A Spectrum Prediction Framework Based on Graph Convolutional Recurrent Neural Network Integrated with Variational Mode Decomposition

Hao SONG [1], Yueshun HE [1], Linlin HE [2], Xiaoyu CAO [1], Yunzhe LIU [1], Jie CHEN [3]

[1] School of Artificial Intelligence and Information Engineering, East China University of Technology, No. 418 Guanglan, 330013 Nanchang, China
[2] School of Software, East China University of Technology, No. 418 Guanglan, 330013 Nanchang, China
[3] School of Geomatics and Spatial Information Engineering, East China University of Technology, No. 418 Guanglan, 330013 Nanchang, China

songhao@ecut.edu.cn, {heys, lilyhe, caoxy, yzliu, codercjie}@ecut.edu.cn

Submitted March 4, 2026 / Accepted May 8, 2026 / Online first May 19, 2026

Abstract. *Wireless spectrum prediction is crucial for dynamic spectrum management and congestion mitigation. However, existing methods often emphasize spatio-temporal feature modeling while neglecting the frequency-domain structure of the spectrum. Moreover, spectrum data inherently exhibits non-stationarity, multi-scale characteristics, and strong spatio-temporal coupling, making it difficult for traditional prediction models to fully capture its complex dynamics, thereby limiting prediction accuracy. To address these issues, this paper proposes a spectrum prediction model that integrates Variational Mode Decomposition (VMD) with a graph convolutional recurrent neural network. First, VMD is employed to adaptively decompose the raw spectrum into multi-scale intrinsic modes, achieving frequency-domain decoupling and noise suppression, thus enhancing the model's sensitivity to spectral details. Second, the Spearman correlation coefficient is utilized to construct a graph topology among frequency bands, and a Graph Convolutional Network (GCN) is applied to extract spatial dependency features, combined with Gated Recurrent Units (GRU) to model temporal evolution patterns. Furthermore, an attention mechanism is introduced to dynamically weight the hidden states, focusing on critical information and improving training efficiency. Experiments on real-world spectrum datasets demonstrate the superior performance of the proposed model, achieving an accuracy of 96.79%, MAE of 0.3342, RMSE of 0.4632, and R^2 of 0.9969.*

Keywords

Spectrum prediction, Variational Mode Decomposition (VMD), Graph Convolutional Network (GCN), Gated Recurrent Unit (GRU), attention mechanism

1. Introduction

With the rapid development of wireless communication technology, the number of connected terminal devices has grown exponentially, and the demand for spectrum resources has been unprecedented. However, as a limited and non-renewable public asset, radio spectrum is becoming increasingly scarce. Therefore, the effective allocation and utilization of these resources have become the key challenges faced by contemporary wireless systems. To alleviate the problem of insufficient spectrum utilization, Mitola introduced the concept of Cognitive Radio (CR) in 1999 [1–3].

By dynamically perceiving the electromagnetic environment, the CR system can intelligently exploit spectrum vulnerabilities, which is a key strategy to alleviate spectrum shortages. Spectrum sensing and spectrum prediction are the fundamental components of this system, with the aim of accurately obtaining the state of the wireless channel. Sensing captures real-time status, while spectrum prediction goes a step further by inferring future channel occupancy status, Power Spectral Density (PSD), and other statistical characteristics based on historical spectrum data [4–7]. This predictive foresight enables the CR system to proactively adjust transmission parameters, thereby minimizing collisions and enhancing overall spectral efficiency.

To achieve accurate spectrum prediction, methods have evolved from traditional statistical approaches to more complex deep learning techniques. Early research mainly relied on statistical regression and probability models, such as Autoregressive Integrated Moving Average (ARIMA) and Hidden Markov Model (HMM) [8–11]. Earlier studies assumed that spectrum occupancy is a stationary process. However, due to the constantly changing wireless environment, these methods perform poorly in highly mobile or dynamic environments and rely heavily on expert knowledge or manual feature engineering. In addition, they are sensitive to environmental noise and unable to simulate complex nonlinear relationships, which limits their application in practical engineering.

Due to its powerful feature extraction capability, deep

DOI: 10.13164/re.2026.0272

learning has become the preferred method for spectrum prediction in the rapid development of artificial intelligence [12], [13]. Although recurrent neural networks (RNNs) were initially used to handle temporal correlations, they have the problem of vanishing gradients [14]. To address this issue, Lindemann et al. implemented a long short-term memory (LSTM) network [15], which significantly enhances prediction accuracy by leveraging historical data. Moving forward, further progress will focus on optimizing the processing of time series. A notable example is the fusion of Singular Spectral Analysis (SSA) and Bidirectional LSTM (BiLSTM) [16], [17], which extracts subsequences representing various components of time series. This hybrid model allocates weights through correlation coefficients, enhancing its adaptability to changing channel environments [18]. Furthermore, in order to improve efficiency, Yuan et al. introduced the Convolutional Gated Recurrent Unit (GRU) model. This method not only maintains a high prediction fidelity, but also accelerates the convergence speed by using fewer parameters compared with the traditional LSTM [16].

However, spectral data not only have temporal continuity but also exhibit significant spatial and frequency-domain correlations. Models that merely focus on the time dimension are often insufficient. To address this issue, researchers have proposed hybrid architectures that integrate Convolutional Neural Networks (CNN) and LSTM to simultaneously capture spatial and temporal features [19], [20]. However, CNN is designed for Euclidean data (grid structure) and may not optimally represent non-Euclidean topological relationships between different frequency bands. GCN have been proven to be significantly superior to CNN in extracting spatial and topological features. Attention Graph Convolutional Recurrent Neural Network (AGCRNN) [21] represents a strong spatiotemporal baseline. Building upon these foundations, recent 2024–2025 European advancements have further extended GNN capabilities in handling complex spatiotemporal and frequency-domain data. Specifically, systematic research demonstrates that spatiotemporal GNNs act as powerful neural operators, effectively capturing dynamic cross-series dependencies through tailored architectural biases [22]. Furthermore, integrating dynamic graph structures has proven critical for model robustness, maintaining prediction accuracy even under severe topology perturbations and environmental noise [23]. Concurrently, contemporary studies have successfully combined GNNs with power spectral density to automatically extract hidden spatial-frequency correlations, highlighting the superiority of message-passing mechanisms in processing frequency-domain signals [24]. Recently, emerging Transformer-based architectures have also revolutionized time-series forecasting. Since Vaswani et al. introduced the foundational self-attention mechanism [25], Transformers have been increasingly adapted for spectrum prediction. For instance, Xu et al. proposed the Dconv-Former [26] to efficiently capture spatial-temporal-spectral features. Furthermore, recent related studies, such as the explainable spectrum prediction approach based on VMD-LSTM [27], have emphasized the critical need for robust models capable of handling severe channel fading and noise. These studies confirm the effectiveness of integrating advanced signal processing techniques with deep learning in this domain.

Despite these global advancements, existing models still encounter considerable obstacles in frequency-domain feature mining. Directly processing raw data creates difficulties for models due to inherent non-stationarity and complex multi-scale features. The primary objective of this paper is to propose a robust forecasting framework capable of explicitly decoupling the non-stationarity of the signal before deep spatiotemporal learning. Compared to previously published works in this domain (such as the aforementioned VMD-LSTM [27], which lacks non-Euclidean spatial modeling), this study essentially advances the field by bridging the gap between advanced signal decomposition and topological deep learning. Specifically, we introduce VMD [28], [29] to decompose the original spectral data into Intrinsic Mode Functions (IMFs). This decomposition effectively separates different frequency components, allowing for noise reduction and smoothing of features through weighted reconstruction. The reconstructed high-quality features are then fed into an Attention-based GCN-GRU framework to model the spatiotemporal dependencies deeply. The main contributions of this paper are as follows:

- To the best of our knowledge, this is the first study to seamlessly integrate Variational Mode Decomposition (VMD) with an Attention-enhanced Graph Convolutional Recurrent Neural Network for spectrum forecasting. By utilizing VMD as a physics-informed preprocessing step, the framework adaptively decouples highly non-stationary raw signals, effectively suppressing high-frequency stochastic noise before deep spatiotemporal feature extraction.
- We design a rigorous graph construction mechanism that replaces traditional linear metrics with the Spearman's rank correlation coefficient. Combined with a global temporal window and a stringent sparsity threshold ($\theta \geq 0.92$), this module accurately captures the hidden, non-linear physical dependencies among diverse frequency bands, providing a highly reliable adjacency matrix for subsequent GCN operations.
- We extensively evaluate the proposed model on two completely distinct, real-world spectrum datasets (spanning Sub-GHz and 5G NR C-band, collected in different geographic locations). The empirical results demonstrate that our framework consistently outperforms existing state-of-the-art baselines, including recent Transformer-based architectures, proving its superior generalization capability and practical deployment feasibility.

The rest of this paper will be organized as follows: Section 2 introduces the design rationale, the overall framework, and the theoretical foundation of the proposed model. In Sec. 3, we detail the architecture, which includes the mathematical formulation of VMD, graph topology

construction, the GCN and GRU modules, and the attention mechanism. Section 4 gives the dataset description and shows all the experimental results and ablation studies. Lastly, Section 5 concludes the paper and suggests some possible directions for further study.

2. Overall System Design

According to the introduction, wireless spectrum data has great non-stationarity and high spatiotemporal correlation. The conventional models that use a single view to extract features tend to neglect the necessity of suppressing noise and performing multi-scale mode decomposition as well as deep spatiotemporal coupling analysis. To address these challenges, this study provides a complete framework including signal decomposition-reconstruction, spatiotemporal modeling, and attention-based prediction. Aimed at overcoming the limitations of existing methods under complicated electromagnetic conditions, the proposed system takes advantage of the collaboration between various special modules. From Fig. 1 we can see that the architecture consists of three different stages: Signal Decomposition Reconstruction, Spatiotemporal Feature Extraction and Prediction Output.

(1) Signal Decomposition and Reconstruction: The model recognizes the non-stationary fluctuation of raw spectrum data and adds a VMD module as a pre-processing unit. Taking advantage of VMD's capability to deal with non-linear signals, the high-dimensional PSD data is adaptively decomposed into several IMFs that have different central frequencies. Then we use a weighted reconstruction method to remove the high-frequency noise components from the signal while keeping its main trend and periodicity, which greatly reduces the learning difficulty for the following neural networks.

