An AI-Driven Multi-Scale Graph Neural Network For Maintenance - Oriented Charging Demand Forecasting In Cyber - Physical Systems

Haixing Guo

State Grid Shaanxi Information and Telecommunication Company, Xi'an, Shaanxi, China.

Corresponding author. E-mail: guohaixing0111@163.com

Received: Jun. 9, 2026; Accepted: Jun. 26, 2026

As a core part of Cyber-Physical Systems (CPS), EV charging networks face limitations in traditional spatiotemporal models due to static spatial representation and poor heterogeneous data fusion. This paper proposes a Multi-Scale Subgraph Spatiotemporal Graph Convolutional Network (MSG-STGCN) for predictive load management. By integrating multi-modal data - including POI semantics and traffic conditions - into a dynamic heterogeneous graph, the model utilizes a condition-triggered mechanism to adaptively capture multi-scale spatial contexts. Experiments on real-world datasets show MSG-STGCN achieves MAE of 0.22 and MAPE of 9.75%, outperforming the baseline AGCRN by 0.07 in MAE and 0.81% in MAPE. Ablation studies further verify the effectiveness of multi-scale subgraph embedding and spatiotemporal attention.

Keywords: Cyber-Physical Systems, AI-Driven Predictive Maintenance, Spatiotemporal Graph Neural Networks

http://dx.doi.org/10.6180/jase.202610_33.059

1. Introduction

In the era of Smart Manufacturing and Industrial IoT, the electric vehicle (EV) network has evolved into a critical, large-scale distributed energy system. Under the "dual carbon" strategy, the surge in EVs has introduced complex, high-intensity operational loads. Efficiently managing this infrastructure is no longer just a logistics challenge but a necessity for maintaining the stability and reliability of the urban energy grid. As a vital component of cyber-physical systems (CPS), the EV charging network requires advanced AI-driven insights to ensure its structural integrity and operational continuity. Beyond operational continuity, charging infrastructure increasingly requires predictive maintenance capabilities. Frequent load fluctuations, localized overload conditions, and demand surges can accelerate equipment degradation, reduce service availability, and increase maintenance costs. Therefore, accurate charging-demand forecasting is not only a prediction task but also a critical enabler of proactive maintenance scheduling, resource allocation, and infrastructure resilience enhancement in smart manufacturing-oriented cyber-physical systems.

Predicting spatiotemporal charging demand is the cornerstone of proactive system maintenance and resource scheduling. However, traditional modeling approaches often struggle with the nonlinear, non-Euclidean nature of urban layouts. While existing Graph Neural Networks (GNNs) capture some spatial interactions, they typically rely on static, single-scale representations. This lack of adaptability prevents systems from responding to the dynamic stress caused by fluctuating demand patterns, leading to potential service failures or grid imbalances in smart urban manufacturing ecosystems. The primary challenge lies in the multi-scale nature of system influences-from micro-level traffic flows to regional socioeconomic dynamics. These heterogeneous data sources (e.g., POI semantics, traffic conditions, and historical loads) act as multi-modal sensor inputs. Current models often fail to capture the complex coupling between these diverse data streams and the dynamic spatial structure, limiting the system's cognitive perception of potential load risks. Furthermore, most existing studies primarily optimize forecasting accuracy while paying limited attention to maintenance-oriented decision support. The relationship between demand prediction and predictive maintenance remains insufficiently explored, particularly regarding how forecasting results can assist in identifying potential overload risks and maintenance-critical regions within charging infrastructures.

To address these issues, we propose MSG-STGCN, which adopts a condition-triggered adaptive scale perception mechanism, a lossless heterogeneous feature integration method and a joint spatiotemporal modeling architecture to capture spatial contexts and identify key nodes and peak-load periods. Experimental results on real-world datasets demonstrate that our framework significantly enhances forecasting robustness and system resilience, providing a sophisticated AI tool for the optimized operation and predictive management of modern smart systems [1]. In summary, our contribution is as follows:

- A maintenance-oriented heterogeneous graph framework is constructed to fuse charging behaviors, traffic conditions, POI semantics and environmental factors into a unified CPS perception model.
- An adaptive multi-scale spatial perception mechanism is proposed to describe urban functional regions for diverse operation and maintenance scenarios.
- The MSG-STGCN architecture is developed to enable load management, proactive maintenance scheduling and infrastructure resilience via reliable demand forecasting.

2. Related work

2.1. Factors Influencing System Operational Loads

The operational load of EV charging infrastructure is driven by a convergence of intrinsic system states and extrinsic environmental variables. From an intrinsic perspective, the stochastic nature of user behavior and vehicle specifications introduces significant uncertainty into the system's demand patterns [2–4]. Extrinsic factors have been analyzed through the lenses of infrastructure configuration, temporal cycles, and spatial semantics. At the component level, the technical specifications of charging piles and pricing strategies directly modulate user interaction with the network [5, 6]. Temporally, cyclic shifts such as seasonal weather and peak periods interfere with the load distribution [7]. Spatially, while prior studies linked demand to urban functional zones [8], they often relied on coarse classifications. There remains a lack of fine-grained modeling of point-of-interest (POI) semantics, which are essential for characterizing the heterogeneous functional roles of different nodes within the smart infrastructure. Furthermore, these systems exhibit dynamic spatiotemporal dependencies. The diffusion of load across the network is governed by service radii [9] and real-time traffic conditions [10].

