Real-Time Corporate Carbon Footprint Estimation Methodology Based on Appliance Identification

Guolong Liu, *Member, IEEE*, Jinjie Liu, *Student Member, IEEE*, Junhua Zhao, *Senior Member, IEEE*, Jing Qiu, *Senior Member, IEEE*, Yiru Mao, Zhanxin Wu, and Fushuan Wen, *Fellow, IEEE*

***Abstract*—Achieving carbon neutrality is widely recognized as the key measure to mitigate climate change. As the basis for achieving carbon neutrality, corporate carbon footprint (CCF) estimation is mainly based on the disclosed information of corporates to roughly estimate the direct carbon emission, but the estimation may not be comprehensive, timely, and accurate. In this article, the CCF estimation problem is formulated and a novel estimation methodology is proposed for the first time to estimate the direct and indirect carbon emissions of factories in real time. An appliance identification method based on the multihead self-attention mechanism and gated recurrent unit is proposed to identify the device states, and then, calculate the corresponding direct carbon emission. The indirect carbon emission is derived from the electricity consumption of the factory and the marginal carbon emission factor of the connected bus. A dataset containing load and device state data from six different industries is released and used to verify the effectiveness of the proposed method. Experiments show that the proposed appliance identification method is significantly superior to the benchmarks in the literature, and the proposed method can achieve a comprehensive and accurate estimation of the minute-level CCF.**

***Index Terms*—Appliance identification, artificial intelligence, carbon neutrality, industrial appliance identification dataset (IAID), real-time corporate carbon footprint (CCF) estimation.**

Manuscript received 23 November 2021; accepted 15 February 2022. Date of publication 25 February 2022; date of current version 13 December 2022. This work was supported in part by the National Natural Science Foundation of China under Grant 71931003, Grant 72061147004, Grant 72171206, and Grant 72192805, and in part by the Shenzhen Institute of Artificial Intelligence and Robotics for Society. Paper no. TII-21-5173. *(Corresponding authors: Junhua Zhao; Jing Qiu.)*

Guolong Liu and Junhua Zhao are with the School of Science and Engineering, The Chinese University of Hong Kong, Shenzhen 518172, China, and also with the Shenzhen Institute of Artificial Intelligence and Robotics for Society, The Chinese University of Hong Kong, Shenzhen 518172, China (e-mail: guolongliu@link.cuhk.edu.cn; zhaojunhua@cuhk.edu.cn).

Jinjie Liu is with the School of Science and Engineering, The Chinese University of Hong Kong, Shenzhen 518172, China, and also with the Shenzhen Research Institute of Big Data, The Chinese University of Hong Kong, Shenzhen 518172, China (e-mail: jinjieliu@link.cuhk.edu.cn).

Jing Qiu is with the School of Electrical and Information Engineering, The University of Sydney, Sydney, NSW 2006, Australia (e-mail: jeremy.qiu@sydney.edu.au).

Yiru Mao and Zhanxin Wu are with the School of Data Science, The Chinese University of Hong Kong, Shenzhen 518172, China (e-mail: yirumao@link.cuhk.edu.cn; zhanxinwu@cuhk.edu.cn).

Fushuan Wen is with the School of Electrical Engineering, Zhejiang University, Hangzhou 310027, China (e-mail: fushuan.wen@gmail.com).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TII.2022.3154467.

Digital Object Identifier 10.1109/TII.2022.3154467

NOMENCLATURE

γ, W, bs	Attention weight, network weight, bias.
$\lfloor \cdot \rfloor$	Operation to obtain integer part of numbers.
ϕ	Number of devices in factory F.
φ	Number of factories or households.
$\vartheta_{\text{ref}}, \vartheta$	Slack bus voltage phase angle, voltage phase angle.
b	Number of carbon footprint estimation time intervals.
B, DE, P_G	Bus admittance matrix, demand, power generation.
C_G	Carbon emission of generators.
C_{direct}	Direct carbon emission.
C_{indirect}	Indirect carbon emission.
C_{total}	Total carbon footprint.
CI, NS	Number of correctly identified subsequences, number of total subsequences.
$\text{CO}_i(P_{G_i})$	Cost function of generating unit G_i.
ECF, ef, θ	Total carbon emission of the power system, carbon emission factor, increased output.
F, D	Factory that needs to estimate its carbon footprint, devices in factory F.
G, ℓ	Generating unit, number of generating unit.
i, j, t, z, τ	Indices.
l, U	Length of subsequence, embedded matrix.
$\text{LN}_{ij}^{\max}$	Maximum power flow on transmission line ij.
n, h, m	Number of GRU layers, hidden states of GRU layers, number of scale dot-product attention blocks.
N_{d_i}	Number of states of device i.
num_h	Number of hidden states in each GRU layer.
Q, K, V	Attention matrices.
RE_{ij}	Reactance of line ij.
RV, EV	Real direct carbon emission, estimated direct carbon emission.
S, α, β	Device states, device direct carbon emission intensity, MCEF of the connected bus.

T	Time period for carbon footprint estimation.
WP	Proportion of subsequences.
$X, \underline{X}, \tilde{\underline{X}}$	Power sequence, power sequence with integer values, power subsequences set.
Y, y_t	Set of device states, state vector.
Z, x, μ, σ	Z-score, element of data, mean, standard deviation.

I. Introduction

IN RECENT years, carbon dioxide-based greenhouse gas (GHG) emissions have led to rising global temperatures, mitigating the resulting climate change has become a common goal worldwide. Considering this background, the Paris Agreement was signed by more than 180 members of the United Nations in 2016 to reduce carbon emissions, with the goal of limiting the global average temperature increase to 1.5 ° C above preindustrial levels. Besides, as the main measure to mitigate climate change, carbon neutrality has been pledged by more than 110 countries, including the United Kingdom, China, Germany, Japan, and Canada. According to the Emission Gap Report 2019 released by the United Nations Environment Programme (UNEP), even if all the unconditional commitments in the Paris Agreement are fulfilled, the global temperature may still rise by 3.2 ° C, which will lead to wider and more damaging climate impacts [1]. The overall global emission reduction efforts must be at least five times the current level to achieve the carbon emission reduction required by the 1.5 ° C target in the next ten years. Therefore, reducing carbon emissions is a very challenging and urgent task.

Carbon emission estimation is the basis for achieving environmental goals, such as climate change mitigation, carbon neutrality, and sustainable development society and can effectively promote the transition of the low-carbon economy. Since industry is the main source of carbon emissions, it is very important to accurately estimate carbon emissions of factories. The CCF estimation can obtain specific emissions to clearly understand the emissions of corporates, thereby providing support for emission reduction. At present, most studies related to the CCF estimation only consider the direct carbon emission in the production process of corporates but not the indirect carbon emission caused by electricity consumption. Besides, the current studies are very rough, and the carbon emission is mainly estimated from corporates' annual reports and other disclosed information [2]. The time granularity is too rough and the accuracy is low, which cannot meet the real-time and accurate carbon emission reduction requirements.

Therefore, real-time CCF estimation is particularly important and valuable. First, it can improve the credibility of carbon emission disclosure information and at the same time enhance corporates' sense of social responsibility. Second, it can achieve real-time and accurate monitoring of carbon emissions to better formulate and adjust the related policies and measures to reduce carbon emissions. Third, it can also monitor and assist the operation and transition toward low-carbon economy more accurately and improve its vitality.

A real-time CCF estimation methodology is proposed in this article, which can obtain the corresponding device states by analyzing the load data of corporates, to achieve accurate and real-time carbon emission estimation for corporates. In the methodology, carbon emissions are divided into two parts: the direct emission emitted during the production process and the indirect emission caused by electricity consumption. For the direct carbon emission part, first, the operation states of its production devices are determined by analyzing the load data of the corporate through appliance identification. Then, the direct carbon emission is obtained through the operation states of devices and the corresponding device direct carbon emission intensities. For the indirect carbon emission, first, the electricity consumption for each estimated time interval is calculated from the load data. Then, the marginal carbon emission factor (MCEF) of the power grid bus where the corporate connected is obtained by solving the direct current optimal power flow (DC-OPF). Finally, the indirect carbon emission can be obtained through the obtained power consumption and MCEF.