(2) Feature extraction and spatiotemporal modeling: In order to correctly obtain the spatial correlations, a spectrum graph topology is built according to Spearman's correlation coefficient. Then a GCN is used to extract spatial features from the reconstructed multi-modal data. GCN can effectively aggregate topological dependency information among different frequency bands, solving the problem that CNN has in processing non-Euclidean data structure. Spatially encoded data is then given to a GRU network, which uses particular gate systems to model both long- and short-term connections and changes over time, producing a thorough mixing of spatiotemporal characteristics.

(3) Prediction Output: Lastly, to solve the problem of losing important information during long sequence forecasting, we add an attention mechanism. It computes weight coefficients for hidden states, which can adjust the weight of different features according to the time step to emphasize the influence of key historical events on the subsequent state and reduce the disturbance caused by unimportant information. This framework naturally combines the latest signal processing technologies with deep learning models to create a novel end-to-end spectrum prediction solution.

3. Prediction Methodology

3.1 Overview of the Prediction Framework

Building on the research framework described above, this part provides the internal structure and data processing flowchart of the proposed spectrum prediction model. Taking historical spectrum data as input, the model successively extracts and combines frequency-domain, spatial, and temporal characteristics via successive stages of hierarchical

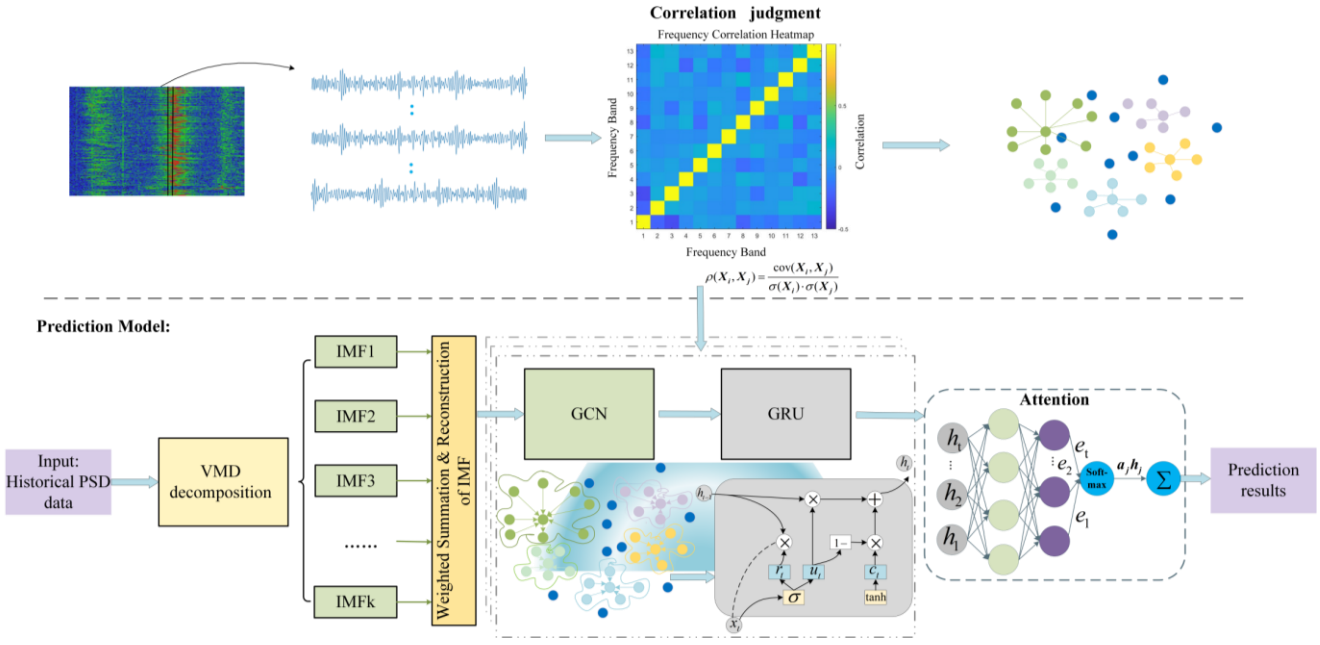

Fig. 1. Overall framework and execution flow.

processing to predict upcoming spectral conditions. Overall architecture and execution workflow are shown in Fig. 1. Core system consists of five closely linked modules:

VMD Module: This module carries out adaptive decomposition and reconstruction on the unprocessed spectrum data to accomplish frequency domain noise reduction and multi-level feature improvement.

Graph Construction Module: According to the reconstituted data, this module evaluates the inter-band correlation and creates the graph structure, providing the basis for spatial modeling.

GCN Module: On the created graph structure, this module extracts spatial dependency features between various frequency bands.

GRU Module: For every frequency band, it models the dynamic development patterns of the time series so as to capture the temporal dependencies.

Attention Mechanism Module: This module adaptively combines spatiotemporal features, allowing the model to concentrate on important information and boost both forecast quality and training speed.

3.2 Variational Mode Decomposition (VMD)

VMD is a fully non-recursive, adaptive signal processing technique capable of decomposing complex nonstationary signals into a set of bandwidth-limited IMFs with distinct center frequencies. Unlike traditional Empirical Mode Decomposition (EMD), VMD achieves adaptive signal separation by addressing the variational problem in the spectral domain, effectively avoiding modal mixing issues and demonstrating superior noise robustness [28]. In this study, VMD is utilized to perform multi-scale decomposition on the historical PSD data of each frequency band. To eliminate the interference of high-frequency noise on prediction accuracy, we identify and remove high-frequency noise components based on the center frequency characteristics of the IMFs. The remaining effective mode components are then aggregated to reconstruct a smoothed feature sequence that retains the core signal trends, serving as the input for the subsequent deep prediction model. The construction of VMD is formulated as a constrained variational optimization problem, aiming to find K mode components $u_k(t)$ and their corresponding center frequencies w_k such that the sum of their estimated bandwidths is minimized. The constrained variational model is defined as shown in (1):

$$\begin{cases} \min\limits_{\{u_k\},\{w_k\}} \left\{ \sum\limits_{k=1}^{K} \left\| \partial_t \left[\left(\delta(t) + \dfrac{\mathrm{j}}{\pi t} \right) * u_k(t) \right] \exp(-\mathrm{j} w_k t) \right\|_2^2 \right\} \\ \text{s.t.} \sum\limits_{k=1}^{K} u_{k(t)=f(t)} \end{cases} \tag{1}$$

where $f(t)$ denotes the original input signal, u_k represents the set of K decomposed mode components, w_k denotes the center frequencies corresponding to each mode, $\delta(t)$ is the Dirac distribution, * represents the convolution operator, ∂t denotes the partial derivative with respect to time, and $\|\cdot\|_2$ represents the L^2 norm. To solve the aforementioned constrained variational problem, a quadratic penalty term α and a Lagrangian multiplier $\lambda(t)$ are introduced to transform the constrained optimization problem into an unconstrained augmented Lagrangian function, defined as in (2):

$$\begin{aligned} L(\{u_k\},\{w_k\},\lambda) = \alpha \sum_k \left\| \partial_t \left[\left(\delta(t) + \frac{\mathrm{j}}{\pi t} \right) \otimes u_k(t) \right] \exp(-\mathrm{j} w_k t) \right\|_2^2 \\ + \left\| f(t) - \sum_k u_k(t) \right\|_2^2 \\ + \left\langle \lambda(t), f(t) - \sum_k u_k(t) \right\rangle \end{aligned} \tag{2}$$

where α is the penalty factor used to control the trade-off between the bandwidth accuracy of the mode decomposition and the reconstruction error. The Alternating Direction Method of Multipliers is employed to iteratively update u_k, w_k, and λ in the spectral domain to search for the saddle point of the augmented Lagrangian function. The iterative update equations for the k-th mode component and its center frequency are given by (3) and (4):

$$\hat{u}_k^{n+1}(w) = \frac{\hat{f}(w) - \sum_{i \neq k} \hat{u}_i(w) + \dfrac{\hat{\lambda}(w)}{2}}{1 + 2\alpha (w - w_k^n)^2}, \tag{3}$$

$$w_k^{n+1} = \frac{\int_0^\infty w \left| \hat{u}_k^{n+1}(w) \right|^2 \mathrm{d}w}{\int_0^\infty \left| \hat{u}_k^{n+1}(w) \right|^2 \mathrm{d}w} \tag{4}$$

where $\hat{u}_k(w)$, $\hat{f}(w)$, and $\hat{\lambda}(t)$ are the Fourier transforms of $u_k(t)$, $f(t)$, and $\lambda(t)$, respectively, and n denotes the iteration count. This optimization process iterates alternately until the convergence condition is met. Ultimately, the original signal is decomposed into K physically meaningful narrowband mode components. During the VMD process, the selection of the number of modes K is critical to the decomposition performance. In this paper, we adopt the RMSE of the final prediction model as the objective metric and determine the optimal K value through experiments to achieve the best balance between noise suppression and feature preservation.

The effectiveness of the VMD module depends crucially on two hyperparameters: the number of intrinsic modes K and the penalty factor α. If K is too small, it leads to under-decomposition (i.e., mode mixing); conversely, if K is too large, the signal is over-decomposed, yielding redundant noise modes that increase computational burden. As detailed in our Parameter Sensitivity Analysis (Section 4.4), the optimal number of modes is empirically determined to be $K = 3$, which achieves the best balance between feature extraction and prediction accuracy. The penalty factor α, which controls the bandwidth of the extracted modes, is set to 2000 to guarantee data fidelity and robust convergence.

Regarding the reconstruction strategy, after decomposing the original non-stationary spectrum signal into $K = 3$ distinct Intrinsic Mode Functions (IMFs), we adopt a direct aggregation (summation) approach. Because VMD is inherently a band-limited decomposition algorithm, restricting the extraction to the 3 most dominant modes naturally leaves out the unmodeled high-frequency residuals, which typically consist of stochastic environmental noise and random channel fading. Therefore, by directly summing these K extracted IMFs, we elegantly synthesize a smoothed, high-fidelity, and denoised spectrum feature sequence. This reconstructed sequence retains the primary temporal trends and physical occupancy patterns while eliminating the need for manual weight tuning, thereby providing a robust and clean input for the subsequent Attention-GCN-GRU module.

