

An Overview on the Application of Graph Neural Networks in Wireless Networks

SHIWEN HE^{1,2,3} (Member, IEEE), SHAWEN XIONG¹, YEYU OU¹, JIAN ZHANG¹ (Member, IEEE),
JIAHENG WANG^{3,4} (Senior Member, IEEE), YONGMING HUANG^{3,4} (Senior Member, IEEE),
AND YAOXUE ZHANG⁵ (Senior Member, IEEE)

¹School of Computer Science and Engineering, Central South University, Changsha 410083, China

²National Mobile Communications Research Laboratory, Southeast University, Nanjing 210096, China

³Pervasive Communication Research Center, Purple Mountain Laboratories, Nanjing 210096, China

⁴National Mobile Communications Research Laboratory, School of Information Science and Engineering, Southeast University, Nanjing 210096, China

⁵Department of Computer Science and Technology, Tsinghua University, Beijing 100084, China

CORRESPONDING AUTHORS: J. ZHANG and Y. HUANG (e-mail: jianzhang@csu.edu.cn; huangym@seu.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 62171474 and Grant 61720106003; in part by the Open Research Fund of National Mobile Communications Research Laboratory Southeast University under Grant 2022D03; and in part by the Natural Science Foundation of Hunan Province under Grant 2020JJ4745.

ABSTRACT In recent years, with the rapid enhancement of computing power, deep learning methods have been widely applied in wireless networks and achieved impressive performance. To effectively exploit the information of graph-structured data as well as contextual information, graph neural networks (GNNs) have been introduced to address a series of optimization problems of wireless networks. In this overview, we first illustrate the construction method of wireless communication graph for various wireless networks and simply introduce the progress of several classical paradigms of GNNs. Then, several applications of GNNs in wireless networks such as resource allocation and several emerging fields, are discussed in detail. Finally, some research trends about the applications of GNNs in wireless communication systems are discussed.

INDEX TERMS Wireless networks, graph neural networks, resource management.

I. INTRODUCTION

THE ADVENT of fifth-generation (5G) wireless communication systems has driven the revolutionary applications extending far beyond smartphones and other mobile devices [1]. Meanwhile, intelligent communication becomes a novel developing trend of future communication systems [2]. Recently, more and more researchers adopt the deep learning (DL) method to study the problems in wireless networks motivating by the successful application of DL in the related fields of computer [3].

A. FROM TRADITIONAL DEEP LEARNING TO GRAPH NEURAL NETWORK

According to the usage of domain knowledge, DL methods can be divided into the data-driven DL method and

the data- and model-driven DL method. The data-driven DL methods without using the domain knowledge generally have poor interpretability and robustness, while having fast inferencing speed compared with the model-based method. A classic work using multi-layer perceptrons (MLPs) is to solve the power control problem of wireless networks by using the MLPs to approximate the weighted minimum mean square error (WMMSE) algorithm [4]. The data- and model-driven DL methods are rising in recent years and its core idea is to retain the main theoretical characteristics of the classical model algorithms, while using the DL methods to partially enhance or replace its related difficult or time-consuming process. Compared with the data-driven DL methods, the data- and model-driven DL methods have better interpretability and robustness with slower inferencing speed.

For example, Bora *et al.* used the generated model from neural networks instead of the standard sparsity model to represent data distributions [5]. Shlezinger *et al.* designed a deep neural network (DNN) to implement the channel-dependent part of Viterbi algorithm, while keeping the unchange of the rest remains [6]. Kim *et al.* studied a family of sequential codes parameterized by recurrent neural network (RNN) architectures [7]. Gregor and LeCun designed a non-linear and parameterized feed-forward architecture with a fixed depth to approximate the optimal sparse code [8]. Some overviews are presented to summary the application of machine learning (ML) or DL with aiming to improve the quality-of-experience (QoE) of wireless networks [9], [10]. The data collected in these tasks is typically represented in the Euclidean domains. Although the existing works using the DL models defined in the Euclidean domains achieve a better performance in small-scale networks, they fail to exploit the underlying topology of wireless networks. Consequently, the performance decreases sharply when the network scale becomes large [11], [12].

In wireless networks, an obvious feature is that the high dynamics of network topology caused by some uncertain factors, e.g., the user mobility, changes in traffic pattern or adjustment of the network resource, etc. In addition, the wireless data may be collected from non-Euclidean domains and represented as graph-structured data with high dimensional features and interdependency among communication devices. These issues bring difficulties to apply directly the learning model defined in Euclidean domains in wireless networks. A straightforward way solving these difficulties is to incorporate the network topology information, which is described as an adjacency matrix depending on the specific node index, into the architecture of neural networks. However, the indices of communication devices in wireless networks may change due to the reallocation of resources and the movement of communication devices, i.e., dynamic graph-structured data. This motivates us to design a novel learning model with taking into account the interdependencies between communication devices and the dynamics of wireless networks [13]. The emerging GNNs enable the graph-structured data to be processed effectively and to use the global parameterization, common system of coordinates, vector space structure, or shift-invariance [14]. In the last few years, many researchers have begun to use GNNs to mine the deep information hidden in the graph-structured data to further improve the abilities of learning and simulating the interaction between nodes.

B. RELATED OVERVIEWS AND CONTRIBUTIONS

Some overviews about the paradigms and applications of GNNs are represented in the past few years. The authors of [15] introduced comprehensively four basic paradigms of GNNs and described the representative models in detail. Differentiating from [15], the authors of [16] further reviewed graph reinforcement learning and graph adversarial methods. The authors of [17] summarized the state-of-the-art of the main models and algorithms of graph learning from

four aspects, i.e., graph signal processing, matrix decomposition, random walk, and deep learning. The aforementioned surveys introduced in detail the characteristics of different GNN paradigms but briefly introduced the application of GNNs in some fields except for wireless networks. The authors of [18] summarized the GNN approaches in terms of the spatial domain and spectral domain, respectively. Furthermore, this work integrates the spatial and spectral domain models into a unified framework. The authors of [19] further discussed the expressive power of GNNs and further summarized the relationships among GNNs, the Weisfeiler-Lehman algorithm, and distributed local algorithms. The authors of [20] illustrated the excellent performance of GNNs depending on the three characteristics of equivariance, stability and transferability, which are further determined by the selection of the optimization objective and technologies, as well as the design of graph filters. The methods using software or hardware to accelerate GNNs were reviewed from the perspective of computation speed in [21]. In order to fill the research on the interpretability of GNNs, the unified methodology and standard testbed for evaluating the interpretability of GNNs were summarized comprehensively in [22]. There are also several overviews on applying the GNNs to solve the problems in the traffic domain [23], power systems [24], and recommender system [25]. In addition, the relationship between GNNs and the latest neural-symbolic computing that aims at integrating the abilities of learning from the environment and of reasoning from what has been learned was introduced in [26].

In this paper, we aim to present a comprehensive overview of the application of GNNs in wireless networks. Meanwhile, we also provide some potential research directions for researchers who are interested in this topic. A detailed organization of this overview is illustrated in Fig. 1. In particular, the main contributions of this overview are summarized as follow.

- The methods of constructing wireless communication graph (WCG) for Mesh/Ad-hoc networks, Cellular networks, or Wireless Local Area Networks (WLANs) are illustrated elaborately.
- Several classical paradigms of GNNs applied in wireless networks are introduced to acquire a better understanding of the concepts and structures of GNNs.
- A comprehensive review of GNNs applied in wireless networks is summarized in terms of the existing directions, e.g., resource allocation and several emerging fields.
- Some challenges and potential research directions are summarized and discussed for the application of GNNs in wireless networks.

The rest of this paper is organized as follows: In Section II, we introduce several basic definitions of graph-structured data and summarize the construction methods of WCG for Mesh/Ad-hoc networks and Cellular networks/WALNs. In Section III, we review several classical paradigms of GNNs that applied in wireless networks. In Section IV, we introduce the application of GNNs in wireless networks. In Section V, we discuss a few valuable directions for the application of

Section I. Introduction		
A. From Traditional Deep Learning to Graph Neural Network	B. Related Overviews and Contributions	
Section II. Definition and Construction of Graph		
A. Definition of Graph	B. Construction of Wireless Communication Graph 1) Mesh/Ad-hoc Networks 2) Cellular Networks/WLANs	
Section III. Paradigms of GNNs	Section IV. Application in Wireless Networks	
A. Graph Convolutional Neural Networks 1) Spectral-based GCNs 2) Spatial-based GCNs B. Graph Attention Networks C. Spatial-Temporal Graph Neural Networks D. Other Hybrid Methods	A. Resource Allocation 1) Power control 2) Link scheduling 3) Others 4) Brief discussion	B. Emerging Fields 1) Channel estimation 2) Traffic prediction 3) Vehicle communication 4) Others 5) Brief discussion
Section V. Key Issues and Future Development		
A. Acquisition of high-quality data	D. Robustness of the learning model	G. Research of GNN-based deep unfolding method
B. Distributed GNN learning model	E. Scalability of the learning model	H. Model design with complex constraints
C. Data privacy issues	F. Cross-layer optimization	
Section VI. Conclusions		

FIGURE 1. Road map of the overview.

TABLE 1. Commonly used notations.

G	Graph
N	The number of nodes
V	The set of nodes
v	A node $v \in V$
$\mathcal{N}(v)$	The set of neighbor nodes of $v \in V$
E	The set of edges
e_{ij}	An edge from node v_i to node v_j
\mathbf{A}	Adjacent matrix
\mathbf{A}^T	The transpose of matrix \mathbf{A}
\mathbf{A}_{ij}	The element of the i -th row and the j -th column of matrix \mathbf{A}
\mathbf{D}	The degree matrix of matrix \mathbf{A}
$\mathbf{D}^{-\frac{1}{2}}$	$-\frac{1}{2}$ power of the elements in the matrix \mathbf{D}
\mathbf{X}	Feature matrix
\mathbf{X}_i	The i -th column of matrix \mathbf{X}
$\mathbf{x}[l]$	Denotes the l -th element of vector \mathbf{x}
d	The dimension of the feature vector
L	The maximum depth of GNNs

GNNs in wireless networks. Finally, we conclude this work in Section VI. For ease of reading, the notations commonly used in graphs are summarized in Table 1.

II. DEFINITION AND CONSTRUCTION OF GRAPH

In this section, we would like to first describe several basic definitions of graph-structured data. Then, the construction methods of WCG for different wireless network scenarios, e.g., Mesh/Ad-hoc networks, Cellular networks, and WLANs, are illustrated in detail.

A. DEFINITION OF GRAPH

Graph-structured data is a kind of non-Euclidean data and is commonly expressed as $G = (V, E)$, where V and E are the sets of nodes and edges, respectively [27]. Let $v_i \in V$ be a node and $e_{ij} = (v_i, v_j) \in E$ be an edge from node v_i to node v_j . The adjacency matrix of a graph is represented as \mathbf{A} . If $e_{ij} \in E$, $\mathbf{A}_{ij} = 1$, otherwise, $\mathbf{A}_{ij} = 0$. The graph is undirected

if \mathbf{A} is symmetric, otherwise, the graph is directed. The degree matrix \mathbf{D} of an undirected graph is a diagonal matrix, where $\mathbf{D}_{ii} = |\mathcal{N}(v_i)|$. The Laplacian matrix of an undirected graph is defined as $\mathbf{L} = \mathbf{D} - \mathbf{A}$. The normalized Laplacian matrix is defined as $\tilde{\mathbf{L}} = \mathbf{I}_N - \mathbf{D}^{-\frac{1}{2}} \mathbf{A} \mathbf{D}^{-\frac{1}{2}}$. Note that the normalized Laplacian matrix $\tilde{\mathbf{L}}$ is a real semi-positive definite matrix. Accordingly, it can be decomposed into $\tilde{\mathbf{L}} = \mathbf{U} \Lambda \mathbf{U}^T$, where \mathbf{U} is the eigenvector matrix and Λ is a diagonal matrix of $[\Lambda]_{ii} = \lambda_i$ with λ_i being the eigenvalue. While for the directed graph, the in-degree and the out-degree matrices are defined as $\mathbf{D}_{jj}^{(in)} = \sum_i \mathbf{A}_{ij}$ and $\mathbf{D}_{ii}^{(out)} = \sum_j \mathbf{A}_{ij}$, respectively. The transition probability matrix \mathbf{P} of a given directed graph is defined as $\mathbf{P} = (\mathbf{D}^{(out)})^{-1} \mathbf{A}$. Accordingly, a symmetric normalized Laplacian of directed graph is defined as $\vec{\mathbf{L}} = \mathbf{I} - \frac{1}{2}(\Phi^{\frac{1}{2}} \mathbf{P} \Phi^{-\frac{1}{2}} + \Phi^{-\frac{1}{2}} \mathbf{P}^T \Phi^{\frac{1}{2}})$, where Φ is generated according to \mathbf{P} and perron vector [28]. It is worth mentioning that, from the reviewing results, there is almost no work to deal with directed graph based on spectral domain, but based on spatial domain. In a graph, each node may have its own attribute feature. The feature matrix of a graph is defined as $\mathbf{X} \in \mathbb{R}^{N \times d}$. If feature matrix \mathbf{X} changes over time, the graph is defined as a spatial-temporal graph.

B. CONSTRUCTION OF WIRELESS COMMUNICATION GRAPH

The first thing of using GNNs is to transform a wireless network into a graph. In general, according to the specific goal of the research, the topology structure of wireless networks may be constructed into an undirected graph or directed graph. According to the types of communication links and communication devices in wireless networks, it can be further constructed into a homogeneous or heterogeneous graph. In the sequel, the construction methods of WCG for various wireless network

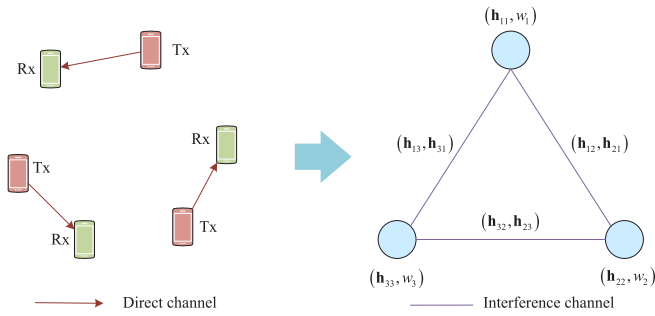


FIGURE 2. Illustration of constructing WCG for homogeneous Mesh/Ad-hoc networks with 3 transceivers [29].

scenarios, e.g., Mesh/Ad-hoc networks, Cellular networks, or WLANs, are illustrated elaborately.