This article has the following technical contributions to real-time industrial carbon emission estimation.

1) The problem of real-time CCF estimation is formulated and an estimation methodology is proposed for the first time, which can realize real time and accurate carbon emission estimation.
2) An appliance identification method is proposed based on the multihead self-attention mechanism and gated recurrent unit (GRU), and its performance is verified by experiments on two residential and one industrial load datasets. The proposed method can realize accurate monitoring of device states and high-precision estimation of the direct carbon emission.
3) A load dataset called industrial appliance identification dataset (IAID), including six industries including steel, metal, chemical, plastic, glass, and textile, is released for further study.

The rest of this article is organized as follows. Related works are presented in Section II. Exploratory data analysis is presented in Section III. The details of the proposed methodology are discussed in Section IV. The experiments and results are presented in Section V. Finally, Section VI concludes this article.

II. Related Works

A. CCFs Estimation

With the growing concern about climate change and the increasing urgency to achieve the global 1.5° C limitation target, carbon emission has become a field of interest to scholars and decision makers. As the first step of carbon management and carbon reduction, carbon accounting covers a series of activities related to carbon emissions, including measuring, calculating, verifying, reporting, etc. [3].

As the core tool of carbon accounting, carbon footprint (CF) has been proposed and widely used by all walks of life, even it has not had a unified definition. Generally, the CF can be defined as the measurement of the total amount of CO_2 emitted directly and indirectly from an entity [2], [4]. The "entity" can be a

process, product, individual, organization, company, corporate, government, etc. [5]. The estimation of CCF can be beneficial to both corporate and society, such as delivering data to carbon reporting, motivating the transition to green business, and so on [6]. The focus of this article is on the corporate-level carbon dioxide emissions estimation. Therefore, the CCF is mainly discussed.

For the CCF, some guidelines and standards are available, which can be found in [7]. Outlined by the Carbon Trust (2006) [7], the estimation process of the CCF can be summarized in the following five steps:

1) specifying objectives and drawing a process map;
2) defining boundaries and determining the calculation methods;
3) collecting data and fixing the emission factors and intensities;
4) calculation;
5) verification and disclosure.

Among them, as the key step, step 2 needs to specify the operational boundaries to identify carbon emission sources for which the corporate is responsible. According to World Resources Institute and World Business Council for Sustainable Development 2004 [8], the carbon emission can be divided into three scopes. Scope 1 includes the direct carbon emission from sources owned and controlled by the corporate itself. Scope 2 refers to the indirect carbon emission from the generation process of the purchased electricity. Scope 3 comprises other indirect emissions from other organizations, including upstream and downstream emissions [5]. Since the reporting of scope 3 is optional and is usually underestimated by companies [9], this article considers the emission from scopes 1 and 2.

The calculation method in step 4 determines the method of collecting data in step 3. In general, there are three kinds of CCF calculation methods: bottom-up, top-down, and hybrid methods [2], [6], [9]. The bottom-up methods also called the process-based analysis methods, which calculate the CF by analyzing each process in which emissions occur. They are more accurate for small firms, but they would be very complex and costly for large firms covering many variants of emissions. The top-down methods are also called the input–output analysis methods. They use the inputs–outputs matrix to obtain the CF, which are straightforward but rough estimation methods. To combine the advantages of the aforementioned two kinds of methods, some hybrid methods are proposed [10]. As the adopted national or regional emission factors are too rough and fixed over a long period, they cannot reflect the real-time emission changes in a small area, there are still some uncertainties and errors in the existing hybrid methods [11].

To solve the problems of the existing methods, some studies use different models and algorithms to estimate CCF. However, most of these studies are specifically designed for a single company or industry, which is not entirely applicable to other industries. In [12], a convolutional neural network bidirectional long short-term memory (CNN-BLSTM) model is proposed to estimate the CCF of steel plants, it is only for a single industry and the results of appliance identification is not accurate enough. Therefore, a CCF estimation methodology based on appliance identification and MCEF is proposed in this article, which is classified as a hybrid method. The proposed methodology can be applied to different industries, and the CCF of six different industries are estimated in experiments.

B. Appliance Load Monitoring

The goal of appliance load monitoring (ALM) is to perform detailed device state sensing and provide disaggregated information about the energy consumed. There are two main types of ALM methods: intrusive load monitoring (ILM) and nonintrusive load monitoring (NILM) [12]. The ILM requires a meter to be installed between each monitored device and the connected socket, while the NILM requires only a meter to be installed at the entry point. Therefore, the NILM has lower installation and maintenance costs and is a more practical solution.

There are some identification methods and applications on the NILM for residential appliances. These methods can be mainly divided into two categories: traditional machine-learning-based methods and deep-learning-based methods. For traditional machine learning-based methods, in [13], artificial neural networks, mainly multilayer perception is used to identify the load events in households. A support vector machine (SVM) and five-nearest neighbors are used to identify different appliances based on features extracted from active and reactive power and power factor [14]. A decision tree based method is designed to monitor household appliances by using the transient signal in [15]. These methods mainly rely on data with a relatively high-frequency sampling frequency, which requires high-frequency meters to collect data. In addition, limited by the learning capabilities of used models, these methods cannot handle large amounts of data with complex appliance characteristics. Conversely, due to the powerful feature extraction capabilities of deep neural networks, deep learning has achieved remarkable results in computer vision, speech recognition, and natural language processing. The applications of deep-learning-based methods in the NILM with the residential environment have been studied by some scholars. In [16], convolutional neural network (CNN) is used to do appliance classification by using voltage–current trajectories. Since the CNN cannot process context information that is an important characteristic in load data, there are some studies trying to solve this problem. A model named two-stream convolutional neural networks (TSCNN) based on two-stream CNN using temporal features [17] and methods based on long short-term memory (LSTM) are proposed to make up for this shortcoming [18], [19]. Most of them are based on high-frequency data, such as 44.1, 16, and 2 kHz. Implementing these methods requires huge bandwidth, massive storage space, and powerful data processing capabilities, which is costly and not feasible in actual situations. There are some studies at lower frequencies, such as 1 Hz. In [12], the CNN-BLSTM is proposed to identify device states in a steel plant based load data with a frequency of 1 Hz. Also, in [20], based on load data with a frequency of 1 Hz, the long short-term memory probabilistic neural network (LSTM-PNN) is proposed to identify the states of household appliances. Nevertheless, the existing ALM studies based on low-frequency

data are mainly limited to residential users or a single industry. Studies on ALM for multiple industries are still minimal.

The aggregated power sequence is a time series, in which the values not only have context information but also have varying degrees of correlation in the temporal dimension. In general, the closer the time is, the higher the correlation. Existing studies extract the context information by using CNN-based or recurrent neural networks (RNN) based methods to process all the values in the sequence uniformly, instead of focusing on the critical information that is more relevant to the target. The methods in these studies are unable to capture the relevance of load data in the temporal dimension, thus failing to achieve more accurate ALM for industrial users. In addition, residential users generally have fixed behavioral characteristics and usage patterns, while the usage patterns of industrial users are mainly affected by factors such as industry characteristics, production plans, and whether the factory uses shift work [21]. Therefore, the industrial NILM requires more specific information and is more difficult to realize than the residential NILM. Therefore, a novel method is needed that can accurately extract contextual information and critical information more related to the target in the time series, so as to achieve more effective appliance identification.

To cope with this challenge, an appliance identification method based on multihead self-attention mechanism and GRU is proposed to accurately identify the states of industrial devices in consideration of usage patterns. As a variant of the RNN, the GRU can capture and memorize long-term information in the time series compared to the standard RNN [22]. Compared with the LSTM, its gate structure is simpler, with similar performance and higher computational efficiency. The multihead self-attention mechanism focuses on and learns critical information that is more closely related to the target, to realize the learning of correlation information between values [23].

