Adaptive Clustered Federated Learning with Seasonal Segmentation

Sungwoong Yeom
Chonnam National University
Gwangju 61186, South Korea
yeomsw0421@gmail.com

Shivani Sanjay Kolekar
Chonnam National University
Gwangju 61186, South Korea
shivanikolekar@gamil.com

Kyungbaek Kim
Chonnam National University
Gwangju 61186, South Korea
kyungbaekkim@jnu.ac.kr

ABSTRACT

Recently, research on federated learning has been actively studied to improve the performance of federated learning by creating clusters with similar characteristics of building electricity demand pattern used as input for federated learning. However, because the electricity demand pattern changes at a certain period in the dynamic industry, it is problematic depending on the parameters of the unchanged deep learning model. Accordingly, therefore, it is necessary to develop a deep learning model that can adaptively learn changing electricity demand patterns and a system that can continuously process large-scale data. In this paper, we propose adaptive clustered federated learning with seasonal segmentation. The proposed method makes segmentation and clustering on repeated electricity demand patterns by seasons, and then performs the adaptive federated learning for each cluster. Through extensive evaluation of actual building electricity demand and weather data, the proposed adaptive clustered federated learning technique shows around 5% lower the mean square error than the clustered federated learning technique.

KEYWORDS

Federated Learning, Building Electricity Demand Prediction, Clustering, LSTM

1 INTRODUCTION

With the increase in renewable power generation and decentralization of the market, the building electricity demand prediction technique is essential research to balance the electricity demand and supply in buildings [1]. Recently, as the number of Internet of Things (IoT) devices installed in buildings increases, research using time series deep learning neural networks has become active to reliably predict nonlinearly observed electricity demand [2, 3]. In order to generalize the time-series deep learning neural network for building electricity demand forecasting, large-scale electricity demand profiles should be collected from various clients.

However, due to the small amount of electricity demand profiles in some buildings, the time series deep learning neural network may not enough to be trained. For buildings with scarce data, building managers have taken a centralized approach, working with multiple buildings to gather power demand profiles [4]. This technique generalizes time-series deep learning neural networks collected in a centralized approach and redistributes them to buildings. However, sending data from the client to the central server is expensive in terms of communication costs. Additionally, if new electricity demand pattern is observed from newly acquired data, the data should be sent to a central server for training [5]. At this time, because the client sends data to the central server, the personal information of client may be infringed.

In order to alleviate this privacy infringement problem and expensive communication cost problem, a federated learning approach that can learn distributed across local buildings has been adopted [6]. The federated learning is a distributed machine learning technique in which buildings participating in model learning cooperate to learn a global model under the control of a central entity. The federated learning can preserve privacy by sharing the weights of model in a local building with a central entity instead of electricity demand. However, if a non-i.i.d (not independent and identically distributed) problem such as statistical heterogeneity by irregular behavior of occupants and weather, the convergence and performance of federated learning can be deteriorated.

In order to alleviate this non-i.i.d problem, a clustered federated learning technique that performs clustering of clients who have similar electricity consumption patterns in advance and then aggregates model updates between clients in the same cluster was proposed [8, 9]. However, because the electricity demand pattern is markedly changed by the seasons, it is recommended to perform clustering again when the season changes [10]. In other words, the building electricity demand prediction model needs constant updating to capture changes in some building physical properties over time due to changes in external factors such as weather or internal factors such as interior furniture and occupancy distribution. A solution to this problem is to develop an adaptive ML-based building model that accurately diagnoses the knowledge of individual learners by using intelligent information technology-based systems such as big data and artificial intelligence, and dynamically adjusts the learning level or type according to the results.