1) MESH/AD-HOC NETWORKS

Suppose there are N transceiver pairs in homogeneous Mesh/Ad-hoc networks. To build a WCG for this kind of wireless network, we view the i -th transceiver pair as the i -th node of WCG, the feature vector of the i -th node includes the direct channel state information (CSI) \mathbf{h}_{ii}^1 and other environmental information, such as the weight w_i of the i -th node. The edge between nodes v_i and v_j in WCG may be undirected or directed. The feature vector of the undirected edge includes the interference CSIs \mathbf{h}_{ij} and \mathbf{h}_{ji} . While the feature vectors of two directed edges between nodes v_i and v_j can include \mathbf{h}_{ij} and \mathbf{h}_{ji} , respectively.² Fig. 2 shows a construction method of WCG for homogeneous Mesh/Ad-hoc networks with 3 transceivers.

In heterogeneous Mesh/Ad-hoc networks, suppose there are N types of communication links. We treat the i -th transceiver pair with link type m as node v_{im} in the WCG. The set of neighboring nodes with link type n of node v_{im} are represented as $\mathcal{N}_{im}^{(n)}$. The feature vector of node v_{im} includes the CSI \mathbf{h}_{imim} of direct link, and other environmental information corresponding to the link of type m . The feature vector of the edge between nodes v_{im} and v_{jn} should also be considered from the perspective of undirected and directed. The feature vector of the undirected edge includes the CSIs, i.e., \mathbf{h}_{imjn} and \mathbf{h}_{jnim} of interference links. While the feature vectors of the directed edges between nodes v_{im} and v_{jn} include the \mathbf{h}_{imjn} and \mathbf{h}_{jnim} , respectively. Fig. 3 illustrates a construction method of WCG for heterogeneous Mesh/Ad-hoc networks with 2 link types, where v_{jn} is the feature vector of node v_{jn} , \mathbf{e}_{jnim} is the feature vector of the edge between nodes v_{jn} and v_{im} , and “Link i_m ” indicates the i -th communication link with type m .

2) CELLULAR NETWORKS/WLANs

In general, Cellular networks/WLANs may consist of M ($M \geq 1$) access points (APs) and N ($N \geq 1$) user

1. It is worth noting that the dimension of \mathbf{h}_{ij} between nodes v_i and v_j is determined by the number of antennas equipped by the transceiver pair.

2. The construction method of the feature vectors of nodes and edges is not limited in the method aforementioned and mentioned later, which can be adjusted and supplied according to the specific research tasks.

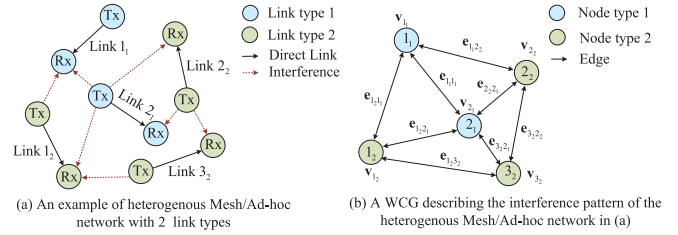


FIGURE 3. Illustration of constructing WCG for heterogeneous Mesh/Ad-hoc networks with 2 link types [30].

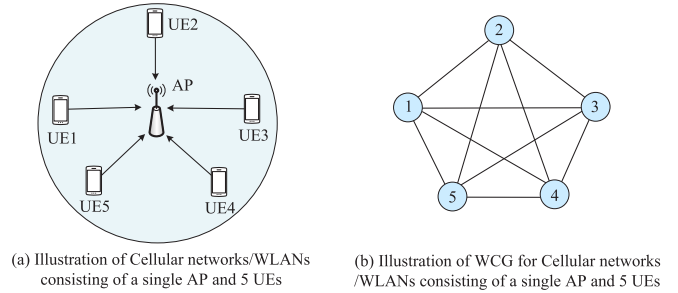


FIGURE 4. Illustration of constructing WCG for Cellular networks/WLANs with a single AP.

equipments (UEs). Considering a simple situation including only one AP, which allocates the resources, such as power control and user association, etc., to the UEs. We treat the i -th UE as a node of WCG should be built while ignoring AP. The feature vector of the i -th node includes the CSI \mathbf{h}_{ii} and other environmental information. The feature vector of the edge between nodes v_i and v_j includes the CSI \mathbf{h}_{ij} and \mathbf{h}_{ji} , etc., which can also be ignored due to all UEs share one AP. An illustration of constructing WCG for Cellular networks/WLANs with a single AP is shown in Fig. 4.

We further consider a more complex situation including multiple APs in Cellular Networks/WLANs, in which one AP may serve multiple UEs, and one UE may also access multiple APs. We first consider the scenario that one UE just accesses one AP and one AP serves multiple UEs. As shown in Fig. 5 (a), UEs 1-4 and UEs 5-7 communicate with AP1 and AP2, while AP2 and AP1 are interfered with UE4 and UE5, respectively. The WCG of this kind of wireless networks can be built in two manners, which are illustrated in Fig. 5 (b) and Fig. 5 (c), respectively. In Fig. 5 (b), there are two types of nodes indicating APs and UEs, respectively. The information of position, channel configuration, and device type are considered to be the feature vector of nodes. The feature vector of the edge between nodes v_i and v_j includes the direct/interference CSI and other link information. In contrast to Fig. 5 (b), the APs are ignored in Fig. 5 (c), which includes only one type of nodes indicating UE. The feature vector of a node includes the position of UE, channel configuration and device type, etc. The direct/interference CSIs and other link information can be considered to be the feature vector of edge. In practice, we can choose the appropriate manner according to the specific research tasks.

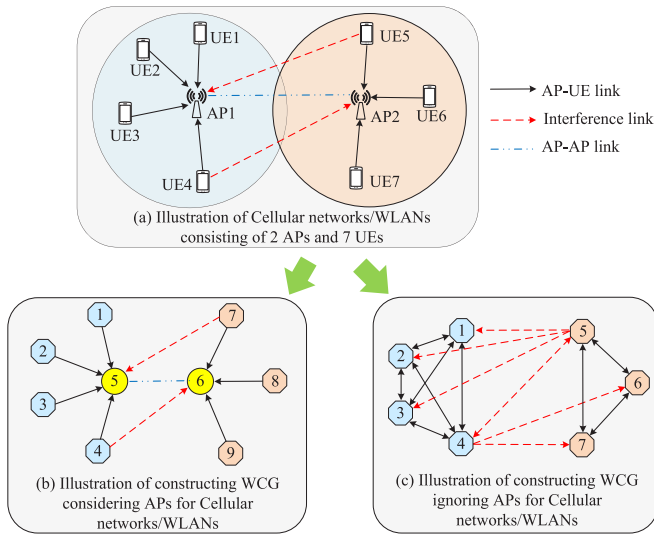


FIGURE 5. Illustration of constructing WCG for Cellular networks/WLANs with multiple APs.

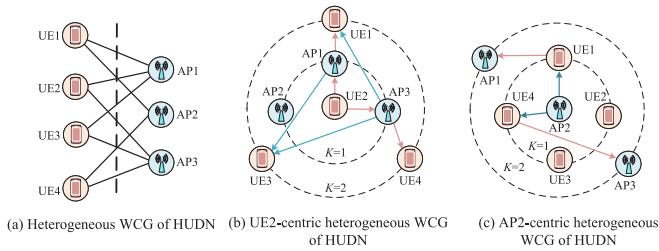


FIGURE 6. Illustration of constructing WCG for ultra-dense Cellular networks/WLANs [31].

In addition to the Cellular networks/WLANs scenarios aforementioned, there are also other scenarios, such as one AP may serve multiple UEs and one UE may access multiple APs, called Heterogeneous Ultra-Dense Network (HUDN) [31]. We treat the UEs and APs as two types of nodes, as shown in Fig. 6 (a). The feature vector of a node indicating UEs includes the CSIs to every AP. However, the feature vector of a node indicating APs includes the CSIs to every UE. The edge of WCG should be built only exists between nodes indicating UEs and nodes indicating APs when the UE can be detected. As shown in Fig. 6 (b), the first-order neighborhood, $K = 1$, of UEs is the APs that have edge connect, the second-order neighborhood, $K = 2$, of UEs is the UEs that connect to the first-order neighborhood. Similarly, the corresponding neighborhood of AP is shown in Fig. 6 (c). The details of designing the GNN model using this WCG can be found in [31], which exploit feature information from first-order and second-order neighborhood.

III. PARADIGMS OF GNNS

In this section, we simply introduce a few classical GNN models, mainly including graph convolutional neural networks, graph attention networks, spatial-temporal graph

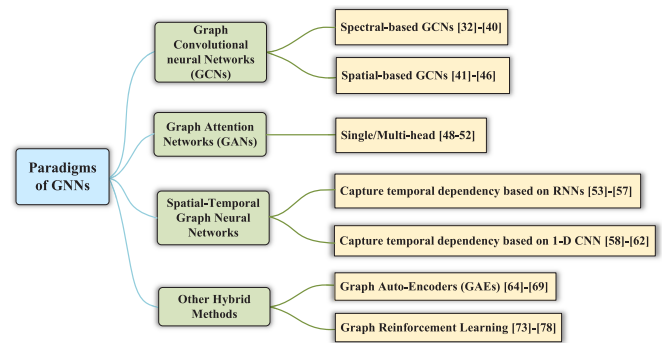


FIGURE 7. Overview of the paradigms of GNNS.

neural networks, and other hybrid methods. The overview of these paradigms is shown in Fig. 7.

A. GRAPH CONVOLUTIONAL NEURAL NETWORKS

Graph convolutional neural networks (GCNs) implement the convolutional operation on graph-structured data, i.e., in non-Euclidean space [14]. The core idea of GCNs is to learn a mapping function, which can combine the neighbor nodes' information with its feature information to generate a new node representation. According to different convolution methods, GCNs can be divided into spectral-based GCNs [32]–[40] and spatial-based GCNs [41]–[46]. In the sequel, we simply introduce several classical models of spectral-based GCNs and spatial-based GCNs, respectively.

1) SPECTRAL-BASED GCNS

Since the number of neighbors may be different for different nodes, a fixed convolutional kernel cannot be used on a graph. To address this problem, the graph-structured data is generally converted to the frequency domain. Specifically, for a given input graph signal $\mathbf{x} \in \mathbb{R}^d$ and a graph filter $\mathbf{g} \in \mathbb{R}^d$, the graph convolution is defined as [15]

$$\begin{aligned} \mathbf{x} *_G \mathbf{g} &= \mathcal{F}^{-1}(\mathcal{F}(\mathbf{x}) \odot \mathcal{F}(\mathbf{g})) \\ &= \mathbf{U}(\mathbf{U}^T \mathbf{x} \odot \mathbf{U}^T \mathbf{g}) = \mathbf{U} \hat{\mathbf{g}} \mathbf{U}^T \mathbf{x}, \end{aligned} \quad (1)$$

where $*_G$ denotes the graph convolution operation, $\mathcal{F}(\mathbf{x}) = \mathbf{U}^T \mathbf{x}$ denotes the graph Fourier transform, $\mathcal{F}^{-1}(\mathcal{F}(\mathbf{x})) = \mathbf{U} \mathcal{F}(\mathbf{x})$ denotes the inverse graph Fourier transform, \odot denotes the Hadamard product, and $\hat{\mathbf{g}} = \text{diag}(\mathbf{U}^T \mathbf{g})$.

Various spectral-based GCNs have been defined by changing $\hat{\mathbf{g}}$. For example, Bruna *et al.* proposed spectral CNN (SpectralCNN) in which $\hat{\mathbf{g}}$ is learnable parameters [32]. However, due to the existing of the eigen-decomposition of $\tilde{\mathbf{L}}$, SpectralCNN faces several challenges, such as low computational efficiency of eigen-decomposition and the learned graph filters cannot be applied in a graph with different structure [27]. To overcome these shortcomings, Defferrard *et al.* proposed Chebnet via redefining the graph filter with Chebyshev polynomials [33]. By constraining the number of parameters, Kipf *et al.* further proposed a model named GCN, which has the ability to overcome the

overfitting, to minimize the number of operations at each layer, i.e.,

$$\mathbf{X}_{*G} = \mathbf{W} \left(\mathbf{I}_N + \mathbf{D}^{-\frac{1}{2}} \mathbf{A} \mathbf{D}^{-\frac{1}{2}} \right) \mathbf{X}, \quad (2)$$

where \mathbf{W} is a learnable weight matrix. In order to tackle the case in which $\mathbf{I}_N + \mathbf{D}^{-\frac{1}{2}} \mathbf{A} \mathbf{D}^{-\frac{1}{2}} \in (0, 2)$ may lead to gradient explosion, the authors further transform $\mathbf{I}_N + \mathbf{D}^{-\frac{1}{2}} \mathbf{A} \mathbf{D}^{-\frac{1}{2}}$ into $\tilde{\mathbf{D}}^{-\frac{1}{2}} \tilde{\mathbf{A}} \tilde{\mathbf{D}}^{-\frac{1}{2}}$, where $\tilde{\mathbf{A}} = \mathbf{A} + \mathbf{I}_N$ and $\tilde{\mathbf{D}}_{ii} = \sum_j \tilde{\mathbf{A}}_{ij}$. Compared to GCN, Chebnet has higher computational complexity, but it has stronger expression ability. Chebnet's K -order convolution operator can cover K steps neighbor nodes of the central node, while GCN only covers the first-order neighbor nodes. However, the perception domain of graph convolution can be expanded by stacking multiple GCN layers, so the flexibility is relative high.

Ma *et al.* proposed a directed graph convolution network based on directed Laplacian, which is defined as [28]

$$\mathbf{Z} = \frac{1}{2} \left(\tilde{\Phi}^{\frac{1}{2}} \tilde{\mathbf{P}} \tilde{\Phi}^{-\frac{1}{2}} + \tilde{\Phi}^{-\frac{1}{2}} \tilde{\mathbf{P}}^T \tilde{\Phi}^{\frac{1}{2}} \right) \mathbf{X} \mathbf{W}, \quad (3)$$

where $\tilde{\mathbf{D}}_{ii}^{(out)} = \sum_j \tilde{\mathbf{A}}_{i,j}$, $\tilde{\mathbf{P}} = (\tilde{\mathbf{D}}^{(out)})^{-1} \tilde{\mathbf{A}}$ and $\tilde{\Phi}$ is calculated based on $\tilde{\mathbf{P}}$. The directed graph filter is approximated by the first-order Chebyshev polynomials.

