UDK 519.217 **DOI** 10.15622/ia.23.1.9

S. QIUSHI, H. YANG, O. PETROSIAN

GRAPH ATTENTION NETWORK ENHANCED POWER ALLOCATION FOR WIRELESS CELLULAR SYSTEM

Qiushi S., Yang H., Petrosian O. **Graph Attention Network Enhanced Power Allocation for Wireless Cellular System.**

Abstract. The importance of an efficient network resource allocation strategy has grown significantly with the rapid advancement of cellular network technology and the widespread use of mobile devices. Efficient resource allocation is crucial for enhancing user services and optimizing network performance. The primary objective is to optimize the power distribution method to maximize the total aggregate rate for all customers within the network. In recent years, graph-based deep learning approaches have shown great promise in addressing the challenge of network resource allocation. Graph neural networks (GNNs) have particularly excelled in handling graph-structured data, benefiting from the inherent topological characteristics of mobile networks. However, many of these methodologies tend to focus predominantly on node characteristics during the learning phase, occasionally overlooking or oversimplifying the importance of edge attributes, which are equally vital as nodes in network modeling. To tackle this limitation, we introduce a novel framework known as the Heterogeneous Edge Feature Enhanced Graph Attention Network (HEGAT). This framework establishes a direct connection between the evolving network topology and the optimal power distribution strategy throughout the learning process. Our proposed HEGAT approach exhibits improved performance and demonstrates significant generalization capabilities, as evidenced by extensive simulation results.

Keywords: MISO, cellular network, edge-feature, graph attention network, power allocation.

1. Introduction. In recent years, significant advancements have been witnessed in the technology employed for data transmission inside wireless networks. The contemporary landscape exhibits a smooth amalgamation of diverse mobile terminals within these networks, including cell phones and wearable gadgets. Consequently, this integration has resulted in a substantial surge in the number of simultaneous users. As a result, the cellular network's topography has witnessed a rise in the concentration of User Equipments (UEs) in highly populated areas, accompanied by a simultaneous reduction in cell size. The cellular environment is now experiencing a significant increase in wireless transmissions due to this phenomenon. The interference problem has become notably apparent as a direct result, manifesting inside individual cells and across cell boundaries. Due to the intricate nature of the circumstances, implementing strategies for interference management and rational power control has increased significance. These strategies mitigate any interference among users and enhance the network's overall efficiency. The gravity of this subject has engendered significant scholarly attention, leading to robust research endeavors to discover more efficacious solutions.

Considerable research endeavors have been devoted to formulating power allocation algorithms for cellular networks, whereas conventional methodologies frequently depend on numerical modeling methodologies. Decentralized and centralized optimization approaches have been prominent in power distribution in cellular networks, primarily due to the intrinsic complexity of this task, which is categorized as an NP-hard problem [1]. Using numerical simulation techniques requires a sequence of repetitive calculations using both implicit and explicit approaches. Noteworthy examples include genetic algorithms (GA), weighted minimum mean square error (WMMSE) [2], fractional programming (FP) [3], branch-and-bound techniques [4], and heuristic algorithms [5]. However, in scenarios involving large-scale networks, these iterative methods often come with extended computational time, rendering them unsuitable for real-time resource allocation. This limitation stems from the rapid fluctuations in the wireless channel's state, which can occur within a brief time span of milliseconds.

Several researchers have proposed the application of deep learning (DL) techniques as an innovative approach to power distribution, including supervised learning methods [6-8]. In contrast to traditional iterative methods, Convolutional Neural Networks (CNNs) and Multi-Layer Perceptrons (MLPs) have emerged as prominent architectural choices for capturing the underlying mapping functions in power control [9]. Once adequately trained, artificial neural networks can efficiently deduce solutions for various scenarios through feed-forward computations. However, it's crucial to recognize that the use of data-driven learning approaches necessitates the development of a benchmark algorithm for training, which can impose constraints on the resulting model's effectiveness. Additionally, while CNN- and MLP-based approaches offer rapid inference capabilities, they may not fully leverage the network's topological characteristics, which could require substantial training data and potentially limit their efficacy.

There have been proposals for GNN-based methodologies aimed at tackling the difficulties above. GNN can collect spatial information that is concealed inside the network topology. Moreover, they can extend this knowledge to other topologies, particularly in dynamic scenarios. The existing body of research demonstrates the efficacy of GNN-based approaches in addressing diverse challenges across various network contexts. The graph convolutional network (GCN) architecture is used to get close to the UWMMSE solution for the power allocation problem in the interference channel [10]. The power allocation technique, developed using GCN, aims to determine the best way to allocate resources in Federated Learning (FL) activities that take place in wireless networks with interference restrictions [11]. Interference

GCN (IGCNet) is introduced as a novel approach for addressing the power control problem in K-user interference channels [12]. In contrast to the conventional GCN, the model learning process is contingent upon incorporating the structural information neighborhood matrix encompassing the entirety of the graph. The Graph Attention Network (GAT) is a neural network model incorporating the widely used attention mechanism to calculate the weights between nodes in a graph [13]. This approach involves utilizing the feature representation of nodes rather than depending just on the structural information of the network, hence facilitating inductive learning for weight computation between nodes.

While previous studies have successfully designed homogeneous or heterogeneous GNNs, these designs have primarily focused on the node update mechanism, often overlooking the valuable information contained within edge characteristics. In conventional research, each transceiver pair in cellular networks has typically been represented as an individual node, with the channel state information of the direct communication connection serving as the corresponding node characteristic. The connections that disrupt communication between various pairs of transceivers are depicted as edges in the model, with the channel state information associated with each edge considered as the relevant edge feature [1]. To address the limitations outlined above, this study introduces the HEGAT framework as a potential solution for multiplexing the downlink of cellular users in Long-Term Evolution (LTE) systems. The primary contributions of this paper are as follows:

– We provide a comprehensive methodology for addressing power control challenges in cellular networks using heterogeneous graph topologies. In this theoretical framework, the communication links originating from the base station and connecting to user equipment are represented as vertices. In contrast, the interference links between different pairs of transceivers are denoted as edges. It's important to note that every node within the cellular network exhibits isomorphism, ensuring fair treatment. However, edges are categorized into two distinct types: intra-cell and inter-cell interference.

– We introduce an innovative method to enhance feature properties. It's crucial to highlight that the optimization of unknown variables primarily occurs within nodes. Consequently, when updating node features, we intentionally give due consideration to the attributes and classifications of the neighboring edges. This approach allows us to extract information from these two sets of adjacent edges using distinct methodologies. Building upon the edge feature enhancement technique, we propose the introduction of a neural network architecture named HEGAT. This architectural design exhibits a high degree

of effectiveness in mapping graph characteristics to node variables, thereby achieving the power allocation objective.

– The simulation results illustrate the effectiveness of the HEGAT approach in addressing power allocation challenges within cellular networks, particularly in the context of power distribution in interference channels. Comparative trials further substantiate that the proposed HEGAT surpasses existing state-of-the-art methods in terms of achieving higher sum rates. Moreover, HEGAT consistently delivers strong performance across a wide range of scenarios, encompassing varying numbers of BSs and UEs, diverse network characteristics, and different transmit power budgets.

The remainder of the paper is organized as follows. Section 2 introduces the related work. Section 3 presents the mathematical formulation of the power allocation problem in wireless cellular networks. Section 4 offers an exhaustive elucidation of the graph representation of the network, alongside a detailed exposition of our novel HEGAT algorithm. Section 5 presents the analytical simulation outcomes. Section 6 presents the discussion and conclusions.

