Source-Free Dynamic Weighted Federated Transfer Learning for State-of-Health Estimation of Lithium-Ion Batteries With Data Privacy

Tengfei Han, Shang Yue, Pu Yang, Ruixu Zhou, and Jianbo Yu, *Member, IEEE*

***Abstract*—Most existing methods for battery state-of-health (SOH) estimation rely on centralized training mode. However, in practical applications, it is difficult for a single user to collect sufficient battery degradation data to train a model. In addition, in order to protect data privacy, users are unwilling to share data, where centralized training mode is not the best choice. To overcome these barriers, a source-free dynamic weighted federated transfer (SF-DWFT) method for battery SOH estimation is proposed, which utilizes a distributed learning paradigm to combine multiple source clients to train a global model while protecting data privacy. First, a Gaussian mixture model is used to model the high-level features extracted by the client model, thus the feature representation can be shared without needing access to the source data; then, in order to reduce the effect of distribution differences between the source and target clients on the global model, a dynamic weighted federated aggregation algorithm is proposed according to their contributions, which is measured by calculating the modified Bhattacharya distance between the source and target clients and testing error. Finally, the effectiveness of SF-DWFT is verified on several battery datasets and it achieves a statistic estimation error of 1.12% on 18 650 lithium-ion battery datasets and 4.46% on NASA battery dataset.**

***Index Terms*—Data privacy, federated learning (FL), Gaussian mixture model (GMM), lithium-ion batteries, state-of-health (SOH).**

I. Introduction

LITHIUM-ION batteries have become an important component of energy storage systems in portable electrical equipment and electric vehicles due to their advantages such as high energy density, long lifetime, and low self-discharge rate [1], [2]. However, batteries will inevitably deteriorate over time, leading to issues such as decreased capacity and power. In fact, when batteries deteriorate, the probability of failure also increases, and it can even cause serious accidents, e.g., fires [3]. Therefore, accurately monitoring and evaluating the state-of-health (SOH) of lithium-ion batteries is of great significance for battery health management.

Currently, significant progress has been made in the research of SOH estimation of lithium-ion batteries. These methods mainly consist of physical model-based approaches and data-driven methods. The former utilizes the internal mechanism of the battery to construct mathematical models that can characterize their degradation process, mainly including equivalent circuit models [4], [5] and electrochemical models [6], [7]. Chen et al. [8] modeled each battery as an equivalent circuit model and assessed the impact of cell parameters on electric vehicle range. However, the modeling process of these methods is sometimes too complex and difficult to apply to real-time SOH evaluation. In contrast, the data-driven approaches are more flexible. The intelligent models such as multilayer perceptron [9] and relevance vector machine [10] can be employed to construct the mapping relationship between the historical monitoring data and the health status of the battery. Recently, deep learning has been widely applied in PHM due to its powerful ability of feature extraction and data fitting. Catelani et al. [11] integrated recurrent neural networks and state space estimation for estimating battery SOH. Ren et al. [12] introduced a SOH and remaining useful life (RUL) prediction model, i.e., Auto-CNN-LSTM, which integrates an improved convolutional neural network (CNN) and long short-term memory neural network (LSTM) for extracting deep-level degradation information from limited data. However, these SOH estimation approaches do not consider the domain shift between training and testing data. In fact, different batteries mostly operate under different conditions, which leads to distribution discrepancy of battery data and concept drift [13]. Transfer learning (TL) is an effective method for this problem. It can transfer the learned knowledge from the source domain to the target domain. Basic TL methods include DaNN [14] and DANN [15]. Some scholars have made improvements on the basis of these methods. Du et al. [16] presented a method combining TL and cycle life prediction techniques for battery aging trajectory prediction. Ye and Yu [17] proposed a novel TL model called DDAN, which constructed a feature alignment method based on CORAL and maximum mean discrepancy (MMD) and extracted domain-invariant features with adversarial learning.

Manuscript received 14 December 2023; revised 11 June 2024; accepted 17 July 2024. Date of publication 29 July 2024; date of current version 11 September 2024. This work was supported in part by the National Key Research and Development Program under Grant 2022YFF0605700, in part by the National Natural Science Foundation of China under Grant 92167107, and in part by the Tencent Fundamental Platform Technology Rhino-Bird Focused Research Program. Recommended for publication by Associate Editor K. Gunawardane. *(Corresponding author: Jianbo Yu.)*

Tengfei Han and Jianbo Yu are with the School of Mechanical Engineering, Tongji University, Shanghai 201804, China (e-mail: han_tf98@tongji.edu.cn; jbyu@tongji.edu.cn).

Shang Yue, Pu Yang, and Ruixu Zhou are with the Tencent Incorporation, Shenzhen 518057, China (e-mail: friday@tencent.com; dowpuyang@tencent.com; ruixuzhou@tencent.com).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TPEL.2024.3434751.

Digital Object Identifier 10.1109/TPEL.2024.3434751

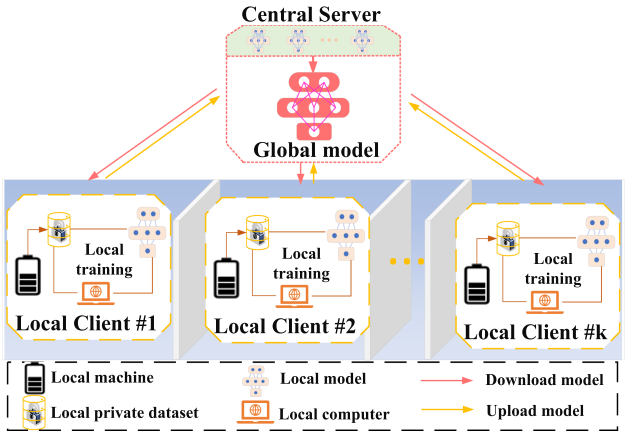

Fig. 1. Framework of FL.

For the problem of SOH estimation of batteries under variable operating conditions, TL methods can achieve desirable results.

Although the SOH estimation methods based on deep learning have achieved great performance, they usually rely on centralized training modes and a large amount of high-quality data. In practical industrial applications, the battery degradation data held by a single user is very limited, making it difficult to independently develop a high-performance SOH estimation model. Moreover, to protect their privacy and interests, users holding similar batteries are usually unwilling to share private data, which leads to data decentralization [18], [19]. Considering data privacy issues, it is impractical to adopt a centralized training mode to obtain an ideal SOH estimation model. Recently, Google proposed a decentralized approach, federated learning (FL) [20], providing an encouraging solution for the aforementioned issues. FL can collaborate with multiple clients to train a deep model while protecting data privacy. During this process, it does not require centralized training of the model. There are two roles in the FL framework, as shown in Fig. 1. One is the user who owns local data and participates in joint training, called the client; the other role is a trusted central server responsible for communicating with clients and aggregating client models. During the communication process between the client and the central server, the parameters of local models are aggregated into a global model without sharing raw data. Therefore, the local data of the client will not be leaked to ensure the data security of each client.

Due to its outstanding performance in privacy protection, FL has attracted widespread attention in these applications, e.g., healthcare and intelligent manufacturing. Liu et al. [21] introduced a federated contrastive learning method for medical image classification problems. Shiri et al. [22] proposed a federated transfer learning (FTL) method with differential privacy techniques to detect and correct artifacts in positron emission tomography images. Some FL methods have been introduced for fault diagnosis and RUL prediction. In terms of fault diagnosis, Zhao et al. [23] developed a multisource domain adaptive FTL method for mechanical fault diagnosis with data privacy preserving. Guo et al. [24] proposed an FL method named FedRUL to predict the RUL of mechanical equipment. This method only uploads low-level features and parameters of the encoder to the central server, which significantly reduces the computational burden on the client. Qing et al. [25] proposed a dynamic weighted FL method on the basis of MMD, which aggregates all client models according to the MMD distance between training and testing clients.

However, to the best of our knowledge, there is currently no scholar conducting research on battery SOH estimation under data privacy, and this issue is a very practical one. Moreover, the aforementioned federated transfer methods still have some limitations: 1) most of them transfer knowledge by sharing high-level features extracted by local models, which still leads to privacy concerns [26], as local data is potentially recovered from the high-level features. 2) the differences among client models have not been fully considered, and most FL methods use the federated average algorithm (FedAvg) to aggregate local models, making it difficult to effectively guide the global model optimize.

In order to address the above-mentioned issues, i.e., data privacy and distribution discrepancy between clients in battery SOH estimation, a source-free dynamic weighted federated transfer learning (SF-DWFT) method is proposed for lithium-ion batteries SOH estimation, where only source client model is provided to the central server, rather than the source data. The main contributions of this study are as follows.

1) A novel FTL method is proposed for the SOH estimation of lithium-ion batteries. It focuses on data privacy issues, which are frequently encountered in battery SOH estimation tasks, but are rarely studied.
2) A dynamic weighted aggregation algorithm based on a modified Bhattacharyya distance and testing error is proposed to achieve an optimal global model for the target client.
3) A Gaussian mixture model (GMM) is used to model the extracted features in the local client training phase. Unlike other federated transfer methods, only parameters of GMM are required to upload, instead of high-level features, which effectively solves the problem of data privacy in the learning procedure. The experimental results on several battery datasets confirm the effectiveness and superiority of SF-DWFT for battery SOH estimation.

The rest of this artcle is organized as follows. SF-DWFT for SOH estimation is presented in Section II. Section III provides the details of experiments and results analysis. Section IV concludes of this article.

II. Methodology

In the existing SOH estimation methods that adopt centralized training modes, the issue of data privacy is rarely considered, while it is a very practical one. To solve this problem, an SF-DWFT learning method, i.e., SF-DWFT, is proposed for SOH estimation under data privacy. The framework of the SF-DWFT is shown in Fig. 2. In order to protect the security of the local dataset, the client only uploads the trained model and feature distribution parameters to the central server during communication. A dynamic weighted federated aggregation method is developed

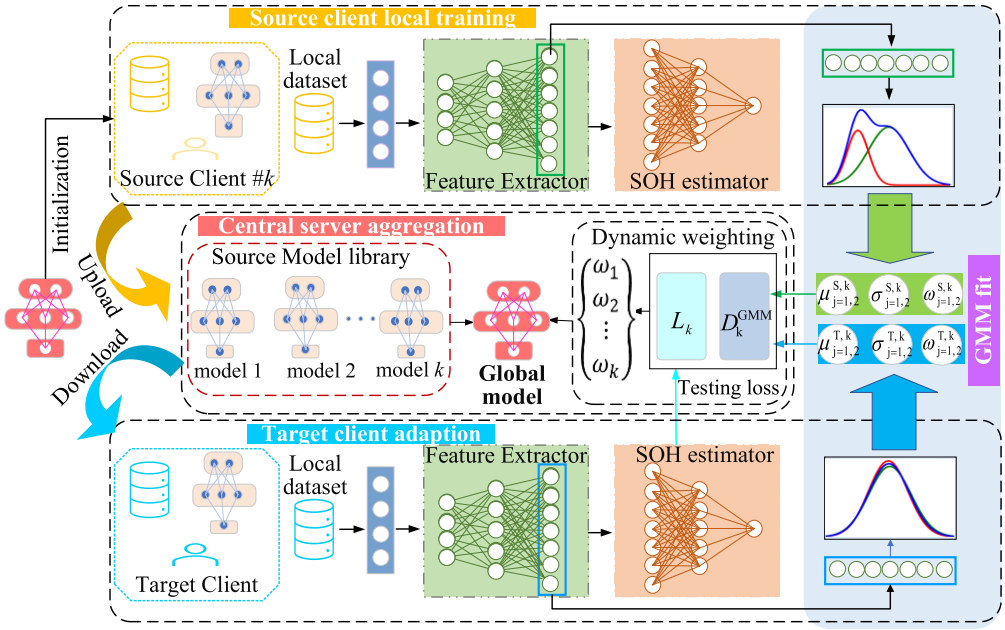

Fig. 2. Network structure of SF-DWFT.

to alleviate the impact of data distribution discrepancy on the global model. The weight of each local model in the aggregation is determined by the distribution discrepancy between the data of the source client and target client, as well as the test error of the local model.

