ISSN : 1226-7244 (Print)
ISSN : 2288-243X (Online)
논문번호 18-03-22

j.inst.Korean.electr.electron.eng.Vol.22,No.3,672~679,September 2018
http://dx.doi.org/10.7471/ikeee.2018.22.3.672

Short-term Load Forecasting of Buildings based on Artificial Neural Network and Clustering Technique

Minh-Duc Ngo*, Sang-Yun Yun*, Joon-Ho Choi*, Seon-Ju Ahn*★

Abstract

Recently, microgrid (MG) has been proposed as one of the most critical solutions for various energy problems. For the optimal and economic operation of MGs, it is very important to forecast the load profile. However, it is not easy to predict the load accurately since the load in a MG is small and highly variable. In this paper, we propose an artificial neural network (ANN) based method to predict the energy use in campus buildings in short-term time series from one hour up to one week. The proposed method analyzes and extracts the features from the historical data of load and temperature to generate the prediction of future energy consumption in the building based on sparsified K-means. To evaluate the performance of the proposed approach, historical load data in hourly resolution collected from the campus buildings were used. The experimental results show that the proposed approach outperforms the conventional forecasting methods.

Key words : Short Term Load Forecasting, Microgrid, Building Load, Clustering, ANN, Feature Selection, Sparsified K-means

Ⅰ. Introduction

In a typical microgrid energy management system (MG-EMS), short term load forecasting (STLF) is one of the critical components, which provides input data to other applications such as optimization and data analysis program in scheduling and operating plans of energy transactions. A large number of load forecasting approaches have been proposed in various papers, which can be categorized into three groups: conventional methods, artificial intelligence (AI) method and hybrid method. The conventional methods are capable of achieving satisfactory results when solving the linear problems. There are several kinds of the conventional approach based on time series models [1]. The regression based approaches including linear regression (LR), multiple linear regression (MLR) [2], regression tree (RT) [3], and support vector regression (SVR) [4] were broadly investigated. Practical load forecasting approaches have to consider both linear and nonlinear factors because the electricity load varies very complicatedly. However, the conventional methods were not able to deal with the nonlinear factors efficiently.

Recently, AI based methods have increased the attention of many researchers and achieved

* Dept. of Electrical Engineering, Chonnam National University
★ Corresponding author
E-mail : sjahn@jnu.ac.kr, Tel : +82-62-530-1738
※ Acknowledgment
This research was supported by Korea Electric Power Corporation (Grant number:R18XA04) and KEPCO Research Institute grant funded by Korea Electric Power Corporation (R16DA11).
Manuscript received Sep. 5, 2018; revised Sep. 14, 2018; accepted Sep. 18, 2018

acceptable forecasting accuracy [5]. Among many AI based methods, artificial neural network (ANN) was widely applied with various network structures such as back propagation neural networks [6], feed forward network [7] and radial basis function neural network [8] for electrical load consumption prediction purposes.

In this paper, we proposed a novel approach based on sparsified K-means clustering and resilient back-propagation ANN (RBP ANN) algorithm. Firstly, all input variables including historical load, time, and weather variables are preprocessed and grouped based on the sparsified K-means clustering algorithm. Then, RBP ANN algorithm was used to forecast the load in each group. The past 24 hours data including time calendar, temperature, and load are investigated to predict the electrical power at the current time. The past 24 hours data is extremely small data in comparison with other published researches which usually use past one week data. The mean absolute percentage error (MAPE) and root mean square error (RMSE) metrics are selected to evaluate the performance of the proposed method. The experimental results show that MAPE and RMSE of the proposed method are smaller than other conventional approaches. The strong correlation between the predicted and real value is depicted by R-squared value.

II. Methodology

This section describes a detailed framework of implementation process from preprocessing data to output predicted forecasting model. All input variables including weather, calendar, and historical load variables are generated and normalized in data pre-processing step to make data feature patterns. Clustering method based on sparsified K-means is adopted to all these features to classify the preprocessed data into different groups. Each data cluster is assigned to the corresponding forecasting model for prediction. The process of the proposed method is summarized in Figure 1.