Recent studies in predictive maintenance have further demonstrated that persistent load concentration and abnormal charging patterns can serve as early indicators of infrastructure deterioration. Consequently, integrating operational load forecasting with maintenance-aware monitoring mechanisms has become an important research direction for intelligent infrastructure management. Despite these insights, two limitations persist: 1) current models lack a synergistic integration mechanism for heterogeneous data sources, failing to exploit the deep interactions between temporal, spatial, and environmental inputs; 2) spatial dependencies are typically modeled via static graph structures, which cannot adapt to the dynamic stress shifts inherent in evolving urban manufacturing environments.

2.2. Evolution of Predictive Modeling in Smart Systems

Charging demand forecasting has evolved from overall load analysis to real-time node-level prediction [11–14]. Existing methodologies are categorized into model-driven and data-driven approaches: **(1)** Model-driven methods often utilize Monte Carlo simulations to estimate demand distributions [15]. While computationally efficient, their reliance on simplified assumptions limits their applicability in complex, real-world cyber-physical systems. **(2)** Data-driven methods learn patterns directly from historical streams. Parametric models (e.g., SVR) are interpretable but struggle with non-stationary and nonlinear behaviors [16, 17]. Conversely, non-parametric models like RNNs offer stronger nonlinear fitting [18, 19] but often neglect the crucial spatial topology of the network.

Recently, Spatiotemporal GNNs have emerged as a powerful tool for modeling the spatial correlations between interconnected system nodes [20, 21]. More recently, transformer-based forecasting architectures, such as PatchTST, STAEformer, and PDFormer have demonstrated superior long-range temporal dependency modeling capabilities. Although these models achieve promising forecasting performance, they generally overlook fine-grained urban semantic structures and maintenance-related spatial dependencies, limiting their applicability in infrastructure-oriented predictive management scenarios. Moreover, existing graph models adopt static single-scale representations and cannot fuse multi-source data effectively, which limits their application in AI-driven predictive maintenance.

3. Preliminaries

Definition 1 (Graph Structure Definition). Let $\mathrm{G} = (\mathrm{V}, \mathrm{E}, \mathrm{A})$ be the heterogeneous graph of charging networks. V is the set of charging stations, E stands for edges between stations, and $\mathrm{A} \in \mathrm{R}^{\mathrm{N} \times \mathrm{N}}$ is the weighted adjacency matrix. Nodes and edges represent charging stations and their spatial/functional dependencies respectively.

Definition 2 (Multi-Scale Subgraph Embedding). For a central node vi $\in$ V, a radiusdependent subgraph is defined as $\mathrm{Gr}(\mathrm{vi}) = (\mathrm{Vr}, \mathrm{Er})$, where Vr denotes the neighboring nodes located within radius r and Er represents the corresponding edge set. Three perception scales are considered, namely $\mathrm{r} \in \{500\ \mathrm{m}, 1000\ \mathrm{m}, 2000\ \mathrm{m}\}$. Each subgraph contains node attributes including historical load, POI features, traffic and environmental attributes.

Definition 3 (Charging Demand Prediction). Let $\mathrm{X_t} \in \mathrm{R}^{\mathrm{N} \times \mathrm{F}}$ denote the node feature matrix at time step t, where N is the number of charging stations and F is the feature dimension. Given a historical observation window $\{\mathrm{X}_{\mathrm{t-T+1}}, \ldots, \mathrm{X_t}\}$ and the corresponding multi-scale graph representation, the objective is to learn a mapping function $f(\cdot)$ that predicts the charging demand vector at the next time step:

$$\hat{X}_{t+1} = f\left(G^r; [X_{t-T+1}, \ldots, X_t]\right) \tag{1}$$

where $\mathrm{X}_{\mathrm{t+1}} \in \mathrm{R}^{\mathrm{N}}$ denotes the predicted charging demand of all stations.

Definition 4 (Maintenance Risk Indicator). Given the predicted charging demand $\mathrm{Y}_{\mathrm{t+1}}$, the maintenance risk level of station vi is defined as $R_i = f(X_i, t+1)$. where R_i represents the predicted operational stress level of the station. Larger values of $\mathrm{R_i}$ indicate a higher probability of overload conditions, accelerated equipment degradation, and maintenance intervention requirements.

4. Methods

4.1. Graph Structure Construction and Spatial Connectivity Modeling

To describe spatial correlations and urban semantic relationships between charging stations, we build a heterogeneous graph integrating multi-source spatial data. Different from traditional geography-only static graphs, the proposed graph incorporates navigation distance, POI distribution, traffic accessibility and external environmental information.