Remark 1: The application of spectral-based GCNs in wireless networks will be introduced in Section IV. The definition of graph filters for spectral-based GCNs usually combines the adjacency matrix \mathbf{A} of wireless network topology and the channel state information \mathbf{H} , which can make full use of the complex wireless information. On the other hand, \mathbf{H} is expressed as the propagation relationship between nodes from the aspect of wireless network environment.

2) SPATIAL-BASED GCNS

The spatial-based graph convolution is similar to the image convolution. The two convolution operations all extract the neighbor information of a node to obtain a richer feature representation of the node or the pixel. The difference between image convolution and spatial-based graph convolution is that the nodes in a graph are unordered while the pixels in an image are irregular, and the number of neighbors of each pixel in an image is limited while the number of neighbors of each node in a graph is not sure. So, spatial-based graph convolution operation cannot use a fixed-size convolution kernel like the image convolution operation. Thus, the key of spatial-based GCNs is to define the convolution operation with different neighborhood numbers and keep local invariance.

The most widely used spatial-based GCNs in wireless networks are message passing neural network (MPNN) and diffusion-convolutional neural networks (DCNNs). MPNN was proposed in [42], which is a unified framework of spatial-based GCNs, and decomposes the spatial-based graph convolution into a message aggregation phase and a

combination phase, i.e.,

$$\mathbf{m}_{v_i}^{(t)} = \sum_{v_j \in \mathcal{N}(v_i)} \mathcal{M}^{(t)} \left(\mathbf{x}_i^{(t-1)}, \mathbf{x}_j^{(t-1)}, \mathbf{e}_{ij} \right), \quad (4a)$$

$$\mathbf{x}_i^{(t)} = \mathcal{U}^{(t)} \left(\mathbf{x}_i^{(t-1)}, \mathbf{m}_{v_i}^{(t)} \right), \quad (4b)$$

where \mathbf{e}_{ij} is the feature vector of the edge between nodes v_i and v_j , $\mathcal{M}^{(t)}(\cdot)$ and $\mathcal{U}^{(t)}(\cdot)$ are the aggregation function and the combination function in the t -th iteration, respectively. $\mathbf{m}_{v_i}^{(t)}$ is the message aggregated from node v_i 's neighbors and $\mathbf{x}_i^{(t)}$ is the hidden state of node v_i in the t -th iteration. It is observed that the computational efficiency of MPNN decreases with the increase of the number of nodes.

Hamilton *et al.* further proposed Graph SAmple and aggreGAtE (GraphSAGE) model via fixing the number of neighbors for message passing to overcome the shortcomings of MPNN [43]. The graph convolution operation of GraphSAGE is implemented by

$$\mathbf{x}_{v_i}^{(t)} = \sigma \left(\mathbf{W}^{(t)} g^{(t)} \left(\mathbf{x}_i^{(t-1)}, \left\{ \mathbf{x}_j^{(t-1)}, \forall j \in S_{\mathcal{N}(v_i)} \right\} \right) \right), \quad (5)$$

where $\mathbf{W}^{(t)}$ and $g^{(t)}(\cdot)$ are a learnable weight matrix and an aggregation function in the t -th layer, respectively. $\sigma(\cdot)$ is a nonlinear activation function. $S_{\mathcal{N}(v_i)}$ is a random sample of the node v_i 's neighbors. The main difference between MPNN and GraphSAGE is that GraphSAGE randomly samples a fixed number of neighbors for each node, while MPNN utilizes all the neighbors of each node. In addition, the diffusion-convolution operation in the DCNNs model builds a potential representation by scanning the diffusion process of each node through the transition probability matrix, i.e., $\mathbf{Z}^{(t)} = \sigma(\mathbf{W}^{(t)} \odot \mathbf{P}^t \mathbf{X})$, where $\mathbf{Z}^{(t)}$ is the hidden state in the t -th layer and \mathbf{P}^t denotes the t power of \mathbf{P} .

Remark 2: The GCN models introduced above meet the graph structures with the same node type and edge type. However, in wireless networks, the types of communication devices and of communication mechanisms between devices may be diverse, that is, the corresponding WCG may be a heterogeneous. Therefore, the design of heterogeneous graph convolution is helpful to learn different types of information and is more suitable for the business needs of practical network scenarios. In Section III, several works on designing graph convolution for heterogeneous WCG are introduced. However, the existing methods are designed based on spatial-based GCNs. There are few works on designing heterogeneous graph convolution based on spectral-based GCNs.

B. GRAPH ATTENTION NETWORKS

One notes that in a graph, different neighbor nodes generally have different influences on the central node. This implies that one needs to distinguish the influences of nodes with a proper means during the design process of the learning model. Attention mechanism [47] has been regarded as an expressive means of information fusion by

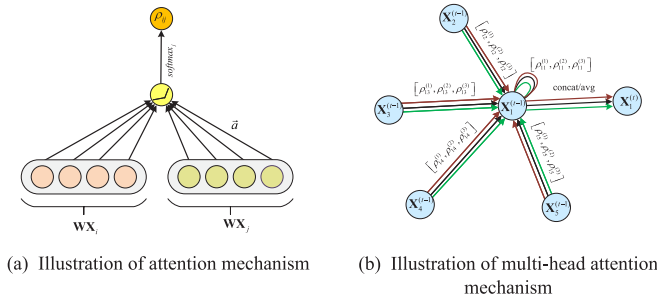


FIGURE 8. Illustration of the implementation details of attention mechanism [48].

assigning weight to given information. In recent years, attention mechanism is also introduced into GNNs, such as, Velickovic *et al.* proposed an attention mechanism based GNN model, i.e., Graph Attention Network (GAT), by adaptively allocating weight to different neighbors in the aggregation operation [48], which is defined as follows

$$\mathbf{X}_i^{(t)} = \sigma \left(\sum_{j \in \mathcal{N}(v_i)} \alpha(\mathbf{X}_i^{(t-1)}, \mathbf{X}_j^{(t-1)}) \mathbf{W}^{(t-1)} \mathbf{X}_j^{(t-1)} \right), \quad (6)$$

where $\alpha(\cdot)$ is the attention mechanism, $\mathbf{X}_i^{(t)}$ is the hidden state of node v_i at the t -th layer. In addition, multi-head attention mechanism is further introduced to improve the expression ability of the attention layer, that is, K independent attention mechanisms can be utilized and then the output are concatenated together, i.e.,

$$\mathbf{X}_i^{(t)} = \parallel_{k=1}^K \sigma \left(\sum_{j \in \mathcal{N}(v_i)} \alpha^k(\mathbf{X}_i^{(t-1)}, \mathbf{X}_j^{(t-1)}) \mathbf{W}^{(t-1)} \mathbf{X}_j^{(t-1)} \right), \quad (7)$$

where \parallel denotes the concatenation operation and $\alpha^k(\cdot)$ is the k -th attention mechanism. The illustration of single-head attention mechanism and multi-head attention mechanism are shown in Fig. 8, where $\rho_{ij}^{(k)}$ denotes the attention weight between node v_i and node v_j obtained by the k -th attention mechanism. Other GNNs models using attention mechanism could be found in [49]–[52].

Remark 3: The essence of the attention mechanism is to weigh the information transmitted to distinguish the importance of different types of information. In wireless networks, especially heterogeneous wireless networks, the environmental information may be diverse, such as communication equipment, communication links, etc., which may affect the problems in different ways. Therefore, it is unreasonable to treat different types of information equally in the design of the learning model. On the contrary, one or more attention mechanisms should be designed according to the relevant prior knowledge to distinguish the influences of different types of information. There are few applications of graph attention mechanism in wireless networks.

C. SPATIAL-TEMPORAL GRAPH NEURAL NETWORKS

Spatial-temporal GNNs (STGNNs) play an important role in dealing with graphs that have dynamic node inputs

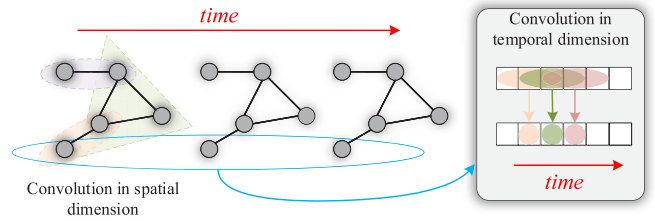


FIGURE 9. Illustration of how ASTGCN captures spatial and temporal dependencies.

while connected nodes are interdependent. There are two categories of STGNNs from the perspective of capturing temporal dependency, i.e., RNN-based methods [53]–[57] and CNN-based methods [58]–[62].

Chen *et al.* utilized the residual recurrent GNN (Res-RGNN) to predict the traffic flow in traffic network [53]. Res-RGNN utilizes the spatial attributes to capture the spatial features with diffusion convolution, while using graph recurrent unit (GRU) to discover the temporal dependency for each node. Specifically, the implementation of RGNN unit at time t is

$$\mathbf{r}^{(t)} = \sigma \left(\Theta_r *_{G} \left[\mathbf{x}^{(t)}, \mathbf{e}^{(t)}, \mathbf{s}^{(t-1)} \right] + \mathbf{b}_r \right), \quad (8a)$$

$$\mathbf{u}^{(t)} = \sigma \left(\Theta_u *_{G} \left[\mathbf{x}^{(t)}, \mathbf{e}^{(t)}, \mathbf{s}^{(t-1)} \right] + \mathbf{b}_u \right), \quad (8b)$$

$$\mathbf{c}^{(t)} = \tanh \left(\Theta_c *_{G} \left[\mathbf{x}^{(t)}, \mathbf{e}^{(t)}, (\mathbf{r}^{(t)} \odot \mathbf{s}^{(t-1)}) \right] + \mathbf{b}_c \right), \quad (8c)$$

$$\mathbf{s}^{(t)} = \mathbf{u}^{(t)} \odot \mathbf{s}^{(t-1)} + (1 - \mathbf{u}^{(t)}) \odot \mathbf{c}^{(t)}, \quad (8d)$$

$$\mathbf{y}^{(t+1)} = \mathbf{W}_o \mathbf{s}^{(t)}, \quad (8e)$$

where $\mathbf{x}^{(t)}$, $\mathbf{e}^{(t)}$ and $\mathbf{s}^{(t)}$ denote the graph signal, external feature and the outputted hidden state at time t , respectively. $\mathbf{r}^{(t)}$ and $\mathbf{u}^{(t)}$ represent the reset gate and update gate at time t , respectively. Θ_r , Θ_u and Θ_c are the learnable graph filters, \mathbf{W}_o is the learned weights of the output layer. $\mathbf{y}^{(t+1)}$ denotes the output at time $t + 1$.

Guo *et al.* proposed an attention mechanism based spatial-temporal GNN (ASTGCN) to predict traffic flow in traffic network [59]. Specifically, ASTGCN utilizes the spectral-based GCNs, i.e., ChebNet [33], to capture the spatial dependency among different nodes in traffic network graph. Meanwhile, one dimension CNN is utilized to capture temporal dependency for each node in time series. The implementation details of capturing spatial and temporal dependencies are illustrated in Fig. 9.

Remark 4: With the development of communication technologies, the wireless network becomes more complex and huge with massive terminals. As a result, the resource management of wireless networks becomes more and more challenging. To improve spectral efficiency, prediction plays an important role in wireless networks. Applying STGNNs to traffic prediction has attracted extensive attention in both academic and industry, which contributes to the resource management of wireless networks. Of course, other directions not involved in this overview also need to consider the

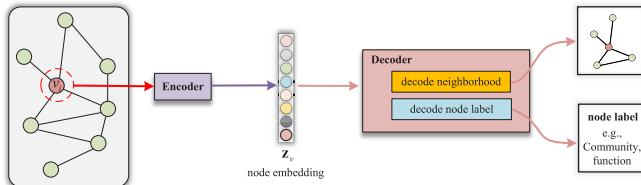


FIGURE 10. Overview of GAEs.

spatial-temporal dependencies during designing the GNNs for wireless networks.

D. OTHER HYBRID METHODS

Inspired by the conventional auto-encoders, graph auto-encoders (GAEs) utilizing GNNs as encoders to learn low-dimensional latent representations (or embeddings) of nodes have been investigated for wireless networks. The goal of encoders in GAEs is to encode the structural information of nodes. While the decoder in GAEs aims at decoding the structural information about the graph from the learned latent representations [63]. The general overview of GAEs is shown in Fig. 10. Specifically, the encoder maps node v to a low-dimensional embedding vector \mathbf{z}_v based on the node's structural information, and the decoder extracts the information interested from the low-dimensional embedding vector. GAEs have been used in many fields by virtue of their concise encoder-decoder structure and efficient encoding ability [64]–[69]. Kipf and Welling proposed the variational GAE using a GCN encoder and a simple inner product decoder, which aims at the link prediction in citation networks [64]. The encoder maps each node to a low-dimensional latent representation using GCN, then a network embedding \mathbf{Z} can be obtained. The decoder computes the pair-wise distance given network embedding and applies a non-linear activation. Finally, the decoder outputs the reconstructed adjacency matrix.

In recent years, reinforcement learning (RL) has been gradually applied to graph-structured tasks, such as graph generation [70], [71], graph classification [72], and graph reasoning tasks [73]–[78], etc. You *et al.* proposed a graph convolutional policy network (GCPN) based on RL and GCNs to address the problem of non-differentiable objective functions and constraints [71]. Graph attention model (GAM) was proposed based on RL and random walks to solve the graph classification task [72]. The generation of random walks was modeled as a partially observable Markov decision process. The RL agent performs two actions at each time step, i.e., predicts the label of input graph and generates the rank vector using designed rank network. The reward is designed as $\mathcal{J}(\theta) = \mathbb{E}_{P(S_{1:L};\theta)} \sum_{l=1}^L r_l$, where $r_l = 1$ if the GAM classified the graph correctly, otherwise, $r_l = -1$. S_l is the environment. Xiong *et al.* proposed a DeepPath model to find the most informative path between two target nodes with the goal of solving the knowledge graph reasoning task [73]. The action of RL agents is to predict the next

node in the path at each step and output a reasoning path in the knowledge graph. The reward functions include the scoring criteria: global accuracy, path efficiency and path diversity.