III. Exploratory Data Analysis

A. IAID Description

The IAID consists of load data and the states of devices contained in six industries: steel, metal, chemical, plastic, glass, and textile. The collection frequency is 1 Hz, and the time length is one month from June 1, 2019 to June 30, 2019. For the steel factory, device types include crusher, electric arc furnace, and filter. For the metal factory, device types include tilting press, electric heat treatment furnace, and filter. For the chemical factory, device types include reaction tank and filter. For the plastic factory, device types include shredder, injection molding machine, and filter. For the glass factory, device types include crusher, electric furnace, and filter. For the textile mill, device types include spinning machine, weaving machine, knitting machine, dyeing machine, and decorating machine. The IAID can be downloaded from.[1]

B. Exploratory Data Analysis

The power profiles of the six industries on the weekend and weekday are shown in Fig. 1. Different industries have different load characteristics, which means that their working hours and patterns are diverse. Besides, some industries have obvious periodic characteristics and similar power fluctuations on the weekday and weekend, which are beneficial for utilizing power profiles to identify devices states by appliance identification. Therefore, there is a need for an appliance identification method that can learn periodic and specific load characteristics from data and accurately infer the operating states of devices.

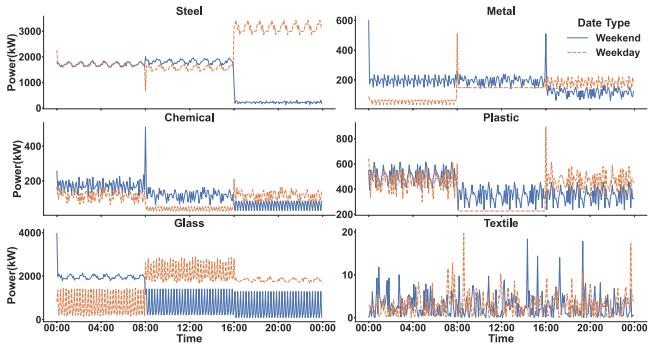

Fig. 1. Power profiles on the weekend and weekday in the IAID.

IV. Real-Time CCF Estimation Methodology

A. Problem Formulation

Real-time CCF estimation is to estimate the carbon emission of a factory in real time. The total CF of a factory in time period T with b estimation time intervals consists of two parts: direct emission and indirect emission. The direct emission part is directly emitted by the devices during the production process, and the indirect emission part is indirectly emitted by power generators during the device power consumption process. Hence, for a factory F with ϕ devices $D = \{d_1, d_2, \ldots, d_\phi\}$, the total CF C_{total} can be represented by $C_{\text{total}} = C_{\text{direct}} + C_{\text{indirect}}$, where C_{direct} is the direct carbon emission of the devices and C_{indirect} is the indirect carbon emission caused by electricity consumption. C_{direct} can be estimated by using

$$C_{\text{direct}} = \sum_{j=1}^{b} \sum_{i=1}^{\phi} \alpha_{S_{i,j}} \times \frac{T}{b} \tag{1}$$

where $\alpha_{S_{i,j}}$ is the direct carbon emission intensity of the device i in time interval j and $S_{i,j}$ is the state of the device i in time interval j. C_{indirect} can be estimated by using

$$C_{\text{indirect}} = \sum_{j=1}^{b} \beta_j \times E_j \times \frac{T}{b} \tag{2}$$

where β_j is the MCEF of the bus where the factory connected in the time interval j and E_j is the electricity consumption of the factory in the time interval j.

Therefore, the key to estimate the CF of a factory in real time is to obtain the device states S, corresponding device direct carbon emission intensities α, MCEF β, and electricity consumption E. Generally, the devices in a factory have their own fixed states, and the same type of factory has the same and similar devices.

[1][Online]. Available: https://www.zhaojunhua.org/dataset/IAID

Therefore, the device states S can be detected through appliance identification and device direct carbon emission intensities α can be collected without too much effort. β can be calculated by solving the DC-OPF and E can be directly read from the smart meter.

B. Optimal Power Flow

Power flow analysis is used to investigate the steady operation of a power system. For a given power system, with known complex power loads and some sets of specifications or restrictions on power generations and voltages, the calculation of the power flow problem is to obtain any unknown bus voltages and unspecified generation, and finally, obtain the complex power flow in the power system. Optimal power flow (OPF) functionally combines the power flow with economic dispatch to minimize cost with realistic equality and inequality constraints [24]. Under the premise of satisfying prevailing system level and unit constraints, the OPF dispatches power generation resources to supply system loads, thereby, achieving the least-cost operation of power systems. To improve the calculation efficiency of the OPF, as an approximation of alternating current optimal power flow (AC-OPF), the DC-OPF is used to obtain the optimal real power dispatch solution of the power system [25]. The DC-OPF problem for a power system is formulated as a mathematical optimization problem as follows:

$$\text{minimize} \quad \sum_{i=1}^{\ell} \text{CO}_i(P_{G_i}) \tag{3}$$

$$\text{s.t.} \quad B \cdot \vartheta = P_G - \text{DE} \tag{4}$$

$$\vartheta_{\text{ref}} = 0 \tag{5}$$

$$\left|\frac{1}{\text{RE}_{ij}}(\vartheta_i - \vartheta_j)\right| \le LN_{ij}^{\max} \text{ for all lines } ij \tag{6}$$

$$P_{G_i}^{\min} \le P_{G_i} \le PG_i{}^{\max} \tag{7}$$

where $\text{CO}_i(P_{G_i})$ is the cost function of the unit G_i, (4) represents the direct current power flow, (5) defines the slack bus voltage phase angle, (6) represents the transmission line power flow limits, and (7) represents the unit active power output limits [26].

C. Appliance Identification Method

The NILM can recognize the states of devices through the data collected by a single meter located at the entry point. An NILM method based on the multihead self-attention mechanism and GRU is proposed for appliance identification, and its structure is shown in Fig. 2.

The method consists of three parts: data embedding, relationship learning, and state identification. The aggregated power sequences of factories are taken as the inputs of the method. In the part of data embedding, the input power sequence X is first converted into $\underline{X}$ with integer values by using (8) to perform integer conversion, where $\min(X)$ is the minimum value in X and $\lfloor\cdot\rfloor$ represents the operation to obtain the integer part of values. Then, $\underline{X}$ is divided into the power subsequences set $\widetilde{\underline{X}}$, and the length of each subsequence is l. After that, each

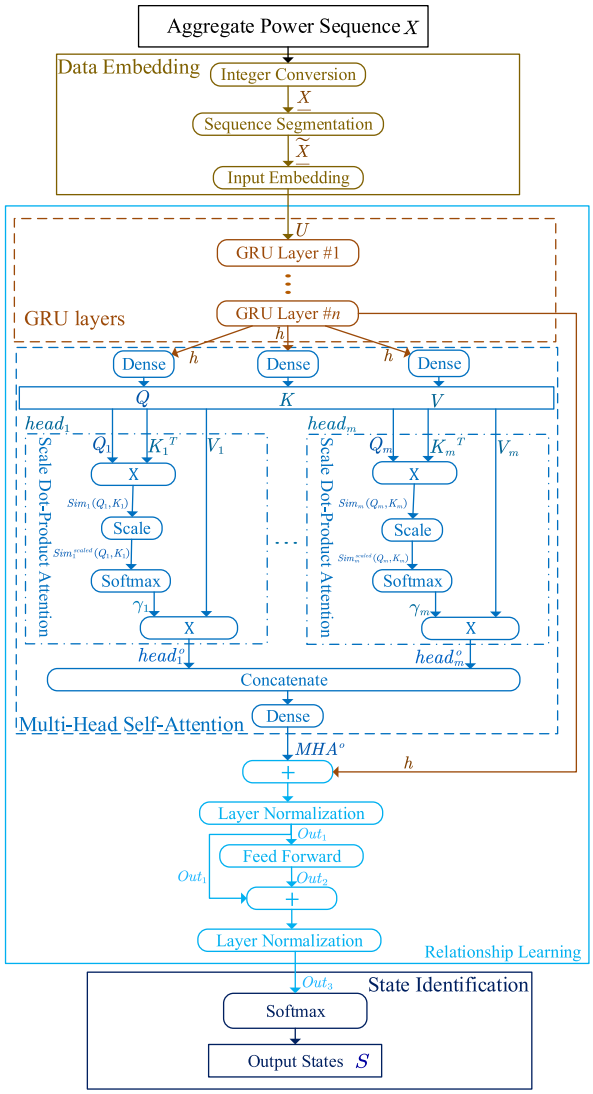

Fig. 2. Structure of the proposed appliance identification method.

