Predicting Power System Frequency Health Index with PMUs and Graph Attention Networks

Koji Yamashita and Nanpeng Yu
Electrical and Computer Engineering
University of California Riverside
Riverside, U.S.A.

Evangelos Farantatos and Lin Zhu
Transmission Operations and Planning R&D Group
Electric Power Research Institute
Knoxville, U.S.A.

***Abstract*—This paper develops a power system frequency health index prediction model using the graph attention network (GAT). A range of pre-contingency grid operating conditions at different loading levels and diverse contingency information serve as training inputs to infer the grid health in terms of system frequency. The frequency nadir is leveraged for the frequency health indicator and its corresponding labels for training samples. The dynamic attention mechanism is incorporated in the graph convolutional network (GCN) model to learn the varying correlation levels between substations/buses effectively. The extended IEEE 118-bus model is adopted to evaluate the frequency health index prediction performance of the proposed GAT and baseline methods. The prediction accuracy of the GAT is 12% higher than that of the baseline method. It is demonstrated that the GAT model is much more robust than the GCN model against erroneous phasor measurement unit (PMU) measurements and partial observability from PMUs.**

***Index Terms*—component, formatting, style, styling, insert**

I. Introduction

The escalating threat of extreme weather events, such as hurricanes and floods, and the increasing risk of cyber-attacks underscores the importance of enhancing power grid resilience. This calls for advanced real-time contingency analysis tools, which complement Energy Management Systems (EMS), and can be accessed in control rooms. Nevertheless, the development of real-time contingency analysis and system health assessment tools faces challenges due to the expanding influence of renewable energy sources, introducing variability and complexity into grid management. The growing importance of stability issues initiated by cyberattacks along with subsequent islanding and reconnection, is another significant technical challenge.

Mainly because the physical grid topology and the branch power flow information may be directly leveraged in the learning process, the graph neural network model has been actively explored in the power system research studies [1]. One track of research is to predict the relation between the contingency scenario and the stability of power systems. Another track of research aims at predicting the dynamic behavior or transient trajectories of the system. Most of the existing research primarily formulates a classification problem and tries to identify whether a power system is stable or unstable. Reference [2] proposes a data-driven transient stability assessment model using the graph convolutional model with a multi-pooling mechanism. Both max pooling and mean pooling are leveraged to improve the stability prediction performance. This model may be used for dynamic security assessment purposes. Reference [3] develops a graph attention model that can predict the dynamic response of synchronous generators using real-time system measurements. The developed model is capable of predicting future dynamic behaviors of the power grid with the input of the first 10-step dynamic response following disturbances. This model may be used for real-time power system control purposes.

Graph neural network models have also been adopted for assessing the other stability health indices. A graph neural network model was developed to predict short-term voltage health index [4], [5]. However, the frequency health index has barely been studied [1], [6]. A key frequency indicator in contingency analysis is the frequency nadir, i.e., the lowest point that the system frequency reaches during a disturbance. Earlier investigations [2], [3] dealt with short-circuit faults only as a contingency scenario, although such faults can potentially impact both transient stability and frequency instability.

This paper delves deeper into this area by developing a specialized graph neural network tailored for analyzing frequency stability following short-circuit faults alongside scenarios involving typical generator or load disconnections. Power system measurements, such as the phasor measurement unit (PMU) data, are not always accessible at all power stations and substations. As a result, the developed machine learning model must demonstrate exceptional robustness when predicting the grid health index using limited measurement data. By developing a graph learning-based system frequency health index prediction algorithm, this paper could enhance not only the system operators' situational awareness but also system resilience.

The main technical contributions of this paper are:

- The development of a graph attention network (GAT)-based frequency health index prediction algorithm by leveraging the topology of the power transmission network, which attains superior accuracy compared to the state-of-the-art health index prediction algorithms.

This project was sponsored by Electric Power Research Institute under agreement 10015026.

- Establishment of comprehensive training and testing dataset with realistic power system configurations, operating conditions, and contingency scenarios using refined time-domain simulation models.
- The proposed GAT-based algorithm demonstrated remarkable reliability and robustness in partially observable grids.

The rest of the paper is organized as follows. Section II presents the proposed graph attention neural network for frequency health index prediction. Section III demonstrates the proposed algorithm in the numerical study using a large-scale IEEE testing system. Section IV concludes the paper.

