Deep Learning Based Radio Resource Management in NOMA Networks: User Association, Subchannel and Power Allocation

Haijun Zhang®, Senior Member, IEEE, Haisen Zhang, Keping Long, Senior Member, IEEE, and George K. Karagiannidis, Fellow, IEEE

Abstract—With the rapid development of future wireless communication, the combination of NOMA technology and millimeter-wave (mmWave) technology has become a research hotspot. The application of NOMA in mmWave heterogeneous networks can meet the diverse needs of users in different applications and scenarios in future communications. In this paper, we propose a machine learning framework to deal with the user association, subchannel and power allocation problems in such a complex scenario. We focus on maximizing the energy efficiency (EE) of the system under the constraints of quality of service (QoS), interference limitation, and power limitation. Specifically, user association is solved through the Lagrange dual decomposition method, while semi-supervised learning and deep neural network (DNN) are used for the subchannel and power allocation, respectively. In particular, unlabeled samples are introduced to improve approximation and generalization ability for subchannel allocation. The simulation indicates that the proposed scheme can achieve higher EE with lower complexity.

Index Terms—Machine learning, resource management, semi-supervised learning, energy efficiency.

I. INTRODUCTION

WITH the explosively growth of users’ service requirements, the use of the limited radio resources to hold more wireless service of next generation wireless networks becomes a challenging problem. Non-orthogonal multiple access (NOMA) is a novel access technique for future communication system, which has been widely concerned because of its high spectral efficiency. The application of NOMA in heterogeneous networks can meet the diverse needs of users in future communications. Resource management in NOMA networks is becoming more and more intriguing to enhance energy efficiency (EE) of systems. Existing research studied several resource management schemes for NOMA networks. Most of the existing method requires high computation and may not be not necessarily feasible in practice.

In existing research, dual decomposition method was widely used to user association in heterogeneous networks [1], [2]. In [3], an optimization scheme for user association, based on graph theory, was proposed. Furthermore, a theoretical mean proportionally fair utility was suggested in [4] to address the problem of user association. In the subchannel allocation problem of NOMA network, existing works often used matching theory. Especially, a two side many-to-many matching algorithm was proposed in [5]–[8], static and dynamic source-destination matching algorithm was proposed in [9]. For the problem of power optimization in NOMA heterogeneous network, there are several power optimization algorithms proposed, such as water-filling algorithm [10], Lagrange dual multiplier method [11], alternating direction method of multipliers (ADMM) algorithm [12], Lyapunov Optimization algorithm [13], [14]. However, these methods are often based on high complexity and need many iterations to converge for BS processing.

As the most promising technology in artificial intelligence, machine learning (ML) could be used to address wireless resource management in wireless network [15]. In this case, the challenge is how to enable ML to assist wireless devices in intelligent learning and decision-making, so as to meet the diverse service demands of the future wireless networks [16], [17].

Recently, as the representative technology of ML, supervised learning has been used to wireless communications. Full-connected deep neural network (DNN) is the most classical model of supervised learning. Its non-linear approximation performance enables DNN scheme to solve many problems in resource management, for instance, beamforming, subchannel allocation and power control. The work in [18] was the classic application of DNN scheme in resource management, which used deep learning to approximate the transmit power policy, while it was shown the universal approximation for power optimization. In [19], a deep learning algorithm in cognitive

Manuscript received November 15, 2019; revised May 20, 2020; accepted June 12, 2020. Date of publication June 23, 2020; date of current version December 30, 2020. This work was supported in part by the National Key R&D Program of China under Grant 2019YFB1803304, in part by the National Natural Science Foundation of China under Grants 61822104 and 61771044, in part by the Beijing Natural Science Foundation under Grants L172025 and L172049, in part by the Fundamental Research Funds for the Central Universities under Grants FRF-TP-19-002C1 and RC1631, in part by Beijing Top Discipline for Artificial Intelligent Science and Engineering, and in part by the University of Science and Technology Beijing. Recommended for acceptance by Prof. Ekram Hossain.

(Corresponding authors: Haijun Zhang and Keping Long.)

