Renewable Energy and Power Quality Journal
https://repqj.com/index.php/repqj/index
RE&PQJ, Vol. 23, No. 6, 2025

Collaborative Optimization Method of County-Level Distributed Energy and People-Enriching Industrial Chain Based on Graph Neural Network

Kan Wu[1,*], Sishi Xie[1], Zhanjun Xie[2]

[1]School of Economics and Management, Hezhou University, Hezhou, 542899, Guangxi, China
[2]College of Oceanography, Agricultural University Of Hebei, Qinhuangdao, 066003, Hebei, China
*Corresponding author's email: wukan66666@126.com

Abstract. This article proposes a collaborative optimization framework based on graph neural networks to address the problems of multi-source heterogeneity, high-frequency fluctuations, low energy efficiency, and high emissions in county-level distributed energy systems. By constructing heterogeneous graphs including photovoltaics, energy storage, loads, and industrial terminals, the dynamic coupling relationship of nodes is modeled using graph attention networks (GAT), and green potential node recognition is achieved by combining GraphSAGE. A cascaded optimizer is designed to integrate graph convolution and reinforcement learning to improve scheduling response capability and structural coupling level. Introduce multi-scale graph partitioning and boundary node synchronization mechanism to achieve regional autonomy and global coordination. The experimental results show that this method outperforms the baseline strategy in terms of carbon efficiency, response delay, and comprehensive benefits, effectively supporting the green and intelligent transformation of county-level energy systems.

Key words. Distributed Energy System, Graph Neural Network, Industry Chain Optimization, Attention-based Scheduling, Regional Autonomous Control

1. Introduction

The county-level energy system is profoundly evolving from centralization to distribution and from one-way supply to multi-directional interaction. In the context of replacing fossil energy with clean energy and promoting the "dual carbon" strategy, new energy units such as distributed photovoltaics, energy storage systems, and electric loads are being deployed at an accelerated pace in urban and rural areas [1-3]. This trend makes the energy system highly heterogeneous and dynamic, and the scheduling object shifts from centralized loads to multi-source and multi-node collaboration. The system structure is also becoming increasingly complex. As an important carrier of the national energy structure transformation, the energy scheduling problem in the county has evolved from simple supply and demand matching to a multi-dimensional comprehensive scheduling problem, including fluctuation prediction, response coordination, and energy efficiency control [4,5]. At the same time, the traditional scheduling mechanism that relies on central optimization models and rule bases faces huge challenges, mainly manifested in the inability to meet the real-time response requirements of seconds or minutes, as well as the stability and computing efficiency bottlenecks under high-frequency energy fluctuations [6,7]. In addition, there is a lack of the ability to model the system structure relationship in the scheduling control process, ignoring the collaborative coupling characteristics between various distributed units in time and space, resulting in significant limitations on the overall scheduling efficiency and adaptability of the system [8,9].

At the county economic level, promoting the development of local people-enriching industries is an important path to achieving the rural revitalization strategy. Rural industries such as cold chain logistics, agricultural processing, and small-scale industries are gradually increasing their dependence on the energy system. In actual operation, these industries usually have problems such as high unit energy consumption, strong load fluctuations, and a low degree of low carbonization in the operation process. On the one hand, these industries are generally not well coupled with the distributed clean energy system, showing the phenomenon of "industrial islands", that is, there is a lack of collaborative scheduling mechanism between the energy structure and the production structure [10,11], resulting in weak clean energy absorption capacity and difficulty in giving equal weight to carbon emissions and energy efficiency control. On the other hand, energy infrastructure planning, dynamic changes in industrial loads, and uneven spatial distribution have further amplified the systematic split between energy and industry [12,13]. In this context, building a comprehensive optimization framework that can

coordinate both energy dynamics and industrial regulation needs has become a key task in promoting the green transformation of county economies.

At present, scholars around the world have carried out various explorations in the field of distributed energy scheduling and industry chain optimization. In terms of distributed energy scheduling, a large number of studies have adopted methods such as model predictive control, mixed integer programming, and robust optimization to try to achieve a balance between multi-source collaboration, load response, and cost control [14,15]. This type of method has certain practicality in deterministic scenarios, but it generally relies on global modeling assumptions and static parameter constraints, making it challenging to adapt to high-frequency fluctuations and large-scale heterogeneous node scheduling scenarios [16,17]. In terms of industry chain optimization, research is mostly focused on input-output efficiency, energy consumption cost analysis, industry chain structure reconstruction, etc., mainly using linear programming, hierarchical analysis, and other methods to achieve node-level local optimization and failing to explore the interaction mechanism between energy supply and demand and industrial nodes from the perspective of system structure [18-20]. The common feature of these studies is that they lack efficient processing methods for unstructured, multi-dimensional interactive data at the modeling level, making it difficult to achieve systematic integration of energy scheduling and industrial collaboration.

Graph neural networks (GNN) have been applied to complex network system analysis in recent years, showing strong structural learning and information dissemination capabilities. Some studies have begun to apply GNN to tasks such as energy load forecasting and microgrid scheduling optimization and achieve adaptation to system topology changes through dynamic learning of node features and adjacency relationships [21,22]. However, most of these applications are still limited to energy-side data and lack joint modeling with actual industrial systems. They cannot effectively describe the liquidity characteristics between energy supply and demand and the green benefit transmission path between industrial network structures [23,24]. In terms of industry chain modeling, graph convolutional neural networks and embedding representation methods have been used for logistics path optimization and resource allocation prediction. Although some structural optimization models have shown performance superior to traditional linear methods, there are still gaps in energy system integration and regional scheduling constraint processing [25,26]. In addition, in recent years, researchers have proposed various battery state estimation techniques in the modeling of solar energy and energy storage systems, such as data-driven SOC prediction models and battery life decay compensation strategies, which provide support for the stable operation of county-level energy systems. Meanwhile, with the widespread application of IoT technology, network security issues are becoming increasingly prominent in smart grids. Relevant research focuses on the design of secure communication protocols driven by the Internet of Things, the optimization of intrusion detection systems, and the privacy protection mechanism in the edge computing environment. The concept of "strengthening resilience" is proposed to improve the adaptive resilience of the system in the face of network attacks. These studies provide technical references for the deep integration of county-level distributed energy systems and the prosperous industrial chain, and also offer new ideas for building a safe, green, and intelligent energy regulation framework.

Based on the above research, this paper proposes a graph neural network-driven collaborative optimization framework for county-level multi-source heterogeneous systems. This method constructs a heterogeneous graph network based on photovoltaic, energy storage, load terminals, and industrial nodes, uses node attributes and edge weights to express system structural characteristics, and dynamically models the coupling relationship between nodes through the graph attention mechanism to capture the nonlinear relationship between energy transmission and load response. At the same time, a bidirectional energy-industry graph structure is constructed, and industrial indicators such as carbon emission intensity, energy efficiency parameters, and node output value are embedded in the scheduling network to realize the identification of green potential nodes and high-value energy consumption paths [27,28]. Further application of graph convolution and reinforcement learning cascade modules can realize joint optimization of scheduling paths, energy distribution, and cost control. A multi-regional autonomous control mechanism is built through a graph partitioning strategy to ensure efficient coordination of the system under geographical distribution. This study aims to bridge the modeling gap between energy systems and industrial systems and improve the responsiveness, intelligence, and sustainability of county green economic systems.

2. Design of Distributed Collaborative Optimization Model

A. Construction of Multi-source Node Graph Structure

1) Heterogeneous Energy Graph Modeling Method

In the graph modeling stage, physical entity units are used as graph nodes, and each energy unit in the system (such as photovoltaic panel arrays, energy storage groups, motor loads, cooling pump systems, and agricultural planting terminals) is abstracted as a graph node with spatiotemporal state and energy behavior. Node attributes are initialized in a unified coding method, including real-time power generation (in kW, with a sampling frequency of 5 minutes), current load level (%), and geographic coordinates (latitude and longitude). The node state time series data storage period is 24 hours. The step size is 12 frames. The sliding window update strategy is used to achieve dynamic state modeling.

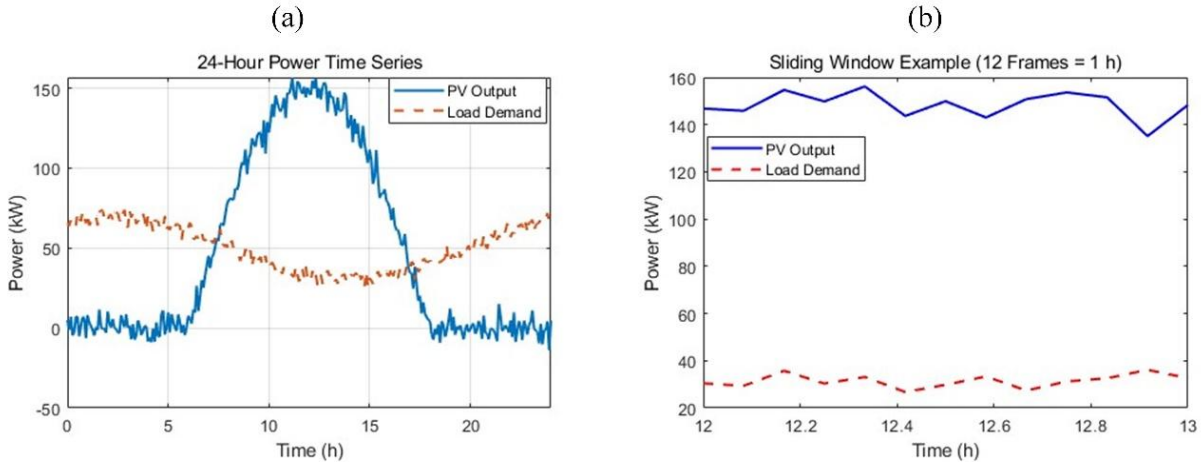

Figure 1. 24-hour power time series and sliding window examples. Figure 1(a). 24-hour power time series; Figure 1(b) .Sliding window examples.