3.3 Construction of Graph Topology via Frequency Correlations

Spectrum data shows both dynamic changes over time and strong connections between different frequencies. Different frequency bands, affected by the same wireless propagation environment or user behavior pattern, frequently exhibit strongly synchronized fluctuations. To effectively mine these latent spatial dependencies, this study formulates the spectrum prediction task as a spatio-temporal feature learning problem on a graph structure. We define the spectrum graph as $\mathbf{G} = (V, E)$, where V represents the set of N sub-band nodes, and E represents the set of edges connecting these bands. To quantify the strength of associations between nodes, the Spearman correlation coefficient is adopted as the similarity metric. For any two frequency bands i and j with historical PSD sequences X_i and X_j, respectively, the correlation coefficient $\rho_{i,j}$ is defined as shown in (5):

$$\rho_{i,j} = \frac{\operatorname{cov}(X_i, X_j)}{\sigma(X_i)\cdot\sigma(X_j)} = \frac{\sum_{t=1}^{T}(x_{i,t}-\overline{x_i})(x_{j,t}-\overline{x_j})}{\sqrt{\sum_{t=1}^{T}(x_{i,t}-\overline{x_i})^2}\sqrt{\sum_{t=1}^{T}(x_{j,t}-\overline{x_j})^2}} \tag{5}$$

where $\operatorname{cov}(X_i, X_j)$ denotes the covariance between band i and band j, while $\sigma(X_i)$ and $\sigma(X_j)$ represent their respective standard deviations, and T is the length of the time series. Based on the calculated coefficient matrix, an adjacency matrix $\mathbf{A}$ is constructed by setting a threshold θ to filter out noise interference and retain only strong correlation connections. When the correlation coefficient exceeds the threshold, a significant dependency is considered to exist between the bands; otherwise, they are regarded as uncorrelated. The construction rule for the adjacency matrix is defined in (6):

$$A_{i,j} = \begin{cases} 1, & \rho_{i,j} \geq \theta \\ 0, & \rho_{i,j} < \theta \end{cases}. \tag{6}$$

In this study, the spatial topological relationships among different frequency bands are modeled as an undirected graph, defined by an adjacency matrix $\mathbf{A}$. The graph construction process is strictly defined as follows:

Correlation Calculation and Time Window: We calculate the correlation coefficient between the historical spectrum sequences of different nodes. The computation utilizes the entire time span of the training dataset as the temporal window to capture globally stable spatial dependencies without data leakage.

Thresholding: To construct a sparse and robust graph, a thresholding mechanism is applied. An edge is established between node i and node j ($A_{ij} = 1$) only if their correlation coefficient r_{ij} exceeds a pre-defined threshold $\theta = 0.92$; otherwise, $A_{ij} = 0$.

Through this process, the model successfully establishes a graph structure characterizing the topological relationships among frequency bands, laying the foundation for the subsequent spatial feature extraction.

3.4 Graph Convolutional Network (GCN)

GCNs are deep learning models capable of operating directly on non-Euclidean graph-structured data. Traditional CNN have trouble dealing with irregular topological links, but GCNs use spectral graph theory to gather and change node features through the graph Laplacian matrix, which helps them catch the spatial dependency features of spectrum information. In this study, we adopt the first-order approximation of the spectral GCN model. Given the graph structure $\mathbf{G}$ and the feature matrix $\mathbf{X}$, the layer-wise propagation rule of the graph convolutional layer is defined as shown in (7):

$$\mathbf{H}^{(l+1)} = \sigma(\tilde{\mathbf{D}}^{-\frac{1}{2}}(\tilde{\mathbf{D}}-\tilde{\mathbf{A}})\tilde{\mathbf{D}}^{-\frac{1}{2}}\mathbf{H}^{(l)}\mathbf{W}^{(l)}) \tag{7}$$

where $\mathbf{H}^{(l)}$ denotes the node feature input at the l-th layer, $\mathbf{W}^{(l)}$ represents the trainable weight matrix for the l-th layer; and $\sigma(\cdot)$ is the non-linear activation function. The Rectified Linear Unit (ReLU) function is selected to enhance the model's non-linear expressive capability. To preserve the node's own feature information while aggregating neighbor information, $\tilde{\mathbf{A}} = \mathbf{A} + \mathbf{I}_N$ represents the adjacency matrix with added self-loops, where $\mathbf{I}_N$ is the identity matrix. $\tilde{\mathbf{D}}$ is the corresponding degree matrix, with its diagonal elements defined as $D_{ii} = \Sigma_j \tilde{A}_{ij}$. By performing the symmetric normalization operation $\tilde{\mathbf{D}}^{-\frac{1}{2}}\tilde{\mathbf{A}}\tilde{\mathbf{D}}^{-\frac{1}{2}}$, we prevent the numerical explosion or vanishing of node features during propagation through multi-layer networks. Considering that an excessively deep network structure may lead to the "over-smoothing" problem –where features of different nodes tend to converge, thereby reducing discriminability [30], [31]– this paper designs a two-layer GCN architecture to extract spatial correlations between frequency bands. This configuration ensures that the receptive field covers relevant neighbor nodes while avoiding a reduction in information transmission efficiency.

3.5 Gated Recurrent Unit (GRU)

Although GCNs effectively extract spatial correlation features of spectrum data, they lack the ability to capture the dynamic laws of spectrum state evolution over time. To address this, the GRU is introduced to process time-series information. As an efficient variant of RNNs, compared to LSTM networks, the GRU merges the cell state with the hidden state and employs a dual-gating mechanism consisting of a reset gate and an update gate. This design not only effectively mitigates the vanishing and exploding gradient problems but also significantly reduces parameter volume and computational complexity, making it highly suitable for spectrum prediction tasks requiring high real-time performance. The structure of the GRU adopted in this paper is shown in Fig. 2.

In the proposed model, the input to the GRU is derived from the spatial feature matrix extracted by the preceding GCN module. The GRU controls the retention and forgetting of the information flow through its gating mechanism. First, the reset gate and the update gate are calculated. The reset gate determines how much information from the previous moment's hidden state needs to be ignored, while the update gate controls how much of the previous state is retained for the current moment. The core calculations are defined as shown in (8) and (9):

$$\mathbf{R}_t = \sigma(\mathbf{W}_r \cdot [\mathbf{Y}'_{\mathrm{GCN}}, \mathbf{h}_{t-1}] + \mathbf{b}_r), \tag{8}$$

$$\mathbf{U}_t = \sigma(\mathbf{W}_u \cdot [\mathbf{Y}'_{\mathrm{GCN}}, \mathbf{h}_{t-1}] + \mathbf{b}_u), \tag{9}$$

subsequently, the candidate hidden state $\tilde{\mathbf{H}}_t$ at the current moment is calculated using the reset gate, as shown in (10). A smaller value of the reset gate implies that more information from the previous moment is disregarded.

$$\tilde{\mathbf{H}}_t = \tanh(\mathbf{W}_h \cdot [\mathbf{Y}'_{\mathrm{GCN}}, (\mathbf{R}_t \odot \mathbf{H}_{t-1})] + \mathbf{b}_h) \tag{10}$$

Finally, the update gate is utilized to perform a weighted fusion of the previous hidden state and the current candidate state to obtain the final output state $\mathbf{H}_t$ at the current moment, as defined in (11):

$$\mathbf{H}_t = \mathbf{U}_t \odot \mathbf{H}_{t-1} + (1 - \mathbf{U}_t) \odot \tilde{\mathbf{H}}_t. \tag{11}$$

In the above equations: $\mathbf{Y}'_{\mathrm{GCN}}$ represents the output features from the graph convolutional layer; $\mathbf{H}_{t-1}$ is the hidden state at the previous time step; $[\cdot, \cdot]$ denotes the matrix concatenation operation; $\odot$ represents the element-wise matrix multiplication (Hadamard product); $\sigma(\cdot)$ is the Sigmoid activation function, mapping values to the (0, 1) interval; $\tanh(\cdot)$ is the hyperbolic tangent activation function; $\mathbf{W}$ represents the trainable weight matrices; and $\mathbf{b}$ denotes the corresponding bias terms.

By using the spatial features obtained from the GCN as the input of the GRU, the model can achieve a good fusion of spatiotemporal features. Band correlation info handled by GCN moves along the GRU's time steps, so the last hidden state includes both how bands relate to each other and how things have changed over time. And this kind of cooperation greatly improves the accuracy and stability of predicting spectra.

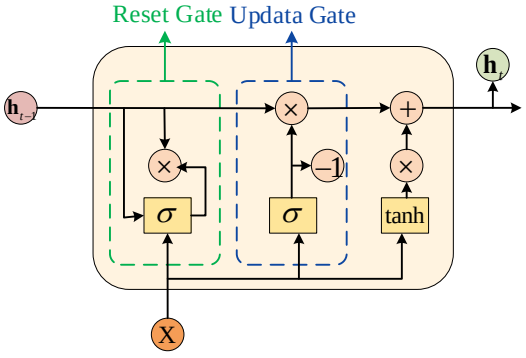

Fig. 2. Schematic diagram of the gated unit structure.

