Federated Learning for Non-intrusive Load Monitoring

Zhaorui Meng*, Xiaozhu Xie, Yanqi Xie

Abstract—**In the realm of non-intrusive load monitoring (NILM), extant deep learning approaches suffer from limitations including inadequate data samples, inadequate model generalization capacity, and insufficient safeguards for data privacy. To overcome these issues, this paper puts forward a novel NILM approach that leverages DeepAR to build a load monitoring model and incorporates federated learning and local fine-tuning methods to develop a non-intrusive load monitoring framework. Utilizing decentralized training, the proposed methodology facilitates iterative updates to model parameters through server-side aggregation, thereby enabling the collaborative construction of a monitoring model whilst maintaining strict confidentiality of individual customer data. The results of experiments conducted on the REDD dataset demonstrate that the approach outlined in this paper can markedly enhance the accuracy of load identification for frequently utilized electrical appliances.**

Index Terms—Non-intrusive load monitoring, DeepAR, federated learning, local fine-tuning.

I. INTRODUCTION

Electric energy consumption constitutes a significant portion of energy usage in both societal production and daily life. Therefore, the implementation of power energy conservation measures represents a crucial component of efforts aimed at reducing energy consumption and associated emissions. As advancements in science and technology continue to be made, and living standards improve, the quantity and diversity of household appliances have grown. Consequently, the proportion of residential electricity consumption within total electricity consumption has also increased. [1]. According to research, the implementation of electricity consumption feedback mechanisms will likely enhance the potential for energy conservation on the residential load side. The use of household load identification technology can substantially aid power supply companies in comprehending the load structure and consumption patterns of residential users, thereby enabling timely adjustments to the power supply plan. Additionally, this technology can assist users in understanding household electricity information, facilitating rational electricity consumption practices on the user side, and promoting overall energy conservation and emission reduction efforts. [2].

The current load monitoring means is to install an independent monitoring device on each load, and to realize the judgment of its running state through the independent monitoring of each load, which consumes a lot of manpower, material resources and financial resources. To solve the above problems, non-intrusive load monitoring method is proposed. NILM does not need to enter the inside of the end user's power system, but only needs to install monitoring equipment at the power entrance. By monitoring the total voltage and total current at the entrance and decomposing it, the running state of each electrical equipment of the end user can be obtained [3].

At the moment, the NILM research mainly focuses on the low frequency characteristics of electrical appliances with active power, because most of monitoring devices can only achieve low frequency sampling of power. In non-intrusive power decomposition problems, the most commonly used method is hidden Markov model. Makonin et al. [4] proposed a load disaggregation algorithm integrated a super-state hidden Markov model and a Viterbi algorithm variant. The algorithm performed extremely well for multi-state load. Bonfigli et al. [5] proposed a NILM algorithm using Additive Factorial Hidden Markov Models framework. Experimental results show that the monitoring effect of this method is better than that of other four comparison methods. Wu et al. [6] presented a Time-Efficient Factorial Hidden Semi-Markov Model to improve the computing efficiency of NILM in real-world scenario. For the past few years, deep learning is gradually applied to NILM field. Kaselimi et al. [7] introduced a deep Long Short-Term Memory (LSTM) neural networks for energy disaggregation. Nolasco et al. [8] proposed a CNN-based framework for multi-label classification in NILM signals. Kong et al. [9] presented a deep convolutional neural networks-based approach to estimate the energy consumption for common multi-functional home appliances.

Although deep learning has been widely applied in NILM field, the following problems still exist in practical application. First of all, deep learning has high requirements on the amount of training data, which may not be met by the electricity consumption data of a single household in practical

Manuscript received February 6, 2023, revised June 26, 2023;

Zhaorui Meng is a senior engineer in School of Computer and Information Engineering, Xiamen University of Technology, Xiamen, Fujian,361024, China (Corresponding author; Email: mengzhaorui@ xmut.edu.cn).

Xiaozhu Xie is a lecturer in School of Computer and Information Engineering, Xiamen University of Technology, Xiamen, Fujian,361024, China (Email: xiaozhuxie@ xmut.edu.cn).

