here \( \mathcal{L}_{SCP} \) is the cross-entropy function for binary classification

\[
\mathcal{L}_{SCP} = -\sum_i \sum_j c_{bj} \log \tilde{c}_{bj} / B
\]

where \([c_{bj}, c_{bj}]\) is the annotated categories while \([\tilde{c}_{bj}, \tilde{c}_{bj}]\) denotes the confidence of the \( b \)th sample on the training set whose total number is \( B \). \( \mathcal{L}_{CGP} \) is the cross-entropy function for multi-label classification as

\[
\mathcal{L}_{CGP} = -\sum_i \sum_j (c_{bj} \log \tilde{c}_{bj} + (1 - c_{bj}) \log (1 - \tilde{c}_{bj})) / B
\]

The optimization object is to minimize the constraints in (8). During the optimization, the learnable parameters of GCN, TC layers as well as SCP and CGP are adjusted.

### 3.3. Performance Metrics

The specific metrics for the SCP involve accuracy (ACC), alarm (MA) rate and false alarm (FA) rate.

| Actual unstable | Actual stable |
|-----------------|---------------|
| Predicted unstable | TN | FN |
| Predicted stable | FP | TP |

\[
\text{ACC} = (TP + TN)/(TP + FP + FN + TN)
\]

\[
\text{MA} = FP/(TN + FP)
\]

\[
\text{FA} = FN/(FN + TP)
\]

Considering that CGP predicts the set of critical generators, we introduce the Jaccard similarity to calculate the distance between the sets. Given any two sets \( s_i, s_j \in \mathbb{N} \), Jaccard similarity is defined as:

\[
J(s_i, s_j) = \frac{|s_i \cap s_j|}{|s_i \cup s_j|} = \frac{|s_i \cap s_j|}{|s_i| + |s_j| - |s_i \cap s_j|}
\]

where \( J \in [0,1] \) and \( J(s_i, s_j) = 1 \). Here, we consider the sample correct only when \( J(\tilde{G}_c, G_c) = 1 \). Similar to ACC, we define Jaccard accuracy (JACC) of all samples.
4. Case Study

4.1. Test System and Model Settings
The studies on proposed scheme is conducted on the IEEE 39 Bus system with 39 buses, 10 generators, 19 loads and 46 transmission lines[13]. The operation states with all transmission lines on service are "Base" cases, while the "N-1" and "N-2" cases are generated by randomly switching off transmission line(s). Besides, the loads are changed based on the basic load level (from 75% to 120%). Set three-phase faults at either end of any transmission lines and clear the faulted line with or without tripping after 0.1s. The stability labels are calculated during TDS of 4s. The "Base" and "N-1" cases form the training set (60%), while the "N-2" cases form the validation (20%) and test set (20%). The best model settings are shown in figure 1.

4.2. Comparison with Baseline Methods

![Figure 3. The performance of the proposed model compared with baselines.](image)

| Model   | ACC(%) | MA(%) | FA(%) | JACC (%) |
|---------|--------|-------|-------|----------|
| SVM     | 90.62  | 7.95  | 9.71  | 67.34    |
| ANN     | 92.97  | 8.06  | 6.79  | 89.20    |
| SAE     | 93.83  | 7.77  | 5.81  | 90.17    |
| CNN     | 92.81  | 11.87 | 6.36  | 88.36    |
| Proposed| 97.49  | 4.97  | 1.96  | 94.82    |

We choose another four methods as baselines. The support vector machine (SVM) and artificial neural network (ANN) are typical shallow networks, while SAE and CNN utilize different DL techniques. In figure 3 and table 2, the size of a circle refers to the values of the metrics. It is clear that SVM performs worst in both tasks. CNN and SAE have a similar performance with ANN, owing to their poor ability to discriminate topology changes. Our TGCN beats all the baselines on account of its excellent superiority in both graphical and temporal learning. The ACC reaches about 97.5%, while JACC is up to 94.8%.

