Transactions of the Korean Nuclear Society Spring Meeting
Jeju, Korea, May 18-19, 2023

Inferring Severe Accident Scenarios in Nuclear Power Plants with Reinforcement Learning (RL) and Supervised Learning (SL) Approaches: Part 1 SL Development

Yeonha Lee[a] , Seok Ho Song[a], Semin Joo[a], Kyusang Song[b], Sung Joong Kim[c] and Jeongik Lee[a*]

[a]Department of Nuclear and Quantum Engineering, N7-1 KAIST, 291 Daehak-ro, Yuseong-gu, Daejeon, Korea 34141
[b]KHNP CRI, 70, Yuseong-daero 1312beon-gil, Yuseong-gu, Daejeon, Korea 34101
[c]Department of Nuclear Engineering, Hanyang University, 222 Wangsimni-ro, Seongdong-gu, Seoul, Korea 04763
*Corresponding author: jeongiklee@kaist.ac.kr

1. Introduction

Severe accidents can lead to severe reactor core degradation and ultimately releasing radioactive material from nuclear power plants to surrounding [1]. Thus, mitigating severe accident consequence is essential to reduce risk of operating a nuclear power plant. Severe Accident Management Guidelines (SAMG) have been developed and used in different countries to mitigate the consequences of severe accidents. However, since SAMG is not a prescriptive procedure, training operators to determine and apply appropriate mitigation strategies under severe accident conditions is necessary.

To train operators under unexpected severe accident conditions, a reinforcement learning (RL) can be utilized to generate variations of severe accident scenarios other than those included in the probabilistic safety assessments (PSA). RL can train an agent that interacts with the environment and learns in the direction of maximizing rewards. The RL agent receives a state and a reward from the environment, learns from the experience and chooses the action with the highest reward. If the reward is in the direction of creating more severe consequences under the given conditions, a new accident scenario can be developed from the RL agent. It is noted that, in this study and companion papers, Part 2 and Part 3, the severity of the accident is measured in terms of the time of reactor pressure vessel (RPV) failure.

The severe accident prediction tool used in the RL environment requires to predict the future state of environment in an accelerated manner. To meet this requirement, an artificial neural network (ANN), a type of model trained by supervised learning (SL) can be utilized. The ANN can approximate non-linear and complex mathematical models, and calculations can be performed very quickly by using the pre-trained network. Therefore, after producing learning data using a severe accident analysis code, the ANN can be trained using the generated dataset.

In this study, it is first assumed that the trained ANN model predicts the thermal-hydraulic (TH) variable with hourly intervals, enabling it to interact with RL for hourly basis in problem time. It is noted that a feed-forward neural network model using only the data of the previous hour as an input was created by the same authors and validated previously [2] to serve the same purpose. However, the authors are proposing that the performance of the ANN model can be improved further by developing a model using information from 3 previous time steps. In other words, if the model uses the previous 3 hours of data to predict the next hour condition instead of 1 hour, the performance can be enhanced. Convolutional neural network (CNN), Long short-term memory (LSTM) and a combination of CNN and LSTM layer models are newly developed for this study, and their performances are compared and presented.

2. Methodology

This section will provide an explanation of the dataset generation process and neural network structure used in this study.

2.1 Dataset Generation

This section explains the process of dataset generation for the SL model, which is the environment of the RL model. The objective of the RL model is to determine the failure time of components to generate the most severe accident scenarios (i.e. making RPV to fail as early as possible). To ensure the ability to predict phenomena in the event of a component failure at any time, the dataset was generated by randomly sampling the failure time of various components. The time when mitigation strategy is in place was also randomly sampled. Each scenario has a length of 72 hours in problem time. The failure probability of component is uniformly assumed and the value is 1/2. The time was sampled an hourly basis.

As shown in Table I, the candidate components for failure were determined to be the ones that failed in a Total Loss of Component Cooling Water (TLOCCW) accident; high-pressure injection pump, low-pressure injection pump, charging pump, containment spray pump, motor-driven auxiliary feedwater pump, heat exchanger, and reactor coolant pump seal. Additionally, as shown in

Table II, three mitigation strategies were also sampled and considered while generating dataset for the SL training; steam generator (SG) feedwater injection (SAMG 1), reactor coolant system depressurization (SAMG 2), and reactor coolant system injection (SAMG 3). 10,679 scenarios were generated and simulated using the MAAP 5.03 code [3] for the reference pressurized water reactor-type nuclear power plant.