2. Related Works

2.1. Edge Enhanced Graph. GNNs have demonstrated strong learning capabilities in tackling challenges associated with graph structures. The need to effectively handle edge information is prevalent in real-world scenarios. Message-passing neural networks (MPNNs) encompass several distinct stages, including message-passing and readout phases. Integral to this paradigm is the incorporation of node attributes. In parallel, the utilization of edge attributes for network characterization is also prevalent, and a technique has been proposed for updating these attributes to predict node features. While MPNNs incorporate edge information during the message-passing phase, their message-passing mechanism lacks the ability to capture knowledge about the topological relationships between nodes and edges. Relational Graph Convolution Networks (RGCN) employ forward-passing rules to enhance the weight matrix with additional edge-related weights [14]. However, empirical studies have shown that computing this simple aggregation doesn't significantly improve performance. Instead, the Edge Feature Graph Neural Network (EGNN) employs an aggregation function to combine node information while training separate attention weights for each feature dimension [15]. But this method may result in the loss of peripheral information. The Convolution with Edge-Node Switching Graph Neural Network (CensNet) utilizes a line graph structure to create an auxiliary graph [16]. This approach involves training the model on both the original graph and the line graph, allowing for the updating of node and edge embeddings. However, using an approximate spectrogram convolution during layer-by-layer propagation renders CensNet

inept in managing extensive directed graphs. Consequently, it is unsuited for wireless networks that are represented as directed graphs

2.2. Heterogenous Networks. Although homogeneous graphs have shown to be highly effective in radio network power control problems involving only one kind of node or edge, it is important to acknowledge that most scenarios are characterized by a diverse range of node and edge types [17]. Heterogeneous graphs have been found to be more advantageous than homogeneous graphs in addressing intricate radio resource management challenges. The Heterogeneous Interference Graph Neural Network (HIGNN) is specifically developed to effectively address diverse network scenarios characterized by heterogeneity [18]. The nodes in the communication network of Device-to-Device (D2D) connections are categorized based on the number of antennas present on the transmitters inside their respective links. The allocation of power is employed in D2D downlink systems. The introduction of the Heterogeneous Ultra-Dense Network (HUDN) aims to address the challenge of resource allocation in communication situations that involve a combination of D2D networks and cellular networks [19]. The communication linkages are considered as nodes, which are classified based on the sorts of devices they establish connections with. The introduction of Heterogeneous Graph Neural Network (HetGNN) aims to explore the Power Allocation strategy in Multi-Cell-Multi-User Systems [20]. The nodes in the network diagram represent the entities BS and UE, while the edges reflect the communication links between them.

One common characteristic seen in the aforementioned publications is their primary emphasis on the weights of edges, while neglecting the differences that arise from different types of edges. However, the focus of these studies is on the incorporation of different node types inside the framework of GNN. In contrast, the present study adopts a divergent methodology. The cellular network is represented as a directed graph, with a distinction made between two types of interfering links: intra-cell and inter-cell. In order to examine the various sorts of edges, we employ the HEGAT architecture. This architectural design is very suitable for effectively capturing and acquiring the intrinsic characteristics of edges, comprising both their properties and classifications.

2.3. Material and Methods. In this subsection we introduce the basic concepts related to resource allocation. MISO (Multiple-Input Single-Output): The MISO cellular network is a wireless communication technology that is employed in cellular networks and several other wireless applications. In a MISO system, the base station or access point is equipped with multiple transmit antennas (inputs), while the user's device, commonly referred to as the UE, is equipped with a single receive antenna (output). One notable attribute

of a MISO system is its capability to enable the base station to concurrently broadcast numerous data streams to a single user equipment, effectively using the spatial diversity offered by the multiple transmit antennas. The presence of spatial diversity has the potential to enhance the dependability and efficiency of wireless communication links, particularly in settings characterized by interference or signal fading. The utilization of MISO technology is prevalent in diverse wireless communication protocols, such as 4G LTE (Long-Term Evolution) and 5G, with the objective of augmenting the capacity and quality of wireless connections.

The Graph Neural Network (GNN): It is an innovative neural network structure capable of inferring the interdependencies between nodes in a graph utilizing propagating messages among the nodes. Hence, GNN have demonstrated their efficacy in tackling the learning task by employing a graphical structure, whereby each vertex incorporates feature information from neighboring vertices to derive a hidden state embedding based on graph perception. In practice, the hidden state embedding of each node is repeatedly updated by aggregating state information from its neighboring nodes. This study employs GNN to model the subnetwork system as a dynamic graph. Next, a two-stage attention method is employed to streamline the graph and identify the probable interference link across subnetworks across different dimensions.

3. System Model And Graph Representation

3.1. System Model. We consider a classical scenario of downlink multicell communication, a massive MISO network with M-antennas BSs and single-antenna UEs. The PA problem in the cellular network is with the setting of interfering multiple-access channels (IMAC) [21]. All BSs within the network coverage area simultaneously serve all UEs. However, since different cells use the same frequency, the UEs are still subject to inter-cell and intra-cell interference. Index the BSs as $\mathcal{N} = \{1,...,N\}$ and UEs as $\mathcal{K} = \{1,...,K\}$. Denote D_{nk} as the set of k-th UE's neighbour UEs in the n-th cell, denote C_n as the set of n-th cell's neighbour cells. Assume that $n \in \mathcal{N}$, $k \in \mathcal{K}$, $k' \in D_{nk}$, $n' \in C_n$, then the received signal of the k-th UE from n-th BS in n-th cell can be formulated by:

$$y_{nk} = \underbrace{g_{n,nk}^H w_{nk}\sqrt{p_{nk}}s_{nk}}_{\text{desired signal}} + \underbrace{\sum_{k'\neq k} g_{n,nk}^H w_{nk'}\sqrt{p_{nk'}}s_{nk'}}_{\text{intra-cell interference}} + \underbrace{\sum_{n'\neq n}\sum_{k'} g_{n',nk}^H w_{n'k'}\sqrt{p_{n'k'}}s_{n'k'}}_{\text{inter-cell interference}} + z_{nk}, \quad (1)$$

where $g_{n,nk}$ denote the the channel response from n-th BS to k-th UE in n-th cell, p_{nk} denotes the corresponding transmit power. $s_{nk} \sim \mathscr{U}(0,1)$ is the transmit signal. $z_{nk} \sim \mathscr{N}\left(0,\sigma^2\right)$ is the additive white Gaussian noise (AWGN). The coordinated beamforming (CB) vector from n-th BS to k-th UE is denoted as w_{nk}. The literature on Coordinated Multi-Point (CoMP) CB has explored multiple schemes. In this study, we have opted to utilize the zero-forcing beamforming scheme [22] to simplify the problem. Then the signal-to-interference-plus-noise ratio (SINR) of k-th UE can be calculated as:

$$\gamma_{nk} = \frac{g_{nk,nk} p_{nk}}{\sum_{k' \neq k} g_{nk,nk'} p_{nk'} + \sum_{n' \neq n} \sum_{k} g_{nk,n'k'} p_{n'k'} + \sigma^2}, \tag{2}$$

where $g_{nk,nk} = \left| g_{n,nk}^H w_{nk} \right|^2$ denote independent channel gain of the desired signal. $g_{nk,nk'} = \left| g_{n,nk}^H w_{nk'} \right|^2$ denote channel gain of intra-cell interference from neighbour UEs in n-th cell. $g_{nk,n'k'} = \left| g_{n',nk}^H w_{n'k'} \right|^2$ denote channel gain of inter-cell interference from neighbour UEs in n-th cell's adjacent cells.