3.6 Attention Mechanism Module

Even though the GRU model can well capture the long-term and short-term dependence of the time sequence, some important information in the first input will also gradually disappear during the transmission process as the time step increases, especially for the long sequences of spectrum data. In order to make the model better understand the global temporal features, a soft attention mechanism is introduced after the GRU layer. This mechanism mimics the human visual attention mechanism by learning adaptively the significance weights of hidden states at various time steps, thus allowing the model to pick out the most crucial information among many past moments that are most relevant to forecasting future spectrum conditions. The structural logic of the soft attention mechanism is illustrated in Fig. 3 and can be specifically divided into three core steps:

(1) Attention Score Calculation: The sequence of hidden states output by the GRU at each moment is input into the attention layer. The energy score of the hidden state at each time step t is calculated by using the Multilayer Perceptron (MLP), thereby measuring the potential importance of the feature at that moment. The calculation is defined as:

$$e_t = \mathbf{v}^{\mathrm{T}} \tanh(\mathbf{W}_a \mathbf{h}_t + \mathbf{b}_a) \tag{12}$$

where $\mathbf{h}_t$ is the output vector of the GRU at the t-th time step; $\mathbf{W}$ and $\mathbf{b}$ represent the weight matrix and bias term of

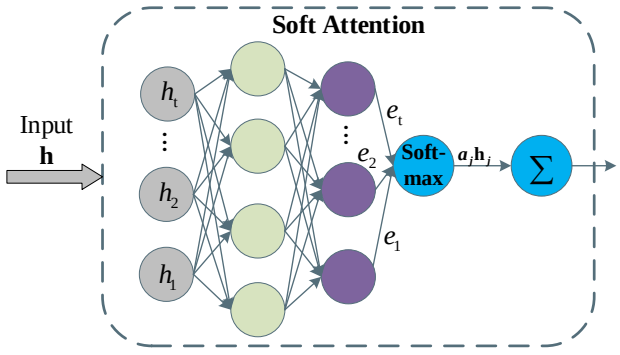

Fig. 3. Schematic diagram of Soft Attention mechanism.

the intermediate layer, respectively; **v** is the weight vector of the output layer used to map the feature to a scalar score; and tanh(·) denotes the hyperbolic tangent activation function.

(2) Weight Normalization: To transform the energy scores into weight coefficients in the form of a probability distribution, the Softmax function is employed to normalize the scores across all time steps. The calculation of the attention weight α_t for the t-th time step is shown in (13):

$$\alpha_t = \text{Softmax}(e_t) = \frac{\exp(e_t)}{\sum_{k=1}^{T} \exp(e_k)} \tag{13}$$

where T denotes the length of the input time series, and $\alpha_t \in [0, 1]$ represents the contribution degree of the information at time t to the final prediction result.

(3) Feature Weighted Fusion: Using the calculated attention weights, a weighted summation of the hidden states at all moments is performed to generate a fused feature vector **C** containing global temporal context information, as shown in (14):

$$\mathbf{C} = \sum_{t=1}^{T} \alpha_t \mathbf{h}_t \ . \tag{14}$$

Ultimately, the vector **C** integrates the spatial features extracted by the GCN, the temporal dynamic features extracted by the GRU, and the key temporal information selected by the attention mechanism. This vector is fed into a Fully Connected layer for linear transformation to obtain the final spectrum prediction result. Based on the detailed descriptions of the key components above, the overall training and prediction procedure of the proposed framework is summarized in Algorithm 1. This framework integrates signal decomposition, spatial topology learning and temporal modeling into a unified end-to-end pipeline.

Algorithm 1 Training and Prediction Procedure of the Proposed Method.

Input: Raw spectrum dataset D, number of VMD modes K, correlation threshold ε, model hyperparameters

Output: Predicted spectrum values $\mathbf{Y}_{\text{pred}}$

1: Stage 1: Multi-scale Decomposition and Denoising
2: **for** each spectrum sequence in D **do**
3: Decompose the non-stationary sequence into IMFs via variational optimization
4: Select effective modes to reconstruct the refined signal sequence, suppressing noise components
5: **end for**
6: Generate the processed feature matrix $\mathbf{X}'$
7: **Stage 2: Graph Topology Learning**
8: Calculate pairwise correlations between frequency bands to measure dependencies
9: Construct the graph adjacency matrix by thresholding the correlation matrix
10: **Stage 3: Spatiotemporal Representation Learning**
11: Initialize neural network parameters Θ
12: **while** training stopping criteria are not met **do**
13: // Forward Propagation Phase
14: **Spatial Modeling:** Aggregate spatial features from neighboring bands using Graph Convolutional Network
15: **Temporal Modeling:** Capture time-series evolution patterns from spatial features using GRU layers
16: **Feature Refinement:** Assign adaptive weights to hidden states via Attention Mechanism to focus on key timestamps
17: **Prediction:** Generate the estimated output $\mathbf{Y}_{\text{pred}}$ through the fully connected layer
18: // Backward Propagation Phase
19: Compute the reconstruction loss (MSE) between $\mathbf{Y}_{\text{pred}}$ and ground truth
20: Optimize model parameters Θ via backpropagation to minimize the loss
21: **end while**
22: **return** The final predicted spectrum $\mathbf{Y}_{\text{pred}}$

4. Experimental Results and Analysis

4.1 Dataset and Parameter Settings

To rigorously verify the validity and stability of the proposed model, we utilized actual spectrum measurements from the open-source spectrum monitoring platform, ElectroSense [32], referred to herein as Dataset 1. We specifically selected the frequency range from 500 MHz to 800 MHz. This sub-GHz band was chosen because it encompasses diverse communication systems, including digital video broadcasting (DVB-T) and cellular networks (e.g., LTE Band 20 and 28). The inherent coexistence among these heterogeneous systems introduces complex PSD fluctuations [33]. Consequently, it exhibits highly dynamic and congested characteristics, providing a challenging and ideal scenario for evaluating spectrum prediction models. The data were collected in Madrid, Spain, spanning a 30-day period from May 28, 2021, to June 28, 2021. The measurements possess a temporal resolution of 15 minutes and a frequency resolution of 2 MHz. After preprocessing, the resulting data sample matrix has dimensions of 151×2880, representing 151 frequency band nodes and 2,880 temporal sampling points, respectively.

To further evaluate the generalization capability of the model across different frequency domains and to mitigate the limitations of single-scenario validation, Dataset 2 was introduced. This dataset was acquired from the Wireless Networks Department at RWTH Aachen University [34]. The raw measurements were conducted in a residential area of Aachen, Germany, utilizing an Agilent E4440 spectrum analyzer. While the original dataset records Power Spectral Density (PSD) across a broad frequency range (20 MHz to 6000 MHz) with a fine temporal resolution of 1.8 seconds, the massive data volume necessitated the extraction of a specific sub-band for intensive analysis.

Therefore, we extracted the 3750 MHz to 3810 MHz band for our experiments. This specific 60 MHz span in the C-band is of particular interest as it is critical for emerging 5G New Radio (NR) networks, exhibiting entirely different occupancy patterns and non-stationary characteristics compared to the sub-GHz band. The extracted data sample matrix has dimensions of 300×6000, corresponding to 300 frequency band nodes (yielding a frequency resolution of 200 kHz) and 6,000 temporal sampling points (representing a continuous observation period of exactly 3 hours).

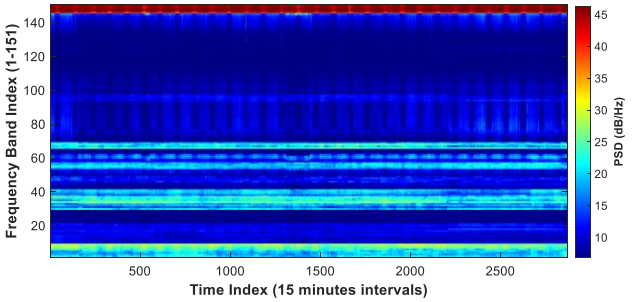

Fig. 4. Graphical Illustration of RF spectrum - Dataset 1.

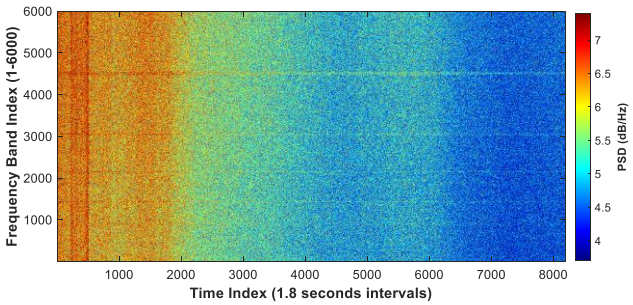

Fig. 5. Graphical illustration of RF spectrum - Dataset 2.

Object	Parameters	Dataset 1	Dataset 2
Dataset	Frequency band	500 MHz – 800 MHz	3750 MHz – 3810 MHz
	Monitoring location	Madrid, Spain	Aachen, Germany
	Hardware equipment	Electrosense sensor	Agilent E4440
	Monitoring time	May 28 – June 28, 2021	3 hours (started July 3, 2007)
	Frequency resolution	2 MHz	0.2 MHz
	Time resolution	15 minutes	1.8 seconds
	The dimensionality of samples	151 × 2880	300 × 6000
Experiment	Dataset split ratio	8 : 2	8 : 2
	Time step	16	16
	Batch size	64	64
	Maximum epochs	1500	1500
	Learning rate	0.0001	0.0001

Tab. 1. Main parameters of dataset and experiment.

To provide an intuitive understanding of the radio environments, Figure 4 and Figure 5 illustrate the raw RF spectra (i.e., PSD distribution) for Dataset 1 and Dataset 2, respectively. These visualizations clearly demonstrate the distinct spatiotemporal occupancy patterns and dynamic fluctuations across different frequency bands, further justifying the necessity of employing graph-based prediction models.

Finally, to prevent model overfitting and accurately assess predictive performance, both datasets are chronologically partitioned into training and test sets at an 8 : 2 ratio. For the forecasting task, a sliding window mechanism with a length of 16 time steps is employed to perform one-step-ahead prediction. This rolling forecast process is executed across the entirety of the test set to ensure a comprehensive and statistically significant evaluation of the model's stability. The detailed configurations, spatiotemporal dimensions for both datasets are systematically summarized in Tab. 1.

4.2 Evaluation Metrics

In order to fully assess the performance of the spectrum prediction model, this paper uses four kinds of indicators that are often adopted by the current spectrum prediction studies, which are Accuracy, MAE, RMSE, and R^2 [35]. The definitions of these evaluation metrics are as follows:

$$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{k=1}^{N}(y_k - \hat{y}_k)}, \tag{15}$$

$$R^2 = 1 - \frac{\sum_{k=1}^{N}(y_k - \hat{y}_k)^2}{\sum_{k=1}^{N}(y_k - \bar{y}_k)^2}, \tag{16}$$

$$\mathrm{MAE} = \frac{1}{N}\sum_{k=1}^{N}\left|y_k - \hat{y}_k\right|, \tag{17}$$

$$\mathrm{Accuracy} = 1 - \frac{\left\|\hat{y}_k - y_k\right\|_{\mathrm{F}}}{\left\|y_k\right\|} \tag{18}$$

where y_k represents the ground truth value at the k-th time step; $\hat{y}_k$ denotes the predicted value at the k-th time step; $\bar{y}_k$ is the mean value of the observed data; N is the total number of prediction samples.