Firstly, based on the urban road network, the navigation distance d_{ii} between charging stations and (v_i, v_i) is obtained by using the API of Autonavi Maps, and the spatial adjacency matrix is constructed by using the Gaussian kernel function:

$$A_{ij}^{(dist)} = \exp\left(-\frac{d_{ij}^2}{\sigma^2}\right), \quad i \neq j \tag{2}$$

Here, σ represents the hyperparameter that controls the extent of distance attenuation. This adjacency matrix provides the basic topological support for the spatial structure. The bandwidth parameter σ is determined using the median-distance heuristic. Specifically, σ is set as the median navigation distance among all station pairs in the training set. It avoids manual selection of spatial influence range and realizes reasonable distance attenuation.

Next, POI data is adopted to supplement spatial semantics beyond geographical relations. Taking the charging station as the center, calculate the distribution density of various types of POI within the radius range of $r_k \in \{500\ \mathrm{m}, 1\ \mathrm{km}, 2\ \mathrm{km}\}$. Obtain the scale-dependent POI feature vector $\mathbf{p}_i^{(r_k)}$. Based on the cosine similarity between POI feature vectors, construct the POI semantic adjacency matrix:

$$A_{ij}^{(POI, r_k)} = \frac{\mathbf{p}_i^{(r_k)} \cdot \mathbf{p}_j^{(r_k)}}{\left\|\mathbf{p}_i^{(r_k)}\right\| \left\|\mathbf{p}_j^{(r_k)}\right\|} \tag{3}$$

Considering the influence of road traffic, we use the real-time traffic flow and average vehicle speed information of main roads to calculate the traffic accessibility index and construct the dynamic traffic adjacency matrix $A^{(\text{traffic})}$:

$$A_{ij}^{(\text{traffic})} = \mathrm{Traffic}\left(v_i, v_j\right) \tag{4}$$

$\mathrm{Traffic}\left(v_i, v_j\right)$ is the weighted accessibility function between stations, which reflects traffic conditions. To integrate heterogeneous spatial dependencies into a unified graph representation, the three adjacency matrices are fused through a weighted aggregation strategy: A_fusion $= \lambda$_d $*$ A_dist $+ \lambda$_p $*$ A_POI $+ \lambda$_t $*$ A_traffic. where A_dist, A_POI, and A traffic denote the distance-based, POI-semantic, and traffic-accessibility adjacency matrices, respectively. λ_d, λ_p, and λ_t are nonnegative weighting coefficients satisfying: λ_ d, λ p and λ_t are nonnegative weights with λ_d $+ \lambda$_p $+ \lambda$_t $= 1$. They are uniformly initialized and optimized via backpropagation during training. The fusion retains geographical proximity, functional similarity and traffic accessibility in graph representation. The fused adjacency matrix is normalized as A_norm = D^{−1/2} (A fusion + I) D^{−1/2} (D :

degree matrix, I: identity matrix) to avoid numerical instability, noise amplification and excessive information propagation in graph propagation. Each node integrates multidimensional features: historical charging load, multi-scale POI density, traffic status and environmental factors (weather, electricity price). The graph structure and node attributes form the input tensor $\mathbf{X} \in \mathbb{R}^{N \times T \times C}$ for multi-scale subgraph embedding and full-graph modeling.

4.2. Multi-Scale Subgraph Generation and Embedding

We propose a multi-scale subgraph embedding mechanism to strengthen the representation of spatial semantics. It builds neighborhood subgraphs with different radii around each station to extract multi-scale urban functional features, which are embedded into node representations for subsequent full-graph modeling. Specifically, given the full graph $\mathrm{G} = (\mathrm{V}, \mathrm{E})$, for any node $v_i \in V$, define its subgraph under the radius $r_k \in \{500\ \mathrm{m}, 1\ \mathrm{km}, 2\ \mathrm{km}\}$:

$$G_i^{r_k} = \left(V_i^{r_k}, E_i^{r_k}\right), \quad \text{where } V_i^{r_k} = \left\{v_j \in V \mid d_{ij} \leq r_k\right\} \quad (5)$$

Each subgraph corresponds to a local functional area at a specific spatial scale. The condition-triggered mechanism adopts a data-driven gating network and lightweight MLP. It estimates the probability of three predefined radii ($500\ \mathrm{m}, 1\ \mathrm{km}, 2\ \mathrm{km}$) based on environmental and temporal features, and optimizes scale thresholds automatically during training. Combined with user charging behavior, three practical rules are defined: users select stations within 2 km in peak hours, nearby stations within 500 m in bad weather, and 1 km range stations when obvious price gaps exist. On each subgraph $G_i^{r_k}$, the model uses graph convolution to extract the multi-scale spatial semantic embedding of node v_i :

$$\mathbf{h}_i^{r_k} = \mathrm{GCN}^{(k)}\left(G_i^{r_k}, \mathbf{X}_{r_k}\right) \quad (6)$$

