

Data-Driven Graph Learning for Transmission Network Partitioning under $N-1$ Contingencies

Mozhi Chen and Liang Du

Villanova University

Villanova, USA

mchen08,liang.du@villanova.edu

Shengyi Wang

University of Arkansas

Little Rock, USA

swang@ualr.edu

Zhenyu Zhao

Imperial College London

London, U.K

z.zhao1@imperial.ac.uk

Abstract—Accurate partitioning of transmission networks under contingency scenarios is critical for ensuring reliable power deliverability in transmission planning and operational management. This paper introduces a data-driven graph partitioning framework designed to generate electrically coherent and scalable zones that remain effective under $N-1$ contingencies. The proposed method leverages Graph Sample and Aggregate (GraphSAGE)-based representation learning techniques to capture spatial and electrical properties of the transmission network, combined with a trainable soft partitioning module optimized through a multi-objective loss. By incorporating contingency-induced system responses during training, the framework produces partitions that align with power flow behavior and electrical distance under stressed conditions. Numerical results on IEEE 39-bus and IEEE 118-bus benchmark systems demonstrate that the proposed method can produce interpretable, balanced, and electrically-meaningful partitions, which highlights its potential for transmission system planning and controllability.

Index Terms—Transmission Network Partitioning, Contingency Analysis, Graph Learning.

I. INTRODUCTION

Efficient contingency analysis is necessary to assess the long-term effect of transmission line failures in order to guarantee the resilience of electricity transmission networks. The $N-1$ contingency criterion, which evaluates the system's ability to withstand the failure of a single transmission component, is widely utilized in operational planning and security assessment. Conventional $N-1$ contingency methods mostly rely on exhaustive simulations, which can be computationally intensive and may not generalize well to evolving grid topologies [1]. Moreover, existing techniques often inadequately use the structural attributes of transmission networks, which inherently influence failure propagation patterns.

The need of adept contingency analysis is underscored by the increasing vulnerability of transmission lines to failures caused by both natural and technical factors. Extreme weather events, such as hurricanes and wildfires, frequently jeopardize transmission infrastructure, while aging equipment and fluctuating power flows exacerbate reliability concerns. Unanticipated operational modifications resulting in overloading can induce thermal stress on lines, increasing the likelihood of cascading failures. Considering these imminent risks, it is imperative to implement real-time contingency assessment methods that efficiently identify failure components and mitigate disruptions before they develop into significant outages.

Despite these challenges, conventional contingency analysis methods primarily rely on snapshot-based approaches that assess predefined contingency scenarios. These scenarios are typically constructed based on heuristic rules, limiting their adaptability to dynamic states. Previous methods commonly employ power flow redistribution models—such as power transfer distribution factor and line outage distribution factors—to estimate the impacts of component failures [2]. However, these techniques often fail to capture the underlying topological structure of the transmission system, thereby reducing their effectiveness in modeling components failure propagation and post-contingency control strategies.

Partitioning the transmission network provides a principled means of decomposing large-scale power systems into electrically coherent regions based on similarities in electrical distance and system response. In the context of contingency analysis, such electrically meaningful partitions help localize disturbance propagation, improve interpretability of system behavior, and support region-level operational assessment by grouping components with strongly coupled electrical characteristics [3]. However, transmission network partitioning has received few attention in learning-based and contingency-aware settings. Existing transmission approaches, such as spectral clustering and modularity-based zoning, are typically based on static system representations or predefined electrical metrics, and do not explicitly account for contingency-induced state changes. Given the critical role of topology in disturbance propagation across transmission grids, there is a pressing need for electrically meaningful and scalable partitioning methods under $N-1$ contingencies, which motivates the proposed graph-based framework.

Recent studies have demonstrated that the topological structure of power grids significantly enhances the ability to analyze and interpret system behavior under contingencies. Building on these insights, topology-aware methods such as topological data analysis and spectral clustering have shown promising results in contingency classification [4] and voltage area partitioning [5]. These advances underscore the potential of graph-based approaches to enhance contingency analysis, motivating the development of machine learning-driven partitioning frameworks.