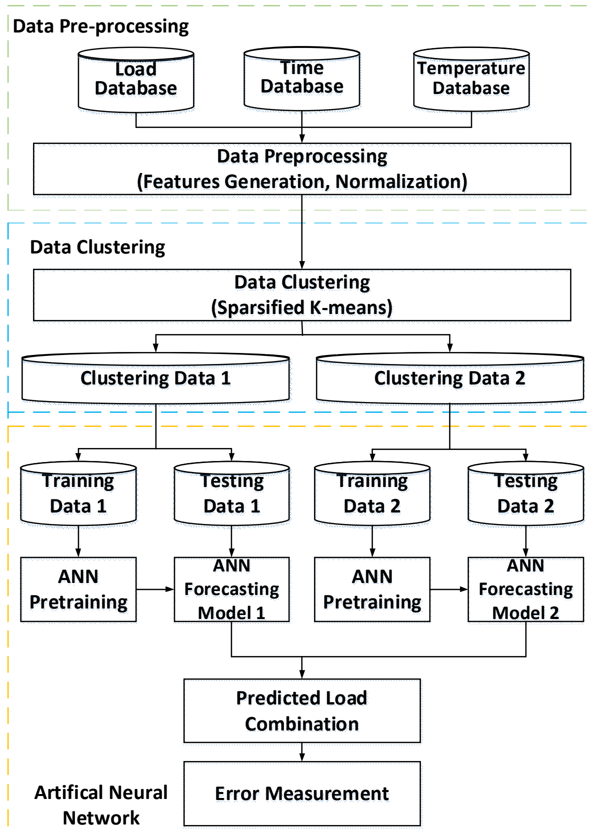

Fig. 1. The proposed load forecasting scheme.

2.1. Data preprocessing

1) Time features

The electricity consumption is critically affected by the time index. The specific time variables such as hour of the day (h = 1, 2⋯ 24), day of the week (d = 1 for Sunday, d = 2 for Monday⋯ d = 7 for Saturday), and month of year (m = 1 for January, m = 2 for February⋯ m = 12 for December) can be coded by their cosine values, as follows:

$$\hat{h} = \cos\left(\frac{\pi . h}{12}\right),\ h = 1, 2, \ldots, 24 \tag{1}$$

$$\hat{d} = \cos\left(\frac{\pi . d}{3.5}\right),\ d = 1, 2, \ldots, 7 \tag{2}$$

$$\hat{m} = \cos\left(\frac{\pi . m}{6}\right),\ m = 1, 2, \ldots, 12 \tag{3}$$

In addition, three binary variables are introduced to classify some variables in more detail. The variable t is used to characterize the time of the day: 0 for the daytime (9:00 am to 18:00 pm) and 1 for the night time (19:00 pm to 8:00 am). The days of week are divided into two groups, i.e. weekends and weekdays, and the variable w is used for this feature. The load of campus building is usually lower on the hot season (from April to October) than the cold season (other months). We used the variable s to denote this seasonal characteristics.

2) Previous load features

The electric loads contain many periodic patterns and the load variation includes autocorrelation. When there is knowledge of previous load values such as from the previous day and from the previous hour, it is easier to predict the load with good accuracy. In this paper, we generate five features from the past recorded load based on the availability of the recorded load dataset from our campus buildings. The features are previous hour load, previous 24 hours load, the average of three closest hour loads, the average of six closest hour load, and the average load of 24 hours before. These variables are denoted by $y(h-1)$, $y(h-24)$, $(\bar{y}_{3h})$, $(\bar{y}_{6h})$, $(\bar{y}_{24h})$, respectively.

3) Temperature features

Temperature is one of the most important factors affecting the load. The main factor causing the forecasting accuracy is that electric demand is not only driven by the temperature of the current hour, but also by the temperature of preceding hours because temperature inside the building reacts slowly to the change of outside temperature. Therefore, we introduce four variables to represent the temperature characteristics: current temperature ($T(h)$), previous hour temperature ($T(h-1)$), average temperature of previous three hours $(\bar{T}_{3h})$, and the temperature at the same hour of previous day ($T(h-24)$).

4) Data normalization

Table 1 summarizes all features used as the input of the forecasting model. All of the features are normalized by using equation (4). Normalization is done to map the data to a uniform scale. Several data normalization techniques, such as min-max, softmax, z-score, are available [9]. Among them, min-max technique preserves all the relationships in the original dataset. Therefore, after computing features variables, the min-max normalization method is applied to scale the input data into the range of [0, 1] as follows.

$$\hat{x} = \frac{x - x_{min}}{x_{max} - x_{min}} \tag{4}$$

Table 1. Summary of selected features.