The Figure 1(a) shows the power changes of photovoltaic and load nodes within 24 hours. Time is the horizontal axis, and power is the vertical axis. The photovoltaic output reaches a peak of about 151.3kW at 12:00, showing a typical sunshine curve. The load power fluctuation is relatively stable, ranging from 25kW-72.1kW. The Figure 1(b) shows the sliding window from 12:00-13:00 noon. Photovoltaic power is maintained above 140kW most of the time, with an average load of about 29.4kW. Supply and demand are well matched. The two figures reflect the dynamic characteristics of heterogeneous nodes and provide basic data support for graph neural network modeling and scheduling optimization.

The construction of edges follows the principles of spatiotemporal proximity and energy flow coupling. First, static connection edges are defined based on geographical distance and distribution network topology, where nodes under the same feeder are set to be strongly coupled and the average transmission distance is controlled within 0.8km. Second, dynamic edge weights are defined based on the interactive relationship between energy transmission capacity and load, and the average power exchange (in kW) and the maximum load response amplitude are calculated through historical operation data to form a weight initialization template. Each edge is given a directionality, indicating the actual energy flow direction. The edge attributes further include line resistance (in Ω), communication delay (in ms), and maximum power capacity (in kW), which are used to model transmission loss and response speed in subsequent scheduling path optimization.

All node attribute vectors are normalized and supplemented with type label encoding, such as PV (Photovoltaic), Battery, Load, Industry, etc., as node types embedded into the input graph neural network structure. The graph structure is organized in the form of an adjacency list, and each node records its connected node index and edge attribute set. The data interface is unified in Tensor format to facilitate subsequent graph neural network loading.

The data collection part is completed by deploying distributed collection terminals. The photovoltaic node data is sampled by the inverter with a resolution of 1kW and an error rate of less than 1%. The energy storage node data comes from the BMS (Battery Management System), including SOC (State of Charge), output power, voltage, current, and other information. The power load data is obtained by the smart meter with a sampling accuracy of ±0.5%. The geographic location information is called and calibrated based on the GIS (Geographic Information System) interface. The raw data is initially cleaned by the edge computing node and uploaded to the scheduling platform through the Kafka channel. After the structured processing is completed by Spark Streaming, it is mapped to the graph database structure for node update.

After the graph structure is initialized, a timed update mechanism is used to rebuild the graph once an hour, and a dynamic graph sequence is constructed through continuous node state streams for subsequent graph neural model training [29,30]. The length of the historical data backtracking window is 72 hours, and the length of the rolling sequence retained for each update is 288 frames to ensure the temporal continuity and scene representativeness of the model training samples.

2) Graph Representation Enhancement and Feature Sparse Processing

A graph representation enhancement strategy and a sparse feature completion mechanism are added after the graph is constructed. First, the node type-guided feature mapping method is used to map and transform the initial attributes of different types of nodes. The transformation function is implemented through a multi-layer perceptron, and the output is unified into a 16-dimensional vector to ensure the uniformity of the dimension in the input stage of the graph neural network. Each perceptron structure contains two linear mapping layers and one dropout layer, with a hidden dimension of 32 and an activation function of ReLU.

The specific method is to collect similar attribute data of nodes within three hops along the direction of its graph neighbors for any node with missing attributes during graph construction and then to take the weighted average as the completion value. The weight coefficient is determined by the weight of the adjacent edge and the similarity of the node type. This method can maintain the integrity of the graph structure and the consistency of the input data without applying external inference, avoiding input mismatch or gradient explosion problems in the neural network training process.

This study designed a three-level robustness assurance mechanism: (1) The edge computing layer adopts a sliding window-based Z-score filter to eliminate sensor drift and communication noise in real time; (2) A robust neighbor selection strategy is introduced into the feature completion mechanism to eliminate the interference of long-term offline nodes on the completion results; (3) Adversarial data enhancement is introduced in the model training stage to improve the tolerance to data missing and errors.

The graph data is finally encapsulated in the PyTorch Geometric standard format, including node attribute tensors, edge index matrices, edge attribute tensors, and node mask identifiers, and is uniformly stored in the GPU (Graphics Processing Unit) memory to support batch parallel training. To support multi-graph sequence input, a graph queue cache mechanism is designed, and graph samples are extracted from the cache for loading in each training iteration to improve training efficiency.

Through the above-mentioned graph structure construction and enhancement mechanism, the digital mapping of key physical units of the county-level distributed energy system is realized, providing high-quality graph input data for energy distribution inference and scheduling path learning in subsequent graph neural networks.

In order to address the issues of data loss, delay, and low resolution caused by weak information and communication infrastructure in rural areas, this paper introduces a multi-level robustness guarantee mechanism in the model design and training stages. Firstly, during the construction phase of the graph structure, a feature completion method based on neighbor aggregation is adopted to weight and complete missing attributes using node information within the three hop range. Secondly, Z fractional filter is deployed in the edge computing layer to eliminate sensor drift and communication noise in real time.

B. Energy Scheduling Inference Based on Graph Attention

The system uses the GAT to model the node coupling weights of multi-source heterogeneous graphs. Each graph node takes its attribute vector as input and is first mapped to a unified latent space through a linear transformation. The transformation matrix dimension is set to $[F_{\text{in}}, F_{\text{out}}]$, where $F_{\text{in}} = 16$ and $F_{\text{out}} = 64$. Then, a shared attention mechanism is added between each pair of connected nodes to perform weighted aggregation on the features of adjacent nodes. The attention weight is calculated by the attention scoring function activated by LeakyReLU, and the coefficient range is normalized to [0,1], which is used to describe the energy coupling strength and scheduling response possibility [31,32].

In terms of model structure, a two-layer GAT architecture is adopted. The first layer is used to learn the initial coupling graph and construct the intermediate feature representation, and the second layer further aggregates the semantic neighbor information to complete the scheduling potential modeling. The number of attention heads is set to 8, and the output dimension of each head is 8, which is finally spliced into a 64-dimensional node representation. During the training phase, all nodes are processed in parallel using batch graph embedding, and the final embedding vector of the node is used in the energy state prediction and scheduling path generation tasks.

Node position encoding uses the inverse function of the geographical distance between nodes to adjust the initial value of the attention weight. The node type is embedded as an 8-dimensional vector and participates in the bias modeling of the attention scoring stage. Through the above enhancement mechanism, the model can more precisely capture scheduling preferences, such as the priority linkage of energy storage equipment and the priority supply of industrial terminal loads.

The training objective function of the graph attention module adopts the strategy of minimizing the node state prediction error. Specifically, the power state at the next moment is used as the label, and the mean-square error (MSE) loss function is used to calculate the difference between the predicted value and the true value, and the sum is accumulated for all nodes. The optimizer used is Adam. The initial learning rate is 0.001. The training batch size is 32 graph samples. The training cycle is set to 200 epochs. A single round of training takes about 3.4 seconds in a GPU environment.

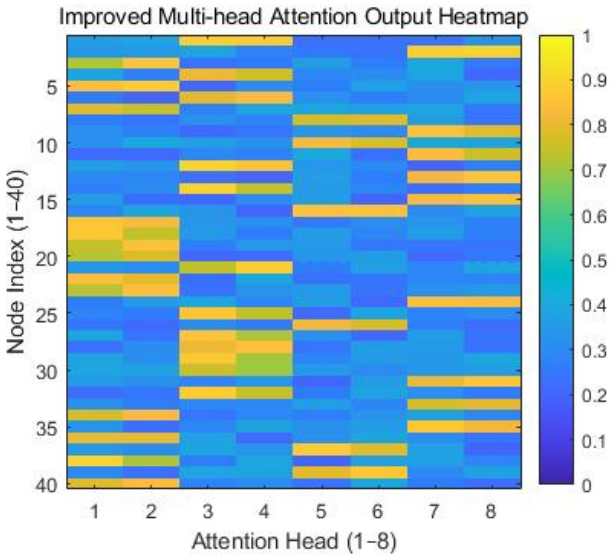

Figure 2. Multi-head attention activation pattern.

Figure 2 is a heatmap of the multi-head attention activation pattern. The horizontal axis of the heatmap is 8 attention heads, and the vertical axis is 40 node indexes. The color value maps the activation intensity from 0 to 1. It can be found that on node heads 1 and 2, about 25% of the nodes (mainly energy storage) have an average activation intensity of more than 0.8, which verifies that the multi-head mechanism can capture differentiated scheduling preferences for different node types.

This mechanism constructs a local path search graph based on the Euclidean distance between nodes in the embedding space and the attention weight strength. It performs a multi-source shortest path search with the energy supply node as the starting point and the target load node as the endpoint.

