A Distributed Computing Algorithm for Electricity Carbon Emission Flow and Carbon Emission Intensity

Xingping Wu, Wei Yang, Ning Zhang, Chunlei Zhou, Jinwei Song, and Chongqing Kang

Abstract—**The calculation of the indirect carbon emission is essential for power system policy making, carbon market development, and power grid planning. The embedded carbon emissions of the electricity system are commonly calculated by carbon emission flow theory. However, the calculation procedure is time-consuming, especially for a country with 500–1000 thousand nodes, making it challenging to obtain nationwide carbon emissions intensity precisely. Additionally, the calculation procedure requires to gather all the grid data with high classified levels from different power grid companies, which can prevent data sharing and cooperation among different companies. This paper proposes a distributed computing algorithm for indirect carbon emission that can reduce the time consumption and provide privacy protection. The core idea is to utilize the sparsity of the nodes' flow matrix of the nationwide grid to partition the computing procedure into parallel sub-procedures executed in multiple terminals. The flow and structure data of the regional grid are transformed irreversibly for privacy protection, when transmitted between terminals. A 1-master-and-N-slave layout is adopted to verify the method. This algorithm is suitable for large grid companies with headquarter and branches in provinces, such as the State Grid Corporation of China.**

Index Terms—**Carbon emission flow, cooperative computing, carbon emission intensity, matrix block partition, power flow tracing, parallel computing, privacy protection.**

I. Introduction

Following the Chinese government's commitment to achieve carbon peak by 2030 and carbon neutrality by 2060 (a.k.a. 'Double Carbon' or '3060' Target) announced in 2020, carbon emission reduction has become one of the top priorities of the government and enterprises [1]. According to statistical data, over 30% of China's carbon emissions stem from the electricity industry [2]. The obligations imposed urge electricity-related enterprises to focus on carbon emission monitoring, reporting, and verifying (MRV) [3]. Accurate calculation of direct and indirect carbon emissions resulting from electric power production and consumption is a critical aspect of this process.

Direct carbon emissions resulting from power production can be calculated by conducting long-term observations and statistics [4]–[6]. On the other hand, indirect carbon emissions resulting from power consumption require sophisticated methods and data from grid companies. As the energy hub, the electricity power grid plays a significant role in transmitting power from plants to customers, and thus, the indirect carbon emissions resulting from customer power consumption can be traced and calculated from flow and marketing data of grid companies [8]. In [9]–[12], carbon emission flow (CEF) is defined as the transmission of carbon emissions in the grid branch with power flow, and a base model is proposed for CEF calculation. The carbon emissions intensity (CEI), defined as unit carbon emissions tCO_2 per MWh of electricity, of loads connected with grid nodes can be calculated by this model.

CEI plays a pivotal role in calculating the carbon emissions of customers. According to the guidelines from the intergovernmental panel on climate change (IPCC) [5], [6] and Chinese Government [7], the indirect carbon emission of electricity power customers can be derived by multiplying the power amount with the CEI of the region from where the power is obtained. However, these CEI factors are often inaccurate because the region is too big, and the factors are updated at a low frequency. The GHG protocol encourages the use of more accurate calculation methods [8]. In [13], it proves that the ECI calculated from long-term static data has a difference of about 10%–30% compared to short time period calculations. Therefore, the need for high-frequency and small region ECI is critical in achieving more accuracy in the future. The base model in [12] and its improved version in [14] enable the calculation of CEIs for small control areas, including

Received: May 15, 2023
Accepted: December 20, 2023
Published Online: March 1, 2024
Jinwei Song (corresponding author) is with the Big Data Center, State Grid Corporation of China, Beijing 100052, China (e-mail: ysjk2000@foxmail.com).
DOI:10.23919/PCMP.2023.000379

transform stations, manual factors, or buildings. This method supports carbon emission and low-carbon optimal dispatch of distribution networks [15], [16]. A life cycle carbon emission calculation method is proposed in [17] that incorporates the CEF calculation. This method can be used to calculate the carbon emissions of a region in support of electricity markets [18], [19]. In addition, references [20]–[22] extend the method to calculate the carbon emissions of multiple energy systems. Also this method can be applied for low-carbon dispatch in power system [23]–[24].

The time resolution of the CEI generated by the model is dependent on the collection frequency of flow data. In theory, it can reach minute-level if the computing procedure runs efficiently. The high resolution and precision of the CEI enable low carbon demand response based on the CEI curve. According to [25], if 5% of loads in the Chinese power grid participate in the low carbon demand response, the carbon reduction achieved could be 20.30, 22.80, 29.40, and 40.10 million tons by the years 2025, 2030, 2035, and 2040, respectively.

Whilst the fundamental model for CEF and CEI calculation is largely completed, practical applications of the model still present some challenges that need to be addressed. When applying the model in practice, the model procedure may consume significant time, particularly when the number of grid nodes is large. In China, a region of 900 000 km² may contain between 50 000 to 100 000 nodes, and over the entire territory, there may be a distribution of 500 000 to 1 000 000 nodes. The traditional CEF and CEI calculation method involves building the CEF balance equation for every node and then solving these equations, with the most time-consuming aspect being the computation of the inverse matrix. The complexity of this step is related to the number of nodes, with the current method being at $O(N^3)$, or it can be reduced to $O(N)$ through the use of Gauss-Seidel iterations. However, the vast number of nodes requires a significant amount of time for the calculation process. Another way to decrease the time consumption is to split the data by regions and compute them separately. Nonetheless, the grids of different regions have connections and form a circular network. For instance, most provincial grids in China are connected with extra-high voltage lines, while some provinces' grids already influence each other. Therefore, the results of the separate calculation method may differ from the integrated method.

