Fault Signal Detection in Nuclear Power Plant Measurements Using Gated Fusion of Temporal and Inter-Variable Experts

Yu Seob Kim, Seung Gyu Cho, Seung Jun Lee*
Ulsan National Institute of Science and Technology, 50, UNIST-gil, Ulsan, 44911
**Corresponding author: sjlee420@unist.ac.kr*

***Keywords :** Nuclear power plant, nuclear fatigue monitoring system, measurement signal, fault signal detection, Temporal convolutional neural network, Transformer

1. Introduction

Measurement signals acquired from instrumentation systems in nuclear power plants are essential information used not only for plant operation and control but also for condition monitoring and integrity assessment of major components and piping. In particular, the Nuclear Fatigue Monitoring System (NuFMS) evaluates cumulative fatigue usage factors of major components and piping using measurement signals such as temperature, pressure, and power obtained from plant monitoring systems. Therefore, the accuracy and reliability of measurement signals directly affect the reliability of fatigue assessment results.

However, measurement signals acquired from actual nuclear power plants may contain various types of spurious signals caused by sensor malfunctions, errors during signal transmission, communication failures, and environmental effects. Such spurious signals can appear in various forms, including spikes, stuck signals, drift, and abnormal fluctuations. If these signals are not properly distinguished from actual process changes, they may degrade the reliability of condition monitoring and integrity assessments based on measurement data.

In our previous study, a Long Short-Term Memory (LSTM)-based classification model was applied to automatically detect spurious signals in nuclear power plant measurements.

Spurious signals in nuclear power plant measurements may need to be identified using different types of information depending on their characteristics and occurrence patterns. For example, abnormalities characterized by distinctive waveform changes, such as abrupt spikes or signals that remain constant for a certain period, can be identified by analyzing the temporal behavior of an individual measurement signal. In contrast, some abnormalities may not exhibit clear anomalous patterns in the waveform itself but may become apparent when the signal is compared with other measurement variables representing the same plant state. In such cases, it is necessary to consider not only the temporal characteristics of an individual signal but also the relationships among multiple measurement variables.

Accordingly, in this study, the information required for spurious signal detection is divided into two categories: temporal characteristics of individual measurement signals and inter-variable relationships among multivariate measurement signals. Because these two types of information represent different characteristics, rather than processing them simultaneously using a single feature extraction structure, separate models are constructed to extract features specialized for each type of information, and the extracted features are subsequently integrated for spurious signal detection.

Based on the expert specialization and selective combination concept of the Mixture-of-Experts (MoE) framework, this study proposes a Dual-Expert-based spurious signal detection model consisting of two parallel experts specialized in different characteristics relevant to spurious signal detection. The Temporal Expert employs a multi-scale Temporal Convolutional Neural Network (CNN) that processes each measurement variable as an independent time series to extract local variations and temporal waveform characteristics. In contrast, the Relation Expert employs Transformer-based self-attention to analyze relationships among multiple measurement variables observed at the same time step.

The features extracted by the two experts are integrated using a learnable gating mechanism. This enables the model to combine the temporal waveform information of individual measurement signals with inter-variable relational information into a unified feature representation and to estimate the probability of a spurious signal for each measurement variable within a given time window.

Through this architecture, this study proposes a Dual-Expert framework that separately analyzes temporal waveform characteristics and inter-variable relationships and selectively integrates the information extracted by each expert, with the aim of developing a spurious signal detection method capable of considering the diverse characteristics of nuclear power plant measurement signals.

2. Proposed Framework

2.1 Overall Framework

The overall architecture of the proposed spurious signal detection framework is illustrated in Fig. 1. The proposed model employs a Dual-Expert architecture to separately analyze two complementary characteristics of nuclear power plant measurement signals: the temporal waveform characteristics of individual measurement variables and the relationships among multiple measurement variables.

The input consists of multivariate time-series data represented as $X \in R^{(B \times T \times C)}$, where B denotes the batch size, T denotes the length of the input time window, and C denotes the number of measurement variables. Unlike a conventional sequence classifier that produces a single prediction for the entire input window, the proposed framework generates an independent spurious-signal prediction for each measurement variable. Accordingly, the final output is represented as $Y_hat \in R^{(B \times C)}$.

As shown in Fig. 1, the proposed framework consists of three main components: a Temporal Expert, a Relation Expert, and a Gated Fusion Module.

The Temporal Expert is designed to extract temporal waveform characteristics from each measurement signal. Each sensor is processed independently using a shared multi-scale Temporal Convolutional Neural Network (CNN), allowing the model to capture both local and relatively long-term waveform patterns without mixing information from different measurement variables. The Temporal Expert also generates point-wise fault evidence for each sensor and time step, indicating the temporal locations that contribute to the detection of a spurious signal.

In parallel, the Relation Expert analyzes relationships among multiple measurement variables using Transformer-based self-attention. At each time step, the measurement variables are treated as individual tokens, and self-attention is performed across the variable dimension. This enables the Relation Expert to extract relational features representing the consistency of each measurement signal with the states of other plant variables. In contrast to the Temporal Expert, the Relation Expert focuses on inter-variable relationships rather than the temporal waveform of an individual signal.

The temporal feature maps generated by the two experts are subsequently converted into sensor-wise window representations through evidence-based temporal pooling. Instead of uniformly averaging all time steps, the pooling process emphasizes time points that provide stronger evidence of abnormal behavior. This allows locally occurring spurious-signal characteristics to be retained in the window-level feature representation.