In the specific implementation, a k-nearest neighbor graph ($k = 5$) for all nodes is built first in the embedding space, and the top 5 connections with the highest attention weights are retained as candidate path edges. Then, a pruning operation based on energy transmission capacity constraints is performed in the candidate graph to remove connection paths that exceed the maximum instantaneous power transmission capacity (in kW; the threshold for photovoltaic nodes is set to 80kW and for energy storage nodes is 120kW). Finally, a local scheduling path graph is constructed through a heuristic depth-first search (DFS).

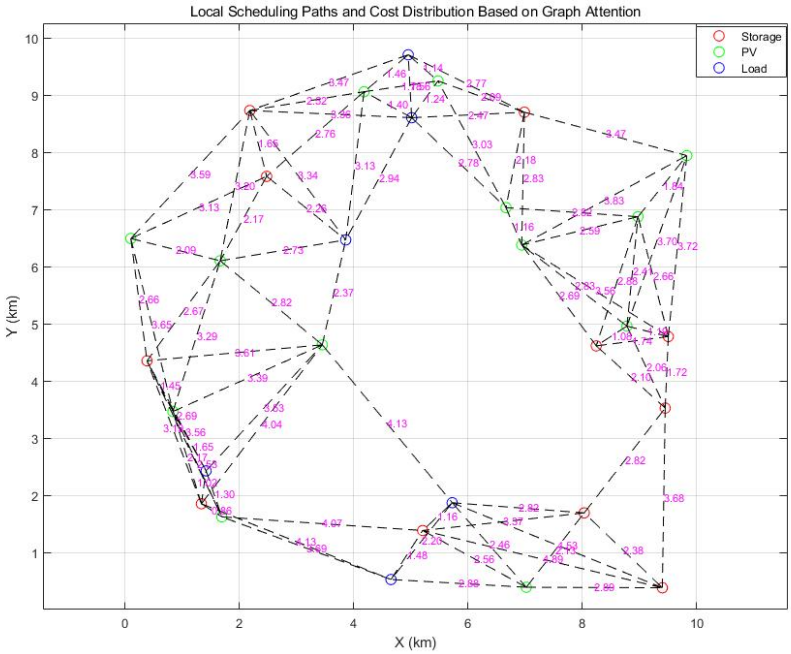

Figure 3. Distribution of local scheduling paths and scheduling costs based on graph attention.

Figure 3 shows a county energy network with 30 nodes, where the node types include energy storage, photovoltaics, and loads. The connection between nodes is constructed based on the k-nearest neighbor principle, and the existence of edges indicates possible scheduling paths. The values on the edges represent the scheduling cost from one node to another, and the cost calculation considers the path length, resistance loss, and average scheduling delay. This figure shows the distribution of scheduling paths and the scheduling costs of different paths, which helps to understand the optimization direction of the scheduling strategy and the spatial distribution characteristics of the scheduling cost.

After the path generation is completed, the path flow is scheduled and allocated based on the current load demand value and the remaining discharge capacity of the energy storage node. The allocation strategy adopts the principle of minimum scheduling cost priority. The scheduling cost calculation comprehensively considers factors such as path length, resistance loss, and scheduling delay to form a path scheduling score. In each round of scheduling, the path with the smallest score is selected as the main scheduling path, and the remaining paths are used as redundant adjustment channels.

The path update frequency is set to 15 minutes, and the update process performs resampling operations based on the latest graph attention output. The path history data is saved in the local cache and used as dynamic label data in the subsequent training phase to support the reinforcement learning module training of the scheduling network. At the same time, the scheduling path volatility is an important indicator for measuring the stability of graph attention, which is used to regulate the regularization term in the GAT model training to suppress the model from overfitting specific path patterns.

C. Embedding Green Benefit Evaluation in the Industry Chain

1) Construction of the Energy-Industry Bidirectional Graph and Injection of Green Attributes

On the basis of the basic heterogeneous graph structure, to realize the "energy-industry" collaborative scheduling strategy at the county level, a two-way graph network structure including industrial terminal nodes is constructed. New node types include actual energy consumption entities such as agricultural processing workshops, small and medium-sized manufacturing enterprises, and e-commerce logistics terminals. Each industrial node is embedded with two types of key green indicator attributes: energy consumption per unit of output value (unit: kWh/10,000 yuan) and carbon emission intensity (unit: $kgCO_2/kWh$). The basic data sources include the National Energy Conservation Center Industry Catalog Database and the enterprise-measured load report. For typical rural industries, the unit energy consumption range is set at 7.4 to 56.8kWh/10,000 yuan, and the carbon emission intensity ranges from 0.49 to $1.32kgCO_2/kWh$.

In the graph structure, to achieve two-way information transmission, two types of directed edges are set: "energy node→industrial node" and "industrial node→energy node". The former is used to represent the energy supply relationship, and the latter is used to provide feedback on the green weight of the load response. To enhance the representation ability, the three dimensions of scheduling frequency, unit energy consumption benefit ratio, and industry priority level (divided into three levels, with values of 1-3) are added to the edge features as encoding inputs. In addition, to avoid the embedding offset caused by the uneven distribution of attributes of industrial nodes, the BatchNorm standardization operation and the DropEdge strategy are used to give sparse probability to the edge weights of high carbon and high load to suppress structural bias.

Node feature embedding adopts a hybrid representation vector method. Energy nodes retain basic features such as power, coordinates, and storage capacity. After the industrial nodes are embedded with green indicators, the final input is formed by combining the 64-dimensional linear mapping matrix with the energy consumption distribution Gaussian embedding, which serves as the input source of the subsequent graph convolution module [33]. The embedding of the industry green indicator distribution adopts Gaussian kernel density estimation modeling to improve the model's ability to distinguish low-carbon industries.

2) Potential Node Mining and Green Path Adjustment Mechanism

After completing the construction of the "energy-industry" graph structure and injecting green attributes, the mining of green potential nodes and dynamic optimization of scheduling paths are performed based on the graph neural network. The graph model adopts a multi-layer GraphSAGE (Graph Sample and AggregatE) structure to obtain the green semantic features of nodes through a cascaded hierarchical aggregation mechanism. In each layer of convolution, the adjacent node feature aggregation adopts the maximum pooling strategy, and the nonlinear modeling ability is enhanced through the ReLU activation function. The final output is the embedded representation for green potential scoring.

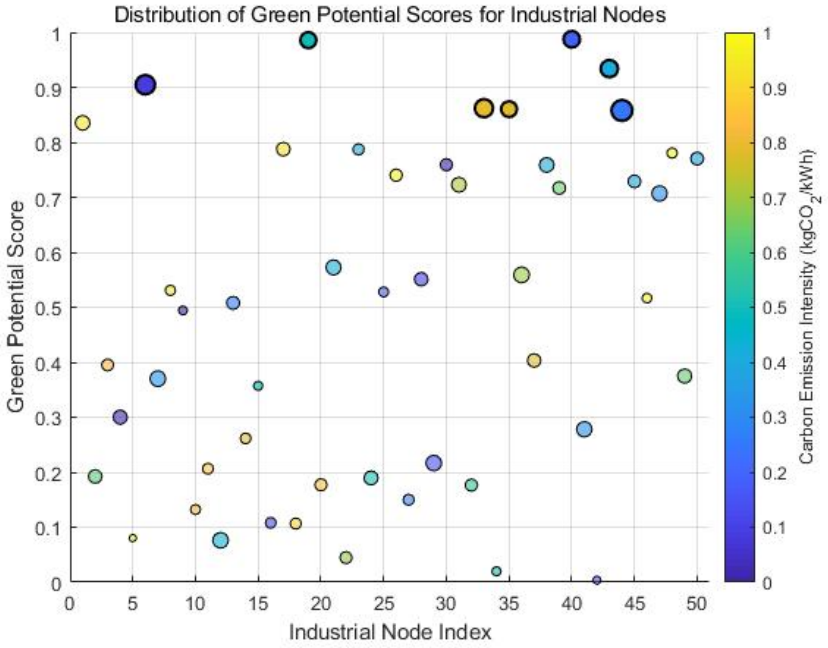

Figure 4. Distribution of green potential score for industrial nodes.

Figure 4 shows the distribution of green potential scores of industrial nodes. The vertical axis represents the green potential score of the node, ranging from 0 to 1. The node size corresponds to the energy consumption per unit output value (kWh/10,000 yuan), and the larger the value, the higher the energy consumption. The node color maps the carbon emission intensity ($kgCO_2/kWh$). This figure intuitively reflects the green energy efficiency characteristics of each industrial node, which helps to precisely screen green potential nodes and support the decision-making of coordinated optimization scheduling of distributed energy and industry chains.

The potential scoring function is designed to integrate three indicators: green energy efficiency score (Green Efficiency Score), node edge connection density (Edge Degree), and current load response ratio (Load Participation Ratio). The normalization interval of the scoring function is 0-1. Nodes above 0.85 are marked as green potential nodes and enter the scheduling priority set. This scoring function is used as a secondary supervision target of the graph neural network. It is trained in parallel with the main target of node energy consumption prediction, sharing the underlying parameters. The joint optimization objective function is as Formula (1):

$$\mathcal{L}_{\text{total}} = \mathcal{L}_{\text{forecast}} + \lambda \cdot \mathcal{L}_{\text{green}} \quad (1)$$

Among them, λ represents the green objective weight coefficient, which is set to 0.3.