The downlink rate of communication link nk can be expressed in terms of normalized bandwidth as:

$$C_{nk} = \log_2 (1 + \gamma_{nk}). \tag{3}$$

The primary aim of this study is to identify the ideal power level that optimizes the overall network sum rate, while adhering to the limitation of a maximum power limit for each transmitter. The provided problem may be expressed as:

$$\begin{aligned} &\max_{p_{nk}, w_{nk}} \sum_{n=1}^{N} \sum_{k=1}^{K} \log_2 (1 + \gamma_{nk}) \\ &\text{s.t. } 0 \leqslant p_{nk} \leqslant p_{\max}, \forall n \in \mathscr{N}, k \in \mathscr{K}. \end{aligned} \tag{4}$$

The objective function presents a challenging obstacle in the form of a nonconvex nonlinear optimization problem, which is further complicated by the presence of constraints. As a result, finding the global optimal solution becomes a complex endeavor [23]. Heuristic algorithms have the capability to approximate solutions that are globally optimum, but at the

cost of substantial computing resources. In order to cater to the demand for real-time applications that require low-complexity solutions, we suggest using deep learning techniques to parameterize these solutions. Although attaining theoretical optimality for learnt solutions may pose challenges, actual evidence continually shows that deep learning methods often produce extremely satisfying performance results.

3.2. Graph Representation. We model the multicell cellular network (Figure 1) as the fully connected weighted directed graph (Figure 2) which can be expressed as $G = (V, E)$. Each BS-to-UE communication link is considered as a node, denoting V as a set of nodes, $V = \{v_{nk}\}, n \in \mathcal{N}, k \in \mathcal{K}$. In contrast, each interfering link is considered as an edge, denoting E as a set of edges, $E = \{e_{nk,ml}\}, n,m \in \mathcal{N}, k,l \in \mathcal{K}, nk \neq ml$. The attributes such as distance, channel information, and weight associated with the communication link are node characteristics. The assigned power is used as the predicted node label. Attributes such as distance and channel information associated with the interfering links are considered edge features. The notations $\phi \to \mathbb{C}^{d_V}$ and $\psi \to \mathbb{C}^{d_E}$ represent the mapping of nodes and edges to their respective features, where d_V and d_E represent the dimension of feature space for node and edge, respectively.

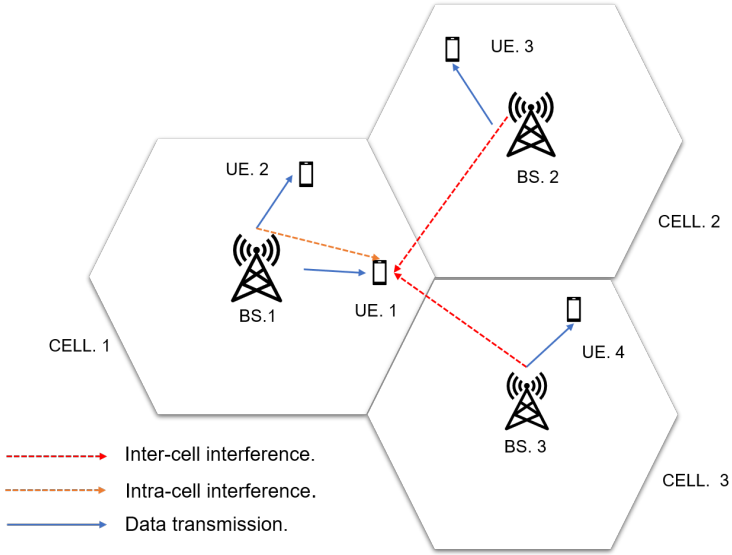

Fig. 1. Example of cellular

To train the GAT model, the graph in Figure 2 is used as input, and the iterations of the neural network are used to learn the vector representation of each network.

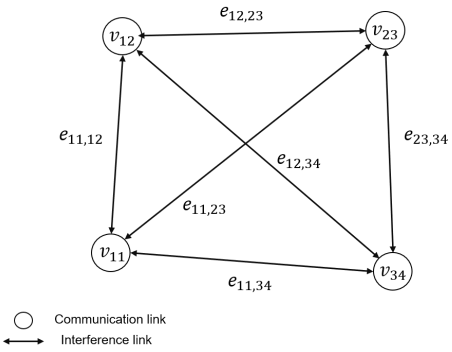

Fig. 2. Graph representation

4. Heterogenous Edge feature Enhanced Graph Attention Network

4.1. Feature Concate. Prior to integrating neighbor information into the heterogeneous graph, it is crucial to recognize that the connections between neighbors and each node, whether through nodes or edges, may possess distinct functionalities and varying degrees of importance in the task of generating node embeddings for resource allocation [24].

In the context of this work, we propose the utilization of node feature attention in conjunction with the augmentation of edge features. The proposed strategic combination is formulated as an effective approach to carefully evaluate the importance of both node-based and edge-based neighbors. By integrating the attention mechanism, it becomes possible to generate node and edge embeddings, effectively capturing the representations of these significant neighboring entities. This methodology enables us to reveal their unique functions and contributions inside the complex framework of resource allocation procedures [25].

The presence of diverse edges within a network introduces a spectrum of variations in their characteristics. As a result, different types of edges may manifest distinctive properties. These features encompass both attributes and types, where attributes are expressed by continuous and intricate variables, while types are denoted as discrete variables. The attribute of the edge $e_{nk,ml}$ denoted by the vector $e^{attr}_{nk,ml}$, and the type of the edge $e_{nk,ml}$ is denoted by the vector $e^{type}_{nk,ml}$. The features of the edge $e_{nk,ml}$ can be represented as

$\mathrm{e}_{nk,ml} = \left[\mathrm{e}^{attr}_{nk,ml} \middle\| \mathrm{e}^{type}_{nk,ml}\right]$, which is obtained by concatenating its transformed node attributes and type information.

4.2. Edge Feature Enhanced Node Attention Layer. In the heterogeneous graph, the definition of node-based neighbors N_{nk} refers to the nodes directly connecting to a particular node v_{nk}. It is essential to acknowledge that the neighbors of a certain node encompass the node itself. Similarly, the edge-based neighbors $E_{nk,ml}$ refers to the set of edge $e_{nk,ml}$ that directly connect to a certain node v_{nk} [26].

As shown in Figure 3, circles represent nodes, and squares represent edges. Take v_{11} for example, the node-based neighbors of node v_{11} are v_{11}, v_{12}, v_{23}, and v_{34}. Note that the node-based neighbors of node v_{11} contain itself; The edge-based neighbors of node v_{11} are $e_{11,12}$, $e_{11,23}$, and $e_{11,34}$.

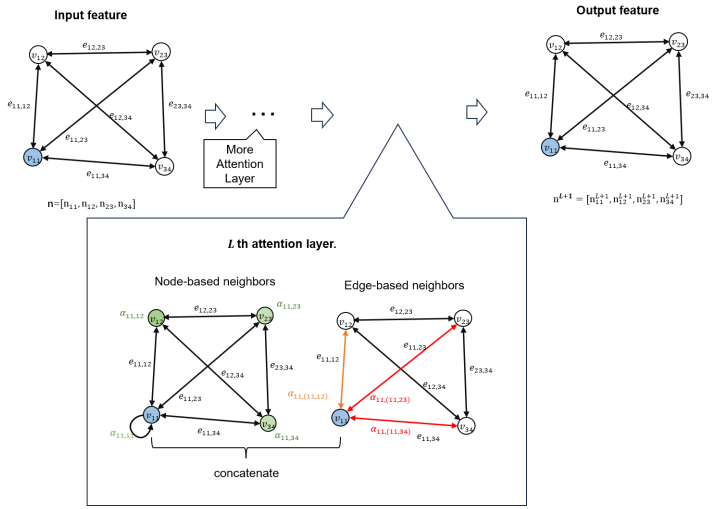

Fig. 3. Architecture of HEGAT. The model consists of multiple HEGAT layers. Each iteration incorporates the features of the first-order neighbors (nodes and edges) to generate new node features

It is necessary to map these vectors to a high-dimensional space using the GAT layer to enhance the network topology information contained in the original low-dimensional features $\mathrm{n_{nk}} = \left[w_{nk}, p_{nk}, g_{nk,nk}\right]$ obtained from the network. The input to the GAT network contains a set of node features denoted as $\mathrm{n} = (\mathrm{n}_{11}, ., \mathrm{n}_{nk}, .., \mathrm{n}_{NK}), \mathrm{n}_{nk} \in R^{d_v}$, where each n_{nk} is the feature of existing BS-to-UE pair v_{nk}.