II. Graph Neural Network Model for Frequency Health Index Prediction

A. Frequency Health Indices and the Corresponding Labels

The frequency health index is derived using the frequency nadir as the lowest frequency level before reaching the post-steady-state frequency, as shown in Fig. 1. The frequency health index is segmented into 5 different categories, with each category defined by specific frequency ranges, as shown in Table I. These categories will serve as the class labels used as the output of the graph neural network. The ordinal encoder [4] is integrated into the output layer of the graph neural network to introduce inductive bias in the learning process by incorporating prior knowledge of ordinal relations among categories. Since grid frequency is generally a global indicator, the frequency health index prediction task is formulated as a graph classification problem.

B. Frequency Health Index Prediction Model Structure

The graph attention network (GAT) [7], [8] is selected as the deep learning model for the frequency health index prediction task for several reasons. As mentioned earlier, while the graph convolutional network (GCN) is highly effective in performing supervised learning on graph-structured data, integrating edge weights into the training process can present challenges. In contrast, the GAT model can effectively learn the critical lines with the strength of inter-correlation between substations/buses. Therefore, the GAT extends the capabilities of the GCN by enhancing its edge-related functionality by incorporating the attention mechanism. The overall architecture of the proposed GCN/GAT model and a *baseline*, multilayer perceptron (MLP) model are shown in Fig. 2. Unlike the GCN/GAT, the MLP cannot directly handle edge features.

There are two categories of inputs to the grid health prediction model. The first category of inputs represents the pre-fault steady-state system information associated with the buses and branches. The second category of inputs is related to the contingency information, which corresponds to the 11-th and 12-th node features in Table II. The voltage drop feature is treated as a synthetic value. For system faults occurring in the middle of a transmission line, the feature is set to 0.5. When the fault occurs inside a transformer, the feature is set to 0.0. Otherwise, it is set at 1.0. The fault duration is normalized to 200 ms, meaning that, for example, 50 ms of the fault duration is represented as 0.25. Note that all the node features, except the 11-th and 12-th features, and all the edge features as listed in Table II, are assumed to be collected from PMUs.

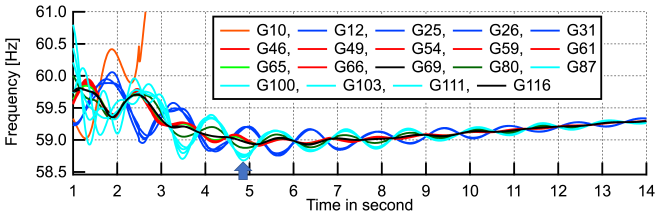

Fig. 1. Example frequency response following a transformer outage at bus 89 in IEEE 118-bus system, indicating its nadir using blue arrow near 5 s.

C. GAT-based Frequency Health Index Prediction Model

Inputs of the GCN/GAT model are first passed through the edge-conditioned convolution (ECC) layer, which updates each node i's feature vector from x_i to x'_i. This is done by activating the sum of the product of the weight matrix and the feature vector and the aggregation of nodal and feature vectors via a parameterized neural network, as shown in (1).

$$\boldsymbol{x}'_i = \mathrm{ReLU}\Big[\boldsymbol{W}\boldsymbol{x}_i + \sum\nolimits_{j\in\mathcal{N}(i)} \boldsymbol{x}_j \cdot h_\Theta(\boldsymbol{e}_{i,j})\Big], \tag{1}$$

where $\boldsymbol{W}$ denotes the weight matrix, h_Θ denotes a neural network (e.g., a multilayer perceptron), $\boldsymbol{x}_i$ denotes the feature vector of the i-th node, and $\boldsymbol{e}_{i,j}$ denotes the edge feature vector for the edge from the source node, i, to the target node, j.

The output of the ECC layer is then used as the input of graph convolutional layers with/without the graph attention mechanism. The node representations $\boldsymbol{x}_i$ are updated at each graph convolutional layer according to (2).

$$\boldsymbol{x}'_i = \mathrm{ReLU}\Big[\widetilde{\boldsymbol{W}}^{\boldsymbol{T}}\sum\nolimits_{j\in\mathcal{N}(i)} \hat{A}_{j,i}\cdot \boldsymbol{x}_j \cdot C\Big],$$

$$C = \left(1+\sum\nolimits_{j\in\mathcal{N}(i)}\hat{A}_{j,i}\right)^{-\frac{1}{2}}\left(1+\sum\nolimits_{i\in\mathcal{N}(j)}\hat{A}_{i,j}\right)^{-\frac{1}{2}} \tag{2}$$