The window-level representations obtained from the Temporal Expert and Relation Expert are then combined using a learnable Gating Mechanism. The gating module determines the relative contribution of the two expert representations according to the characteristics of the input signals. The fused representation therefore incorporates both the temporal waveform information of individual sensors and the relational information among multiple measurement variables.

Finally, the fused representation is passed through a sensor-wise classification head to estimate whether each measurement variable contains a spurious signal within the input window. Through this parallel expert architecture, the proposed framework separately extracts temporal and relational characteristics and integrates them through adaptive gated fusion for spurious signal detection.

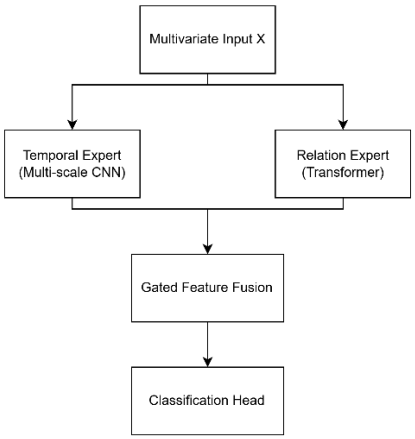

Fig. 1. Overall architecture of the proposed Dual-Expert spurious signal detection framework

3. Experiments and Results

3.1 Experimental Setup

To evaluate the proposed Dual-Expert spurious-signal detection model, experiments were conducted using 22 measurement variables from Kori Units 3 and 4. Data from Kori Unit 3 collected from 2006 to 2019 were used for model training, while Kori Unit 3 data from 2020 to 2024 were used for validation. Kori Unit 4 data collected from 2006 to 2024 were used as the test dataset to evaluate the performance of the trained model on measurement data from a different unit.

Each input sample consisted of a sequence of 120 time steps. Sliding windows were generated using a stride of 10 for the training dataset and a stride of 30 for the validation and test datasets. The resulting datasets consisted of 736,261 training windows, 87,693 validation windows, and 333,117 test windows.

The latent feature dimension of the model was set to 128. The Temporal Expert employed multi-scale convolutional kernels of 3, 7, 15, and 31, followed by residual convolutional blocks with dilation factors of 1, 2, 4, and 8. The Relation Expert consisted of three Transformer encoder layers with four attention heads and a feed-forward dimension of 256. Top-5 evidence pooling was applied to obtain representative window-level features from both experts. The model was trained

with a batch size of 256 and a learning rate of 1 × 10^-4 for up to 10 epochs. The total number of trainable parameters was 908,229.

The detection performance was evaluated using precision, recall, and F1-score. The reported F1-score was calculated using micro averaging across all measurement variables, where the true positives, false positives, and false negatives from all sensors were aggregated before calculating the overall score.

3.2 Experimental Results

Among the evaluated epochs, epoch 9 showed the highest validation F1-score. On the Kori Unit 3 validation dataset, the proposed model achieved a precision of 0.936, a recall of 0.824, and an F1-score of 0.877.

When applied to the Kori Unit 4 test dataset, the model achieved a precision of 0.787, a recall of 0.858, and an F1-score of 0.821. Although the test dataset was obtained from a different unit from that used for training, the model maintained an F1-score above 0.8.

Dataset	Precision	Recall	F1-score
Kori Unit 3 Validation	0.936	0.824	0.877
Kori Unit 4 Test	0.787	0.858	0.821

4. Conclusion

In this study, a Dual-Expert framework was proposed for spurious-signal detection in nuclear power plant measurements. The proposed model consists of a Temporal Expert for extracting temporal waveform characteristics of individual measurement signals and a Relation Expert for capturing relationships among multiple measurement variables. The features extracted by the two experts are integrated through a learnable gating mechanism to provide sensor-wise spurious-signal predictions.

The model was trained and validated using measurement data from Kori Unit 3 and subsequently evaluated using Kori Unit 4 data. The proposed model achieved a precision of 0.936, a recall of 0.824, and an F1-score of 0.877 on the Kori Unit 3 validation dataset. When applied to the Kori Unit 4 test dataset, the model achieved a precision of 0.787, a recall of 0.858, and an F1-score of 0.821. These results show that the proposed framework can detect spurious signals from multivariate nuclear power plant measurement data and maintains reasonable detection performance when applied to data from a different unit.

However, the sensor-wise evaluation showed differences in detection performance among measurement variables, indicating that further improvement is required for signals with relatively low detection performance. Future work will investigate the individual contributions of the Temporal and Relation Experts through comparative experiments and improve the robustness of the model for measurement variables with different signal characteristics.

ACKNOWLEDGEMENTS

This work was supported by KOREA HYDRO & NUCLEAR POWER CO., LTD (No. 2024-Tech-09)

This work was supported by Korea Institute of Energy Technology Evaluation and Planning(KETEP) grant funded by the Korea government(MOTIE)(RS-2024-00403194, Next-Generation Nuclear Technology Creation IP-R&D Talent (Human Resources) Development Project)

REFERENCES

[1] Korea Atomic Energy Research Institute.. "Development of fatigue monitoring system in nuclear power plants" (KAERI/RR-1307/93). (1994)
[2] Sebastian Zug. " Fault-Handling in Networked Sensor Systems ".
[3] Sojin Shin. "Multi-sensor data-based anomaly detection and diagnosis of a pumped storage hydropower plant" (2023)