SF-DWFT mainly consists of three parts when implementing battery SOH estimation: client local training, distribution discrepancy calculation, and central server aggregation. First, the local client model is trained separately by the local data. After the training is completed, a GMM is employed to model the extracted features. Then, all source clients upload the local models and the parameters of GMMs to the central server. Finally, the improved Bhattacharyya distance is proposed to measure the distribution difference between the source clients and target client, and the weights of the local models are determined based on the distribution discrepancy and testing error.

In addition, SF-DWFT is proposed based on the following assumptions.

1) Multiple clients collaboratively train the SOH estimation model. The training data in the source client is labeled, while the target client holds a small amount of labeled data.
2) In global communication, only model parameters, gradients, losses, and feature parameters can be passed between the local client and the central server. There is no direct contact between clients.
3) The local and global models share an identical network structure.

A. Client Local Training

1) Local Model Training: To facilitate model aggregation in the central server, the local client and central server adopt the same network structure in SF-DWFT, which is mainly composed of a feature generator and SOH estimator.

The degradation data of lithium-ion batteries are a type of time series data, as they are collected during the working process. Thus, it is important to accurately extract the temporal information from the sensor data. Bidirectional gate recurrent unit (BGRU) is a network composed of GRUs [27] stacked together, which can simultaneously extract past and future information. In order to enhance the feature learning ability of BGRU, multiple BGRUs are densely connected to construct DBGRU [17]. The ith BGRU unit takes the extracted features of all previous BGRUs, $[H_0, H_1, \ldots, H_{i-1}]$ as input

$$H_i = \text{BGRU}\left([H_0, H_1, \ldots, H_{i-1}]\right) \tag{1}$$

where H_i represents the output of DBGRU, $[H_0, H_1, \ldots, H_{i-1}]$ is the concatenation of the features extracted by BGRU_1, $\text{BGRU}_2, \ldots,$ BGRU_{i-1}.

The SOH estimator is composed of two fully connected (FC) layers. The relationship between the extracted high-level feature and SOH can be learned by the FC layers. The SOH estimator is defined by

$$y = f\left(\omega H + b\right) \tag{2}$$

where y is the estimated SOH value, ω is the weight matrix of FC layer, H represents the extracted features, $f(\cdot)$ stands for the activation function, and b is the bias. After the basic SOH prediction model is constructed, the local models are initialized based on the global model downloaded from the central server. Then, each client trains the local model using local privacy data.

In particular, let D_k^S denotes the labeled dataset of the kth source client C_k^S, $D_k^S = \{X_k^S,\ Y_k^S\} = \{(x_{k,i}^S, y_{k,i}^S)\}_{i=1}^{n_k^S}$, $k = 1, 2, \ldots, K$, n_k^S is the number of samples. The X_k^S and Y_k^S

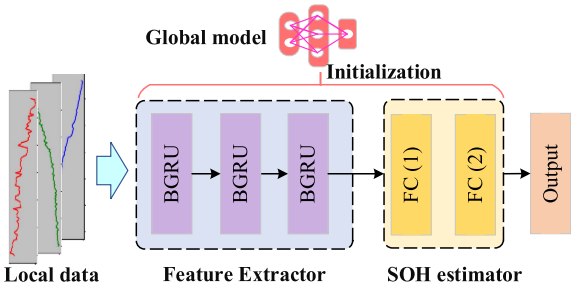

Fig. 3. Procedure of local model training.

represent the input dataset of the local model and the corresponding SOH values of batteries, respectively. The labeled data samples held by the target client C^T are denoted as D^T, $D^T = \{X^T, Y^T\} = \{(x_i^T, y_i^T)\}_{i=1}^{n^T}$, and $D^{T'}$ is the unlabeled datasets, $D^{T'} = \{X^{T'}\} = \{(x_i^{T'})\}_{i=1}^{n^{T'}}$. Here, the number of labeled data is very small. X^T and Y^T represent the input dataset and corresponding SOH of the target client, respectively. D^T participates in the central server model aggregation, while the unlabeled dataset $D^{T'}$ is used to test the estimation performance of SF-DWFT. The M_i^G and $M_{k,i}^S$, $i = 1, 2, \ldots, N_{\text{round}}$, are the global model and the local model of the kth client after the ith round communication, respectively. Fig. 3 shows the procedure of local model training. First, each source client initializes the local model with the downloaded global model M_i^G. Then, the local model of each client is trained individually. For all client models $M_{k,i}^S$, $k = 1, 2, \ldots, K$, the mean square error (MSE) is set as the loss function

$$L_{C_k^S} = \frac{1}{n_k^S} \sum_{i=1}^{n_k^S} \left(y_{k,i}^S - y_{k,i,\text{pre}}^S\right)^2 \tag{3}$$

where $y_{k,i}^S$ and $y_{k,i,\text{pre}}^S$ are the actual and estimated SOH of the ith input data from the kth source client, respectively, n_k^S is the number of samples held by the k th source client.

2) GMM Fit: In TL, the high-level features are often shared to measure the distribution discrepancy between the source and target domain. However, this may lead to local data leakage [28]. In [29], features can be encoded as a Gaussian, so the shared data only includes the mean and standard deviation of the features. Here, a GMM is adopted to describe the features extracted by the local model, as for complex tasks, a simple Gaussian distribution makes it difficult to fully characterize the feature. In the process of global communication, only the parameters of GMM, i.e., the weight, mean, and standard deviation of each Gaussian component are shared with the server. By the definition of the GMM model, these parameters only contain summary statistics of the extracted features and do not contain any identifying information for specific subjects, which can protect data privacy effectively. In particular, for source client C_k^S, the feature $F_{k,i}^S$ extracted by $M_{k,i}^S$ can be expressed as

$$F_{k,i}^S \sim \sum_{j=1}^{m} \pi_{i,j}^{S,k} N\left(X^S; \mu_{i,j}^{S,k}, \sigma_{i,j}^{S,k^2}\right) \tag{4}$$

where m represents the number of components of GMM, $\mu_{i,j}^{S,k}$ and $\sigma_{i,j}^{S,k^2}$ represent the mean and variance of the jth component of the feature $F_{k,i}^S$, respectively, and $\pi_{i,j}^{S,k}$ is the weight of the jth component of GMM (the sum of $\pi_{i,j}^{S,k}$ is 1). Here, the number of components of GMM for each client is assumed to be the same. Thus, the GMM for the feature extracted by the kth client can be defined by the parameters

$$\Theta_{k,i}^S = \left\{\pi_{i,j}^{S,k}, \mu_{i,j}^{S,k}, \sigma_{i,j}^{S,k^2}\right\}, \ j = 1, \ldots, m. \tag{5}$$

To find the optimal parameters for the GMM, the likelihood function of unknown parameters is optimized

$$l\left(\Theta_{k,i}^S\right) = \sum_{n=1}^{n_k^S} \log\left(\sum_{j=1}^{m} \pi_{i,j}^{S,k} \mathbb{N}\left(X_k^S; \mu_{i,j}^{S,k}, \sigma_{i,j}^{S,k^2}\right)\right) \tag{6}$$

for the feature of each client k. To obtain the parameters of GMM, the expectation maximum is employed to update the likelihood function [30].

For the kth source client, the parameter $\{\pi_{i,j}^{S,k}, \mu_{i,j}^{S,k}, \sigma_{i,j}^{S,k^2}\}_{j=1}^m$ can be used to fully characterize its feature information. These parameters do not leak the specific information of local data, only describe the statistics of the whole data samples. Then, each source client uploads the local model $M_{k,i}^S$ and the parameters of the GMM to the central server. After that, the target client will download $M_{k,i}^S$ from the central server. The target feature vector $F_{k,i}^T$ can be achieved by the source model $M_{k,i}^S$. Next, $F_{k,i}^T$ is described as a GMM, and can be defined by the parameters

$$\Theta_{k,i}^T = \left\{\pi_{i,j}^{T,k}, \mu_{i,j}^{T,k}, \sigma_{i,j}^{T,k^2}\right\}_{j=1}^{m}. \tag{7}$$

Then, $\{\Theta_{k,i}^T\}_{k=1}^K$ will be uploaded to the central server for the following distribution discrepancy calculation.

B. Distribution Discrepancy Calculation

The central server, after acquiring the GMM parameters for both the source and target client features, utilizes these parameters to directly compute the distribution difference between each source client and the target client. A modified Bhattacharyya distance is used as an indicator to measure the distribution difference.

The Bhattacharyya distance is a metric that measures the similarity between two probability distributions. It is often used in TL, especially in domain adaptation. The Bhattacharyya distance is intimately linked with the Bhattacharyya coefficient, which quantifies the overlap between two statistical samples or overall distributions [31]. For two continuous probability distributions R and W, the Bhattacharyya coefficient $B_C(R, W)$ and the Bhattacharyya distance $B_D(R, W)$ are defined as

$$B_C(R, W) = \int \sqrt{R(x) \times W(x)}\, dx \tag{8}$$

$$B_D(R, W) = -\ln\left(B_C(R, W)\right) \tag{9}$$

when $R \sim N(\mu_R, \sigma_R^2)$ and $W \sim N(\mu_W, \sigma_W^2)$, $B_D(R, W)$ can be simplified as

$$B_D\ (R, W) = \frac{1}{4}\frac{(\mu_R - \mu_W)^2}{\sigma_R^2 + \sigma_W^2} + \frac{1}{2}\ln\left(\frac{\sigma_R^2 + \sigma_W^2}{2\sigma_R\sigma_W}\right). \quad (10)$$

However, there is no closed-form solution for GMM. Inspired by [32], an approximation for the Bhattacharyya distance between GMMs is adopted, which pays more attention to the difference between the corresponding pairs of Gaussians in two GMMs. The approximation for the Bhattacharyya distance can be expressed as

$$B_{D,i}{}^{\text{GMM}}\left(\Theta_{k,i}^S, \Theta_{k,i}^T\right) = \sum_{j=1}^{m} \pi_{i,j}^{S,k}\pi_{i,j}^{T,k}\left(\frac{1}{4}\frac{\left(\mu_{i,j}^{S,k} - \mu_{i,j}^{T,k}\right)^2}{\sigma_{i,j}^{S,k^2} + \sigma_{i,j}^{T,k^2}} + \frac{1}{2}\ln\left(\frac{\sigma_{i,j}^{S,k^2} + \sigma_{i,j}^{T,k^2}}{2\sigma_{i,j}^{S,k}\sigma_{i,j}^{T,k}}\right)\right) \quad (11)$$

where $\Theta_{k,i}^S$ and $\Theta_{k,i}^T$ are the parameters of GMMs for the source and target clients, respectively, and m represents the number of components of GMM. It is worth noting that the source client shares the same distribution with the target client, i.e., $\Theta_{k,i}^S = \Theta_{k,i}^T$, $B_{D,i}^{\text{GMM}}\left(\Theta_{k,i}^S, \Theta_{k,i}^T\right) = 0$.