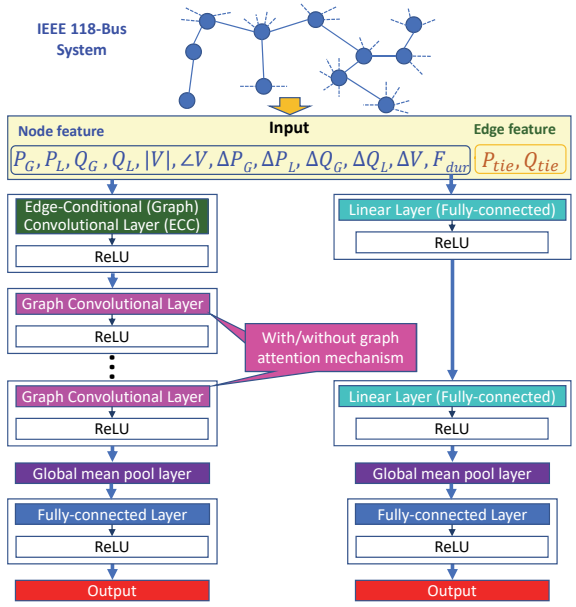

Fig. 2. GCN/GAT model (left) and MLP model (right).

TABLE I
LABELS SET FOR EACH FREQUENCY HEALTH INDEX

Class	Label 0	Label 1	Label 2	Label 3	Label 4
Frequency Range (Hz)	(59.975, ∞]	(59.95, 59.975]	(59.85, 59.95]	(59.70, 59.85]	($-\infty$, 59.7]
Number of Contingencies	13,217	2,581	3,309	3,402	2,716

TABLE II
INPUTS OF HEALTH INDEX PREDICTION MODEL

	Node/Branch feature	Feature type
1	Active power output, P_G	Node
2	Reactive power output, Q_G	Node
3	Active power load, P_L	Node
4	Reactive power load, Q_L	Node
5	Voltage magnitude, $\|V\|$	Node
6	Voltage angle, $\angle V$	Node
7	Active power output deviation, ΔP_G	Node
8	Active power load deviation, ΔP_L	Node
9	Reactive power output deviation, ΔQ_G	Node
10	Reactive power load deviation, ΔQ_L	Node
11	Voltage drop (3 simplified levels), ΔV	Node
12	Fault duration, F_{dur}	Node
13	Active power transfer, P_{tie}	Branch
14	Reactive power transfer, Q_{tie}	Branch

where $\hat{A}_{i,j}$ denotes an ij-th element of $\hat{A} = A + I$, where A is the adjacency matrix. It is noted that $A_{i,j}$ equals 1 if node i is connected to node j and 0 otherwise.

The GAT model [7] uses a refined approach to analyze graph-structured data with the help of the attention mechanism, enabling it to concentrate on specific nodes and their interconnected relationships (represented by edges). The GCN model leverages pre-assigned *unity* edge weights between nodes, while the GAT model trains the edge weights dynamically using the attention mechanism. The nodal representations, x_i, are updated in each of the graph layers as shown below:

$$x_i' = \alpha_{i,i}\boldsymbol{W_s}x_i + \sum\nolimits_{j\in N(i),j\neq i} \alpha_{i,j}\boldsymbol{W_i}x_j \tag{3}$$

In (3), the weight matrix, $\mathbf{W}$, is separately expressed with $\mathbf{W}_s$ and $\mathbf{W}_i$ that correspond to the self-loop elements and other elements. The attention coefficient, $\alpha_{i,j}$ is calculated from:

$$\alpha_{i,j} = \text{softmax}(e_{i,j}) = \frac{e^{e_{i,j}}}{\sum_{k\in N_i} e^{e_{i,k}}} \tag{4}$$

$$e_{i,j} = a(\boldsymbol{W h_i}, \boldsymbol{W h_j}) = a^T \text{LeakyReLU}(\boldsymbol{W}\cdot[\boldsymbol{h_i}||\boldsymbol{h_j}]), \tag{5}$$

where $a \in \mathbb{R}^{2F'}$, $\boldsymbol{W} \in \mathbb{R}^{F'\times F}$, and $||$ denotes a vector concatenation. F denotes the number of features at each node. $\boldsymbol{h}$ denotes a set of node features. (5) proposed in [8] is the result of shifting a^T outside the original LeakyReLU($\cdot$) that is proposed in [7]. (5) enables the algorithm to handle the attention mechanism dynamically, unlike [6] which deals with the attention mechanism only in a static manner.