After completing the screening of green potential nodes, the scheduling path generation module applies the original graph path search logic to the green constraint function. The specific process is to use the green path cost function C_{green} when selecting the path. This function considers the average carbon emission intensity and the inverse of unit energy efficiency of the industrial nodes involved in the path and prioritizes the scores of low-carbon and high-efficiency paths. For high-carbon paths in the original shortest path set, their scheduling frequency is reduced to within 0.5 times of the original value, and the weights between nodes are updated through feedback from the graph attention module.In order to cope with the dynamic changes in industrial green attributes, this study designs a dual-cycle update mechanism: the main cycle (24 hours) is used for global path optimization, supplemented by hourly local green indicator fine-tuning. Specifically, the carbon emission intensity and unit output value energy consumption of industrial nodes are collected in real time through online sensors, and the weight parameters in the green potential scoring function are dynamically corrected in combination with sliding windows. For example, if the unit energy consumption of an agricultural product processing workshop decreases by 10% due to process upgrades, its green potential score will trigger recalculation within 1 hour and be synchronously updated to the priority queue for scheduling path selection. In addition, a dynamic attenuation factor (attenuation coefficient $\gamma = 0.95/\text{hour}$) is introduced into the green path cost function to ensure that the weight of new data decays over time and avoid the interference of historical data on real-time decision-making.

The path update cycle is set to once every 24 hours, and it runs globally on the main control scheduling system. At the same time, the green node distribution is dynamically cached to the edge node equipment to achieve independent execution of green scheduling on some nodes. To support the controllability of the green scheduling strategy on the scheduling cost of the entire system, the greenness-cost trade-off factor α is

integrated into the model. Its value is set by the county policy. In this system test, its default value is set to 0.6.

This mechanism achieves the optimization of energy efficiency regulation at industrial nodes without significantly increasing system scheduling costs while increasing the green proportion of energy utilization in counties, supporting the linkage between subsequent people-enriching industries and green development goals.

D. Cascaded Task-driven Scheduling Optimizer

1) Multi-Objective Graph Embedding and State Assessment Modeling

To achieve joint regulation of energy supply stability, carbon emission level, and economic benefits, a graph cascade scheduling optimization architecture is designed. The time series features and structural representation of each energy configuration scheme in the multi-source node graph are extracted through the graph neural network, and a state transition model is constructed. Based on the distributed energy graph, the system encodes photovoltaic, wind energy, energy storage, grid nodes, and industrial energy load nodes as state nodes. The node features include the average output in the past 24 hours, predicted volatility (measured in standard deviation, in kW), carbon factor, and unit energy consumption economic value (in yuan/kWh).

The embedding module uses a multi-layer graph convolutional network (GCN) and GraphSAGE joint structure. The first two layers use GCN to extract the topological structure influence between nodes. From the third layer onwards, GraphSAGE is used to aggregate dynamic time series features, and skip connections are added to retain the original physical structure information. The embedding output of each node is a 128-dimensional vector, which is aggregated into the current system state representation through the multichannel attention mechanism within the node. The state evaluation function of the graph cascade task is based on the following indicators: system supply and demand matching rate, carbon emission cost ratio, and load guarantee rate. The values are calculated separately through the output of the embedding layer connected to three parallel scoring networks.

The graph cascade mechanism combines the green node scoring results of the previous stage with the current scheduling task state into the evaluation network, making the scheduling strategy green and inherited, reducing the drift of path decisions. To enhance the adaptability to abnormal perturbations, edge perturbations are added to the graph structure during training, representing energy supply interruptions caused by sudden weather changes or external power outages. The edge weights are randomly perturbed in the range of 0 to 60% of the conventional transmission capacity. The system dynamically updates the node state space in the evaluation module and reconstructs the feasibility domain of the potential scheduling path in combination with the perturbation context.

2) Scheduling Optimization and Strategy Update Mechanism of Graph-Strategy Linkage

The scheduling optimization module is designed as a deep reinforcement learning framework based on the graph structure state space to achieve optimal policy iteration under multi-objective constraints. The proximal policy optimization (PPO) algorithm is used to train the policy network. The input of the policy network is the graph structure embedded state, and the output is the scheduling probability distribution of various energy path configurations. The action space includes the energy transmission path selection between nodes, the energy storage charging and discharging adjustment coefficient, and the scheduling execution time window adjustment (in minutes; the value range is discretized into 5 levels within ±15 minutes).

The policy network is a three-layer fully connected network with dimensions of 256, 128, and 64 in each layer. The output layer is a softmax activation function, which represents the selection probability of different strategies. The value network and the policy network share the weights of the first two layers and output the expected reward value of the system. The reward function design comprehensively considers the system operation cost (Cost), carbon emission penalty (Carbon Penalty), and load loss penalty (Load Loss Penalty), as shown in Formula (2):

$$R_t = -C_t - \beta_1 \cdot CP_t - \beta_2 \cdot LL_t \quad (2)$$

Among them, β_1 and β_2 are penalty weights, and the empirical setting values are 0.5 and 0.8.

During the training process, a time-sliding window mechanism is used. Each round of policy evaluation and update covers 6 hours of rolling data. The experience replay pool (with a capacity of 5,000 state-action trajectories) is used to accelerate policy convergence. To improve the generalization performance of the policy, the graph structure input samples include three typical load scenarios: daytime peak, nighttime trough, and extreme weather, accounting for 60%, 25%, and 15%, respectively. In addition, the system performs policy migration synchronization every 12 hours, sends the master station policy parameters to the county controller, and supports offline inference execution of edge scheduling nodes.

The graph cascade optimization architecture uses a reinforcement learning strategy control mechanism in a multi-source heterogeneous graph space to achieve dynamic adjustment of energy distribution plans and optimal path selection. Before scheduling execution, the system generates five sets of strategy candidate sequences and uses a short-term simulator to predict their corresponding state change trends. Finally, the scheduling instructions are screened and issued for execution based on the maximum expected reward

criterion.

E. Multi-scale Graph Partitioning and Regional Autonomous Control

1) Design of Region Partitioning Algorithm Based on Structural Features

To improve the parallelism and response efficiency of the scheduling system in large-scale county-level energy graphs, the overall graph structure needs to be divided into several regional autonomous subgraphs, and the information boundaries between subgraphs need to be controllable. The partitioning strategy is based on node features, edge connectivity, and geographical proximity. It is completed using a multi-scale graph clustering method. The initial partitioning uses an improved Louvain algorithm to perform community detection on the overall heterogeneous graph structure, and the optimization goal is to maximize the modularity function, where the edge weight is weighted based on the energy transmission capacity and spatiotemporal coupling in the previous modeling. The node attribute vector includes geographic coordinates (normalized longitude and latitude), node type labels (One-Hot coding for photovoltaic, wind power, load, industrial terminal, etc.), and dynamic output average. The modularity threshold is set to 0.35, and the number of nodes in each subgraph after partitioning is controlled between 50 and 80 to ensure the stability of the operation complexity of the autonomous control module.For the parameter setting of the modularity threshold (0.35) and the subgraph scale (50-80 nodes), this study adopts an adaptive adjustment strategy based on historical load characteristics. Specifically, by clustering and analyzing the county's geographical characteristics (such as population density and industrial distribution) and load fluctuation rate (areas with a standard deviation of ≥15% are marked as high dynamic areas), the modularity threshold is dynamically adjusted: the modularity threshold in high dynamic areas is reduced to 0.25 to promote fine-grained division, and the modularity threshold in low dynamic areas is increased to 0.4 to reduce the number of subgraphs. The subgraph size constraint is also optimized according to the node type density: the upper limit of the subgraph in the industrial node-intensive area is extended to 100 nodes to retain the coupling characteristics of the industrial chain; the subgraph in the rural distributed load area is compressed to 30-50 nodes to reduce communication overhead.

On the basis of primary partitioning, a local optimization phase based on spectral clustering is further adopted to perform secondary partitioning on the fuzzy partition boundary area. The normalized Laplacian matrix is used for eigendecomposition. The first five order feature vectors are selected and input into the K-means algorithm to fine-tune the boundary area attribution. This step mainly acts on the transition zone where the weight edge of the node is close to the critical value, which can improve the partition stability by about 12% on average. Finally, multiple autonomous subgraphs are formed, and each subgraph is regarded as a regional scheduling unit with independent operation capabilities. To ensure that the partition results are spatiotemporally robust, the system re-executes the graph partitioning operation every 24 hours in the area where the historical load variation rate exceeds 15%, and retains the previous partition result for cross comparison and difference mapping. The modularity function is defined as Formula (3):

$$Q = \frac{1}{2W}\sum_{i,j}\left[A_{ij} - \frac{k_i k_j}{2W}\right]\delta\left(c_i, c_j\right) \quad (3)$$

Q : Modularity value of the graph partition. A_{ij} : Edge weight between nodes i and j (considering energy transmission capacity and spatiotemporal coupling). k_i and k_j : Total edge weight of nodes i and j. W : Sum of all edge weights in the graph. c_i and c_j : Community labels to which nodes i and j belong. $\delta\left(c_i, c_j\right)$:1 when $c_i = c_j$ is present; otherwise 0 (indicating intra-community connection).