Table I. Selected failed components based on TLOCCW accident scenario

Failed Components
High pressure injection pumps
Low pressure injection pumps
Containment spray system pumps
Charging pumps
Motor-driven auxiliary feed water pumps
Heat exchangers
Reactor coolant pump seal

Table II. Selected mitigation strategies for TLOCCW accident scenario

Mitigation Strategy	
SAMG-1 Injection into Steam Generator	External injection
SAMG-2 Depressurize Reactor Coolant System	Safety depressurization system valve
SAMG-3 Injection into Reactor Coolant System	External injection

2.2 Structure of neural network

This section outlines the design of a neural network model capable of predicting thermal hydraulic (TH) variables for every hour. A total of 7 TH variables were chosen to be predicted with the trained ANN model. These variables are all observable from the main control room by an operator, and they are summarized in Table III. Moreover, the selected variables are important to the reactor pressure vessel failure.

Table III Selected TH variables for neural network model

Target TH Variables
Primary system pressure
Cold leg temperature
Hot leg temperature
Reactor vessel water level
Steam Generator pressure
Steam Generator water level
Max Core Exit Temperature

The neural network takes not only the TH variables as input parameters but also binary values indicating whether a component has failed at that time and whether mitigation strategies have been implemented at that time as well. This is shown in Fig. **1**. The newly developed model utilizes information covering three previous hours to predict the 7 TH variables for the next hour.

As shown in Table IV, three neural network models were tested in this study: 1D CNN, LSTM, and a combination of the two. A convolutional neural network (CNN) is known for the rapid learning speed, as it calculates weights by extracting features from only a portion of the entire dataset. One-dimensional CNN (1D CNN) are commonly used for time series prediction as they learn by extracting features from nearby temporal data. In contrast, a Long Short-Term Memory (LSTM) network is a type of recurrent neural network (RNN) specifically designed for time series data and aims to enable long- and short-term memory capabilities. Therefore, both 1D CNN and LSTM are commonly used for time series prediction, making them suitable choices for this study.

Before training three networks, the input and the output of the dataset were normalized for data preprocessing. The dataset was then divided into training, validation, and testing sets with a ratio of 7:2:1, respectively.

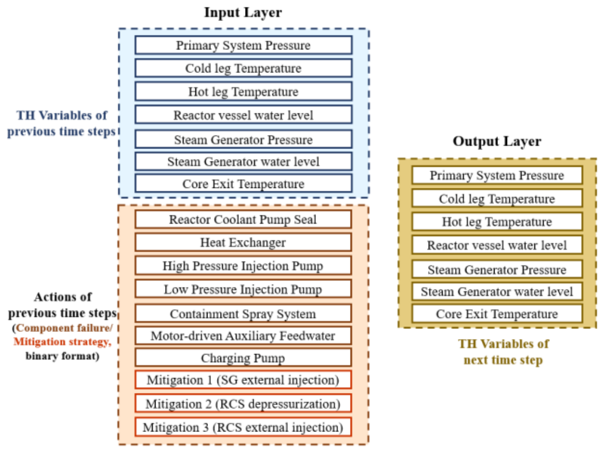

Fig. 1 Structure of input and output layer of SL model

Table IV Structure of neural network models

	CNN model	LSTM model	Combined model
Hidden layer structure	1D CNN	LSTM	1D CNN
	Dense	LSTM	LSTM
			LSTM

2.3 Performance Metrics

For model comparison, the mean absolute error (MAE) was used to evaluate regression performance. The mean absolute error equation, shown in Eqn. (1), calculates the average of the absolute differences between the actual and predicted values.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|y_i - \hat{y}_i\right| \qquad (1)$$

3. Results and Discussions

The training of all three models was successful, and mean absolute error (MAE) values are obtained. The same hyperparameters were used for all three models, and they are presented in Table V. As the MAE becomes smaller, the difference between the predicted and the actual data becomes smaller, indicating better regression performance. Therefore, the combined model shows the best regression performance, followed by CNN and LSTM. The reason why the combined 1D CNN and LSTM model demonstrates better performance is likely attributed to the fact that passing through the CNN layer reduces the length of the sequence. This enables the LSTM to effectively detect longer patterns.

This can be observed from Table VI, where the MAE results for the validation and the test sets are presented. Fig. 2 to Fig. 4 show the scatter plot of primary system pressure, cold leg temperature, and SG water level from the combined model, which has the least error. R^2 values are also shown in the graph, and it can be seen that all of them show very high values of 0.99 or more.