value of the subsequence in $\widetilde{\underline{X}}$ is embedded into a g-dimensional hyperspace using the word embedding technique to obtain the vector representation. Then, the vector representations are used to replace the corresponding values. Therefore, the subsequence set $\widetilde{\underline{X}}$ is represented by the matrix U as

$$\underline{X} = \left\lfloor \frac{X}{\min(X)} \right\rfloor. \tag{8}$$

In the part of relationship learning, the embedded matrix U is taken as the input. First, there are n GRU layers to extract the context information in these vectors, and finally, output the information as hidden states h. Then, a multihead self-attention mechanism module (MHA) composed of m scale dot-product attention blocks is used to learn the internal connections between the vectors. In the MHA, h is used as the input, and then, mapped into Q, K, and V by using a dense layer, respectively. Q, K, and V are equally divided into m parts by column as the inputs of scale dot-product attention blocks. In zth block of MHA,

the similarity $\mathrm{Sim}_z(Q_z, K_z)$ between Q_z and K_z is calculated by (9). $\mathrm{Sim}_z(Q_z, K_z)$ is scaled to $\mathrm{Sim}_z{}^{\mathrm{scaled}}(Q_z, K_z)$ through (10) where $\dim_{kz}$ represents the number of columns in K_z. Afterward, $\mathrm{Sim}_z{}^{\mathrm{scaled}}(Q_z, K_z)$ is normalized by the softmax function to obtain the weight γ_z. Therefore, the output head_z^o of the zth head can be calculated by using (11). Then, the outputs of the m heads are concatenated and a dense layer is used to map the result of concatenation to the output MHA^o of MHA. MHA^o and h are summed as shown in (12), and a normalization layer is used to perform layer normalization on the summation result to get Out_1. Then, a fully connected feedforward network with two linear transformations is applied to Out_1 as shown in (13), where W_1 and W_2 represent the weights and bs_1 and bs_2 represent the biases. Out_1 and Out_2 are summed, and layer normalization is performed on the result to get the output Out_3 of relationship learning. In the part of state identification, Out_3 is used as the input of a dense layer with softmax as the activation function, so the corresponding probabilities of the device in different states under a given power subsequence are obtained. Finally, the states corresponding to the probabilities greater than the threshold are taken as the operation states of the device, thereby forming the device states set S.

$$\mathrm{Sim}_z(Q_z, K_z) = Q_z {K_z}^T \tag{9}$$

$$\mathrm{Sim}_z^{\mathrm{scaled}}(Q_z, K_z) = \frac{\mathrm{Sim}_z(Q_z, K_z)}{\sqrt{\dim_{kz}}} \tag{10}$$

$$\mathrm{head}_z^o = \gamma_z V_z \tag{11}$$

$$\mathrm{Out}_1 = \mathrm{MHA}^o + h \tag{12}$$

$$\mathrm{Out}_2 = f(\mathrm{Out}_1) = \max(0, \mathrm{Out}_1 W_1 + bs_1) W_2 + bs_2. \tag{13}$$

D. Marginal Carbon Emission Factor (MCEF)

The indirect carbon emission from electricity consumption accounted for a large proportion of industrial carbon emission, so it is important to measure the carbon emission from electricity consumption. There are mainly two metrics defined in existing studies, average carbon emission factor (ACEF) and MCEF [27]–[29]. The ACEF has been proved to be less accurate due to its neglect of marginal property and can even get the misestimate error up to 100%. By contrast, the MCEF can be specific in spatial and temporal dimensions so that it is more accurate and commonly used [30]. Although there is no unified formula definition, the meanings of the MCEF mentioned in studies like [27]–[30] are similar. The MCEF reflects the impact of bus load changes on carbon emissions, which is the total emission change caused by marginal generators. For the MCEF at bus BUS_k, it is the change in carbon emission of the power system caused by an infinitesimal load change at BUS_k. It is denoted by the following formula:

$$\mathrm{MCEF}_{\mathrm{BUS}_k} = \frac{\partial \mathrm{ECF}}{\partial \mathrm{LD}_{\mathrm{BUS}_k}} \tag{14}$$

where ECF is the total carbon emission of the power system and $\mathrm{LD}_{\mathrm{BUS}_k}$ is the load demand at BUS_k [28]. In this article, the MCEF of BUS_k at time t is calculated according to [31], as follows:

$$\mathrm{MCEF}_{\mathrm{BUS}_k,t} = \sum ef_{p,t} \times \theta_{p,t} \tag{15}$$

where $ef_{p,t}$ represents the carbon emission factor of the pth generator, and $\theta_{p,t}$ is the increased output of the pth generator due to a unit load increased at BUS_k.

There are two kinds of commonly used methods to estimate the MCEF. One is the statistical methods based on historical data, and the most effective method among them is the regression model. The other is the dispatch model based on mathematical definition and OPF [29]. The main limitation of statistical methods is that a large amount of high-resolution data it relies on may not be available [32]. Since the dispatch model does not have this limitation and is more practical, the DC-OPF is used to calculate the MCEF in this article.

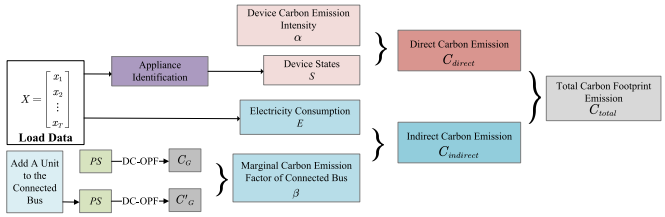

Fig. 3. Methodology of the real-time CCF estimation.

E. Real-Time CCF Estimation Methodology

The real-time CCF estimation methodology is shown in Fig. 3. For a given time period T with b intervals, first, the load data are used as the input of appliance identification and electricity consumption. By utilizing the proposed method of appliance identification, the device states S can be obtained. Together with the device direct carbon emission intensities α, the direct carbon emission C_{direct} can be estimated. For the indirect carbon emission C_{indirect} caused by electricity consumption, given a power system PS, the DC-OPF is solved to obtain the generator dispatch plan. Together with corresponding generator carbon emission factors, the carbon emission C_G of generators is obtained. Then, a unit is added to the load of the bus where the factory is connected, the DC-OPF is solved again to obtain the new dispatch plan. Based on the new dispatch plan and corresponding generator carbon emission factors, the new carbon emission C'_G of generators is obtained. Next, the MCEF β of the connected bus can be calculated by using the following formula:

$$\beta = C'_G - C_G. \tag{16}$$

The electricity consumption can be directly calculated from the load data. Therefore, the indirect carbon emission C_{indirect} can be obtained by using (2). Finally, the total carbon emission C_{total} of the factory is the sum of C_{direct} and C_{indirect}.

In practical applications, some historical load data are first used to train the appliance identification models on the cloud (data center), and these models will be stored on the cloud after the training process is completed. Then, the load data collected by smart meters installed at the industrial users is transmitted to the cloud. The load data are used to calculate electricity

consumption and taken as the input of appliance identification models to obtain device states in the corresponding time interval. Based on the given power system topology, the MCEFs of each bus are obtained by solving the DC-OPF twice at a regular time interval (the time interval in this article is 5 min) in the cloud. In experiments, appliance identification models are trained offline using graphic processing units (GPUs), and then, transmitted to the cloud. The server on the cloud is configured with a single core and eight Gigabytes of memory. The time consumption for data collection and transmission to the cloud is less than 0.5 s during the CCF estimation process. The time consumption of appliance identification models to obtain the device states is less than 1 s. Solving DC-OPF twice consumes 5 s. The entire estimation process takes about 6 s. In a short time interval, the MCEF of a bus generally does not change significantly. If it is assumed that the MCEF of a bus remains unchanged in a fixed time interval, the DC-OPF only needs to be solved twice in this time interval. The entire calculation time will be less than 2 s during the rest of this time interval. In addition, in the case of higher computing performance, the time consumption will be shorter. Therefore, the CCF estimation methodology can be regarded as a real-time estimation methodology of the CCF.