Here, X_{rk} represents the corresponding node feature of the subgraph. This step extracts multi-scale POI structural features for each station, helping nodes understand the urban functional environment they are in. To unify the feature representations at multiple scales, we introduce a scale-level attention mechanism to fuse embeddings at different scales:

$$\mathbf{h}_i = \sum_{k=1}^{K} \alpha_k \cdot \mathbf{h}_i^{r_k}, \quad \alpha_k = \frac{\exp\left(\mathbf{w}^\top \tanh\left(W_k \mathbf{h}_i^{r_k}\right)\right)}{\sum_{m=1}^{K} \exp\left(\mathbf{w}^\top \tanh\left(W_m \mathbf{h}_i^{r_m}\right)\right)} \quad (7)$$

Here, α_k represents the importance weight of scale r_k . The fused feature $\mathbf{h}_i$ is then input into the full graph GCN model as the spatial representation of node $\mathrm{v_i}$. Subgraph-based embedding better captures multi-scale urban functional features and provides sufficient prior information for charging demand forecasting. The scale selection follows domain-based rules.

All multi-modal data are collected hourly, synchronized via timestamps, and missing values are filled by linear interpolation.

4.3. Design of Multi-Scale Graph Convolution Module

Urban charging networks show strong spatial heterogeneity, and single-scale modeling fails to capture complex spatial semantics. We adopt a "local embedding first, then full-graph modeling" paradigm. Here, D denotes the latent embedding dimension. For each station, we build three subgraphs ($500\ \mathrm{m}, 1\ \mathrm{km}, 2\ \mathrm{km}$) based on navigation distance $\mathrm{d_{ij}}$. Apply the graph convolution operator on each subgraph to encode the local structure of node $\mathrm{v_i}$:

$$\mathbf{h}_i^{(r_k)} = \sigma\left(\sum_{j \in \mathcal{N}_i^{(r_k)}} \frac{1}{\sqrt{d_i d_j}} W^{(r_k)} \mathbf{x}_j\right), \quad k = 1, 2, 3 \quad (8)$$

Here, $\mathrm{x_j}$ represents the initial input features of adjacent nodes (including POI, traffic, load, etc.), $W^{(r_k)}$ is the scale dependent weight matrix, $\sigma(\cdot)$ is the nonlinear activation function, and $\mathrm{d_i}$ is the node degree. Graph convolution uses mean aggregation, followed by linear transformation with Wr and ReLU activation. Considering that semantic contributions of different scales have dynamic differences in different scenarios, this paper further introduces a scale-level attention mechanism to integrate multi-scale spatial semantics:

$$\mathbf{h}_i = \sum_{k=1}^{\infty} \alpha_k \cdot \mathbf{h}_i^{(r_k)}, \quad \alpha_k = \frac{\exp\left(\mathbf{w}^\top \tanh\left(W_k \mathbf{h}_i^{(r_k)}\right)\right)}{\sum_{m=1}^{3} \exp\left(\mathbf{w}^\top \tanh\left(W_m \mathbf{h}_i^{(r_m)}\right)\right)} \quad (9)$$

Here, α_k represents the importance weights under the subgraph r_k, which are automatically learned by the attention network. The $\mathrm{h_i}$ obtained through fusion is the multiscale static spatial representation of node v_i, which will subsequently be injected as input into the overall urban map structure to participate in the dynamic modeling of the entire map. Attention coefficients α_k. reflect the relevance of each spatial scale to forecasting. This multi-scale design improves the model's ability to perceive regional differences.

4.4. Spatiotemporal Attention Mechanism Modeling

Charging load has obvious spatio-temporal coupling properties. We adopt a joint spatiotemporal attention mechanism in full-graph modeling to focus on key spatial regions and time periods. **(1) Modeling of Spatial attention mechanism.** Traditional graph convolution uses static adjacency matrices with fixed aggregation matrices. We add spatial attention to adaptively adjust node weights and highlight influential nodes. For node representations $\mathrm{H} \in \mathrm{R}^{\mathrm{N} \times \mathrm{D}}$, query, key, and value matrices are generated as: $\mathrm{Q} = \mathrm{HW_Q}, \mathrm{K} = \mathrm{HW_K}, \mathrm{V} = \mathrm{HW_V}$, where $\mathrm{W_Q}, \mathrm{W_K}$, and $\mathrm{W_V}$ are learnable projection matrices. Given the embedding features $\mathrm{h_i}$ and $\mathrm{h_j}$ of the node for $(\mathrm{v_i}, \mathrm{v_j})$, define their spatial attention weights as:

$$S_{ij} = \frac{\exp\left(\mathrm{LeakyReLU}\left(\mathbf{a}^\top\left[W_q h_i \| W_k h_j\right]\right)\right)}{\sum_{j \in \mathcal{N}(i)} \exp\left(\mathrm{LeakyReLU}\left(\mathbf{a}^\top\left[W_q h_i \| W_k h_j\right]\right)\right)} \tag{10}$$