Features	Selected features	Denoted
Time Features	Normalized hours	$\hat{h}$
	Normalized days	$\hat{d}$
	Normalized months	$\hat{m}$
	Time of days	t
	Weekend and weekday	w
	Season	s
Temperature features	Current temperature	$T(h)$
	Previous hour temperature	$T(h-1)$
	Average temperature of three past hours	$(\bar{T}_{3h})$
	Temperature at 24 hours before	$T(h-24)$
Previous load features	Previous hour load	$y(h-1)$
	Previous 24 hours load	$y(h-24)$
	Average of power during three closest past hours	$(\bar{y}_{3h})$
	Average of power during six closest past hours	$(\bar{y}_{6h})$
	Average of power during 24 closest past hours	$(\bar{y}_{24h})$

2.2. Data Clustering Method

The load forecasting problem becomes challenging due to the high diversity of training dataset. Therefore, dividing the training datasets into smaller subgroups based on their common characteristic is considered as a good method to improve the forecasting accuracy. Selecting the features to apply clustering method is crucial steps, which can lead to process in different ways.

One of the most common method for clustering task with a large amount of data set is K-means clustering approach. The performance function of K-means is the minimum distance from the observations to the centroid of the closest cluster. For the ideal solution, we must find the real centroid of each cluster, but in ordinary K-means algorithm, the centroids will be approximated randomly and then iteratively refined until converge to the global optimum. This will suffer from computational burden and sometimes yields a poor result.

In order to deal with large-scale data which does not require incoherence and distributional assumptions on the data, Farhad and Stephen proposed an advanced compression scheme for accelerating K-means clustering namely, sparsified K-means. This method provides guarantees for principal component analysis in terms of the covariance matrix, and guarantees for K-means in terms of the errors in the center estimators at a given step [10]. The detailed process of sparsified K-means approach is summarized in the references [10] and [11]. In our process, the processed features are treated as the input of sparsified K-means. Each group will be trained separately using neural network.

2.3. ANN Model Development

ANN has been one of the most effective solutions for the prediction of energy consumption in the buildings. A typical ANN comprises three layers including input layer, hidden layer, and output layer as shown in Figure 2.

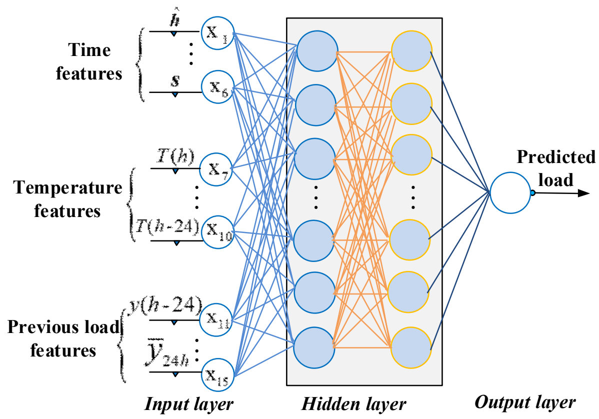

Fig. 2. Inputs-outputs forecasting model.

The activation of a neuron is calculated by summation of the weighted inputs as in equation (5).

$$y = f\left(\sum (w_{ij}.x_{ij})\right) \tag{5}$$

Where y is the output of the neuron, x_{ij} is the input to that neuron, w_{ij} is the weight of the connection of the input to the neuron and f is the transfer function. Sigmoid function is selected as the transfer function of ANN due to the ability to handle the nonlinear problems. Sigmoid function is shown in equation (6).

$$f(x) = \frac{1}{1+e^{-x}} \tag{6}$$

The output of the output layer is calculated by equation (7). The output is the summation of all output at the last hidden layer with the same transfer function.

$$Y = f\left(\sum (w_j.y_j)\right) \tag{7}$$

The training process of an ANN is the changing weight vector to reduce the error between the real output and the predicted output, which is formulated in equation (8).

$$E = \frac{1}{2}\sum (Y_p - Y_r)^2 \tag{8}$$

Where Yp is the predicted output, Yr is the real output, and E is the total error. The most popular training algorithm is the back propagation method which is based on the individual weight update $\Delta_{ij}^{(t)}$ rule as shown in equation (9) [12].

$$\Delta_{ij}^{(t)} = \begin{cases} -\Delta_{ij}^{(t)},\ if\ \dfrac{\partial E^{(t)}}{\partial X_{ij}} > 0 \\ +\Delta_{ij}^{(t)},\ if\ \dfrac{\partial E^{(t)}}{\partial X_{ij}} < 0 \\ 0,\ \text{otherwise} \end{cases} \tag{9}$$

Where the $\dfrac{\partial E^{(t)}}{\partial X_{ij}}$ express the sum of gradient information for all the patterns at the time t and $\Delta_{ij}^{(t)}$ is considered as the direction of updating.