4.3 Performance Comparison and Analysis

Among the aforementioned evaluation metrics, Accuracy measures the overall similarity between the predicted values and the ground truth; a value closer to 1 indicates higher prediction precision. MAE quantifies the average magnitude of prediction deviations, where a smaller value implies greater model robustness. RMSE is particularly sensitive to large errors and reflects the average deviation between predicted and observed values. R^2 assesses the goodness of fit regarding the variance of the true values, with a value approaching 1 representing a superior fit [36], [37]. To verify the effectiveness and superiority of the proposed framework, we selected several representative and state-of-the-art modeling approaches currently used in the field of spectrum prediction as baselines. These benchmark models can be categorized into three groups:

(1) Temporal Sequence Models: including the classic Recurrent Neural Network variant GRU [38] and the self-attention-based Transformer [25].

(2) Spatiotemporal Hybrid Models: including CNN+LSTM [31] and GCN+LSTM [39], which sequentially combine spatial feature extractors with temporal predictors.

(3) Advanced Graph and Attention-driven Models: including the graph attention-based network AGCRNN [21], the cutting-edge adaptive spatial-temporal graph neural

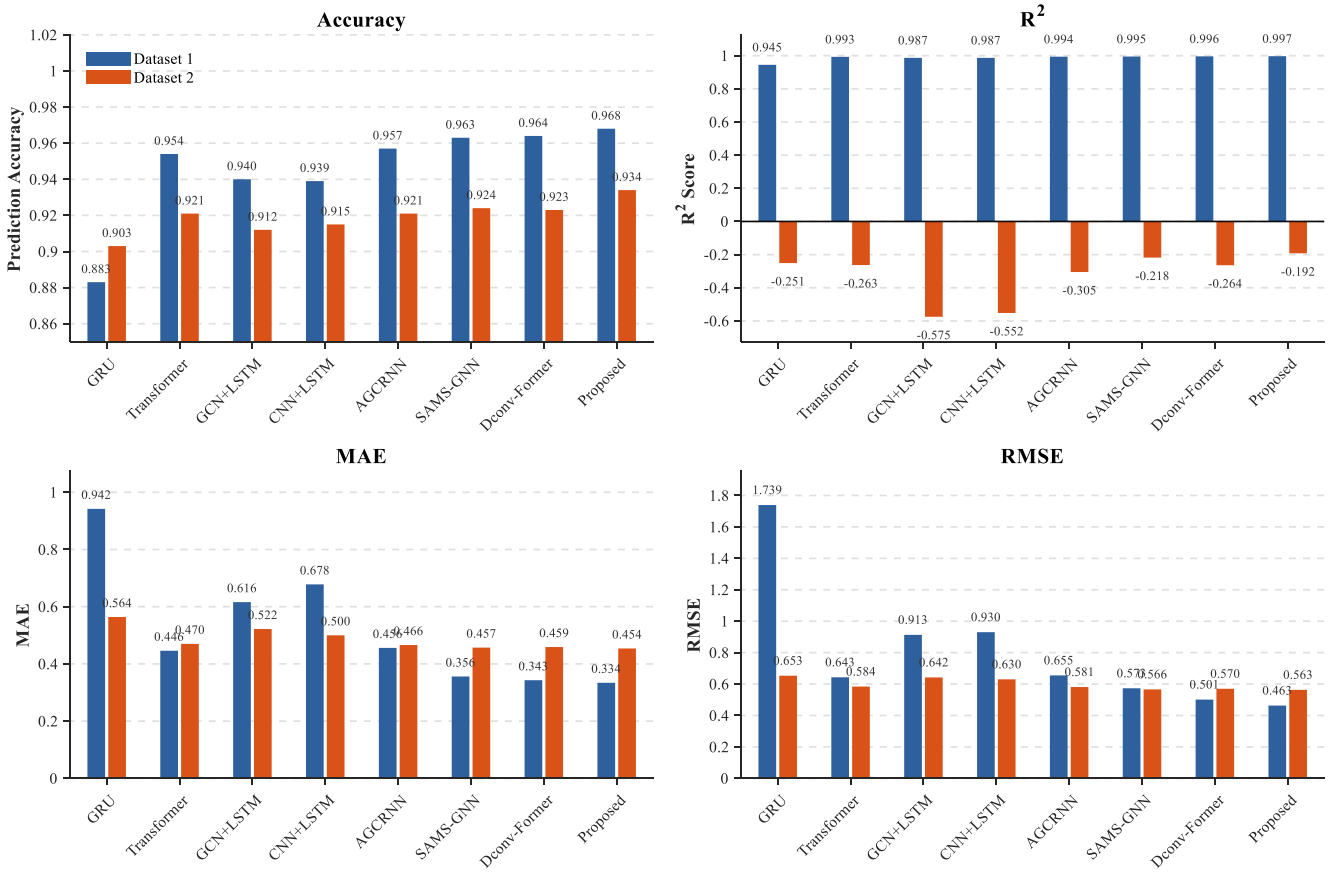

Fig. 6. Performance comparison of different models on Dataset 1 and Dataset 2.

network SAMS-GNN [40], and the dynamically convolutional Transformer Dconv-Former [26].

The comprehensive comparative results of all models evaluated on Dataset 1 and Dataset 2 are presented in Fig. 6. As shown in Fig. 6, traditional temporal models (GRU) exhibit the worst performance, as they treat frequency bands as isolated time series, entirely ignoring the cross-band interference inherent in radio environments. While spatiotemporal fusion models (GCN+LSTM, CNN+LSTM) improve upon this by incorporating spatial dependencies, their performance remains suboptimal. This is primarily because conventional GNNs rely on static or linear adjacency matrices, which fail to capture the highly non-linear, dynamic topology caused by transient channel fading. In contrast, our proposed framework utilizes a Spearman-based non-linear graph, allowing our proposed model to track transient correlations across the spectrum, significantly reducing the MAE and RMSE compared to standard GCN+LSTM architectures.

A critical observation from the results is the comparison with state-of-the-art Transformer-based models (Transformer [25] and Dconv-Former [26]). While Dconv-Former achieves highly competitive results on Dataset 1 (Accuracy: 0.9639) by leveraging global receptive fields through self-attention, it is still strictly outperformed by the proposed model (Accuracy: 0.9679). The fundamental reason lies in the susceptibility of pure attention mechanisms to high-frequency stochastic noise. In highly congested spectrum bands, sudden bursts of noise can easily distract the attention map, leading to suboptimal weight allocation. Our model effectively circumvents this bottleneck by employing VMD as a physics-informed preprocessing step. By adaptively decomposing the raw signal into intrinsic modes, VMD explicitly strips away high-frequency environmental noise, feeding only the predictable, essential frequency components into the deep learning backend. This structural bias ensures cleaner feature representations than the raw inputs used by Dconv-Former.

The superiority of the proposed framework is further magnified when transitioning from the relatively stable Sub-GHz broadcast band (Dataset 1) to the highly dynamic 5G NR C-band (Dataset 2). As illustrated in Fig. 6, the extreme burstiness of the C-band causes a severe performance degradation across all baseline models (noticeable drops in R^2 scores and increased MAE). Notably, models lacking adaptive noise suppression (such as GCN+LSTM and CNN+LSTM) experience the most significant Accuracy collapse. Conversely, the proposed model demonstrates exceptional robustness, maintaining the highest prediction accuracy (0.934) and the lowest error metrics (MAE: 0.454). This cross-domain validation conclusively proves that the deep coupling of VMD (for noise isolation), GCN (for non-linear topological mapping), and Attention-GRU (for temporal feature weighting) provides a highly resilient architecture capable of handling the severe non-stationarity of diverse real-world electromagnetic environments.

4.4 Parameter Sensitivity Analysis

In the VMD process, the selection of the number of modes *K*, directly determines the separation quality of the spectrum data components, which in turn affects the prediction accuracy of the model. To study the optimal decomposition level, this research conducted a comparative experiment with *K* varying from 2 to 7 and a step size of 1.

As shown in Fig. 7, with the change of the number of modal components *K*, the model performance shows a distinct fluctuating trend. The experimental results show that all error indicators reach the global minimum when $K = 3$, and the prediction performance is the best. From the perspective of signal processing, this phenomenon can be attributed to the inherent physical composition of radio spectrum signals. When $K = 3$, VMD effectively decouples the original spectrum into three components with different functions: Macro trends, cyclical fluctuations and random noise. This configuration achieves an ideal balance between signal resolution and spectral sparsity. On the contrary, an excessive amount of *K* can lead to "over-decomposition", that is, a single inherent pattern is artificially split into multiple sub-components. This redundant information increases the "redundancy entropy" of the input features, forcing the subsequent GCN-GRU module to process highly correlated but physically meaningless data. This fragmentation not only complicates the spatio-temporal modeling process but also amplifies high-frequency noise, ultimately leading to overfitting and seriously reducing the generalization ability of the model. Therefore, $K = 3$ represents the optimal structural risk minimization point, achieving a compromise between feature completeness and modeling complexity.

To explore the influence of different correlation thresholds during the graph construction process on the model performance, we conducted parameter optimization within the range of 0.70 to 0.96, with a step size of 0.02.

As shown in Figs. 8 and 9, the experimental results indicate that the model performance is highly sensitive to the strictness of the inter-band filtering mechanism. However, the relationship between the correlation threshold and the prediction accuracy is non-monotonic. When the relevant threshold is set to 0.92, both the accuracy rate and R^2 reach their peaks, RMSE and MAE drop to their respective minimum values, and the comprehensive performance reaches the best.