Another significant challenge that arises with the practical application of the CEF and CEI model is the sensitivity of the grid structure data, which is considered vital for a company. Gathering data from various nodes, branches (or lines), and electricity flow data across the entire country may require the collection of data from multiple companies, possibly leading to privacy breaches. This imposes significant challenges to data security and management, which can hinder data sharing among different companies, even within the branch companies of a group enterprise.

To address this concern, this paper proposes a parallel computing algorithm that reduces the calculation time while protecting data privacy during data transmission between computing terminals. The proposed algorithm yields results that are equal to traditional methods mathematically. Comparing with the existing researches, the contributions of this paper are threefold.

1) A distributed computing framework is proposed that enables the massive calculation procedure partitioning in multi-terminal. This framework significantly reduces the computation time consumption.

2) An algorithm is proposed to enable the privacy protection between terminals by transforming the flow and structure data of the regional grid in irreversible way.

3) The frameworks and the algorithms feasibility are proved with practical data from State Grid Corporation of China.

This paper introduces the traditional computing method for the base mode and outlines the problem of CEI calculation in Section II. The proposed distributed computing framework is introduced in Section III. The implementation algorithm is introduced in Section IV while Section V verifies its efficacy. Finally, Section VI provides the summary and conclusion of the paper.

II. Problem Formulation

In [26]–[28], the flow trace method, which is capable of calculating the impact of power sources and loads, is proposed, while the traditional CEF and CEI calculation method is a specific application of the flow trace method. In the flow trace method, the flow of electricity energy in one node is divided proportionally among the lines in which the energy flows out of the node. For instance, as illustrated in Fig. 1, the energy that flows from line 2 into line 3 can be computed as: $P_{2\to 3} = P_3 \times \dfrac{P_1}{P_1 + P_2} = 64$ MW.

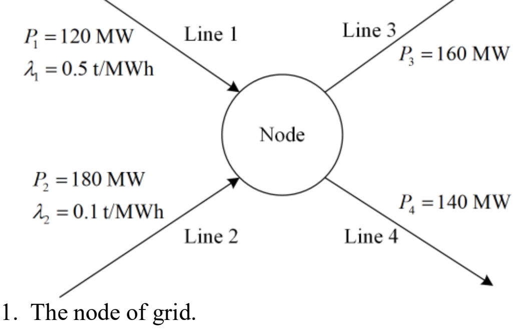

Fig. 1. The node of grid.

where P represents the power; λ is the CEI of the lines.

The carbon emission in the out branch of node is portioned in the same way. In Fig. 1, the carbon emission C in line 1 in t hour is $c_1 = P_1\lambda_1 t$, and for line 3, it is:

$$C_3 = P_3 \frac{P_1\lambda_1 + P_2\lambda_2}{P_1 + P_2} t = P_3\lambda_{\text{node}} t = P_3\lambda_3 t \tag{1}$$

where λ_{node} is considered as the CEI of the node. Equation (1) illustrates that the CEI of the out lines equal to the ECEI of the node. In practical, the node may be connected with generators and loads, e.g., as Fig. 2 shows, node 2, which may represent transformer substation of some solar power plants, connects with generators and loads.

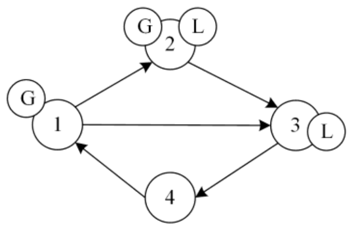

Fig. 2. A grid model with 4 nodes.

For every node in the grid, it is assumed that the carbon emission and electricity energy flowing in the node equal to the out amount and loads consumption without considering the loss, i.e.:

$$E_i + \sum_j P_{j,i} = \sum_k P_{i,k} + L_i \tag{2}$$

and

$$E_i\varepsilon_i + \sum_j P_{j,i}\lambda_j = \left(\sum_k P_{i,k} + L_i\right)\lambda_i \tag{3}$$

where E_i symbolizes the power produced by generators that are connected to node i, while L_i represents the power consumed by the loads; $P_{i,j}$ represents the power transmitted from i to j, while λ_i is the CEI of node i and ε_i denotes the carbon emission factor of the generators that are connected to node i; notably, ε_i is a static number that is unique to power stations within a specific region. For example, the china carbon market delineates a reference value for generator carbon emissions (refer to Table I), which can be used to calculate the ECEI of the Chinese grid. On the other hand, for regions outside of this scope, the carbon emission factor of generators must be obtained from power stations and follow the recommendations prescribed by the intergovernmental panel on climate change (IPCC).

TABLE I
CARBON EMISSION REFERENCE VALUE OF DIFFERENT TYPE OF GENERATORS

Type of generator	Carbon emission reference value (tCO_2/MWh)
Coal generator over 300 MW	0.8177
Coal generator not over 300 MW	0.8729
Other coal generator	0.9303
Gas power generator	0.3901

Equations (2) and (3) illustrate that the time variable t has been eliminated from both sides. From (2) and (3), the carbon emission balance equation of one node can be derived as:

$$E_i\varepsilon_i + \sum_j P_{j,i}\lambda_j = \left(E_i + \sum_j P_{j,i}\right)\lambda_i \tag{4}$$