V. Experiments

A. Experimental Setup

Three datasets, including AMPds2 [33], U.K.-DALE [34], and IAID, are used to verify the performance of the proposed appliance identification method. AMPds2 and U.K.-DALE are residential load datasets, and IAID is an industrial load dataset. In AMPds2, aggregate- and appliance-level load data, including electrical characteristics such as voltage, current, and power, are collected at a frequency of 1 Hz. The temporal coverage is from 2012 to 2014. In U.K.-DALE, the load data of five households is recorded. The temporal coverage is from August 2015 to April 2017. Since the benchmarks of appliance identification require voltage–current (V-I) trajectory or voltage, current, power, etc., only the data of houses 1, 2, and 5 with a sampling frequency of 1 Hz is used for experiments. In the IAID, the load and device state data of six factories are used for experiments. The factories are located at different buses of the IEEE 30-bus system for solving the DC-OPF.

In the original datasets, there are some problems with data quality, such as data loss, duplicate sampling values, and abnormal data. These problems will greatly affect the experimental results, the data cannot be used directly and need to be preprocessed. First, for data loss, mean imputation introduced in [35] is used to fill in the missing values. Then, the duplicate sample values are detected and deleted based on timestamps. Next, the Z-score is used to detect abnormal values and the calculation of Z_i is shown as follows:

$$Z_i = \frac{x_i - \mu}{\sigma} \tag{17}$$

where x_i, μ, and σ represents the ith element, mean, and standard deviation of X, respectively. If the score Z_i is greater than the threshold Z_{thr}, x_i is determined as an outlier and it will be replaced by the mean of two adjacent elements, where Z_{thr} is set to three. In the experiments, each device has a corresponding multiple-classification or binary-classification identification model to identify its states. The part of appliance identification consists of these models. As a prerequisite for supervised learning, these identification models require the given input $X = \{x_1, x_2, \ldots, x_t, \ldots, x_T\}$ and the label $Y = \{y_1, y_2, \ldots, y_t, \ldots, y_T\}$, where Y is the set of device states. The N_{d_i} states of the device d_i are encoded as $0, 1, \ldots, N_{d_i} - 1$. The state vector $y_t = (y_t(1), y_t(2), \ldots, y_t(i), \ldots, y_t(\phi))$ is composed of the states of ϕ devices at time t, where $y_t(i)$ is the encoded state of ith device at time t. As a time series, the load data are first converted into integers, and then, divided into the set of subsequences $\underline{\widetilde{X}}$ each with a time length of 5 min. For device states set Y, in U.K.-DALE and AMPds2, the states of devices are divided into two states: ON and OFF. If the power consumption of the device in the 5-min interval is 0, then its state is OFF; otherwise, its state is ON. In the IAID, the operation state of each device with the longest duration within 5 min corresponding to X is regarded as the state within the time interval. That is, each subsequence in X has a corresponding encoded device state vector of the length ϕ as its label. Since voltage and current are not collected individually in the IAID and cannot meet the requirements of all appliance identification benchmarks, U.K.-DALE and AMPds2 are used for performance comparison of appliance identification methods. The IAID is used to verify the performance of the proposed appliance identification method and CCF estimation methodology. In particular, unlike some appliance identification benchmarks that require multiple electrical characteristics, such as voltage and current, the proposed appliance identification method only uses power consumption data as the model input. After the aforementioned processing steps, 75% of the data are used as the training set to train the appliance identification models, 5% of the data are used as the validation set to verify the models, and the remaining 20% of the data are used as the test set to evaluate the performance of the trained models.

It is generally recognized that deep neural networks have powerful information extraction and learning capabilities, but the network structure, number of layers, hidden units, and other parameters are hard to determine. For the task in this article, the training time of the deep neural network in appliance identification needs to be as short as possible while ensuring accuracy. Therefore, there is a tradeoff between the training time and accuracy. With the tradeoff as the goal, the IAID is used to determine the optimal parameters and its validation set is used to guide the selection of the optimal network structure and parameters through experiments. The number of GRU layers, the number of hidden states in each GRU layer, and the number of scale dot-product attention blocks are considered at the same time for optimal parameter selection. The number of GRU layers is selected from 1–5 in order. The number of hidden states in each GRU layer is selected from 32, 64, 128, 256, and 512 in order. Since the number of scale dot-product attention blocks needs to be divisible by the number of hidden states in each GRU layer, it is selected from 2, 4, 8, and 16 in order. In the experiment, the corresponding values are sequentially taken from the number of GRU layers, the number of hidden states in each GRU layer,

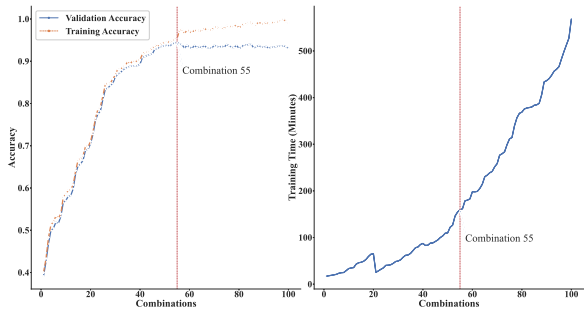

Fig. 4. Training time and accuracy of different combinations.

and the number of scale dot product attention blocks to form a new combination (n, num_h, m) where num_h represents the number of hidden states in each GRU layer. Therefore, there are 100 possible combinations. For example, combination 1 is $(1, 32, 2)$, combination 2 is $(1, 32, 4)$, combination 50 is $(3, 128, 4)$, and combination 100 is $(5, 512, 16)$. Fig. 4 shows the training time and accuracy of these combinations. As the model complexity increases, the training time of the model increases faster and faster. The validation accuracy first rises rapidly, and then, maintains a small range of fluctuations. The training accuracy first rises quickly with the validation accuracy; when the validation accuracy reaches the peak, it continues to increase to near one. Obviously, combination 55 is the optimal parameter combination.

After obtaining the optimal combination, the corresponding parameters can be determined. For input embedding, there are 256 dimensions to represent each unique power value. There are three GRU layers with 256 hidden states for learning the context information of the data, that is, $n = 3$. There are eight heads in MHA for learning the internal relationships of the data, that is, the value of m is eight. For the feedforward layer, there are 256 hidden nodes. The dropout technique is used in the proposed network to avoid overfitting, a problem that is particularly easy to occur in the training process of deep neural networks [36]. The dropout rate is set to 0.2, which means that 20% of the connections between and within hidden layers will be randomly disconnected. Adam is chosen as the optimization algorithm and categorical cross entropy is used as the loss function [37]. The minibatch size is set to 32 and the initial learning rate (LR) is set to 0.001. To make the proposed network learn the information in the data better, a dynamic LR step decay method is used in the training process. It uses the loss loss_{val} on the validation set as the monitoring indicator, and the LR will be adjusted to one-fifth of its own if the change of loss_{val} in ten consecutive epochs is less than 0.0001. Since the time interval of appliance identification is 5 min, to maintain consistency, the frequency of the MCEF calculation is once every 5 min. The experiments are implemented in Keras 2.4.3 with Tensorflow 2.3.0 as backend, and executed on a GPU cluster with Ubuntu Server 18.04 64-bits, two GTX 1080Ti, a 40-core CPU and 128-GB RAM.