This optimal threshold of 0.92 represents a key trade-off between topological connectivity and spatial noise suppression. A lower threshold will generate an overly dense adjacency matrix, introducing redundant weakly correlated edges as spatial interference during the GCN feature aggregation process. Specifically, the significant performance improvement observed when the threshold increases from 0.70 to 0.80 is primarily because the Spearman correlation values for multiple frequency bands are concentrated within this interval. A lower threshold setting erroneously connects physically independent frequency bands that host uncorrelated wireless services. This

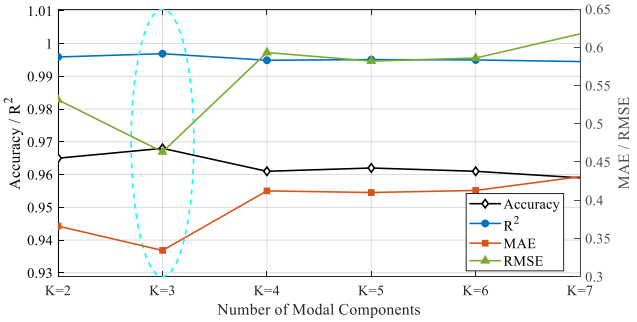

Fig. 7. Effect of the number of modal components on the performance of the model.

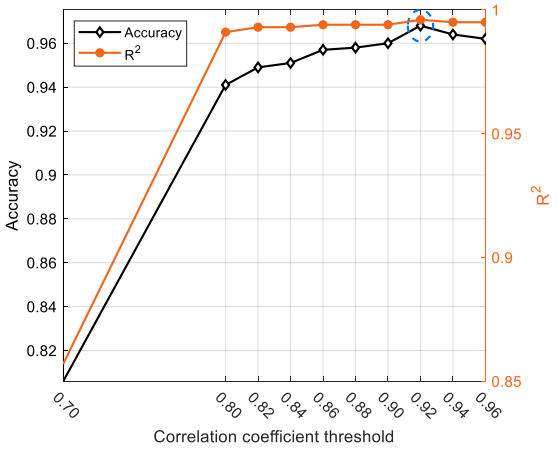

Fig. 8. Effect of varying correlation thresholds on model Accuracy and R^2.

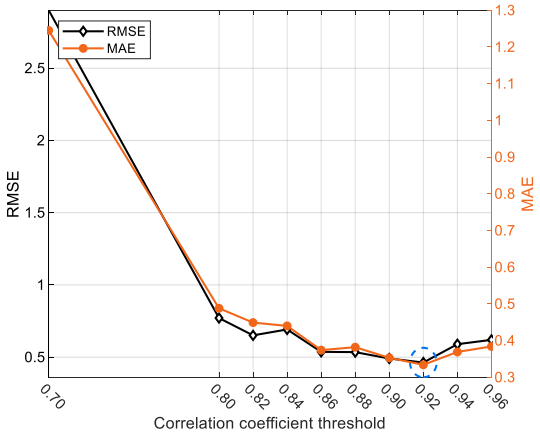

Fig. 9. Effect of varying correlation thresholds on model MAE and RMSE.

forces the model to aggregate irrelevant out-of-band traffic patterns, thereby causing severe cross-band interference. On the contrary, when the threshold exceeds 0.92, the graph structure becomes overly sparse, resulting in the exclusion of frequency bands with inherent physical correlations. This leads to the loss of the basic spatial background and weakens the model's ability to capture the complex spatiotemporal coupling of the radio environment. Therefore, the 0.92 threshold effectively retains the most informative interband dependencies while filtering out random interferences, thereby maximizing the gain of spatial modeling components.

4.5 Convergence Analysis and Visualization

To study the convergence speed and the trend of prediction error of each model during the convergence process, we conducted a comparative experiment of loss value and time between the mainstream model and the model proposed in this paper. The variation of the loss value over time is shown in Fig. 10.

As shown in Fig. 10, the proposed model exhibits superior convergence characteristics, outperforming all baseline models in terms of both convergence speed and final accuracy. Notably, while the Transformer-based models (Transformer and Dconv-Former) were trained for a longer duration, their final loss values still remain higher than that of our proposed model, which stabilizes at 0.385 much earlier. In contrast, mainstream models such as AGCRNN and GRU+GCN converge to higher loss levels (0.642 and 0.909, respectively) despite similar training cycles.

Theoretically speaking, the superior convergence performance of this architecture can be attributed to the integration of VMD and GCN modeling. This model decomposes non-stationary spectral data into multiple stabilizer components through VMD, effectively reducing the complexity of the underlying objective function, making the loss curve smoother and the gradient descent speed faster. Furthermore, spatial modeling based on GCN offers a structural inductive bias that captures interband dependencies more effectively than traditional cyclic or convolutional structures, which often struggle with high-dimensional time-frequency couflers. Finally, the combination of attention mechanisms allows for adaptive weighting of key features, thereby minimizing modeling residuals and preventing the model from falling into suboptimal local minima. These factors jointly ensure that the proposed model achieves rapid convergence and at the same time guarantees high fidelity suitable for complex radioenvironments.

To more intuitively demonstrate the prediction capability of the proposed model across different scenarios, prediction results on the test set from two distinct datasets are visualized. Figures 11–13 present the results from Dataset 1, covering the bands 618–620 MHz, 652–654 MHz, and 784–786 MHz. To further verify the model's performance in higher frequency ranges, Figures 14–16 display the results from Dataset 2, specifically for the bands 3806.6–3806.8 MHz, 3775.2–3775.4 MHz, and from 3759.4 to 3759.6 MHz.

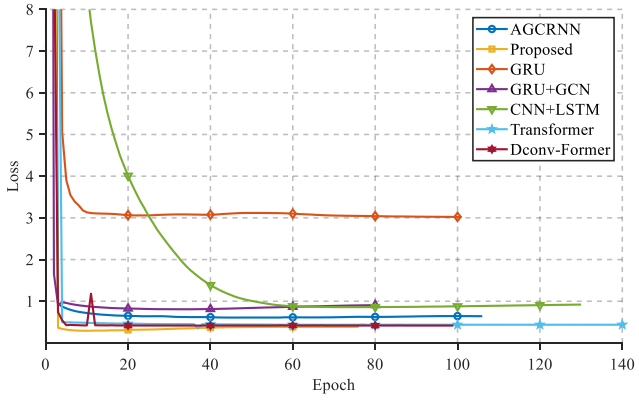

Fig. 10. Loss curve over training epochs.

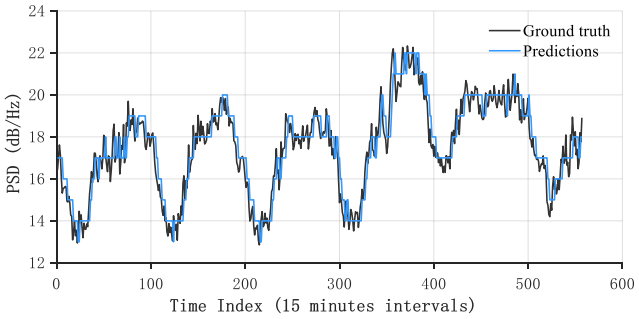

Fig. 11. Predicted spectrum results for the 618–620 MHz test set.

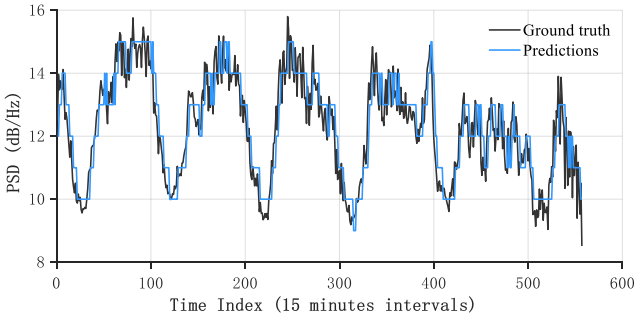

Fig. 12. Predicted spectrum results for the 652–654 MHz test set.

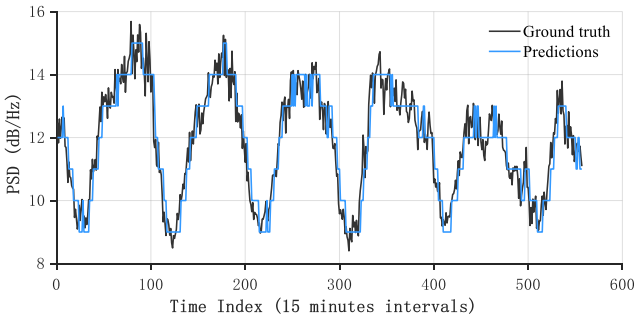

Fig. 13. Predicted spectrum results for the 784–786 MHz test set.

In the visualization results, the black curves represent the ground truth values, while the blue curves represent the model's predicted values. Across the three selected test bands,

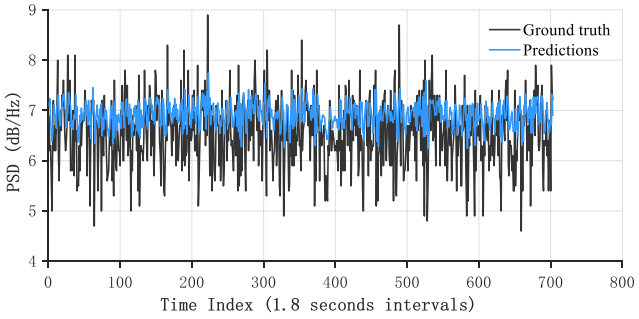

Fig. 14. Predicted spectrum results for the 3806.6–3806.8 MHz test set.

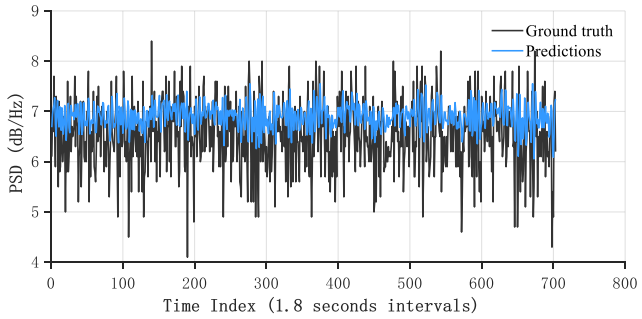

Fig. 15. Predicted spectrum results for the 3775.2–3775.4 MHz test set.