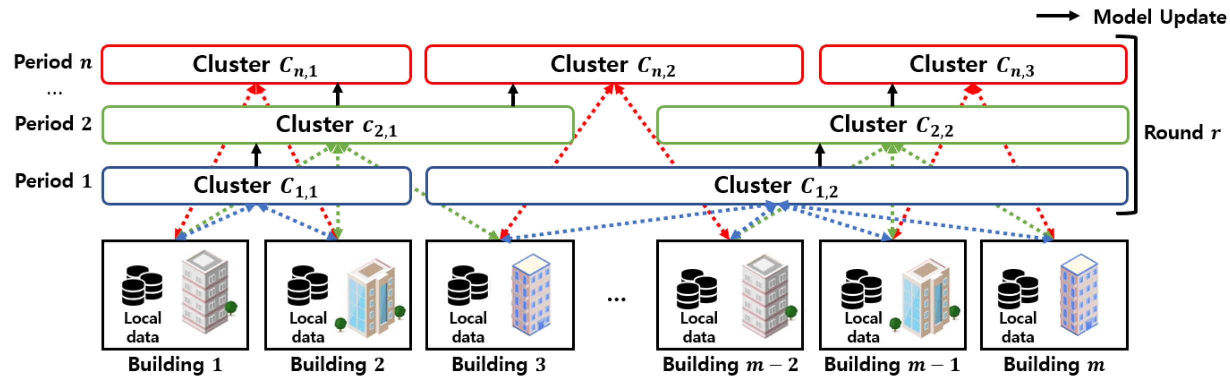

Figure 1: Adaptive Clustered Federated Learning System Architecture

In this paper, we propose adaptive federated learning with seasonal segmentation and clustering of electricity demand patterns. The proposed method has segmentation and clustering for extracting electricity demand patterns observed seasonality by period, and then the federated learning is performed for each cluster. At this time, the adaptive learning is performed to continuously update periodic features. In order to verify the effectiveness of the proposed framework, we evaluated the mean square error of the proposed method according to the clustering result based on actual building electricity demand data and weather data.

2 RELATED WORK

In order to improve the short-term energy consumption prediction techniques, Long Short-Term Memory(LSTM) is showing high performance in predicting electricity demand [11]. However, because LSTM can guarantee performance only when large-scale data is used for training, it is a complicated approach from a computational viewpoint. So far, in order to centrally train the model to generalize, deep learning neural network-based solutions have a common aspect that requires a centralized entity who collects energy profiles from customers.

For this, the neural network training pipeline collects client data for model training from the central server. Sending data from client to server is expensive in terms of communication cost. In addition, in order to adapt the model to new demand patterns in the future, newly acquired data should be transmitted to the central server [5]. In this case, since the clients transmit data itself to the central server, the personal information of client may be infringed. To overcome these communication cost problems and privacy infringement problems, the federated learning can be applied [12], which allows clients to cooperatively train deep neural networks on their own local data without the need for centralization. However, training neural networks with federated learning usually have problems with non-i.i.d data [13], where clients involve different data distributions. As a result, global models created by aggregating different client model updates have poor convergence and poor performance. To alleviate this non-i.i.d problem, a technique for clustering electricity consumers with similar properties is needed.

In order to improve the performance of federated learning for electricity demand forecasting, clustering techniques which groups residential customers according to the similarity of their energy consumption patterns is being studied [8]. When training deep learning neural networks using only consumer pattern identified in the same cluster, the performance of deep learning neural networks is greatly improved. Moreover, because of electricity patterns of consumers are affected by weather (temperature and humidity) [14, 15] and calendar (day or month)[16], it is necessary to perform clustering with each feature [17]. In order to capture changes in electricity demand patterns caused by external and internal factors such as weather and occupancy distribution over time, continuous update of the clustered federated learning is required. According to the results, dynamically adaptive learning is required according to the clustering situation. In this paper, we propose adaptive clustered federated learning with seasonal segmentation.

3 ADAPTIVE CLUSTERED FEDERATED LEARNING SYSTEM

The framework contains buildings M and a cloud G. Let $B = \{b_1, b_2, .., b_m\}$ denote the set of buildings. Each building collects an electricity demand profile consisting of the electricity demand, internal temperature, and internal humidity. The electricity demand profile collected in the m^{th} building is represented as D_m. We assume that federated learning task for cluster C will be assigned to the cloud G. In the cloud G, the clustered federated learning is performed by pre-configuring the clusters C that shows a similar electricity demand pattern. However, the data of the electricity demand profile is repeated by period, and clusters configured for each period can be different. After performing clustering by period, an adaptive machine learning is needed for re-learning the weights of the previous model. In this paper, we propose adaptive clustered federated learning system. Figs. 1 shows the adaptive clustered federated learning system architecture. First, the proposed system has segmentation and clustering using electricity demand patterns by period in which seasonality is prominent. Then the clustered federated learning is performed using segmentation and clustering results on a periodic basis.