B. Metrics and Benchmarks for Appliance Identification

Since appliance identification is realized by classification models, accuracy is mainly considered for evaluating the model performance. The state proportions of some devices in the data are not balanced, and the accuracy cannot fully reflect the actual performance of the models, so the weighted-average $F1$-score is also used as an evaluation metric. For a given device i with N_{d_i} unique states, the accuracy Accuracy_i is defined in (18), where $\text{CI}_{i,\tau}$ is the number of subsequences in the state τ that are correctly identified and $\text{NS}_{i,\tau}$ is the total number of subsequences in the state τ that need to be identified. For a factory or household with ϕ devices, the single user accuracy $\text{Accuracy}_{\text{user}}$ is defined in (19). For a dataset with φ factories or households, the dataset total accuracy $\text{Accuracy}_{\text{dataset}}$ is defined in (20). The precision $\text{Precision}_{i,\tau}$ and recall $\text{Recall}_{i,\tau}$ of the state τ in the device i are defined in (21) and (22), respectively, where $\text{NI}_{i,\tau}$ represents the number of subsequences identified as state τ. Therefore, the corresponding weighted-average $F1$-score $F1_i$ of the device i can be obtained by (23), where $\text{WP}_{i,\tau}$ is the proportion of subsequences in the state τ.

$$\text{Accuracy}_i = \frac{\sum_{\tau=1}^{N_{d_i}} \text{CI}_{i,\tau}}{\sum_{\tau=1}^{N_{d_i}} \text{NS}_{i,\tau}} \tag{18}$$

$$\text{Accuracy}_{user} = \frac{1}{\phi} \sum_{i=1}^{\phi} \text{Accuracy}_i \tag{19}$$

$$\text{Accuracy}_{dataset} = \frac{1}{\varphi} \sum_{i=1}^{\varphi} \text{Accuracy}_{\text{user}} \tag{20}$$

$$\text{Precision}_{i,\tau} = \frac{\text{CI}_{i,\tau}}{\text{NS}_{i,\tau}} \tag{21}$$

$$\text{Recall}_{i,\tau} = \text{CI}_{i,\tau} \times \text{NI}_{i,\tau} \tag{22}$$

$$F1_i = \sum_{\tau=1}^{N_{d_i}} 2 \times \frac{\text{Precision}_{i,\tau} \times \text{Recall}_{i,\tau}}{\text{Precision}_{i,\tau} + \text{Recall}_{i,\tau}} \times \text{WP}_{i,\tau}. \tag{23}$$

For appliance identification, there are some existing NILM methods, including the K-nearest neighbor (KNN), SVM, decision tree, multiple-layer feedforward artificial neural network with the backpropagation training algorithm (BP-ANN), CNN, CNN-BLSTM, LSTM-PNN, and TSCNN, they are used as benchmarks for performance comparison [12], [17], [20], [38]. The parameter settings of these benchmark methods are consistent with those in corresponding references.

C. Appliance Identification Results

The dataset total accuracy of different methods is shown in Fig. 5. It can be clearly seen that the proposed method outperforms the eight benchmarks, which can achieve more accurate device state monitoring. The proposed method realizes an accuracy of over 90% in U.K.-DALE and AMPds2. Fig. 6 shows the $F1$-score of different methods in U.K.-DALE and AMPds2. The proposed method has the highest mean, median, maximum, and minimum values. The $F1$-score distribution range is also smaller than the existing methods and there are no outliers, which indicates that the proposed method is superior to eight benchmarks in accuracy and stability. Since some of the input features required by LSTM-PNN and TSCNN are not collected in IAID, the remaining six benchmark methods are used for performance comparison with the proposed method.

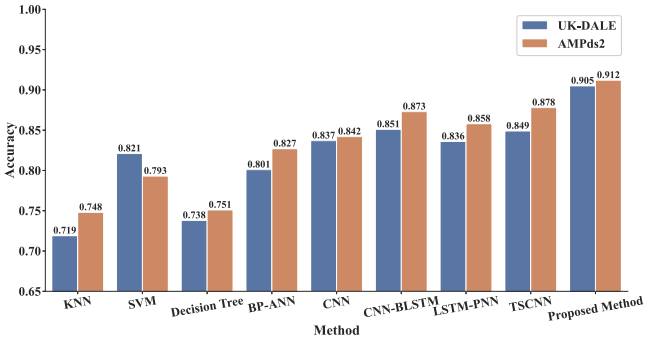

Fig. 5. Dataset total accuracy of different methods in U.K.-DALE and AMPds2.

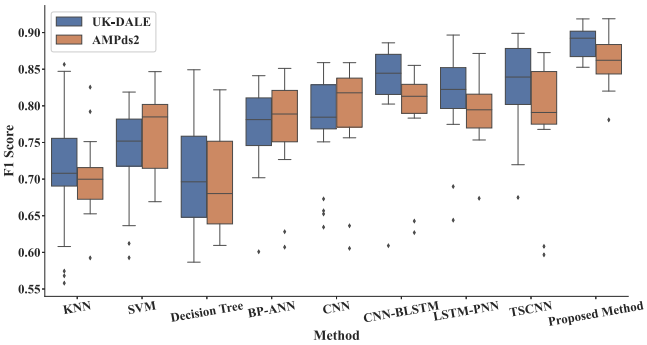

Fig. 6. Box plot of F1 score in U.K.-DALE and AMPds2.

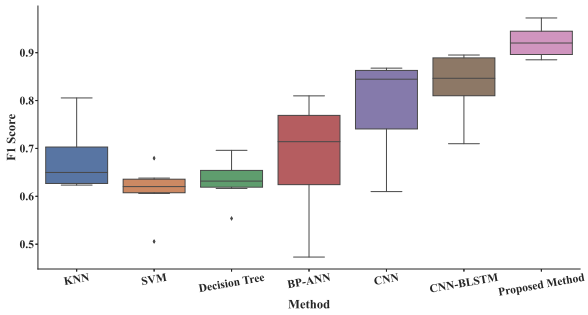

Fig. 7. Box plot of the F1 score in the IAID.

Fig. 7 shows the $F1$-score of different methods in the IAID. The proposed method is also significantly superior to the six benchmarks in industrial appliance identification and has a better stability and identification accuracy. Therefore, the proposed method can learn the critical characteristics in residential and industrial environments with different power consumption characteristics and patterns to identify device states accurately.

D. Real-Time CCF Estimation Results

After obtaining the states S of the devices in the six factories by using appliance identification models, the carbon emissions

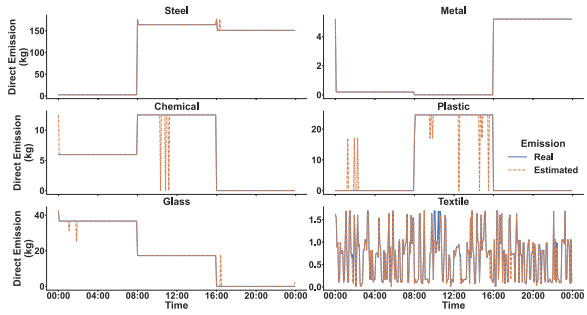

Fig. 8. Real and estimated direct carbon emission.

of the six factories are estimated separately. For the direct carbon emission of a given factory, it is determined by its device states S and corresponding device direct carbon emission intensity α, and α is given in the IAID. To quantify the accuracy of appliance identification on the estimation of the direct carbon emission, mean absolute percentage error (MAPE) is used for the performance evaluation, which is shown as follows:

$$\text{MAPE} = \frac{1}{\text{NF}} \sum_{nf=1}^{\text{NF}} \left| \frac{\text{RV}_{nf} - \text{EV}_{nf}}{\text{RV}_{nf}} \right| \times 100\% \qquad (24)$$

where NF is the number of factories that needed to be estimated, in the experiments, it is six. RV_{nf} and EV_{nf} represents the real and estimated direct carbon emission of the factory nf, respectively. The MAPE value between the real and the estimated direct carbon emission is 3.61%. The real and estimated direct carbon emission profiles are shown in Fig. 8, which indicate that direct carbon emissions over some time intervals are estimated incorrectly. The main reason is that the states of some devices that directly emit carbon dioxide in corresponding time intervals are misidentified. It is worth noting that the operating states of industrial devices are continuous in most cases, so postprocessing can be applied to appliance identification results for improving the identification accuracy and reducing the estimation error of direct carbon emission. In postprocessing, those states that are different from their neighboring states are first detected, and then, replaced by their neighboring states. After postprocessing, the MAPE value between the estimated value and the actual value drops to 1.92%, which means that the proposed method will only lead to a deviation of less than 2%. It can be seen from Fig. 8 that the emission profile of the textile mill is significantly different and fluctuating compared with several other factories. The textile mill has a vast difference between its production mode and other industries, its production devices, such as knitting machines, are mainly operated manually. The devices switch between different operation states frequently, which leads to the carbon emission profiles of this factory having large fluctuations.