Among them, $\mathrm{W_q}$ and $\mathrm{W_k}$ are the query and key mapping weights respectively, **a** is the attention vector, and $[\cdot\|\cdot]$ represents the vector concatenation operation. The attention matrix **S** is used to perform weighted correction on the adjacency structure of graph convolution to achieve spatial path reinforcement. **(2) Modeling of the temporal attention mechanism.** In terms of the time dimension, the model needs to identify which historical time slices have the greatest impact on the results when predicting the current moment load. This paper uses a weighted convergence strategy to dynamically model the time dimension:

$$\alpha_t = \frac{\exp\left(\mathrm{score}\left(H_t, q\right)\right)}{\sum_{s=1}^{T_h} \exp\left(\mathrm{score}\left(H_s, q\right)\right)}, \quad \widetilde{H} = \sum_{t=1}^{T_h} \alpha_t H_t \tag{11}$$

4.5. Overall Architecture of MSG-STGCN

We build the MSG-STGCN framework combining multi-scale subgraph embedding and spatio-temporal attention for charging demand forecasting, as shown in Fig. 1. It inherits the strengths of GNNs and enhances node semantic representation via subgraph-level spatial perception.

Firstly, multi-scale subgraphs around each station are built, and graph convolution encodes POI semantics within 500 m, 1 km and 2 km ranges. The fused static spatial embedding is used as node representation for subsequent dynamic modeling. Next, multilayer graph convolution models static embeddings and dynamic load sequences. The dualchannel attention adaptively adjusts spatial adjacency weights and identifies key time slices to model charging patterns. All modalities are converted into 64 -dimensional latent vectors for balanced multimodal fusion. The framework follows the pipeline of static structure modeling, spatio-temporal interaction and prediction. It improves node embedding quality and multi-source data

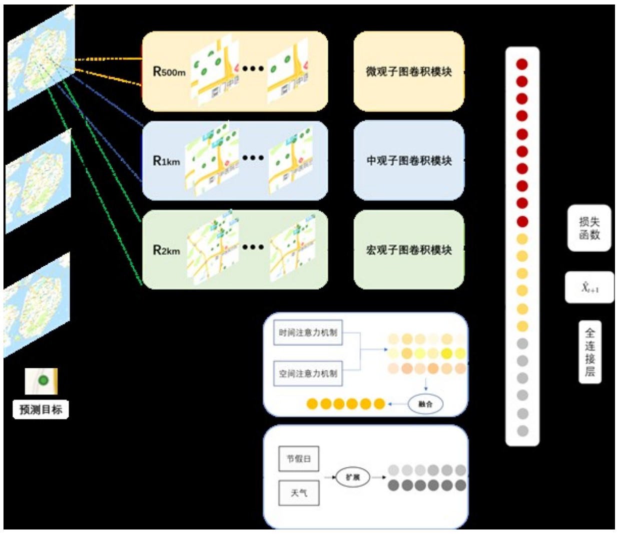

Fig. 1. Framework diagram of MSG-STGCN.

fusion capability for robust and interpretable charging demand forecasting. Given N (station number), E (edge number) and K (spatial scale number), the complexity of multi-scale GCN is $\mathrm{O}(\mathrm{K}|\mathrm{E}|)$ and the attention module adds $\mathrm{O}\left(\mathrm{N}^2\right)$ complexity. The sparse graph ensures feasible deployment for city-scale scenarios.

5. Results and discussion

5.1. tatistical description of the characteristics of the Xiamen dataset.

We adopt a private Xiamen charging dataset covering 2022–2024. Over 600 original stations are filtered, and 547 commercial stations are used for modeling. The dataset consists of charging transactions, inter-station navigation distances, station attributes, meteorological records, time labels and POI distribution data within 3 km of each station. Traffic and meteorological data are sourced from official platforms. All data are synchronized with hourly timestamps. Raw data contains missing values, duplicates and outliers caused by device errors. We remove duplicates and invalid samples, and process load outliers beyond three standard deviations. About 90% of data is retained after cleaning. The statistical characteristics of the Xiamen dataset are shown in Table 1. The site section navigation distance encoding matrix is shown in Fig. 2.

5.2. Data Processing

We process the three-year multi-source data via load aggregation, outlier correction, feature alignment, subgraph encoding and sample construction. All charging records are aggregated into hourly load series, and numerical features are normalized with Z-Score. Temporal and meteorological

Table 1. Statistical description of the characteristics of the Xiamen dataset.

Variable type	Variable name	Data volume	mean	variance
Basic statistics	Number of sites	547	-	-
	Charging transaction volume	10,963,154	0.585	1.345
	Charging fee (yuan per hour)	14,221,910	3.472	0.936
Site attributes	Whether parking charges	547	0.604	0.487
	Fast pile filling number	547	9.412	11.581
	Slow pile filling number	547	4.205	6.873
	Navigation Distance between stations	81,778	21,505.4	11,675.7
Environment	Weekday/holiday labels	26,304	-	-
	Hour label	26,304	-	-
	Temperature	26,304	23.518	5.688