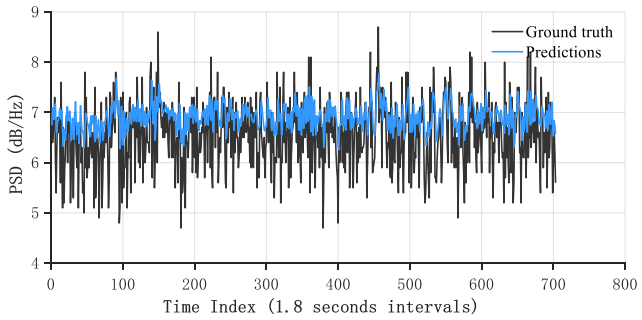

Fig. 16. Predicted spectrum results for the 3759.4–3759.6 MHz test set.

the predicted curves highly overlap with the ground truth curves. Even at points of abrupt change, the model maintains excellent trend-tracking capability and strong fitting precision, demonstrating superior ability in capturing temporal dynamics and expressing frequency-domain features.

4.6 Ablation Study

To further quantify the contribution of each key component within the proposed model, systematic ablation studies were conducted. Using the GRU as a baseline, the study explores the sources of performance gains by incrementally integrating VMD, GCN-based spatial modeling, and the Attention mechanism. The numerical results of the evaluation metrics are presented in Tab. 2.

The baseline GRU yields the highest error (MAE: 0.9418) as it treats all historical moments equally and ignores inter-band interference. Introducing the Attention mechanism (GRU+Attention) reduces MAE by 38.40% (0.5801), effectively forcing the network to focus on transient, high-energy traffic spikes to address the highly bursty nature of spectrum occupancy. Integrating the GCN (GCN+GRU+Attention) further decreases MAE to 0.4878 by capturing non-linear inter-band dependencies, such as cross-channel interference and spectrum leakage, which pure temporal models overlook. However, the most profound leap occurs with VMD. Compared to GCN+GRU+Attention, adding VMD further reduces MAE by 31.48% (0.3342) and RMSE by 35.06% (0.4632). This proves that feeding raw, highly non-stationary signals directly into deep networks causes severe fitting bottlenecks due to stochastic environmental noise. VMD acts as a physics-informed filter, adaptively deconstructing chaotic time-frequency mixtures into smooth IMFs to drastically reduce prediction space complexity. Ultimately, the proposed framework achieves the best overall performance (Accuracy: 0.9679, R^2: 0.9969), unequivocally validating that isolating noise via VMD, mapping cross-band interference via GCN, and dynamically weighting historical bursts via Attention-GRU forms a highly effective and resilient paradigm for real-world spectrum forecasting.

Methods	Accuracy	R^2	MAE	RMSE
GRU	0.8829	0.9452	0.9418	1.7394
GRU+Attention	0.9426	0.9885	0.5801	0.8699
VMD+GRU+Attention	0.9546	0.9928	0.4719	0.6884
GCN+GRU	0.9443	0.9892	0.5682	0.8506
GCN+GRU+Attention	0.9525	0.9927	0.4878	0.7133
Proposed	0.9679	0.9969	0.3342	0.4632

Tab. 2. Numerical comparison of evaluation metrics for ablation studies model.

4.7 Computational Complexity and Runtime Analysis

To evaluate the practical feasibility of the proposed framework, we briefly analyze its computational complexity and actual runtime. The overall architecture consists of VMD preprocessing and the AGCRNN prediction module.

For the data preprocessing phase, the complexity of VMD is approximately $O(K \cdot T \log T)$ per frequency band, where K is the number of intrinsic modes and T is the sequence length. While VMD introduces additional computational overhead, it is primarily performed offline or can be parallelized across frequency nodes, minimizing its impact on real-time execution. For the deep learning prediction module, the GCN operation incurs a complexity of $O(|E|d)$, where $|E|$ is the number of edges and d is the feature dimension. The GRU and attention mechanisms require $O(Nd^2)$ and $O(NT_{\text{in}}^2)$, respectively, where N is the number of frequency bands.

In terms of actual runtime, we benchmarked the inference speed on an NVIDIA RTX 5000 GPU. The software environment was configured with Python 3.9.25 and PyTorch 2.5.0, utilizing CUDA 12.1 and cuDNN 90100 for hardware acceleration, alongside essential data-handling libraries including NumPy 2.0.1, Pandas 2.3.3, and SciPy 1.13.1. Taking Dataset 1 as an example, the total inference time for the entire test set (comprising 576 temporal snapshots) is merely 0.2502 seconds. Consequently, the average inference time for a single forward prediction step is approximately 0.43 milliseconds. This sub-millisecond execution time conclusively demonstrates that the proposed model is highly efficient during the inference phase and fully satisfies the stringent low-latency requirements of real-world dynamic spectrum tracking.

5. Conclusion

Aiming at the challenges of suboptimal spectral pre-

diction accuracy caused by the inherent non-stationarity of spectral data and strong spatio-temporal coupling, a graph convolutional recurrent neural network model combined with VMD is proposed. Our approach does not directly process the raw data but first decomposes the spectrum into isolated frequency components through VMD. Then, each pattern undergoes specialized smoothing and denoising processing, and is reassembled through a weighted reconstruction scheme. This strategy significantly enhances the system's sensitivity to high-frequency noise and complex spectral patterns. By simultaneously capturing the nuances in the frequency domain and the evolution over time, this model ensures a more comprehensive understanding of the data, thereby achieving better prediction results. Through extensive experiments, including ablation analysis and comparisons with leading baseline models, we verified the proposed architecture.

Practically, the proposed framework holds significant potential for real-world dynamic spectrum access and cognitive radio networks. For example, in a dense smart city Internet of Things environment, numerous devices intermittently access the spectrum, causing severe interference and non-stationary traffic patterns. By deploying our model at the base station or spectrum management center, the system can accurately predict spectrum availability in advance. This capability allows the network to proactively allocate frequency resources, avoid potential transmission collisions, and significantly improve overall spectral efficiency in congested environments.

While the current study utilizes Spearman correlation coefficients to construct the adjacency matrix, our future work will focus on incorporating adaptive graph-learning mechanisms to dynamically mine complex inter-band relationships, thereby further enhancing prediction precision and exploring its efficient deployment in real-world cognitive radio networks.

The source code and datasets used to support the findings of this study are available from the corresponding author upon reasonable request.

Acknowledgments

This work was supported by the Jiangxi Province 03 Special Project and 5G Project (Grant No. 20232ABC03A09) and the Graduate Innovation Fund of the East China University of Technology (Grant No. DHYC-2025145).

References

[1] DEHOS, C., GONZALEZ, J. L., DE DOMENICO, A., et al. Millimeter-wave access and backhauling: The solution to the exponential data traffic increase in 5G mobile communications systems. *IEEE Communications Magazine*, 2014, vol. 52, no. 9, p. 88–95. DOI: 10.1109/MCOM.2014.6894457

[2] TAMIZHELAKKIYA, K., GAUNI, S., CHANDHAR, P. Transfer learning based location-aided modulation classification in indoor environments for cognitive radio applications. *Radioengineering*, 2023, vol. 32, no. 4, p. 531–541. DOI: 10.13164/re.2023.0531

[3] RABIU, E. O., AKANDE, D. O., ADEYEMO, Z. K., et al. Three-state hidden Markov model for spectrum prediction in cognitive radio networks. *ABUAD Journal of Engineering Research and Development*, 2024, vol. 7, no. 2, p. 425–435. DOI: 10.53982/ajerd.2024.0702.40-j

[4] CHAE, K., PARK, J., KIM, Y. Rethinking autocorrelation for deep spectrum sensing in cognitive radio networks. *IEEE Internet of Things Journal*, 2023, vol. 10, no. 1, p. 31–41. DOI: 10.1109/JIOT.2022.3200968

[5] PAN, G., LI, J. LI, M. Multi-channel multi-step spectrum prediction using transformer and stacked Bi-LSTM. *China Communications*, 2025, vol. 22, no. 5, p. 1–13. DOI: 10.23919/JCC.ja.2022-0667

[6] XUE, W. J., FU, N., GAO, Y. L. Spectrum prediction algorithm based on GCN-LSTM (in Chinese). *Radio Communications Technology*, 2023, vol. 49, no. 2, p. 203–208.

[7] LI, S., SUN, Y., ZHANG, H., et al. MTF2N: Multi-channel temporal-frequency fusion network for spectrum prediction. In *Proceedings of the IEEE Global Communications Conference (GLOBECOM)*. Rio de Janeiro (Brazil), 2022, p. 4703–4709. DOI: 10.1109/GLOBECOM48099.2022.10001101

[8] RADHAKRISHNAN, N., KANDEEPAN, S., YU, X., et al. Performance analysis of long short-term memory-based Markovian spectrum prediction. *IEEE Access*, 2021, vol. 9, p. 149582–149595. DOI: 10.1109/ACCESS.2021.3125725

[9] AYGÜL, M. A., ÇIRPAN, H. A., ARSLAN, H. Machine learning-based spectrum occupancy prediction: a comprehensive survey. *Frontiers in Communications and Networks*, 2025, vol. 6, p. 1–21. DOI: 10.3389/frcmn.2025.1482698

[10] KUMAR, A., GAUR, N., CHAKRAVARTY, S., et al. Analysis of spectrum sensing using deep learning algorithms: CNNs and RNNs. *Ain Shams Engineering Journal*, 2024, vol. 15, no. 3, p. 1 to 11. DOI: 10.1016/j.asej.2023.102505

[11] SOLANKI, S., DEHALWAR, V., CHOUDHARY, J., et al. Spectrum sensing in cognitive radio using CNN-RNN and transfer learning. *IEEE Access*, 2022, vol. 10, p. 113482–113492. DOI: 10.1109/ACCESS.2022.3216877

[12] DUBININ, I., EFFENBERGER, F. Fading memory as inductive bias in residual recurrent networks. *Neural Networks*, 2024, vol. 173, p. 1–15. DOI: 10.1016/j.neunet.2024.106179