Theoretically, GAT can use all nodes other than the central node to calculate the similarity with that central node. The dimensions of the node and edge feature spaces may vary. To enhance the expression of the node features, the layer of parameter-sharing neural networks is added to linearly transform the features (whether for nodes or edges), and the network is denoted by Att_{node} and Att_{edge}. Define $W_{\mathrm{n}} \in \mathbb{R}^{d_V \times d'_V}$ and $W_{\mathrm{e}} \in \mathbb{R}^{d_E \times d'_E}$ as the learnable weight matrix transforming input features linearly into high-level features, the dimension for the feature vector of the node is made to change from d_V to d'_V, the dimension for the feature vector of the edge is made to change from d_E to d'_E. The self-attention mechanism is used to calculate the similarity between the central node and the neighbor nodes, where the similarity is calculated by a layer of the neural network the parameters are denoted by β^T, and the two transformed feature vectors are fed into this network after stitching. The importance coefficient of node v_{ml} to node v_{nk} is expressed as:

$$\mathrm{c}_{nk,ml} = Att_{node}\left(W_{\mathrm{n}}\mathrm{n}_{nk}, W_{\mathrm{n}}\mathrm{n}_{ml}\right). \tag{5}$$

The importance coefficient of edge $e_{nk,ml}$ to node v_{nk} is expressed as:

$$\mathrm{c}_{nk,(nk,ml)} = Att_{edge}\left(W_{\mathrm{n}}\mathrm{n}_{nk}, W_{\mathrm{e}}\mathrm{e}_{ml}\right). \tag{6}$$

In this context, we apply the edge feature enhanced attention mechanism to calculate the attention weights associated with the neighbors of each one-degree node and edge. This calculation considers not only the features of the neighboring nodes, but also incorporates the features of the adjacent edges. Throughout this process, the features are jointly combined, with their respective parameters governed by the attention vector β^T and the activation function *LeakyReLU* . The resulting weight which is normalized using the softmax function for node v_{ml} is expressed as:

$$\alpha_{nk,ml} = softmax_{ml}\left(\mathrm{c}_{nk,ml}\right) = \left(\frac{\exp\left(\sigma\left(\beta_n^T\left[W_{\mathrm{n}}\mathrm{n}_{nk} \| W_{\mathrm{n}}\mathrm{n}_{ml}\right]\right)\right)}{\sum_{rs\in\mathcal{N}_j}\exp\left(\sigma\left(\beta_n^T\left[W_{\mathrm{n}}\mathrm{n}_{nk} \| W_{\mathrm{n}}\mathrm{n}_{rs}\right]\right)\right)}\right), \tag{7}$$

where $\beta_n^T \in \mathbb{R}^{2d'_V}$ denotes the attention vector for node-based neighbors. The attention mechanism is implemented as a single-layer feedforward neural network, which is characterized by the parameters W and β. This neural network utilizes the *LeakyReLU* nonlinearity with a negative slope of 0.2. The variable $\alpha_{nk,ml}$ refers to the attention score that is used to quantify the significance of a neighboring node v_{ml} concerning node v_{nk}.

Similarly, the resulting weights which are normalized using the softmax function foredge $e_{nk,ml}$ is expressed as:

$$\begin{aligned}\alpha_{nk,(nk,ml)} &= softmax_{(nk,ml)}\left(\mathrm{c}_{nk,(nk,ml)}\right)\\ &= \left(\frac{\exp\left(\sigma\left(\beta_e^T\left[W_{\mathrm{n}}\mathrm{n}_{nk}\left\|W_{\mathrm{e}}\mathrm{e}_{nk,ml}\right.\right]\right)\right)}{\sum_{(rs,tu)\in\mathscr{E}_i}\exp\left(\sigma\left(a_e^T\left[W_{\mathrm{n}}\mathrm{n}_{nk}\left\|W_{\mathrm{e}}\mathrm{e}_{rs,tu}\right.\right]\right)\right)}\right),\end{aligned} \tag{8}$$

where $\beta_e^T \in \mathbb{R}^{d'_V+d'_E}$ denotes the attention vector for edge-based neighbors. The variable $\alpha_{nk,(nk,ml)}$ refers to the attention score that used to quantify the significance of a neighboring edge $e_{nk,ml}$ concerning node v_{nk}.

Hence, the node-based neighbors' embedding of node v_{nk} is computed by aggregating the features of its first-order neighbors and weighted by the attention score:

$$\mathrm{n}_{nk}\left[\mathscr{N}_{nk}\right] = \sigma\left(\sum_{ml\in\mathscr{N}_{nk}}\alpha_{nk,ml}W_{\mathrm{n}}\mathrm{n}_{ml}\right), \tag{9}$$

where σ is the activation function, n_{nk}^{L+1} is the new feature of each vertex v_{nk} aggregated with neighborhood information.

Similarly, the edge-based neighbors' embedding of node v_{nk} is expressed as:

$$\mathrm{n}_{nk}\left[\mathscr{E}_{nk}\right] = \sigma\left(\sum_{(nk,ml)\in\mathscr{E}_{nk}}\alpha_{nk,(nk,ml)}W_{\mathrm{e}}\mathrm{e}_{(nk,ml)}\right). \tag{10}$$

To make the learning process of the self-attention mechanism more robust, we use a multi-headed attention mechanism, which can be viewed as multiple single-headed attentions executed independently in parallel and averaged as the output. Then $\mathrm{n}_{nk}^{L+1}\left[\mathscr{N}_{nk}\right]$ and $\mathrm{n}_{nk}^{L+1}\left[\mathscr{E}_{nk}\right]$ can be calculated as:

$$\mathrm{n}_{nk}\left[\mathscr{N}_{nk}\right] = \sigma\left(\frac{1}{Q}\sum_{q=1}^{Q}\sum_{ml\in\mathscr{N}_{nk}}\alpha_{nk,ml}^{q}W_{\mathrm{n}}^{q}\mathrm{n}_{ml}\right), \tag{11}$$

$$\mathrm{n}_{nk}\left[\mathscr{E}_{nk}\right] = \sigma\left(\frac{1}{Q}\sum_{q=1}^{Q}\sum_{(nk,ml)\in\mathscr{E}_{nk}}\alpha_{nk,(nk,ml)}^{q}W_{\mathrm{e}}^{q}\mathrm{e}_{(nk,ml)}\right), \tag{12}$$

where the variable Q denotes the number of attention heads, while W^q signifies the weight matrix shared by the q th attention head. The aggregation function

can be classified as either tandem or average. Specifically, when conducting multiple tasks at the final stage, the method of averaging is employed.

At last, by integrating the edge-based neighbors' embedding $\mathrm{n}_{nk}[\mathscr{E}_{nk}]$ and node-based neighbors' embedding $\mathrm{n}_{nk}[\mathscr{N}_{nk}]$, the embedding of node v_{nk} in the $(L+1)$-th layer is expressed as:

$$\mathrm{n}_{nk}^{L+1} = concat\left(\mathrm{n}_{nk}^{L}[\mathscr{N}_{nk}] \,\|\, \mathrm{n}_{nk}^{L}[\mathscr{E}_{nk}]\right), \tag{13}$$

where *concat* denotes the operation of concatenation.