Combining all node equations yield:

$$\begin{bmatrix} E_1 + \sum_j P_{j,1} & -P_{2,1} & \cdots & -P_{N,1} \\ -P_{1,2} & E_1 + \sum_j P_{j,2} & \cdots & -P_{N,2} \\ \vdots & \vdots & & \vdots \\ -P_{1,N} & -P_{2,N} & \cdots & E_N + \sum_j P_{j,N} \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_N \end{bmatrix} = \begin{bmatrix} E_1\varepsilon_1 \\ E_2\varepsilon_2 \\ \vdots \\ E_N\varepsilon_N \end{bmatrix} \tag{5}$$

where N is the count of nodes. Let

$$\boldsymbol{E} = \begin{bmatrix} E_1 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & E_N \end{bmatrix} \tag{6}$$

and

$$\boldsymbol{P} = \begin{bmatrix} 0 & P_{2,1} & \cdots & P_{N,1} \\ P_{1,2} & 0 & \cdots & P_{N,2} \\ \vdots & \vdots & & \vdots \\ P_{1,N} & P_{2,N} & \cdots & 0 \end{bmatrix} \tag{7}$$

where $\boldsymbol{E}$ and $\boldsymbol{P}$ represent the matrixes that are formed by generated power E_i and transmitted power $P_{i,j}$, equation (5) is reformed as:

$$(\boldsymbol{E} + diag(\boldsymbol{Pi}) - \boldsymbol{P})\boldsymbol{\lambda} = \boldsymbol{c} \tag{8}$$

where $\boldsymbol{c} = (C_1, C_2, \cdots, C_N)^{\mathrm{T}}$ is a vector formed by carbon emission; $C_i = E_i\varepsilon_i$ and $\boldsymbol{i} = (1, \cdots, 1)^{\mathrm{T}}$. Equation (8) can be changed into a compact form as:

$$\boldsymbol{M}\lambda = \boldsymbol{c} \tag{9}$$

The ECEIs in λ is calculated by:

$$\lambda = \boldsymbol{M}^{-1}\boldsymbol{c} \tag{10}$$

where $\boldsymbol{M}$ is the matrix formed by power flow information.

One crucial step involves obtaining the inverse matrix of $\boldsymbol{M}$, which has a time complexity of $O(N^3)$. To solve this challenging problem, various iteration methods such as Gauss-Seidel or Jacobi methods can be applied to (9). If λ fully converges in L iterations, the time complexity of the iteration method should be $O(LN^2)$. However, calculating all the ECEIs of the grid nodes across an entire country presents additional complications, given that the total number of nodes could range between 500 000–1000 000.

It is also important to acknowledge the potential implications of collecting node information, which could compromise high-level security classifications originated from diverse electricity grid companies or subsidiary corporations. This, in turn, can lead to difficulties in data sharing. Therefore, a distributed computing framework is proposed based on the partitioning of the matrix that is presented in (9).

III. The Distributed Computing Framework

Equation (9) can be partitioned by region. In order to better introduce the proposed distributed computing framework, several definitions are given as follows.

Def.1 Inner nodes: the grid nodes in a region without connection with other regions' nodes.

Def.2 Edge nodes: the grid nodes connect with other region's nodes.

Fig. 3 is a sample graph of two regions (provinces). The grids of the two regions are connected by lines and edge nodes.

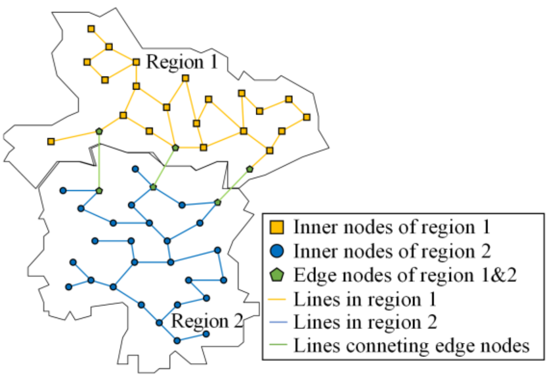

Fig. 3. Illustrative example for a two-region grid.

Matrix $\boldsymbol{M}$ in (9) can be partition in $(K+1)\times(K+1)$ blocks. Where K is the region number, e.g., $K=34$ for China. Thus, there is:

$$\begin{bmatrix} \boldsymbol{M}_1+\boldsymbol{I}_{\mathrm{e},1} & \cdots & \boldsymbol{0} & \boldsymbol{M}_{\mathrm{e},1} \\ \vdots & & \vdots & \vdots \\ \boldsymbol{0} & \cdots & \boldsymbol{M}_K+\boldsymbol{I}_{\mathrm{e},K} & \boldsymbol{M}_{\mathrm{e},K} \\ \boldsymbol{M}_{1,\mathrm{e}} & \cdots & \boldsymbol{M}_{K,\mathrm{e}} & \boldsymbol{M}_{\mathrm{e}} \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_K \\ \lambda_{\mathrm{e}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{c}_1 \\ \vdots \\ \boldsymbol{c}_K \\ \boldsymbol{c}_{\mathrm{e}} \end{bmatrix} \tag{11}$$

where $\boldsymbol{M}_k$ is formed by the flow information of the inner nodes in region k; and $\boldsymbol{I}_{\mathrm{e},k}=diag\{\boldsymbol{M}_{\mathrm{e},k}\boldsymbol{I}\}$; $\boldsymbol{M}_{\mathrm{e}}$ is the flow information between all edge nodes which form the edge nodes set; matrix $\boldsymbol{M}_{\mathrm{e},k}$ ($\boldsymbol{M}_{k,\mathrm{e}}$) represents the flow information from edge nodes set (inner node of region k) to inner node of region k (edge nodes set). Also, (11) can be written in a compact form as:

$$\begin{bmatrix} \boldsymbol{M}_{\mathrm{in}} & \boldsymbol{M}_{\mathrm{e,in}} \\ \boldsymbol{M}_{\mathrm{in,e}} & \boldsymbol{M}_{\mathrm{e}} \end{bmatrix} \begin{bmatrix} \lambda_{\mathrm{in}} \\ \lambda_{\mathrm{e}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{c}_{\mathrm{in}} \\ \boldsymbol{c}_{\mathrm{e}} \end{bmatrix} \tag{12}$$