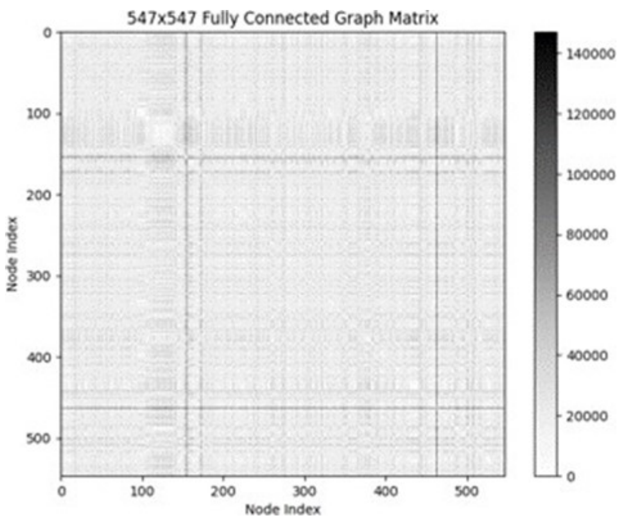

Fig. 2. Charging station node distance matrix.

features are fused, while spatial data is used to build full graphs and multi-scale subgraphs with station attributes encoded. Multi-scale subgraphs are built and encoded via graph convolution, and embeddings are fused to obtain node representations for subsequent modeling. We adopt a 168-hour sliding window: 7-day historical data is used to predict the load of the next 24 hours, which captures temporal periodicity.

5.3. Experimental Setup

The three-year data is divided into training (70%), validation (15%) and test (15%) sets by date. We use Adam optimizer with an initial learning rate of 0.001 , batch size of 64 and maximum 200 epochs. The hidden embedding dimension D is set to 64 , the spatiotemporal attention mechanism utilizes a multi-head configuration with 4 heads, and a dropout rate of 0.2 is applied across all layers. The loss function is Weighted MSE, with load-related weights $w_{i,t}$ to emphasize prediction accuracy in high-load periods and high-density stations, enhancing the model's fitting ability in urban core areas and peak hours. Its form is as follows:

$$\mathcal{L}_{\text{WMSE}} = \frac{1}{N \cdot T} \sum_{i=1}^{N} \sum_{t=1}^{T} w_{i,t} \cdot \left(\hat{y}_i^t - y_i^t\right)^2 \tag{12}$$

Among them, the setting of $w_{i,t}$ is constructed by combining multiple factors such as historical load intensity and time period characteristics. Both the model training and testing are based on the sequential input method of the sliding window: taking the data of each station for 7 consecutive days (168 hours) as input, the charging load on the 8 th day (24 hours) is predicted. The specific network structure and training parameters are shown in Table 2.

All experiments were run on the NVIDIA RTX 3090 GPU platform and implemented using the PyTorch 2.0 framework. The random seed was fixed at 42 to ensure reproducibility. The model occupies 4.2 GB GPU memory during training and has an inference latency below 15 ms per batch, supporting real-time urban deployment. All experiments are repeated five times. We use MAE and MAPE (Eq. (13)) for performance evaluation.

$$\text{MAE} = \frac{1}{N \cdot T} \sum_{i=1}^{N} \sum_{t=1}^{T} \left|\hat{y}_i^t - y_i^t\right|, \ \text{MAPE} = \frac{100\%}{N \cdot T} \sum_{i=1}^{N} \sum_{t=1}^{T} \left|\frac{\hat{y}_i^t - y_i^t}{y_i^t + \epsilon}\right| \tag{13}$$

Here, N represents the number of predicted sites, T is the prediction time step (24 hours), and ϵ is a very small positive number introduced to avoid division by zero errors. The weighted loss function assigns larger penalties to high-demand stations: w_i $= 1 + \alpha^*$L_i, where Li denotes the average historical charging load of station i and α is a tunable weighting coefficient. This strategy encourages the model to prioritize operationally critical stations whose forecasting errors may result in larger maintenance and service risks.

5.4. Comparative Experiment

We compare with eight mainstream baselines, including traditional time-series models and various graph neural

Table 2. Graph neural network model structure and training parameters.

Parameter Classification	Parameter Name	Parameter Setting
Network module parameters	Subgraph convolution module	multi-scale GCN, activation function: ReLU
	Subgraph fusion mechanism	attention-weighted fusion + hierarchical nested
	Full-image convolution module	Single-layer GCN, activation function: ReLU
	Time modeling module	Convolutional gating unit
	Spatio-temporal attention mechanism	Spatio-temporal attention, weight sharing
	Output layer	Prediction: 24 hours, activation function: Linear
Training phase parameters	Batch Size	64
	Epoch	200
	Early Stopping	15
	Optimizer	Adam
	Learning Rate	0.001
	Loss Function	Weighted mean square error

Table 3. Comparison of the prediction performance of different models.

Model	MAE	MAPE (%)
ARIMA	0.57 ± 0.02	18.52 ± 0.03
SVR	0.52 ± 0.01	16.01 ± 0.02
LSTM	0.44 ± 0.01	13.44 ± 0.005
GCN	0.43 ± 0.02	12.58 ± 0.01
STGCN	0.35 ± 0.01	11.77 ± 0.01
ASTGCN	0.33 ± 0.005	10.98 ± 0.005
AGCRN	0.29 ± 0.01	10.56 ± 0.005
MSG-STGCN (ours)	0.22 ± 0.005	9.75 ± 0.005

networks. All models are tested under identical experimental settings, and results are presented in Table 3. MSG-STGCN achieves the lowest MAE and MAPE among all methods. Compared with AGCRN, it reduces MAE by 0.07 and MAPE by 0.81%. Traditional models lack spatial modeling capability, while existing GNNs cannot fully capture multi-scale spatial semantics and urban POI features. The results verify the superior accuracy and robustness of our model for charging demand forecasting.