Remark 5: The main advantage of GAEs is to mine the topological information in the graph, and then learn an effective low dimensional feature vector representation for each node or the whole graph. This feature vector representation can reflect the characteristic that can separate from other nodes or graphs to a certain extent. Although RL has been widely used in wireless networks, the application of RL in GNNs is still in its infancy. Generally speaking, the introduction of RL can enable GNNs to achieve approximate optimal performance without the prior information of the environment, and have independent exploration and optimal decision-making capabilities. Therefore, the introduction of RL into wireless communication technology has important practical significance.

IV. APPLICATIONS IN WIRELESS NETWORKS

In this section, we focus on introducing comprehensively the application of GNNs in wireless networks. As shown in Table 2, the applications of GNNs in wireless networks mainly cover resource allocation and a few emerging fields. The commonly used algorithms are illustrated in Fig. 11 for each research direction in wireless networks.

A. RESOURCE ALLOCATION

Resource allocation is one of the key issues for wireless communication systems. Applying GNNs to study the problem of resource allocation mainly focuses on power control, link scheduling, channel allocation, and spectrum allocation, etc.

1) POWER CONTROL

A large amount of works has studied the power control problem using traditional optimization methods and DNNs. Unfortunately, the traditional optimization methods face high computational complexity. On the other hand, with the expansion of the wireless network scale, the scalability and generalization of DNNs will become worse. Motivated by these observations, many researchers utilize the GNNs, which have the natural characteristics of solving the problem with graph-structured data, to investigate the power control problem in wireless networks.

To develop scalable methods to solve the power control problem in wireless networks, Shen *et al.* proposed an interference graph convolutional neural network (IGCNet) based on MPNN for K -user interference channels [29]. In particular, the K -user interference channels are modeled as a complete WCG with node and edge labels, as shown in Fig. 2. The aggregation and combination rules of IGCNet are designed as follows

$$\gamma_{j,i}^{(t)} = \text{MLP1}\left(h_{ji}, h_{ij}, w_i, h_{jj}, \beta_j^{(t-1)}\right), \quad (9a)$$

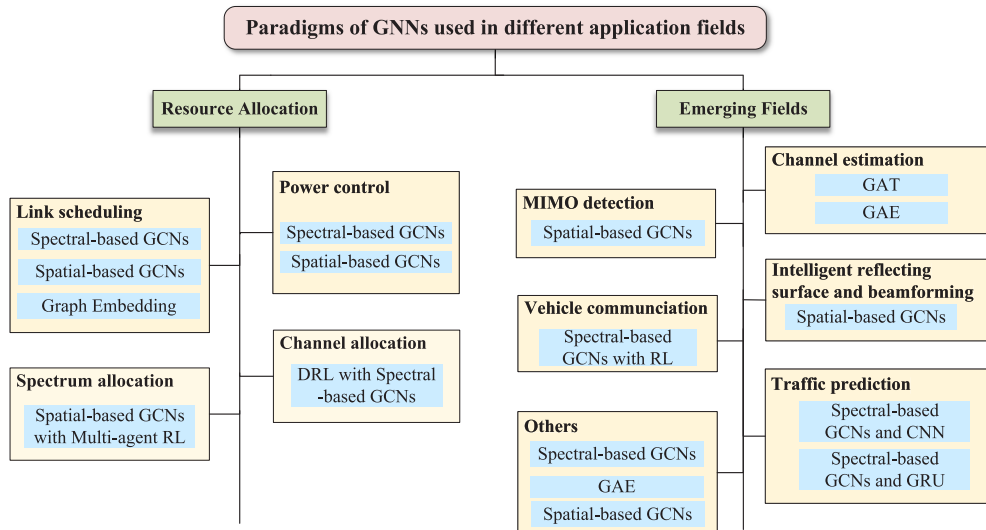


FIGURE 11. Paradigms of GNNs used in different application fields.

$$\alpha_i^{(t)} = \text{CONCAT} \left(\text{MAX}_{j \in \mathcal{N}(v_i)} (\gamma_{j,i}^{(t)}), \sum_{j \in \mathcal{N}(v_i)} \gamma_{j,i}^{(t)} \right), \quad (9b)$$

$$\beta_i^{(t)} = \text{MLP2}(\alpha_i^{(t)}, h_{ii}, \beta_i^{(t-1)}, w_i), \quad (9c)$$

where $\text{MAX}(\cdot)$ is to take the largest value in a set, MLP1 and MLP2 represent two different MLPs, CONCAT denotes the operation of vector concatenations. $\gamma_{j,i}^{(t)}$ denotes the feature vector of the edge connecting node v_j and node v_i in the t -th iteration. $\alpha_i^{(t)}$ is the aggregated information from the neighbor nodes to the central node v_i , and $\beta_i^{(t)}$ is the updated hidden representation of node v_i in the t -th iteration. The IGCNet is trained in an unsupervised manner to learn the optimal power control.

In addition, a family of neural networks, i.e., message passing graph neural networks (MPGNNs), is designed to solve the problem of radio resource management in wireless networks [79]. It demonstrates that MPGNNs satisfy the permutation equivariance property and have the ability to address the resource management problem of large-scale wireless networks while enjoying a high computational efficiency. To guarantee an effective implementation, this work further proposed a wireless channel graph convolution network (WGCN) belonging to the MPGNNs class. The effectiveness of WGCN is evaluated with respect to the power control and beamforming problems. It demonstrates that WGCN matches or outperforms the classic optimization-based algorithms and does not need domain knowledge and has significant speedups. However, MPGNNs just consider the problem with simple constraints or without constraints, complex resource constraints need to be further considered.

To solve the power allocation problem for device-to-device (D2D) wireless networks, Spectral-based GCNs was employed in [80]–[82]. The proposed model, i.e., the random edge graph neural networks (REGNN), performs the

convolutions over a random graph formed by the fading interference patterns in wireless networks. The authors further presented an unsupervised model-free primal-dual learning algorithm to train the weights of the REGNN to overcome the difficulties incurred by the constrained objective function. Additionally, REGNN is utilized to solve the problem of power control in decentralized wireless networks [83]. To adapt the time-varying topologies, the first-order meta-learning is adopted to adapt the new network configurations with a few shots exploiting the data obtained from multiple topologies. The problem of down-link power control in wireless networks over a single shared wireless medium is investigated and addressed by using spectral-based GCNs and primal-dual learning [84]. The main highlight of the works aforementioned is to solve the resource management problem under the complex constraints via the primal-dual learning method in homogeneous wireless networks.

Compared with the homogeneous wireless networks, it is more challenging to design the GNNs-based learning mechanism for the resource allocation problem in heterogeneous wireless networks. Zhang *et al.* focused on addressing the problem of power control or beamforming using MPNN in heterogeneous D2D networks [30]. This work considers a heterogeneous D2D network with two types of links, in which each kind of link holds different features, as depicted in Fig. 3. In particular, let $r = (n, m)$ be the interference from link type n to link type m , the update rules in relation (n, m) is defined as follows

$$\mathbf{e}_{j_{n,m}}[L] = \phi_{(n,m)}^e(\mathbf{v}_{j_n}[L-1], \mathbf{e}_{j_{n,m}}[0]), \quad (10a)$$

$$\mathbf{v}_{i_m}^{(n)} = \phi_{(n,m)}^v \left(\mathbf{v}_{i_m}[L-1], \max_{j \in \mathcal{N}_{i_m}^{(n)}} \mathbf{e}_{j_{n,m}}[L] \right). \quad (10b)$$

where ϕ_r^e and ϕ_r^v are an edge update function and a node update function of relation r , respectively. The aggregation

TABLE 2. Applications of GNNs in wireless networks.

Area	Year	Application	Wireless network type	Algorithm	Scheme	
Resource Allocation	2019	Power control in ad-hoc wireless networks	Mesh/Ad-hoc	Spatial-based GCNs	Y. Shen <i>et al.</i> [29]	
	2021	Radio resource management in mmWave networks	Mesh/Ad-hoc	Spatial-based GCNs	Y. Shen <i>et al.</i> [79]	
	2019	Power allocation with a set of transmitter receiver pairs in a large scale wireless network	Mesh/Ad-hoc	Spectral-based GCNs	M. Eisen <i>et al.</i> [80]	
	2020	Power allocation in wireless networks	Ad-hoc/Cellular	Spectral-based GCNs	M. Eisen <i>et al.</i> [81]	
	2020	Power control in ad-hoc wireless networks	Mesh/Ad-hoc	Spectral-based GCNs	M. Eisen <i>et al.</i> [82]	
	2021	Power control in decentralized wireless networks	Mesh/Ad-hoc	Spectral-based GCNs	I. Nikoloska <i>et al.</i> [83]	
	2020	Downlink power control in wireless networks	Mesh/Ad-hoc	Spectral-based GCNs	N. Naderializadeh <i>et al.</i> [84]	
	2021	Power control/beamforming in heterogeneous D2D networks	Mesh/Ad-hoc	Spatial-based GCNs	X. Zhang <i>et al.</i> [30]	
	2021	Power control in multi-cell cellular networks	Cellular	Spatial-based GCNs	J. Guo <i>et al.</i> [85]	
	2021	Joint user association and power allocation in heterogeneous ultra dense network	Cellular	Spatial-based GCNs	X. Zhang <i>et al.</i> [31]	
	2020	Resource allocation in free space optical (FSO) fronthaul networks	Mesh/Ad-hoc	Spectral-based GCNs	Z. Gao <i>et al.</i> [86]	
	2020	Resource allocation problems under asynchronous wireless network setting	Mesh/Ad-hoc	Spectral-based GCNs	Z. Wang <i>et al.</i> [87], [88]	
	2021	Resource allocation in wireless IoT networks	Mesh/Ad-hoc	Spatial-based GCNs	T. Chen <i>et al.</i> [89]	
	2020	Power allocation in a single-hop ad-hoc wireless network	Mesh/Ad-hoc	Spectral-based GCNs	A. Chowdhury <i>et al.</i> [90], [91]	
	2019	Link scheduling in D2D networks	Mesh/Ad-hoc	Graph embedding	W. M. Lee <i>et al.</i> [92]	
	2020	Schedule transmission for wireless networks in a distributed manner	Mesh/Ad-hoc	Spectral-based GCNs	Z. Zhao <i>et al.</i> [93]	
	2019	Temporal link prediction in various network systems	Mesh/Ad-hoc	Spectral-based GCNs	K. Lei <i>et al.</i> [94]	
	2021	Joint link scheduling and beam selection in ultra-dense D2D mmWave communication networks	Mesh/Ad-hoc	Spatial-based GCNs	S. He <i>et al.</i> [95]	
	Emerging Fields	2020	Channel allocation for densely deployed WLANs	WLANs	DRL with Spectral-based GCN	K. Nakashima <i>et al.</i> [96]
		2020	Spectrum allocation in vehicle-to-everything (V2X) networks	Mesh/Ad-hoc	Spatial-based GCNs with Multi-agent RL	Z. He <i>et al.</i> [97]
2021		AP selection for Cell-Free massive MIMO systems	Cellular	Spatial-based GCNs	V. Ranasinghe <i>et al.</i> [98]	
2021		Intelligent reflecting surface and beamforming	Cellular	Spatial-based GCNs	T. Jiang <i>et al.</i> [99]	
2020		Channel estimation for wireless networks	Cellular	GAT	K. Tekb y k <i>et al.</i> [100]	
2020		Massive MIMO detection in wireless communication	Mesh/Ad-hoc	Spatial-based GCNs	A. Scotti <i>et al.</i> [101]	
2020		Channel tracking for the massive MIMO networks	Cellular	GAE	Y. Yan <i>et al.</i> [102]	
2020		Cellular traffic prediction	Cellular	Spectral-based GCNs and CNN with GLU	S. Zhao <i>et al.</i> [103]	
2020		Satellite traffic prediction	Mesh/Ad-hoc	Spectral-based GCNs with graph GRU	L. Yang <i>et al.</i> [104]	
2020		Multiagent cooperative control for CAV networks	Mesh/Ad-hoc	Spectral-based GCNs with RL	J. Dong <i>et al.</i> [105]	
2020		Active traffic management for CAV networks	Mesh/Ad-hoc	Spectral-based GCNs with RL	P. Y. J. Ha <i>et al.</i> [106]	
2020		Efficient point cloud processing	Mesh/Ad-hoc	Dynamic GCNs	J. Shao <i>et al.</i> [107]	
2020		Point cloud delivery	Mesh/Ad-hoc	GAE	T. Fujihashi <i>et al.</i> [108]	
2021		3D object detection	Mesh/Ad-hoc	3D GNN	C. S. Jeong <i>et al.</i> [109]	
2019		Throughput maximization for UAV assisted ground networks	Mesh/Ad-hoc	Spectral-based GCNs	S. Lohani <i>et al.</i> [110]	
2021		Wireless network localization	-	Spectral-based GCNs	W. Yan <i>et al.</i> [111]	
2021		Performance prediction in Next-Generation WLANs	WLANs	Spatial-based GCNs	P. Soto <i>et al.</i> [112]	
2021	Routing in small satellite networks	Mesh/Ad-hoc	Spectral-based GCNs	M. Liu <i>et al.</i> [113]		
2021	Decentralized control in wireless communication systems	Mesh/Ad-hoc	Spectral-based GCNs	M. Lee <i>et al.</i> [114]		

rules is given by

$$\mathbf{v}_{i_m}[l] = \rho_m^{v \rightarrow v} \left(\left\{ \mathbf{v}_{i_m}^{(n)}[l] \right\}_n \right) = \frac{1}{c_{i,m}} \sum_n \mathbf{v}_{i_m}^{(n)}[l], \quad (11)$$

where $c_{i,m}$ is the number of relations causing interference to link i_m . $\rho_m^{v \rightarrow v}(\cdot)$ is the aggregation function of node to node with link type m .

Similarly, Guo and Yang considered the power control problem in multi-cell cellular networks [85]. Specifically, this work models the cellular networks as a heterogeneous graph, i.e., wireless interference graph, and then proposed a heterogeneous GNN (HetGNN) based on spatial-based GCNs, called PGNN, to learn the power control policy

in multi-cell cellular networks. Inspired by the finding that the parameter sharing scheme determines the invariance or equivalence relationship, the optimal power control policy has a combination of different PI and PE properties that existing heterogeneous GNNs do not satisfy [115]. Additionally, Zhang *et al.* considered the joint user association and power control problem in HUDNs [31]. The HUDNs are also modeled as a heterogeneous graph, which is shown in Fig. 6. A heterogeneous GraphSAGE (HGSAGE) that extended from GraphSAGE [43], is used to extract the latent node representations. To embrace both the generalization of the learning algorithm and the higher performance of HUDNs, the learning process of HUDNs is divided into two phases. The first

phase of HUDNs learns a representation with a tremendous generalized ability to suit any scenario with different user distributions in an off-line manner. The second phase of HUDNs is to finely tune the parameters of GNN online to further improve the performance for quasi-static user distribution.