In this paper, we propose to use two hidden layers neural networks with the resilient back propagation training algorithm. The inputs-outputs architecture of forecasting model employed in this study is shown in Figure 2. As mentioned in this section, each group of data will be trained separately. The performance was investigated with various number of neural at each hidden layers from 20 to 30. We chose 23 neural for each hidden layer, which gives the best prediction performance.

III. Experimental Results and Discussion

The proposed methodology is tested on some campus building datasets. The evaluation metrics most frequently used to assess the performance of a model are MAPE and RMSE defined in equations (10) and (11), respectively.

$$MAPE = \frac{100}{T}\sum_{t=1}^{T}\left|\frac{y_r(t) - y_f(t)}{y_r(t)}\right| \tag{10}$$

$$RMSE = \sqrt{\sum_{t=1}^{N}\frac{\left(y_r(t) - y_f(t)\right)^2}{N}} \tag{11}$$

Where $y_r(t)$ is the real value of load at hour t and $y_f(t)$ is the predicted value of load at hour t. T is the total number of hours. N is the number of observed points.

3.1. Comparison with other forecasting methods

We evaluate the performance of different forecasting methods including MLR, SVM, RT, ANN without K-means, ANN combined with K-means for only load.

1) One week prediction results

Figure 3 shows the one week prediction results among methods, Figure 4 shows the load patterns of actual data and predicted load of proposed method. Table 2 summarizes the accuracy of all methods deployed in this paper for each day of the first week from 1/3/2015 to 7/3/2015. The daily and weekly average MAPE using the proposed method are the lower than those of other methods. The best approach reached an MAPE of 1.40%. Taking a closer look at the errors of all day in one week among the methods, we found that the prosed method not only archives the best accuracy in each day but also obtains smallest spreading errors with the range from 1.40% to 3.0%.

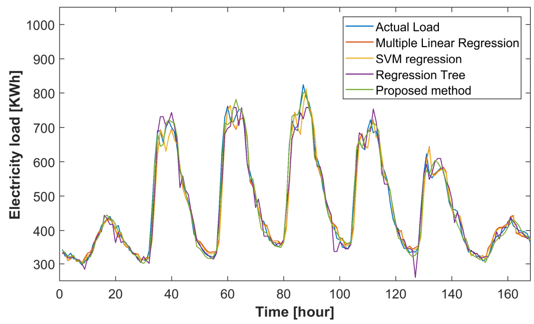

Fig. 3. One week load forecasting results with different methods.

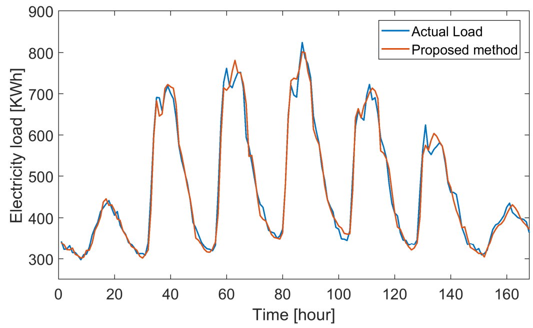

Fig. 4. One week actual load and predicted load pattern of proposed method.

As shown in Figure 5, the correlation between real load and predicted load of the proposed method in one week is extremely high with the R-square correlation of approximately 0.99.

Table 2. MAPE of one week prediction among methods.

Method	MLR	SVR	RT	Proposed method
Day 1	2.61	2.37	2.94	2.06
Day 2	4.67	4.72	4.65	3.01
Day 3	4.72	4.56	3.35	2.80
Day 4	4.56	4.57	3.80	2.27
Day 5	4.44	4.33	4.57	2.96
Day 6	4.41	4.28	4.43	2.61
Day 7	3.34	2.85	2.42	1.40
Average	**4.11**	**3.95**	**3.92**	**2.45**

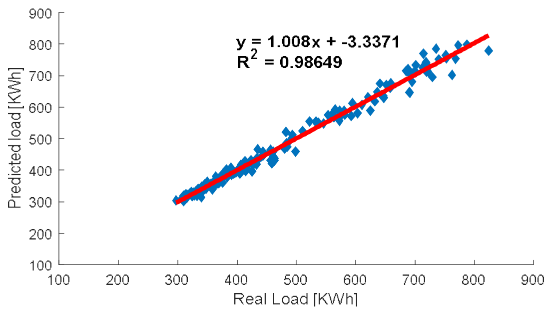

Fig. 5. Correlation between real load and predicted load.