Yanqi Xie is a lecturer in School of Computer and Information Engineering, Xiamen University of Technology, Xiamen, Fujian,361024, China (Email: yqxie@ xmut.edu.cn).

.

application. Secondly, the generalization ability of the model only aiming at single family training is poor, and it is difficult to obtain good monitoring effect on the data of other families. Finally, there are privacy concerns such as data breaches if data from other households are obtained in order to train better generalized models. In response to these problems, federated learning [10] comes into being, which can utilize multiple computing nodes for efficient machine learning without compromising user data in a legal and compliant manner. Federated learning can reduce data traffic while protecting data privacy. Currently, federated learning has been used by researchers in some scenarios that need to consider data privacy and reduce communication overhead. Feng et al. [11] proposed a human mobility prediction framework via federated learning. Wang et al. [12] designed a federated learning framework in order to provide a better learning parameter exchange method for mobile edge computing. Hu et al. [13] presented a federated learning method for urban environment sensing. Pfohl et al. [14] proposed a federated learning framework for electronic health records. Yan et al. [15] proposed a federal learning application scheme in the field of financial credit risk management. All the above federated learning applications have achieved good results, which proves that the federated learning method is practical and feasible.

In view of the above problems of insufficient data required for the construction of load monitoring model, poor model generalization ability and privacy involved in data sharing, this paper proposes a non-intrusive load monitoring method based on DeepAR model and federated learning, which implements collaborative training of the model under the premise of protecting the privacy of each customer's data. Moreover, the monitoring accuracy and generalization ability of non-intrusive load monitoring model are improved effectively.

II. FEDERATED LEARNING-BASED NON-INTRUSIVE LOAD MONITORING MODEL

2.1 Method overview

The non-intrusive load monitoring model based on federated learning can not only protect the privacy of each customer's data, but also use the load data resources owned by each customer to train the load monitoring model cooperatively, effectively improving the prediction accuracy and multi-scenario generalization ability of the model.

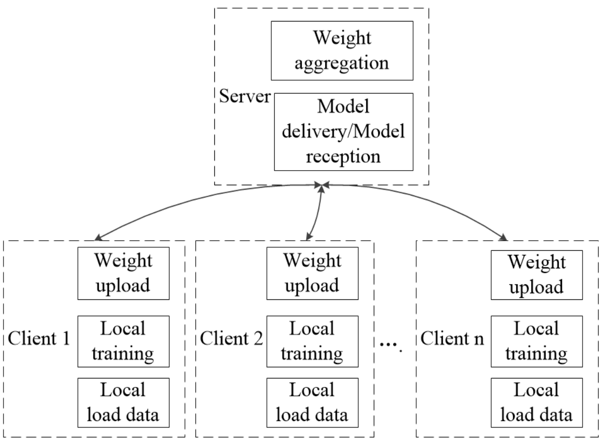

Fig. 1. The overall system architecture of the method in this paper

The overall system architecture of this method is shown in Fig. 1. In terms of structure, the whole method consists of server, communication network, local load monitoring client and corresponding load monitoring data. Among them, the load monitoring model built based on DeepAR model is deployed on the server and each load monitoring client.

The overall process mainly includes five steps: load monitoring model delivery, local training of each client, upload of each client model, model aggregation and local fine-tuning of each client. The specific steps are shown in Algorithm 1.

Algorithm 1. Non-intrusive load monitoring based on federated learning and local fine-tuning

Input: Number of iterations T, various clients involved in load monitoring L

Output: A non-intrusive load monitoring model trained by the method presented in this paper

Step 1. Preprocess each customer's electricity data, including missing value, outlier value and data normalization

Step 2. Set the communication address between the server and each local server

Step 3. Enable communication service

Step 4. The server initializes a global DeepAR model and obtains the initial model weight parameters W

Step 5. Set up all local clients l_k $(k = 1,2,\dots,n), l_k \in L$ to participate in training

Step 6. for t =1, 2, …, T, do:

Step 7. for $l_k \in L$ do:

Step 8. l_k downloads W from server

Step 9. l_k train DeepAR model W_k with local electricity data

Step 10. l_k uploads W_k to server

Step 11. end for

Step 12. The central server aggregates local models from different customers to update the global shared model

Step 13. Repeat the above steps until the model accuracy reaches the required standard or iterates to the specified number of times T

Step 14. Based on the global model, local fine-tuning is performed to generate the final load monitoring model of each client

2.2. DeepAR

DeepAR is a time series prediction method based on deep learning proposed by Salinas et al. [16], whose goal is to simulate conditional probability distribution $P(Z_{i,t_0:T}|Z_{i,1:t_0-1},x_{i,1:T})$.The future series $Z_{i,t_0:T}$ is modeled according to the past time series$Z_{i,1:t_0-1}$ and covariable $x_{i,1:T}$, where t_0 is the time division point, $Z_{i,t}$ represents the value of time series i at time t.

DeepAR is an autoregressive RNN time series model, which is a cyclic neural network (using LSTM or GRU units) with hidden states. DeepAR learns periodic representations and is based on covariates across time series. When obtaining highly complex, group-dependent representations, only a small amount of data processing needs to be carried out manually.

The conditional probability distribution used by the DeepAR model can be written in the following likelihood form:

$$Q_\Theta(Z_{i,t_0:T}|Z_{i,1:t_0-1},x_{i,1:T}) = \prod_{t=t_0}^{T} Q_\Theta(Z_{i,t}|Z_{i,1:t-1},x_{i,1:T}) = \prod_{t=t_0}^{T} l(Z_{i,t}|\theta(h_{i,t},\Theta)) \quad (1)$$

where $h_{i,t} = h(h_{i,t-1}, z_{i,t-1}, x_{i,t}, \Theta)$ is the output of an autoregressive recurrent network composed of multi-layer RNNs.

As shown in Fig. 2, DeepAR input the hidden layer $h_{i,t-1}$ and $z_{i,t-1}$ at the previous time and the known information $x_{i,t}$ at the current time, and the hidden layer $h_{i,t}$ at this time can be obtained. Then, $h_{i,t}$ is converted into the parameter of the given distribution through neural network $\theta(\cdot)$. After the distribution is determined, likelihood $l(Z_{i,t}|\theta(h_{i,t},\Theta))$ can be calculated and the predicted probability distribution can be finally obtained.

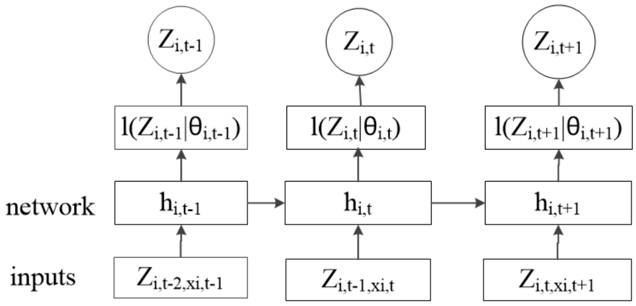

Fig. 2. Diagram of DeepAR network

2.3 Federated aggregation

In federated learning, the central server receives the results from multiple clients, aggregates them, and then sends the aggregated results to each client, the process is called federated aggregation. Because FedAvg federated aggregation algorithm has higher communication efficiency, the method in this paper adopts FedAvg federated aggregation algorithm, and the specific steps are shown in Algorithm 2.

Algorithm 2. Federated Averaging Algorithm

Input: T is the maximum number of iterations, η is the learning rate, k is the client number, n_k is the amount of data of the k client and $N = \sum_k n_k$

Step 1. Initialize an DeepAR model W_0
Step 2. While r < T do
Step 3. Select subset
Step 4. for client k in K do
Step 5. k receives model w_r
Step 6. k computes average gradient g_k with SGD
Step 7. k updates local model $w_{r+1}^k \leftarrow w_r^k - \eta g_k$
Step 8. k sends updated model to server
Step 9. End for
Step 10. Server computes new global model $w_{r+1}^k \leftarrow \sum_{k=0}^{K} \frac{n_k}{N} w_{r+1}^k$
Step 11. End while

2.4. Local fine-tuning

The global model generated by federated learning has better generalization ability, and local fine-tuning based on the global model can improve the prediction ability of the model on specific clients. In addition, local fine-tuning only needs fewer iterations to get better prediction results.