3.1 Seasonal Segmentation Clustering

In the presence of dynamic changes in electricity demand patterns, clustering can identify similar patterns, but clustering is impractical due to the high dimensions, characteristic correlation, and noise of long time series data. Changes in daily, weekly, and monthly electricity consumption of consumption data at different aggregation levels were subdivided to identify consumer consumption behavior at different times of the day, along with seasonal influences. The electricity demand pattern is repeated, and the clustering result is different for each period. This period P can be day, week, or month. For seasonal clustering, the electricity demand profile is segmented by period. The number of clusters by period is represented as follows:

$$N = |D_m|/P \tag{1}$$

To cluster electricity demand profiles segmented by daily, weekly and monthly, we use a simple, efficient and scalable K-means algorithm. The clustering result of the n^{th} period is C_n. The cluster id of the m^{th} building is represented as $c_{n,k}$, which is the cluster id of the k^{th} cluster in the n^{th} period. The final clustering id of the m^{th} building is the most frequently observed cluster id.

3.2 STL-LSTM based Electricity Demand Prediction

The proposed electricity demand prediction method estimates electricity demand based on LSTM, a time series neural network model. The proposed model consists of 3 LSTM layers and 3 Dense layers. Of these three Dense layers, the first two layers use a Rectified Linear Unit (ReLU) activation function, and the last layer uses a linear activation function. Each layer contains 10 nodes. The proposed LSTM model predicts the electricity demand for the next 1 hour by using 10 consecutive electricity demand information from a specific time to 9 hours before. To ensure the stability of the prediction for the observed electricity demand, the input vector for the proposed LSTM model is composed of six domains of the electricity demand profiles D : electricity demand, indoor temperature, indoor humidity, month index, hour index, and day type. The proposed LSTM model is trained as an input vector to predict electricity demand.

The first, second and third domains of the input vector are electricity demand, room temperature and room relative humidity. Each domain feature is decomposed into seasonality, trend, and residual through STL decomposition, which is then constructed as an input vector. The fourth and fifth domains of the input vector are hourly and monthly indices. The monthly index has a value between 1 and 12, and the hourly index has a value between 1 and 24. Finally, the sixth domain of the input vector is the type of day of the week. The day type embeds the relationship between day characteristics and electricity demand through one-hot encoding by considering eight categories including holidays. The proposed model is trained through federated learning.

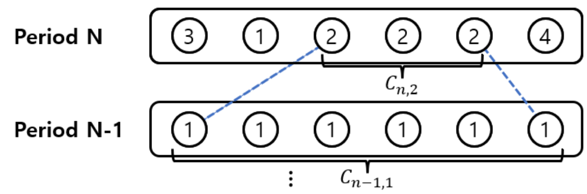

Figure 2: Cluster Selection by Period

3.3 Adaptive Clustered Federated Learning

The clustered federated learning is the method that shares the weight of the deep learning model of buildings included in the cluster and average it. However, the proposed adaptive clustered federated learning requires that the buildings contained in the cluster differ by period, and the method transfer the model weights of the period $n-1$ to the model weights of the period n. In other words, in order to update the model weight, the next clusters should be selected for each cluster. Figs. 2 shows cluster selection by period. The id of the 6 buildings belonging to the 1st cluster $C_{n-1,1}$ in the period $n-1$ is different in the period n. In the period n, 2 with the largest number is selected. This cluster selection algorithm is performed when federated learning proceeds.