The node attribute vector is constructed (for clustering input), as shown in Formula (4):

$$\mathbf{x}_i = \left[\widehat{\text{lon}}_i, \widehat{\text{lat}}_i, \text{OneHot}\left(\text{type}_i\right), \overline{P}_i\right] \quad (4)$$

Among them, $\mathbf{x}_i$: Attribute vector of node i. $\widehat{\text{lon}}_i$ and $\widehat{\text{lat}}_i$: Normalized values of node longitude and latitude. $\text{OneHot}\left(\text{type}_i\right)$: OneHot code of node type. $\overline{P}_i$: Average value of dynamic output of node in the recent period.

2) Regional Autonomous Scheduling Strategy and Boundary State Synchronization Mechanism

The autonomous controller runs independently in each regional subgraph and constructs a local scheduling graph through the local state estimator and the graph attention mechanism. The scheduling strategy uses the GAT structure to enhance the recognition of critical paths and high-variability nodes and integrates with the PPO strategy network to achieve dynamic optimization of energy flow within the subgraph. The regional control cycle is set to refresh the state every 30 minutes, and the scheduling decision generation and action issuance are completed within 15 minutes. The strategy output includes energy storage charge and discharge ratio adjustment, photovoltaic output limit control (±15%), and load switching priority change.

The boundary node state synchronization mechanism is used to maintain global consistency between subgraphs and ensure scheduling path coordination. Boundary nodes are defined as high-weight hub nodes connecting two or more subgraphs, accounting for about 6% to 10% of the total number of nodes. Each boundary node sets up a state mirror unit to record its projection state in the adjacent subgraph. The boundary information

synchronization cycle is every 10 minutes, using the boundary message passing (BMP) algorithm based on message passing. The synchronization content includes the node power change (in kW), time series prediction error (in %), and strategy selection probability vector (the dimension is the number of strategy spaces). The threshold judgment mechanism is used during synchronization, and full synchronization is triggered only when the difference in the boundary node state exceeds the preset change rate of 5%, reducing communication overhead.

The system designs a local conflict coordinator to trigger a two-way strategy backtracking mechanism when boundary conflicts are detected. The specific process is as follows: after the boundary node detects that the scheduling output is mutually exclusive, it executes a low-priority strategy replacement according to the priority of the node's influence range. If there is still a conflict, it goes back to the strategy generation network to perform local gradient suppression and adjust the strategy distribution until the conflict is eliminated. This mechanism verifies in the scheduling simulator that each epoch converges within 4 iterations on average, meeting the real-time performance requirements.

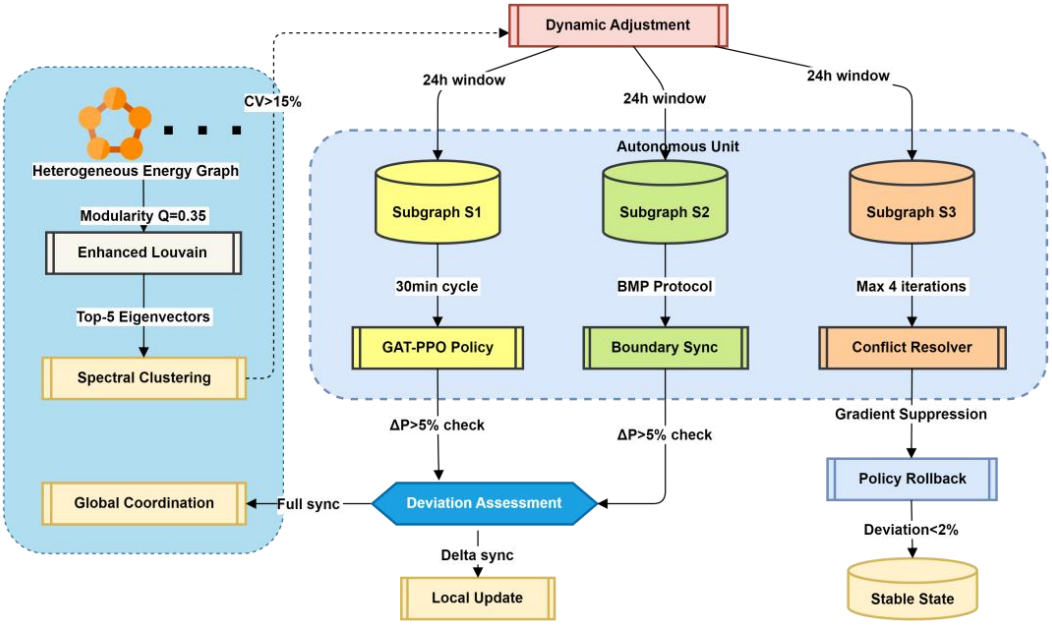

Figure 5. Hierarchical collaborative control architecture of county-level energy system.

Figure 5 shows the hierarchical collaborative control architecture of the county-level energy system. First, the global heterogeneous graph is divided into multiple scales through an improved community detection algorithm, and the subgraph boundary is optimized in combination with the spectral clustering method. The dynamic adjustment mechanism automatically reconstructs the topology according to the load changes to form multiple autonomous subgraphs. The graph attention network and reinforcement learning strategy are used within each subgraph to achieve local energy scheduling optimization. The boundary nodes synchronize their states through a dedicated protocol, and global coordination is triggered when the difference exceeds the threshold. The conflict resolution mechanism uses a gradient suppression method to quickly converge the system state, maintain overall stability while ensuring regional autonomy, and achieve an effective balance between distributed control and global optimization.

By building a stable and efficient graph partitioning and regional autonomous system, this study significantly improves the scalability and local response speed of the system while ensuring the stable operation of the county-level energy graph structure. The autonomous subgraphs achieve global perception and local adaptive coordination through boundary nodes, providing a structural basis for the subsequent dynamic perturbation response and the linkage of energy use strategies of the people-enriching industry chain.

In order to enhance the adaptability of the model to different counties, the following mechanisms were introduced in this study:

(1) Dynamic regional division mechanism: automatically adjust regional boundaries based on the geographical characteristics (population density) and load fluctuation rate of the county;

(2) The configurability of the green potential scoring function: the scoring weights can be adjusted according to the carbon emission targets and industry types of different counties;

(3) Incremental update mechanism of graph attention module: supports updating only node embedding vectors without retraining the entire model;

(4) Edge deployment and lightweight processing: The model can be deployed locally in the county, supporting real-time updates and policy migration.

The boundary node synchronization and policy update mechanism mentioned in this article has low computational and communication overhead in practice. The system adopts a boundary message passing algorithm based on message passing, which only triggers synchronization when the node state difference exceeds the preset threshold (5%), significantly reducing communication frequency. The synchronization content mainly includes node power variation, prediction error, and strategy selection probability vector, with a single communication data volume controlled within 1KB. The strategy update adopts a lightweight model parameter synchronization method, with a full update every 12 hours. The edge nodes support local strategy inference, significantly reducing the computing pressure on the central node and ensuring the real-time and scalability of the system.

3. Task-driven Collaborative Evaluation Indicator System

The experimental evaluation in this chapter is based on the actual operation data set of a typical county in China throughout 2023. The data set covers 40 distributed energy nodes and 30 people-enriching industry nodes. The data collection frequency is 5 minutes/time, including key attributes such as dynamic power generation power, load intensity, real-time carbon emission intensity ($kgCO_2/kWh$), and energy consumption per unit output value (kWh/10,000 yuan). 12 groups of typical scenario data are specially included: sunny/cloudy photovoltaic fluctuation scenarios (accounting for 60%), night valley scenarios (25%), and extreme weather disturbance scenarios (15%) to fully verify the robustness of the model. The original data is cleaned by the Kafka-Spark edge computing pipeline and stored in the graph database.

A. Energy-Industry System Carbon Efficiency Index

The carbon efficiency index is evaluated by the level of industrial output value corresponding to unit carbon emissions. First, the energy input and carbon emission intensity parameters of each industrial node are collected and normalized with the output value data within the scheduling cycle. Then, a node carbon efficiency matrix is constructed to aggregate the regional energy consumption and output ratio, reflecting the carbon utilization efficiency level brought by different energy configurations to industrial output. The index is used to compare the green collaboration performance of the system under different strategies.

(a) (b)

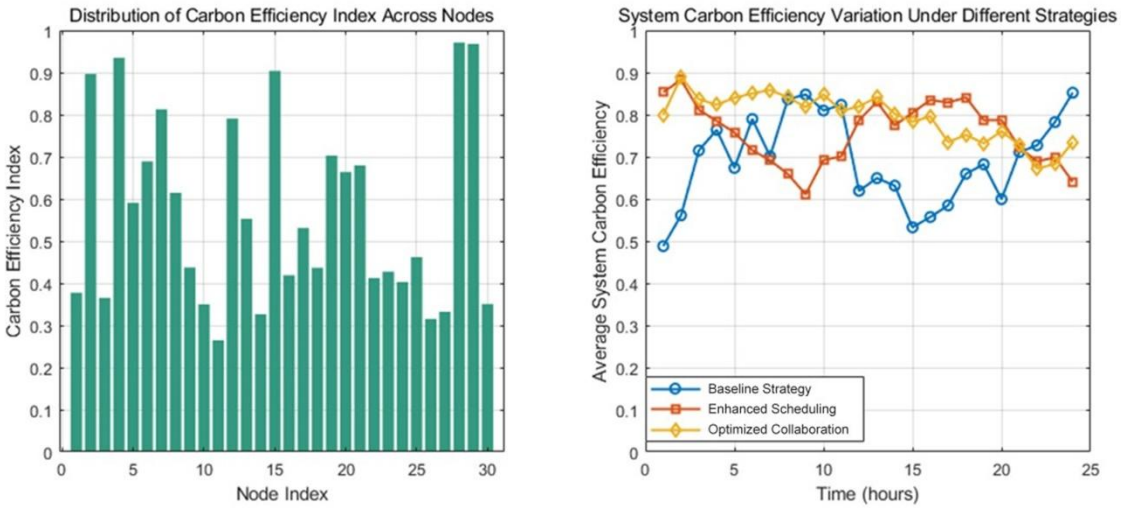

Figure 6. Cross-node carbon efficiency evaluation and scheduling strategy. Figure 6 (a) Cross node carbon efficiency evaluation; Figure 6 (b) Scheduling strategy.