where matrix $\boldsymbol{M}$ is a sparse matrix, while (11) and (12) portion and reform $\boldsymbol{M}$, remark as $\boldsymbol{M}_{\mathrm{R}}$, in order to use this character. Most blocks below and above the diagonal are 0 matrixes in $\boldsymbol{M}_{\mathrm{R}}$, except the last row and column blocks. Because all the inner nodes do not connect with other regions' inner nodes. The key step is to get the inverse matrix of $\boldsymbol{M}_{\mathrm{R}}$. According to the matrix block inverse theorem, there is:

$$\boldsymbol{M}_{\mathrm{R}}^{-1} = \begin{bmatrix} \boldsymbol{M}_{\mathrm{in}} & \boldsymbol{M}_{\mathrm{e,in}} \\ \boldsymbol{M}_{\mathrm{in,e}} & \boldsymbol{M}_{\mathrm{e}} \end{bmatrix}^{-1} = \begin{bmatrix} \boldsymbol{M}_{\mathrm{in}}^{-1} + \boldsymbol{M}_{\mathrm{in}}^{-1}\boldsymbol{M}_{\mathrm{e,in}}\boldsymbol{M}_{\varepsilon}\boldsymbol{M}_{\mathrm{in,e}}\boldsymbol{M}_{\mathrm{in}}^{-1} & -\boldsymbol{M}_{\mathrm{in}}^{-1}\boldsymbol{M}_{\mathrm{e,in}}\boldsymbol{M}_{\varepsilon} \\ -\boldsymbol{M}_{\varepsilon}\boldsymbol{M}_{\mathrm{in,e}}\boldsymbol{M}_{\mathrm{in}}^{-1} & \boldsymbol{M}_{\varepsilon} \end{bmatrix} \tag{13}$$

then

$$\boldsymbol{M}_{\varepsilon} = (\boldsymbol{M}_{\mathrm{e}} - \boldsymbol{M}_{\mathrm{in,e}}\boldsymbol{M}_{\mathrm{in}}^{-1}\boldsymbol{M}_{\mathrm{e,in}})^{-1} = (\boldsymbol{M}_{\mathrm{e}} - \sum_k \boldsymbol{M}_{i,\mathrm{e}}(\boldsymbol{M}_1+\boldsymbol{I}_{\mathrm{e},1})^{-1}\boldsymbol{M}_{\mathrm{e},i})^{-1} = (\boldsymbol{M}_{\mathrm{e}} - \sum_k \boldsymbol{R}_k)^{-1} \tag{14}$$

where $\boldsymbol{M}_{\varepsilon}$ is the parts of inverse matrix of $\boldsymbol{M}_{\mathrm{R}}$ that related to $\boldsymbol{M}_{\mathrm{e}}$, and $\boldsymbol{R}_k = \boldsymbol{M}_{k,\mathrm{e}}(\boldsymbol{M}_1+\boldsymbol{I}_{\mathrm{e},1})^{-1}\boldsymbol{M}_{\mathrm{e},i}$. From (11), there is

$$\boldsymbol{M}_{\mathrm{in}}^{-1} = \begin{bmatrix} (\boldsymbol{M}_1+\boldsymbol{I}_{\mathrm{e},1})^{-1} & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & (\boldsymbol{M}_K+\boldsymbol{I}_{\mathrm{e},K})^{-1} \end{bmatrix} \tag{15}$$

where $\boldsymbol{M}_{\mathrm{in}}$ is the matrix formed by all inner nodes.

Equation (15) uses the characters of block diagonal matrix to simplify the solving procedure. From (12) and (13), there is:

$$\lambda_{\mathrm{e}} = \boldsymbol{M}_{\varepsilon}\boldsymbol{c}_{\mathrm{e}} - \boldsymbol{M}_{\varepsilon}\boldsymbol{M}_{\mathrm{in,e}}\boldsymbol{M}_{\mathrm{in}}^{-1}\boldsymbol{c}_{\mathrm{in}} = \boldsymbol{M}_{\varepsilon}\boldsymbol{c}_{\mathrm{e}} - \boldsymbol{M}_{\varepsilon}\sum_k \boldsymbol{M}_{k,\mathrm{e}}(\boldsymbol{M}_k+\boldsymbol{I}_{\mathrm{e},k})^{-1}\boldsymbol{c}_k = \boldsymbol{M}_{\varepsilon}\boldsymbol{c}_{\mathrm{e}} - \boldsymbol{M}_{\varepsilon}\sum_k \boldsymbol{Q}_k \tag{16}$$

where $\boldsymbol{Q}_k = \boldsymbol{M}_{k,\mathrm{e}}(\boldsymbol{M}_k+\boldsymbol{I}_{\mathrm{e},k})^{-1}\boldsymbol{c}_k$. From (11), there is:

$(\boldsymbol{M}_i+\boldsymbol{I}_{\mathrm{e},i})\lambda_i + \boldsymbol{M}_{\mathrm{e,in}}\lambda_{\mathrm{e}} = \boldsymbol{c}_i$.

and then

$$\lambda_i = (\boldsymbol{M}_i+\boldsymbol{I}_{\mathrm{e},i})^{-1}\boldsymbol{c}_i - (\boldsymbol{M}_i+\boldsymbol{I}_{\mathrm{e},i})^{-1}\boldsymbol{M}_{\mathrm{e},i}\lambda_{\mathrm{e}} \tag{17}$$

In (14) and (16), matrix $\boldsymbol{O}_i = \boldsymbol{M}_{i,\mathrm{e}}(\boldsymbol{M}_i+\boldsymbol{I}_{\mathrm{e},i})^{-1}$ contains the flow information of region i, which can be obtained from the power dispatching department of region company. For a star mode structure company, e.g., State Grid Corporation of China, the power dispatching has different levels. The dispatching department in headquarter is in charge of the power transmission between regions (provinces). The power transmission in the region is arranged by province companies in regions (provinces). The flow information of one province is unknown from other province companies. For example, as displayed in Fig. 3, $\boldsymbol{M}_k$, $\boldsymbol{M}_{\mathrm{e},k}$, $\boldsymbol{M}_{k,\mathrm{e}}$, $\boldsymbol{I}_{\mathrm{e},k}$ and $\boldsymbol{c}_{\mathrm{e}}$ can be obtained from yellow square nodes in region A and blue circle nodes in region B by the province companies in the two regions. $\boldsymbol{M}_{\mathrm{e}}$ and $\boldsymbol{c}_{\mathrm{e}}$ can be obtained from the green rhombus nodes by the headquarter.