In the federated learning, the cloud G initializes the weight of global model ω_0^g, parameters such as learning rate μ, and batch size b. The global model ω^g is transmitted to each building and learned using the electricity demand profiles $D_{m,n}$ corresponding to period n. The formula for updating model learned in the i^{th} building using the global model weight $\omega_r^{g_{n,k}}$ of k^{th} cluster in n^{th} period is as follows:

$$\omega_{r+1}^{m_i} = \begin{cases} \mu \cdot \nabla \mathrm{loss}(\omega_0^g; b), & r = 0 \\ \mu \cdot \nabla \mathrm{loss}(\omega_r^{g_{n,k}}; b), & \text{else} \end{cases} \tag{2}$$

The weight of model $\omega_r^{m_i}$ learned in the i^{th} building at round r is transmitted to the cloud G. The cloud G aggregates updated model weights $\omega_r^{m_1}, \dots, \omega_r^{m_i}$ from 1 to m buildings and learns them through FedAvg. In this case, $|D_{m_i}|$ means the size of the data collected in the i^{th} building, and the weight of the global model $\omega_r^{m_i}$ at round r is averaged through FedAvg. The formula of FedAvg is as follows:

$$\omega_{r+1}^{g_{n,k}} = \sum_{m_i \in C_{n,k}}^{C_{n,k}} \frac{|D_{m_i}| \cdot \omega_{r+1}^{m_i}}{|D_{m_i}|} \tag{3}$$

4 Evaluation

In order to evaluate the proposed adaptive clustered federated learning method for building electricity demand prediction, we compare the performance of the adaptive clustered federated learning system the clustered federated learning and the basic federated learning. We use the building electricity demand data set of CityLearn[18] pre-calculated by CitySim[19], a building energy simulator, for the experiment. The CityLearn environment consists of 4 anonymized climate zones and 9 building clusters per climate zone. Each building data set consists of electricity demand data measured hourly over a year. This experiment is performed on a machine with two NVIDIA GeForce RTX 2080 SUPER installed, based on a Docker container, with a cluster consisting of one Chief node and 36 Worker nodes.

4.1 Comparison of Clustering Results per Period

In this chapter, we identify representative electricity demand profiles by period and determine the optimal number of clusters. First, Calinski Harabasz Index(CHI) and Davies-Bouldin Index (DBI) were used to identify the most suitable number of clusters. This is a commonly used measurement to assess how well a data set is segmented.

Table. 1 shows the comparison of cluster validity indices DBI and CHI according to the number of clusters using the K-means algorithm. DBI captures the separation and compactness of all pairs of data clusters and returns a system-wide similarity measure of each cluster compared with its most similar neighbor. That is, the lower the value of DBI, the better the clustering. CHI knowns as the variance ratio criterion that measures how well clusters are defined. It returns the CHI score calculated by the ratio of average inter-cluster and intra-cluster sum of squares. That is, the lower the DBI, the better the clustering. That is, the higher the value of CHI, the better the clustering. Accordingly, when the number of clusters is 8, efficient clustering is achieved.

Table. 2 shows the list of buildings that make up a cluster when there are 8 number of clusters. Z and B stand for the climate zone and the building id, respectively. In Table. 2, Clusters 5, 6, 7, and 8 consist of buildings that belong to different climate zones whereas clusters 1, 2, 3, and 4 consist of buildings that belong to the identical climate zones. Therefore, the performance of clustered federated learning is compared between clusters configured buildings in the identical climate zone and different climate zones. However, as the length of the electricity demand profile increases, the performance of clustering can be lower.

Figure. 3 shows the representative shape of demand profiles by period. First, we analyze the electricity demand pattern that appears by day period. In Figure. 3 Day (a) and (b), it shows the demand pattern for electricity on weekdays, and the difference between the two figures varies depending on the characteristics of the building. In Figure. 3 Day (c), it shows the demand pattern for electricity on Saturdays. Day (c) shows the relatively lower electricity demand compared to Day (a) and (b). In Figure. 3 Day (d), it shows the demand pattern for electricity on Sundays and holidays. Electricity demand patterns on holidays and Sundays are less and more irregular than other electricity demand patterns. In other words, the clustering by day period is correlated with the day type, and more detailed clustering is possible. In Figure. 3 Month (a), (b), (c), and (d), it shows the monthly electricity demand pattern.