C. Central Server Aggregation

FedAvg is widely used in current FL methods, where a simple averaging method is employed to aggregate client models. However, this aggregation strategy has certain drawbacks. It assumes that the contribution of each source client to the global model is the same. In fact, the working conditions of the battery in each client are different, and some source clients always have low-quality data. These low-quality data can have a negative impact on the global model. Thus, it is necessary to develop a dynamic weighted aggregation algorithm that optimizes the global model in favor of the target client by adjusting the weight of each client model.

For the shortcomings of the simple average aggregation method, a dynamic weighted aggregation algorithm based on the modified Bhattacharya distance between the source and target clients and testing error is proposed in this study to guide the global model optimization. For simplicity, we assume that there is one target client. The process of the proposed dynamic weighted aggregation algorithm is as follows.

Step 1: In $i-1$ round of global communication, the central server assigns the initialized global model M_{i-1}^G to all source clients, and the source client $\{C_k^S\}_{k=1}^K$ initialize the local model $\{M_{k,i}^S\}_{k=1}^K$ using M_{i-1}^G. Then, the initialized local models are trained with the local data and the updated local models $\{M_{k,i}^S\}_{k=1}^K$ can be obtained after E epochs of the iterative update.

Step 2: The updated local models $M_{k,i}^S$ are employed to extract the high-level features of the local data D_k^S, i.e., $F_{k,i}^S$, and GMMs with the parameters $\Theta_{k,i}^S$ are used to characterize $F_{k,i}^S$ to avoid direct contact with the local dataset. Then $M_{k,i}^S$ and $\Theta_{k,i}^S$ are both uploaded to the central server.

Step 3: The central server receives all local models $\{M_{k,i}^S\}_{k=1}^K$ and the parameters of GMMs uploaded by the source client, and then distributes all source models to the target client. The estimation accuracy of all source models, i.e., $\{L_{k,i}\}_{k=1}^K$ is evaluated using the labeled data D^T. Meanwhile, the source models are employed to extract the features of D^T. Then the features are modeled as GMMs with the parameters $\{\Theta_{k,i}^T\}_{k=1}^K$. $\{L_{k,i}\}_{k=1}^K$ and $\{\Theta_{k,i}^T\}_{k=1}^K$ are all sent to the central server for the following distribution discrepancy calculation.

Step 4: In the central server, the Bhattacharyya distance $B_{D,i}^{\text{GMM}}(\Theta_{k,i}^S, \Theta_{k,i}^T)$ is calculated with $\{\Theta_{k,i}^S\}_{k=1}^K$ and $\{\Theta_{k,i}^T\}_{k=1}^K$ by (11).

Step 5: Central server calculates the weight of each source model based on the contribution of each source client to the global model, and the contribution value (CV) is defined as

$$\text{CV}_i^k = \frac{1}{L_{k,i} + \alpha B_{D,i}^{\text{GMM}}\left(\Theta_{k,i}^S, \Theta_{k,i}^T\right)} \quad (12)$$

where CV_i^k is determined by the testing accuracy $\text{Loss}_{k,i}$ and the distribution discrepancy $B_{D,i}^{\text{GMM}}(\Theta_{k,i}^S, \Theta_{k,i}^T)$. α is a parameter that is used to balance $\text{Loss}_{k,i}$ and $B_{D,i}^{\text{GMM}}(\Theta_{k,i}^S, \Theta_{k,i}^T)$, and α is set 0.1 in this study. The smaller $\text{Loss}_{k,i}$ and $B_{D,i}^{\text{GMM}}(\Theta_{k,i}^S, \Theta_{k,i}^T)$ indicate that the source model $M_{k,i}^S$ has a significant contribution to the global model.

Based on the calculated CV, the weight of each source model can be expressed by

$$\omega_i^k = \frac{\text{CV}_i^k}{\sum_{k=1}^K \text{CV}_i^k}. \quad (13)$$

Step 6: After the weight of each source model is obtained, all source models are aggregated into the global model by

$$M_i^G = \sum_{k=1}^{K} \omega_i^k\ M_i^k. \quad (14)$$

At this point, the *i*th round of global communication is finished.

When the distribution difference between the source and target clients is calculated based on GMM parameters, only the overall distribution parameters need to be shared. This can avoid direct contact with client private data and effectively protect data privacy. In addition, the weight of each source client is determined according to its contribution to the global model, which is mainly measured by the testing error and the distribution distance between the source client and the target client. A small testing error indicates that the source model can effectively learn the relationship between the extracted features and SOH values, which is positive for the global model. The smaller the distribution distance, the higher the feature similarity between the source client and the target client. Thus, the contribution of

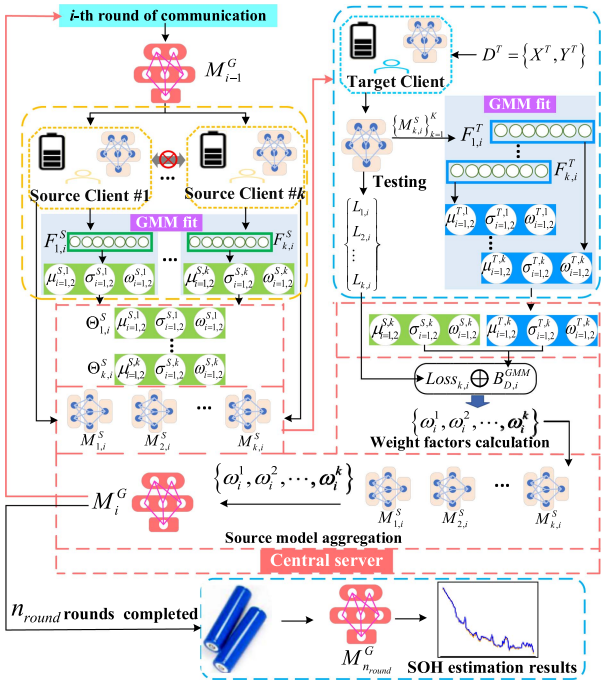

Fig. 4. Procedure of SF-DWFT-based SOH estimation.

the corresponding client model to the global model is greater. Giving larger weight to client models with greater contributions can guide the global model optimize in a direction that is more conducive to the target client. For clients with smaller contributions, smaller weights can also suppress its effect on the global model.

D. Procedure of SF-DWFT-Based SOH Estimation

The existing methods for SOH estimation are mostly based on the centralized training mode, but it is exceedingly challenging for a single user to collect enough data to meet the requirements of deep model training. Also, users are unwilling to share data with others to protect their local private data. In order to solve this issue, an SF-DWFT method for battery SOH estimation is proposed in this study. The FL framework is utilized to extract rich information from the source client without accessing the source client data. SF-DWFT mainly consists of two steps: client local training and central server aggregation. The flowchart of SF-DWFT for battery SOH estimation is shown in Fig. 4.

Client-local training consists of source client-local training and target client testing. First, the global model is downloaded from the central server for initializing the local model. The local private data is used for the model training, then the trained model is employed to extract features from the source data. To avoid direct contact with features, the extracted features are characterized as GMMs. Once the source model training is finished, source clients upload the trained model and the parameters of GMMs. Then, the target client downloads all source models from the central server and extract features of the target data with source models. A small number of labeled samples are used to test the prediction error of source models. Finally, the target client sends the test errors of each source model, and the parameters of GMMs back to the central server. The process of local model training on the client side is presented in detail in Algorithm 1.

In the stage of central server aggregation, a dynamic weighted aggregation method is proposed based on the contributions of source models to the global model, and the contributions of source models are quantified by the testing error and the distribution discrepancy, as (12). The distribution difference between source and target clients can be calculated with the parameters of GMMs uploaded by the source and target clients. The process of the global model aggregation is presented thoroughly in Algorithm 2. After N_{round} rounds of global communication, the

TABLE I
DETAILED INFORMATION OF THE SELECTED BATTERY

Item	Specification	Battery Sample
Rated Capacity (Ah)	0.8	
Standard Voltage (V)	3.7	
Maximum Voltage (V)	4.2	

Algorithm 1: Client Local Training.

Input: Global model M_i^G, K, source client data $\{D_k^S\}_{k=1}^K$, target labeled data D^T

Source clients:
for communication round $i = 1, 2, \ldots, N_{\text{round}}$ **do**
  for each source client $k = 1, 2, \ldots, K$ **do**
    download M_i^G for local model $M_{k,i}^S$ initialization
    train local model $M_{k,i}^S$ with local data D_k^S
    extract high-level features $F_{k,i}^S$ of local data D_k^S by $M_{k,i}^S$
    model $F_{k,i}^S$ as a GMM with the parameter $\Theta_{k,i}^S$
    upload the updated $\{M_{k,i}^S\}_{k=1}^K$, and $\{\Theta_{k,i}^S\}_{k=1}^K$
  end for
end for
Target client:
for communication round $i = 1, 2, \ldots, N_{\text{round}}$ **do**
  download $\{M_{k,i}^S\}_{k=1}^K$ from central server
  for each source model $M_{k,i}^S$, $k = 1, 2, \ldots, K$ **do**
    extract high-level features $F_{k,i}^T$ of the target data D_k^T using $M_{k,i}^S$
    model the feature $F_{k,i}^T$ as a GMM with the parameters $\Theta_{k,i}^T$
    test $M_{k,i}^S$ with the target data D^T and obtain test error $L_{k,i}$
    upload $\{\Theta_{k,i}^T\}_{k=1}^K$, and $\{L_{k,i}\}_{k=1}^K$
  end for
end for

Output: The updated source models $\{M_{k,i}^S\}_{k=1}^K$, the parameters of the source GMMs $\{\Theta_{k,i}^S\}_{k=1}^K$, the parameters of the target GMMs $\{\Theta_{k,i}^T\}_{k=1}^K$, and the test error $\{L_{k,i}\}_{k=1}^K$.

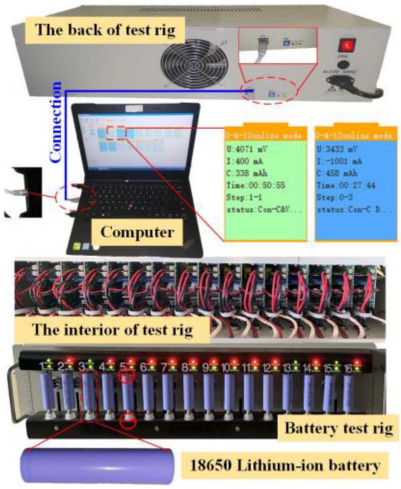

Fig. 5. Test rig of 18 650 lithium-ion battery.