To address the limitations of traditional transmission zoning methods, we propose a data-driven Graph Partitioning frame-

work for power systems operating under $N-1$ contingency scenarios. The proposed model utilizes contingency simulation data to generate electrically coherent network partitions, rather than depending on static topology. It integrates GraphSAGE-based graph representation learning with a trainable soft partitioning module, optimized via a multi-objective loss that promotes electrical cohesiveness, balanced zone dimensions, and structural connection. Numerical validation results on the IEEE benchmark systems indicate that the proposed model surpasses traditional clustering techniques (e.g., k -means, Spectral Clustering), producing more balanced and electrically coherent partitions for transmission planning and control.

II. PRELIMINARIES

A transmission network can be modeled as a graph $\mathcal{G} = (\mathcal{N}, \mathcal{E})$, where buses correspond to nodes $\mathcal{N}$ and transmission lines correspond to edges $\mathcal{E}$. Let $n = |\mathcal{N}|$ and $m = |\mathcal{E}|$. Each line $(i, j) \in \mathcal{E}$ is characterized by its susceptance B_{ij} , and all line susceptances are collected into a diagonal matrix $B \in \mathbb{R}^{m \times m}$. Throughout this paper, the terms *bus/node* and *line/edge* are used interchangeably. Graph partitioning aims to divide the network into K electrically coherent zones while minimizing the connections between them. Formally, given an undirected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, the objective is to identify disjoint subsets $\{V_1, \dots, V_K\}$ such that nodes within each subset are strongly connected, whereas couplings between different subsets remain weak. To formalize this, the cut function is defined as the total weight of edges crossing between a subset V_k and its complement $\bar{V}_k$:

$$\text{cut}(V_k, \bar{V}_k) = \sum_{i \in V_k, j \in \bar{V}_k, (i,j) \in \mathcal{E}} A_{ij}, \quad (1)$$

where A is the adjacency matrix of the graph and A_{ij} is the weight of the edge between node i and node j . However, directly minimizing the cut value can lead to trivial solutions that isolate small sets of nodes. A normalization term, the volume of a subset V_k , is then defined as:

$$\text{vol}(V_k) = \sum_{i \in V_k, j \in \mathcal{V}, (i,j) \in \mathcal{E}} A_{ij}. \quad (2)$$

The normalized cut objective measures the quality of such a partition by penalizing the sum of edge weights between clusters relative to the total edge weights within each cluster [6], [7]. To avoid trivial partitions that isolate small sets of nodes, the cut is normalized by the volume of each subset. Formally, the Ncut function is defined as [6], [7]:

$$\text{Ncut}(V_1, \dots, V_K) = \sum_{k=1}^K \frac{\text{cut}(V_k, \bar{V}_k)}{\text{vol}(V_k)}. \quad (3)$$

Let $A \in \mathbb{R}^{N \times N}$ denote the adjacency matrix and $D = \text{diag}(A\mathbf{1}) \in \mathbb{R}^{N \times N}$ be the degree matrix. The node-to-cluster assignments are encoded using a clustering indicator matrix $M \in \{0, 1\}^{N \times K}$, where $M_{ik} = 1$ indicates that node i belongs to cluster k , and each node is assigned to one cluster. The Ncut objective can then be equivalently formulated by [8]:

$$\max_{M \in \{0, 1\}^{N \times K}} \frac{1}{K} \sum_{k=1}^K \frac{\mathbf{M}_k^\top \mathbf{A} \mathbf{M}_k}{\mathbf{M}_k^\top \mathbf{D} \mathbf{M}_k} \quad \text{s.t. } \mathbf{M} \mathbf{1}_K = \mathbf{1}_N, \quad (4)$$

where $\mathbf{M}_k$ is the k -th column of $\mathbf{C}$, indicating the nodes in cluster k . This objective balances inter-cluster sparsity with intra-cluster compactness, providing a principled criterion for graph decomposition that avoids unbalanced partitions.

