An accident diagnosis algorithm with untrained accident identification

Jaemin Yang, Jonghyun Kim*

Department of Nuclear Engineering, Chosun University, 309 Pilmun-daero, Dong-gu, Gwangju 501-709, Republic of Korea
*Corresponding author: jonghyun.kim@chosun.ac.kr

1. Introduction

All actions at nuclear power plants (NPPs) are carried in accordance with procedures. However, in the case of an accident or a transient, there are several factors that can cause human errors. First of all, many indicators and alerts are generated at the same time that can aggravate interpretation of the situation. In addition, due to conditions such as time pressure or sudden change of parameters, diagnosis is known as one of difficult tasks for operators. These features can lead not only to the delay in effective response but also to more severe consequences from inappropriate response [1, 2].

Under this circumstance, there are a variety of approaches for operator assistance systems and algorithms to reduce the burdens of operators. Among them long short term memory (LSTM) is considered as one of the best artificial intelligence (AI) techniques for solving pattern recognition problems and nonlinear problems. In addition, the LSTM can handle long time-sequential data corresponding to the time-dominant dynamic feature of the NPP [3].

Handling cases that are out of coverage of system is an important issue in developing operator support systems. If the system tries to produce any result for situations that it cannot actually diagnose and then provides a wrong result for operators, this may lead to operator's inappropriate situation awareness. Therefore, for untrained accidents, it is more desirable to answer "unknown" or "don't know" to the operators than to provide a wrong diagnosis result.

This study suggests an accident diagnosis algorithm by combining the LSTM and Auto Encoder (AE). This algorithm is developed to perform the diagnosis of accident as well as the identification of untrained events. The LSTM is used for identifying an accident, while the AE is to determine whether a situation is trained or not. The algorithm is also implemented with a compact nuclear simulator (CNS) based on the Westinghouse 930MWe three loops pressurized water reactor (PWR) as a test bed.

2. Methodologies

This section describes the LSTM and AE for developing an accident diagnosis algorithm. LSTM is applied to implement the network modeling of the accident detection algorithm. In addition, AE is used in conjunction with LSTM to design the untrained accident identification function.

2.1 LSTM

The LSTM, which improves the limitations of recurrent neural networks (RNNs) that cannot deal with long sequences, can reflect the dynamic time-series feature of NPPs, thus, it is suitable for modeling the accident diagnosis algorithm in NPPs. A most distinctive feature of LSTM, compared to conventional RNNs, is the gate structure. The gate structure consists of input gate, forget gate and output gate. The output from the input is regulated by how much it will be reflected through the input gate (i_t), how much forget it will be through the forget gate (f_t), and how much it will be output through the output gate (o_t). Equation (1) to (4) represent each gating logic and input modulation (g_t), respectively.

$$g_t = \phi(W_g \cdot [h_{t-1}, x_t] + b_g) \qquad (1)$$
$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \qquad (2)$$
$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \qquad (3)$$
$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \qquad (4)$$

The input conditioning node is represented in equation (1) and tanh activation function is denoted by ($\emptyset$). Equation (2) represents the input gate and the sigmoid activation function is denoted by σ. The output value (0 or 1) will be calculated by this activation function. The forget gate and output gate are represented in equation (3) and (4). Through this structure of gating logics, the effect of previous state information on the current state can be reflected appropriately, the information associated with the current input can be updated, and the level of impact on the output can be determined.

2.2 Auto Encoder

To deal with untrained or unknown accidents, this study applies AE, i.e., one of the prominent unsupervised learning methods. AE learns functions that approximate input values to represent output values well. Fig. 1 shows the structure of AE. It consists of an encoder that can encode the input to the hidden layer and a decoder that can decode the encoded hidden unit (i.e., latent variable) and outputs the same size representing the input.
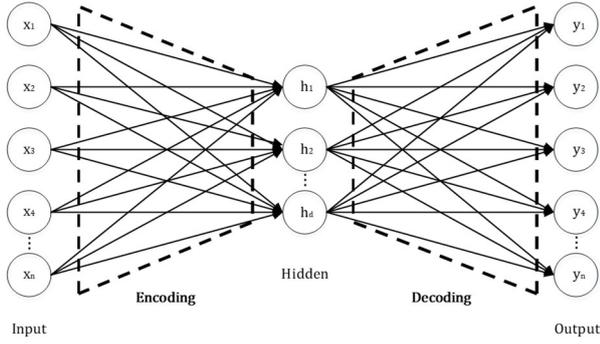
Fig. 1. A structure of auto encoder.