The Figure 6(a) shows the carbon efficiency index of each industrial node. The X-axis is the node number, and the Y-axis is the carbon efficiency value (0-1). Nodes 28 and 29 have the highest values, reaching 0.9711 and 0.9682, respectively, which have significant green potential. The Figure 6(b) shows the changes in the average carbon efficiency of the system under the three scheduling strategies. The optimized collaboration strategy performs best, with a peak value of 0.89, which is better than the baseline strategy overall. The results show that reasonable scheduling can improve the carbon efficiency of the system, and giving priority to supporting efficient green nodes is conducive to achieving industrial low-carbon collaboration.

B. Node Scheduling Response Delay

This indicator is used to quantify the average time from when the node state changes to when the scheduling policy is updated. The evaluation process includes recording the state mutation timestamps of all key nodes in each epoch of scheduling and tracking the execution time triggered by the policy output in the graph network. All response delay data are aggregated to calculate the weighted average response cycle of the node to evaluate the control sensitivity of the system under sudden load or environmental changes.

Table 1. Performance differences of different types of nodes in scheduling response

Node Type	Avg Response Delay (ms)	Max Response Delay (ms)	Weighted Avg Response Delay (ms)
PV Node	120	145	118
Storage Node	95	110	97
Flexible Load Node	80	100	82
Rigid Load Node	130	160	128
Industry Terminal Node	105	130	107

Table 1 shows the performance differences of different types of nodes in scheduling response, and the comparison reflects the sensitivity and worst performance of the system driven by the graph neural network under sudden load or environmental perturbation. The average response delay of the photovoltaic node is 120ms, and the maximum delay is 145ms, which means that although it can quickly trigger the strategy in high-priority scheduling, it occasionally has large delays. The average delay of the energy storage node is 95ms, and the maximum delay is 110ms. The weighted delay is 97ms. The results highlight its stability in second-level response. The average delay of the rigid load node is 130ms, and the maximum is 160ms. The weighted delay is 128ms, reflecting the bottleneck effect of the strategy update of this type of node. The average delay of the industrial terminal node is 105ms, and the maximum is 130ms, indicating that the scheduling constraint of the industrial load is between the two. Overall, the flexible load node has a higher scheduling sensitivity, while the rigid load node is the key constraint point of the system response.

C. Network Collaboration Stability Indicator

The network collaboration stability is evaluated by measuring the fluctuation amplitude of the graph embedding representation after multiple epochs of scheduling iterations. After each epoch of graph network forward propagation, the embedding vector of each node is extracted, and the change of Euclidean distance between consecutive moments is recorded. The mean and variance of the embedding change of all nodes are statistically analyzed to reflect the structural convergence of the system and the consistency of the whole network behavior during the strategy adjustment process, which effectively measures robust performance.

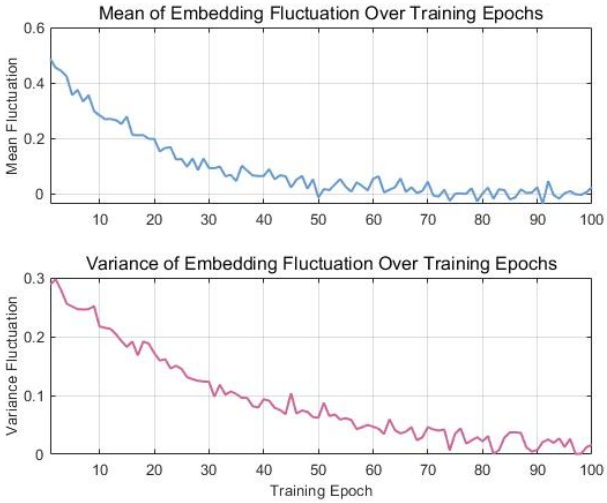

Figure 7. Embedding fluctuation mean and variance during graph neural network training.

Figure 7 shows the evolution trend of the embedding fluctuation mean and variance during the training of the graph neural network, which is used to evaluate the network collaboration stability of the system. The X-axis is the training epoch (1 to 100). The Y-axis of the upper figure represents the node embedding fluctuation mean, and the Y-axis of the lower figure represents the embedding variance. It can be found that the fluctuation mean gradually decreases from the initial approximately 0.5, stabilizes after the 60th epoch, and remains below 0.1, indicating that the node representation gradually converges after multiple epochs of scheduling optimization. The variance indicator also shows a significant downward trend, from the initial 0.3 to within 0.05 after the 80th epoch, indicating that the consistency of the embedding fluctuation of the entire network is enhanced. This result shows that the constructed scheduling strategy has strong robustness and structural convergence in the graph space, supporting the system in achieving stable energy flow collaborative regulation under dynamic perturbations.

D. System Benefit Composite Score

The composite score integrates the results of multiple objective functions, including unit energy cost, carbon

emission fines, and the increase in the output value of each industry. The evaluation method constructs a weighted linear combination based on the normalized score vector. The weight of each dimension is preset according to the county planning goal. For example, when the carbon emission reduction goal is the main one, the weight is tilted towards the carbon indicator. The final score is used for horizontal comparison and priority ranking among scheduling strategies.

Table 2. Comprehensive benefit scores of four scheduling strategies.

Strategy	Cost Score	Carbon Penalty Score	Output Value Score	Cost Weighted (0.3)	Carbon Weighted (0.4)	Revenue Weighted (0.3)	Composite Score
Baseline Strategy	0.6	0.55	0.65	0.18	0.22	0.195	0.595
Collaborative Optimization	0.75	0.8	0.9	0.225	0.32	0.27	0.815
Reinforcement Learning	0.7	0.7	0.85	0.21	0.28	0.255	0.745
Hybrid Strategy	0.65	0.75	0.8	0.195	0.3	0.24	0.735

Table 2 compares the comprehensive benefit scores of the four scheduling strategies, strengthening green collaborative optimization across the three dimensions of cost, carbon emissions, and output value. The baseline strategy is defined as a traditional model predictive control framework, whose objective function only optimizes the scheduling cost and does not explicitly constrain carbon emissions and industry collaboration. The composite score of the baseline strategy is only 0.595, including a cost score of 0.60, a carbon penalty score of 0.55, and an output value score of 0.65. In contrast, the collaborative optimization strategy proposed in this paper scores a high score of 0.80 in the carbon penalty dimension, with a weighted carbon penalty contribution of 0.32, and the final composite score climbs to 0.815, significantly ahead of other solutions. The reinforcement learning strategy and hybrid strategy obtain composite scores of 0.745 and 0.735, respectively, indicating that they are also feasible in balancing costs and carbon emissions, but they are still inferior to the comprehensive driving effect of the collaborative optimization strategy on improving low-carbon output value.

E. Comparison of Energy Configuration Elastic Response

This indicator is used to measure the adaptability of the system to maintain the target output capacity under different energy ratio adjustments. By constructing a set of simulated scheduling scenarios, the main energy types (such as photovoltaics, wind power, and energy storage) are perturbed proportionally, and the changes in system output value and carbon emission response are observed under fixed industrial load conditions.

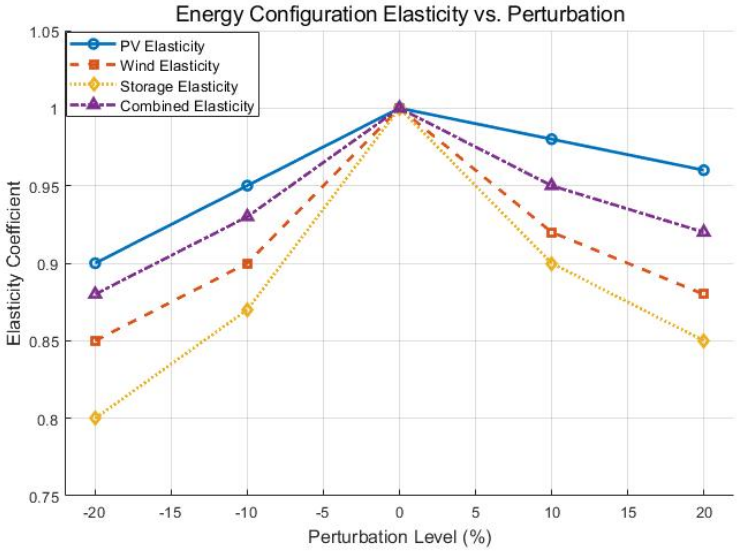

Figure 8. Comparison of elastic response of multi-energy system configurations under different perturbation levels.