III. EXPERIMENTS

3.1. Experimental environment

In this paper, multiple computers are used to simulate each customer participating in federated learning, and then verify the improvement of prediction accuracy and generalization ability of the proposed method. Six computers were used to simulate six customers. All computers were equipped with Intel Xeon platinum 8124 CPU, RTX 3080 GPU, and 128 GB memory. A computer with the same configuration serves as the server side. The devices can communicate point to point.

In this paper, REDD dataset [17] is selected for the experiment. This dataset contains the electricity consumption data of 6 American households, and the low-frequency power data of this dataset is used for load monitoring. The model uses the first 70% of each household's electricity consumption data for training, the next 20% as the verification set, and the last 10% as the test dataset.

In this paper, four typical household electrical equipment including dish washer, refrigerator, washing machine and lights are selected for testing.

The activation thresholds of each electrical appliance are shown in Table 1.

Table 1. The activation threshold of each appliance.

Appliance	Activation threshold(W)
Dish washer	10
Refrigerator	50
Washing machine	20
Lights	10

3.2. Evaluation metrics

In order to comprehensively evaluate model performance, precision (PRE) and mean average absolute error (MAE) are selected as evaluation metrics in this paper. Specific formulas are as follows:

$$PRE = \frac{TP}{TP+FP} \quad (2)$$

where TP is the number of sequences in which both the model prediction result and the actual load are running states, and FP is the number of sequences in which the model prediction result is not in the running state but the actual load is in the running state.

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \quad (3)$$

where N is total sample input number. y_i is the real power of electric appliance at time t, and $\hat{y}_i$ is the model decomposition power.

3.3. Experimental results

Table 2-Table 7 shows the comparison of PRE and RMSE of experimental results of each model from household 1 to household 6. Based on the presented tables, it is evident that the proposed algorithm in this paper yields the minimum RMSE and the maximum PRE in various electrical appliance experiments when the frequency of use is relatively high, denoted by RMSE values greater than 1. This indicates that combining federated learning and local fine-tuning leads to the most effective load monitoring outcome. When the RMSE value is less than 1, comparing the values of PRE and RMSE

Table 2. Comparison of Precision and RMSE of different appliances in each model in household 1. Bold indicates the smallest RMSE or the largest PRE.

Model	Dish washer		Refrigerator		Washing machine		Lights	
	PRE	RMSE	PRE	RMSE	PRE	RMSE	PRE	RMSE
DeepAR	0.9774	15.13	0.9957	26.17	0.4119	83.89	0.9998	**0.9706**
DeepAR(Centralized)	0.9860	15.05	0.9960	26.02	0.4573	81.56	0.9998	1.1034
Federated learning	0.9866	16.1	0.9960	28.05	0.428	94.62	0.9998	1.2018
Federated learning+ local fine-tuning	**0.9867**	**14.95**	**0.9971**	**25.97**	**0.7165**	**80.57**	**0.9998**	1.56

Table 3. Comparison of Precision and RMSE of different appliances in each model in household 2. Bold indicates the smallest RMSE or the largest PRE.

Model	Dish washer		Refrigerator		Washing machine		Lights	
	PRE	RMSE	PRE	RMSE	PRE	RMSE	PRE	RMSE
DeepAR	0.9785	23.74	0.9976	22.86	0	0.3550	0.9922	1.3819
DeepAR(Centralized)	0.9815	16.03	0.9973	22.25	0	**0.3543**	0.9928	1.3723
Federated learning	0.9865	16.75	0.9972	22.36	0	0.4452	0.9927	1.4204
Federated learning+ local fine-tuning	**0.9875**	**12.85**	**0.9977**	**21.41**	0	0.3830	**0.9991**	**1.3673**

Table 4. Comparison of Precision and RMSE of different appliances in each model in household 3. Bold indicates the smallest RMSE or the largest PRE. N/A indicates not available.