Gao *et al.* investigated the optimal power assignment and node selection based on the instantaneous channel state information of the links in free space optical (FSO) fronthaul networks [86]. Spectral-based GCNs are utilized to exploit the FSO network structure with small-scale training parameters. Then, a primal-dual learning algorithm is developed to train the GNN in a model-free manner. Wang *et al.* addressed the asynchronous decentralized wireless resource allocation problem with a novel unsupervised learning approach [87], [88]. Specifically, the interference patterns between transmitting devices are modeled as a graph to capture the asynchrony patterns via the activation of the edges on a highly granular time scale. A decentralized learning architecture, i.e., the aggregation graph neural networks (Agg-GNNs) is designed based on the graph representation of interference and asynchrony. Chen *et al.* proposed a spatial-based GCNs based framework to address the high complexity of the practical implementation of wireless Internet of things (IoT) networks [89]. The effectiveness of the framework is evaluated by the link scheduling in D2D networks and the joint channel and power allocation in D2D underlaid cellular networks.

The methods proposed by the aforementioned works all are based on data-driven neural networks with poor interpretability and scalability. Inspired by the algorithmic unfolding of the iterative WMMSE, i.e., unfolded WMMSE (UWMMSE), Chowdhury *et al.* proposed a data- and model-driven neural architecture to solve the power allocation problem in a single-hop Ad-hoc wireless network [90]. The optimization problem that should be solved is

$$\min_{\mathbf{w}, \mathbf{a}, \mathbf{b}} \sum_{i=1}^M (w_i q_i - \log w_i), \quad (12a)$$

$$s.t. \quad q_i = (1 - a_i h_{ii} b_i)^2 + \sigma^2 a_i^2 + \sum_{i \neq j} a_i^2 h_{ij}^2 b_j^2, \quad (12b)$$

$$b_i^2 \leq p_{max}, \quad (12c)$$

where $\mathbf{w} = [w_1, w_2, \dots, w_M]^T$, $\mathbf{a} = [a_1, a_2, \dots, a_M]^T$, $\mathbf{b} = [b_1, b_2, \dots, b_M]^T$ are vectors of optimization variables. The allocated power is computed by a function $\mathbf{p} = \Phi(\mathbf{H}; \theta_\vartheta, \theta_\nu)$ of the channel state matrix through a layered architecture Φ with trainable weights θ_ϑ and θ_ν . Precisely, setting $\mathbf{b}^{(0)} = \sqrt{p_{max}} \mathbf{1}$, the t -th layer of UWMMSE is implemented as follows

$$\vartheta_i^{(t)} = \Psi(\mathbf{H}; \theta_\vartheta^{(t)}), \quad \nu_i^{(t)} = \Psi(\mathbf{H}; \theta_\nu^{(t)}), \quad (13a)$$

$$a_i^{(t)} = \frac{h_{ii} \nu_i^{(t-1)}}{\sigma^2 + \sum_j h_{ij}^2 b_j^{(t-1)} b_j^{(t-1)}}, \quad (13b)$$

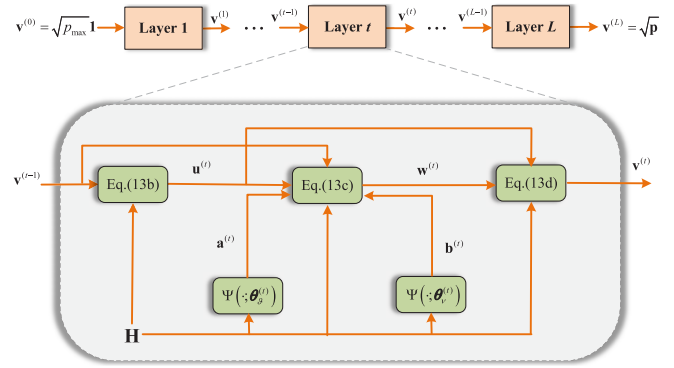


FIGURE 12. Illustration of UWMMSE [90].

$$w_i^{(t)} = \frac{\vartheta_i^{(t)}}{1 - a_i^{(t)} h_{ii} b_i^{(t-1)}} + \nu_i^{(t)}, \quad (13c)$$

$$b_i^{(t)} = \alpha \left(\frac{a_i^{(t)} h_{ii} w_i^{(t)}}{\sum_j h_{ji}^2 a_j^{(t)} a_j^{(t)} w_j^{(t)}} \right), \quad (13d)$$

and the output power is determined as $\mathbf{p} = \Phi(\mathbf{H}; \theta_\vartheta, \theta_\nu) = (\mathbf{b}^{(L)})^2$. $\alpha(z) := [z]_0^{\sqrt{p_{max}}}$ simply ensures that $b_i^{(t)} \in [0, \sqrt{p_{max}}]$. The function Ψ parameterized by θ_ϑ and θ_ν is chosen to be spectral-based GCNs. The whole workflow of UWMMSE is shown in Fig. 12, which has better interpretability and scalability compared with the data-driven learning models. Numerical experiments demonstrate that UWMMSE not only significantly reduced the computational complexity, but also improved the performance compared to the conventional WMMSE [91].

2) LINK SCHEDULING

Although the overall performance of wireless networks can be improved via power control, it is not enough to eliminate the strong interference for ultra-dense wireless networks. Link scheduling is regarded as an effective means to further improve the performance of wireless networks.

To overcome the high computational complexity of the traditional optimization methods and eliminate the costly channel estimation, Lee *et al.* proposed a novel DL-based graph embedding method to implement the link scheduling in D2D networks [92]. In detail, this work firstly models the D2D network as a fully connected directed graph, then computes a low-dimensional feature vector based on the distances of both communication and interference links without requiring the accurate channel state information for each node. Finally, a multi-layer classifier is utilized to learn the scheduling policy in a supervised and unsupervised manner, respectively. Numerical results show that this method can achieve near-optimal performance compared with the state-of-the-art methods but with a small number of training samples, and has competitive generalization and scalability.

A distributed scheduling scheme was proposed to overcome the difficulty encountered in solving the maximum weighted independent set (MWIS) problem for wireless

networks [93]. The authors proposed a distributed MWIS solver based on spectral-based GCNs for link scheduling by combining the learning capabilities of GCNs and the efficiency of greedy MWIS solvers. The proposed solver achieves superior performance over greedy baselines with minimum increase in complexity, and generalizes well across different types of graphs and utility distributions.

Besides, the information of the dynamics, the topology structure and evolutionary patterns of dynamic networks can be fully exploited to improve the temporal link prediction performance. In dynamic network scenarios, Lei *et al.* introduced a novel non-linear GCN-GAN model by leveraging the benefits of spectral-based GCNs, LSTM as well as the GANs to tackle the challenging temporal link prediction task [94]. While in ultra-dense D2D mmWave communication networks, in order to effectively control the interference between communication pairs, spatial-based GCNs and primal-dual learning are utilized to solve the problem of joint beam selection and link activation across a set of communication pairs via inactivating part communication pairs [95].

3) OTHERS

To improve the spectral efficiency in densely deployed WLANs, Nakashima *et al.* proposed a deep RL model using spectral-based GCNs for channel allocation [96]. The idea behind their work is that the objective function is modeled as a parametric function of topologies, channels and communication quality. He *et al.* studied the spectrum allocation via learning the low dimensional representations of a graph by modeling the V2X network as a graph, where each vehicle-to-vehicle link is a node in the graph [97]. According to the learned characteristics, multi-agent RL is used for spectrum allocation. DQN is used to learn to optimize the total capacity of the V2X network. In [98], GraphSAGE is employed to predict the potential links between APs for cell-free massive MIMO.

4) BRIEF DISCUSSION

The application of GNNs in resource allocation mainly focuses on power allocation and link scheduling, while there is less work related to spectrum allocation and channel allocation. In addition, almost all works adopt GCNs and a few of them introduce RL in terms of the paradigms of GNNs. Although these works have achieved good results, there are also some issues needed to be considered. For one thing, resource allocation tasks in some wireless networks, such as intelligent factories and intelligent transportation, etc., may have many simple or complex constraints needed to be handled. Most of the existing works directly use projection strategy for simple constraints, while Lagrange dual learning framework is used for complex constraints. Although Lagrange dual learning framework can deal with complex constraints, it can not guarantee the complete satisfaction of constraints, and the training efficiency is not ideal. For

another thing, in some wireless network scenarios with delay-sensitive traffics, the designed model should have low time complexity on the basis of achieving certain performance. A small amount of works has discussed the processing delay of the designed model, but there is still a certain distance from practical application. Therefore, the problem of model complexity is a direction worthy of exploration and research. On the other hand, most of the work is to solve the optimization problem of a single network layer. With the development of information and communication technologies, joint resource allocation at different network layers, i.e., cross-layer optimization, is a potential direction for the design of learning methods based on GNNs. The advantage of cross-layer optimization is that it can comprehensively utilize the information between different network layers, and then it may get performance improvement potentially.

B. EMERGING FIELDS

GNNs are also applied in other wireless networks scenarios. Although these studies are in their infancy, the results of the existing works show that GNNs have a good application prospect in these directions.

1) CHANNEL ESTIMATION

Accurate CSI is important for guaranteeing the performance of massive MIMO high-dynamic networks. However, traditional solutions rely so much on hypothetical statistical models that they are hard to adapt the high-dynamic network environment. To overcome this defect, many researchers use DL to estimate channel of wireless communication system in recent years, but the ability of DL to extract spatial dependency is limited. However, GNNs have advantages in spatial dependency mining, so GNNs have been applied in this field in recent years.

The estimation of channels between the intelligent reflecting surface (IRS), the base station (BS), and the users is necessary for the optimal tuning of phase shifters at the IRS. Jiang *et al.* proposed a DL model to configure the IRS and beamforming at the BS such that the system utility function is maximized directly based on the received pilots instead of the channel coefficients [99]. Specifically, spatial-based GCNs is utilized to directly map the received pilots to the beamformers at the BS and the reflective pattern at the IRS. While in [100], GAT is employed to solve the channel estimation for the two-way backhaul link of high-altitude platform stations with reconfigurable intelligent surfaces. Numerical results show that for the full-duplex channel estimation, the performance of the GAT estimator is better than the least-squares. Moreover, numerical results also show that even if the training data does not include all changes, the GAT estimator is robust to hardware impairments and small-scale fading characteristics changes. As a further case of channel control, Scotti *et al.* considered the inference task of massive MIMO detection under time-varying channels and higher-order quadrature amplitude

modulation and proposed a message-passing solution based on GNNs, i.e., MIMO-GNN [101].

Yang *et al.* proposed a new channel tracking method based on GAE [102]. Specifically, the channel tracking framework is designed as

$$\bar{e}'_{i,j}(t) = \text{MLP}_e^{\text{coder}}(\bar{e}_{i,j}(t), \bar{v}_i(t), \bar{v}_j(t)), \quad (14a)$$

$$\bar{v}'_i(t) = \text{MLP}_v^{\text{coder}}\left(\bar{v}_i(t), \sum_{\bar{v}_j \in \mathcal{N}(\bar{v}_i(t))} \bar{e}'_{i,j}(t)\right), \quad (14b)$$

$$\hat{e}_{i,j}(t) = \text{MLP}_e^{\text{decoder}}(\bar{e}'_{i,j}(t)), \quad (14c)$$

$$\hat{v}_i(t) = \text{MLP}_v^{\text{decoder}}(\bar{v}'_i(t)), \quad (14d)$$

where $\bar{v}_i(t)$ and $\bar{e}_{i,j}(n)$ are the resultant node feature of node i and edge feature between nodes i and j at time t , respectively. $\text{MLP}_v^{\text{coder}}$ and $\text{MLP}_e^{\text{coder}}$ are coders for node and edge, respectively. In contrary, $\text{MLP}_v^{\text{decoder}}$ and $\text{MLP}_e^{\text{decoder}}$ are decoders. It's not hard to find that the designed framework combines MPNN and codec. Numerical results confirm that the GNN-based scheme outperforms the feed-forward neural network in terms of the MSE.

2) TRAFFIC PREDICTION

Generally speaking, effective resource management can improve the utilization of network resources. In addition, if one can predict the required resource of future wireless traffic, resource management will become more flexible in wireless networks. However, the high spatial-temporal interdependencies make traffic prediction more challenging. There are fewer works on wireless network traffic prediction using GNNs, mainly including cellular network traffic prediction and satellite network traffic prediction.

To improve the accuracy of cellular traffic prediction, Zhao *et al.* proposed a new Spatio-Temporal GCNs incorporating Handover information (STGCN-HO) prediction model using the transition probability matrix of the handover graph [103]. STGCN-HO builds a stacked residual neural network structure that combines spectral-based GCNs and CNN with gated linear units [116] to capture the spatial and temporal interdependencies of traffic. Compared with RNN, STGCN-HO has a faster training speed due to the use of CNN, and has the ability to train or predict cell or base stations with the information collected from the entire graph at the same time. In addition, compared with CNN grid, STGCN-HO can predict both base stations and the cells within the base stations. While in reference to the satellite network traffic prediction, Yang *et al.* pointed out that the traditional network traffic prediction model could not effectively extract the spatio-temporal characteristics of network traffic. Therefore, they proposed a network traffic prediction model GCN-GRU via combining the spectral-based GCNs with GRU [104]. Specifically, GCN-GRU model utilizes the spectral-based GCN to extract the spatial characteristics of the satellite network traffic, and utilizes GRU model to extract the temporal characteristics of the satellite network

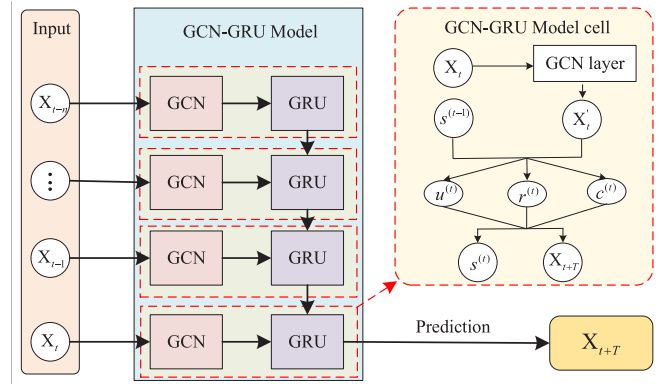


FIGURE 13. Illustration of GCN-GRU.