Haijun Zhang, Haisen Zhang, and Keping Long are with the Institute of Artificial Intelligence, Beijing Advanced Innovation Center for Materials Genome Engineering, Beijing Engineering and Technology Research Center for Convergence Networks and Ubiquitous Services, University of Science and Technology Beijing, Beijing 100083, China (e-mail: haijunzhang@ieee.org; z-haisen@qq.com; longkeping@ustb.edu.cn).

George K. Karagiannidis is with the Aristotle University of Thessaloniki, Thessaloniki 54124, Greece (e-mail: geokarag@ieee.org).

Digital Object Identifier 10.1109/TNSE.2020.3004333
A deep learning-based framework is proposed, which can solve team decision problems of power control and provides a basis for the data parallel distributed neural network in resource management. In [21], a convolutional neural network was proposed for power control, which can achieve higher EE and spectrum efficiency in a much lower computation time, due to sharing the convolution kernel parameters and the sparsity of inter-layer connections in hidden layers.

As shown in the previous study [18]–[21], DNN can be used without explicitly solving the problem of complex optimal control strategy of communication system. As an intelligent resource management tool, deep learning is used to solve various problems, such as user connection, wireless access selection, frequency allocation, power control and intelligent beamforming. In addition, considering that the training DNN model with medium number of layers can be used in lower calculation time [22], it is very suitable for real-time operation. Compared with the conventional distributed optimization technology, the resource management algorithm based on deep learning can understand the wireless network state and network user state in real time, so as to adjust the resource management scheme in real time. This kind of intelligent decision-making is very important for most Internet of things and beyond 5th generation (B5G) services, especially those that need real-time, low latency operation, such as automatic driving and UAV control [23]–[26].

Nowadays, there are some literatures on DL-based resource management in noma network. The research of [27] used deep recurrent neural network to solve the resource allocation for the NOMA heterogeneous IoT network. However, the authors did not consider the influence of high power consumption on EE. The work in [28] proposed the power allocation scheme for NOMA system via exhaustive search method and DNN model. However, it is unrealistic to obtain training sets by exhaustive methods when the network environment becomes more complex. The research of [29] proposed a deep reinforcement learning framework to allocate channel in a near optimal way. However, the size of state space will limit the performance of the algorithm in this work.

In our work, we study the problem of wireless resource management aiming at maximizing EE considering power consumption, and the data set obtained by gradient iteration algorithm has advantages in complexity.

The effect of deep learning algorithm largely depends on the label samples. This will inevitably be affected by the process of existing algorithms. In resource management based on ML, it is simple to obtain massive unlabeled samples, but in this case we need to use expensive and very long computation to get the output of each sample. Therefore, we try to improve the learning performance and reduce the dependence on the reference algorithm. Thus, semi-supervised learning is introduced to study this optimization problem, that is, training through a finite number of labeled samples generated by numerical iterative algorithms, and these labeled samples are also combined with massive unlabeled samples. Semi-supervised learning can avoid the waste of data and resources and solve the problem of weak generalization ability of supervised learning model.

Co-training is proposed in [30] as a methods of semi-supervised learning, which initializes more than two learners, based on clustering or popular hypothesis. In the process of learning, the labeled data with the highest confidence after labeling is selected, and the labeled data is put into each other after labeling in order to update the model.

According to the author’s knowledge, the issue of deep learning based user association, subchannel and power allocation has not been yet well investigated in NOMA networks. The main contributions of this paper can be summarized as follow:

- A deep learning-based framework is proposed, which is mainly used to deal with energy-efficient user association, subchannel and power allocation in NOMA mmWave heterogeneous networks.
- Under the premise of guaranteeing quality of service (QoS), interference limitation and power limitation, we concentrate on maximizing the EE of the network. Specifically, sample data are generated by iterative algorithm, and ML schemes are adopted in the decision-making stage.
- The Lagrange dual decomposition method based user association scheme is proposed, while semi-supervised learning and DNN are used to address the subchannel allocation and power control in NOMA heterogeneous networks.
- Extensive simulations reveal that by using unlabeled data, the proposed method can optimize EE and reduce the dependence on the reference algorithm.