[13] CULLEN, A. C., RUBINSTEIN, B. I. P., KANDEEPAN, S., et al. Predicting dynamic spectrum allocation: a review covering simulation, modelling, and prediction. *Artificial Intelligence Review*, 2023, vol. 56, p. 10921–10959. DOI: 10.1007/s10462-023-10449-9

[14] BEN, C., PENG, Y., ZHANG, X., et al. Multi-band spectrum prediction method via singular spectrum analysis and A-BILSTM. In *Proceedings of the 23rd IEEE International Conference on Communication Technology (ICCT)*. Wuxi (China), 2023, p. 645 to 650. DOI: 10.1109/ICCT59356.2023.10419464

[15] LINDEMANN, B., MÜLLER, T., VIETZ, H., et al. A survey on long short-term memory networks for time series prediction. *Procedia CIRP*, 2021, vol. 99, p. 650–655. DOI: 10.1016/j.procir.2021.03.088

[16] YUAN, L., NIE, L., HAO, Y. Communication spectrum prediction method based on convolutional gated recurrent unit network. *Scientific Reports*, 2024, vol. 14, no. 1, p. 1–17. DOI: 10.1038/s41598-024-56311-y

[17] GUO, X., XU, Y., SUN, J., et al. Dynamic graph temporal-frequency dual-channel network for multi-band spectrum

prediction. *IEEE Communications Letters*, 2024, vol. 28, no. 12, p. 2940–2944. DOI: 10.1109/LCOMM.2024.3451536

[18] SHAWEL, B. S., WOLDEGEBREAL, D. H., POLLIN, S. Convolutional LSTM-based long-term spectrum prediction for dynamic spectrum access. In *Proceedings of the 27th European Signal Processing Conference (EUSIPCO)*. A Coruna (Spain), 2019, p. 1–5. DOI: 10.23919/EUSIPCO.2019.8902956

[19] JIN, D., YU, Z., HUO, C., et al. Universal graph convolutional networks. *Advances in Neural Information Processing Systems*, 2021, vol. 34, p. 10654–10664.

[20] GU, N. R., ZHANG, J. Z., JI, X. Spectrum prediction method based on a time–frequency fusion attention network (in Chinese). *Journal of Chongqing University of Posts and Telecommunications (Natural Science Edition)*, 2025, vol. 37, no. 3, p. 305–314.

[21] ZHANG, X., GUO, L., BEN, C., et al. A-GCRNN: Attention graph convolution recurrent neural network for multi-band spectrum prediction. *IEEE Transactions on Vehicular Technology*, 2024, vol. 73, no. 2, p. 2978–2982. DOI: 10.1109/TVT.2023.3315450

[22] CINI, A., MARISCA, I., ZAMBON, D., et al. Graph deep learning for time series forecasting. *ACM Computing Surveys*, 2025, vol. 57, no. 12, p. 1–34. DOI: 10.1145/3742784

[23] DAKHMOUCHE, R., LUNATI, I., GORJI, H. Network system forecasting despite topology perturbation. In *Proceedings of the ICML 2025 Workshop on Scaling Up Intervention Models*. Vienna (Austria), 2025.

[24] DEL PRIORE, E., LAMPANI, L. Real-time damage detection and localization on aerospace structures using graph neural networks. *Journal of Sensor and Actuator Networks*, 2025, vol. 14, no. 5, p. 1–26. DOI: 10.3390/jsan14050089

[25] VASWANI, A., SHAZEER, N., PARMAR, N., et al. Attention is all you need. In *Proceedings of the 31st Conference on Neural Information Processing Systems (NeurIPS)*. Long Beach (USA), 2017, p. 1–11.

[26] XU, H., LI, J., WU, Q., et al. Dconv-Former: Efficient transformer for spatial-temporal–spectral spectrum prediction. *IEEE Wireless Communications Letters*, 2025, vol. 14, no. 8, p. 2272–2276. DOI: 10.1109/LWC.2025.3530981

[27] XU, W. ZHANG, J. SU, Z. Explainable spectrum prediction based on VMD-LSTM, *Radioengineering*, 2026, vol. 35, no. 1, p. 15–25. DOI: 10.13164/re.2026.0015

[28] CHEN, T., GAO, S., ZHENG, S., et al. EMD and VMD empowered deep learning for radio modulation recognition. *IEEE Transactions on Cognitive Communications and Networking*, 2022, vol. 9, no. 1, p. 43–57. DOI: 10.1109/TCCN.2022.3218694

[29] DRAGOMIRETSKIY, K., ZOSSO, D. Variational mode decomposition. *IEEE Transactions on Signal Processing*, 2013, vol. 62, no. 3, p. 531–543. DOI: 10.1109/TSP.2013.2288675

[30] RASHID, Y., BHAT, J. I. OlapGN: A multi-layered graph convolution network-based model for locating influential nodes in graph networks. *Knowledge-Based Systems*, 2024, vol. 283, p. 1 to 10. DOI: 10.1016/j.knosys.2023.111163

[31] CHIANG, W. L., LIU, X., SI, S., et al. Cluster-GCN: An efficient algorithm for training deep and large graph convolutional networks. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining.* Anchorage (USA), 2019, p. 257–266. DOI: 10.1145/3292500.3330925

[32] RAJENDRAN, S., CALVO-PALOMINO, R., FUCHS, M., et al. Electrosense: Open and big spectrum data. *IEEE Communications Magazine*, 2017, vol. 56, no. 1, p. 210–217. DOI: 10.1109/MCOM.2017.1700200

[33] POLAK, L., KALLER, O., KLOZAR, L., et al. Coexistence between DVB-T/T2 and LTE standards in common frequency bands. *Wireless Personal Communications*, 2016, vol. 88, p. 669 to 684. DOI: 10.1007/s11277-016-3191-2

[34] RIIHIJÄRVI, J. Empirical modelling of spectrum use and evaluation of adaptive spectrum sensing in dynamic spectrum access networks. *Ph.D. Dissertation*. RWTH Aachen University, Aachen (Germany), 2010.

[35] XU, X., YONEDA, M. Multitask air-quality prediction based on LSTM-autoencoder model. *IEEE Transactions on Cybernetics*, 2019, vol. 51, no. 5, p. 2577–2586. DOI: 10.1109/TCYB.2019.2945999

[36] WANG, L., HU, J., ZHANG, C., et al. Deep learning models for spectrum prediction: A review. *IEEE Sensors Journal*, 2024, vol. 24, no. 18, p. 28553–28575. DOI: 10.1109/JSEN.2024.3416738

[37] XIAO, W., WANG, X. Spatial-temporal dynamic graph convolutional neural network for traffic prediction. *IEEE Access*, 2023, vol. 11, p. 97920–97929. DOI: 10.1109/ACCESS.2023.3312534

[38] TUSHA, A., KAPLAN, B., CIRPAN, H. A., et al. Intelligent spectrum occupancy prediction for realistic measurements: GRU based approach. In *Proceedings of the 10th International Black Sea Conference on Communications and Networking (BlackSeaCom)*. Sofia (Bulgaria), 2022, p. 179–184. DOI: 10.1109/BlackSeaCom54372.2022.9858237

[39] ZHANG, H., TIAN, Q., HAN, Y. Multi-channel spectrum prediction algorithm based on GCN and LSTM. In *Proceedings of the IEEE 96th Vehicular Technology Conference (VTC2022-Fall)*. London (UK), 2022, p. 1–5. DOI: 10.1109/VTC2022-Fall57202.2022.10013030

[40] ZHANG, X., PENG, Y., HUANG, H., et al. SAMS-GNN: Self-adaptive multi-scale graph neural network for multi-band spectrum prediction. *IEEE Transactions on Cognitive Communications and Networking*, 2025, vol. 11, no. 3, p. 1442–1451. DOI: 10.1109/TCCN.2024.3483202

About the Authors ...

Hao SONG is a master's student at the School of Artificial Intelligence and Information Engineering, East China University of Technology. He received his bachelor's degree in Software Engineering from Henan University of Engineering in 2024. He is currently pursuing a master's degree in New-Generation Electronic Information Technology (including Quantum Technology, etc.) at East China University of Technology. His research interests include wireless spectrum sensing and spectrum prediction.

Yueshun HE, born in September 1971 in Yongzhou, Hunan Province, China, is a Professor and currently serves as the Deputy Director of the Institute of Advanced Studies in Resources and Environment at East China University of Technology. Prof. He was promoted to Associate Professor in 2003 and began supervising master's students that same year. He was subsequently promoted to Full Professor in 2007. He earned his Ph.D. in Computer Application Technology from Nanjing University of Aeronautics and Astronautics in 2010. With extensive experience in teaching and research within the field of computer science and technology, his primary research interests include big data and intelligent information processing, cyberspace security, artificial intelligence, as well as wireless sensor networks and Internet of Things (IoT) technology.

Linlin HE is a lecturer at the East China University of Technology. She received her master's degree in Cryptography from Xidian University in 2007. Her research interests include cyberspace security and intelligent information processing.

Xiaoyu CAO is a master's student at the School of Artificial Intelligence and Information Engineering, East China University of Technology. He received his bachelor's degree in Information and Computational Science from Hubei Polytechnic University in 2023. He is currently pursuing a master's degree in New-Generation Electronic Information Technology (including Quantum Technology, etc.) at East China University of Technology. His research interests include wireless spectrum sensing and spectrum prediction.

Yunzhe LIU is a master's student at the School of Artificial Intelligence and Information Engineering, East China University of Technology. He received his bachelor's degree in Software Engineering from the East China University of Technology in 2023. He is currently pursuing a master's degree in Computer Science and Technology at the East China University of Technology. His research interests include wireless spectrum sensing and spectrum prediction.

Jie CHEN is a doctoral student at the School of Surveying, Mapping and Information Engineering, East China University of Technology. He received his bachelor's degree from the East China University of Technology in 2021 and his master's degree in Computer Science and Technology from the East China University of Technology in 2024. He is currently pursuing a doctoral degree in Surveying Engineering at the School of Surveying, Mapping and Information Engineering, East China University of Technology. His research interests include few-shot learning, image processing, and spectrum sensing.