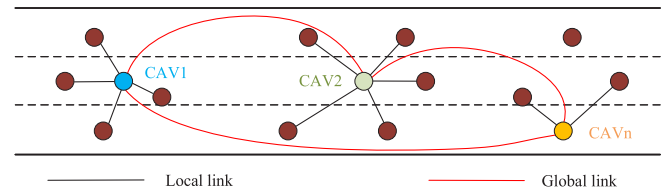


FIGURE 14. Illustration of graph construction for CAVs.

traffic, and finally predict satellite network traffic through the fully connected layer. The illustration of GCN-GRU is shown in Fig. 13, where $r^{(t)}$, $u^{(t)}$, $c^{(t)}$ and $s^{(t)}$ correspond to (8a)-(8d) and X_t is the input feature at time t .

3) VEHICLE COMMUNICATION

Recently, GNNs have been applied to control the connected autonomous vehicles (CAVs) lane changing decisions for a road segment, to mitigate the highway bottleneck congestion, and to allocate spectrum in V2X networks. Dong *et al.* proposed a DL model that combines spectral-based GCN and a deep Q network to control multiple CAVs to make cooperative lane change decisions [105]. The graph construction methods of CAVs is shown in Fig. 14. There is a state $S^{(t)}$ that is considered as a triplet at time t , i.e., $S^{(t)} = \{\mathbf{X}^{(t)}, \mathbf{A}^{(t)}, \mathbf{M}^{(t)}\}$, where $\mathbf{X}^{(t)}$, $\mathbf{A}^{(t)}$ and $\mathbf{M}^{(t)}$ denote the node feature matrix, adjacent matrix and a mask matrix that document the index of autonomous vehicles at time t , respectively. From the perspective of CAV operations, the proposed model not only enables CAV to successfully carry out lane changes to meet its personal intention of merging from the prescribed ramp, but also guarantees safety and efficiency. Similarly, RL algorithms are employed to train CAV driving behaviors, which can be used to relieve highway bottleneck congestion [106].

4) OTHERS

In addition to the aforementioned related works, GNNs are also used to solve other problems in wireless networks. Shao *et al.* proposed Branchy-GNN using the branch network and source-channel coding to reduce the computational cost and intermediate feature transmission overhead for efficient point cloud processing [107]. Fujihashi *et al.* proposed a

novel soft point cloud transmission scheme that combines GNN-based point cloud coding and near-analog modulation for future wireless streaming of holographic and three-dimensional data [108]. Jeong *et al.* provided AR services via three-dimensional GNN using cameras and sensors on mobile devices [109]. Lohani and Glasser designed a model based on spectral-based GCNs to maximize the throughput of unmanned aerial vehicle (UAV)-assisted ground networks [110]. Spectral-based GCNs were utilized to solve the network localization problem of a wireless network in two-dimensional space [111]. In [112], spatial-based GCNs was first used to predict the achieved throughput in highly dense WLANs using channel bonding. Using the topology extraction ability of GNN, spectral-based GCNs based learning routing scheme was proposed to implement onboard routing in small satellite networks [113]. Lee *et al.* analyzed and enhanced the robustness of the decentralized GNN in different wireless communication systems, making the prediction results not only accurate but also robust to transmission errors [114].

5) BRIEF DISCUSSION

Emerging fields where GNNs are applied mainly include channel estimation, channel tracking, MIMO detection, traffic prediction, vehicle communication, point clouds, and so on. There is not much work in these fields, but some fields are worthy of further exploration. For instance, the traffic prediction task not only plays a pivotal role in the management of wireless network resources but also has higher requirements for data collection and acquisition. On the one hand, traffic prediction generally requires long-term historical data. It also involves issues such as the granularity of data collection, the level of data collection, and data privacy, which in turn brings varying degrees of difficulties. On the other hand, for scenarios with high requirements for traffic prediction service delay, the model designed is required to be as low as possible in complexity, so as to achieve real-time prediction capabilities. In addition, in terms of the application of the paradigms of GNNs in emerging fields, GCNs are frequently utilized. While there are also a small number of new paradigms of GNNs applied, such as GAT, GAE, and several generalized GCNs. According to the characteristics of different business needs, we can measure and compare different GNNs paradigms, and then adopt the best GNNs paradigm to better solve the problems faced.

V. KEY ISSUES AND FUTURE DEVELOPMENT

Although GNNs have made some progress in the application of wireless networks, some key issues need to be further studied in-depth. Accordingly, some ongoing or future research directions that are worth exploring are summarized as follows:

A. ACQUISITION OF HIGH-QUALITY DATA

One of the fundamental elements of the data-driven DL method is to obtain a lot of training and testing data of

wireless networks. The higher the quality of data, the better the training of the model. Though many researchers have studied the application of DL in wireless networks, most of the datasets used in the existing work are generated by numerical simulation, which is somewhat different from the real data of wireless networks. Furthermore, unlike the successful application of DL in image processing and social networks, etc, there is a lack of publicly recognized data set for the physical layer and media access layer of wireless networks. Therefore, it is urgent to collect and construct opening wireless network datasets for method comparison and performance verification.

Wireless networks have some unique characteristics, such as high dynamic, heterogeneous terminals, and non-uniform, resulting in many difficulties in obtaining the data of wireless networks, especially the real-time communication data of the physical layer and media access control layer. The data of wireless networks can generally be obtained from spectrum measurement instruments, base station, core network equipment, user terminals, and so on. Different data acquisition devices may be provided by different manufacturers, who may define different data extraction formats, feature names, and data calculation methods, and may also be different in the time granularity of extraction. This makes it very challenging to collect massive data in the real wireless network, especially at the physical layer and media access control layer. Meanwhile, the construction of opening testing and training datasets is still a very urgent and challenging task for the successful application of DL and GNN, etc, in wireless networks.

B. DISTRIBUTED GNN LEARNING MODEL

As we all know, the goal of whether the data-driven DL methods or the data- and model-driven methods is to learn the super parameters of NNs that depend on the super-computing power. Furthermore, the learning abilities of the DL methods are proportional to the network complexity. However, in wireless networks, the computing capacity of communication nodes is very limited, especially for the battery-powered lightweight devices. How to design a proper DL model is a challenging and opening problem for the battery-powered lightweight devices. Fortunately, distributed ML is regarded as an effective and efficient technology to balance the performance and the computational resource and to reduce the required amount of training and testing data. However, several issues are needed to be considered when designing a GNN-based distributed learning model, such as the split of the GNN-based learning model, the parameter updating strategy, the integration of the results of each distributed running node, etc. In addition, the convergence of the whole distributed learning model is also should be guaranteed.

C. DATA PRIVACY ISSUES

Generally speaking, the acquisition of the data of wireless networks inevitably touches the user's privacy. However,

privacy protection is one of the core issues in the field of information and communications technology in the future, and the process of data acquisition is cumbersome. Consequently, the acquisition of real network data will face many difficulties. On the other hand, distributed learning model needs data interaction, which also involves data privacy. In some scenarios with high requirements for data privacy protection, data interaction security needs to be ensured. Some methods can be used to solve data privacy issues, such as secure multi-party computing and trusted execution environments, etc. Another popular solution is to use the federated learning, which has the outstanding advantage of privacy protection.

D. ROBUSTNESS OF THE LEARNING MODEL

Although DL methods have been successfully applied in many fields, their effectiveness depends on high-quality training datasets. When the training dataset contains significant complex noise, anomaly intrusion, category imbalance, etc., its effectiveness cannot be guaranteed. This implies that the robustness of the learning model should be considered when designing the DL model for wireless networks. Generally, the robustness of the model can be discussed from two aspects, i.e., malicious or abnormal data and the defense ability of the model. On the one hand, when a malicious attacker attacks the model with input data, it can generally be defended by adversarial training, input transformation, gradient shielding, and detection and rejection. On the other hand, the model's ability to resist attacks can be improved by improving the model itself and combining it with other security technologies.

E. SCALABILITY OF THE LEARNING MODEL

In wireless networks, network scale, network state, and network service data volume may be highly dynamic, which requires stronger scalability of the designed model. For example, the changes in network scale (such as changes in the number of communication devices) and network state (such as changes in the location and attributes of network communication nodes) are enough to bring great challenges to the design of the learning model. Fortunately, GNNs have the ability to address these problems based on the existing research work, but the current work is still in its infancy and needs further research. On the other hand, with the high dynamic change of network business data volume, such as the sharp increase of business data volume, the real-time processing capability of the model should be enhanced accordingly.

F. CROSS-LAYER OPTIMIZATION

Most of the existing works have covered different individual network layers in terms of GNNs applications. These algorithms generally cannot obtain the global optimal solution from the perspective of whole communication networks. More recently, with the development of information and

communication technologies, joint resource allocation at different network layers, i.e., cross-layer optimization, is regarded as a potential direction for further improving the performance of learning methods. The advantage of cross-layer optimization is that the information at multiple network layers can be fully explored and exploited to design learning methods. Although cross-layer optimization has the potential to improve performance, it introduces more optimization parameters. Meanwhile, the objective function will be more complicated, which in turn requires higher computing power and may introduce relatively large processing delays. Parameter abstraction can reduce the complexity to a certain extent, but it may reduce the optimality of the generated configuration.

G. RESEARCH OF GNN-BASED DEEP UNFOLDING METHOD

One shortcoming of the data-driven learning method is the lack of interpretability due to the NN-based methods treat the learning of mapping between the input and output as a black box, which weakens the domain knowledge. To tackle the shortcoming, an emerging direction is the algorithm unfolding (unrolling) aiming at combining the knowledge of data and domain fields. Motivated by the successful application of algorithm unfolding solving some classical problems like UWMMSE, recently, deep unfolding is regarded as an effective combination to not only effectively utilize the interpretability and scalability of model-driven algorithms, but also use the expressive power of data-driven methods. Although deep unfolding has been paid a lot of attention from both the industry and academia, the problem to be solved has only a few simple constraints. Meanwhile, there is a little work to deeply study the GNN-based unfolding models to solve the optimization problem in the wireless network under practical constraints. This implies that there are still a lot of works to be explored for wireless networks.

H. MODEL DESIGN WITH COMPLEX CONSTRAINTS

So far, a large amount of work focuses on designing data-driven or data- and model-driven DL models aiming to solve the optimization problem of wireless network, however, these problem investigated is generally formulated subjecting to simpler constraints. In practical wireless communication, the transmission schemes may be investigated under some complex constraints, such as the per-antenna or per-base station power constraint, the quality of service constraint, delivery latency, etc. How to design an effective and efficient DL model under complex constraints is still an opening and challenging issue. Although Lagrange dual learning framework has been used to deal with such problems, there are still some problems with this method. On the one hand, the convergence speed and the effectiveness of the Lagrange dual learning framework are closely related to the updating step-sizes of the Lagrange multipliers and the model parameters, while the update step-sizes of the Lagrange multipliers and the model parameters are coupled with each other, which

makes it difficult to determine the approximate update step-sizes to maximize the performance of the DL model. On the other hand, the DL models trained based on Lagrange dual learning framework are generally difficult to fully satisfy constraints (especially for the problems with large-scale complex constraints), and the generalization performance of models is also questionable.

VI. CONCLUSION

In this paper, we first illustrated the construction method of WCG for various wireless networks. Then, we simply introduced several classical paradigms of GNNs that have been applied in wireless networks, and made a classified introduction for the application of GNNs in wireless networks, mainly including resource allocation and several emerging fields. From the overview results, the application of GNNs in wireless networks is still in its infancy. Many challenging problems are needed to be further solved and improved. Finally, based on the existing results, several key issues and research directions are summarized for participants interested in this domain.