If every province company transmits matrix $\boldsymbol{R}_k$ and $\boldsymbol{Q}_k$ to the headquarter, the headquarter can calculate the ECEIs of edge nodes by (14) and (16). The headquarter then broadcasts ECEIs of edge nodes to province companies after calculation. The province companies can get the ECEIs of inner nodes by (17). Matrix $\boldsymbol{R}_k$ and $\boldsymbol{Q}_k$ are the transformations of flow information of a region (province), while the flow information has been compressed in $\boldsymbol{R}_k$ and $\boldsymbol{Q}_k$, so they cannot be restored to original flow information. Therefore, even the headquarter does not know the flow information of province company. The procedure of the information transmission is displayed in Fig. 4.

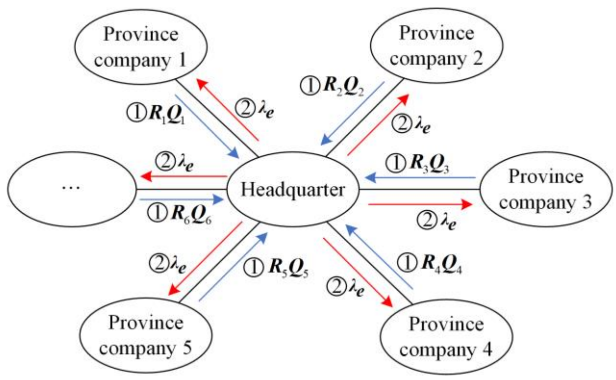

Fig. 4. Information transmission of headquarter and province companies in star structure organization.

IV. Algorithm Implementation

This section describes the implementation of the computing algorithm in a distributed computing framework, as depicted in Fig. 4. The framework adopts a 1-master-N-slave structure, consisting of a single headquarter computing terminal and K province companies. The algorithm's flow chart is illustrated in Fig. 5. The headquarter and province companies follow different procedures, and their cooperation is crucial to the success of the calculation.

The headquarter computing terminal focuses on the ECEI of edge nodes, which requires the results from province companies' terminals. The ECEI of edge nodes is broadcasted to the province companies' terminals, and the transmission of ECEI between computing terminals is reliant on the TCP/IP protocol. It is necessary to synchronize the terminals to ensure that the headquarter terminal receives all the results from the province terminals. Additionally, the flow information data must be in the same tidal current section in the headquarter and all province companies. Hence, the transmission frame includes a time segment, and the headquarter terminal checks the segment to ensure data consistency. When the terminals communicate, the time stamp which marks the

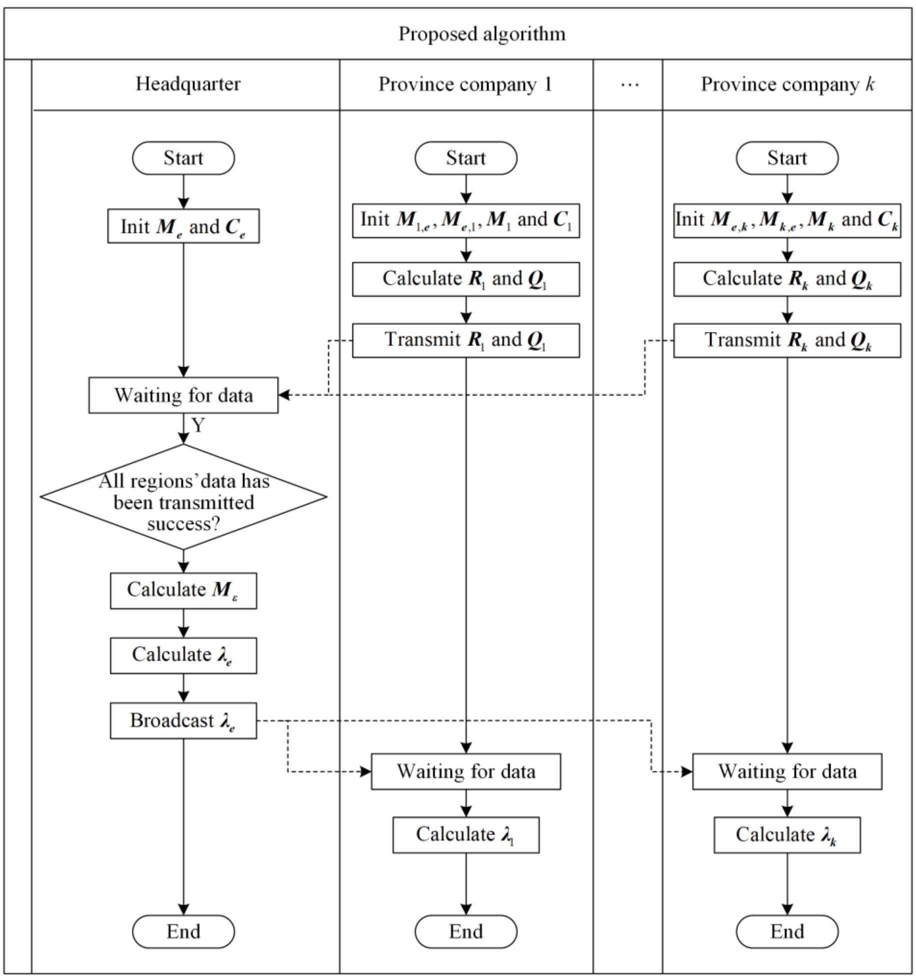

Fig. 5. Flow chart for the proposed method.

tidal current section time will also be transmitted in the frame. When the headquarter receives these frames from the province companies, the headquarter terminal compares the timestamps and send re-send commands to those province companies who have different timestamps to the headquarter's. For some network problems, the communications between some province companies and headquarter may not be available, though it is suggested to continue the calculation process and to obtain the results from the headquarter and those province companies without communication problems.