III. PROBLEM FORMULATION

The objective of this work is to develop a learning-based framework for contingency-aware transmission network partitioning, ensuring that partitions exhibit strong intra-cluster electrical connectivity while minimizing outer-cluster interactions. While the Normalized Cut (Ncut) objective offers a mathematically sound criterion for graph partitioning, its direct optimization is computationally demanding due to the discrete nature of partition assignments. More importantly, existing partitioning methods often lack access to or ignore the electrical and operational characteristics of the transmission network, resulting in partitions that are not meaningful for power system applications. Previous approaches rely on purely topological or heuristic structures, without incorporating system information such as power flows, voltage magnitudes, or contingency responses—limiting their applicability in transmission-level planning [9], which highlight the need for a learning-based approach that can leverage actual data, scale efficiently to large networks, and generate interpretable and robust partitions.

A. Graph-based Learning for Partitioning

Instead of directly minimizing the Ncut objective using spectral methods, we reformulate the problem as a node embedding learning task. Given a transmission network represented as $\mathcal{G} = (\mathcal{N}, \mathcal{E})$, we define a learnable function $\mathcal{F}$ that maps node features to partition assignments:

$$\mathcal{F} : X \rightarrow P, \quad P \in \mathbb{R}^{n \times K}, \quad (5)$$

where $X \in \mathbb{R}^{n \times d}$ represents the input node features, including electrical properties such as active power, voltage magnitudes, and phase angle. P is a soft partitioning matrix, where each entry P_{ik} represents the probability of node i belonging to partition k . This formulation allows for a continuous relaxation of the partitioning problem, making it differentiable and trainable via standard optimization techniques.

B. Dataset Generation

To enable data-driven graph partitioning, the dataset is generated using time-series simulations of the transmission system under contingency conditions. In each simulation, a single transmission line is disconnected at the initial time step, corresponding to an $N-1$ contingency. The subsequent system response is simulated over a fixed time horizon, yielding time-series measurements for each contingency scenario. In this manner, data are generated for all feasible $N-1$ line outage cases that do not result in power flow failure. The total number of contingency samples is therefore $m \times T$, where m denotes the number of valid $N-1$ contingencies and T represents the number of time steps per simulation. In the final dataset, the normal operating condition is sampled more frequently than contingency cases, with a ratio of 9:1, allowing the model to

learn a robust representation of nominal system behavior while still capturing the effects of contingencies.

IV. METHODOLOGY

This paper proposes a data-driven graph partitioning framework is shown in Fig. 1. First, the graph embedding module applies GraphSAGE to extract spatial and electrical features from the transmission network. Second, the graph partitioning module utilizes a trainable mechanism to assign nodes to soft partitions based on the learned node embeddings.

A. Graph-Based Feature Representation

Consider a transmission network modeled by a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where each node is associated with a feature matrix $X \in \mathbb{R}^{N \times d}$ that encodes electrical operating states, including active power injections, voltage magnitudes, and phase angles. These features reflect both steady-state conditions and contingency-induced responses obtained from power flow simulations. To extract spatial and electrical dependencies arising from power flow coupling and network topology, we employ the Graph Sample and Aggregate (GraphSAGE) framework [10]. GraphSAGE learns node embeddings through neighborhood sampling and aggregation, allowing electrically interacting buses to exchange information and enabling inductive generalization to unseen operating conditions. The node representation $h_v^{(l+1)}$ at layer $l+1$ is given by

$$h_v^{(l+1)} = \sigma \left(W^{(l)} \cdot \text{AGGREGATE}^{(l)} \left(\{h_v^{(l)}\} \cup \{h_u^{(l)}, \forall u \in \mathcal{N}(v)\} \right) \right), \quad (6)$$

where $h_v^{(l)}$ denotes the representation of node v at layer l with $h_v^{(0)} = x_v$, and $\mathcal{N}(v)$ represents electrically and topologically adjacent nodes. The aggregation operator $\text{AGGREGATE}^{(l)}$, implemented using a pooling function in this work, captures the influence of neighboring buses through shared electrical behavior. Here, $W^{(l)}$ is a trainable weight matrix and $\sigma(\cdot)$ is an activation function (e.g., ReLU).