Upon the integration of the previously discussed components, the resulting node representation becomes accessible, thus facilitating its application in a range of subsequent tasks. HEGAT is trained through a supervised learning methodology that considers task-specific attributes of nodes. The model's weights are effectively refined using back-propagation and gradient descent techniques. The objective of this process is to minimize cross-entropy, ultimately yielding meaningful node embeddings for heterogeneous networks.

4.3. Training Samples Generation. The current research study makes a suggestion for a semi-supervised training strategy that may be used for the HEGAT-based architecture. This method requires a significant quantity of labeled training data. According to a study that compared many stochastic search-based algorithms to discover the estimated optimum power allocation, the Particle Swarm Optimization (PSO) approach outperformed all others in a range of application circumstances. This was determined by comparing the algorithms' performance. A collection of candidate solutions for the optimization issue that is being considered in this work is first generated at random. These candidate solutions are then moved throughout the search space according to rules that have been defined, and the search is guided by optimal positions that are already known to exist. The goal of this process is to identify the best option for allocating resources. When new prospective solutions are found, their locations are mapped out and used to influence the movement of those that are close by. After a number of iterations, the PSO approach produces outcomes that are very close to being optimum. These replies may be used as labels for the expected graph locations, and we record the properties of the sample graph, including information about the graph's structure and channels, at the positions corresponding to those labels.

The graph neural network we use learns resource allocation strategies through semi-supervised learning. We run the simulated environment and record a graph representation of the network realization, including the graph topology and channel states. The graph representation of each time point

is regarded as an independent sample, regardless of the dependence of the user's state on the front and back time, and the users in the network are randomly distributed. Among them, the power allocation and beamforming vector as optimization variables are used as features of each vertex and are also variables predicted by the graph neural network. The loss function L is the negative expectation of utility function over different channel realizations: $\mathscr{L}(\theta) = -E_H\left[\sum_{i,m} B\log(1+\gamma_{nk})\right]$, which updates the model parameters θ of HEGAT by backpropagation of the neural network. This is an optimization process of a static graph. The graph neural network learns the mapping relationship between the topology of the graph and the channel state to the optimal resource allocation strategy. Essentially graph neural networks are data-driven optimization methods, and whether or not the sample contains as comprehensive a network realization as possible determines the upper bound of the model.

5. Illustrative Results

5.1. Simulation Environment Setting. Through the use of numerical simulations, this part evaluates the effectiveness of the proposed GAT framework for solving the challenge of maximizing the total rate. We evaluate the suggested methodology in comparison to a variety of standards derived from past research. The computer that was utilized to do the simulation studies for this research has a Central Processing Unit (CPU) with a speed of 3.19 GHz and 32 Gigabytes of Random Access Memory (RAM). In the recommended method, Python 3.6 is utilized as the platform for the simulation, while Pytorch is the tool of choice for the development of the neural network. Table 1 presents both the cellular network and the neural network's corresponding parameters for your perusal.

Table 1. Simulation parameters

Number of BS	9
Average users per cell	4
The Doppler frequency	10 HZ
Minimum allocated power	5 dBm
Maximum allocated power	38 dBm
Inner space distance	0.01 km
Half cell-to-cell distance	1 km
Noise power spectral density	− 114 dBm/Hz
Learning rate	0.0001
Number of layers	3
Optimizer	ADAM
Batch size	32
Epoch count	300

5.2. Performance Comparison. In order to evaluate the efficacy of the proposed HEGAT, a thorough comparison study is conducted, encompassing many cutting-edge baseline approaches. The present analysis incorporates many strategies, such as network embedding approaches and techniques based on graph neural networks.

– GAT: GAT incorporates an attention mechanism that operates on homogeneous graphs. In this study, we conduct an evaluation of several meta-paths and provide the findings on the optimal performance achieved.

– GCN: The GCN architecture proposed in [11] captures the deep features of topology in the network. The topology and channel information are concatenated as input features, and then the model is trained offline in a semi-supervised manner.

– Multi-Layer Perception (MLP): A fully connected neural network is designed to build a mapping from channel information to power allocation through the data of the Euclidean structure. This approach has been studied in [27].

The performance evaluation of HEGAT is conducted as the network scale expands, in comparison to benchmarks. As illustrated in Figure 4, the average sum rate per UE achieved by HEGAT exhibits notable enhancements of 9.14%, 9.68%, and 10.2% respectively, as opposed to the leading benchmark GAT, with the increase in the number of cell pairs from 4 to 16.

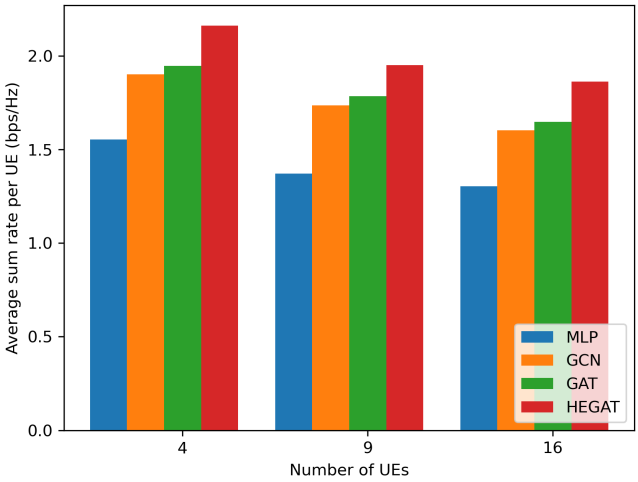

Fig. 4. Average sum rate per UE versus cellular network scale

Notably, HEGAT demonstrates robust generalization capabilities while simultaneously achieving optimal performance. The effectiveness of the proposed edge feature enhanced mechanism can be attributed to its ability to proficiently extract features from the channel state of the edge attributes, resulting in a further augmentation of the effectiveness of the initial node update approach.

The performance of each algorithm for various cellular network densities is shown in Figure 5. The simulation takes into consideration the inherent fluctuations in user density, reflecting the real-world intricacies of time and spatial factors. The user density is subject to variations contingent upon the number of UEs present within a specific cell, and this value ranges from 1 to 6. As the user density escalates, a concomitant decline in the average total rate becomes evident. This consistent trend is observed across all algorithms, implying a shared behavior among them. To assess the performance of the proposed technique in high-density networks, we change the number of UEs in the cell while maintaining the other parameters constant (N=16). The average sum rate per UE drops when network density rises as a result of the rise in interference. It is important to note that the GAT algorithm continues to perform better than other algorithms at boosting the overall network rate.

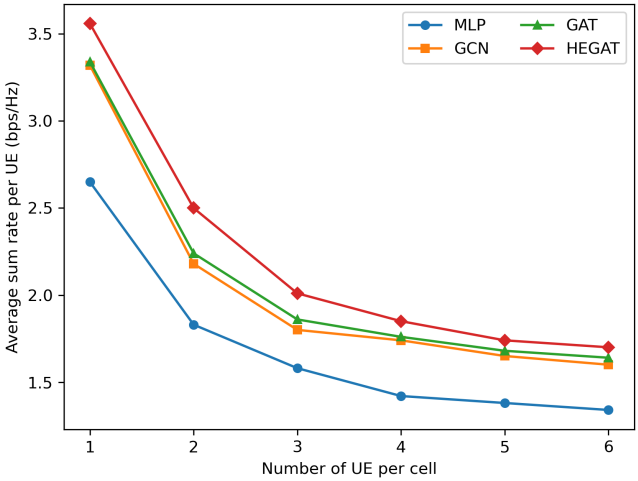

Fig. 5. Average sum rate per UE versus network density

Figure 6 illustrates the performance of the algorithms with respect to different cellular network radii. In this section, we explore the variability of the range between half a BS degree and a full BS degree, which is represented as R_{max}. In general, it is observed that a decrease in cell range has a tendency to amplify both intra-cell and inter-cell interference, resulting in a decrease in the average total rate. The aforementioned tendency is notably apparent in the overall rate performance, as the random and maximal power tactics demonstrate the least effectiveness. The mean rate per UE increases as the radius increases within the range of 0.2 km to 0.4 km for the half-cell length. The mean rate per UE does not significantly increase beyond a distance of 0.4 km, which corresponds to half the length of a cell. The data clearly indicates that the influence of cell radius on the overall rate enhancement has significantly lessened. The performance of the HEGAT algorithm across different network settings illustrates its effectiveness and reliability.