Figure 8 is a comparison of the elastic response of multi-energy system configurations under different perturbation levels, where the horizontal axis is ±20% perturbation and the vertical axis is the elasticity coefficient. Figure 8 shows the changes in the elasticity coefficients of photovoltaic, wind power, energy storage, and their combined configurations under different perturbation levels, reflecting the system's

anti-perturbation ability. All types reach the maximum elasticity of 1.00 at 0 perturbation. The elasticity coefficient of the energy storage drops to 0.80 at -20% perturbation, and that of the photovoltaic is 0.90, showing that photovoltaic stability is higher. The elasticity coefficient of the combined configuration maintains 0.93 and 0.95, respectively, under ±10% perturbation, indicating that multi-source collaborative scheduling can significantly improve the system's elasticity and robustness, verifying the effectiveness of the proposed mechanism.

F. Regional Autonomous Collaboration Efficiency

To evaluate the collaboration effect of the multi-scale graph partitioning strategy in actual scheduling, five typical regional subgraphs are selected and quantitatively analyzed from three dimensions: synchronization delay, message frequency, and collaboration success rate, to reveal the response efficiency and collaboration stability of the graph neural scheduling mechanism in regional autonomy.

Table 3. Comparison of regional subgraph autonomous collaboration performance.

Subgraph Region	Synchronization Delay (ms)	Message Frequency (times/min)	Collaboration Success Rate (%)
Subgraph A	150	12	95
Subgraph B	180	10	92
Subgraph C	140	14	97
Subgraph D	165	11	94
Subgraph E	155	13	96

Table 3 shows the collaboration performance indicators of the five subgraphs under the multi-scale region graph partitioning. Overall, subgraph C performs best, with the lowest synchronization delay (140ms) and the highest collaboration success rate (97%), indicating that its graph structure partitioning and policy inference are the most efficient. Subgraph B has problems such as high synchronization delay and low message frequency, and its collaboration success rate is relatively low (92%), indicating that its node relationship density may not be conducive to high-frequency collaboration. The overall data reflects the dynamic adaptability differences of the regional autonomous mechanism under GNN collaborative control, which is helpful to further optimize the graph structure and scheduling algorithm.

G. Model Expansion Performance

The performance under graph scale expansion is shown in Table 4.

Table 4. Performance under Scale Expansion.

Number of nodes in the graph	Average training time (per epoch)	Inference delay (ms)	GPU memory usage (MB)
50	2.3s	68	420
80	3.1s	79	610
100	4.2s	93	850
150	6.8s	125	1200
200	10.5s	168	1800

As the number of nodes in the graph increases, the training time and inference delay show a non-linear increase, but still meet the deployment requirements at the county edge within 100 nodes. For nodes above 150, the graph structure needs to be optimized or a lightweight strategy should be adopted. The overall architecture has certain scalability and is suitable for medium-sized county-level energy system applications.

The results of the ablation experiment are shown in Table 5.

Table 5. Ablation Experiment.

Module combination	Carbon efficiency index	Weighted response delay (ms)
Complete model (GAT+GraphSAGE+PPO)	0.89	97
Remove GAT	0.78	115
Remove GraphSAGE	0.82	105

Removing GAT and GraphSAGE in sequence resulted in a decrease in carbon efficiency index and response delay, indicating that each module made significant contributions to system performance. Among them, GAT plays a key role in graph attention modeling, GraphSAGE plays a crucial role in green potential

mining, and PPO plays a key role in multi-objective scheduling optimization. Module collaboration can achieve optimal performance.

The performance of the model in extreme environmental scenarios such as prolonged cloudy days and system failures is shown in Table 6.

Table 6. Extreme Environmental Scenarios.

Scenario Type	Decrease in photovoltaic output (prolonged cloudy days)	Energy storage system malfunction
Scene Description	The photovoltaic power generation has decreased to 30% of the normal level	Energy storage node crashes or output is limited
Scheduling response delay (ms)	115	142
Carbon Efficiency Index (0-1)	0.81	0.74
Green path availability rate (%)	86%	75%
System recovery time (min)	8	15

In extreme environmental scenarios, the model still exhibits strong robustness and adaptability. For example, in long-term cloudy conditions, carbon efficiency and green path availability remain at a high level, verifying the practicality and fault tolerance of the model under complex disturbances.

The performance comparison of different graph neural network architectures in county-level energy system tasks is shown in Table 7.

Table 7. Performance comparison of different graph neural network architectures in county-level energy system tasks.

Model architecture	Average response delay (ms)	Carbon efficiency index	Comprehensive score
This paper model	97	0.89	0.815
GIN	112	0.83	0.76
Graph Transformers	105	0.86	0.79
GCN	118	0.8	0.735
GraphUNet	125	0.78	0.71

GAT dynamically models the energy coupling relationship between nodes through attention mechanism, improving scheduling response speed and carbon efficiency; GraphSAGE supports efficient aggregation and green potential mining under large-scale graph structures, adapting to the needs of county-level energy industry collaborative optimization. The combination of the two is better than GIN, Graph Transformers and other architectures in response delay, carbon efficiency and comprehensive score, taking into account performance and deployability, and is more suitable for rural edge computing environment.

4. Conclusion

The distributed energy dispatch and county industrial chain collaborative optimization framework based on graph neural network relies on graph attention mechanism, graph cascade optimizer and multi-scale autonomous control method to achieve efficient coordination of energy dynamic allocation and industrial green transformation. Experimental results show that the proposed model performs well in improving energy utilization efficiency, reducing carbon emission intensity and enhancing system stability, and has good practical value and promotion potential. There are still certain limitations in algorithm adaptability, regional difference adaptability and system integration stability. In particular, the generalization ability needs to be further improved in the face of cross-regional dispatch, complex environmental disturbances and long-term dynamic feedback conditions. Although the framework proposed in this paper integrates complex components such as GAT, GraphSAGE and PPO, it achieves functional decoupling through modular design. Each module can be deployed and upgraded independently, reducing the overall maintenance difficulty. For example, the graph attention module only needs to update the node embedding vector regularly without retraining the entire model; the regional autonomous control subgraph can be run locally to reduce dependence on central computing power. In addition, the model parameters have been lightweight processed and verified to be real-time on county-level edge devices. For scenarios with insufficient operation and maintenance capabilities, it is recommended to adopt a pre-trained model + incremental update strategy, which only requires basic IT personnel to complete daily maintenance. This article focuses on interpretability for non-technical personnel in model output design, ensuring that energy managers can understand and trust system recommendations. By introducing standardized green potential scoring, visualization of path scheduling costs, display of regional division mechanisms, and multi-objective comprehensive

scoring mechanisms, the system output not only has technical guidance significance, but also facilitates practical scheduling operations for rural energy managers. The framework proposed in this study has a good modular structure and parameter adjustability, which can adapt to county-level energy systems under different geographical and industrial backgrounds by adjusting the input characteristics of green potential rating weights, regional partitioning strategies, and graph attention mechanisms. In the future, the introduction of federated learning mechanisms can achieve collaborative training across multiple counties, further enhancing the model's generalization ability.

Acknowledgment

None

Data Availability Statement

The county-level distributed energy system and prosperous industry chain operation data used in this study are sourced from the 2023 national typical county-level actual operation dataset, covering 40 distributed energy nodes and 30 prosperous industry nodes. The data collection frequency is 5 minutes per time, including key attributes such as dynamic power generation rate, load intensity, carbon emission intensity, and energy consumption per unit output value. The original data is cleaned by Kafka Spark edge computing pipeline and stored in the graph database. Due to privacy concerns related to local energy infrastructure, the complete raw data has not been made public to the public.

Consent to Publish

The manuscript has neither been previously published nor is under consideration by any other journal. The authors have all approved the content of the paper.

Funding

This work was supported byHezhou University doctoral research startupfund project "Research on the optimization of industries for enrichingpeople in Guangxi karst environment"(2024BSQD12).Guangxi ZhuangAutonomous Region's philosophy and Social Sciences Planning researchproject "Research on the theoretical logic and practical path of greenindustries for enriching people promoting the development of new qualityproductive forces in Guangxi"(24MZF018).

Author Contribution

[Kan Wu]: Developed and planned the study, performed experiments, and interpreted results. Edited and refined the manuscript with a focus on critical intellectual contributions.

[Sishi Xie]: Participated in collecting, assessing, and interpreting the date. Made significant contributions to date interpretation and manuscript preparation.

[Zhanjun Xie]: Provided substantial intellectual input during the drafting and revision of the manuscript.

Conflicts of Interest

The authors declare that they have no financial conflicts of interest.