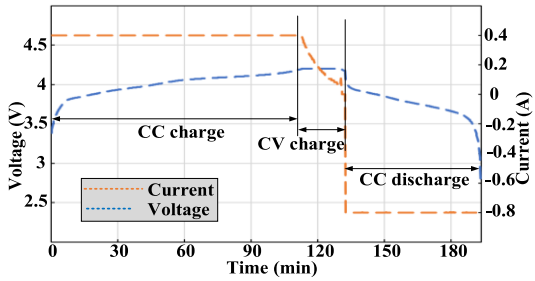

Fig. 6. Charging and discharging process in one cycle of battery.

TABLE II
DETAILS OF 18 650 LITHIUM-ION BATTERY

Condition	CC (A)	CV(V)	DC (A)	Battery No.
1	0.4	4.2	0.4	B5, B6, B7, B8
2	0.4	4.2	0.6	B9, B10, B11, B12
3	0.4	4.2	1.6	B13, B14, B15, B16

Algorithm 2: Central Server Aggregation.

Input: The updated source model $\{M_{k,i}^S\}_{k=1}^K$, the GMMs of source clients $\{\Theta_{k,i}^S\}_{k=1}^K$, the GMMs of the target client $\{\Theta_{k,i}^T\}_{k=1}^K$, the test error $\{L_{k,i}\}_{k=1}^K$.

Central server:
  Initialize the global model M_0^G
  for communication round $i = 1, 2, \ldots, N_{\text{round}}$ **do**
    distribute global model M_i^G to each source client
    receive $\{M_{k,i}^S\}_{k=1}^K$, and $\{\Theta_{k,i}^S\}_{k=1}^K$
    receive $\{\Theta_{k,i}^T\}_{k=1}^K$, and $\{L_{k,i}\}_{k=1}^K$
    // Distribution discrepancy calculation
    calculate the distribution discrepancy $B_{D,i}^{\text{GMM}}$ between $\Theta_{k,i}^S$ and $\Theta_{k,i}^T$ by Eq. (11)
    // Contribution of each client calculation
    calculate the contribution value CV_i^k by Eq. (12)
    // Weight factors calculation
    calculate the weight of each source model by Eq. (13)
    // Source models aggregation
    aggregate source models into the global model M_i^G by Eq. (14)
  end for

Output: The global model M^G

trained global model M_i^G is downloaded by the target client for battery SOH estimation.

III. Experimental and Result Analysis

A. Case 1: 18 650 Lithium-Ion Battery Datasets

In this section, the effectiveness of SF-DWFT for battery SOH estimation is validated on the experimental datasets of commercial 18 650 lithium-ion batteries. The datasets are collected on the battery test rig, as shown in Fig. 5. The detailed parameters of 18 650 lithium-ion battery are listed in Table I. The experiments record the degradation of batteries during cyclic charging and discharging processes. Fig. 6 shows the detailed experiment processes in one cycle. Regarding the charging process, the constant current constant voltage (CC-CV) mode is used to charge the batteries. The batteries are first charged at a 0.5C constant current until the voltage reached 4.2 V. Then, the voltage keeps unchanged until the charging current dropped to 0.05 A. A CC mode is adopted to discharge the batteries until the voltage reaches a cut-off voltage, which is 2.75 V. In this study, to verify the effectiveness of SF-DWFT, three subdatasets from three different working conditions are considered. Table II presents detailed information on these datasets, where dc represents discharging current. The dataset for each working condition contains four batteries. These batteries are all charged with the same CC-CV charging mode. The difference between three different working conditions is the constant current rate in the discharging process and the constant current rate is set to 0.5, 0.75, and 2C, respectively.

In the experiment, the voltage and current data of the battery were recorded each 10 s, and after each cycle of charging and discharging was finished, the direct current internal resistance value of the battery was recorded. In this study, for better describing the degradation process of batteries, the internal resistance, discharging time, and the mean of the voltage and current for each cycle of discharging are selected as the input of SF-DWFT. Before the data is input into the network, data normalization is necessary. The normalization is defined as follows:

$$x_{\text{norm}} = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \tag{15}$$

where x_{norm} is the normalized data, $x_{\max}$ and $x_{\min}$ are the maximum and minimum values of the input data x, respectively. Battery capacity is widely used as an indicator to measure the SOH of batteries [17]. Fig. 7 shows the capacity degradation process of batteries B5, B6, B7, and B8. When the battery capacity drops to a certain extent, the battery will fail. Thus, the expression for SOH is as follows:

$$\text{SOH} = \frac{\text{Capacity}_t}{\text{Capacity}_0} \times 100\% \tag{16}$$

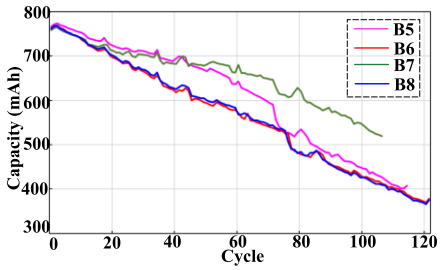

Fig. 7. Capacity degradation process of four batteries (B5, B6, B7, and B8).

TABLE III
DETAILS OF THE LOCAL DATA OF SOURCE CLIENTS AND TARGET CLIENTS IN THESE EXPERIMENTS

Task	Transfer direction	Source client	Target client
D1	C1+C2→C3	C1, C2	C3
D2	C1+C3→C2	C1, C3	C2
D3	C2+C3→C1	C2, C3	C1

where Capacity_t represents the battery capacity at the t th cycle. Capacity_0 represents the rated capacity.

1) Experimental Settings: It is assumed that the FL system has two source clients and one target client, and it can also be easily extended to other situations. Considering that the working condition of batteries held by multiple users are different, the battery datasets under three operating conditions are used as local datasets for two source clients and one target client respectively to simulate the scenario of FL. The task of the system is to combine multiple source clients to construct an SOH estimation model for the target client. In these experiments, the detailed setup of tasks is listed in Table III, in which C1 denotes the dataset collected under working condition 1.

All experiments are conducted on a laptop with an Inter I9-13900 CPU, 16G RAM, and a GeForce RTX 4060, and the SOH estimation models were constructed with torch 2.1.0 in python 3.8. Under the same experimental settings, SF-DWFT is compared with three kinds of methods, including Only local-training methods, Centralized TL methods, and FL methods.

1) Only local-training methods

These methods only use local datasets for training and have no connection with other clients.

Baseline: Each source client trains the model using local data separately. The central server computes a simple average of the results from each client. This method often yields the worst results.

OneByOne: The central server communicates with all source clients in sequence. This learning scheme is likely to cause catastrophic forgetting.

2) Centralized TL methods

A centralized training mode is adopted, and the private data can be shared during the training process. They include MFSAN [33] and MDAN [34], which have shown excellent performance in multisource domain adaptation and provide an upper bound for FTL methods.

MFSAN: It has a two-stage alignment framework consisting of domain-specific distribution alignment and domain-specific classifier alignment to realize knowledge transfer.

TABLE IV
PARAMETER SETUP OF THE BASIC MODEL

Structure	Parameters	Output
Input	Multiple features	[4×1]
BGRU1	Number of hidden nodes =4-32	[64×1]
BGRU2	Number of hidden nodes =64-32	[64×1]
BGRU3	Number of hidden nodes =64-32	[64×1]
SOH estimator	Number of hidden nodes =64-16-1	[1]

TABLE V
OTHER HYPERPARAMETER SETTINGS OF SF-DWFT

Parameter	Batch size	Local epoch	Learning rate	Communication rounds
	20	50	0.001	20

MDAN: A mixed metric combining MMD and CORAL distance is used to achieve domain adaptation.

3) FLmethods

FL methods include FedAvg [20] and some state-of-the-art FL methods, including FedMMD [25], FedGA [35] and FedIIR [36]. FedMMD, FedGA, and FedIIR are all proposed in the past two years, and achieved good performance.

FedAvg: FedAvg is a typical FL method that does not consider the impact of distribution differences between domain clients and target clients on the global model. A simple averaging method is used to aggregate all client models.

FedMMD: FedMMD is a dynamic weighted FL method based on MMD, which allows clients to upload extracted advanced features. MMD distance is used to measure the distribution discrepancy.

FedGA: FedGA dynamically adjusts aggregation weights and optimizes global objectives through a generalized adjustment method.

FedIIR: FedIIR implicitly realizes out-of-distribution generalization through inter-gradient alignment between clients.

For a fair comparison, SF-DWFT and all comparison methods share the same backbone. The detailed network structure is shown in Table IV. The feature extractor consists of three stacked BGRUs with hidden nodes set to 32. The SOH estimator uses a "FC (64)-FC (16)-FC (1)" structure to output SOH estimation. The performance of SF-DWFT is affected by batch size, learning rate, and communication rounds. Table V lists the settings for other hyperparameters of SF-DWFT. To ensure that all comparison methods are tuned to their respective optimal, the learning rate is adjusted based on changes in loss dynamically during the training process. The Adam optimizer is used for parameter optimization by all the models. For Baseline, OneByOne and centralized TL methods, the training epoch is set to 100. For FL methods, to ensure the convergency of both local models and the global model, the training epoch of local models is set to 50 and the maximum of the global round is 20 through multiple experiments and adjustments. During the training process, the learning rate is dynamically adjusted based on the training loss to accelerate model convergence. The initial learning rate is set to 0.001. To avoid the impact of randomness on the experiment, all the models run five times, and the average of five recorded results is taken as the final result.

TABLE VI
ESTIMATION RESULTS OF SF-DWFT AND OTHER COMPARISON METHODS

Model		D1		D2		D3		Avg	
		RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE
Only local-training methods	Baseline	0.1231	0.1117	0.0995	0.0827	0.0596	0.0546	0.0940	0.0830
	OneByOne	0.1046	0.0898	0.1012	0.0828	0.0519	0.0440	0.0869	0.0722
Centralized TL methods	MFSAN [32]	0.0106	0.0089	0.0155	0.0140	0.0110	0.0102	0.0123	0.0110
	MDAN [33]	0.0110	0.0093	0.0151	0.0139	0.0125	0.0114	0.0128	0.0115
FL methods	FedAvg [20]	0.0745	0.0663	0.0842	0.0652	0.0563	0.0385	0.0716	0.0566
	FedMMD [25]	0.0123	0.0128	0.0238	0.0259	0.0112	0.0088	0.0157	0.0158
	FedGA [34]	0.0227	0.0192	0.0462	0.0378	0.0303	0.0257	0.0330	0.0275
	FedIIR [35]	0.0254	0.0215	0.0268	0.0215	0.0202	0.0163	0.0241	0.0197
	SF-DWFT	0.0116	0.0091	0.0182	0.0142	0.0126	0.0104	0.0141	0.0112

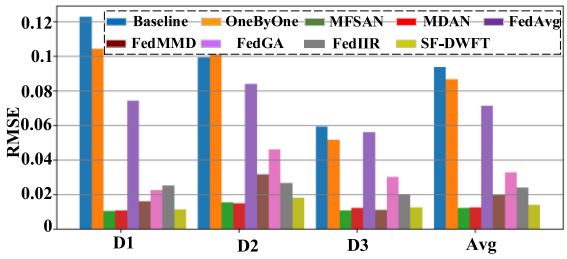

Fig. 8. RMSE results of SF-DWFT and other methods in all tasks.