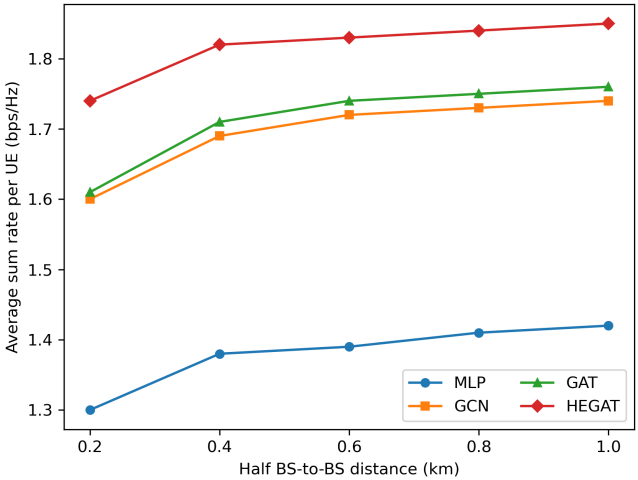

Fig. 6. Average sum rate per UE versus cell length

As shown in Figure 7, the HEGAT demonstrates a runtime that is similar to that of previous benchmark algorithms. The comparable runtime of these approaches may be ascribed to the utilization of the same network architecture and input characteristics. However, the strategy that has been suggested exhibits a higher level of time consumption in comparison to alternative benchmark techniques. The reason for this is the increased amount of time needed to

get and combine the embedding information of both nodes and edges. Every node iteratively enhances its own embedding characteristics together with the embedding features of its neighboring nodes and edges. When the quantity of BS reaches 16, HEGAT exhibits a computational duration of 0.01 seconds, which is notably below the required timeframe for decision-making in wireless systems, namely 0.02 seconds. This solution satisfactorily addresses the need of making real-time judgments. Therefore, the proposed methodology exhibits effectiveness in enabling immediate implementation into wireless networks.

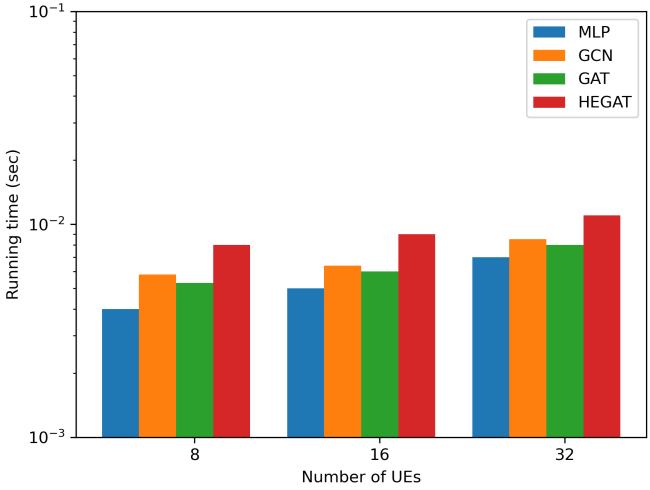

Fig. 7. Computation time comparison

As shown in Figure 8, the doppler frequency, denoted as f_d, is a crucial quantity that is closely associated with the phenomena of small-scale fading. The incorporation of real-time data into our suggested data-driven methodology for instantaneous power distribution may encounter obstacles arising from the swift variations in signal intensity. The observed oscillations have the potential to result in a decrease in overall performance. The Doppler frequency is seen to occur within a frequency range that extends from 4 Hz to 18 Hz. The simulation results presented in Figure 8 reveal a steady decrease in the average sum rate associated with the data-driven algorithm as the values of f_d span throughout this frequency range. This fascinating discovery suggests that the data-driven model has the ability to withstand the changes offered by f_d.

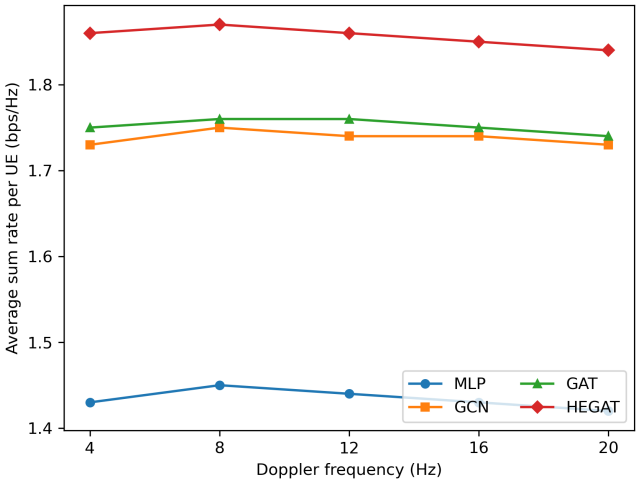

Fig. 8. Average sum rate per UE versus doppler frequency

6. Discussion and Conclusion

6.1. Discussion. From the perspective of the convolution operator, the graph neural network can be divided into two categories, spectral methods and spatial methods.

Spectral methods: e.g., GCN, depend on Laplacian eigenbasis of the graph, which is hard to calculate for a large graph. Its main drawbacks: 1) the edge weights are fixed when fusing, not flexible enough; 2) poor generalizability because it is a full graph convolutional fusion, the whole graph does gradient update, and when the graph is larger, such an approach is too slow and inappropriate.

Spatial methods: e.g., GAT, perform information aggregation only on the local neighborhood, avoiding heavy calculation of Laplacian eigenbasis. The HEGAT algorithm proposed in this paper is based on the attention aggregation mechanism of GAT, and the performance degradation is weaker and more stable than that of the spectral graph-based neural network when the graph structure is changed.

Different graph tasks may have different preferences for node and edge features. There are some graphs where node features have a greater impact on the graph and some graph edges where edge features have a greater impact on the graph. Whereas, in the graph representation of cellular networks, direct

links from BS-to-UE are considered as nodes and channel gain of the desired signal is considered as node features. The interference links between different BS-to-UEs are regarded as edges, and channel gain of interference is regarded as node features. While the latter is computed from the former and the topology of the graph, the edge characteristics are strongly correlated with the node characteristics. So essentially the graph representation in cellular networks is still a node-sensitive task. Edge features are used as a complement to enhance the performance of the graph neural network.

6.2. Conclusion. This study introduces a framework called HEGAT, which has been developed to tackle the issue of power regulation in cellular networks. The cellular network is represented as a heterogeneous directed graph, and the network parameters of HEGAT are trained using a semi-supervised learning technique. The HEGAT framework successfully incorporates node and edge properties, resulting in a notable improvement in node and edge embeddings across various neural network layers. This technique effectively captures the significance of adjacent nodes and edges inside the network. The experimental findings derived from thorough studies conducted on huge datasets of mobile traffic provide evidence of the almost ideal performance of our proposed strategy. Moreover, the system demonstrates robust generalization abilities across many network conditions. In anticipation of future endeavors, our research agenda includes the investigation of complex heterogeneous network settings, enhancement of current approaches, and the incorporation of multi-source datasets to attain higher levels of prediction accuracy. In future work, we will continue research in this area to verify the versatility of HEGAT in different network scenarios.