![Figure 4. The verification of TC layer.](image)

4.3. Verification of TC Layer
In this subsection, we compared the proposed model with one without TC layer. As figure 4 shows, the TC layer enables higher accuracy in both tasks, where more improvement is observed in the critical generator prediction, almost 0.9% in JACC. It means the expressive temporal representations can help the model deal with the fine-grained task.
4.4. Training and Inference Efficiency

Table 2. The model efficiency.

| Training time (s) | Inference time(s/batch) | TDS time (s/batch) |
|-------------------|-------------------------|-------------------|
| 196.2             | 0.045                   | 360.96            |

The efficiency of a model is also a key factor to be considered. We list the training and inference time in table 2. Obviously, our model exploits the parameter sharing mechanism such that the optimization time consumption is relatively low within 4min. The online inference is far more satisfactory in contrast with TDS. The former finishes a batch in 0.1s whereas the latter needs over 6min. This validates the superior efficiency of our model over TDS in real-time TCS tasks.

5. Conclusion

In this paper, we introduce a TGCN scheme to address two TCS tasks, stability classification and critical generator predictions based on merely electrical information at $t_0$ and $t_{0+}$. The popular GCN approach is utilized to achieve graph learning, while a new TC layer aims at encoding the time-evolving nature of the transient process. Two downstream networks share the same graph embedding module and make inference parallelly. Test results on the IEEE 39 Bus system demonstrate the outstanding performance of TGCN in scenarios with different topology disturbances, fault locations or even clearing modes.

Acknowledgments

This work is supported by the China Southern Power Grid Research Project (ZDKJXM20180084) and National Natural Science Foundation (52077080).

References

[1] Kundur P, Balu N J, Lauby M G. Power system stability and control[M]. New York: McGraw-hill, 1994.
[2] Sobbouhi A R, Vahedi A. Transient stability prediction of power system; a review on methods, classification and considerations[J]. Electric Power Systems Research, 2021, 190:106853.
[3] Yi A, Mz B, JJ A, et al. Intelligent online catastrophe assessment and preventive control via a stacked denoising autoencoder - ScienceDirect[J]. Neurocomputing, 2020, 380:306-320.
[4] Yu J, Hill D J, Lam A Y S, et al. Intelligent Time-Adaptive Transient Stability Assessment System[J]. IEEE Transactions on Power Systems, 2016, PP(99):1-1.
[5] Yu J, Hill D J, Lam A. Delay aware transient stability assessment with synchrophasor recovery and prediction framework[J]. Neurocomputing, 2018, 322.
[6] Gupta A, Gurrala G, Sastry P S. An Online Power System Stability Monitoring System Using Convolutional Neural Networks[J]. IEEE Transactions on Power Systems, 2018.
[7] Zhu L, Hill D, Lu C. Hierarchical Deep Learning Machine for Power System Online Transient Stability Prediction[J]. IEEE Transactions on Power Systems, 2020.
[8] Ishizaki T, Chakrabortty A, Imura J I. Graph-Theoretic Analysis of Power Systems[J]. Proceedings of the IEEE, 2018, 106(5):931-952.
[9] Kip F T N, Welling M. Semi-Supervised Classification with Graph Convolutional Networks[C]/International Conference on Learning Representations. 2017.
[10] Chen K, Hu J, Zhang Y, et al. Fault Location in Power Distribution Systems via Deep Graph Convolutional Networks[J]. IEEE Journal on Selected Areas in Communications, 2020, 38(1):119-131.
[11] Kim C, Kim K, Balaprakash P, et al. Graph Convolutional Neural Networks for Optimal Load Shedding under Line Contingency[C]/ 2019 IEEE Power & Energy Society General Meeting (PESGM). IEEE, 2019.
[12] Yu J, Hill D J, Li V, et al. Synchrophasor Recovery and Prediction: A Graph-Based Deep Learning Approach[J]. IEEE Internet of Things Journal, 2019, PP(99):1-1.
[13] Huang J, Guan L, Su Y, et al. Recurrent Graph Convolutional Network-Based Multi-Task Transient Stability Assessment Framework in Power System[J]. IEEE Access, 2020, PP(99):1-1.