According to the input $x \in R^n$ and the hidden representation $h(x) \in R^m$, these are described in Equation (5), in which the nonlinear activation function is denoted by f(z). The logistic sigmoid function $f(z) = \frac{1}{1+\exp(-z)}$ is applied. $W_e \in m \times n$ and $b_e \in R^m$ mean a weight matrix and a bias vector, respectively. The hidden representation of the network output and reconstruction $y \in R^n$ are described in Equation (6), where $W_d \in n \times m$ and $b_d \in R^n$ mean a weight matrix and bias vector, respectively.

$$h(x) = f(W_e x + b_e) \qquad (5)$$
$$y = f(W_d h(x) + b_d) \qquad (6)$$

After performing the process of extracting and reconstructing feature expressions from the input data via the encoder and decoder, the parameters $\theta(W_e, b_e)$, $\theta'(W_e, b_e)$ are optimized to minimize the loss function (L) as in Equation (7). As a loss function, the square of the error between inputs and outputs are used as

described in Equation (8). According to these equations the algorithm is trained to decrease the loss by regulating the parameters.

$$(\theta, \theta') = argmin_{\theta, \theta'} \frac{1}{n} \sum_{i=1}^{n} L(x, y) \qquad (7)$$
$$L(x, y) = \| x - y \|^2 \qquad (8)$$

3. Accident Diagnosis Algorithm

3.1 Algorithm modeling

An accident diagnosis algorithm is developed for diagnosing accidents as well as handling untrained events. Fig. 2 shows the overview of the algorithm for accident diagnosis and the untrained accident identification. For the accident diagnosis, taking into account a certain number of NPP input sequences, the algorithm can diagnose an accident by capturing a pattern (i.e., NPP trend). Input variables are selected based on procedures, considering their importance that can affect the system availability. Through the network comprised of LSTM layers, the final diagnostic value is regulated and outputted via the output layer.

For the untrained accident identification, inputs are compressed as latent variables through the encoder, and then these are represented via the decoder based on latent variables. The reconstruction errors from differences between represented outputs and inputs are calculated, and then it can be identified whether it is trained considering the threshold value. If the reconstruction error is higher than the threshold value, then it is classified as untrained data. In this study, the maximum value of errors (i.e., maximal reconstruction error during training) is utilized as threshold.
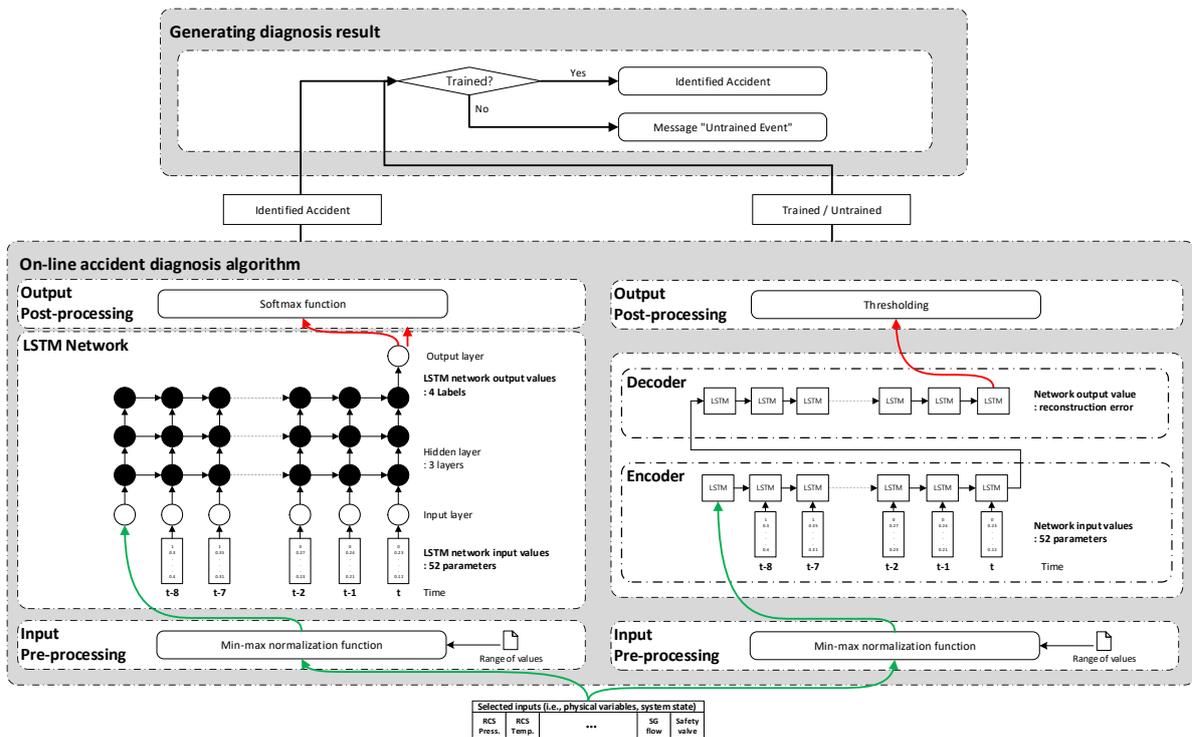
Fig. 2. Overview process of the algorithm

3.2 Training of the network

The algorithm is trained and implemented using the CNS developed by Korea Atomic Energy Research Institute. The algorithm consists of three hidden layers of 64 batch sizes and 100 epochs are trained based on 52 scenarios (i.e., 9,549 datasets, about 20% are used for validation sets) without main steam line break (MSLB). Table 1 and table 2 describe training scenarios and test scenarios used for the algorithm.

