Prediction of Reactor Vessel Water Level in Severe Accident Circumstances Using LSTM

Young Do Koo [a], Ji Hun Park [a], and Man Gyun Na [a*]
[a]*Department of Nuclear Engineering, Chosun Univ., 309 Pilmun-daero, Dong-gu, Gwangju 61452*
[*]*Corresponding author: magyna@chosun.ac.kr*

1. Introduction

Many studies that suggested operator support systems for nuclear power plants (NPPs) have been being carried out. Several operator support systems were designed for effective actions and mitigation in an abnormal state or an accident circumstance. Among them, the systems showed its capability for tasks such as fault detection, diagnosis of an abnormal state or an accident, and prediction of safety-related factors in NPPs by deploying artificial intelligence algorithms. Furthermore, with the introduction of machine learning methods from conventional support vector machines [1], fuzzy neural networks [2] to state-of-the-art deep learning methods with feedforward deep neural network (DNN) [3] or recurrent neural network (RNN) [4] architectures, many methods were able to be applied to various NPP factors, and therefore their performances were shown and being advanced.

In an effort to predict a safety-critical NPP factor, long-short term memory (LSTM) neural network [5], of which structure is based on RNNs', was used to predict reactor vessel (RV) water level under postulated severe accident circumstances of the NPPs. The LSTM was utilized in the study due to the fact that it is well known for its better stability for time series prediction in large-scale networks, and less vulnerable to vanishing gradient problem than typical RNNs.

For application to the LSTM and establishment of a prediction model, modular accident analysis program (MAAP) code [6] was used to obtain simulated data. The data were comprised of time-dependent values of variables gained by simulating the severe accident circumstances originated from postulated loss-of-coolant accidents (LOCAs) and steam generator tube rupture (SGTR).

In this paper, prediction performance of an established LSTM model is presented when limited instrumentation signals from the aforementioned simulated data were applied. In addition, prediction performance of two deep learning methods for a NPP factor can be assessed by comparing the proposed LSTM model with the DNN model designed in previous study [7].

2. Deep Learning Methods

Effectiveness of each deep learning method differs depending on its inherent characteristics although most of the deep learning methods are based on artificial neural network (ANN) structure inspired by inter-connection between neuronal cells in the human brain. The RNNs, basic structure of the LSTM, are dynamic networks with RNN cells.

2.1 Recurrent Neural Networks

In RNNs, the reason why the word '*recurrent*' is used is due to the loop that each RNN cell perform same process and its current outputs are affected by results previously calculated as described in Fig. 1. Since RNN cells share its values calculated at every time step, thus, RNNs or RNN-based networks are considered as more proper methods to process time series data, that is, sequential data.

In Fig. 1, hidden state, h_t, is computed using inputs x_t, hidden state at previous time step h_{t-1}, and nonlinear activation function (e.g. hyperbolic tangent) expressed as Eq. (1). Once the hidden state, h_t, arrives in the output layer by forward propagation, an overall output of the RNNs, $\hat{y}_t$, is computed using Eq. (2).

$$h_t = f\left(W_{xh}x_t + W_{hh}h_{t-1} + b_h\right) \tag{1}$$

where W_{xh} is weights transferred from input layer to hidden layer and W_{hh} is weights shared from previous to next hidden cells.

$$\hat{y}_t = W_{hy}h_t \tag{2}$$

where W_{hy} is weights from hidden layer to output layer.

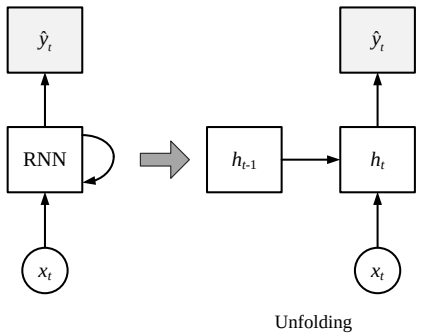

Fig. 1. Simple Vanilla Recurrent Neural Networks [8]

To optimize the deep learning methods, gradient descent and backpropagation algorithms are generally used. These are utilized to update weights and overall

output by propagating errors backward in an input layer direction until a point on a curve, where loss computed using error between target and predicted values is minimized, is finally found. Unlike a typical ANN-based method (e.g. DNNs), however, backpropagation through time (BPTT) is utilized on behalf of existing backpropagation algorithm. This is because each RNN cell depends on gradients of the hidden state at previous time step as well as current time step.

