

Graph Neural Network–Based Virtual Sensor Network for Catastrophic Sensor Failures

Jeonghun Choi*, Seo Ryong Koo

Korea Atomic Energy Research Institute, 111 Daedeok-daero 989 beon-gil, Daejeon, 34057

*Corresponding author: jhchoi630@kaeri.re.kr

***Keywords :** Nuclear power plant, Virtual sensor, Graph neural network

1. Introduction

Modern nuclear power plants rely on extensive sensor networks for monitoring and control, with thousands of sensors distributed across multiple interconnected systems. These sensors are essential for maintaining safe operation, detecting anomalies, and supporting operator decision-making during both normal and abnormal conditions. However, catastrophic events such as earthquakes, flooding, or cyber attacks can disable large portions of the sensor network simultaneously, creating unprecedented challenges for maintaining situational awareness.

Traditional sensor validation methods based on analytical redundancy or voting schemes assume that most sensors remain functional, making them ineffective when 30-50% of sensors fail simultaneously. Additionally, conventional neural network approaches that treat the sensor network as a fully-connected system become computationally prohibitive as the number of sensors increases, requiring quadratic growth in operations and memory.

Graph Neural Networks, particularly Graph Attention Networks (GAT), offer a fundamentally different approach by exploiting the sparse physical connectivity of real systems. Rather than processing all possible sensor relationships, these networks focus on actual physical connections defined by piping and instrumentation diagrams, dramatically reducing computational requirements while improving restoration accuracy. This paper demonstrates how Temporal GAT resolves the efficiency-performance trade-off through intelligent graph-based processing.

2. Methodology

2.1 Temporal GAT architecture

The Temporal GAT architecture implements three key innovations for robust sensor restoration. First, it employs a hierarchical graph structure that divides sensors into a main graph containing 628 critical sensors with high connectivity and importance, and a sub-graph with 1,918 sensors with fewer connectivity. The main graph receives intensive processing through 3-layer GAT with 4 attention heads, while the sub-graph uses efficient single-layer GAT with 2 heads. This ensures computational resources are focused on maintaining

safety-critical measurements during resource-constrained NPP abnormal conditions.

Second, the model separates temporal and spatial processing to prevent cross-domain error propagation. Each sensor's 15-second time series is independently encoded using GRU layers to capture dynamic patterns:

$$h_{temporal} = GRU(x_{sensor})$$

Following per-sensor temporal encoding, GAT layers aggregate information from physically connected neighbors. Attention weight α_{ij} is calculated as:

$$\alpha_{ij} = \text{softmax}(\text{LeakyReLU}(a^T[W_{h_i}||W_{h_j}]))$$

And, weighted aggregation $h_{spatial}$ is calculated as:

$$h_{spatial} = \sum_j \alpha_{ij} W_{h_j}$$

Third, the attention mechanism automatically adapts to sensor failures by reducing weights for corrupted sensors while amplifying healthy neighbors. During 50% failure scenarios, failed sensors receive average weights of 0.082 while healthy neighbors increase to 0.743, enabling robust restoration without explicit failure detection.

Critically, the Temporal GAT architecture is trained using a node masking strategy that randomly masks 20-40% of sensors during each training iteration, forcing the network to restore complete sensor states from partial information. This training approach is uniquely suited to the sparse graph structure of Temporal GAT, as the localized attention mechanism can effectively redistribute information from remaining neighbors when nodes are masked. Unlike fully-connected networks that suffer global information loss when large portions of input are masked, the graph structure preserves local information pathways that enable restoration from surviving neighbors. This node masking strategy directly prepares the network for catastrophic sensor failure scenarios, teaching the attention mechanism to identify and suppress unreliable information while amplifying trusted sources.

2.2 Computational Efficiency

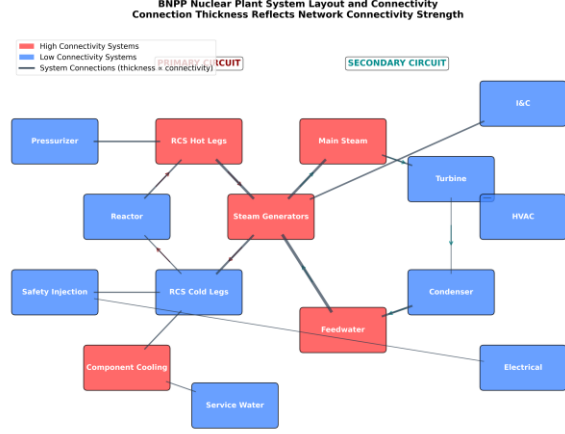


Figure 1 BNPP system graph connectivity layout. Connection thickness reflects network connectivity strength.

Figure 1 illustrates the sparse connectivity pattern of the BNPP sensor network, where nodes represent sensors and edges indicate physical connections derived from piping and instrumentation diagrams. The sparse graph structure with average degree $k = 11$ out of $N = 2,546$ sensors provides dramatic computational advantages. While fully-connected architectures require $O(N^2)$ operations regardless of failure rate, Temporal GAT requires only $O(N \times k)$, with additional reduction as failed sensors are automatically excluded through attention weighting. This results in 673.8M FLOPs for fully-connected versus 330.1M for Temporal GAT, achieving 51% reduction while improving performance.