Table 1: Comparison of cluster validity indices of K-means according to the number of clusters.

Num. Cluster	Method	DBI	CHI
5	K-means	2.03	13.14
6	K-means	1.98	13.75
7	K-means	1.92	14.63
8	K-means	**1.83**	**14.74**

Table 2: List of Buildings by Cluster when there are 8 number of Clusters

Cluster id	Building List
1	Z1B1, Z2B1, Z3B1, Z4B1
2	Z1B2, Z2B2, Z3B2, Z4B2
3	Z1B3, Z2B3, Z3B3, Z4B3
4	Z1B4, Z2B4, Z3B4, Z4B4
5	Z1B5, Z1B6, Z1B7, Z1B8, Z1B9
6	Z2B5, Z2B6, Z2B7, Z2B8, Z2B9
7	Z3B5, Z3B6, Z3B7, Z3B8, Z3B9
8	Z4B5, Z4B6, Z4B7, Z4B8, Z4B9

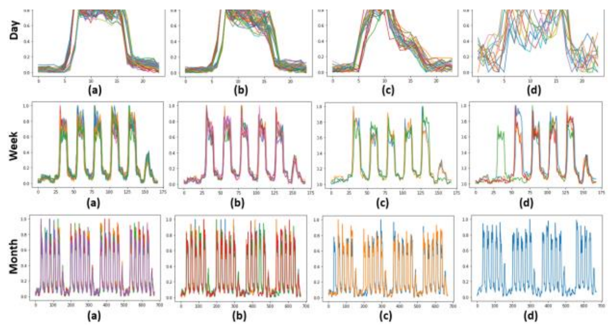

Figure 3: Seasonal Segmentation Clustering Result of Electricity Demand Profiles by Period: Day, Week, Month

And we analyze the electricity demand pattern that appears by week period and month period. In Figure. 3 Week (a), (b), and (c), it shows a common weekly electricity demand pattern. Whereas Figure. 3 Week (d) shows the electricity demand pattern including relatively a lot of holidays. However, it is difficult to visually identify the monthly electricity demand pattern whether clustering is done properly. That is, as the period increases, it may be seen that the result of clustering is suspicious.

4.2 Performance Comparison of Different Federated Learning

In this chapter, we compare the performance of different federated learning methods. The performance of federated learning can be affected by the characteristics and profiles of buildings

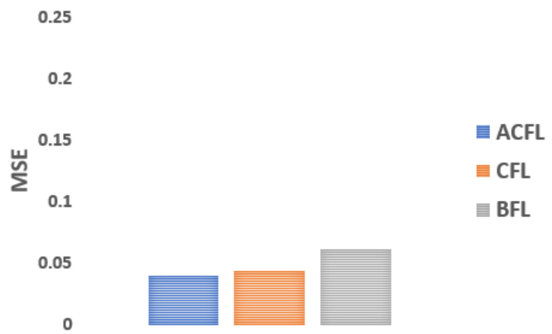

Figure 4: Performance Comparison of Different Federated Learning Methods with Clusters in Identical Climate Zones

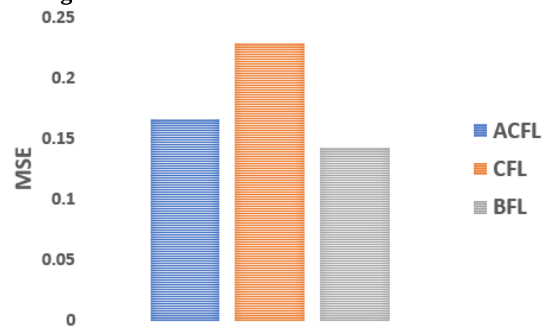

Figure 5: Performance Comparison of Different Federated Learning Methods with Clusters in Different Climate Zone