2) One day prediction results

Based on the prediction time range, the more short-time requires instant and high accurate prediction results. Table 3 and Figure 6 show the performances and load profile of one day ahead forecast using various methods. It can be observed that the proposed method not only get the highest accuracy for a whole day but also reach the smallest error at every hour.

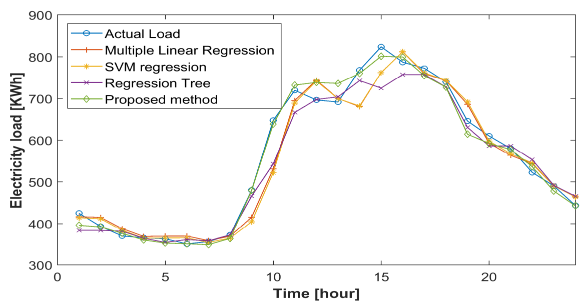

Fig. 6. One day load profile forecasted with different methods.

Table 3. The evaluation of one day predicted results.

Method	MAPE	RMSE
RT	2.939	14.417
MLR	2.610	11.587
SVR	2.37	10.836
Proposed method	**2.061**	**8.861**

3.2. Performance of forecasting model for one year of several building datasets

To evaluate the performance of the proposed forecasting method, we tested the forecasting performance of the proposed model for the one year of several buildings and mapping the errors by box-plot of hourly and daily errors distribution as shown in Figures 7 and 8, respectively. More specifically, in Figure 7, approximately 85% of hourly errors has distribution less than MAPE of 4% in building 6 and Student Hall dataset. In High School and Building3_5 datasets, 85% of hourly errors lie under 5% of MAPE. The results of day-ahead forecast in Building 6 and Student Hall shows that 90% of daily errors are in the range from 1% to 3.5%. In case of High School and Building3_5, around 80% of daily errors are less than 4%. All results prove that well-combined clustering method and neural network result in high accuracy and robust performance of the forecasting model.

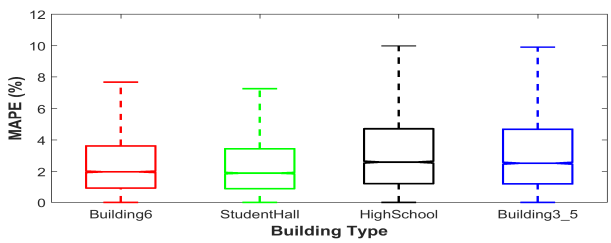

Fig. 7. Distribution of hourly error for one year of several building datasets.

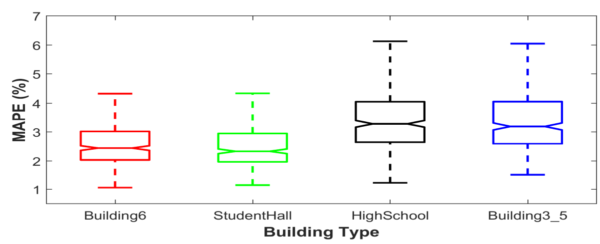

Fig. 8. Distribution of daily error for one year of several building datasets.

Ⅳ. Conclusion

This paper described an artificial neural network approach combined with a clustering method and effective input feature selection to predict short-term energy consumption of buildings with highly nonlinear load profiles. The proposed method is computationally simple and also suitable for analyzing a large set of data whose pattern changes over time. The forecasting model was applied to several campus building loads. The hourly and daily MAPE of the proposed method showed the smallest value comparing to various existing methods. The accuracy and robustness of the proposed method have also been proved from the test results of various building datasets.