D. Ordinal Encoder

The labels in each health index can be described as an ordinal variable. For example, the difference between system health for labels 4 and 5 is much smaller than that of levels 4 and 1. Thus, ordinal encoding can be adopted to introduce prior knowledge (also known as inductive bias) that represents the natural ordering between the labels in each health index.

Algorithm 1 Classification with Ordinal Encoder

for every epoch **do**
  K logits L_i are derived at the last layer
  $s_i = sigmoid(L_i), \forall 1 \leq i \leq K$
  $L_i' = \sum_{j\leq i}\log(s_j) + \sum_{i<j}\log(1-s_j), \forall 1 \leq i \leq K$
  $p_i' =$softmax$(L_i'), \forall 1 \leq i \leq K$
end for

TABLE III
CONTINGENCY SCENARIO

Type	Equipment (number of cases)	Fault duration
Line outage	Transmission line (163)	50, 200, 367 ms
Transformer outage	Unit transformer (55) Load transformer (91) Tie transformer (9)	50. 200, 367 ms
Synchronous machine outage	Synchrnous generator (19) Synchronous condensers (36)	N/A (0 ms)

To implement the ordinal encoder, we take the output logits L_i from the last fully-connected layer and run it through Algorithm 1 to derive the probability that the health index belongs to each labels. Since the grid health index prediction task is formulated as a classification problem, the cross-entropy loss function is selected for gradient calculations.

III. NUMERICAL STUDY

A. Case Study Setup

1) Test System and Contingency Scenarios: The IEEE 118-bus system [9] is adopted as the fundamental testing environment (see Fig. 3). To make the system more realistic, step-up/down transformers and additional units at power stations are included, referring to [10]. These modifications led to a larger system with 264 buses and 325 branches. The hypothesized disturbances are N-1 contingencies, such as unit trips, load sheddings, transformer trips, and line trips, as shown in Table III. The fault durations are set to 50, 200, and 367 ms, considering the typical fault identification time with and without intentional delay at transmission lines and the breaker failure clearing time [11]. In total, there are 25,225 contingency cases (373 outages and 25 different loading profiles) for this test system. The entire cases are further split into 15,145 (60%), 2,520 (10%), and 7,560 (30%) cases for training, validation, and testing, respectively.

2) Scenarios of System Operating Conditions: The number of units at power stations and fuel types is inferred from [10] and [12] and illustrated in Table IV. The small capacity units are all treated as run-of-river hydropower facilities. 1-4 units are assumed to be connected to each power station. Based on [9], [10], the following steps are employed to generate a wide range of system operating conditions:

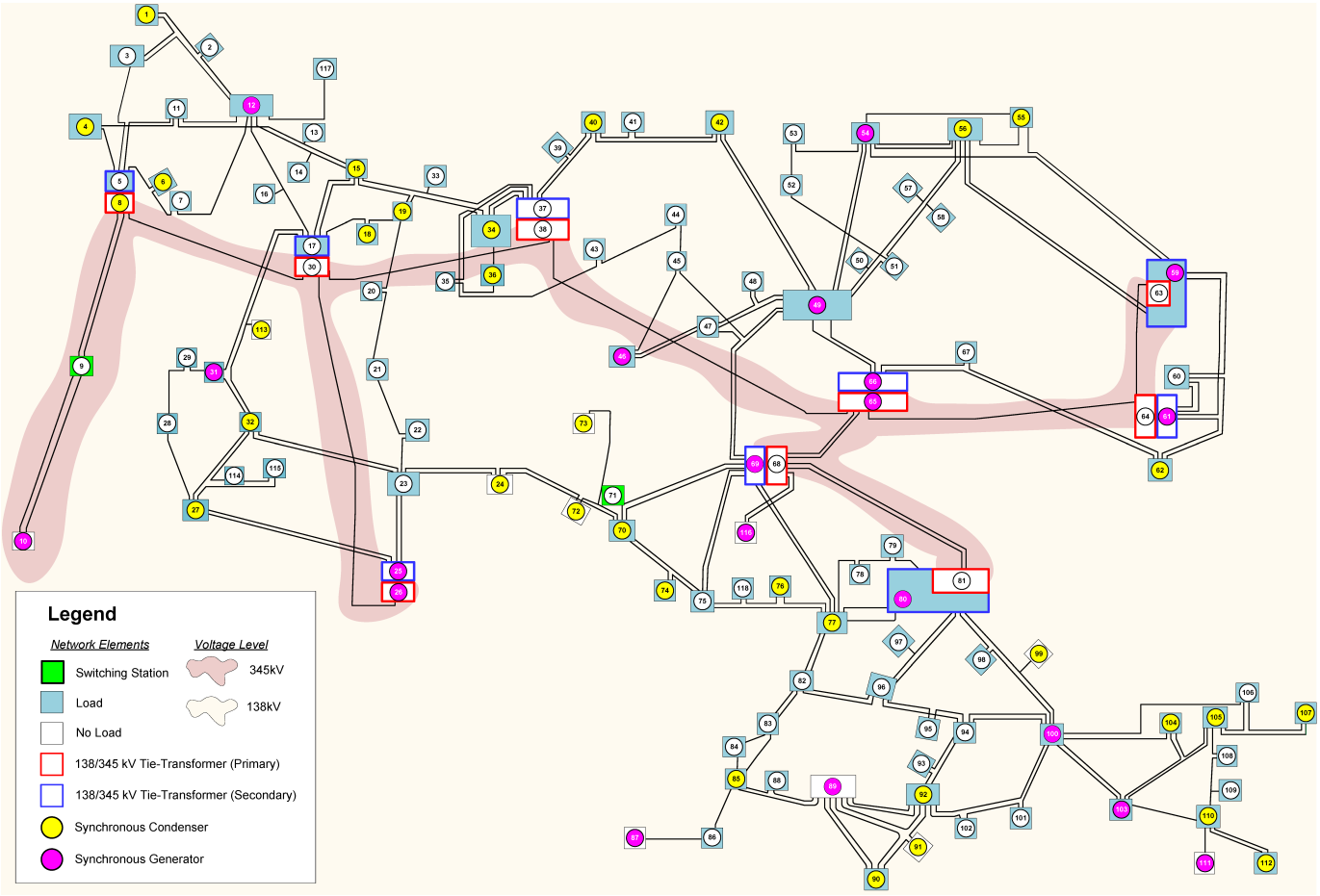

Fig. 3. IEEE 118-Bus system model.

a) Loading conditions: Different loading scenarios are generated by adjusting the overall system demand from 40% to 100% of the peak demand, increasing in increments of 5%.

b) Power station-level dispatch: Advanced gas combined cycle (AGCC) power units are fuel efficient. Thus, their active power outputs are fixed at over 90% for all units at the power station. When the demand level lowers, e.g., below 50%, 3 out of 5 AGCC power stations reduce the number of connected units without changing active power output. The run-of-river hydropower station is treated as a constant power output unit regardless of the loading level. The rest of the power stations, such as the coal/gas-fired and hydropower, are treated as variable active power units.

c) Unit-level dispatch scenarios: Two unit-level dispatch scenarios are employed: even unit dispatch, where all units at a power station alter the active power output with the same percentage, and uneven unit dispatch, where one unit's active power output is adjusted while others remain unchanged.

On-load tap changers, shunt capacitors/reactors, and the reactive power output from synchronous generators/condensers are individually adjusted to satisfy all bus voltage level constraints for a specified demand.

3) Dynamic Models: The dynamic model is built for a round rotor generator with one damper winding circuit for each of the D-axis and Q-axis with saturation characteristics. The inertia time constant, H, is set to 4 seconds and 3 seconds for all synchronous generators and synchronous condensers, respectively, considering the presence of the rotating mass of turbines in generators and its non-existence in condensers. A static load model with frequency-dependent characteristics (2%/%Hz for active power load) is employed to consider the induction motor behavior indirectly.

4) Control and Protection Models: Generator controller models cover an IEEE standard steam turbine-governor controller with a droop of 4% and an IEEE standard static exciter model with an over-excitation limiter, but no power system stabilizer. These models are derived based on the parameters outlined in [10], [13]. Frequency protections that consist of an underfrequency load shedding scheme with three-step cascaded trips and over-/under-frequency relays for synchronous generators/condensers are thoroughly implemented into the simulation model. These parameters were sourced from the specifications provided in [14].

5) PMU Coverage: In consideration of the uneven deployment of PMUs within transmission grids, where PMUs, despite their extensive use, may not be present in every substation, the study aims to assess the algorithm's robustness under limited observability comprehensively. To achieve this goal, 24