B. Graph Partitioning module

This module manages partitioning the graph, taking in node embeddings and generating the probability that each node belongs to partitions $S_1, S_2, \dots, S_g$. Given node embeddings H , a partitioning module with fully connected layers is

$$Z = \text{ReLU}(HW_1 + b_1), C = \text{softmax}(ZW_2 + b_2), \quad (7)$$

where $C \in \mathbb{R}^{N \times k}$ represents first-stage node partition probabilities. To enforce spatial coherence, we apply the network Laplacian L to promote topology-aware propagation. This encourages nodes that are topologically close to exhibit similar partitioning behavior. Specifically, L is used to smooth the initial soft partitioning matrix C by $P = LC$, where P is the final soft partitioning matrix, in which each entry P_{ik} represents the probability of node i belonging to partition k .

C. Loss Function

To ensure that the learned partitions are topologically coherent and electrically meaningful, we adopt a multi-objective loss function $\mathcal{L} = \alpha \mathcal{L}_{\text{Ncut}} + \beta \mathcal{L}_{\text{ortho}} + \gamma \mathcal{L}_{\text{conn}}$, with

$$\mathcal{L}_{\text{Ncut}} = -\frac{\text{Tr}(P^T AP)}{\text{Tr}(P^T DP)}, \quad (8)$$

$$\mathcal{L}_{\text{ortho}} = \left\| \frac{P^T P}{\|P^T P\|_F} - I_K \right\|_F, \quad (9)$$

$$\mathcal{L}_{\text{conn}} = \sum_{k=1}^K \frac{C_k - 1}{|V_k|}, \quad (10)$$

where α, β, γ are weighting hyperparameters. $S \in \mathbb{R}^{n \times K}$ is soft partition assignment matrix, A is adjacency matrix, D is corresponding degree matrix, C_k denotes the number of connected nodes within the subgraph induced by partition k , and $|V_k|$ is the number of nodes assigned to that partition.

The normalized cut term $\mathcal{L}_{\text{Ncut}}$ encourages low inter-partition connectivity, promoting modular and electrically cohesive partitions. To mitigate the smoothing effect induced by information propagation in Graph Neural Networks, the orthogonality term $\mathcal{L}_{\text{ortho}}$ is introduced to prevent a degenerate minimum caused by $\mathcal{L}_{\text{Ncut}}$ where all nodes are assigned equally across all clusters. This term encourages cluster assignments to be distinct and of similar size. Finally, A connectivity loss $\mathcal{L}_{\text{conn}}$ ensures that each partition induces a connected subgraph.

This combined loss function allows the model to generate partitions that are structurally coherent, electrically interpretable, and suitable for downstream applications such as contingency analysis or graph-based modeling.

D. Evaluation of Power System Clustering

To quantitatively assess the quality of the learned transmission network partitions, this work follows the state-of-the-practice [3] and utilizes five evaluation indices, which are designed to capture the electrical and topological coherence of the clustering outcomes.

1) *Electrical Cohesiveness Index (ECI)*: Measures the extent to which buses within a cluster are electrically proximate. Higher values indicate that nodes within the same partition experience similar phase angle responses to power injections, reflecting strong intra-cluster electrical ties. It is defined as:

$$\text{ECI} = 1 - \frac{\sum_{i=1}^n \sum_{j \in B_i} e_{ij}}{\sum_{i=1}^n \sum_{j=1}^n e_{ij}}, \quad (11)$$

where B_i is the set of buses in the same cluster as bus i , and e_{ij} is the electrical distance between buses i and j . The electrical distance e_{ij} is defined as:

$$e_{ij} = (J_{P\theta}^+)_{ii} - (J_{P\theta}^+)_{ij} - (J_{P\theta}^+)_{ji} + (J_{P\theta}^+)_{jj}, \quad (12)$$

where $J_{P\theta}^+$ is Moore–Penrose pseudoinverse of real power–phase angle sensitivity matrix $J_{P\theta}$, which corresponds to bus susceptance matrix under DC power flow assumptions.