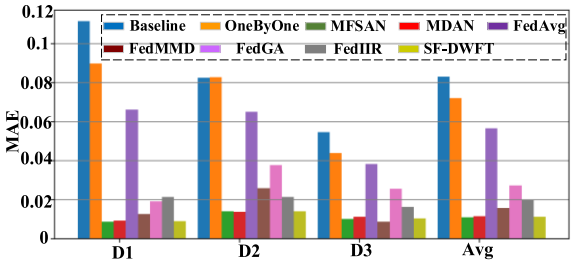

Fig. 9. MAE results of SF-DWFT and other methods in all tasks.

In order to quantify the estimation performance of the model, root-mean-square error (RMSE) and mean absolute error (MAE) are employed to evaluate the performance

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\mathrm{SOH}^{i}_{\mathrm{actual}} - \mathrm{SOH}^{i}_{\mathrm{pre}}\right)^{2}} \tag{17}$$

$$\mathrm{MAE} = \frac{1}{N}\sum_{i=1}^{N}\left|\mathrm{SOH}^{i}_{\mathrm{actual}} - \mathrm{SOH}^{i}_{\mathrm{pre}}\right| \tag{18}$$

where $\mathrm{SOH}^{i}_{\mathrm{actual}}$ and $\mathrm{SOH}^{i}_{\mathrm{pre}}$ are the ith actual and predicted SOH values, respectively, and N is the number of data sample.

2) Results Analysis: Figs. 8, 9, and Table VI show the SOH estimation results of SF-DWFT and other comparison methods. First, it can be found that the only local-training methods have the worst prediction performance, which indicates that a single client is difficult to train an ideal model alone with limited local data. Second, MFSAN achieves the highest estimation accuracy, which is an upper bound for the FTL methods. Generally, the centralized TL methods achieve the highest estimation accuracy in almost all tasks. This is because MFSAN and MDAN utilize TL to fuse multiple source domain data, which can effectively improve the generalization ability. However, this method has a drawback that data privacy is not considered during training. Finally, compared to FL methods, e.g., FedAvg, FedMMD, FedGA, and FedIIR, SF-DWFT shows the best performance of SOH estimation. Particularly, SF-DWFT has 80.3% higher (0.0716→0.0141) estimation accuracy than FedAvg, and 10.2% (0.0157→0.0141), 57.2% (0.0330→0.0141) and 41.4% (0.0241→0.0141) higher accuracy than FedMMD, FedGA, and FedIIR, respectively. Moreover, Fig. 10 presents the SOH estimation results of SFDWFT and comparison methods on B14 in task D1, B11 in task D2, and B7 in task D3. It can be found that SF-DWFT is stable and achieves high accuracy in all tasks. The results indicate that SF-DWFT takes into account the distribution differences between the source and target clients, and dynamically adjusts the aggregation weights of the source

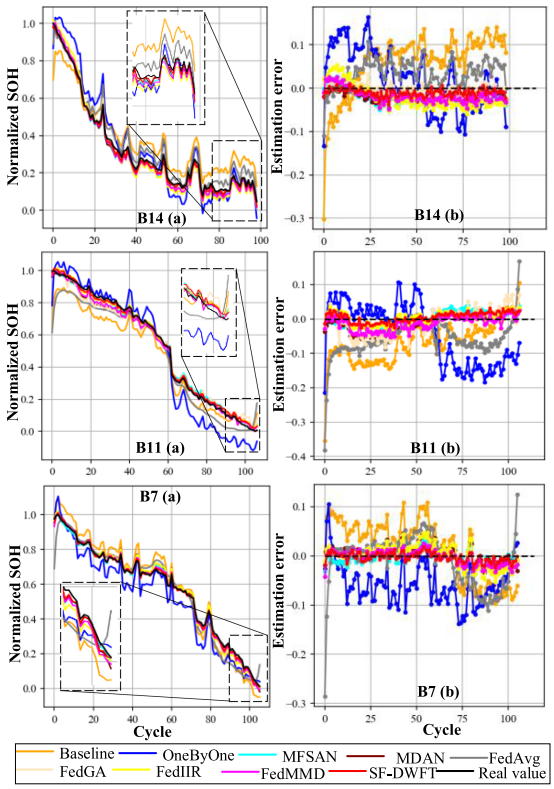

Fig. 10. SOH estimation results for B14 in task D1, B11 in task D2, and B7 in task D3. (a) Estimated SOH. (b) Estimated error.

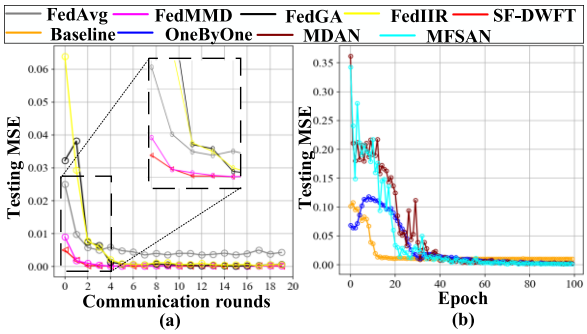

Fig. 11. Testing MSE values of SF-DWFT and comparison methods in different communication rounds and epochs.

TABLE VII
SOH ESTIMATION RESULTS OF SD-DWFT WITH DIFFERENT NUMBER OF GMM COMPONENTS

Number of components		m=1	m=2	m=3
D1	RMSE	0.0156	0.0116	0.0172
	MAE	0.0095	0.0091	0.0141
D2	RMSE	0.0205	0.0182	0.0209
	MAE	0.0167	0.0142	0.0161
D3	RMSE	0.0146	0.0126	0.0139
	MAE	0.0131	0.0104	0.0108
Avg	RMSE	0.0169	**0.0141**	0.0173
	MAE	0.0131	**0.0112**	0.0136

model, which can guide the global model to optimize in a direction favorable to the target client.

Fig. 11 presents the testing MSEs of SF-DWFT and other FL methods in different communication rounds in Task 1. It can be observed from Fig. 11(a) that after three rounds of global communication, the estimation accuracy of SF-DWFT will be stable. Compared with other FL methods, SF-DWFT requires minimal global communication to achieve convergence, which is of practical significance. Also, the frequent communication between servers and clients is likely to cause privacy leakage [38]. After eight global rounds, other FL methods reach convergence. Fig. 11(b) presents the testing MSEs of the only local-training methods and centralized TL methods. It can be found that after 40 epochs, baseline, OneByOne, MDAN, and MFSAN all reach convergence. Thus, it is reasonable to set the maximum global round and maximum epoch to 20 and 100, respectively, which can ensure that SF-DWFT and other comparison methods converge and achieve their respective optimal.

3) Parameter Sensitivity Analysis: In local training phase, each client utilizes a GMM to model the extracted features and uploads the parameters of the GMM to the central server instead of the extracted high-level features for data privacy. The number of components of GMM is an important parameter that affects GMM fitting on the features. Table VII shows the results of SF-DWFT with different numbers of GMM components. As shown in Table VII, when the GMM with two components is employed to model the extracted feature, SF-DWFT achieves the best performance. However, the advantage of two components over one component and three components is not very obvious, which indicates that SF-DWFT has limited sensitivity to the number of components.

TABLE VIII
BATTERY DATASET DIVISION FOR THREE APPLICATION SCENARIOS

	Source client 1	Source client 2	Target client
Scenario 1	B5	B6	B7
Scenario 2	B5	B10	B7
Scenario 3	B5	B6	B10

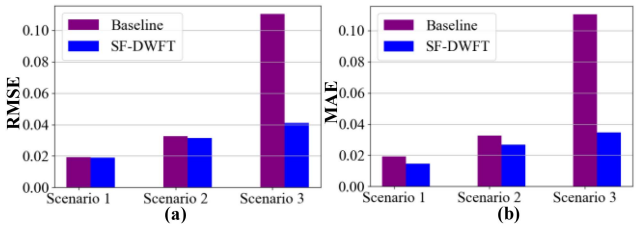

Fig. 12. Results of SF-DWFT and Baseline in three application scenarios.

4) Analysis of the Impact of Different Scenarios: In this study, it is assumed that battery datasets from source clients and target clients are collected under different working conditions. To verify the applicability of SF-DWFT in different scenarios, the simulation experiments consider various scenarios. The specific experimental scenario settings are shown in Table VIII. Scenario 1 simulates that the batteries held by source and target clients run under the same operating conditions. Scenario 2 simulates that the data held by the source client is collected under different working conditions, and the data held by the target client is collected under the working conditions similar to one of the source clients. Scenario 3 denotes the situation that the data held by source clients is from the same operating conditions, while the target client is from another operating conditions.

The estimation results of SF-DWFT and Baseline in three application scenarios are shown in Fig. 12. It can be seen that in Scenario 1 and 2, the estimation accuracy of SF-DWFT and Baseline is similar, while SF-DWFT is slightly better than baseline. This is because the source client always has data with a similar distribution to the target client data in Scenario 1 and 2. In Scenario 3, SF-DWFT has excellent estimation performance, while the results of baseline are not satisfactory. The battery data of the target client and source client are collected under different operating conditions in scenario 3, which results in significant differences in data quality and statistical distribution. SF-DWFT reduces the impact of different operating conditions on the model by assigning larger weights to clients with greater contributions. Hence, SF-DWFT can be applied to different application scenarios and achieve satisfactory performance.

5) Visualization Analysis: Generally, due to the influence of temperature and loading profiles, there is a distribution difference between different lithium-ion batteries. To address this issue, SF-DWFT is proposed to reduce the impact of distribution discrepancy on battery SOH estimation. In order to demonstrate the performance of SF-DWFT in feature alignment, the distributions of the different batteries at the beginning and end of global communication are visualized. Fig. 13 shows the probability density distribution of B5, B9, and B13 before and after communication. It can be observed that there are significant differences in the distribution of batteries from different

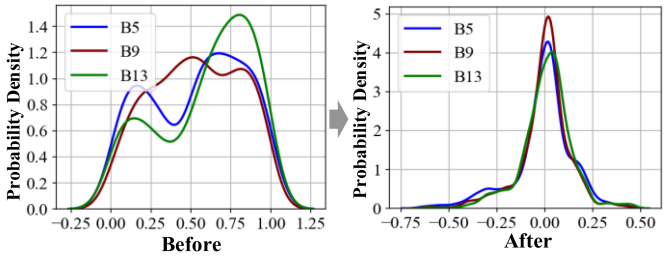

Fig. 13. Probability density distribution of battery B5, B9, and B13 before and after global communication.

TABLE IX
DETAILED DESCRIPTION ABOUT BATTERY BATTERIES UNDER SUBZERO TEMPERATURE

Condition	C1	C2	C3
Battery No.	B2, B3, B4, B6	B7, B8, B9, B10	B11, B12, B13, B14
CV (mV)	4200	4200	4200
CC (mA)	2400	2400	2400
Charging rate	2C	2C	2C
DC (mA)	2400	1200	1800
Discharging rate	2C	1C	1.5C

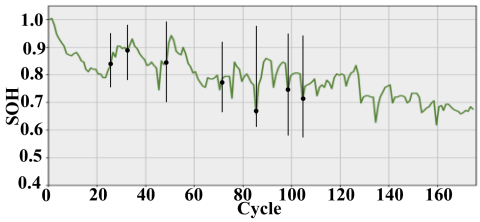

Fig. 14. Capacity degradation process of battery B13.

operating conditions, which indicates a significant domain-shift of batteries from different source clients. After the global communication is completed, the distribution of different batteries is very similar. This demonstrates that SF-DWFT is effective in reducing distribution discrepancy between source and target clients.