TABLE IV
GENERATION RESOURCE WITH TWO DISPATCHING SCENARIOS

Bus number	Fuel type	Unit capacity	# of connected units for demand level	
			Even dispatch	Uneven dispatch
10	Coal	248 MW	2: [40%, 100%]	3: [90%, 100%] 2: [65%, 85%] 1: [40%, 60%]
12	AGCC	23.5 MW	4: [40%, 100%]	4: [40%, 100%]
25	Coal	215 MW	1: [40%, 100%]	1: [40%, 100%]
26	Coal	153 MW	2: [40%, 100%]	2: [65%, 100%] 1: [40%, 60%]
31	ROR	8.0 MW	1: [40%, 100%]	1: [40%, 100%]
46	Gas	10.9 MW	2: [40%, 100%]	2: [65%, 100%] 1: [40%, 60%]
49	AGCC	150 MW	1: [40%, 100%]	1: [40%, 100%]
54	Coal	40 MW	3: [40%, 100%]	2: [50%, 100%] 1: [40%, 45%]
59	AGCC	115 MW	2: [55%, 100%] 1: [40%, 50%]	2: [50%, 100%] 1: [40%, 45%]
61	AGCC	119 MW	2: [55%, 100%] 1: [40%, 50%]	2: [55%, 100%] 1: [40%, 50%]
65	Coal	220 MW	2: [40%, 100%]	2: [65%, 100%] 1: [40%, 60%]
66	Coal	238 MW	2: [40%, 100%]	2: [65%, 100%] 1: [40%, 60%]
69	Hydro	115 MW	1: [40%, 100%]	1: [40%, 100%]
80	Coal	162 MW	2: [40%, 100%]	2: [65%, 100%] 1: [40%, 60%]
87	ROR	12.5 MW	1: [40%, 100%]	1: [40%, 100%]
89	Coal	238 MW	3: [40%, 100%]	3: [85%, 100%] 2: [60%, 80%] 1: [40%, 55%]
100-1	AGCC	100 MW	1: [40%, 100%]	1: [40%, 100%]
100-2	AGCC	119 MW	2: [55%, 100%] 1: [40%, 50%]	2: [55%, 100%] 1: [40%, 50%]
103	Hydro	75 MW	1: [40%, 100%]	1: [40%, 100%]
111	Gas	67 MW	1: [40%, 100%]	1: [40%, 100%]
116	Gas	217 MW	1: [40%, 100%]	1: [40%, 100%]

AGCC: Advanced Gas Combined Cycle, ROR: Run-of-River.

distinct configurations representing diverse scenarios of PMU partial coverage were examined. These configurations detail six variations outlined in Table V, each presenting different types of connected power equipment and varying instances of missing PMUs. Moreover, three unique imputation methods are individually implemented with both the GCN and GAT models, collectively forming comprehensive sets of scenarios.

B. Frequency Health Index Prediction Results

1) Model Performance: The validation losses and testing accuracies obtained from the MLP, GCN, and GAT models are displayed in Table VI. The validation loss decreased when embracing the ordinal encoder for each individual model. The GCN and GAT models achieved a lower validation loss and much higher testing accuracy than the MLP model. The GAT model attained around 1% higher testing accuracy compared to the GCN model. Note that the hyperparameters, e.g., learning rate and the number of hidden layers and neurons, were fine-tuned and chosen based on validation dataset performance.

2) GCN- and GAT- based Frequency Health Index Prediction with Partial PMU Coverage: This study investigated the impact of diverse PMU locations, considering various power equipment, on frequency health index prediction. It also assessed prediction sensitivity to reduced PMU coverage across substations, indicating limited observability. The investigation

TABLE V
PARTIAL PMU COVERAGE SCENARIO

# of missing PMU	Power equipment	Remark
2 PMUs	Switching station	—
9 PMUs	Tie-transformer	Secondary-side only
18 PMUs	Tie-transformer	Both sides
19 PMUs	Synchronous generator	—
35 PMUs	Synchronous condenser	—
53 PMUs	Step-down transformer	Load bus with no unit

TABLE VI
FREQ. HEALTH INDEX PREDICTION PERFORMANCE

Machine learning model	Validation loss ($\times 10^{-3}$)	Testing accuracy	# of hidden channels	# of hidden layers	Learning rate of Adam
MLP	1.3382	0.8192	64	4	0.0001
GCN	0.4419	0.9302	64	5	0.0001
GAT	0.4762	0.9392	64	4	0.0001

Number of heads (i.e., number of attention mechanisms) in GAT model: 4.

aimed to understand how different PMU locations and reduced PMU coverage influence prediction accuracy.

a) Pseudo Measurements from the State Estimator: In the absence of PMUs at specific bus locations, the GCN and GAT models utilized pseudo measurements derived from the state estimator. Notably, state estimator outputs, updated in intervals of $0.5 \sim 5$ minutes, had a much longer sampling rate compared to the standard PMU measurement with a reporting rate of 33.33 ms. Consequently, the state estimator-derived pseudo measurements are inherently noisy and delayed measurements. To emulate this characteristic in experiments, the Gaussian noise with a standard deviation ranging from $1/30$ to $1/3$ is introduced into dynamic simulation results. For instance, a standard deviation of $1/30$ is assumed to correspond to a $\pm 1\%$ measurement error.