The proposed algorithm can distribute most of the calculation process to the terminals of the province companies. If the average number of every region (province) is n and there are l edge nodes, the time complex of the proposed algorithm will be $O(Kn^3+l^3)$, where K is the number of regions (provinces). Usually, the count of edge nodes is far less than the inner nodes. Therefore, comparing to the complexity of traditional method, $O((Kn+l)^3)$, the proposed algorithm significantly reduces the time complexity. If the k province companies run the process synchronously, the time cost will be reduced to $O(n^3+l^3)$. As this algorithm needs remote communication, the time consumption should include transmission delay. Nevertheless, the delay time is typically far less than the computing time.

V. Case Study

To verify the proposed algorithm, three datasets are adopted. Dataset-1 contains three provinces' flow data with 910 nodes in one period, which is used to prove that the results of the proposed algorithm are equal to the traditional method. Dataset-2 is used for testing the time complexity and contains 6300 stimulate data. Dataset-3 contains 2050 nodes and is divided into 28 part randomly, with every part presenting a province company that has 180 to 270 nodes. Dataset-3 is to test the adaptability of the proposed algorithm in the scenario that the 27 province companies are calculating and communicating with the headquarter.

In Dataset-1, the three provinces' flow data contain 910 nodes that cover voltage level from 220 kV to 1000 kV. The edge nodes of the three provinces are similar. The verification procedure uses the traditional method and the proposed method to calculate all λ_i of the three provinces. The results are then compared and if the difference of the results is lower than threshold T which is very small and is considered to the error from computer, it means that the two methods have the same results. The difference is defined as:

$$D=\frac{1}{N}\sum_{i=1}^{N}\left|\lambda_{1,i}-\lambda_{2,i}\right| \tag{18}$$

where $\lambda_{1,i}$ and $\lambda_{2,i}$ are the ECEIs of i calculated by the traditional method and the proposed method; N is the selected nodes number, which can be random, and $T=0.000\,01$.

The results of selected nodes from those of the 910 nodes in Table II show that the two methods have the same results, with zero difference between the two.

Three group simulation data are adopted in Fig. 6, which are divided into four parts, i.e. the three-region data with 2000 inner nodes per region and 300 edge nodes. The time consumptions of the proposed method are shown in Table III. As seen, the speed of the proposed method is much faster than the traditional one.

TABLE II
COMPARISON OF THE TWO METHODS FROM SELECTED NODES

Code of nodes	Results of traditional method (tCO_2/MWh)	Result of proposed method (tCO_2/MWh)
Node 1	0.523 95	0.523 95
Node 2	0.817 70	0.817 70
Node 3	0.690 63	0.690 63
Node 4	0.552 84	0.552 84
Node 5	0.137 59	0.137 59
Node 6	0.606 12	0.606 12
Node 7	0.425 75	0.425 75
Node 8	0.257 76	0.257 76
Node 9	0.609 54	0.609 54
Node 10	0.264 84	0.264 84
Node 11	0.325 40	0.325 40
Node 12	0.325 40	0.325 40

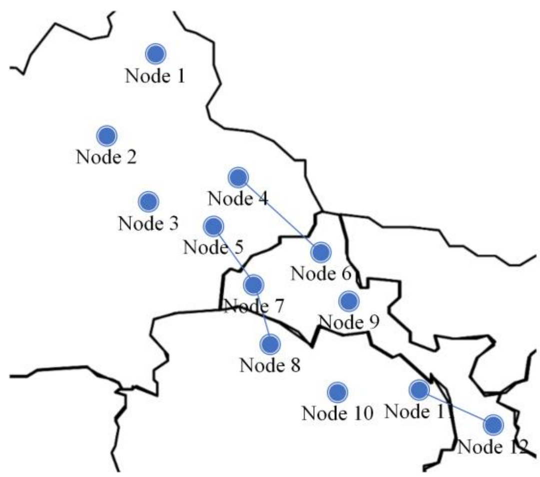

Fig. 6. The selected nodes in three provinces in Table II

TABLE III
TIME CONSUMPTION OF THE TRADITIONAL METHOD AND THE PROPOSED METHOD

Nodes	Traditional method (s)	Proposed method (s)		
		Head-quarter	Province companies	Total
1000 inner nodes × 3 region + 200 edge nodes	1.825	0.014	0.022	0.036
2000 inner nodes × 3 region + 300 edge nodes	8.623	0.018	0.259	0.277
3000 inner nodes × 3 region + 400 edge nodes	24.355	0.189	1.05	1.239

Table III illustrates that the proposed method significantly reduces the time consumption. In Table IV, the times of the proposed method are shown with headquarter's times and branch companies' times. All the branch companies' times are the same because the simulate data are generated with the same number of nodes. In practical calculation, the time consumption of branch companies' side is determined by the company that has the most nodes, i.e., the bottleneck of the consumption time is the maximum number of the inner nodes. Some calculation procedures of the headquarter and province companies can be run parallel. The total times in Table III are the suprema of the proposed algorithm. As seen, the suprema are far less than the time consumptions of the traditional method. The time complex of the traditional method and the proposed method are $O((Kn+l)^3)$ and $O(Kn^3+l^3)$, respectively.