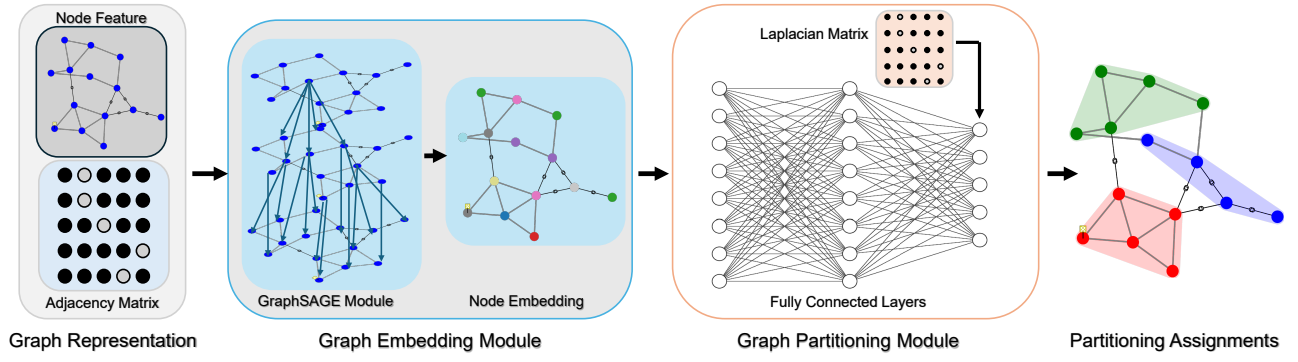


Fig. 1. Overview of proposed graph partitioning model.

2) *Between-Cluster Connectedness Index (BCCI)*: Assesses how loosely connected clusters are to each other. A high BCCI value implies that inter-cluster electrical distances are large, promoting isolation between zones and minimizing undesired loop flows. It is calculated as:

$$BCCI = 1 - \frac{\sum_{i=1}^n \sum_{j \notin B_i} \frac{1}{e_{ij}}}{\sum_{i=1}^n \sum_{j=1, j \neq i}^n \frac{1}{e_{ij}}}. \quad (13)$$

3) *Cluster Count Index (CCI)*: Evaluates how close the number of generated clusters is to a desired or user-specified number c^* . It penalizes over-segmentation or under-segmentation using a log-normal-based formula:

$$CCI = \exp\left(-\frac{(\ln c - \ln c^*)^2}{2\sigma^2}\right), \quad (14)$$

where c is the number of clusters, and $\sigma = w \ln n$, with w being a tuning parameter.

4) *Cluster Size Index (CSI)*: Measures how evenly the cluster sizes are distributed relative to an ideal size $s^* = n/p^*$. It also uses a log-normal-based formula:

$$CSI = \exp\left(-\frac{(\ln \bar{s} - \ln s^*)^2}{2\sigma^2}\right), \quad (15)$$

where $\bar{s}$ is the weighted average cluster size and $\sigma = w \ln n$, with w being a tuning parameter.

5) *Cluster Connectedness (CC)*: A binary indicator defined as $CC = 1$ if all clusters are connected, and $CC = 0$ otherwise.

6) *Overall Fitness*: To quantify the overall quality of a given partitioning result, a composite fitness score that multiplicatively combines all key evaluation indices is defined as

$$f = ECI^a \cdot BCCI^b \cdot CCI^c \cdot CSI^d \cdot CC, \quad (16)$$

where a, b, c , and d are user-defined scalar weights that determine the relative contribution of each component to the overall fitness [3]. In our experiments, we set $a = b = c = d = 1$ to treat all indices equally. The binary term $CC \in \{0, 1\}$ ensures that only fully connected cluster configurations are considered as final solutions.

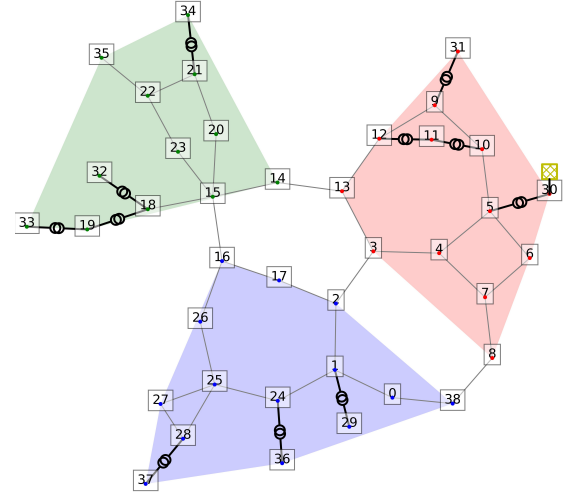


Fig. 2. Partition result of proposed model for the IEEE 39 bus system.