References

[1] J. N. Zagrajek and R. Weron, "Modeling electricity loads in California: ARMA model with hyperbolic noise," *Signal Processing*, Vol.82, No.12, pp.1903-1915, 2002.
DOI:10.1016/S0165-1684(02)00318-3
[2] V. M. Bianco and S. Nardini, "Linear regression models to forecast electricity consumption in Italy," *Energy Sources, Part B: Economics, Planning and Policy*, Vol.8, No.1, pp.86-93, 2013. DOI:10.1080/15567240903289549
[3] G. Dudek, "Multivariate Regression Tree for Pattern-Based Forecasting Time Series with Multiple Seasonal Cycles," *Proc. of 38th International Conference on Information Systems Architecture and Technology -ISAT 2017*, pp.85-94, 2017.
DOI:10.1007/978-3-319-67220-5_8
[4] R. K. Jain, K. M. Smith, P. J. Culligan and J. E. Taylor, "Forecasting energy consumption of multi-family residential buildings using support vector regression: Investigating the impact of temporal and spatial monitoring granularity on performance accuracy," *Applied Energy*, Vol.123, pp.168-178, 2014. DOI:10.1016/j.apenergy.2014.02.057
[5] M. A. Daut, M. Y. Hassan, H. Abdullah, H. A. Rahman, M. P. Abdullah and F. Hussin, "Building electrical energy consumption forecasting analysis using conventional and artificial intelligence methods: A review," *Renewable and Sustainable Energy Reviews*, 70, pp.1108-1118, 2017.
DOI:10.1016/j.rser.2016.12.015
[6] N. Ding, C. Benoit, G. Foggia, Y. Besanger and F. Wurtz, "Neural Network-Based Model Design for Short-Term Load Forecast in Distribution Systems," *IEEE Transactions on Power Systems*, Vol.31, No.1, pp.72-81, 2016.
DOI:10.1109/TPWRS.2015.2390132
[7] F. Rodrigues, C. Cardeira and J. M. F Calado, "The Daily and Hourly Energy Consumption and Load Forecasting Using Artificial Neural Network Method: A Case Study Using a Set of 93 Households in Portugal," *Energy Procedia*, Vol.62, pp.220-229, 2014. DOI:10.1016/j.egypro.2014.12.383
[8] W. Mai, C. Y. Chung, T. Wu and H. Huang, "Electric load forecasting for large office building based on radial basis function neural network," *In Proceedings of the 2014 IEEE PES General Meeting Conference & Exposition*, pp.1-5, 2014.
DOI:10.1109/PESGM.2014.6939378
[9] B. S. Kwon, R. J. Park and K. B. Song, "Analysis of Short-Term Load Forecasting Accuracy Based on Various Normalization Methods," *Journal of Korean Institute of Illuminating and Electrical Installation Engineers*, Vol.32, No.06, pp.30-33, 2018.
[10] P. A. Farhad and B. Stephen, "Preconditioned Data Sparsification for Big Data with Applications to PCA and K-means," *IEEE Transactions on Information Theory*, Vol.63, No.5, pp.2954-2974, 2017. DOI:10.1109/TIT.2017.2672725
[11] A. David and S. Vassilvitskii, "k-means++: The Advantages of Careful Seeding," *Proceeding of the 18th Annual ACM-SIAM Symposium on Discrete Algorithms*, 2007.
[12] M. Riedmiller and H. Braun, "A direct adaptive method for faster backpropagation learning: The RPROP algorithm," *In Proceedings of IEEE International Conference on Neural Networks*, 1993. DOI:10.1109/ICNN.1993.298623

BIOGRAPHY

Minh-Duc Ngo (Member)

He received the B.S. and M.S. degree in electrical engineering from Thai Nguyen University, Vietnam in 2005 and 2009. He is currently a PhD student in Chonnam National University. His research interests are Electrical Drive System, Machine Learning and Microgrid.

Sang-Yun Yun (Member)

He received the B.S., M.S. and Ph.D degree in electrical engineering from Soongsil University in 1996, 1998 and 2002. He worked as a Senior Researcher at LS Industrial System Co. in 2009. He worked as a Senior Researcher at Korea Electric Power Research Institute until 2016. He is currently a professor in Chonnam National University. His research interest is Transmission and Distribution Operation System.

Joon-Ho Choi (Member)

He received the B.S., M.S. and Ph.D degree in electrical engineering from Soongsil University in 1996, 1998 and 2002. He is currently a professor in Chonnam National University. His research interests are Distribution Automation System and Modeling, Smart grid operation algorithm.

Seon-Ju Ahn (Member)

He received the B.S., M.S. and Ph.D degree in electrical engineering from Seoul National University in 2002, 2004 and 2009. He is currently a professor in Chonnam National University. His research interests are Distributed Generation, Smart grid, and Microgrid.