TABLE IV
TIME CONSUMPTION OF THE PROPOSED METHOD WITH DATASET-3

Nodes	Traditional method (s)	Proposed method (s)		
		Headquarter	Province companies	Total
180-270 nodes × 27 region +50 edge nodes	8.510	0.012	0.013	0.025

Fig. 7 displays the trends of the two methods with the number of nodes increasing. It is clear that the run time of the proposed method increases more slowly than the traditional method.

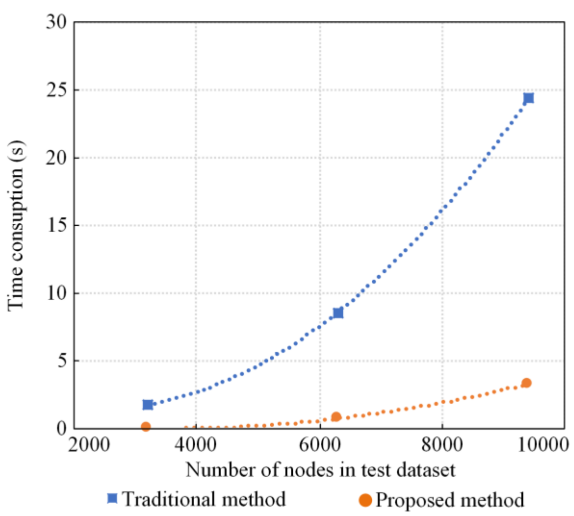

Fig. 7. The trends of time consumption of the traditional method and the proposed method.

When the data $\boldsymbol{R}_k$ and $\boldsymbol{Q}_k$ are transmitted to the headquarter, the flow information of one region has been compressed. For example, the province 1's inner nodes flow matrix $\boldsymbol{M}_k$ has 2000×2000 element, but $\boldsymbol{R}_k$ and $\boldsymbol{Q}_k$ have 300×300 after transformation. Thus, it is impossible to recovery the flow information of a region from $\boldsymbol{R}_k$ and $\boldsymbol{Q}_k$.

To validate that the proposed algorithm is able to adapt to the cooperative computing scenario in an electricity corporation, such as State Grid Corporation of China, 6063 500 kV and 220 kV nodes from 27 province companies are selected. The results of the proposed and traditional methods are the same. Because the single province company only selected about 180 to 270 nodes, the average time consumption of the province company is 0.013 s in Table IV. Thus, it is proved that the proposed algorithm can be applied in multi-branch corporation.

VI. CONCLUSION

In this paper, an algorithm is proposed for the calculation of both ECEI and ECEF by capitalizing on the sparse nature of the flow matrix to reduce overall computation time consumption. The algorithm supports multi-terminal calculations, which can be achieved through distributed mechanisms. Moreover, information exchange between terminals is restructured, thereby making it impossible to retrieve the flow information of every region. This characteristic preserves privacy and encourages data sharing across different companies. The method is particularly suited for large-grid companies that have numerous subsidiaries based in different regions, such as the State Grid Corporation of China.

ACKNOWLEDGMENT

Not applicable.

AUTHORS' CONTRIBUTIONS

Xinping Wu: project chief leader, literature review, idea proposing. Wei Yang: one of the project leaders, resource coordination and experiment design. Ning Zhang: algorithm review, experiment design, writing reviewing. Chunlei Zhou: one of the project leaders, experiment design, sample data detection and writing reviewing. Jinwei Song: problem modeling, mathmatic proving, code implementation. Chongqing Kang: paper writing and reviewing. All authors read and approved the final manuscript.

FUNDING

This work is supported by the Science and Technology Project of State Grid Cooperation of China (No. 5700-202290184A-1-1-ZN).

AVAILABILITY OF DATA AND MATERIALS

Not applicable.

DECLARATIONS

Competing interests: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this article.

AUTHORS' INFORMATION

Xinping Wu professor-level senior engineer, director of Big Data Center of China Network. In 2000, he was awarded special government allowance by The State Council. In 2004, he was selected as one of the first national candidates of the New Century Project of Ten Million Talents. He is a member of the Chinese Society of Electrical Engineering, who mainly engaged in information technology, power grid dispatching automation technology research and application. Recently, he focuses on the research and application of new technologies such as big data, artificial intelligence and knowledge graph.

Wei Yang senior engineer, deputy director of State Network Big Data Center. He mainly engaged in information technology research and application work. Recently, he focuses on the research and application of new technologies such as big data, artificial intelligence and knowledge graph.

Ning Zhang received the B.S. and Ph.D. degrees from the Department of Electrical Engineering, Tsinghua University, Beijing, China, in 2007 and 2012, respectively. He is currently an associate professor with Tsinghua University. His research interests include multiple energy system integration, renewable energy, and power system planning and operation.

Chunlei Zhou master, senior engineer. Engaged in information communication, operation monitoring, data management, data application, data security, energy big data center and other professional fields.

Jinwei Song PhD, senior engineer, corresponding author, engaged in big data and artificial intelligence research, recently engaged in power system carbon emission calculation technology research.

Chongqing Kang received the Ph.D. degree in electrical engineering from the Department of Electrical Engineering, Tsinghua University, Beijing, China, in 1997. He is currently a professor with Tsinghua University. His current research interests include low-carbon electricity, power system planning, power markets, power system reliability, and load forecasting.

REFERENCES

[1] T. Jiang, Y. Yu, and A. Jahanger, "Structural emissions reduction of China's power and heating industry under the goal of "double carbon": a perspective from input-output analysis," *Sustainable Production and Consumption*, vol. 31, pp. 346-356, 2022.

[2] National Energy Administration, "China ETS," *China Emissions Trading Scheme*, pp. 1-112, Apr. 2021.