Model	Dish washer		Refrigerator		Washing machine		Lights	
	PRE	RMSE	PRE	RMSE	PRE	RMSE	PRE	RMSE
DeepAR	N/A	**0.3861**	0.9939	17.52	0.9807	263.05	N/A	0.2765
DeepAR(Centralized)	N/A	0.4215	0.9940	17.45	0.9813	267.03	N/A	0.3213
Federated learning	N/A	0.7705	0.9940	17.52	0.9656	391.8	N/A	0.4714
Federated learning+ local fine-tuning	N/A	0.3904	**0.9942**	**17.35**	**0.9839**	**261.52**	N/A	**0.2647**

Table 5. Comparison of Precision and RMSE of different appliances in each model in household 4. Bold indicates the smallest RMSE or the largest PRE. N/A indicates not available.

Model	Dish washer		Refrigerator		Washing machine		Lights	
	PRE	RMSE	PRE	RMSE	PRE	RMSE	PRE	RMSE
DeepAR	0	0.4831	N/A	N/A	0	0.5860	0.9989	2.7696
DeepAR(Centralized)	0	0.4634	N/A	N/A	0	0.5532	0.9989	2.6423
Federated learning	0	0.5108	N/A	N/A	0	**0.4496**	0.9989	2.6603
Federated learning+ local fine-tuning	0	**0.4245**	N/A	N/A	0	0.8033	**0.9989**	**2.6020**

Table 6. Comparison of Precision and RMSE of different appliances in each model in household 5. Bold indicates the smallest RMSE or the largest PRE. N/A indicates not available.

Model	Dish washer		Refrigerator		Washing machine		Lights	
	PRE	RMSE	PRE	RMSE	PRE	RMSE	PRE	RMSE
DeepAR	N/A	**0.5116**	0.9955	8.199	N/A	**0.0487**	N/A	**0.0010**
DeepAR(Centralized)	N/A	0.5234	0.9958	8.032	N/A	0.0975	N/A	0.0382
Federated learning	N/A	0.6366	0.9954	8.058	N/A	0.1353	N/A	0.0432
Federated learning+ local fine-tuning	N/A	0.5197	**0.9965**	**7.84**	N/A	0.1152	N/A	0.0026

becomes challenging. This is due to the relatively low frequency of appliance use, as well as the low decomposition power of each appliance. Consequently, obtaining 0 values for TP and TR becomes more likely. The presented tables further reveal that the accuracy of the federated learning algorithm may not surpass that of DeepAR and DeepAR (Centralized) if local fine-tuning is not employed. The latter approach involves training the model on aggregated data

Table 7. Comparison of Precision and RMSE of different appliances in each model in household 6. Bold indicates the smallest RMSE or the largest PRE. N/A indicates not available.

Model	Dish washer		Refrigerator		Washing machine		Lights	
	PRE	RMSE	PRE	RMSE	PRE	RMSE	PRE	RMSE
DeepAR	N/A	**0.3017**	0.9964	6.1792	N/A	0.4950	1	3.7481
DeepAR(Centralized)	N/A	0.3532	0.9975	6.8823	N/A	0.4632	1	3.7012
Federated learning	N/A	0.3765	0.9982	7.89	N/A	**0.4437**	1	4.274
Federated learning+ local fine-tuning	N/A	0.3184	**0.9983**	**6.129**	N/A	0.5231	**1**	**3.669**

from all households. Interestingly, in the majority of experiments conducted, the differences in the RMSE and PRE values among these three algorithms are minimal.