6) SOH Estimation Under Subzero Temperature: Temperature has a significant influence on the degradation process of batteries. To verify the effectiveness of SF-DWFT under subzero temperature conditions, three sets of battery degradation data with different loads are collected under −9 °C temperature. Nickel manganese cobalt batteries with a rated capacity of 1200 mAh were used in the experiment. During the charging and discharging process, a CC-CV mode is adopted to charge the battery, and then a CC mode is used to discharge the battery. Three different loading conditions constitute three operating conditions of the battery. Each operating condition contains four batteries. Table IX provides a detailed description of three operating conditions. During the experiment, batteries are randomly placed on hold. The ambient temperature of the battery remains unchanged during this period. Fig. 14 shows the degradation process of battery B13. It can be observed that under subzero temperature, the capacity regeneration phenomenon is more obvious, especially after batteries are placed on hold for a period of time, which poses a challenge to the health assessment

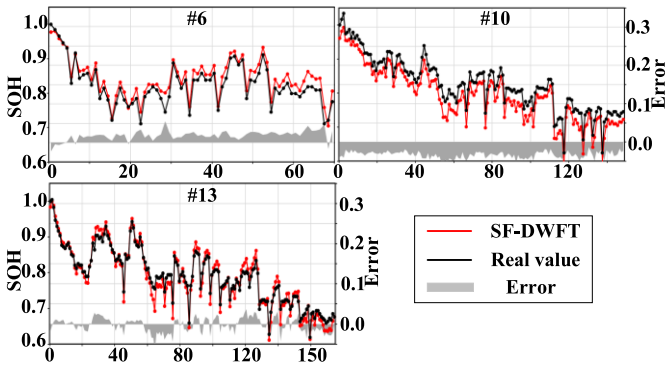

Fig. 15. SOH estimation results of SF-DWFT on B6 in task C2+C3→C1, B10 in task C1+C3→C2, and B13 in task C1+C2→C3.

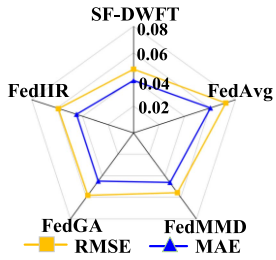

Fig. 16. Comparison results between SF-DWFT and other FL methods on batteries under subzero temperature.

of batteries. The partial charging time points after batteries are placed on hold have been marked in Fig. 14. The data of voltage, current, time, and internal resistance during the discharge stage are used as input of SF-DWFT for battery SOH estimation. The data collection and preprocessing are similar to those of 18 650 lithium-ion batteries.

According to the different directions of data migration, the experimental tasks are divided into three tasks, including C1+C2→C3, C1+C3→C2, and C2+C3→C1. The datasets from two of the working conditions are considered as local data for the source clients, and the data from the other working conditions is used as the test data for the target client. Fig. 15 shows the estimation results of SF-DWFT on B6, B10, and B13. It can be found that the estimation error of SF-DWFT is controlled within 5%. Thus, SF-DWFT can effectively deal with the capacity regeneration phenomenon of batteries during charging and discharging, and accurately evaluate battery SOH. Fig. 16 presents the comparison results between SF-DWFT and other FL methods, from which it can be seen that SF-DWFT achieves the minimum RMSE and MSE values. Thus, SF-DWFT can achieve high estimation accuracy at low temperatures.

7) Computational Efficiency Analysis: To further verify the implement-ability of SF-DWFT in practical industrial applications, the computational efficiency of SF-DWFT and other comparison methods is quantitatively provided from two aspects: offline training and real-time estimation. For offline training, the average training time in each epoch and global aggregation time is calculated. In terms of real-time evaluation, to simulate the real data flow, all test samples are one by one estimated by the

TABLE X
COMPUTATIONAL COST AND ESTIMATION TIME OF SF-DWFT AND COMPARISON METHODS

Methods	Local training time (ms)	Aggregation time (ms)	Estimation time (ms)
MFSAN	48.87	/	0.21
MDAN	50.51	/	0.16
FedAvg	46.55	2.01	0.22
FedMMD	56.91	9.35	0.32
FedGA	13.77	11.25	0.84
FedIIR	46.51	2.58	0.32
SF-DWFT	48.40	6.39	0.22

TABLE XI
COMPUTATIONAL COST OF SF-DWFT ON DIFFERENT DEVICES

Device	GPU	CPU	Estimation time (ms)
IPC1	RTX4060	i9-13900HX	0.219
IPC2	RTX3060	R7 6800HX	0.560
IPC3	RTX4060	R9 7945HX	0.231
IPC4	GTX1050	i5-7300HQ	1.218
IST1	RTX3090	12th i9-12900K	0.725
IST2	RTX4090	Xeon Gold 5218R	0.211
IST3	RTX4090	12th i9-12900K	0.216
IST: Industrial work station; IPC: Industrial personal computer			

TABLE XII
DETAILED DESCRIPTION ABOUT NASA BATTERIES

Condition	CC (A)	AT (°C)	DC (A)	Battery No.
C1	1.5	24	2	#5, #6, #7, #18
C2	1.5	43	4	#29, #30, #31, #32
C3	1.5	4	1	#45, #46, #47, #48

trained SF-DWFT, and the average estimation time is recorded. Table X lists the detailed comparison results, which illustrates that SF-DWFT has good performance in both offline training and real-time estimation and can effectively achieve real-time health assessment. Table XI presents an online estimation cost on various industrial computational devices. The online estimation time on IPC4, which has relatively poor hardware, is only 1.218 ms. Thus, SF-DWFT can achieve good computational efficiency across different devices.

B. Case 2: NASA Battery Datasets

To further verify the performance of SF-DWFT, experimental studies are conducted on battery datasets from the NASA Ames prognostics center of excellence [38]. This dataset includes a set of batteries that operate with different operating profiles at different ambient temperatures. Three sets of battery data under different working conditions are selected as the research subjects. The specific information of the dataset is shown in Table XII, where AT represents the ambient temperature, CC represents the charging current and DC represents the discharging current. Each operating condition includes four batteries. The difference between these three working conditions is the different ambient temperatures and discharge rates. In condition 1, #5, #6, #7, and #18 batteries are run under an ambient temperature of 24 °C. These batteries are first charged at a 1.5 A constant current until the voltage reached 4.2 V. A CC mode with 2 A is adopted to discharge the batteries until the voltage reaches a cut-off voltage and the cut-off voltages for these four batteries are 2.7, 2.5, 2.2, and 2.5 V, respectively. To better describe the degradation process of batteries, the measured current, voltage, charging current, charging voltage, temperature, and time are used as the inputs of SF-DWFT. According to the directions of knowledge transfer between source clients and the target client, there are three tasks in the experiment. The battery datasets under two of the operating conditions are respectively used as local data for two source clients, while the battery dataset under the other operating conditions is used as local data for the target client. The specific experimental settings are shown in Table XIII.

TABLE XIII
DETAILS OF THE LOCAL DATA OF SOURCE CLIENTS AND TARGET CLIENT IN THE EXPERIMENTS OF NASA BATTERIES

Task	Transfer direction	Source client	Target client
T1	C1+C2→C3	C1, C2	C3
T2	C1+C3→C2	C1, C3	C2
T3	C2+C3→C1	C2, C3	C1

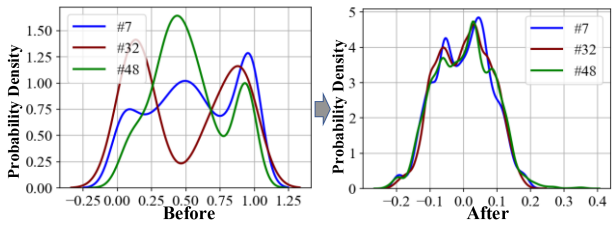

Fig. 17. Probability density distribution of battery #7, #32, and #48 before and after global communication.

Fig. 17 shows the probability density distribution of # 7, # 32, and # 48 before and after global communication. These batteries are from different operating conditions. It can be found that SF-DWFT can achieve feature alignment and effectively reduce the distribution discrepancy between source and target clients.

Fig. 18 presents the specific SOH estimation results of #45 in task T1, #39 in task T2, and #7 in task T3, and Table XIV shows the comparison results of SF-DWFT and other methods. It can be easily found that SF-DWFT can achieve great performance and estimate SOH accurately. Compared to centralized TL methods, SF-DWFT achieves a little higher value of RMSE and MAE. Among only local-training methods and FL methods, the RMSE and MAE of SF-DWFT are the smallest, which demonstrates that SF-DWFT can also achieve high SOH estimation accuracy when protecting data privacy.

In this section, to further verify the effectiveness of SF-DWFT under dynamic loading profiles, a random walk (RW) battery dataset [39] is used. The dataset records the degradation of batteries in a dynamic charging and discharging conditions. The detailed description of the dataset is shown in Table XV. In condition 1, four batteries are charged for a random duration, and the specific charging time is randomly selected from the set {0.5 h, 1 h, 1.5 h, 2 h, or charge until full}. Then, the battery is discharged with a random discharge current ranging from 0.5 to 4 A. For condition 2, the charging process of batteries remains unchanged. The batteries are charged to 4.2 V with a constant current and voltage mode, and then discharged to 3.2 V with a random discharge current. Similarly, for condition 3, a current sequence between -4.5 and 4.5 A is used to continuously charge and discharge the battery, and negative current

TABLE XIV
ESTIMATION RESULTS OF SD-DWFT AND OTHER COMPARISON METHODS ON NASA DATASETS

Model		T1		T2		T3		Avg	
		RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE
Only local-training methods	Baseline	0.3711	0.3612	0.2485	0.2148	0.2555	0.2215	0.2917	0.2658
	OneByOne	0.1053	0.0977	0.1340	0.1114	0.1359	0.1191	0.1250	0.1094
Centralized TL methods	MFSAN	0.0684	0.0552	0.0510	0.0374	0.0585	0.0460	0.0593	0.0462
	MDAN	0.0665	0.0601	0.0437	0.0340	0.0580	0.0444	0.0560	0.0461
FL methods	FedAvg	0.1525	0.1373	0.0984	0.0781	0.1346	0.1108	0.1285	0.1087
	FedMMD	0.0890	0.0674	0.0684	0.0543	0.0785	0.0571	0.0786	0.0596
	FedGA	0.0749	0.0604	0.0704	0.0581	0.0994	0.0733	0.0815	0.0639
	FedIIR	0.0710	0.0574	0.0937	0.0799	0.1040	0.0720	0.0895	0.0697
	SF-DWFT	0.0840	0.0591	0.0536	0.0417	0.0493	0.0330	0.0623	0.0446

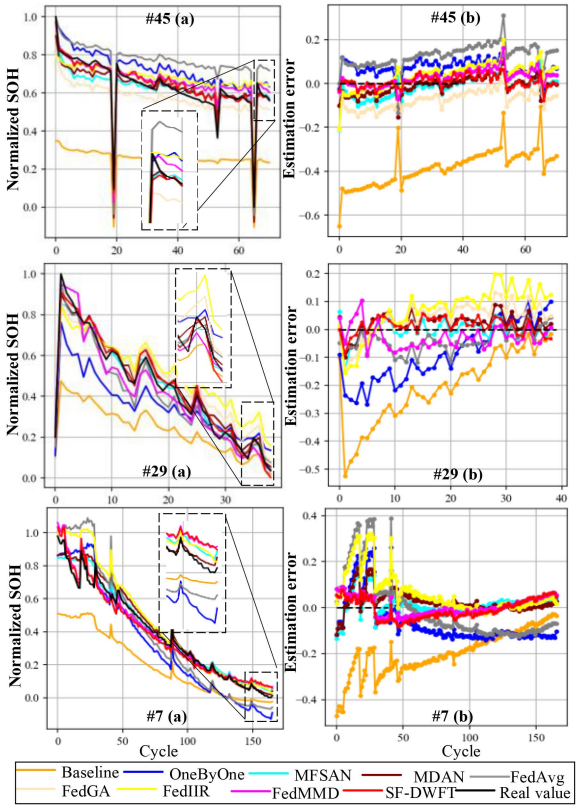

Fig. 18. SOH estimation results for #45 in task T1, #29 in task T2, and #7 in task T3. (a) Estimated SOH. (b) Estimated error.