5.5. Ablation Experiment

We conduct ablation experiments to validate the effectiveness of two core modules: Spatio-Temporal Attention (STA) and Multi-scale Subgraph Embedding (MSSE). Three model variants are built: w/o STA, w/o MSSE, and w/o both modules (degenerated to STGCN). All variants are tested under the same settings, with results shown in Table 4.

Prediction performance degrades after removing any module. The model suffers the largest accuracy drop when both modules are eliminated. This demonstrates that multi-scale subgraph embedding and spatio-temporal attention jointly enable the model to capture complex spatial and temporal characteristics, and the two modules are indispensable for improving prediction adaptability and generalization.

Table 4. Experimental Results of model structure ablation (Xiamen Dataset).

Model structure	MAE	MAPE (%)
MSG-STGCN	0.22	9.75
w/o STA	0.31	10.85
w/o MSSE	0.28	11.42
w/o MSSE & STA	0.35	11.77

5.6. Discussion

Charging demand forecasting provides practical guidance for predictive maintenance, resource scheduling and risk management of charging infrastructures. Benefiting from multi-scale spatial modeling and heterogeneous data fusion, MSG-STGCN can track dynamic load changes. It helps identify high-risk stations for preventive maintenance, locate key maintenance areas and optimize resource allocation, so as to improve infrastructure resilience and reduce maintenance costs. Specifically, the predicted charging demand Yt+1 is mapped to the operational stress index Ri by calculating the ratio of predicted load to the transformer station's rated safety capacity. When Ri exceeds a critical threshold, the system automatically triggers a high-risk alert, enabling operators to shift load or dispatch maintenance crews to inspect aging charging piles before actual hardware failures occur.

6. Conclusion

This study presented MSG-STGCN, a maintenance-oriented spatiotemporal graph learning framework for charging-demand forecasting in large-scale charging infrastructures. By integrating heterogeneous urban information and adaptive multi-scale spatial perception, the proposed method effectively captures complex spatial interactions and temporal demand dynamics, and outperforms mainstream forecasting models in experiments. The proposed framework achieved the lowest forecasting errors across all evaluation metrics and maintained stable performance under peak-demand, holiday, and extreme-weather scenarios. Beyond forecasting accuracy, the proposed framework provides actionable information for predictive maintenance and operational management. Accurate demand prediction enables early identification of overload risks, facilitates preventive maintenance scheduling, and supports resource allocation decisions for charging-service operators. Despite its effectiveness, the proposed framework has several limitations. First, the experiments were conducted using data collected from a single city, which may limit generalizability to other urban environments. Second, the adaptive scale-selection strategy currently relies on rule-based triggering conditions derived from domain knowledge. Third, the proprietary nature of the dataset restricts complete public accessibility and independent replication.

Future research will focus on four directions. First, graph-transformer architectures will be investigated to further enhance long-range spatiotemporal dependency modeling. Second, reinforcement-learning-based maintenance decision optimization will be explored to directly connect forecasting outputs with maintenance scheduling actions. Third, cross-city transfer-learning strategies will be developed to improve model generalization across heterogeneous urban environments. Finally, online learning mechanisms will be introduced to support real-time adaptation under continuously evolving operational conditions.

Acknowledgments

This paper is supported by the Science and Technology Project of State Grid Shaanxi Electric Power Co., LTD. (5226TX25000H).

References

[1] Z. Lv, M. Wu, Z. Song, and S. Zhou, (2022) *"Special Issue: Operation and control of Modular DC Transformer Applied to Energy Internet"* **Mechatron. Syst. Control. *50***: URL: https://api.semanticscholar.org/CorpusID:247170808.

[2] Y. Bing, W. Lifang, and L. Chenglin, (2013) *"Large-scale Electric Vehicle Charging Demand and Its Influencing Factors"* **Transactions of the China Electrotechnical Society**: 7. DOI: 10.19595/j.cnki.1000-6753.tces.2013.02.003.

[3] P. Adsule and M. Manoj, (2025) *"A stated choice study to assess charging and travel decisions of electric car users considering car attributes and trip chaining complexity – Case New Delhi, India"* **Transportation Research Part A: Policy and Practice *192***: 104366. DOI: 10.1016/j.tra.2024.104366.

[4] J. Huber, D. Dann, and C. Weinhardt, (2020) *"Probabilistic forecasts of time and energy flexibility in battery electric vehicle charging"* **Applied Energy *262***: 114525. DOI: 10.1016/j.apenergy.2020.114525.