Table V Hyperparameters of neural network

Loss	Mean squared error
Optimizer	Adam
Epochs	500 with early stopping
Conv1d filters	100
LSTM units	100

Table VI Mean absolute error of valid and test set

	CNN model	LSTM model	Combined model
Valid	0.01013	0.01425	0.00822
Test	0.01029	0.01444	0.00843

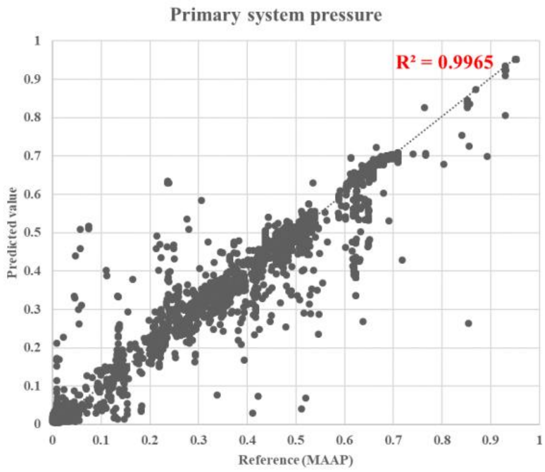

Fig. 2 Scatter plot of primary system pressure from combined model

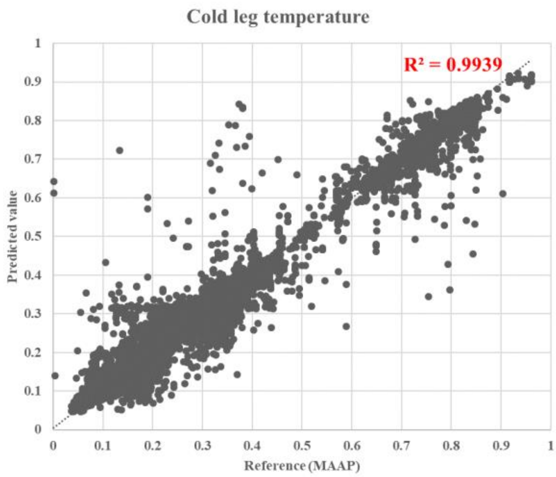

Fig. 3 Scatter plot of cold leg temperature from combined model

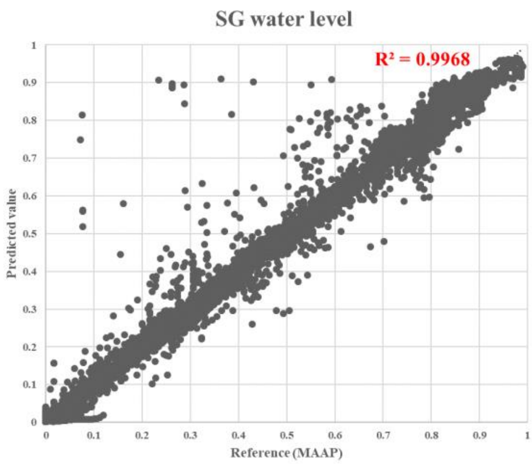

Fig. 4 Scatter plot of SG water level from combined model

4. Summary and Conclusions

Three neural network models were constructed to predict TH variables on an hourly basis from datasets representing over 10,000 different severe accident scenarios. These scenarios can be viewed as a subset or variations of TLOCCW of a conventional large PWR. The performances of three models were compared using mean absolute error (MAE). The model with the smallest MAE was a combination of CNN and LSTM, suggesting that combining CNN with LSTM leads to the development of better performing model than using LSTM alone. Scatter plots for each variable predicted from the combined model showed very good regression performance, indicating that the trained SL can be an excellent environment for the RL agent to be trained. Part 2 and Part 3 companion papers will be focusing on the RL agent training using the SL model generated from this study.

ACKNLOGEMENT

This work was supported by KOREA HYDRO & NUCLEAR POWER CO., LTD (No. 2020-Tech-01).

REFERENCES

[1] INTERNATIONAL ATOMIC ENERGY AGENCY, IAEA Nuclear Safety and Security Glossary, Non-serial Publications , IAEA, Vienna (2022).

[2] Yeonha Lee, "Development of accelerated prediction method using artificial neural network for Nuclear Power Plant Severe Accident application," Ms. Thesis, Korea Advanced Institute of Science and Technology (2022).

[3] EPRI, “Modular Accident Analysis Program (MAAP5) Version 5.03 – Windows,” Fauske & Associates, Inc, August 2014.