TABLE XV
RW BATTERY DATASETS DESCRIPTION

Condition	Battery Num.	RW mode
1	RW #1, RW #2, RW #7, RW #8	Randomized charging time length (0.5 hours-charge until full) and randomized discharging current (0.5A-4A)
2	RW #3, RW #4, RW #5, RW #6	Randomized discharging current (0.5A-4A)
3	RW #9, RW #10, RW #11, RW #12	Randomized sequence currents (-4.5A-4.5A)

represents discharge, positive current represents charging. The discharge voltage, current, temperature, and duration of batteries are selected as input of the model to estimate the battery SOH.

According to the direction of data transfer, the experiment consists of three tasks, i.e., C1+C2→C3, C1+C3→C2, C2+C3→C1. Table XVI presents the comparison results of

TABLE XVI
RMSE AND MAE RESULTS OF SF-DWFT AND OTHER FL METHODS ON RW BATTERY DATASETS

Tasks		SF-DWFT	FedAvg	FedMMD	FedGA	FedIIR
C1+C2→C3	RMSE	0.0398	0.0693	0.0133	0.0514	0.0584
	MAE	0.0304	0.0595	0.0085	0.0424	0.0515
C1+C3→C2	RMSE	0.0140	0.0362	0.0276	0.0194	0.0350
	MAE	0.0110	0.0300	0.0203	0.0159	0.0293
C2+C3→C1	RMSE	0.0222	0.0668	0.0452	0.0281	0.0365
	MAE	0.0178	0.0554	0.0431	0.0221	0.0301
Avg	RMSE	**0.0253**	0.0574	0.0287	0.0329	0.0433
	MAE	**0.0197**	0.0483	0.0239	0.0268	0.0369

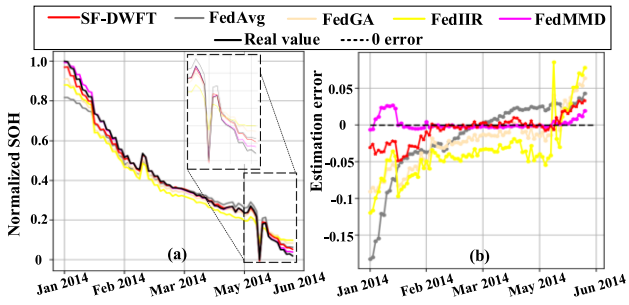

Fig. 19. SOH estimation results for RW #10 in task C1+C2→C3. (a) Estimated SOH. (b) Estimated error.

TABLE XVII
DETAILED INFORMATION OF BATTERY PACKS AND THE DATASETS UNDER THREE DIFFERENT OPERATION CONDITIONS

Battery pack	Rated Capacity	Maximum Voltage	Cell Chemistry
	3600mAh	4200mV	NCM
Condition	**Battery pack No.**	**CV (mV)**	**DC (mA)**
C1	BP#1, BP#2, BP#3	4200	3600
C2	BP#4, BP#5, BP#6	4200	5400
C3	BP#7, BP#8, BP#9	4200	6300

SF-DWFT and other FL methods. It can be found that the RMSE and MAE of SF-DWFT are the smallest among these FL methods, which illustrates that the dynamic weighted aggregation algorithm can effectively mitigate the influence of distribution discrepancy on the global model, and SF-DWFT can also achieve high evaluation accuracy under dynamic load conditions. Fig. 19 shows the estimation results of RW #10 in task C1+C2→C3. The results illustrate that during the degradation process, SF-DWFT evaluates the SOH of battery RW #10 with a maximum

TABLE XVIII
ESTIMATION RESULTS OF SF-DWFT AND OTHER COMPARISON METHODS ON BATTERY PACK DATASETS

Model		C1+C2→C3		C1+C3→C2		C2+C3→C1		Avg	
		RMSE	MAE	RMSE	MAE	RMSE	MAE	RMSE	MAE
Only local-training methods	Baseline	0.0560	0.0452	0.0414	0.0326	0.0730	0.0546	0.0568	0.0441
	OneByOne	0.0460	0.0406	0.0359	0.0306	0.0525	0.0421	0.0448	0.0377
Centralized TL methods	MFSAN	0.0118	0.0092	0.0211	0.0130	0.0133	0.0073	0.0154	0.0098
	MDAN	0.0136	0.0111	0.0152	0.0171	0.0119	0.0109	0.0135	0.0130
FL methods	FedAvg	0.0448	0.0352	0.0385	0.0327	0.0480	0.0380	0.0437	0.0353
	FedMMD	0.0215	0.0198	0.0183	0.0123	0.0299	0.0245	0.0232	0.0188
	FedGA	0.0130	0.0097	0.0244	0.0193	0.0464	0.0361	0.0279	0.0217
	FedIIR	0.0243	0.0210	0.0192	0.0165	0.0460	0.0340	0.0298	0.0238
	SF-DWFT	0.0231	0.0155	0.0162	0.0114	0.0206	0.0186	0.0199	0.0151

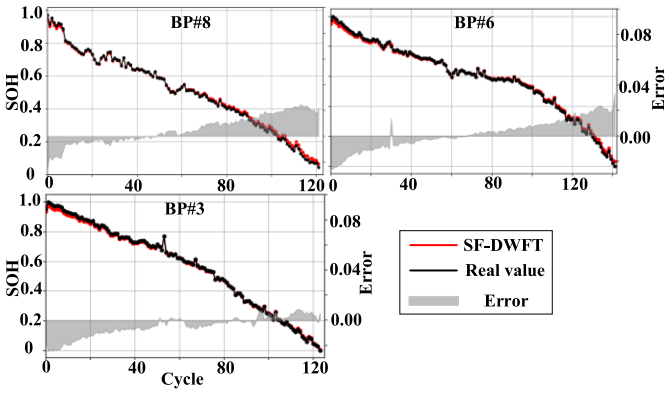

Fig. 20. SOH estimation results of SF-DWFT on BP#8 in task C1+C2→C3, BP#6 in task C1+C3→C2, and BP#3 in task C2+C3→C1.

error of 0.05. Thus, SF-DWFT also can achieve good SOH estimation performance under dynamic loading profiles.

C. Case 3: Battery Pack Datasets

This section considers a battery pack dataset to confirm the effectiveness of SF-DWFT on battery packs. Each battery pack is composed of three 1200 mAh lithium batteries in parallel. The experiment process of these battery packs is similar to that of 18 650 lithium-ion batteries. When the capacity of the battery pack drops to 80%, i.e., 1880 mAh, the battery pack has failed. The degradation data of the battery pack is recorded under three different operating conditions, and each condition contains three battery packs. The specific parameters of battery packs and detailed information about the datasets under different operating conditions are shown in Table XVII. Four types of sensor signals (i.e., the discharging voltage, current, time, and internal resistance) are used as inputs for SOH estimation of battery packs.

Fig. 20 presents the SOH estimation results of SF-DWFT on BP#8 in task C1+C2→C3, BP#6 in task C1+C3→C2, and BP#3 in task C2+C3→C1. It is clear that SF-DWFT can accurately evaluate battery pack SOH throughout the entire process of degradation, with an error almost within 2%, which is of great significance for the health maintenance of batteries. Moreover, Table XVIII shows the comparison results of SF-DWFT with other methods. It can be found that the RMSE values of SF-DWFT are 0.0045 and 0.0064 higher than that of MFSAN and MDAN. This is mainly because, in order to protect data privacy, some representation information of battery packs is lost during the global communication. Among all FL methods and only local-training methods, SF-DWFT obtains the smallest RMSE and MAE values, which indicates that SF-DWFT can achieve good SOH estimation while protecting the data privacy of local clients. This proves the effectiveness of SF-DWFT for SOH estimation of battery packs.

IV. CONCLUSION

In this study, a novel FL model, i.e., SF-DWFT is proposed for battery SOH estimation, which focuses on the issues of data privacy and distribution discrepancy between clients in practical applications. A dynamic weighted federated aggregation algorithm is proposed and the weight of each client is determined by their own contribution to the global model, which effectively reduces the negative impact of distribution discrepancy between clients on model estimation performance. It should be pointed out that a GMM model is used to characterize the extracted features during the local client training phase. This thereby transmits feature distribution between the client and the central server avoiding data privacy violations. Finally, several battery datasets were considered to validate the estimation performance of SF-DWFT. The RMSE of SF-DWFT on 18 650 Lithium-Ion battery datasets and NASA datasets is 0.0141 and 0.0623, respectively. Moreover, the comparison results demonstrate that SF-DWFT can achieve higher estimation accuracy while protecting data privacy, which is significant for engineering applications. In future study, further research on the heterogeneity issues in FL will be conducted.

REFERENCES

[1] Y. Qin, S. Adams, and C. Yuen, "A transfer learning-based state of charge estimation for lithium-ion battery at varying ambient temperatures," *IEEE Trans. Ind. Inform.*, vol. 17, no. 11, pp. 7304–7315, Nov. 2021.
[2] Z. Lyu, R. Gao, and L. Chen, "Li-ion battery state of health estimation and remaining useful life prediction through a model-data-fusion method," *IEEE Trans. Power Electron.*, vol. 36, no. 6, pp. 6228–6240, Jun. 2021.
[3] Z. Bao, J. Nie, H. Lin, J. Jiang, Z. He, and M. Gao, "A global–local context embedding learning based sequence-free framework for state of health estimation of lithium-ion battery," *Energy*, vol. 282, 2023, Art. no. 128306.
[4] B. Sah and P. Kumar, "Charger integrated coestimation of parameters and states of battery," *IEEE Trans. Power Electron.*, vol. 38, no. 7, pp. 7923–7932, Jul. 2023.
[5] Z. Song, X. Wu, X. Li, J. Sun, H. F. Hofmann, and J. Hou, "Current profile optimization for combined state of charge and state of health estimation of lithium-ion battery based on cramer-rao bound analysis," *IEEE Trans. Power Electron.*, vol. 34, no. 7, pp. 7067–7078, Jul. 2019.