Despite information sharing which is an advantage of the RNNs, the RNNs can be vulnerable to vanishing gradient problem in multiple cells due to its repetitive multiplying process.

2.2 Long-Short Term Memory Neural Networks

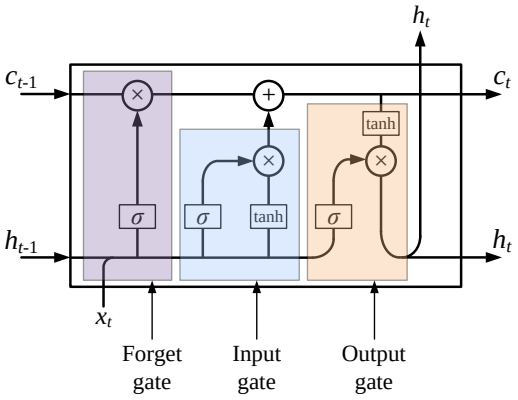

Fig. 2. Long-short Term Memory and its Memory Cell [9]

To deal with the vanishing gradient in the RNN structure, one of the variations, LSTM, has been presented. The LSTM used in the study is the same as the entire RNN structure and weight flow while architecture of each cell differs from each other. The main characteristic of the LSTM is the usage of 'memory cell' containing cell state and three types of gates (forget, input, and output) as shown in Fig. 2.

The cell state performing like a conveyer belt is determined by cell state at previous time step c_{t-1}, and the forget and input gates expressed as Eqs. (3) and (4), and propagated to the next cells. Three types of the gates use the sigmoid function to determine whether a cell saves the information or not. If an output of the sigmoid function is 0, the information is forgotten.

$$f_t = \sigma\left(W_{xh,f}x_t + W_{hh,f}h_{t-1} + b_{h,f}\right) \tag{3}$$

$$i_t = \sigma\left(W_{xh,i}x_t + W_{hh,i}h_{t-1} + b_{h,i}\right) \tag{4}$$

The hidden state at current time step is calculated using the computed cell state, the nonlinear activation function, and the output gate of Eq. (5), and then transferred into the next memory cells.

$$o_t = \sigma\left(W_{xh,o}x_t + W_{hh,o}h_{t-1} + b_{h,o}\right) \tag{5}$$

Although computing time for the LSTM tends to increase on account of complex architecture of the cell, the LSTM can solve long term dependency problem (refer to Fig. 3).

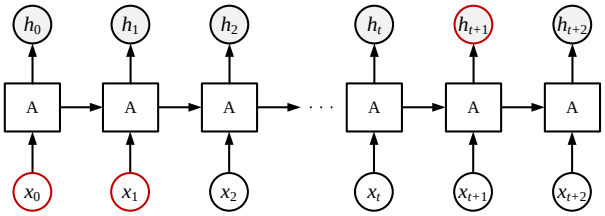

Fig. 3. Long Term Dependency Problem [9]

3. Applied Data for Training LSTM Model

3.1 Simulation Data for Assumed Accidents

To obtain the simulated data, the MAAP code was utilized to simulate assumed LOCAs in hot-leg and cold-leg, and SGTR. Each postulated accident was under the circumstance that high-pressure and low-pressure safety injection did not function normally. In addition, its break (or rupture) sizes were variously divided from small to double-ended guillotine break.

Among all the simulated data on the postulated accidents, the number of accident simulations applied to the LSTM model is 600, and these were separated 200 for each accident location, respectively. Specifically, in case of LOCAs in hot-leg and cold-leg, 170 were for larger break sizes and 30 were for smaller sizes, respectively. For SGTR data, 100 simulated data were for smaller and larger rupture sizes, respectively.

3.2 Applied Simulated Signal Values to LSTM

Trend for the desired factors of the NPP systems can be outputted when the assumed NPP accidents were simulated using the MAAP code. The trend on the variables were expressed as time-integrated values after the reactor trip indicated as Eq. (6). The applied data simulated according to every LOCA break or SGTR sizes contain the time-integrated values for various signals from elapsed time after the reactor trip to 3 days.

$$x_j = \int_{t_s}^{t_s+\Delta t} g_j(t)dt, \quad j = 1,\ 2,\ \cdots,\ m \tag{6}$$

To predict the RV water level, only a few types of the simulated signals were applied to the LSTM owing to the assumption that acquiring instrumentation signals was limited. Elapsed time after the reactor trip and containment pressure from the MAAP code, and the

estimated LOCA break size (or SGTR size) were applied, which are the same as the previous study [7].