The testing losses of the GCN and GAT models in the six scenarios are displayed in the top two subgraphs of Fig. 4. When the pseudo-PMU measurement error is 3% or lower, both models exhibit a minimal increase in testing loss, regardless of missing PMU locations and model variations. However, as measurement noise surpasses 3%, both models show increased losses. Notably, the GAT model displays significantly lower sensitivity to increasing pseudo-PMU measurement error compared to the GCN model, signifying superior robustness despite measurement noise.

Examining testing accuracies in the bottom two subgraphs of Fig. 4 , a decline is observed when measurement error exceeds 3%. However, the GAT model exhibits a less pronounced decrease in accuracy ($0.0 \sim 9.3\%$) compared to the GCN model ($0.0 \sim 25\%$), further emphasizing its superior robustness against measurement noise. This robustness stems from the GAT model's enhanced ability to learn the inter-correlations between nodes efficiently.

Both GCN and GAT models become most vulnerable when no PMUs are present at power stations with synchronous generators, meaning that PMU locations are more important than the PMU coverage fraction/area in terms of frequency health index prediction performance.

b) Zero/Peak Value Imputation for Substation with no PMUs: Figure 5 illustrates the testing losses and accuracies of the

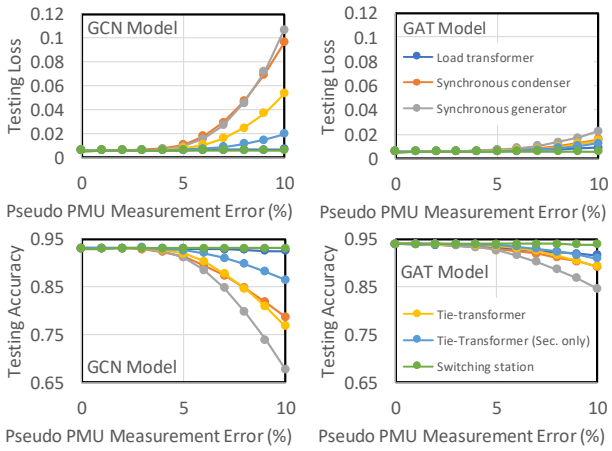

Fig. 4. Testing loss and accuracy across 6 PMU missing scenarios.

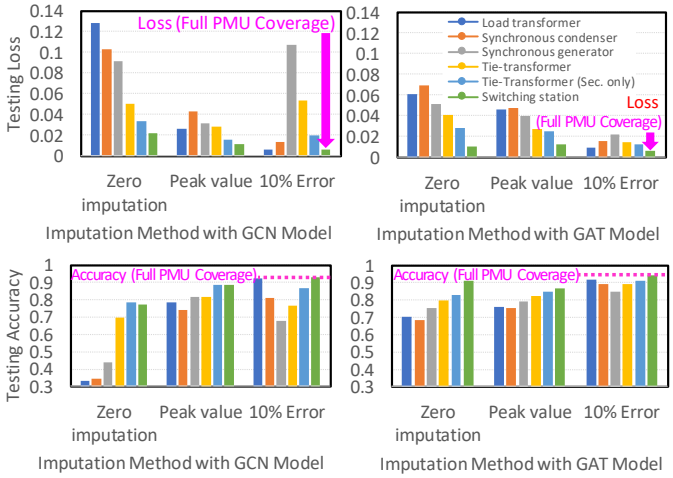

Fig. 5. Comparison of testing loss and accuracy across 3 imputation methods.

GCN and GAT models using two naive imputation methods, including those at the 10% pseudo-PMU measurement error, for better contrast. The GAT model with the peak value replacement strategy overall demonstrates much better health index prediction performance ($68 \sim 91\%$) compared to the GCN model ($33 \sim 78\%$). Conversely, the GAT model with the zero imputation strategy overall presents slightly lower health index prediction performance ($76 \sim 87\%$) compared to the GCN model ($74 \sim 88\%$). Figure 5 also indicates that testing accuracies of both models decrease as the PMU installation density decreases, meaning that the PMU installation ratio is more critical than PMU locations for these naive methods.