[5] Y. Kim and S. Kim, (2021) *"Forecasting Charging Demand of Electric Vehicles Using Time-Series Models"* **Energies *14***(5): DOI: 10.3390/en14051487. URL: https://www.mdpi.com/1996-1073/14/5/1487.

[6] Y. Zhang, X. Luo, Z. Wang, Q. Mai, and R. Deng, (2022) *"Estimating charging demand from the perspective of choice behavior: A framework combining rule-based algorithm and hybrid choice model"* **Journal of Cleaner Production *376***: 134262. DOI: 10.1016/j.jclepro.2022.134262.

[7] Q. Xing, Z. Chen, Z. Zhang, X. Huang, Z. Leng, K. Sun, Y. Chen, and H. Wang, (2019) *"Charging Demand Forecasting Model for Electric Vehicles Based on Online Ride-Hailing Trip Data"* **IEEE Access 7**: 137390–137409. DOI: 10.1109/ACCESS.2019.2940597.

[8] Y. Xiong, J. Gan, B. An, C. Miao, and A. L. C. Bazzan, (2018) *"Optimal Electric Vehicle Fast Charging Station Placement Based on Game Theoretical Framework"* **IEEE Transactions on Intelligent Transportation Systems *19***(8): 2493–2504. DOI: 10.1109/TITS.2017.2754382.

[9] G. Alface, J. C. Ferreira, and R. Pereira, (2019) *"Electric Vehicle Charging Process and Parking Guidance App"* **Energies *12***(11): DOI: 10.3390/en12112123.

[10] L. Zenglu, G. Qiang, and N. Jiajia, (2020) *"Research on Online Promotion Strategies of Dual-Channel Manufacturers Based on Risk Avoidance"* **Chinese Journal of Management Science *28***(07): 112–121. DOI: 10.16381/j.cnki.issn1003-207x.2020.07.011.

[11] Z. Xue and S. Li, (2006) *"Multi-Model Modelling and Predictive Control Based on Local Model Networks"* **Control. Intell. Syst. *34***: URL: https://api.semanticscholar.org/CorpusID:20880051.

[12] S. Rahman and G. Shrestha, (1993) *"An investigation into the impact of electric vehicle load on the electric utility distribution system"* **IEEE Transactions on Power Delivery *8***(2): 591–597. DOI: 10.1109/61.216865.

[13] Z. Wu, J. Hu, X. Ai, and G. Yang, (2021) *"Data-driven approaches for optimizing EV aggregator power profile in energy and reserve market"* **International Journal of Electrical Power & Energy Systems *129***: 106808. DOI: 10.1016/j.ijepes.2021.106808.

[14] M. Sedraoui, (2008) *"Application of the multivariable predictive control on a distillation column using the optimization methods"* **Control Intell. Syst. *36***(2): 111–118. URL: https://dl.acm.org/doi/10.5555/1739794.1739795.

[15] J. H. Moon, H. N. Gwon, G. R. Jo, W. Y. Choi, and K. S. Kook, (2020) *"Stochastic Modeling Method of Plug-in Electric Vehicle Charging Demand for Korean Transmission System Planning"* **Energies *13***(17): DOI: 10.3390/en13174404.

[16] E. Xydas, C. Marmaras, L. M. Cipcigan, N. Jenkins, S. Carroll, and M. Barker, (2016) *"A data-driven approach for characterising the charging demand of electric vehicles: A UK case study"* **Applied Energy *162***: 763–771. DOI: 10.1016/j.apenergy.2015.10.151.

[17] M. H. Amini, A. Kargarian, and O. Karabasoglu, (2016) *"ARIMA-based decoupled time series forecasting of electric vehicle charging demand for stochastic power system operation"* **Electric Power Systems Research *140***: 378–390. URL: https://api.semanticscholar.org/CorpusID:113824600.

[18] K. L. López, C. Gagné, and M.-A. Gardner, (2019) *"Demand-Side Management Using Deep Learning for Smart Charging of Electric Vehicles"* **IEEE Transactions on Smart Grid *10***(3): 2683–2691. DOI: 10.1109/TSG.2018.2808247.

[19] M. Boulakhbar, M. Farag, K. Benabdelaziz, T. Kousksou, and M. Zazi, (2022) *"A Deep Learning Approach for Prediction of Electrical Vehicle Charging Stations Power Demand in Regulated Electricity Markets: The Case of Morocco"* **Cleaner Energy Systems**: URL: https://api.semanticscholar.org/CorpusID:253391727.

[20] G. Kim, S. Kang, G. Park, and B.-C. Min, (2023) *"Electric vehicle battery state of charge prediction based on graph convolutional network. International Journal of Automotive Technology"* **International Journal of Automotive Technology *24***(6): 1519–1530. DOI: 10.1007/s12239-023-0122-6.

[21] H. J. Kim and M. K. Kim, (2024) *"Spatial-Temporal Graph Convolutional-Based Recurrent Network for Electric Vehicle Charging Stations Demand Forecasting in Energy Market"* **IEEE Transactions on Smart Grid *15***(4): 3979–3993. DOI: 10.1109/TSG.2024.3368419.