References

1. Jiang W. Graph-based deep learning for communication networks: A survey. Computer Communications. 2022. vol. 185. pp. 40–54.
2. Shi Q., Razaviyayn M., Luo Z.Q., He C. An iteratively weighted MMSE approach to distributed sum-utility maximization for a MIMO interfering broadcast channel. IEEE Transactions on Signal Processing. 2011. vol. 59. no. 9. pp. 4331–4340.
3. Shen K., Yu W. Fractional programming for communication systems – Part I: Power control and beamforming. IEEE Transactions on Signal Processing. 2018. vol. 66. no. 10. pp. 2616–2630.
4. Feiten A., Mathar R., Reyer M., Rate and power allocation for multiuser OFDM:An effective heuristic verified by branch-and-bound. IEEE Transactions on Wireless Communications. 2008. vol. 7. no. 1. pp. 60–64.
5. Sun Q., Wu H., Petrosian O. Optimal Power Allocation Based on MetaheuristicAlgorithms in Wireless Network. Mathematics. 2022. vol. 10(18). no. 3336.
6. Sun H., Chen X., Shi Q., Hong M., Fu X., Sidiropoulos N.D. Learning to optimize: Training deep neural networks for interference management. IEEE Transactions on Signal Processing. 2018. vol. 66. no. 20. pp. 5438–5453.

7. Liang F., Shen C., Yu W. and Wu F. Towards optimal power control via ensemblingdeep neural networks. IEEE Transactions on Communications. 2019. vol. 68. no. 3. pp. 1760–1776.
8. Lee W., Kim M., Cho D.H. Deep learning based transmit power control in underlaid device-to-device communication. IEEE Systems Journal. 2018. vol. 13. no. 3. pp. 2551–2554.
9. Liao X., Shi J., Li Z., Zhang L., Xia B. A model-driven deep reinforcement learning heuristic algorithm for resource allocation in ultra-dense cellular networks. IEEE Transactions on Vehicular Technology. 2019. vol. 69. no. 1. pp. 983–997.
10. Shen Y., Shi Y., Zhang J., Letaief K.B. A graph neural network approach for scalable wireless power control. IEEE Globecom Workshops (GC Wkshps). 2019. pp. 1–6.
11. Chowdhury A., Verma G., Rao C., Swami A., Segarra S. Unfolding WMMSE using graph neural networks for efficient power allocation. IEEE Transactions on Wireless Communications. 2021. vol. 20. no. 9. pp. 6004–6017.
12. Li B., Swami A., Segarra S. Power allocation for wireless federated learning using graph neural networks. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). 2022. pp. 5243–5247.
13. Wang X., Ji H., Shi C., Wang B., Ye Y., Cui P., Yu P.S. Heterogeneous graph attention network. The world wide web conference. 2019. pp. 2022–2032.
14. Busbridge D., Sherburn D., Cavallo P., Hammerla N.Y. Relational graph attention networks. arXiv preprint. 2019. arXiv:1904.05811.
15. Gong L., Cheng Q. Exploiting edge features for graph neural networks. Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2019. pp. 9211–9219.
16. Jiang X., Ji P., Li S. CensNet: Convolution with Edge-Node Switching in Graph Neural Networks. IJCAI. 2019. pp. 2656–2662.
17. Fu X., Zhang J., Meng Z., King I. Magnn: Metapath aggregated graph neural network for heterogeneous graph embedding. Proceedings of The Web Conference. 2020. pp. 2331–2341.
18. Zhang X., Zhao H., Xiong J., Liu X., Zhou L., Wei J. Scalable power control/beamforming in heterogeneous wireless networks with graph neural networks. IEEE Global Communications Conference (GLOBECOM). 2021. pp. 01–06.
19. Guo J., Yang C. Learning power allocation for multi-cell-multi-user systems with heterogeneous graph neural networks. IEEE Transactions on Wireless Communications 2021. vol. 21. no. 2. pp. 884–897.
20. Khodmi A., Rejeb S.B., Agoulmine N., Choukair Z. A joint power allocation and user association based on non-cooperative game theory in an heterogeneous ultra-dense network. IEEE Access. 2019. vol. 7. pp. 111790–111800.
21. Challita U., Saad W., Bettstetter C. Interference management for cellular-connected UAVs: A deep reinforcement learning approach. IEEE Transactions on Wireless Communications. 2019. vol. 18. no. 4. pp. 2125–2140.
22. Nguyen L.D., Tuan H.D., Duong T.Q., Poor H.V. Multi-user regularized zero-forcing beamforming. IEEE Transactions on Signal Processing. 2019. vol. 67. no. 11. pp. 2839–2853.
23. Li Y., Han S., Yang C. Multicell power control under rate constraints with deep learning. IEEE Transactions on Wireless Communications. 2021. vol. 20. no. 12. pp. 7813–7825.
24. Wang X., Ji H., Shi C., Wang B., Ye Y., Cui P., Yu P.S. Heterogeneous graph attention network. The world wide web conference. 2019. pp. 2022–2032.
25. Yang Y., Li D. Nenn: Incorporate node and edge features in graph neural networks. Asian conference on machine learning. 2020. pp. 593–608.

26. Chen J., Chen H. Edge-featured graph attention network. arXiv preprint. 2021. arXiv:2101.07671.

27. Kim J., Park J., Noh J., Cho S. Autonomous power allocation based on distributed deep learning for device-to-device communication underlaying cellular network. IEEE access. 2020. vol. 8. pp. 107853–107864.

Qiushi Sun — Graduate student, Faculty of applied mathematics and control process, St. Petersburg State University. Research interests: mathematical statistics, system control, intelligent manufacturing. The number of publications — 3. st059656@student.spbu.ru; 35, Universitetskiy Av., 198504, St. Petersburg, Peterhof, Russia; office phone: +7(911)740-8019.

Yang He — Graduate student, Faculty of applied mathematics and control process, St. Petersburg State University. Research interests: mathematical statistics, system control, intelligent manufacturing. The number of publications — 1. hy1186867324@outlook.com; 35, Universitetskiy Av., 198504, St. Petersburg, Peterhof, Russia; office phone: +7(812)428-4530.

Petrosian Ovanes — Ph.D., Professor of the department, Department of mathematical modeling of energy systems, St. Petersburg State University. Research interests: machine learning, reinforcement Learning, time series analysis, graph theory, discrete probability models and stochastic processes and game theory, operations research. The number of publications — 81. o.petrosyan@spbu.ru; 35, Universitetskiy Av., 198504, St. Petersburg, Peterhof, Russia; office phone: +7(812)428-4530.

Acknowledgements. This work was supported by St. Petersburg State University (grant 94-06-2114).

DOI 10.15622/ia.23.1.9

С. Цюши, Х. Ян, О.Л. Петросян

РАСПРЕДЕЛЕНИЕ МОЩНОСТИ В БЕСПРОВОДНОЙ СОТОВОЙ СИСТЕМЕ С ПРИМЕНЕНИЕМ ГРАФОВОЙ СЕТИ ВНИМАНИЯ

Цюши С., Ян Х., Петросян О.Л. **Распределение мощности в беспроводной сотовой системе с применением графовой сети внимания.**

Аннотация. С быстрым развитием технологии сотовых сетей и распространением мобильных устройств эффективная политика распределения сетевых ресурсов становится все более важной для улучшения пользовательских услуг и производительности сети. Наша цель – максимизировать суммарную мощность всех пользователей сети путем нахождения оптимальной схемы распределения мощности. В последние годы методы глубокого обучения на основе графов продемонстрировали большой потенциал для решения проблемы распределения сетевых ресурсов. Из-за топологической природы мобильных сетей графовые нейронные сети (GNN) могут лучше работать с данными, структурированными в виде графов. Однако большинство из этих методов фокусируются только на узловых функциях в процессе обучения и часто игнорируют или упрощают граничные функции, которые играют не менее важную роль, чем узлы. Чтобы решить эту проблему, мы предлагаем дизайн сети с расширенным графическим вниманием (HEGAT), который напрямую связывает изменяющуюся топологию сети и наилучший способ распределения мощности во время обучения. Обширные результаты моделирования подтверждают превосходную производительность и надежные возможности обобщения, демонстрируемые предлагаемой методологией HEGAT.