[3] A. Runa, Z. Zhang, and H. Zhang, "Carbon emission peak and carbon neutrality under the new target and vision," in *International Conference on Advanced Electrical Equipment and Reliable Operation* (AEERO), 2021, pp. 1-5.

[4] China Ministry of Ecological Environment, "Chinese power generation enterprises guidelines for accounting methods and reporting of greenhouse gas emissions," *China Ministry of Ecological Environment*, ed., 2021.

[5] Intergovernmental Panel on Climate Change, "Guidelines for National Greenhouse Gas Inventories (2006 Ed.)," 2006.

[6] Intergovernmental Panel on Climate Change, "Guidelines for national greenhouse gas inventories (2019Ed.)," 2019.

[7] China Ministry of Ecological Environment, "Guidelines for the preparation of provincial greenhouse gas emission inventories," *C. M. o. E. Environment, ed*., 2015.

[8] World Resource Institution, "GHG protocol scope 2 guidance," *World Resource Institution*, 2015.

[9] T. Zhou, C. Kang, and Q. Xu *et al*., "Analysis on distribution characteristics and mechanisms of carbon emission flow in electric power networks," *Automation of Electric Power Systems*, vol. 36, no. 15, pp. 39-44, Aug. 2012. (in Chinese)

[10] T. Zhou, C. Kang, and Q. Xu, "Preliminary investigation on a method for carbon emission flow calculation of power system," *Automation of Electric Power Systems*, vol. 36, no. 11, pp. 44-49, Jun. 2012 . (in Chinese)

[11] T. Zhou, C. Kang, and Q. Xu, "Preliminary investigation on power system carbon emission flow," *Automation of Electric Power Systems*, vol. 36, no. 7, pp. 38-43, Apr. 2012. (in Chinese)

[12] C. Kang, T. Zhou, and Q. Chen, *et al*., "Carbon emission flow from generation to demand: a network-based model," *IEEE Transactions on Smart Grid*, vol. 6, no. 5, pp. 2386-2394, May 2015.

[13] G. J. Miller, K. Novan, and A. Jenn, "Hourly accounting of carbon emissions from electricity consumption," *Environmental Research Letters*, vol. 17, 2022.

[14] C. Yang, J. Liu, and H. Liao, *et al*., "An improved carbon emission flow method for the power grid with prosumers," *Energy Reports*, vol. 9, pp. 114-121, 2023.

[15] S. Shen, Y. Dan, and M. Qi *et al*., "Efficient whole-process carbon intensity calculation method for power users in active distribution networks," *Frontiers in Energy Research*, vol. 10, 2022.

[16] A. Borchers and T. Pieler, "Flexible low-carbon optimal dispatch of honeycombed active distribution network," *Genes (Basel)*, vol. 1, no. 3, pp. 413-426, Nov. 2010.

[17] G. Chen, B. Chen, and H. Zhou *et al*., "Life cycle carbon emission flow analysis for electricity supply system: a case study of China," *Energy Policy*, vol. 61, pp. 1276-1284, Oct. 2013.

[18] B. Tranberg, O. Corradi, and B. Lajoie *et al*., "Real-time carbon accounting method for the European electricity markets," *Energy Strategy Reviews*, vol. 26, pp. 4, Nov. 2019.

[19] J. Xin, R. Fan, and S. Zheng *et al*., "Low-carbon benefit simulation and evaluation system for smart distribution

grid and its application in Jiangxi Province," *Power System Protection and Control*, vol. 42, no. 7, pp. 86-90, Apr. 2014. (in Chinese)

[20] Y. Cheng, N. Zhang, and Y. Wang *et al*., "Modeling carbon emission flow in multiple energy systems," *IEEE Transactions on Smart Grid*, vol. 10, no. 4, pp. 3562-3574, Jul. 2019.

[21] Y. Cheng, N. Zhang, and C. Kang, "Carbon emission flow: from electricity network to multiple energy systems," *Global Energy Interconnection*, 2018.

[22] J. Huang, W. Duan, and Q. Zhou *et al*., "Methodology for carbon emission flow calculation of integrated energy systems," in *2022 the 4th International Conference on Clean Energy and Electrical Systems (CEES 2022)*, Tokyo, Japan, Apr. 2022, pp. 8-13.

[23] M. Qin, Y. Yang, and X. Zhao *et al*., "Low-carbon economic multi-objective dispatch of integrated energy system considering the price fluctuation of natural gas and carbon emission accounting," *Protection and Control of Modern Power Systems*, vol. 8, no. 4, pp. 1013-1030, Oct. 2023.

[24] H. Chen, W. Mao, and R. Zhang *et al*., "Low-carbon optimal scheduling of a power system source-load considering coordination based on carbon emission flow theory," *Power System Protection and Control*, vol. 49, no. 10, pp. 1-11, May 2021. (in Chinese)

[25] Y. Li, N. Zhang, and E. Du *et al*., "Mechanism study and benefit analysis on power system low carbon demand response based on carbon emission flow," *Proceedings of the CSEE*, vol. 42, no. 8, pp. 2830-2841, Apr. 2022. (in Chinese)

[26] F. Ahsan, N. H. Dana, and S. K. Sarker *et al*., "Data-driven next-generation smart grid towards sustainable energy evolution: techniques and technology review," *Protection and Control of Modern Power Systems*, vol. 8, no.3, pp. 696-737, Jul. 2023.

[27] J. Bialek, and D. B. Tam, "Tracing the generators output," in *International Conference on Opportunities & Advances in International Electric Power Gene*ration, 1996, pp. 4-10.

[28] D. Kirschen, R. Allan, and G. Strbac, "Contributions of individual generators to loads and flows," *IEEE Transactions on Power Systems*, vol. 12, no. 1, pp. 52-60, Feb. 1997.