V. NUMERICAL VALIDATIONS

This section demonstrates the effectiveness of the proposed graph partitioning framework on V-A. IEEE 39-bus system and V-B. IEEE 118-bus system. The objective is to evaluate the capability of the data-driven partitioning model to generate electrically meaningful and balanced zones under realistic operating and contingency conditions. Performance comparisons are made against conventional clustering methods such as k -means and spectral clustering. For both systems, the number of voltage partitions is fixed at three. The graph neural network (GNN) adopts a two-layer GraphSAGE architecture to capture local and global topological dependencies.

All numerical validations are implemented in PyTorch and executed on an NVIDIA GeForce RTX 4070 Ti Super GPU. The datasets are generated using pandapower [11], encompassing both normal operating states and $N-1$ contingency cases to ensure robustness and generalization of the learned partitions.

A. IEEE 39-Bus System

Table I summarizes the performance comparison between the proposed model, k -means, and spectral clustering. The proposed method achieves the highest electrical cohesion (ECI

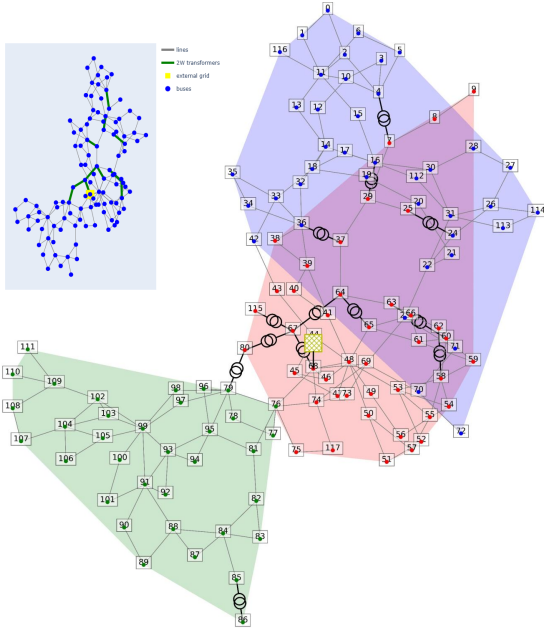


Fig. 3. Partition result of proposed model for the IEEE 118 bus system.

TABLE I
COMPARISON OF THE PROPOSED METHOD WITH THE LITERATURE BY
VARIOUS METRICS ON THE IEEE 39-BUS BENCHMARK

Method	ECI	BCCI	CCI	CSI	Fitness
<i>k</i> -means	0.6761	0.6476	1.0000	0.3555	0.1512
Spectral Clustering	0.7142	0.5086	1.0000	0.9251	0.3360
Proposed	0.7692	0.4994	1.0000	0.9998	0.3841

= 0.7692) and an almost perfect cluster size index (CSI = 0.9998), outperforming both benchmarks. Although the boundary cluster connectivity index (BCCI) is slightly lower than that of *k*-means, the proposed approach attains the best overall fitness (0.3841). As illustrated in Fig. 2, the resulting clusters align with electrically proximal zones and follow the network topology, forming partitions connected by only a few critical tie-lines. This indicates that the model effectively minimizes inter-zone bridges while preserving strong intra-zone connectivity, ensuring robustness.