The direct, indirect, and total carbon emission profiles of six factories are shown in Fig. 9. At around 7 A.M., the total carbon emission of most factories dropped sharply. Taking steel and glass as examples, their total emissions have dropped significantly, while direct emissions remain unchanged. The main

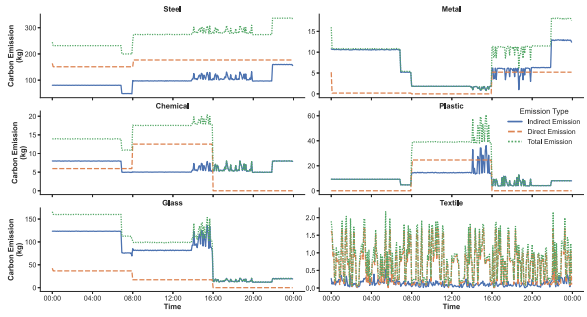

Fig. 9. Carbon emission results of factories.

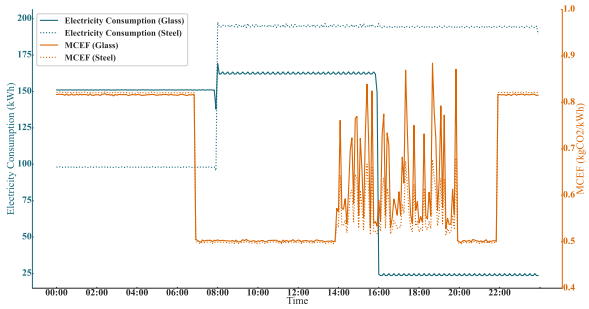

Fig. 10. Electricity consumption and MCEF of steel and glass.

reason is that the MCEF of the two factories suddenly decreased, leading to a rapid decline in indirect emissions. Fig. 10 indicates the electricity consumption and MCEF of the two factories. The MCEF is constantly changing over time. Even there are almost no changes in electricity consumption, the indirect carbon emission calculated based on it will also change. Besides, due to the higher proportion of renewable energy sources integrated into the power grid in the daytime, the corresponding MCEF is also relatively lower.

The experimental results show that the proposed methodology can realize real-time CCF estimation of factories. The real-time dynamic estimation of the direct carbon emission in a factory can be realized based on the device states identified by appliance identification. Increasing the sampling frequency of data and shortening the identification time interval can also achieve a higher frequency estimation of the direct carbon emission. By adding a unit load on the bus where the factory connected and solving DC-OPF, the MCEF of the factory can be obtained to realize real-time estimation of the indirect carbon emission. The frequency of the indirect carbon emission estimation can also be increased by increasing the frequency of solving the DC-OPF and collecting load data.

VI. Conclusion

In this article, a real-time CCF estimation methodology was proposed for estimating the direct and indirect carbon emissions of industrial users. A dataset namely IAID contained load, and device state data in six industries was also collected and released for further studies. In the proposed methodology, the direct carbon emission of a factory was determined by the device states obtained from appliance identification and the corresponding device direct emission intensities. An appliance identification method based on multihead self-attention mechanism and GRU was proposed to identify the states of industrial devices. The indirect carbon emission was determined by the MCEF of the bus where the factory connected and the electricity consumption within the corresponding time interval. Experimental results showed that the proposed appliance identification method was significantly superior to the benchmark methods in both residential and industrial datasets. The proposed method can achieve a direct carbon emission estimation difference of less than 2% after postprocessing. The effectiveness of the proposed methodology was verified by conducting experiments on factories in six industries, simulation results proved that the CCF can be accurately estimated in real time.

References

[1] A. Olhoff and J. Christensen, "*Emissions gap report 2019*," United Nations Environment Programme, Nairobi, Kenya, 2019.
[2] R. Waheed, S. Sarwar, and C. Wei, "The survey of economic growth, energy consumption and carbon emission," *Energy Rep.*, vol. 5, pp. 1103–1115, 2019.
[3] R. L. Burritt, S. Schaltegger, and D. Zvezdov, "Carbon management accounting: Explaining practice in leading german companies," *Australian Accounting Rev.*, vol. 21, no. 1, pp. 80–98, 2011.
[4] J. Yang *et al.*, "Key index framework for quantitative sustainability assessment of energy infrastructures in a smart city: An example of Western Sydney," *Energy Convers. Econ.*, vol. 1, no. 3, pp. 221–237, 2020.
[5] M. Csutora and G. Harangozo, "Twenty years of carbon accounting and auditing—A review and outlook," *Soc. Economy*, vol. 39, no. 4, pp. 459–480, 2017.
[6] C. Szigeti and G. Harangozó, "Online carbon calculators-corporate carbon footprint analysis in practice," in *Proc. MAC-MME*, 2016, pp. 262–272.
[7] A. McKinnon, M. Browne, A. Whiteing, and M. Piecyk, *Green Logistics: Improving the Environmental Sustainability of Logistics*. London, U.K.: Kogan, 2015.
[8] J. Ranganathan, L. Corbier, P. Bhatia, S. Schmitz, P. Gage, and K. Oren, "The greenhouse gas protocol: A corporate accounting and reporting standard (Revised Edition)," World Resour. Inst. World Bus. Council Sustain. Development, Washington, DC, USA, 2004.
[9] O. Durojaye, T. Laseinde, and I. Oluwafemi, "A descriptive review of carbon footprint," in *Proc. Int. Conf. Human Syst. Eng. Des., Future Trends Appl.*, 2019, pp. 960–968.
[10] B. Goldhammer, C. Busse, and T. Busch, "Estimating corporate carbon footprints with externally available data," *J. Ind. Ecol.*, vol. 21, no. 5, pp. 1165–1179, 2017.
[11] A. Navarro, R. Puig, and P. Fullana-I Palmer, "Product vs corporate carbon footprint: Some methodological issues. A case study and review on the wine sector," *Sci. Total Environ.*, vol. 581, pp. 722–733, 2017.
[12] G. Liu, J. Liu, J. Zhao, F. Wen, and Y. Xue, "A real-time estimation framework of carbon emissions in steel plants based on load identification," in *Proc. Int. Conf. Smart Grids Energy Syst.*, 2020, pp. 988–993.
[13] H.-H. Chang, K.-L. Lian, Y.-C. Su, and W.-J. Lee, "Power-spectrum-based wavelet transform for nonintrusive demand monitoring and load identification," *IEEE Trans. Ind. Appl.*, vol. 50, no. 3, pp. 2081–2089, May/Jun. 2014.
[14] M. Figueiredo, A. De Almeida, and B. Ribeiro, "Home electrical signal disaggregation for non-intrusive load monitoring (NILM) systems," *Neurocomputing*, vol. 96, pp. 66–73, 2012.
[15] T.-T.-H. Le *et al.*, "Non-intrusive load monitoring based on novel transient signal in household appliances with low sampling rate," *Energies*, vol. 11, no. 12, 2018, Art. no. 3409.
[16] L. De Baets, J. Ruyssinck, C. Develder, T. Dhaene, and D. Deschrijver, "Appliance classification using vi trajectories and convolutional neural networks," *Energy Build.*, vol. 158, pp. 32–36, 2018.
[17] J. Chen, X. Wang, X. Zhang, and W. Zhang, "Temporal and spectral feature learning with two-stream convolutional neural networks for appliance recognition in NILM," *IEEE Trans. Smart Grid*, vol. 13, no. 1, pp. 762–772, Jan. 2022.
[18] J.-G. Kim and B. Lee, "Appliance classification by power signal analysis based on multi-feature combination multi-layer LSTM," *Energies*, vol. 12, no. 14, 2019, Art. no. 2804.