Table I: Scenarios used for training and test

Initiating Events	Number
Loss of coolant accident (LOCA)	40
Steam generator tube rupture (SGTR)	12
Main steam line break (MSLB)	0
Total	52

4. Results

The algorithm is validated with two test scenarios. The malfunction is injected at 10 second. Fig. 3 shows the accident diagnosis result with untrained accident identification. In case of Fig.3, which is one of the trained scenarios (i.e., LOCA), figure (a) shows an accident diagnosis result and it shows numerically stable diagnosis result according to time. Also, figure (b) shows the untrained accident identification result, and reconstruction errors are below than threshold, which is a maximum value of trained reconstruction errors.
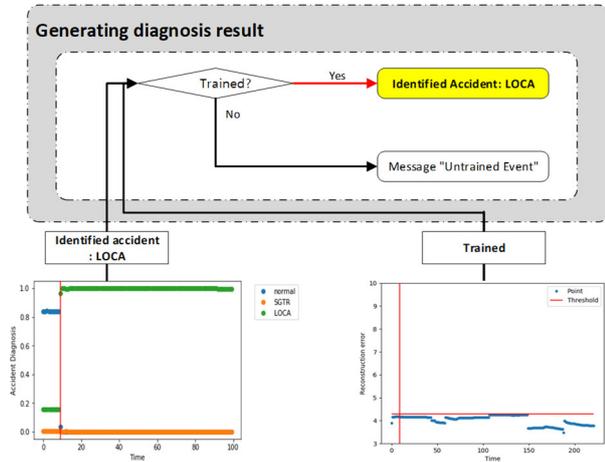
Fig. 3. Trained case, LOCA in loop2_40cm^2

However, Fig. 4 represent accident diagnosis and untrained accident identification results under the untrained accident (i.e., MSLB). Figure (a) shows dynamically changing diagnosis result, also figure (b) shows the result that the latter part of reconstruction errors are higher than the threshold.
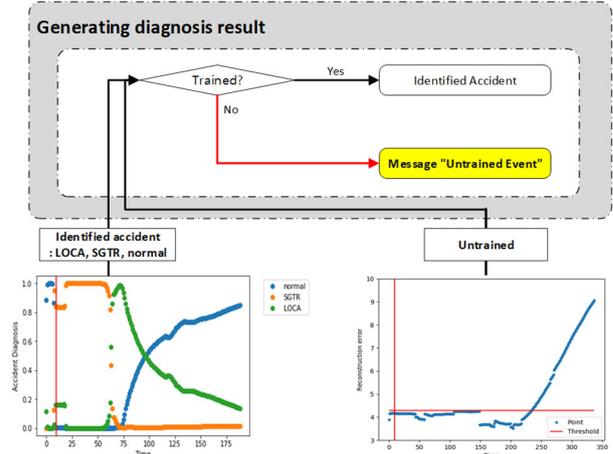
Fig. 4. Untrained case, MSLB in loop2_900cm^2

5. Conclusion

This study proposes an algorithm for accident diagnosis with untrained accident identification to lesson burden of operators under emergency. As a result of accident diagnosis, it shows applicability of the algorithm for accident diagnosis allowing "don't know" response.

ACKNOWLEDGEMENT

REFERENCES

[1] Meister, D., Cognitive behavior of nuclear reactor operators. International journal of industrial ergonomics, 16(2): p. 109-122, 1995.
[2] Safety-related decision making at a nuclear power plant. Nuclear engineering and design, 185(2-3): p. 335-345, 1998.
[3] Yang, J. and J. Kim, An accident diagnosis algorithm using long short-term memory. Nuclear Engineering and Technology, 50(4): p. 582-588, 2018.
[4] Pascual, D.G., Artificial intelligence tools: decision support systems in condition monitoring and diagnosis. Crc Press, 2015.
[5] Bartal, Y., J. Lin, and R.E. Uhrig, Nuclear power plant transient diagnostics using artificial neural networks that allow "don't-know" classifications. Nuclear Technology, 110(3): p. 436-449, 1995.
[6] Antoˆnio, C.d.A., A.S. Martinez, and R. Schirru, A neural model for transient identification in dynamic processes with "don't know" response. Annals of Nuclear Energy, 30(13): p. 1365-1381, 2003.