REFERENCES

- [1] S. He *et al.*, "A survey of millimeter-wave communication: Physical-layer technology specifications and enabling transmission technologies," *Proc. IEEE*, vol. 109, no. 10, pp. 1666–1705, Oct. 2021.
- [2] Q. Mao, F. Hu, and Q. Hao, "Deep learning for intelligent wireless networks: A comprehensive survey," *IEEE Commun. Surveys Tuts.*, vol. 20, no. 4, pp. 2595–2621, 4th Quart., 2018.
- [3] C. Zhang, P. Patras, and H. Haddadi, "Deep learning in mobile and wireless networking: A survey," *IEEE Commun. Surveys Tuts.*, vol. 21, no. 3, pp. 2224–2287, 3rd Quart., 2019.
- [4] H. Sun, X. Chen, Q. Shi, M. Hong, X. Fu, and N. D. Sidiropoulos, "Learning to optimize: Training deep neural networks for interference management," *IEEE Trans. Signal Process.*, vol. 66, no. 20, pp. 5438–5453, Oct. 2018.
- [5] A. Bora, A. Jalal, E. Price, and A. G. Dimakis, "Compressed sensing using generative models," in *Proc. Int. Conf. Mach. Learn. (ICML)*, Mar. 2017, pp. 537–546.
- [6] N. Shlezinger, N. Farsad, Y. C. Eldar, and A. J. Goldsmith, "ViterbiNet: A deep learning based viterbi algorithm for symbol detection," *IEEE Trans. Wireless Commun.*, vol. 19, no. 5, pp. 3319–3331, May 2020.
- [7] H. Kim, Y. Jiang, R. B. Rana, S. Kannan, S. Oh, and P. Viswanath, "Communication algorithms via deep learning," in *Proc. Int. Conf. Learn. Represent.*, Feb. 2018, pp. 1–17.
- [8] K. Gregor and Y. LeCun, "Learning fast approximations of sparse coding," in *Proc. 27th Int. Conf. Mach. Learn. (ICML)*, Jul. 2010, pp. 399–406.
- [9] M. Kulin, T. Kazaz, E. De Poorter, and I. Moerman, "A survey on machine learning-based performance improvement of wireless networks: PHY, MAC and network layer," *Electronics*, vol. 10, no. 3, pp. 1–63, Mar. 2021.
- [10] F. Tang, B. Mao, Y. Kawamoto, and N. Kato, "Survey on machine learning for intelligent end-to-end communication toward 6G: From network access, routing to traffic control and streaming adaptation," *IEEE Commun. Surveys Tuts.*, vol. 23, no. 3, pp. 1578–1598, 3rd Quart., 2021.
- [11] F. Liang, C. Shen, W. Yu, and F. Wu, "Towards optimal power control via ensembling deep neural networks," *IEEE Trans. Commun.*, vol. 68, no. 3, pp. 1760–1776, Mar. 2020.
- [12] W. Lee, M. Kim, and D.-H. Cho, "Deep power control: Transmit power control scheme based on convolutional neural network," *IEEE Commun. Lett.*, vol. 22, no. 6, pp. 1276–1279, Jun. 2018.
- [13] S. Zhang, B. Yin, and Y. Cheng, "Topology aware deep learning for wireless network optimization," Jan. 2021, *arXiv:1912.08336*.
- [14] M. M. Bronstein, J. Bruna, Y. LeCun, A. Szlam, and P. Vandergheynst, "Geometric deep learning: Going beyond euclidean data," *IEEE Signal Process. Mag.*, vol. 34, no. 4, pp. 18–42, Jul. 2017.
- [15] Z. Wu, S. Pan, F. Chen, G. Long, C. Zhang, and P. S. Yu, "A comprehensive survey on graph neural networks," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 1, pp. 4–24, Jan. 2021.
- [16] Z. Zhang, P. Cui, and W. Zhu, "Deep learning on graphs: A survey," *IEEE Trans. Knowl. Data Eng.*, to be published.
- [17] F. Xia *et al.*, "Graph learning: A survey," *IEEE Trans. Artif. Intell.*, vol. 2, no. 2, pp. 109–127, Apr. 2021.
- [18] Z. Chen *et al.*, "Bridging the gap between spatial and spectral domains: A survey on graph neural networks," Feb. 2020, *arXiv:2002.11867*.
- [19] R. Sato, "A survey on the expressive power of graph neural networks," Mar. 2020, *arXiv:2003.04078*.
- [20] L. Ruiz, F. Gama, and A. Ribeiro, "Graph neural networks: Architectures, stability, and transferability," *Proc. IEEE*, vol. 109, no. 5, pp. 660–682, May 2021.
- [21] S. Abadal, A. Jain, R. Guirado, J. López-Alonso, and E. Alarcón, "Computing graph neural networks: A survey from algorithms to accelerators," Sep. 2020, *arXiv:2010.00130*.
- [22] H. Yuan, H. Yu, S. Gui, and S. Ji, "Explainability in graph neural networks: A taxonomic survey," Dec. 2020, *arXiv:2012.15445*.
- [23] J. Ye, J. Zhao, K. Ye, and C. Xu, "How to build a graph-based deep learning architecture in traffic domain: A survey," *IEEE Intell. Transp. Syst. Mag.*, early access, Dec. 29, 2020, doi: [10.1109/TITS.2020.3043250](https://doi.org/10.1109/TITS.2020.3043250).
- [24] W. Liao, B. Bak-Jensen, J. R. Pillai, Y. Wang, and Y. Wang, "A review of graph neural networks and their applications in power systems," Jan. 2021, *arXiv:2101.10025*.
- [25] Y. Gao, Y.-F. Li, Y. Lin, H. Gao, and L. Khan, "Deep learning on knowledge graph for recommender system: A survey," Mar. 2020, *arXiv:2004.00387*.
- [26] L. C. Lamb, A. Garcez, M. Gori, M. Prates, P. Avelar, and M. Vardi, "Graph neural networks meet neural-symbolic computing: A survey and perspective," Feb. 2020, *arXiv:2003.00330*.
- [27] Z. Wu, S. Pan, F. Chen, G. Long, C. Zhang, and P. S. Yu, "A comprehensive survey on graph neural networks," Jan. 2019, *arXiv:1901.00596*.
- [28] Y. Ma, J. Hao, Y. Yang, H. Li, J. Jin, and G. Chen, "Spectral-based graph convolutional network for directed graphs," Jul. 2019, *arXiv:1907.08990*.
- [29] Y. Shen, Y. Shi, J. Zhang, and K. B. Letaief, "A graph neural network approach for scalable wireless power control," in *Proc. IEEE Globecom Workshops (GC Wkshps)*, Dec. 2019, pp. 1–6.
- [30] X. Zhang, H. Zhao, J. Xiong, L. Zhou, and J. Wei, "Scalable power control/beamforming in heterogeneous wireless networks with graph neural networks," Apr. 2021, *arXiv:2104.05463*.
- [31] X. Zhang, H. Zhao, J. Xiong, L. Zhou, and J. Wei, "Joint user association and power allocation in heterogeneous ultra dense network via semi-supervised representation learning," Mar. 2021, *arXiv:2103.15367*.
- [32] J. Bruna *et al.*, "Spectral networks and deep locally connected networks on graphs," in *Proc. Int. Conf. Learn. Represent. (ICLR)*, Jan. 2014, pp. 1–14.
- [33] M. Defferrard, X. Bresson, and P. Vandergheynst, "Convolutional neural networks on graphs with fast localized spectral filtering," in *Proc. Int. Conf. Neural Inf. Process. Syst. (NIPS)*, Jun. 2016, pp. 3844–3852.
- [34] T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," Oct. 2016, *arXiv:1609.02907*.
- [35] M. Henaff, J. Bruna, and Y. LeCun, "Deep convolutional networks on graph-structured data," Jun. 2015, *arXiv:1506.05163*.
- [36] R. Li, S. Wang, F. Zhu, and J. Huang, "Adaptive graph convolutional neural networks," Jan. 2018, *arXiv:1801.03226*.
- [37] R. Levie, F. Monti, X. Bresson, and M. M. Bronstein, "CayleyNets: Graph convolutional neural networks with complex rational spectral filters," *IEEE Trans. Signal Process.*, vol. 67, no. 1, pp. 97–109, Jan. 2019.
- [38] S. Zhang, H. Zhang, S. Cui, and Z. Ding, "From spectrum wavelet to vertex propagation: Graph convolutional networks based on Taylor approximation," Nov. 2020, *arXiv:2007.00730*.

- [39] S. Fu, W. Liu, D. Tao, Y. Zhou, and L. Nie, "HesGCN: Hessian graph convolutional networks for semi-supervised classification," *Inf. Sci.*, vol. 514, no. 4, pp. 484–498, Apr. 2020.
- [40] M. Xu, W. Dai, C. Li, J. Zou, H. Xiong, and P. Frossard, "Spectral graph convolutional networks with lifting-based adaptive graph wavelets," Aug. 2021, *arXiv:2108.01660*.
- [41] D. Duvenaud *et al.*, "Convolutional networks on graphs for learning molecular fingerprints," Sep. 2015, *arXiv:1509.09292*.
- [42] J. Gilmer, S. S. Schoenholz, P. F. Riley, O. Vinyals, and G. E. Dahl, "Neural message passing for quantum chemistry," in *Proc. Int. Conf. Mach. Learn. (ICML)*, Apr. 2017, pp. 1263–1272.
- [43] W. L. Hamilton, R. Ying, and J. Leskovec, "Inductive representation learning on large graphs," in *Proc. Conf. Workshop Neural Inf. Process. Syst. (NIPS)*, Jun. 2017, pp. 1025–1035.
- [44] J. Atwood and D. Towsley, "Diffusion-convolutional neural networks," in *Proc. Neural Inf. Process. Syst. (NIPS)*, Nov. 2016, pp. 1993–2001.
- [45] H. Pei, B. Wei, K. C.-C. Chang, Y. Lei, and B. Yang, "Geom-GCN: Geometric graph convolutional networks," in *Proc. Int. Conf. Learn. Represent.*, Sep. 2019, pp. 1–12.
- [46] T. Danel *et al.*, "Spatial graph convolutional networks," in *Proc. Int. Conf. Neural Inf. Process.*, Nov. 2020, pp. 668–675.
- [47] A. Vaswani *et al.*, "Attention is all you need," in *Proc. Conf. Workshop Neural Inf. Process. Syst. (NIPS)*, Jun. 2017, pp. 6000–6010.
- [48] P. Velickovic, G. Cucurull, A. Casanova, A. Romero, P. Liò, and Y. Bengio, "Graph attention networks," in *Proc. Int. Conf. Learn. Represent. (ICLR)*, Oct. 2018, pp. 1–12.
- [49] H. Zhou, T. Young, M. Huang, H. Zhao, J. Xu, and X. Zhu, "Commonsense knowledge aware conversation generation with graph attention," in *Proc. IJCAI*, Jul. 2018, pp. 4623–4629.
- [50] S. Yang, G. Li, and Y. Yu, "Dynamic graph attention for referring expression comprehension," in *Proc. IEEE/CVF Int. Conf. Comput. Vis.*, 2019, pp. 4644–4653.
- [51] L. Li, Z. Gan, Y. Cheng, and J. Liu, "Relation-aware graph attention network for visual question answering," in *Proc. IEEE/CVF Int. Conf. Comput. Vis.*, 2019, pp. 10313–10322.
- [52] X. Wang *et al.*, "Heterogeneous graph attention network," in *Proc. World Wide Web Conf.*, 2019, pp. 2022–2032.
- [53] C. Chen *et al.*, "Gated residual recurrent graph neural networks for traffic prediction," *Proc. AAAI Conf. Artif. Intell.*, vol. 33, no. 1, pp. 485–492, Jul. 2019.
- [54] X. Geng *et al.*, "Spatiotemporal multi-graph convolution network for ride-hailing demand forecasting," *Proc. AAAI Conf. Artif. Intell.*, vol. 33, no. 1, pp. 3656–3663, Jan. 2019.
- [55] B. Yu, H. Yin, and Z. Zhu, "ST-UNet: A spatio-temporal U-network for graph-structured time series modeling," Jul. 2019, *arXiv:1903.05631*.
- [56] K. Guo *et al.*, "Optimized graph convolution recurrent neural network for traffic prediction," *IEEE Intell. Transp. Syst. Mag.*, vol. 22, no. 2, pp. 1138–1149, Feb. 2021.
- [57] Y. Li, R. Yu, C. Shahabi, and Y. Liu, "Diffusion convolutional recurrent neural network: Data-driven traffic forecasting," in *Proc. ICML*, Feb. 2018, pp. 1–16.
- [58] B. Yu, H. Yin, and Z. Zhu, "Spatio-temporal graph convolutional networks: A deep learning framework for traffic forecasting," Jul. 2017, *arXiv:1709.04875*.
- [59] S. Guo, Y. Lin, N. Feng, C. Song, and H. Wan, "Attention based spatial-temporal graph convolutional networks for traffic flow forecasting," *Proc. AAAI Conf. Artif. Intell.*, vol. 33, no. 1, pp. 922–929, Jul. 2019.
- [60] L. Ge, H. Li, J. Liu, and A. Zhou, "Temporal graph convolutional networks for traffic speed prediction considering external factors," in *Proc. 20th IEEE Int. Conf. Mobile Data Manag. (MDM)*, Jun. 2019, pp. 234–242.
- [61] S. Fang, Q. Zhang, G. Meng, S. Xiang, and C. Pan, "GSTNet: Global spatial-temporal network for traffic flow prediction," in *Proc. IJCAI*, Aug. 2019, pp. 2286–2293.
- [62] Z. Wu, S. Pan, G. Long, J. Jiang, and C. Zhang, "Graph wavenet for deep spatial-temporal graph modeling," May 2019, *arXiv:1906.00121*.
- [63] W. L. Hamilton, R. Ying, and J. Leskovec, "Representation learning on graphs: Methods and applications," Sep. 2017, *arXiv:1709.05584*.
- [64] T. N. Kipf and M. Welling, "Variational graph auto-encoders," Nov. 2016, *arXiv:1611.07308*.
- [65] S. Pan, R. Hu, G. Long, J. Jiang, L. Yao, and C. Zhang, "Adversarially regularized graph autoencoder for graph embedding," Feb. 2018, *arXiv:1802.04407*.
- [66] J. J. Choong, X. Liu, and T. Murata, "Optimizing variational graph autoencoder for community detection," in *Proc. IEEE Int. Conf. Big Data (Big Data)*, Dec. 2019, pp. 5353–5358.
- [67] T. H. Do, D. M. Nguyen, and N. Deligiannis, "Graph auto-encoder for graph signal denoising," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, May 2020, pp. 3322–3326.
- [68] S. Fan, X. Wang, C. Shi, E. Lu, K. Lin, and B. Wang, "One2multi graph autoencoder for multi-view graph clustering," in *Proc. Web Conf.*, Apr. 2020, pp. 3070–3076.
- [69] I. Li, A. Fabbri, S. Hingmire, and D. Radev, "R-VGAE: Relational-variational graph autoencoder for unsupervised prerequisite chain learning," in *Proc. 28th Int. Conf. Comput. Lang.*, Dec. 2020, pp. 1147–1157.
- [70] M. Svetlik, M. Leonetti, J. Sinapov, R. Shah, N. Walker, and P. Stone, "Automatic curriculum graph generation for reinforcement learning agents," *Proc. AAAI Conf. Artif. Intell.*, vol. 31, no. 1, Jan. 2017, pp. 2590–2596.
- [71] J. You, B. Liu, R. Ying, V. Pande, and J. Leskovec, "Graph convolutional policy network for goal-directed molecular graph generation," in *Proc. Int. Conf. Neural Inf. Process. Syst. (NIPS)*, Jun. 2018, pp. 6412–6422.
- [72] J. B. Lee, R. Rossi, and X. Kong, "Graph classification using structural attention," in *Proc. ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining (KDD)*, Jul. 2018, pp. 1666–1674.
- [73] W. Xiong, T. Hoang, and W. Y. Wang, "DeepPath: A reinforcement learning method for knowledge graph reasoning," in *Proc. Conf. Empirical Methods Nat. Lang. Process. (EMNLP)*, Jul. 2017, pp. 564–573.
- [74] Q. Wang, Y. Hao, and J. Cao, "ADRL: An attention-based deep reinforcement learning framework for knowledge graph reasoning," *Knowl. Based Syst.*, vol. 197, no. 6, Jun. 2020, Art. no. 105910.
- [75] G. Wan, S. Pan, C. Gong, C. Zhou, G. Haffari, "Reasoning like human: Hierarchical reinforcement learning for knowledge graph reasoning," in *Proc. IJCAI*, Jul. 2020, pp. 1926–1932.
- [76] P. Tiwari, H. Zhu, and H. Pandey, "DAPath: distance-aware knowledge graph reasoning based on deep reinforcement learning," *Neural Netw.*, vol. 135, no. 3, pp. 1–12, Mar. 2021.
- [77] S. Li, H. Wang, R. Pan, and M. Mao, "MemoryPath: A deep reinforcement learning framework for incorporating memory component into knowledge graph reasoning," *Neurocomputing*, vol. 419, no. 5, pp. 273–286, Jan. 2021.
- [78] L. Chen, J. Cui, X. Tang, Y. Qian, Y. Li, and Y. Zhang, "RLPath: A knowledge graph link prediction method using reinforcement learning based attentive relation path searching and representation learning," *Appl. Intell.*, to be published.
- [79] Y. Shen, Y. Shi, J. Zhang, and K. B. Letaief, "Graph neural networks for scalable radio resource management: Architecture design and theoretical analysis," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 1, pp. 101–115, Nov. 2020.
- [80] M. Eisen and A. Ribeiro, "Large scale wireless power allocation with graph neural networks," in *Proc. IEEE Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Jul. 2019, pp. 1–5.
- [81] M. Eisen and A. Ribeiro, "Optimal wireless resource allocation with random edge graph neural networks," *IEEE Trans. Signal Process.*, vol. 68, no. 4, pp. 2977–2991, Apr. 2020.
- [82] M. Eisen and A. Ribeiro, "Transferable policies for large scale wireless networks with graph neural networks," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, May 2020, pp. 5040–5044.
- [83] I. Nikoloska and O. Simeone, "Fast power control adaptation via meta-learning for random edge graph neural networks," May 2021, *arXiv:2105.00459*.
- [84] N. Naderializadeh, M. Eisen, and A. Ribeiro, "Wireless power control via counterfactual optimization of graph neural networks," in *Proc. IEEE Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Feb. 2020, pp. 1–5.
- [85] J. Guo and C. Yang, "Learning power control for cellular systems with heterogeneous graph neural network," in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, May 2021, pp. 1–6.