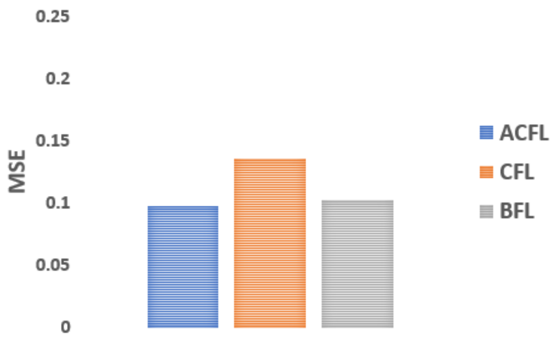

Figure 6: Performance Comparison of Different Federated Learning Methods with All Buildings

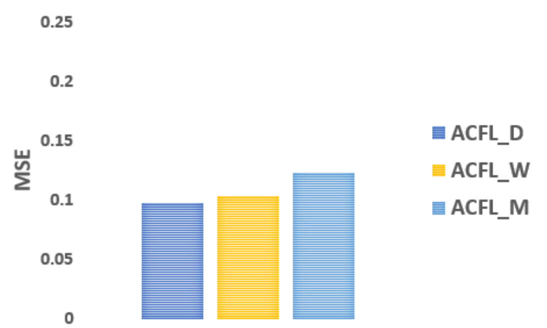

Figure 7: Performance Comparison of Adaptive Clustered Federated Learning Methods with Different Seasonal Clustering by Period

participating in federated learning because they learn different climate and electricity demand patterns located in each building. The first 10 months of each dataset are used to learn the LSTM model and the next 2 months are used for testing. For LSTM model learning, we use Adam Optimizer and set a batch size and the number of rounds to 70 and 100. And we compare the performance of the proposed adaptive clustered federated learning (ACFL), the clustered federated learning (CFL), and the basic federated learning (BFL) using the mean square error (MSE).

Figure. 4 shows the performance comparison of the different federated learning method with clusters in identical climate zones. Because buildings located in the different climate zone as well as in identical climate zones participate in BFL, the MSE of BFL is relatively higher than other methods. On the other hand, because only buildings located in the identical climate zone participate in learning, the MSE of CFL is relatively small. The proposed ACFL method shows the lowest MSE compared to other methods because it performs more detailed clustering compared to CFL. However, the performance of federated learning can be lower when buildings located in other climate zones participate in the federated learning.

Figure. 5 shows the performance comparison of the different federated learning method with clusters in different climate zones. Because buildings located in all climate zones can participate in the BFL and learn various characteristics, the MSE of BFL is relatively lower than other methods. On the other hand, because only buildings participate in learning in different climate zones, the MSE of CFL is increased. However, the proposed ACFL shows a relatively low MSE compared to CFL because it involves similar patterns of buildings by period in federated learning. In order to grasp the final result at a glance, it is necessary to combine it with the results of Figures 4 and 5.

Figure. 6 shows the performance comparison of the different federated learning method participated in 36 buildings. The proposed ACFL shows approximately 0.05 lower MSE compared to CFL and similar MSE to BFL. Because the proposed ACFL has different performance depending on the period, it is necessary to compare the performance of the ACFL according to the period.

Figure. 7 shows the performance comparison of the adaptive clustered federated learning method with different period. As shown in Figure. 3, because the clustering result of method using day period is better than methods using other periods, ACFL_D has lower MSEs compared to ACFL_W and ACFL_M.

5 CONCLUSIONS

In this paper, we propose the adaptive clustered federated learning in order to improve federated learning performance. with seasonal segmentation. To adaptively learn federated learning models by period, the proposed method uses pre-clustering results using K-means with seasonal segmentation. The proposed method with buildings in Identical climate zone shows the highest performance and shows similar performance to basic federated learning with buildings in different climate zones. However, because of frequent federated learning, the proposed method is not efficient in terms of network performance. In the future, in order to improve the network efficiency of federated learning, the federated clustering technique will be applied to federated learning.