Ключевые слова: МИСО, сотовая сеть, пограничная функция, сеть графического внимания, распределение мощности.

Литература

1. Jiang W. Graph-based deep learning for communication networks: A survey. Computer Communications. 2022. vol. 185. pp. 40–54.
2. Shi Q., Razaviyayn M., Luo Z.Q., He C. An iteratively weighted MMSE approach to distributed sum-utility maximization for a MIMO interfering broadcast channel. IEEE Transactions on Signal Processing. 2011. vol. 59. no. 9. pp. 4331–4340.
3. Shen K., Yu W. Fractional programming for communication systems – Part I: Power control and beamforming. IEEE Transactions on Signal Processing. 2018. vol. 66. no. 10. pp. 2616–2630.
4. Feiten A., Mathar R., Reyer M., Rate and power allocation for multiuser OFDM:An effective heuristic verified by branch-and-bound. IEEE Transactions on Wireless Communications. 2008. vol. 7. no. 1. pp. 60–64.
5. Sun Q., Wu H., Petrosian O. Optimal Power Allocation Based on MetaheuristicAlgorithms in Wireless Network. Mathematics. 2022. vol. 10(18). no. 3336.
6. Sun H., Chen X., Shi Q., Hong M., Fu X., Sidiropoulos N.D. Learning to optimize: Training deep neural networks for interference management. IEEE Transactions on Signal Processing. 2018. vol. 66. no. 20. pp. 5438–5453.
7. Liang F., Shen C., Yu W. and Wu F. Towards optimal power control via ensemblingdeep neural networks. IEEE Transactions on Communications. 2019. vol. 68. no. 3. pp. 1760–1776.

8. Lee W., Kim M., Cho D.H. Deep learning based transmit power control in underlaid device-to-device communication. IEEE Systems Journal. 2018. vol. 13. no. 3. pp. 2551–2554.
9. Liao X., Shi J., Li Z., Zhang L., Xia B. A model-driven deep reinforcement learning heuristic algorithm for resource allocation in ultra-dense cellular networks. IEEE Transactions on Vehicular Technology. 2019. vol. 69. no. 1. pp. 983–997.
10. Shen Y., Shi Y., Zhang J., Letaief K.B. A graph neural network approach for scalable wireless power control. IEEE Globecom Workshops (GC Wkshps). 2019. pp. 1–6.
11. Chowdhury A., Verma G., Rao C., Swami A., Segarra S. Unfolding WMMSE using graph neural networks for efficient power allocation. IEEE Transactions on Wireless Communications. 2021. vol. 20. no. 9. pp. 6004–6017.
12. Li B., Swami A., Segarra S. Power allocation for wireless federated learning using graph neural networks. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). 2022. pp. 5243–5247.
13. Wang X., Ji H., Shi C., Wang B., Ye Y., Cui P., Yu P.S. Heterogeneous graph attention network. The world wide web conference. 2019. pp. 2022–2032.
14. Busbridge D., Sherburn D., Cavallo P., Hammerla N.Y. Relational graph attention networks. arXiv preprint. 2019. arXiv:1904.05811.
15. Gong L., Cheng Q. Exploiting edge features for graph neural networks. Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2019. pp. 9211–9219.
16. Jiang X., Ji P., Li S. CensNet: Convolution with Edge-Node Switching in Graph Neural Networks. IJCAI. 2019. pp. 2656–2662.
17. Fu X., Zhang J., Meng Z., King I. Magnn: Metapath aggregated graph neural network for heterogeneous graph embedding. Proceedings of The Web Conference. 2020. pp. 2331–2341.
18. Zhang X., Zhao H., Xiong J., Liu X., Zhou L., Wei J. Scalable power control/beamforming in heterogeneous wireless networks with graph neural networks. IEEE Global Communications Conference (GLOBECOM). 2021. pp. 01–06.
19. Guo J., Yang C. Learning power allocation for multi-cell-multi-user systems with heterogeneous graph neural networks. IEEE Transactions on Wireless Communications 2021. vol. 21. no. 2. pp. 884–897.
20. Khodmi A., Rejeb S.B., Agoulmine N., Choukair Z. A joint power allocation and user association based on non-cooperative game theory in an heterogeneous ultra-dense network. IEEE Access. 2019. vol. 7. pp. 111790–111800.
21. Challita U., Saad W., Bettstetter C. Interference management for cellular-connected UAVs: A deep reinforcement learning approach. IEEE Transactions on Wireless Communications. 2019. vol. 18. no. 4. pp. 2125–2140.
22. Nguyen L.D., Tuan H.D., Duong T.Q., Poor H.V. Multi-user regularized zero-forcing beamforming. IEEE Transactions on Signal Processing. 2019. vol. 67. no. 11. pp. 2839–2853.
23. Li Y., Han S., Yang C. Multicell power control under rate constraints with deep learning. IEEE Transactions on Wireless Communications. 2021. vol. 20. no. 12. pp. 7813–7825.
24. Wang X., Ji H., Shi C., Wang B., Ye Y., Cui P., Yu P.S. Heterogeneous graph attention network. The world wide web conference. 2019. pp. 2022–2032.
25. Yang Y., Li D. Nenn: Incorporate node and edge features in graph neural networks. Asian conference on machine learning. 2020. pp. 593–608.
26. Chen J., Chen H. Edge-featured graph attention network. arXiv preprint. 2021. arXiv:2101.07671.

27. Kim J., Park J., Noh J., Cho S. Autonomous power allocation based on distributed deep learning for device-to-device communication underlaying cellular network. IEEE access. 2020. vol. 8. pp. 107853–107864.

Цюши Сунь — аспирант, факультет прикладной математики и процессов управления, Санкт-Петербургский государственный университет. Область научных интересов: математическая статистика, системное управление, интеллектуальное производство. Число научных публикаций — 3. st059656@student.spbu.ru; Университетский проспект, 35, 198504, Санкт-Петербург, Петергоф, Россия; р.т.: +7(911)740-8019.

Ян Хе — аспирант, факультет прикладной математики и процессов управления, Санкт-Петербургский государственный университет. Область научных интересов: математическая статистика, системное управление, интеллектуальное производство. Число научных публикаций — 1. hy1186867324@outlook.com; Университетский проспект, 35, 198504, Санкт-Петербург, Петергоф, Россия; р.т.: +7(812)428-4530.

Петросян Ованес Леонович — канд. техн. наук, профессор кафедры, кафедра математического моделирования энергетических систем, Санкт-Петербургский государственный университет. Область научных интересов: электронное машинное обучение, обучение с подкреплением, анализ временных рядов, теория графов, дискретные вероятностные модели и стохастические процессы, а также теория игр, исследование операций. Число научных публикаций — 81. o.petrosyan@spbu.ru; Университетский проспект, 35, 198504, Санкт-Петербург, Петергоф, Россия; р.т.: +7(812)428-453.

Поддержка исследований. Эта работа была поддержана Санкт-Петербургским государственным университетом (грант 94-06-2114).