[6] Y. Xu, C. Jiang, J. Peng, X.-L. Wu, L. Xiao, and X. Li, "Fault prognosis method for solid oxide fuel cells based on mechanism degradation process model and particle filtering," *IEEE Trans. Power Electron.*, vol. 38, no. 6, pp. 6831–6840, Jun. 2023.
[7] J. Peng, J. Meng, X. Du, L. Cai, and D. Store, "A fast impedance measurement method for lithium-ion battery using power spectrum property," *IEEE Trans. Ind. Inform.*, vol. 19, no. 7, pp. 8253–8261, Jul. 2023.
[8] J. Chen, Z. Zhou, Z. Zhou, X. Wang, and B. Liaw, "Impact of battery cell imbalance on electric vehicle range," *Green Energy Intell. Tansp.*, vol. 1, no. 3, 2022, Art. no. 100025.
[9] Y. Zhu, Z. Dong, Z. Cheng, X. Huang, Y. Dong, and Z. Zhang, "Neural network extended state-observer for energy system monitoring," *Energy*, vol. 263, 2023, Art. no. 125736.
[10] M. Jiao, D. Wang, and J. Qiu, "A GRU-RNN based momentum optimized algorithm for SOC estimation," *J. Power Sources*, vol. 459, 2020, Art. no. 228051.
[11] M. Catelani, L. Ciani, R. Fantacci, G. Patrizi, and B. Picano, "Remaining useful life estimation for prognostics of lithium-ion batteries based on recurrent neural network," *IEEE Trans. Instrum. Meas.*, vol. 70, 2021, Art. no. 3524611.
[12] L. Ren, J. Dong, X. Wang, Z. Meng, L. Zhao, and M. J. Deen, "A data-driven auto-CNN-LSTM prediction model for lithium-ion battery remaining useful life," *IEEE Trans. Ind. Inform.*, vol. 17, no. 5, pp. 3478–3487, May 2021.
[13] H. Sheng, B. Ray, S. Kayamboo, X. Xu, and S. Wang, "Battery health estimation based on multidomain transfer learning," *IEEE Trans. Power Electron.*, vol. 39, no. 4, pp. 4758–4770, Apr. 2024.
[14] M. Ghifary, W. B. Kleijn, and M. Zhang, "Domain adaptive neural networks for object recognition," in *Proc. Pacific Rim Int. Conf. Artif. Intell.*, 2014, pp. 898–904.
[15] U. Y. Ganin and V. Lempitsky, "Unsupervised domain adaptation by backpropagation," in *Proc. 32nd Int. Conf. Mach. Learn., PMLR*, 2015, vol. 37, pp. 1180–1189.
[16] J. Du, C. Zhang, S. Li, L. Zhang, and W. Zhang, "Two-stage aging trajectory prediction of LFP lithium-ion battery based on transfer learning with the cycle life prediction," *Green Energy Intell. Tansp.*, vol. 1, no. 1, 2022, Art. no. 100008.
[17] Z. Ye and J. Yu, "State-of-health estimation for lithium-ion batteries using domain adversarial transfer learning," *IEEE Trans. Power Electron.*, vol. 37, no. 3, pp. 3528–3543, Mar. 2022.
[18] L. Zhao, J. Li, Q. Li, and F. Li, "A federated learning framework for detecting false data injection attacks in solar farms," *IEEE Trans. Power Electron.*, vol. 37, no. 3, pp. 2496–2501, Mar. 2022.
[19] X. Chen, H. Wang, S. Lu, J. Xu, and R. Yan, "Remaining useful life prediction of turbofan engine using global health degradation representation in federated learning," *Rel. Eng. System Saf.*, vol. 239, 2023, Art. no. 109511.
[20] H. McMahan, E. Moore, D. Ramage, S. Hampson, and B. Arcas, "Communication-efficient learning of deep networks from decentralized data," in *Proc. Artif. Intell. Statist.*, 2017, pp. 1273–1282.
[21] Z. Liu, F. Wu, Y. Wang, M. Yang, and X. Pan, "FedCL: Federated contrastive learning for multi-center medical image classification," *Pattern Recognit.*, vol. 143, 2023, Art. no. 109739.
[22] I. Shiri et al., "Differential privacy preserved federated transfer learning for multi-institutional 68Ga-PET image artefact detection and disentanglement," *Eur. J. Nucl. Med. Mol. Imag.*, vol. 51, pp. 40–53, 2023.
[23] K. Zhao, J. Hu, H. Shao, and J. Hu, "Federated multi-source domain adversarial adaptation framework for machinery fault diagnosis with data privacy," *Rel. Eng. System Saf.*, vol. 236, 2023, Art. no. 109246.
[24] L. Guo, Y. Yu, M. Qian, R. Zhang, H. Gao, and Z. Cheng, "FedRUL: A new federated learning method for edge-cloud collaboration based remaining useful life prediction of machines," *IEEE/ASME Trans. Mechatron.*, vol. 28, no. 1, pp. 350–359, Feb. 2023.
[25] Y. Qing, J. Zhou, H. Pu, X. Zhang, and Y. Mao, "Dynamic weighted federated remaining useful life prediction approach for rotating machinery," *Mech. Syst. Signal Process.*, vol. 202, 2023, Art. no. 110688.
[26] H. Feng et al., "Kd3a: Unsupervised multi-source decentralized domain adaptation via knowledge distillation," in *Proc. Int. Conf. Mach. Learn.*, 2021, vol. 139, pp. 3274–3283.
[27] Z. Yang, D. Yang, C. Dyer, X. He, A. Smola, and E. Hovy, "Hierarchical attention networks for document classification," in *Proc. Conf. North Amer. Chapter Assoc. Comput. Linguistics, Hum. Lang. Technol.*, 2016, pp. 1480–1489.
[28] Q. Feng, C. Guo, F. Benitez-Quiroz, and A. M. Martinez, "When do gans replicate? On the choice of dataset size," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2021, pp. 6681–6690.
[29] N. Dinsdale et al., "Fedharmony: Unlearning scanner bias with distributed data," in *Proc. Int. Conf. Med. Image Comput. Comput. Assist. Intervention*, 2022, vol. 13438, pp. 695–704.
[30] A. Dempster, N. Laird, and D. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *J. Roy. Stat. Soc., Ser. B.*, vol. 39, no. 1, pp. 1–38, 1977.
[31] S. M. Ali and S. D. Slivery, "A general class of coefficients of divergence of one distribution from another," *J. Roy. Stat. Soc., Ser. B.*, vol. 28, no. 1, pp. 131–142, 1966.
[32] N. K Dinsdale, M. Jenkinson, and A. I. Namburete, "SFHarmony: Source free domain adaptation for distributed neuroimaging analysis," in *Proc. IEEE Int. Conf. Comput. Vis.*, Mar. 2023, pp. 11460–11471.
[33] Y. Zhu, F. Zhuang, and D. Wang, "Aligning domain-specific distribution and classifier for cross-domain classification from multiple sources," in *Proc. AAAI Conf. Artif. Intell.*, 2019, pp. 5989–5996.
[34] Y. Ding, P. Ding, X. Zhao, Y. Cao, and M. Jia, "Transfer learning for remaining useful life prediction across operating conditions based on multisource domain adaptation," *IEEE/ASME Trans. Mechatron.*, vol. 27, no. 5, pp. 4143–4152, Oct. 2022.
[35] R. Zhang, Q. Xu, J. Yao, Y. Zhang, Q. Tian, and Y. Wang, "Federated domain generalization with generalization adjustment," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2023, pp. 3954–3963.
[36] Y. Guo, K. Guo, X. Cao, T. Wu, and Y. Chang, "Out-of-distribution generalization of federated learning via implicit invariant relationships," in *Proc. Int. Conf. Mach. Learn.*, 2023, pp. 11905–11933.
[37] Z. Wang, M. Song, Z. Zhang, Y. Song, Q. Wang, and H. Qi, "Beyond inferring class representatives: User-level privacy leakage from federated learning," in *Proc. IEEE Conf. Comput. Commun.*, 2019, pp. 2512–2520.
[38] B. Saha and K. Goebel, "Battery data set," NASA Ames Prognostics Data Repository, 2007. [Online]. Available: http://ti.arc.nasa.gov/tech/dash/pcoe/prognostic-data-repository/
[39] B. Bole, C. Kulkarni, and M. Daigle, "Adaptation of an electrochemistry-based li-ion battery model to account for deterioration observed under randomized use," in *Proc. Annu. Conf. PHM Soc.*, 2014, pp. 502–510.

Tengfei Han received the B.Eng. degree in industrial engineering from the School of Mechanical and Electrical Engineering, Soochow University, Suzhou, China, in 2020, and the M.Eng. degree in management science and engineering from the School of Economics and Management, Nanjing Forestry University, Nanjing, China, in 2023. He is currently working toward the Ph.D. degree in industrial engineering with the School of Mechanical Engineering, Tongji University, Shanghai, China.

His current research interests include battery SOH estimation and deep learning.

Shang Yue received the B.S. degree in computer science from Huazhong University of Science and Technology, Wuhan, China, in 2003.

In 2003, he joined Tencent Holding Limited, Shenzhen, China, where he is currently an R&D Director in Tencent IDC. His current research interests include IDC system modeling and simulation, production scheduling, fault diagnosis, and machine learning.

Pu Yang received the B.Eng. and M.Sc. degrees in power system & automation from Huazhong University of Science and Technology, Wuhan, China, in 2000 and 2003, respectively.

He worked for Emerson Network Power China from 2003 to 2015, as Chief Architect of data center infrastructure management. He started up Huayuan IoT company for industry 4.0 in 2015, and joined Tencent in 2022, where he is the Tech Lead of IDC platform R&D Center. His current research interests include AI-driven PUE optimization, battery predictive maintenance, system dynamic reliability, and smart sensor and network.

Ruixu Zhou received the B.Eng. degree in electrical engineering from Wuhan University, Wuhan, China, in 2013, and the M.Sc. and Ph.D. degrees in electrical engineering from Tsinghua University, Beijing, China, in 2016 and 2022, respectively.

He joined Tencent in 2022, where he is currently a Researcher with Platform R&D Center of Tencent IDC. His research interests include pattern recognition, modern signal analysis and processing, data mining in power system, artificial intelligence, power system reliability, etc.

Jianbo Yu (Member, IEEE) received the B.Eng. degree from the Department of Industrial Engineering, Zhejiang University of Technology, Zhejiang, China, in 2002, the M.Eng. degree from the Department of Mechanical Automation Engineering, Shanghai University, Shanghai, China, 2005, and the Ph.D. degree from the Department of Industrial Engineering, Shanghai Jiaotong University, Shanghai, China, 2009.

From 2009 to 2013, he was an Associate Professor with Shanghai University, Shanghai, China. Since 2016, he worked as a Professor with the School of Mechanical Engineering, Tongji University, Shanghai, China. His research interests include machine maintenance and machine learning.

Dr. Yu is an Associate Editor of IEEE TRANSACTIONS ON INSTRUMENTATION MEASUREMENT and an Editorial Board Member of *Sensors*, and *Industrial Engineering and Management*.