ACKNOWLEDGMENTS

This research was supported by the Bio & Medical Technology Development Program of the National Research Foundation (NRF)& funded by the Korean government (MSIT) (NRF-2019M3E5D1A02067961). This research was supported by the MSIT(Ministry of Science and ICT), Korea, under the Innovative Human Resource Development for Local Intellectualization support program(IITP-2022-RS-2022-00156287) supervised by the IITP(Institute for Information & communications Technology Planning & Evaluation). This work was supported by Institute for Information & communications Technology Planning & Evaluation(IITP) grant funded by the Korea government(MSIT) (No.2022-0-01203, Regional strategic Industry convergence security core talent training business).

REFERENCES

[1] Aslam, Sheraz, et al. "A survey on deep learning methods for power load and renewable energy forecasting in smart microgrids." Renewable and Sustainable Energy Reviews 144 (2021): 110992.

[2] Ibrahim, Ibrahim Anwar, and M. J. Hossain. "LSTM neural network model for ultra-short-term distribution zone substation peak demand prediction." 2020 IEEE Power & Energy Society General Meeting (PESGM). IEEE, 2020.

[3] Ramokone, Agnes, Olawale Popoola, and Ayodele Temitope. "Load and energy profile development using occupancy and interactions/activities in the residential building." 2020 International SAUPEC/RobMech/PRASA Conference. IEEE, 2020.

[4] Yuan, Shuang, et al. "A Framework for the Automatic Integration and Diagnosis of Building Energy Consumption Data." Sensors 21.4 (2021): 1395.

[5] Berriel, Rodrigo F., et al. "Monthly energy consumption forecast: A deep learning approach." 2017 International Joint Conference on Neural Networks (IJCNN). IEEE, 2017.

[6] Yang, Qiang, et al. "Federated machine learning: Concept and applications." ACM Transactions on Intelligent Systems and Technology (TIST) 10.2 (2019): 1-19.

[7] Luo, Bing, et al. "Cost-effective federated learning design." IEEE INFOCOM 2021-IEEE Conference on Computer Communications. IEEE, 2021.

[8] Tun, Ye Lin, et al. "Federated learning based energy demand prediction with clustered aggregation." 2021 IEEE International Conference on Big Data and Smart Computing (BigComp). IEEE, 2021.

[9] Savi, Marco, and Fabrizio Olivadese. "Short-term energy consumption forecasting at the edge: A federated learning approach." IEEE Access 9 (2021): 95949-95969.

[10] De Cian, Enrica, Elisa Lanzi, and Roberto Roson. "Seasonal temperature variations and energy demand." Climatic Change 116.3 (2013): 805-825.

[11] Wang, Jian Qi, Yu Du, and Jing Wang. "LSTM based long-term energy consumption prediction with periodicity." Energy 197 (2020): 117197.

[12] Konečný, Jakub, et al. "Federated learning: Strategies for improving communication efficiency." arXiv preprint arXiv:1610.05492 (2016).

[13] Zhu, Hangyu, et al. "Federated learning on non-IID data: A survey." Neurocomputing 465 (2021): 371-390.

[14] Fikru, Mahelet G., and Luis Gautier. "The impact of weather variation on energy consumption in residential houses." Applied Energy 144 (2015): 19-30.

[15] Quilumba, Franklin L., et al. "Using smart meter data to improve the accuracy of intraday load forecasting considering customer behavior similarities." IEEE Transactions on Smart Grid 6.2 (2014): 911-918.

[16] Lusis, Peter, et al. "Short-term residential load forecasting: Impact of calendar effects and forecast granularity." Applied energy 205 (2017): 654-669.

[17] Kim, Inhee, and Jaehee Kim. "Multivariate time series clustering of electricity consumption data." The Korean Data & Information Science Society 32.3 (2021): 569-584.

[18] Vázquez-Canteli, José R., et al. "CityLearn v1. 0: An OpenAI gym environment for demand response with deep reinforcement learning." Proceedings of the 6th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation. 2019.

[19] Robinson, Darren, et al. "CitySim: Comprehensive micro-simulation of resource flows for sustainable urban planning." Proceedings of the Eleventh International IBPSA Conference. No. CONF. 2009.