Algorithm for the Detection of Signal Failure in the Emergency Situation Using VAE-LSTM

Younhee Choi, Gyeongmin Yoon, Subong Lee , Jonghyun Kim *
Department of Nuclear Engineering, Chosun University, 309 Pilmun-daero, Dong-gu, Gwangju, 61452, Republic of Korea.
Corresponding author: jonghyun.kim@chosun.ac.kr*

1. Introduction

The condition of sensors is critical because the decision for control action, either by the operator or by the automatic controller, depends on the plant state reflected by sensors [1]. Recently, the interest in autonomous control is increasing and then the reliability of signal becomes more important for the success of it. For this reason, on-line monitoring (OLM) techniques of sensors and signals have been an active research area in nuclear power plants (NPPs).

Many researches have been performed for the signal validation so far. The methods for signal validation can be divided into model based or data-driven approaches. The model-based approach was applied in early studies of this topic by understanding the physical mechanism of the system and presenting an accurate model to detect the failure of signal. Zhao proposed a physical model based approach combining the simulation model and Principal Component Analysis (PCA) to compensate for the disadvantages of the physical model and defects of Helical Coil Steam Generator Systems [2]. This method was applied to the case of a secondary side tube blockage. Ning and Chou proposed a mathematical mode base for the sensor defect detection, and applied it to NPP's systems [3].

Data-driven approaches are using historical operation data without presenting an accurate model. Examples includes Principal Component Analysis (PCA), Multivariate State Estimation Techniques (MSET), Auto-Associative Kernel Regression (AAKR), Artificial Neural Networks (ANN). Pantoni proposed multiple failure signal validation in NPPs using ANN in accident scenarios [4]. Holbert used Process Empirical Modeling (PEM) and Fuzzy ANN for determining signal failure [5]. Eryurek and Upadhyaya proposed an adaptive back-propagation network for signal validation of NPPs [6]. Nabeshima et al. proposed a methodology to detect anomaly signals using Associative ANN for the typical normal operational patterns in NPPs [7]. Choi et al. proposed a signal validation method using a supervised learning for NPPs in an emergency situation [8].

However, those approaches have several issues in detecting sensor and signal failures. First, some of proposed methods were focused on the failure detection in the steady state [3, 5-7] in which the signal is not changing rapidly. For instance, in an emergency situation in which the signal is changing dramatically over time, those methods are not applicable. Second, many data-driven approaches applied the supervised learning and thus are able only to detect trained failures [4-8]. Since there are thousands of signals in NPPs and then different types of failures, it is practically impossible to train all the failure types of signals.

In this light, this study suggests an algorithm to detect signal failures in the emergency situation using the Variational Auto Encoder (VAE)-Long Short Term Memory (LSTM) which is an unsupervised learning method. First, a structure of algorithm to detect signal failure is developed by using the VAE-LSTM. Then, the training and test data are collected by using a compact nuclear simulator (CNS) for the loss of coolant accident (LOCA). Then, an algorithm using thresholds of reconstruction error has been developed. Finally, the suggested algorithm has been tested.

2. Methodology

2.1 VAE

VAE is a modified version of Auto-Encoder (AE), one of representative unsupervised learnings that are trained to reconstruct input data. The structure of VAE is shown in Fig. 1.

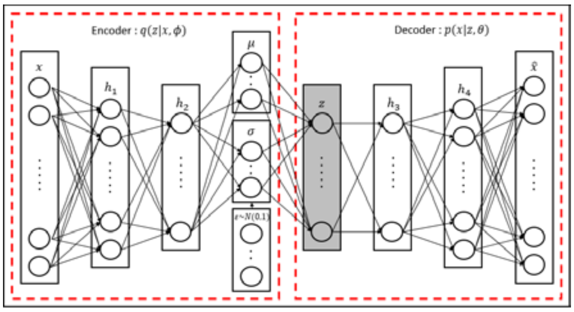


Fig. 1. The architecture of VAE

VAE adds a variational constrain that the latent variable z is subject to a normal distribution and the decoder starts with sampling from the distribution. Accordingly, VAE can map the training data to a normal distribution and generate new samples from the distribution [9]. Encoder and Decoder are given by probabilistic function $q(z|x,\phi)$ and $p(x|z,\theta)$. q is the approximate posterior called adversarial model or encoder, while p is the likelihood of x given z called generative model or decoder [9, 10]. VAE is a special unsupervised learning generative model trying to reconstruct its input data with $\hat{x}$ as close as possible by minimizing [11].

2.2 LSTM

LSTM is based on the Recurrent Neural Network (RNN) and capable of learning long-short term dependency problem [8]. The key to LSTM is the cell state (C_t), which enables the information to flow along is unchanged. The distinctive feature of the LSTM structure is the gate structure, which consists of the forget gate, the input gate, and the output gate. The structure of LSTM is shown in Fig. 2.