Even though LOCA break or SGTR size cannot be accurately recognized in actual accident circumstance, these factors have to be determined. Since low estimation errors for the break sizes were shown in previous studies for LOCA diagnosis [10]-[12], the estimated LOCA size was considered as an applicable input for the LSTM in this study.

In each accident case according to break locations and break sizes, approximately 95% simulated signal values were used as 'training data' and the others were 'test data' to verify the trained LSTM model.

4. Prediction Results of LSTM Model

4.1 Utilized LSTM Model for RV Water Level Prediction

Prediction performance of the established LSTM model was presented using root mean square error (RMSE) and it was compared with result of the DNN model of the previous study. Several hyper-parameters making the LSTM model for RV water level prediction optimal were used as follows:

- No. of stacked LSTM layers: 7
- Hidden sizes: 8
- Sequence length: 12
- Activation function: softsign
- Optimization function: Adamoptimizer
- Learning rate: 0.002

4.2 Performance of the LSTM Model

Table I: Prediction Performance of the LSTM Model

Break size	Accident location	Training data	Test data
		RMSE (m)	RMSE (m)
Small	Hot-leg LOCA	0.07	0.11
	Cold-leg LOCA	0.10	0.07
	SGTR	0.07	0.09
Large	Hot-leg LOCA	0.08	0.08
	Cold-leg LOCA	0.11	0.10
	SGTR	0.10	0.12

Among the whole obtained data, the simulated data on specific break size cases were selected as the test data set. In case of hot-leg and cold-leg LOCAs, simulated signal values on less cases were considered as the test data since the number of the simulated data are relatively smaller. On the other hand, the larger number of break size cases were used for large break LOCAs and SGTR.

Table I shows the prediction results for each data set using the established LSTM model. Figs. 4-6 also indicate the performance of the proposed model for the test data in LOCA and SGTR with small break size cases. In Figs. 4-6, black circles (target) are well tracked by red-colored crosses (predicted) except for several missed points.

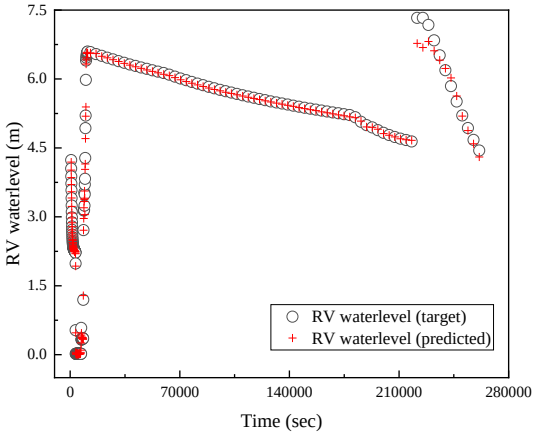

Fig. 4. Prediction of RV Water Level by LSTM Model (small hot-leg LOCA)

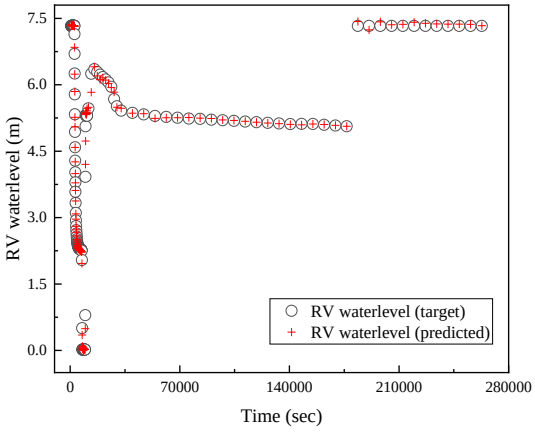

Fig. 5. Prediction of RV Water Level by LSTM Model (small cold-leg LOCA)

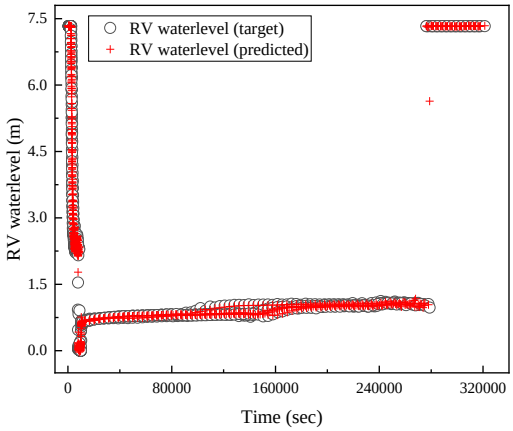

Fig. 6. Prediction of RV Water Level by LSTM Model (small SGTR)

4.2 Comparison of Prediction Results with DNN Model

The prediction results of the LSTM and DNN models were compared in Table II in order to assess the prediction performance on the RV water level. The DNN model used typical ANN framework and its hidden layers were optimized by genetic algorithm (GA) [13] in the previous study.