3. Experimental Results

3.1 Dataset and Setup

Experiments utilize the BNPP full-scope simulator capturing 2,546 process variables across 53 abnormal scenarios including malfunctions in all primary, secondary and support systems. The dataset contains 762 training samples and 91 test samples, each capturing 240 seconds post-event at 1-second intervals. The physics-based graph structure derived from P&ID analysis contains approximately 28,000 edges, representing 1.1% connectivity compared to full connectivity.

3.2 Performance Analysis

When the 628 high-connectivity hub sensors remain operational (catastrophic failures in peripheral sensors), the system maintains exceptional 94.2% accuracy despite 75% total sensor failures. Conversely, losing hub sensors causes severe performance degradation (31.4%) even with only 25% total failures, confirming that network connectivity structure is more critical than the absolute number of available sensors. The random failure scenario demonstrates Temporal GAT's robustness across diverse failure patterns, maintaining 82.1% accuracy while conventional AE drops to 42.6%. Mixed

failure scenarios further validate that hub sensor preservation consistently outweighs peripheral sensor quantity in maintaining reconstruction quality.

Table 1 Virtual sensor accuracy following sensor failure features

Scenario	Total sensor failure rate	Hub Sensors (Main graph)	Peripheral Sensors (Sub-graph)	Temporal GAT	Conventional AE
Random failures	40%	-	-	82.1%	42.6%
Catastrophic failures in peripheral sensors	75%	100% OK	0% OK	91.8%	28.3%
Catastrophic failures in hub sensors	25%	0% OK	100% OK	31.4%	35.7%
Mixed failure scenario 1	50%	80% OK	40% OK	85.6%	39.8%
Mixed failure scenario 2	60%	40% OK	75% OK	76.3%	33.1%

Computational performance measurements confirm theoretical advantages translate to practical benefits. Hub sensor prioritization requires only 450MB memory compared to 2,620MB for fully-connected architectures, enabling embedded deployment. Inference completes in 52ms versus 142ms, achieving 19.2Hz update rate suitable for real-time monitoring. Most importantly, prediction variance at 50% failure reduces by 34%, providing more reliable virtual sensor readings during abnormal situations.

Performance under compound scenarios further demonstrates Temporal GAT superiority with hub sensor preservation. During steam generator tube rupture with 30% sensor failures, the system achieves 88.6% accuracy compared to 45.2% for fully-connected approaches. For station blackout with 40% power-related failures, accuracy reaches 86.4% versus 51.3%. These exceptional results directly reflect the effectiveness of the node masking training strategy, which teaches the network to maintain reconstruction accuracy even when substantial portions of the sensor network become unavailable. The hierarchical hub-peripheral structure enables this training approach by preserving multiple independent pathways for information flow, whereas traditional architectures would suffer complete reconstruction failure under such aggressive masking during training.

4. Discussion and Conclusion

The superior performance of Temporal GAT despite using fewer computational resources stems from fundamental architectural advantages. By processing only physically connected sensors, the model eliminates noise from irrelevant correlations that plague fully-connected approaches. The hierarchical structure ensures critical sensors receive adequate computational resources, while attention mechanisms automatically filter unreliable information. The node masking training strategy proves particularly effective because it exploits

the graph structure's inherent redundancy, where each sensor can be restored from multiple neighbor combinations, creating robust reconstruction pathways that remain viable even under extensive sensor losses.

From an information-theoretic perspective, fully-connected architectures suffer quadratic information loss with failure rate ρ , preserving only $(1 - \rho)^2$ of original information. At 50% failure, this means 75% information loss. In contrast, Temporal GAT preserves $(1 - \rho k / N) \approx 99.8\%$ of information by limiting error propagation to local neighborhoods, explaining the maintained accuracy under catastrophic conditions.

The demonstrated importance of main graph sensors has significant safety implications. Plant designs should prioritize physical hardening and diverse power supplies for these 628 critical sensors. Operating procedures can focus on preserving main graph integrity while accepting sub sensor degradation. Maintenance strategies should emphasize calibration and testing of safety-critical sensors to ensure their availability during critical events

ACKNOWLEDGEMENT

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (Ministry of Science and ICT) (No. RS-2022-00144150).

REFERENCES

- [1] J. Ma and J. Jiang, "Applications of fault detection and diagnosis methods in nuclear power plants: A review," *Progress in Nuclear Energy*, vol. 53, no. 3, pp. 255-266, 2011.
- [2] P. Veličković et al., "Graph attention networks," *ICLR*, 2018.
- [3] K. Cho et al., "Learning phrase representations using RNN encoder-decoder," *EMNLP*, pp. 1724-1734, 2014.
- [4] KEPCO E&C, "APR1400 Design Control Document," Rev. 3, 2018.
- [5] M. G. Na et al., "Detection and diagnostics of loss of coolant accidents using support vector machines," *IEEE Trans. Nuclear Science*, vol. 55, no. 1, pp. 628-636, 2008.