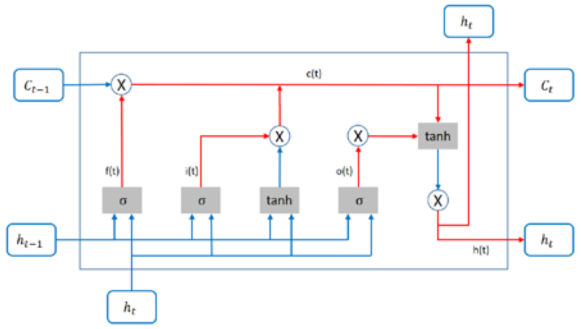


Fig. 2. The architecture of LSTM

c_{t-1}: Initial cell state(Input)
c_t: Final cell state(Output)
h: hidden layer activation
x: current input

The functions are computed by Equations (1) to (5) as below:

$$i_t = \sigma(x_t W_{xi} + h_{t-1} W_{hi} + b_i) \quad (1)$$
$$f_t = \sigma(x_t W_{xt} + h_{t-1} W_{hf} + b_f) \quad (2)$$
$$o_t = \sigma(x_t W_{xo} + h_{t-1} W_{ho} + b_o) \quad (3)$$
$$c_t = f_t c_{t-1} + i_t tanh(x_t W_{xc} + h_{t-1} W_{hc} + b_c) \quad (4)$$
$$h_t = o_t tanh(c_t) \quad (5)$$

where W is the weight matrix of each gate and b is the bias.

At the forget gate (f_t), it reflects some of the previous cell state (C_{t-1}) for the cell state (C_t). It is kept or discarded according to the previous output and the present value. The input gate (i_t) adjusts the value after the input signal (x_t) has passed through the complete connection layer of tanh as an activation function. Finally, the input signal (x_t) passes through the output gate. The output gate (o_t) considers past and modified input data, by adjusting the input signal (x_t) to the tanh and making the output signal.

3. Algorithm for the Detection of Signal Failures

An algorithm for the detection of signal failures in the emergency situation has been suggested as shown in Fig. 3. The algorithm consists of four steps: Input Preprocessing, signal reconstruction by using VAE-LSTM, Output Post-processing, and determination of signal failure.

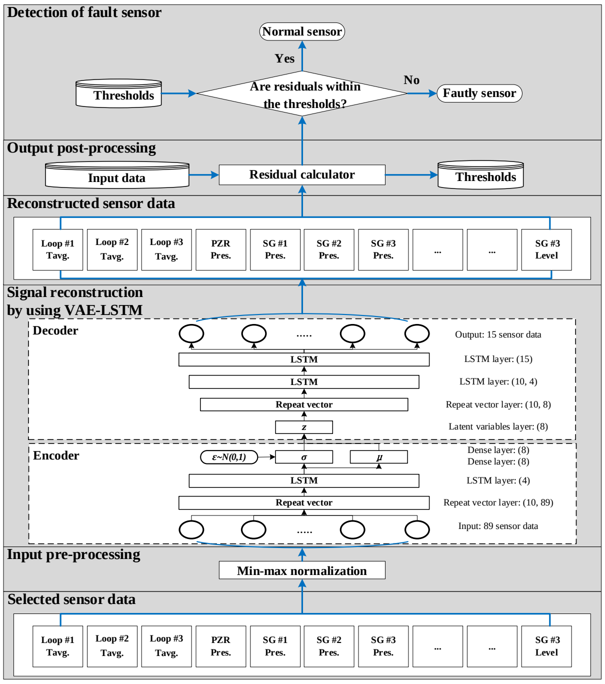


Fig. 3. Algorithm for signal validation

3.1 Input Pre-processing

In this step, a normalization of input values was performed to improve the performance by converting the signal value. As the signal values have different scales, normalization can prevent convergence at the local minimum. The min-max normalization method is applied. The maximum and minimum values are determined from the training data and the input is calibrated within the range of 0 to 1 through Eq. 6.

$$X_{norm} = \frac{(x - x_{min})}{(x_{max} - x_{min})} \quad (6)$$

3.2 Signal Reconstruction

In this step, the reconstructed data are generated by using the VAE-LSTM. This step attempts to produce the reconstructed data which is similar to the input data.

3.3 Output Post-processing

The post-processing step compares a normalization of input data and a normalization of reconstructed data and generates a reconstruction error. The error is calculated by Eq. 7.

$$\text{Reconstruction error} = \left(x_{input} - x_{reconstruction}\right)^2 \quad (7)$$

3.4 Determination of Signal Failures

This step determines the signal failure based on the threshold. If the reconstruction error exceeds the threshold, this signal is determined to be a failure.

4. Experiments

4.1 Data set

The training and test data were collected using CNS. The data represent the normal values of signal in the emergency situation. The total 81,000 data were collected for 54 scenarios in the LOCA, as shown in Table II. Ninety percent (90%) and ten percent (10%) of collected data are used for training and testing, respectively.

Table II: The database used for network training

Initiating Situation	Number of scenarios	Number of training sets
Cold/hot leg loss of coolant accident (LOCA)	54	81,000

4.2 Training

The VAE-LSTM model was trained so that the model produces the reconstructed value as same as the input value. Total 89 inputs including process parameters and component states were selected through correlation analysis for the target signals. The outputs are 15 signal that indicate main process parameters in the LOCA scenario.