- [86] Z. Gao, M. Eisen, and A. Ribeiro, "Resource allocation via graph neural networks in free space optical fronthaul networks," Jun. 2020, *arXiv:2006.15005*.
- [87] Z. Wang, M. Eisen, and A. Ribeiro, "Unsupervised learning for asynchronous resource allocation in ad-hoc wireless networks," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, Nov. 2021, pp. 8143–8147.
- [88] Z. Wang, M. Eisen, and A. Ribeiro, "Learning decentralized wireless resource allocations with graph neural networks," Jul. 2021, *arXiv:2107.01489*.
- [89] T. Chen, X. Zhang, M. You, G. Zheng, and S. Lambotharan, "A GNN based supervised learning framework for resource allocation in wireless IoT networks," *IEEE Internet Things J.*, early access, Jun. 22, 2021, doi: [10.1109/JIOT.2021.3091551](https://doi.org/10.1109/JIOT.2021.3091551).
- [90] A. Chowdhury, G. Verma, C. Rao, A. Swami, and S. Segarra, "Efficient power allocation using graph neural networks and deep algorithm unfolding," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process (ICASSP)*, May 2021, pp. 4725–4729.
- [91] A. Chowdhury, G. Verma, C. Rao, A. Swami, and S. Segarra, "Unfolding WMMSE using graph neural networks for efficient power allocation," *IEEE Trans. Wireless Commun.*, vol. 20, no. 9, pp. 6004–6017, Apr. 2021.
- [92] M. Lee, G. Yu, and G. Y. Li, "Graph embedding based wireless link scheduling with few training samples," *IEEE Trans. Wireless Commun.*, vol. 20, no. 4, pp. 2282–2294, Dec. 2021.
- [93] Z. Zhao, G. Verma, C. Rao, A. Swami, and S. Segarra, "Distributed scheduling using graph neural networks," in *Proc. IEEE Inter. Conf. Acoust. Speech Signal Process. (ICASSP)*, Nov. 2021, pp. 4720–4724.
- [94] K. Lei, M. Qin, B. Bai, G. Zhang, and M. Yang "GCN-GAN: A non-linear temporal link prediction model for weighted dynamic networks," in *Proc. IEEE Int. Conf. Comput. Commun. (INFOCOM)*, Jan. 2019, pp. 388–396.
- [95] S. He, S. Xiong, W. Zhang, Y. Yang, J. Ren, and Y. Huang, "GBLinks: GNN-based beam selection and link activation for ultra-dense D2D mmWave networks," Jul. 2021, *arXiv:2107.02412*.
- [96] K. Nakashima, S. Kamiya, K. Ohtsu, K. Yamamoto, T. Nishio, and M. Morikura, "Deep reinforcement learning-based channel allocation for wireless LANs with graph convolutional networks," *IEEE Access*, vol. 8, pp. 31823–31834, 2020.
- [97] Z. He, L. Wang, H. Ye, G. Y. Li, and B.-H. F. Juang, "Resource allocation based on graph neural networks in vehicular communications," in *Proc. IEEE Global Commun. Conf. (GlobeCom)*, Dec. 2020, pp. 1–5.
- [98] V. Ranasinghe, N. Rajatheva, and M. Latva-Aho, "Graph neural network based access point selection for cell-free massive MIMO systems," Jul. 2021, *arXiv:2107.02884*.
- [99] T. Jiang, H. V. Cheng, and W. Yu, "Learning to reflect and to beamform for intelligent reflecting surface with implicit channel estimation," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 7, pp. 1931–1945, May 2021.
- [100] K. Tekbiyik, G. K. Kurt, C. Huang, A. R. Ekti, and H. Yanikomeroglu, "Channel estimation for full-duplex RIS-assisted HAPS backhauling with graph attention networks," Oct. 2020, *arXiv:2010.12004*.
- [101] A. Scotti, N. N. Moghadam, D. Liu, K. Gafvert, and J. Huang, "Graph neural networks for massive MIMO detection," Jul. 2020, *arXiv:2007.05703*.
- [102] Y. Yang, S. Zhang, F. Gao, J. Ma, and O. A. Dobre, "Graph neural network-based channel tracking for massive MIMO networks," *IEEE Commun. Lett.*, vol. 24, no. 8, pp. 1747–1751, Aug. 2020.
- [103] S. Zhao *et al.*, "Cellular network traffic prediction incorporating handover: A graph convolutional approach," in *Proc. Annu. IEEE Int. Conf. Sens. Commun. Netw. (SECON)*, Jun. 2020, pp. 1–9.
- [104] L. Yang, X. Gu, and H. Shi, "A novel satellite network traffic prediction method based on GCN-GRU," in *Proc. Int. Conf. Wireless Commun. Signal Process. (IC-WCSP)*, Oct. 2020, pp. 718–723.
- [105] J. Dong, S. Chen, P. Y. J. Ha, Y. Li, and S. Labi, "A DRL-based multiagent cooperative control framework for CAV networks: A graphic convolution Q network," Oct. 2020, *arXiv:2010.05437*.
- [106] P. Y. J. Ha, S. Chen, J. Dong, R. Du, Y. Li, and S. Labi, "Leveraging the capabilities of connected and autonomous vehicles and multi-agent reinforcement learning to mitigate highway bottleneck congestion," Oct. 2020, *arXiv:2010.05436*.
- [107] J. Shao, H. Zhang, Y. Mao, and J. Zhang, "Branchy-GNN: A device-edge co-inference framework for efficient point cloud processing," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, May 2021, pp. 8488–8492.
- [108] T. Fujihashi, T. Koike-Akino, S. Chen, and T. Watanabe, "Wireless 3D point cloud delivery using deep graph neural networks," Jun. 2020, *arXiv:2006.09835*.
- [109] C.-S. Jeong, J.-S. Kim, D.-K. Kim, S.-C. Kwon, and K.-D. Jung, "AR anchor system using mobile based 3D GNN detection," *J. Broadcast. Electron. Media*, vol. 13, no. 1, pp. 54–60, Feb. 2021.
- [110] S. Lohani and R. T. Glasser, "Robust free space oam communications with unsupervised machine learning," in *Proc. Frontiers Opt. (FIO)*, Sep. 2019, Art. no. FTU5B.3.
- [111] W. Yan, D. Jin, Z. Lin, and F. Yin, "Graph neural network for large-scale network localization," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process. (ICASSP)*, Oct. 2021, pp. 5250–5254.
- [112] P. Soto *et al.*, "ATARI: A graph convolutional neural network approach for performance prediction in next-generation WLANs," *Sensors*, vol. 21, no. 13, p. 4321, May 2021.
- [113] M. Liu, J. Li, and H. Lu, "Routing in small satellite networks: A GNN-based learning approach," Aug. 2021, *arXiv:2108.08523*.
- [114] M. Lee, G. Yu, and H. Dai, "Decentralized inference with graph neural networks in wireless communication systems," Apr. 2021, *arXiv:2104.09027*.
- [115] S. Ravanbakhsh, J. Schneider, and B. Póczos, "Equivariance through parameter-sharing," in *Proc. Int. Conf. Mach. Learn. (ICML)*, Feb. 2017, pp. 2892–2901.
- [116] Y. N. Dauphin, A. Fan, M. Auli, and D. Grangier, "Language modeling with gated convolutional networks," in *Proc. Int. Conf. Mach. Learn. (ICML)*, Dec. 2017, pp. 933–941.



SHIWEN HE (Member, IEEE) received the M.S. degree in computational mathematics from the Chengdu University of Technology, Chengdu, China, in 2009, and the Ph.D. degree in information and communication engineering from Southeast University, Nanjing, China, in 2013.

Since April 2018, he has been a Professor with the School of Computer Science and Engineering, Central South University, Changsha, China. From November 2015 to March 2018, he has been an Associate Research Fellow with the School of Information Science and Engineering, Southeast University, where he was a Postdoctoral Researcher with the State Key Laboratory of Millimeter Waves, Department of Radio Engineering, from September 2013 to October 2015. He has authored or coauthored more than 125 technical publications and held 105 invention patents granted. His research interests include wireless communications and networking, distributed learning and optimization computation theory, and big data analytics.

Prof. He was the recipient of the Contribution Award for the development of the IEEE 802.11aj Standard published on April 18, 2018. He has been a Technical Associate Editor for the IEEE 802.11aj from December 2015 to July 2018 and submitted more than 20 technical proposals to IEEE standards. He has served as an Associate Editor for IEEE SIGNAL PROCESSING LETTERS and *EURASIP Journal on Wireless Communications* and has been a TPC Member of various conferences, including Globecom, ICC, ICDC, and WCSP.



SHAOWEN XIONG received the B.S. degree in information and computing sciences from the Hunan University of Technology, Zhuzhou, China, in 2019. He is currently pursuing the M.S. degree in computer science and technology with Central South University, Changsha, China. His current research interests include graph neural networks and its application on wireless communication networks.



YEYU OU received the B.S. degree in engineering mechanics from the Hehai university, Nanjing, China, in 2020. She is currently pursuing the M.S. degree in computer technology with the Central South University, Changsha, China. Her current research interests include graph neural networks and big data analysis.



JIAN ZHANG (Member, IEEE) received the B.Eng. degree in computer science from the National University of Defense Technology in 1998, and the M.Eng. and Ph.D. degrees in computer science from Central South University in 2002 and 2007, respectively.

He is currently an Associate Professor with the School of Computer Science and Engineering, Central South University. He has published over 30 peer-viewed journal papers, and is holding 12 granted and filed patents. His research interests

include optimization theory, cyberspace security, cloud computing, and cognitive radio technology.



JIAHENG WANG (Senior Member, IEEE) received the Ph.D. degree in electronic and computer engineering from the Hong Kong University of Science and Technology, Hong Kong, in 2010, and the B.E. and M.S. degrees from the Southeast University, Nanjing, China, in 2001 and 2006, respectively.

He is currently a Full Professor with the National Mobile Communications Research Laboratory, School of Information Science and Engineering, Southeast University, Nanjing. From

2010 to 2011, he was with the Signal Processing Laboratory, KTH Royal Institute of Technology, Stockholm, Sweden. He also held visiting positions with the Friedrich–Alexander University of Erlangen–Nurnberg, Germany, and the University of Macau, Macau, China. He has published more than 100 articles on international journals and conferences. His research interests are mainly on wireless communications and networks.

Prof. Wang is a recipient of the Humboldt Fellowship for Experienced Researchers. He served as an Associate Editor for the IEEE SIGNAL PROCESSING LETTERS from 2014 to 2018. Since 2018, he has been serving as a Senior Area Editor for the IEEE SIGNAL PROCESSING LETTERS.



YONGMING HUANG (Senior Member, IEEE) received the B.S. and M.S. degrees from Nanjing University, China, in 2000 and 2003, respectively, and the Ph.D. degree in electrical engineering from Southeast University, China, in 2007.

Since March 2007, he has been a Faculty with the National Mobile Communication Research Laboratory, School of Information Science and Engineering, Southeast University, China, where he is currently a Full Professor. From 2008 to 2009, he was visiting the Signal Processing Lab,

Electrical Engineering, Royal Institute of Technology, Stockholm, Sweden. He has published over 200 peer-reviewed papers, hold over 60 invention patents. His current research interests include MIMO wireless communications, cooperative wireless communications, and millimeter wave wireless communications. He submitted around 20 technical contributions to IEEE standards, and was awarded a certificate of appreciation for outstanding contribution to the development of IEEE standard 802.11aj. He has served as an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING, and a Guest Editor for the IEEE JOURNAL SELECTED AREAS IN COMMUNICATIONS. He is serving as an Editor-at-Large for the IEEE OPEN JOURNAL OF THE COMMUNICATIONS SOCIETY, and an Associate Editor for the IEEE WIRELESS COMMUNICATIONS LETTERS.



YAOXUE ZHANG (Senior Member, IEEE) received the B.Sc. degree from the Northwest Institute of Telecommunication Engineering, China, in 1982, and the Ph.D. degree in computer networking from Tohoku University, Japan, in 1989.

He is currently a Professor with the School of Computer Science and Engineering, Central South University, China, and also a Professor with the Department of Computer Science and Technology, Tsinghua University, China. He has published over 200 technical papers in international journals and conferences, as well as nine monographs and text-books. His research interests include computer networking, operating systems, ubiquitous/pervasive computing, transparent computing, and big data. He is currently serving as the Editor-in-Chief of *Chinese Journal of Electronics*. He is a Fellow of the Chinese Academy of Engineering.

He is currently serving as the Editor-in-Chief of *Chinese Journal of Electronics*. He is a Fellow of the Chinese Academy of Engineering.