References

[1] W. Liao, B. Bak-Jensen, J.R. Pillai, Y.L. Wang, Y.S. Wang. A review of graph neural networks and their applications in power systems. Journal of Modern Power Systems and Clean Energy, 2021, 10(2), 345-360. DOI: 10.48550/arXiv.2101.10025

[2] Y. Wu, H.N. Dai, H. Tang. Graph neural networks for anomaly detection in industrial Internet of Things. IEEE Internet of Things Journal, 2021, 9(12), 9214-9231. DOI: 10.1109/JIOT.2021.3094295

[3] H.Y. Wu, M.H. Wang, Z. Xu, Y.W. Jia. Graph attention enabled convolutional network for distribution system probabilistic power flow. IEEE Transactions on Industry Applications, 2022, 58(6), 7068-7078. DOI: 10.1109/TIA.2022.3202159

[4] A. Protogerou, S. Papadopoulos, A, Drosou, D. Tzovaras, L. Refanidis. A graph neural network method for distributed anomaly detection in IoT. Evolving Systems, 2021, 12(1), 19-36. DOI: 10.1007/s12530-020-09347-0

[5] J.H. Cai, W. Liang, X. Li, K.C. Li, Z.W. Gui, M.K. Khan. GTxChain: A secure IoT smart blockchain architecture based on graph neural network. IEEE Internet of Things Journal, 2023, 10(24), 21502-21514. DOI: 10.1109/JIOT.2023.3296469

[6] Y.H. Zhu, Y.Z. Zhou, W. Wei, L.Q. Zhang. Real-time cascading failure isk evaluation with high penetration of renewable energy based on a graph convolutional network. IEEE Transactions on Power Systems, 2022, 38(5), 4122-4133. DOI: 10.1109/TPWRS.2022.3213800

[7] G.M. Dong, M.Y. Tang, Z.Y. Wang, J.C. Gao, S.K. Guo, L.H. Cai. Graph neural networks in IoT: A survey. ACM Transactions on Sensor Networks, 2023, 19(2), 1-50. DOI: 10.1145/3565973

[8] E.E. Kosasih, F. Margaroli, S. Gelli, A. Aziz, N. Wildgoose, A. Brintrup. Towards knowledge graph reasoning for supply chain risk management using graph neural networks. International Journal of Production Research, 2024, 62(15), 5596-5612. DOI: 10.1080/00207543.2022.2100841

[9] E.E. Kosasih, A. Brintrup. A machine learning approach for predicting hidden links in supply chain with graph neural networks. International Journal of Production Research, 2022, 60(17), 5380-5393. DOI: 10.1080/00207543.2021.1956697.

[10] X.K. Zhou, J.Y. Wu, W. Liang, K.I. Wang, Z. Yan, L.T. Yang. Reconstructed graph neural network with knowledge distillation for lightweight anomaly detection. IEEE Transactions on Neural Networks and Learning Systems, 2024, 35(9), 11817-11828. DOI: 10.1109/TNNLS.2024.3389714

[11] Y.X. Shao, H.Z. Li, X.Z. Gu, H.B. Yin, Y.W. Li, X.P. Miao, et al. Distributed graph neural network training: A survey. ACM Computing Surveys, 2024, 56(8), 1-39. DOI: 10.1145/3648358

[12] J.Y. Han, J.X. Wang, Z.H. He, Q. An, Y.Y. Song, A. Mujeeb, et al. Hydrogen-powered smart grid resilience.

Energy conversion and economics, 2023, 4(2), 89-104. DOI: 10.1049/enc2.12083

[13] V.Veerasamy, L.P.M.I. Sampath, S. Singh, H.D. Nguyen, H.B. Gooi. Blockchain-based decentralized frequency control of microgrids using federated learning fractional-order recurrent neural network. IEEE transactions on smart grid, 2023, 15(1), 1089-1102. DOI: 10.1109/TSG.2023.3267503

[14] C. Ju, T.K. Trinh. A Machine Learning Approach to Supply Chain Vulnerability Early Warning System: Evidence from US Semiconductor Industry. Journal of Advanced Computing Systems, 2023, 3(11), 21-35. DOI: 10.69987/JACS.2023.31103

[15] S. Park, W.B. Chen, D. Han, M. Tanneau, P.V. Hentenryck. Confidence-aware graph neural networks for learning reliability assessment commitments. IEEE Transactions on Power Systems, 2023, 39(2), 3839-3850. DOI: 10.1109/TPWRS.2023.3298735

[16] A. Presekal, A. Ştefanov, V.S. Rajkumar, P. Palensky. Attack graph model for cyber-physical power systems using hybrid deep learning. IEEE Transactions on Smart Grid, 2023, 14(5), 4007-4020. DOI: 10.1109/TSG.2023.3237011

[17] H.Y. Lin, M.Y. Yan, X.C. Ye, D.R. Fan, S.R. Pan,W.G. Chen. A comprehensive survey on distributed training of graph neural networks. Proceedings of the IEEE, 2023, 111(12), 1572-1606. DOI: 10.1109/JPROC.2023.3337442

[18] K. Zkik, A. Sebbar, O. Fadi, S. Kamble, A. Belhadi. Securing blockchain-based crowdfunding platforms: an integrated graph neural networks and machine learning approach. Electronic Commerce Research, 2024, 24(1), 497-533. DOI: 10.1007/s10660-023-09702-8

[19] M. Zhang, Z. Zhen, N. Liu, H.J. Zhao, Y.Q. Sun, C.G. Feng. Optimal graph structure based short-term solar PV power forecasting method considering surrounding spatio-temporal correlations. IEEE Transactions on Industry Applications, 2022, 59(1), 345-357. DOI: 10.1109/TIA.2022.3213008

[20] Z.W. Chen, H.B. Ke, J.M. Xu, T. Peng, C.H. Yang. Multichannel domain adaptation graph convolutional networks-based fault diagnosis method and with its application. IEEE Transactions on Industrial Informatics, 2022, 19(6), 7790-7800. DOI: 10.1109/TII.2022.3224988

[21] E.E. Kosasih, A. Brintrup. Towards trustworthy AI for link prediction in supply chain knowledge graph: a neurosymbolic reasoning approach. International Journal of Production Research, 2025, 63(6), 2268-2290. DOI: 10.1080/00207543.2024.2399713

[22] Y. Song, D.Y. Tang, J.S. Yu, Z. Yu, X. Li. Short-term forecasting based on graph convolution networks and multiresolution convolution neural networks for wind power. IEEE Transactions on Industrial Informatics, 2022, 19(2), 1691-1702. DOI:10.1109/TII.2022.3176821

[23] J.S. Kumar, B. Archana, K. Muralidharan, V.S. Kumar. Graph Theory: Modelling and Analyzing Complex System. Metallurgical and Materials Engineering, 2025, 31(3), 70-77. DOI: 10.63278/1320

[24] E. Mohammadi, M. Alizadeh, M. Asgarimoghaddam, X.Y. Wang, M.G. Simões. A review on application of artificial intelligence techniques in microgrids. IEEE Journal of Emerging and Selected Topics in Industrial Electronics, 2022, 3(4), 878-890. DOI: 10.1109/JESTIE.2022.3198504

[25] K. Choudhary, B. DeCost, L. Major, K. Butler, J. Thiyagalingam, F. Tavazza. Unified graph neural network force-field for the periodic table: solid state applications. Digital Discovery, 2023, 2(2), 346-355. DOI:10.1039/D2DD00096B

[26] T. Mortlock, D. Muthirayan, S.Y. Yu, P.P. Khargonekar, M.A.A. Faruque. Graph learning for cognitive digital twins in manufacturing systems. IEEE Transactions on Emerging Topics in Computing, 2021, 10(1), 34-45. DOI: 10.1109/TETC.2021.3132251

[27] Z.G. Liu, P. Qian, X.Y. Wang, Y. Zhuang, L. Qiu, X. Wang. Combining graph neural networks with expert knowledge for smart contract vulnerability detection. IEEE Transactions on Knowledge and Data Engineering, 2021, 35(2), 1296-1310. DOI: 10.1109/TKDE.2021.3095196

[28] C. Gao, Y. Zheng, N. Li, Y.F. Li, Y.R. Qin, J.H. Piao, et al. A survey of graph neural networks for recommender systems: Challenges, methods, and directions. ACM Transactions on Recommender Systems, 2023, 1(1), 1-51. DOI: 10.1145/3568022.

[29] B.N. Huang, Y.S. Li, F.N. Zhan, Q.Y. Sun, H.G Zhang. A distributed robust economic dispatch strategy for integrated energy system considering cyber-attacks. IEEE Transactions on Industrial Informatics, 2021, 18(2), 880-890. DOI: 10.1109/TII.2021.3077509.

[30] E.E. Kosasih, E. Papadakis, G. Baryannis, A. Brintrup. A review of explainable artificial intelligence in supply chain management using neurosymbolic approaches. International Journal of Production Research, 2024, 62(4), 1510-1540. DOI: 10.1080/00207543.2023.22816631

[31] S.W. Wu, F. Sun, W.T. Zhang, X. Xie, B. Cui. Graph neural networks in recommender systems: a survey. ACM Computing Surveys, 2022, 55(5), 1-37. DOI: 10.1145/3535101

[32] A.M. Kettner, L. Reyes-Chamorro, J.K.M. Becker, Z.X. Zou, M. Liserre, M. Paolone. Harmonic power-flow study of polyphase grids with converter-interfaced distributed energy resources—Part I: Modeling framework and algorithm. IEEE Transactions on Smart Grid, 2021, 13(1), 458-469. DOI: 10.1109/TSG.2021.3120108

[33] L. Alrahis, S. Patnaik, M.A. Hanif, M. Shafique, O. Sinanoglu. PoisonedGNN: Backdoor Attack on Graph Neural Networks-based Hardware Security Systems. IEEE Transactions on Computers, 2023, 72(10), 2822-2834. DOI: 10.1109/TC.2023.3271126