4.3 Threshold

The upper line of the threshold of reconstruction errors to detect signal failures are determined from the reconstruction errors in the training. The threshold is calculated by the method suggested by Shewhart [12]. The center line (CL) and upper control limit (UCL) are calculated as follows:

$$\mathrm{UCL} = \mu + 3\sigma \tag{8}$$
$$\mathrm{CL} = \mu \tag{9}$$

The μ and σ refer to the mean value and standard deviation of reconstruction errors. The UCL is used as the criteria to determine the signal failure. The signal is detected to be faulty when the output value generated by the model exceeds the threshold. Table I shows an example of UCLs for four signals.

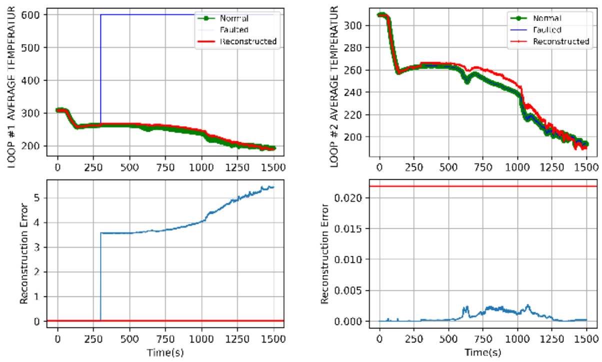
Table I: Threshold

	Loop1 Tavg	Loop2 Tavg	PZR pressure	SG1 pressure
UCL	0.0263	0.0218	0.0271	0.002

4.4 Result

This study verified the suggested algorithm using the test data. Three types of failures are considered: stuck at high, low, and current value.

Fig. 4 presents an example of signal failure detections for the “stuck at high (600°C)” of RCS Loop #1 Coldleg temperature in the LOCA scenarios. The upper of Fig. 4 (a) shows that the faulty signal is injected, while the lower indicates that the reconstruction error exceeds the threshold and thus the algorithm detects the signal failure. Fig. 4 (b) shows how the algorithm works for the normal signals. The reconstruction error for the RCS Loop #2 Coldleg temperature does not exceed the threshold.



a. RCS Loop #1 Coldleg Temperature

b. RCS Loop #2 Coldleg Temperature

Fig. 4. VAE-LSTM output when stuck at the RCS loop1 Tavg high degrees

5. Conclusions

this study suggested an algorithm to detect signal failures in the emergency situation using the VAE-LSTM which is an unsupervised learning method. The data were collected from the CNS for the LOCA scenario. The algorithm was trained and tested for 15 signals.

Acknowledgment

This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Science, ICT & Future Planning(N01190021-06)

REFERENCES

[1] S. Lee, J. Kim, Signal Validation Algorithm using Deep Learning Methods, Korean Nuclear Society Spring Meeting, 2020.

[2] K. Zhao, B. R. Upadhyaya, Model based approach for fault detection and isolation of helical coil steam generator systems using principal component analysis, IEEE Transactions on Nuclear Science, pp.2343-2352, 2006.
[3] J. N. Ning, H. P. Chou, Construction and Evaluation of Fault Detection Network for Signal Validation, IEEE Transaction on Nuclear Science, pp.943–947, 1992.
[4] P. F. Fantoni, A. Mazzola, Multiple-failure signal validation in nuclear power plants using artificial neural networks. Nuclear technology, 113(3), 368-374, 1996
[5] KE. Holbert ,B. R. Upadhyaya, Empirical process modeling technique for signal validation, Annals of Nuclear Energy, pp.387-403, 1994.
[6] E. Eryurek, B. R. Upadhyaya, Sensor Validation for Power Plants using Adaptive Backpropagation Neural Network, IEEE Transactions on Nuclear. Science, 37(2), pp.1040–1047, 1990.
[7] K. Nabeshima, T. Suzudo, K. Suzuki, Real-time nuclear power plant monitoring with neural network. Journal of nuclear science and technology, pp.93-100, 1998.
[8] J. Choi, S. Lee, Consistency Index-Based Sensor Fault Detection System for Nuclear Power Plant Emergency Situations Using an LSTM Network, Sensors, Vol.20, no.6, 2020.
[9] R. Yao, C . Liu, L. Zhang, P. Peng, Unsupervised Anomaly Detection Using Variational Auto-Encoder based Feature Extraction, IEEE International Conference on Prognostics and Health Management (ICPHM), pp.1-7, 2019.
[10] A. Geron, Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow, O'Reilly Media, Inc., Released 2019.
[11] J. Pereira, M. Silveira, Unsupervised anomaly detection in energy time series data using variational recurrent autoencoders with attention, IEEE International Conference on Machine Learning and Applications (ICMLA), pp.1275-1282, 2018.
[12] W.A. Shewhart, Economic Control of Quality of Manufactured Product, New York: Van Nostrand, 1931.