IV. Conclusion

This paper develops a graph attention network (GAT)–based frequency health index prediction algorithm to monitor a transmission system in online control and operations. By feeding system topology information in the neural network and introducing ordinal encoding and an attention mechanism, the proposed GAT-based algorithm achieves a 12% improvement in testing accuracy over the baseline algorithm. The proposed GAT-based frequency health index prediction algorithm can also monitor transmission grids with partial PMU coverage. The proposed GAT-based algorithm is more robust than the GCN model when facing missing PMU measurements and measurement errors. Specifically, the GAT model, with a zero imputation strategy, yielded 9.7-37% higher testing accuracy than the GCN model. The GAT model also achieved up to 17% higher testing accuracy than the GCN model at 10% error of pseudo measurements obtained from the output of the state estimator (which is one of the most appropriate replacements for the PMU measurements). These findings complement existing literature by showing GAT's effectiveness in improving frequency health index predictions, offering a valuable approach for future research and applications.

This research work will be expanded by considering N-2 or more contingencies and other types of system health indices, such as angle stability-related health index, to enhance the real-world utility of our approach. We also plan to investigate different types of measurement noise, validate the proposed algorithm on real-world data, and conduct comparative analyses to strengthen its practical value and broaden its applicability.

References

[1] W. Liao, B. Bak-Jensen, J. R. Pillai, Y. Wang, and Y. Wang, "A review of graph neural networks and their applications in power system," *Journal of Modern Power Systems and Clean Energy*, vol. 10, no. 2, pp. 345–360, Mar. 2022.

[2] J. Huang, L. Guan, Y. Su, H. Yao, M. Guo, and Z. Zhong, "System-scale-free transient contingency screening scheme based on steady-state information: A pooling-ensemble multi-graph learning approach," *IEEE Trans. Power Syst.*, vol. 37, no. 1, pp. 294–305, Jan. 2022.

[3] T. Zhao, M. Yue, and J. Wang, "Structure-informed graph learning of networked dependencies for online prediction of power system transient dynamics," *IEEE Trans. Power Syst.*, vol. 37, no. 6, pp. 4885–4895, Jan. 2022.

[4] K. Yamashita, J. Qin, N. Yu, E. Farantatos, and L. Zhu, "Predicting power system voltage health index with PMUs and graph convolutional networks," in *2023 IEEE Power Energy Soc. General Meeting*, 2023.

[5] S. Gu, J. Qiao, W. Shi, F. Yang, X. Zhou, and Z. Zhao, "Multi-task transient stability assessment of power system based on graph neural network with interpretable attribution analysis," *Energy Reports*, vol. 9, pp. 930–942, 2023.

[6] S. Yang, B. Vaagensmith, and D. Patra, "Power grid contingency analysis with machine learning: A brief survey and prospects," in *Proc. of 2018 Int. Jt. Conf. Neural Netw. (IJCNN)*, Oct. 2020.

[7] P. Velickovic, G. Cucurull, A. Casanova, A. Romero, P. Liò, and Y. Bengio, "Graph attention networks," in *6th Int. Conf. Learn. Represent. ICLR 2018, Vancouver, BC, Canada, April 30 - May 3, 2018, Conference Track Proceedings*, 2018.

[8] S. Brody, U. Alon, and E. Yahav, "How attentive are graph attention networks?" in *10th Int. Conf. Learn. Represent. ICLR 2022, Virtual, April 25 - 29, 2022, Conference Track Proceedings*, 2022.

[9] Power Systems Test Case Archive. (Aug., 1993) 118 bus power flow test case. University of Washington. [Online]. Available: https://labs.ece.uw.edu/pstca/pf118/pg_tca118bus.htm

[10] CIGRE WG C4.503, "Power system test cases for EMT-type simulation studies," CIGRE, Tech. Rep. TB736, Aug. 2018.

[11] NERC System Protection and Control Task Force, "Protection system reliability: Redundancy of protection system elements," NERC, Tech. Rep., Nov. 2008.

[12] A. Anderson, S. Kincic, B. Jefferson, B. Mcgary, C. Fallon, D. Ciesielski, J. Wenskovitch, and Y. Chen, "A real-time operation manual for the IEEE 118 bus transmission model," PNNL, Tech. Rep. 33499, Sep. 2022.

[13] IEEE TF on Turbine-Governor Modeling, "Dynamic models for tuebine-governors in power system studies," IEEE, Tech. Rep. PES-TR1, Jan. 2013.

[14] ERCOT, "ERCOT nodal operating guides – section 2: System operations and control requirements," May 2019.