Preliminary Study on Energy-Guided Diffusion for Leakage Signal Augmentation in Nuclear Power Plant Secondary Systems

DongYun Cho[a,b*], You-Rak Choi[a*]
[a]*Nuclear System Integrity Sensing & Diagnosis Division, Korea Atomic Energy Research Institute (KAERI), Daejeon,* South Korea
[b]*University Science and Technology (UST), Daejeon, South Korea*
[*]*Corresponding author:* yrchoi@kaeri.re.kr

***Keywords :* nuclear power plant, energy-guided diffusion, leakage signal augmentation**

1. Introduction

The rapid growth of artificial intelligence (AI) and data centres is expected to greatly raise worldwide electricity consumption. Among various energy sources, nuclear power is gaining interest as a major, carbon-free energy option capable of functioning 24-hours. As a result, ensuring the safety and reliability of operating nuclear power plants (NPPs) has become more critical.

In the secondary systems of NPPs, data-driven AI models, especially the supervised learning models, have shown strong performance in fault diagnosis. Nonetheless, these techniques inherently necessitate abundant labeled data. Although normal operational data is plentiful, abnormal data, such as leakage signals, is difficult to collect. This leads to data imbalance that constrains model performance and real-world applicability.

To address this problem, this study investigates the generative AI, particularly energy-guided diffusion, to augment leakage data. The suggested method intends to produce leakage signals that align with the statistical properties of real-world leakage data while maintaining physically meaningful signal characteristics.

2. Background

Diffusion models have recently surfaced as robust generative frameworks capable of understanding sophisticated signal properties. Denoising Diffusion Probabilistic Models (DDPM) produce samples by gradually introducing noise to data during the training and learning to reverse this process through repeated denoising [1]. This approach facilitates reliable training and generates high-quality samples in diverse fields.

Denoising Diffusion Implicit Models (DDIM) build upon DDPM by introducing deterministic and non-Markovian properties to the sampling [2]. DDIM enhances computational efficiency by decreasing sampling steps while preserving the generation quality.

Diffusion frameworks can include extra constraints during sampling to steer generation towards particular traits. Among various strategies, this study employs energy-based guidance [3]. Such controllable sampling is particularly important for leakage signal synthesis, where maintaining signal characteristics is vital.

3. Methods and Results

The study employs ultrasonic acoustic signals gathered from a wireless sensor placed in a laboratory setting that simulates a pressurized pipe loop fitted with a pressurizer. The experimental arrangement allows for the collection of both normal and leakage signals. Normal signals exhibit relatively stable spectral distribution, whereas leakage signals demonstrate intensity amplification throughout the entire distribution, Fig. 1.

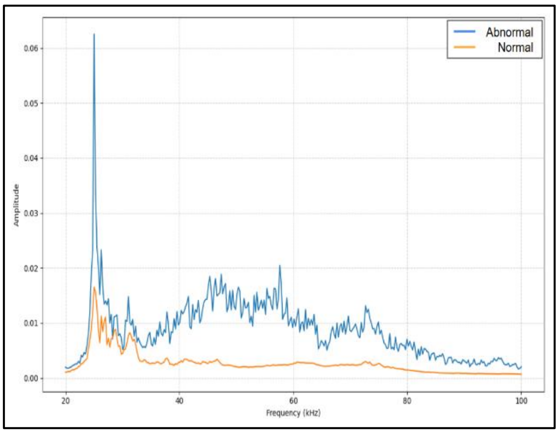

Fig. 1. Visual representations of normal and leakage data in the frequency domain.

This study adopts the DDIM diffusion framework, and the model is trained exclusively on normal signals to understand the baseline behaviour. The signals are transformed into fixed-size representations suitable for the diffusion architecture while preserving their spectral characteristics.

For leakage signal augmentation, the energy-based guidance is integrated into the DDIM reverse update:

$$x_{t-1} = x_{t-1}^{DDIM} - \lambda \nabla E(\hat{x}_0) \quad \textbf{(1)}$$

where x_{t-1}^{DDIM} denotes the standard DDIM update, $\hat{x}_0$ indicates the predicted clean signal by the diffusion model, and $E(.\ )$ is an energy function that captures leakage characteristics. The gradient of E represents the direction that promotes leakage-like properties, while λ controls the guidance strength during each sampling step.