B. IEEE 118-Bus System

To further evaluate scalability and generalization, the proposed framework is applied to the larger IEEE 118-bus system, maintaining three voltage partitions. The same experimental settings and GNN architecture are used for consistency. As shown in Table II, the proposed model again achieves the best electrical cohesion (ECI = 0.7642) and near-perfect cluster size balance (CSI = 0.9996). Note that, despite a marginally lower BCCI, the overall fitness reaches the highest 0.3667, confirming its scalability and effectiveness in generating electrically meaningful partitions for larger networks. Moreover, Figure 3 visualizes the learned partitions, where the learned clusters form clear and well-defined boundaries along weakly coupled tie-lines, showing that the model captures electrical proximity rather than purely geometric structure. These results validate

TABLE II
COMPARISON OF THE PROPOSED METHOD WITH THE LITERATURE BY
VARIOUS METRICS ON THE IEEE 118-BUS BENCHMARK

Method	ECI	BCCI	CCI	CSI	Fitness
<i>k</i> -means	0.7067	0.5814	1.0000	0.7328	0.2984
Spectral Clustering	0.6865	0.4806	1.0000	0.9496	0.3140
Proposed	0.7642	0.4801	1.0000	0.9996	0.3667

robustness and adaptability of the proposed model to preserve electrical proximity while generating topologically meaningful and balanced zones at larger scale.

VI. CONCLUSIONS

This paper presents a learning-based framework for contingency-aware transmission network partitioning, aimed at improving system zoning and deliverability under stressed conditions. By integrating graph representation learning with a soft partitioning module and multi-objective loss, the model generates partitions that reflect the electrical attribute of the system. Unlike traditional clustering methods, our approach explicitly learns from contingency data to ensure the resulting zones are both topologically coherent and operationally meaningful. To quantitatively evaluate the quality of the learned transmission network partitions, we employ five metrics that capture both electrical and topological coherence. Numerical validation results on IEEE 39 and 118 bus benchmark systems demonstrate effectiveness of the proposed framework for network partitioning, supporting applications such as transmission planning, regional control, and enhanced grid resilience.

REFERENCES

- [1] H. Yang, X. Shen, W. Gao, X. Ding, Y. Du, and Q. Li, "Optimal planning for offshore wind farm electrical collector system considering $n - 1$ criterion," *IEEE Transactions on Sustainable Energy*, pp. 1–13, 2025.
- [2] K. Strunz, S. Schilling, M. Kuschke, and A. Czerwinski, "Fast contingency analysis and control for overload mitigation in integrated high voltage ac and multi-terminal hvdc grids," *IEEE Transactions on Power Systems*, vol. 39, no. 4, pp. 5575–5590, 2023.
- [3] E. Cotilla-Sanchez *et al.*, "Multi-attribute partitioning of power networks based on electrical distance," *IEEE Transactions on Power Systems*, vol. 28, no. 4, pp. 4979–4987, 2013.
- [4] Y. Chen, R. A. Jacob, Y. R. Gel, J. Zhang, and H. V. Poor, "Learning power grid outages with higher-order topological neural networks," *IEEE Transactions on Power Systems*, vol. 39, no. 1, pp. 720–732, 2024.
- [5] S. Wang, L. Du, L. Zhu, and Y. Li, "Neural spectral clustering based voltage area partition of active distribution systems," in *2024 IEEE 63rd Conference on Decision and Control (CDC)*, 2024, pp. 3495–3500.
- [6] J. Shi and J. Malik, "Normalized cuts and image segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 8, pp. 888–905, 2000.
- [7] A. Nazi, W. Hang, A. Goldie, S. Ravi, and A. Mirhoseini, "Gap: Generalizable approximate graph partitioning framework," *arXiv preprint arXiv:1903.00614*, 2019.
- [8] Yu and Shi, "Multiclass spectral clustering," in *Proceedings Ninth IEEE International Conference on Computer Vision*, 2003, pp. 313–319 vol.1.
- [9] E. Cotilla-Sanchez, P. D. Hines, C. Barrows, and S. Blumsack, "Comparing the topological and electrical structure of the north american electric power infrastructure," *IEEE Systems Journal*, vol. 6, no. 4, pp. 616–626, 2012.
- [10] W. Hamilton, Z. Ying, and J. Leskovec, "Inductive representation learning on large graphs," *Advances in neural information processing systems*, vol. 30, 2017.
- [11] L. Thurner *et al.*, "pandapower — an open-source python tool for convenient modeling, analysis, and optimization of electric power systems," *IEEE Trans. Power Systems*, vol. 33, no. 6, pp. 6510–6521, Nov 2018.