IV. CONCLUSION

This paper introduces a novel non-intrusive load monitoring approach, which leverages federated learning and local fine-tuning techniques. This method enables data isolation and facilitates the collaborative construction of a universal monitoring model across multiple data nodes whilst guaranteeing user privacy protection. By optimizing the global model through local fine-tuning, the accuracy of load monitoring is significantly enhanced. Additionally, this study establishes a basis for future research to optimize the federated learning algorithm, with the objective of further enhancing the accuracy of NILM while maintaining robust data security measures.

REFERENCES

[1] Rajendiran, G., et al. "Energy management using non-intrusive load monitoring techniques - State-of-the-art and future research directions," Sustainable Cities and Society, vol. 62, no.1, 2020.
[2] Wahba, Asmaa, Reda El-khoribi, and Shereen Taie. "A New Hybrid Model for Energy Consumption Prediction Based on Grey Wolf Optimization," IAENG International Journal of Computer Science, vol. 49, no.2, pp469-481,2022.
[3] Hosseini, S. S, et al. "Non-intrusive load monitoring through home energy management systems: A comprehensive review," Renewable and Sustainable Energy Reviews, vol. 79, pp1266-1274,2017.
[4] Makonin, Stephen, et al. "Exploiting HMM sparsity to perform online real-time nonintrusive load monitoring," IEEE Transactions on smart grid, vol. 7, no.6, pp2575-2585,2015
[5] Bonfigli, Roberto, et al. "Non-intrusive load monitoring by using active and reactive power in additive Factorial Hidden Markov Models," Applied Energy, vol. 208, pp1590-1607,2017.
[6] Wu, Z., et al. "A time-efficient factorial hidden Semi-Markov model for non-intrusive load monitoring," Electric Power Systems Research. vol. 199, no.1, pp1590-1607,2021.
[7] Kaselimi, M., et al. "Context Aware Energy Disaggregation Using Adaptive Bidirectional LSTM Models," IEEE Transactions on Smart Grid, vol. 11, no.4, pp3054-3067,2020.
[8] Nolasco, Lds, A. E. Lazzaretti, and B. M. Mulinari "DeepDFML-NILM: A New CNN-Based Architecture for Detection, Feature Extraction and Multi-Label Classification in NILM Signals," IEEE sensors journal, vol. 22, no.1, pp501-509,2022.
[9] Kong, W., et al. "A Practical Solution for Non-Intrusive Type II Load Monitoring Based on Deep Learning and Post-Processing," IEEE Transactions on Smart Grid, vol. 11, no.1, pp148-160,2020.
[10] Zhu, Hangyu, Haoyu Zhang, and Yaochu Jin. "From federated learning to federated neural architecture search: a survey," Complex & Intelligent Systems, vol. 7, no.1, pp639-657,2021.
[11] Feng, J. et al. "PMF: A Privacy-preserving Human Mobility Prediction Framework via Federated Learning," Proc. ACM Interact. Mob. Wearable Ubiquitous Technol, vol. 4, no.1, pp1-21,2020.
[12] Wang, X, et al. "In-Edge AI: Intelligentizing Mobile Edge Computing, Caching and Communication by Federated Learning," IEEE Network, vol. 33, no.5, pp156-165,2019.
[13] Hu, B., et al. "Federated Region-Learning: An Edge Computing Based Framework for Urban Environment Sensing," 2018 IEEE Global Communications Conference (GLOBECOM) IEEE, 2019.
[14] Pfohl, S. R, A. M. Dai, and K. Heller. "Federated and Differentially Private Learning for Electronic Health Records," ArXiv abs/1911.05861, 2019.
[15] Li, Yan, and Guihua Wen. "Research and Practice of Financial Credit Risk Management Based on Federated Learning," Engineering Letters, vol. 31, no.1, pp271-278,2023.
[16] Salinas, David, et al. "DeepAR: Probabilistic forecasting with autoregressive recurrent networks," International Journal of Forecasting, vol. 36, no.3, pp1181-1191,2020.
[17] Kolter, J. Zico, and Matthew J. Johnson. "REDD: A public data set for energy disaggregation research," Workshop on data mining applications in sustainability (SIGKDD), San Diego, CA. vol. 25, no. Citeseer. 2011.