3.1 Specific Region Weighted (SRW) Sampling

SRW sampling adds an emphasis on leakage-relevant frequency regions. These regions are heuristically selected based on prior analysis of normal and leakage signals: low (0-50th bin), middle (51-160th bin), and high (161-320th bin) regions. Leakage behaviours is also defined by band-wise energy and peak amplification. These statistical properties are derived from reference leakage signals collected by the wireless sensor. The energy function is then developed to measure the difference between these desired statistics and those of the generated signal.

During the sampling process, the gradient of this region weighted energy is calculated with respect to the projected clean signal, indicating a local spectral shift towards leakage behaviour, Fig. 2.

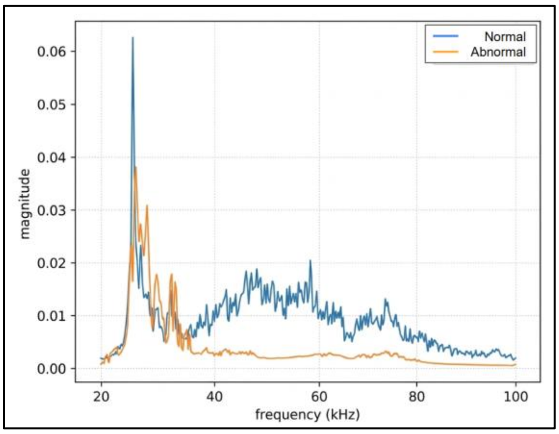

Fig. 2. SRW sampling result compared to real leakage signal.

SRW sampling successfully generates amplified peak in the low range, creating features that resemble genuine leakage signals. However, inconsistencies exist outside the low range, where the created spectrum retains patterns similar to normal signal characteristics. This suggests that region-focused guidance is effective for local feature sampling, but falls short in replicating global leakage properties.

3.2 Global L2 Sampling

Global L2 energy guidance characterizes energy as the L2 distance between the statistics of the generated signal and the reference leakage signal across the entire distribution. This formulation works throughout the spectrum rather than focusing on particular frequency bands like SRW sampling.

During sampling, the gradient of the L2 energy is integrated into the DDIM reverse process, and the result is illustrated in Fig. 3.

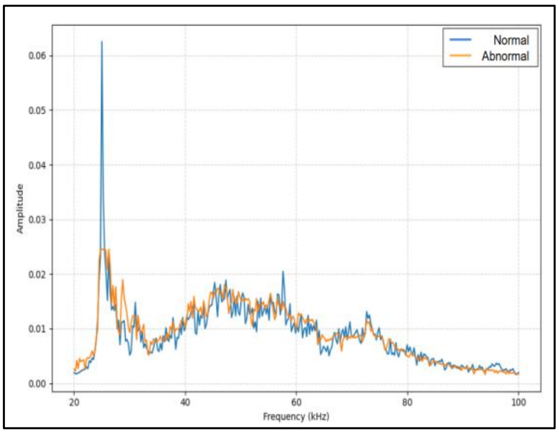

Fig. 3. Global L2 sampling result compared to leakage signal.

Global L2 guidance generates a signal that closely matches the general leakage pattern across frequency bands. However, the peak in the low range is underrepresented, suggesting restricted influence over particular local characteristics.

4. Conclusions

The results show a balance between the accuracy of local features and the overall spectral distribution. Although SRW sampling is effective in capturing peak characteristics in the low region, global L2 guidance offers a better resemblance of the overall leakage signal distribution. To leverage both strengths, future research will focus on a multi-conditional energy function that dynamically shifts gradients from structural alignment to local detail matching based on the diffusion timestep. This preliminary study suggests a promising path for generating high-fidelity leakage signals and addressing data scarcity in nuclear fault diagnosis.

REFERENCES

[1] J. Ho, A. Jain, and P. Abbeel, Denoising Diffusion Probabilistic Models, In *Advances in Neural Information Processing Systems* (*NeurIPS*), 2020.
[2] J. Song, C. Meng, and S. Ermon, Denoising Diffusion Implicit Models, In *International Conference on Learning Representations* (*ICLR*), 2021.
[3] J. Yu, Y. Wang, C. Zhao, B. Ghanem, and J. Zhang, FreeDoM: Training-Free Energy-Guided Conditional Diffusion Model, http://arxiv.org/abs/2303.09833, 2023.