[19] S. Heo *et al.*, "Toward load identification based on the Hilbert transform and sequence to sequence long short-term memory," *IEEE Trans. Smart Grid*, vol. 12, no. 4, pp. 3252–3264, Jul. 2021.
[20] Z. Zhou, Y. Xiang, H. Xu, Z. Yi, D. Shi, and Z. Wang, "A novel transfer learning-based intelligent nonintrusive load-monitoring with limited measurements," *IEEE Trans. Instrum. Meas.*, vol. 70, 2021, Art no. 2500508.
[21] E. Holmegaard and M. B. Kjaergaard, "NILM in an industrial setting: A load characterization and algorithm evaluation," in *Proc. IEEE Int. Conf. Smart Comput.*, 2016, pp. 1–8.
[22] A. F. M. Agarap, "A neural network architecture combining gated recurrent unit (GRU) and support vector machine (SVM) for intrusion detection in network traffic data," in *Proc. 10th Int. Conf. Mach. Learn. Comput.*, 2018, pp. 26–30.
[23] A. Vaswani *et al.*, "Attention is all you need," in *Proc. Adv. Neural Inf. Process. Syst.*, 2017, pp. 5998–6008.
[24] B. H. Chowdhury and S. Rahman, "A review of recent advances in economic dispatch," *IEEE Trans. Power Syst.*, vol. 5, no. 4, pp. 1248–1259, Nov. 1990.
[25] Y. Wang, L. Wu, and S. Wang, "A fully-decentralized consensus-based ADMM approach for DC-OPF with demand response," *IEEE Trans. Smart Grid*, vol. 8, no. 6, pp. 2637–2647, Nov. 2017.
[26] A. G. Bakirtzis and P. N. Biskas, "A decentralized solution to the DC-OPF of interconnected power systems," *IEEE Trans. Power Syst.*, vol. 18, no. 3, pp. 1007–1013, Aug. 2003.
[27] W. Schram, I. Lampropoulos, T. AlSkaif, and W. van Sark, "On the use of average versus marginal emission factors," in *Proc. 8th Int. Conf. Smart Cities Green ICT Syst.*, 2019, pp. 187–193.
[28] J. Hu, C. Jiang, H. Cong, and Y. He, "A low-carbon dispatch of power system incorporating active distribution networks based on locational marginal emission," *IEEJ Trans. Elect. Electron. Eng.*, vol. 13, no. 1, pp. 38–46, 2018.
[29] X. Yu *et al.*, "Research on locational marginal emission based on probabilistic power flow," in *Proc. IEEE Asia Power Energy Eng. Conf.*, 2019, pp. 285–290.
[30] C. N. Smith and E. Hittinger, "Using marginal emission factors to improve estimates of emission benefits from appliance efficiency upgrades," *Energy Efficiency*, vol. 12, no. 3, pp. 585–600, 2019.
[31] C. Wang, Y. Wang, C. J. Miller, and J. Lin, "Estimating hourly marginal emission in real time for PJM market area using a machine learning approach," in *Proc. IEEE Power Energy Soc. Gen. Meeting*, 2016, pp. 1–5.
[32] T. A. Deetjen and I. L. Azevedo, "Reduced-order dispatch model for simulating marginal emissions factors for the United States power sector," *Environ. Sci. Technol.*, vol. 53, no. 17, pp. 10506–10513, 2019.
[33] S. Makonin, B. Ellert, I. V. Bajić, and F. Popowich, "Electricity, water, and natural gas consumption of a residential house in Canada from 2012 to 2014," *Sci. Data*, vol. 3, no. 1, pp. 1–12, 2016.
[34] J. Kelly and W. Knottenbelt, "The UK-dale dataset, domestic appliance-level electricity demand and whole-house demand from five UK homes," *Sci. Data*, vol. 2, no. 1, pp. 1–14, 2015.
[35] A. R. T. Donders, G. J. van Der Heijden, T. Stijnen, and K. G. Moons, "A gentle introduction to imputation of missing values," *J. Clin. Epidemiol.*, vol. 59, no. 10, pp. 1087–1091, 2006.
[36] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A simple way to prevent neural networks from overfitting," *J. Mach. Learn. Res.*, vol. 15, no. 1, pp. 1929–1958, 2014.
[37] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *Proc. 2nd Int. Conf. Learn. Representations*, 2014.
[38] G. Liu, J. Gu, J. Zhao, F. Wen, and G. Liang, "Super resolution perception for smart meter data," *Inf. Sci.*, vol. 526, pp. 263–273, 2020.

Guolong Liu (Member, IEEE) received the Ph.D. degree in computer and information engineering from The Chinese University of Hong Kong, Shenzhen, China, in 2021.

He is currently a Research Fellow with the School of Science and Engineering, The Chinese University of Hong Kong. He is also a Research Fellow with the Shenzhen Institute of Artificial Intelligence and Robotics for Society, Shenzhen, China. His research interests include artificial intelligence, data analytics in smart grids, low-carbon transition, and load modeling.

Jinjie Liu (Student Member, IEEE) received the B.S. degree in electrical engineering and automation from North China Electric Power University, Beijing, China, in 2018. She is currently working toward the Ph.D. degree in computer and information engineering with the School of Science and Engineering, The Chinese University of Hong Kong, Shenzhen, China.

She is with the Shenzhen Research Institute of Big Data, Shenzhen, China. Her research interests include low-carbon electricity, smart grid, carbon emission estimation, and carbon market.

Junhua Zhao (Senior Member, IEEE) received the Ph.D. degree in electrical engineering from the University of Queensland, Brisbane, QLD, Australia, in 2007.

He was a Senior Lecturer with the University of Newcastle, Callaghan, NSW, Australia, and also with the Center for Intelligent Electricity Networks, University of Newcastle. He is currently an Associate Professor with the School of Science and Engineering, The Chinese University of Hong Kong, Shenzhen, China. He is also the Director with the Energy Market and Finance Lab, Shenzhen Finance Institute, The Chinese University of Hong Kong. His research interests include power system analysis and computation, smart grid, electricity market, data mining, and artificial intelligence.

Prof. Zhao was the recipient of the Young Scientist of the Future Award from ADC Forum in 2016, the IEEE PES General Meeting Best Paper Award in 2014, and the Science and Technology Progress Award (second prize) from the Hunan Provincial government, China, in 2013. He is the Editorial Board Member for the *IET Energy Conversion and Economics* and *Electric Power Components and Systems*.

Jing Qiu (Senior Member, IEEE) received the B.Eng. degree in control engineering from Shandong University, Jinan, China, in 2008, the M.Sc. degree in environmental policy and management, majoring in carbon financing in the power sector from The University of Manchester, Manchester, U.K., in 2010, and the Ph.D. degree in electrical engineering from The University of Newcastle, Callaghan, NSW, Australia, in 2014.

He is currently a Senior Lecturer in electrical engineering with the University of Sydney, Sydney, NSW, Australia. His research interests include power system operation and planning, energy economics, electricity markets, and risk management.

Dr. Qiu is the Editorial Board Member for the *IET Energy Conversion and Economics*.

Yiru Mao is currently working toward the B.S. degree majored in computer science and engineering with the Chinese University of Hong Kong, Shenzhen, China.

She is an undergraduate Research Assistant with Energy Internet Laboratory, The Chinese University of Hong Kong. Her research interests include super-resolution perception and short-term load forecasting.

Zhanxin Wu is currently working toward the B.S. degree majored in computer science and engineering with the Chinese University of Hong Kong, Shenzhen, China.

She is an undergraduate Research Assistant with the Energy Internet Laboratory, The Chinese University of Hong Kong. Her current research interests include load disaggregation and load forecasting.

Fushuan Wen (Fellow, IEEE) received the B.E. and M.E. degrees in electrical engineering and automation from Tianjin University, Tianjin, China, in 1985 and 1988, respectively, and the Ph.D. degree in electrical engineering from Zhejiang University, Hangzhou, China, in 1991.

He joined the faculty of Zhejiang University in 1991, and has been a Full Professor of Electrical Engineering since 1997. He had been a University Distinguished Professor of Electrical Engineering, the Deputy Dean with the School of Electrical Engineering, and the Director with the Institute of Power Economics and Electricity Markets, South China University of Technology, Guangzhou, China, from 2005 to 2009. He is currently a Professor of Electrical Engineering with the Department of Electrical Power Engineering and Mechatronics, Tallinn University of Technology, taking leave from Zhejiang University. His research interests include power industry restructuring, power system alarm processing, fault diagnosis and restoration strategies, as well as smart grids and electric vehicles.

Prof. Wen is the Editor-in-Chief for *IET Energy Conversion and Economics*, and a Deputy Editor-in-Chief for *Automation of Electric Power Systems*. He also serves as the Editor, Subject Editor, and Associate Editor for a few international journals.