As described in Table II, RMSE of the DNN model is slightly lower than that of the LSTM model in general. It is considered that this is because the DNN model was with network structure efficiently optimized by the suitable optimization technique. Thus, additional study is needed to apply several well-known optimization methods in addition to the GA to the LSTM since the proposed LSTM model also is expected to have low RMSE.

Table II: Comparison of Performance between LSTM and DNN Models

Break size	Accident type	LSTM model	DNN model
		RMSE for test data (m)	RMSE for test data (m)
Small	Hot-leg LOCA	0.11	0.11
	Cold-leg LOCA	0.07	0.09
	SGTR	0.09	0.03
Large	Hot-leg LOCA	0.08	0.04
	Cold-leg LOCA	0.10	0.13
	SGTR	0.12	0.07

5. Conclusions

This study was carried out to predict one of the safety-critical variables, reactor vessel (RV) water level, using the long-short term memory (LSTM) model. Furthermore, its prediction result was compared to that of the deep neural network (DNN) model of the previous study to assess the performance of deep learning methods for the prediction of the nuclear power plant (NPP) factor. The DNN model has shown better accuracy for RV water level prediction than the LSTM model in this study.

Generally, many efforts (e.g. selecting proper network structure and hyper-parameters) are necessary to establish an optimal deep learning model. Therefore, an LSTM model with enhanced prediction performance will be developed if appropriate parameters are found in additional study.

REFERENCES

[1] V. Vapnik, The Nature of Statistical Learning Theory, Springer, New York, 1995.
[2] M. G. Na, On-line Estimation of DNB Protection Limit via a Fuzzy Neural Network, Nuclear Engineering and Technology, Vol.30, pp.222-234, 1998.
[3] Y. LeCun, Y. Bengio, and G. Hinton, Deep Learning, Nature, Vol.521, pp.436-444, 2015.
[4] I. Goodfellow, Y. Bengio, and A. Courville, Deep Learning, MIT Press, Massachusetts, 2016.
[5] S. Hochreiter and J. Schmidhuber, Long Short-term Memory, Neural Computation, Vol.9, pp.1735-1780, 1997.
[6] R. Henry et al., MAAP4: Modular Accident Analysis Program for the LWR Power Plants, User's Manual, Fauske and Associates Inc., 1994-2005.
[7] Y. D. Koo, Y. J. An, C.-H. Kim, and M. G. Na, Nuclear Reactor Vessel Water Level Prediction During Severe Accidents Using Deep Neural Networks, Nuclear Engineering and Technology, In Press.
[8] F.-F. Li, J. Johnson, and S. Yeung, CS231n 2017: Lecture 10: Recurrent Neural Networks, Stanford University School of Engineering, 2017.
[9] C. Colah, understanding LSTM Networks, http://colah.github.io/posts/2015-08-Understanding-LSTMs/, 2015.
[10] M. G. Na, W. S. Park, and D. H. Lim, Detection and Diagnostics of Loss of Coolant Accidents Using Support Vector Machines, IEEE Transactions Nuclear Science, Vol.55, pp.628-636, 2008.
[11] S. H. Lee, Y. G. No, M. G. Na, K. I. Ahn, and S. Y. Park, Diagnostics of Loss of Coolant Accidents Using SVC and GMDH Models, IEEE Transactions Nuclear Science, Vol.58 pp.267-276, 2011.
[12] K. H. Yoo, Y. D. Koo, and M. G. Na, Identification of LOCA and Estimation of Its Break Size by Multiconnected Support Vector Machines, IEEE Transactions Nuclear Science Vol.64, pp.2610-2617, 2017.
[13] D. Goldberg, Genetic Algorithms in Search, Optimization, and Machine Learning, Addison-Wesley, Massachusetts, 1989.