The rest of the paper is organized as follows. The system model is provided in Section II. Section II provides the system model. The deep learning for wireless resource management is proposed in Section III, a large number of simulation experiments were carried out in Section IV, and the performance of the proposed algorithm was evaluated. Section V summarizes the paper.

II. SYSTEM MODEL AND PROBLEM FORMULATION

A. System Model

In this paper, a NOMA based mmWave heterogeneous wireless networks with an intelligent control center (Fig. 1) is studied, where small cells are distributed uniformly in a macrocell. We concentrate on issues of downlink user association, subchannel allocation and power control. Denote $B = \{1, 2, \ldots, b, \ldots, B\}$ as the set of all BSs including macro BS and small BSs, $\mathcal{N} = \{1, 2, \ldots, N\}$ is the set of all subchannel, $\mathcal{M} = \{1, 2, \ldots, m, \ldots, M\}$ is the set of users, and $X = \{x_{1,1}, x_{1,2}, \ldots, x_{b,m}\}$ is used to represent the user association strategy between BS $b$ and user $m$. If user $m$ is assigned to the BS $b$, then, $x_{b,m} = 1$. Otherwise, $x_{b,m} = 0$.

It is assumed that a subchannel in NOMA can be utilized by up to two users through successive interference cancelation (SIC). The set $S = \{s_{1,1}^1, s_{1,2}^1, \ldots, s_{b,m}^n\}$ stands for the subchannel allocation strategy between user $m$ and subchannel $n$.
of BS $b$. If the subchannel $n$ of BS $b$ is occupied by the user $m$, then, $s_{b,m}^n = 1$. Otherwise, $s_{b,m}^n = 0$.

The signal to interference plus noise ratio (SINR) of user $m$ on subchannel $n$ of BS $b$ can be given as

$$
\gamma_{b,m}^n = \frac{s_{b,m}^n p_{b,m}^n g_{b,m}^n}{g_{b,m}^n \sum_{r=m+1}^M s_{b,r}^n p_{b,r}^n + \sum_{j=1,j\neq b}^B \sum_{r=1}^M s_{j,r}^n p_{j,r}^n g_{j,m}^n + \sigma^2},
$$

where $p_{b,m}^n$ is the power of user $m$ on subchannel $n$ of BS $b$, and $g_{b,m}^n$ is the gain of user $m$ on subchannel $n$ of BS $b$ and $\sigma^2$ is the variance of AWGN. The capacity of user $m$ on subchannel $n$ of BS $b$ can be expressed as

$$
e_{b,m}^n = \frac{W}{\log_2 (1 + \gamma_{b,m}^n)}, \forall b \in \mathcal{B},
$$

where $W$ is the system bandwidth.

The total rate of the network is

$$R(X, S, P) = \sum_{b=1}^B \sum_{m=1}^M \sum_{n=1}^N x_{b,m} e_{b,m}^n,
$$

and the total transmit power is

$$U(X, S, P) = \sum_{b=1}^B \sum_{m=1}^M \sum_{n=1}^N x_{b,m} s_{b,m} p_{b,m} + \sum_{b=1}^B p_{c,b},
$$

where $p_{c,b}$ denotes the circuit power of BS $b$.

The total EE of the system can be formulated as

$$EE = \frac{R(X, S, P)}{U(X, S, P)}.
$$

### B. Problem Formulation

It's assumed that each user can be serviced only by one BS at time. Thus,

$$\sum_{b \in \mathcal{B}} x_{b,m} = 1, \forall m \in \mathcal{M}.
$$

This work also assumes that at most two users can occupy one subchannel of each BS. Therefore hold that

$$\sum_{m \in \mathcal{M}} s_{b,m}^n \leq 2, \forall n \in \mathcal{N}, b \in \mathcal{B},
$$

with the constraint:

$$\sum_{m \in \mathcal{M}} x_{b,m} \leq K_b, \forall b \in \mathcal{B},
$$

where $K_b$ is the the maximum number of users assigned to BS $b$. The total power constraint is:

$$\sum_{m=1}^M \sum_{n=1}^N x_{b,m} s_{b,m} p_{b,m} \leq p_{\text{max}}, \forall b \in \mathcal{B},
$$

where $p_{\text{max}}$ is the maximum transmit power of BS $b$, and the QoS constraint is:

$$\sum_{j=1,j \neq b}^B \sum_{n=1}^N s_{j,r}^n p_{j,r}^n g_{j,m}^n \leq I_b, \forall b \in \mathcal{B},
$$

which ensures that each user has the minimum QoS requirement of $R_t$. Finally, the cross-tier interference constraint can be written as:

$$\sum_{m=1}^M \sum_{n=1}^N x_{b,m} s_{b,m} c_{b,m}^n \geq R_t, \forall m \in \mathcal{M},
$$

where $I_b$ denotes the maximum interference constraint.

Aiming to achieve load-balance and maximum EE, the problem of user association, subchannel and power optimization can be formulated as:

$$\max_{(X,S,P)} \sum_{b=1}^B \sum_{m=1}^M \sum_{n=1}^N x_{b,m} \log \left( \frac{W}{\sum_{m \in \mathcal{M}} x_{b,m} \log_2 (1 + \gamma_{b,m}^n)} \right)
$$

s.t.  

C1 : $\sum_{b \in \mathcal{B}} x_{b,m} = 1, \forall m \in \mathcal{M},$

C2 : $\sum_{m \in \mathcal{M}} s_{b,m}^n \leq 2, \forall n \in \mathcal{N}, b \in \mathcal{B},$

C3 : $\sum_{m \in \mathcal{M}} x_{b,m} = K_b, \forall b \in \mathcal{B},$

C4 : $\sum_{m=1}^M \sum_{n=1}^N x_{b,m} s_{b,m} p_{b,m} \leq p_{\text{max}}, \forall b \in \mathcal{B},$

C5 : $\sum_{m=1}^M \sum_{n=1}^N x_{b,m} s_{b,m} c_{b,m}^n \geq R_t, \forall m \in \mathcal{M},$

C6 : $\sum_{j=1,j \neq b}^B \sum_{n=1}^N s_{j,r}^n p_{j,r}^n g_{j,m}^n \leq I_b, \forall b \in \mathcal{B}.$
C4 is the maximum power constraint on each BS, C5 is the QoS constraint on each user and C6 is the cross-tier interference constraint on each BS.

III. MACHINE LEARNING FOR WIRELESS RESOURCE MANAGEMENT

A. User Association

In this subsection, an user association scheme based on the Lagrange dual decomposition method is proposed.

The Lagrangian dual function for user association can be written as

\[
\begin{align*}
\mathcal{D} (\mu, \lambda, \nu, \tau) &= \max_{X,S,P} L(X,S,P) \quad \text{s.t.} \quad \mu \geq 0, \lambda \geq 0, \nu \geq 0, \tau \geq 0, \\
&= \sum_{b \in B} \log (\frac{1 + \gamma^u_{h,b}}{\lambda_b(t) s^u_{h,b,m} n^u_{h,b,m} - \nu_c(t) \sum_{n=1}^N \sum_{b \in B, b \neq j} g^u_{h,b,m} p^u_{b,m}(t) - \tau_b(t) \sum_{n=1}^N \sum_{b \in B, b \neq j} g^u_{h,b,m} p^u_{b,m}(t)}) \\
&= \sum_{b \in B} \log (\frac{1 + \gamma^u_{h,b}}{\lambda_b(t) s^u_{h,b,m} n^u_{h,b,m} - \nu_c(t) \sum_{n=1}^N \sum_{b \in B, b \neq j} g^u_{h,b,m} p^u_{b,m}(t) - \tau_b(t) \sum_{n=1}^N \sum_{b \in B, b \neq j} g^u_{h,b,m} p^u_{b,m}(t)}) \\
&= \sum_{b \in B} \log (\frac{1 + \gamma^u_{h,b}}{\lambda_b(t) s^u_{h,b,m} n^u_{h,b,m} - \nu_c(t) \sum_{n=1}^N \sum_{b \in B, b \neq j} g^u_{h,b,m} p^u_{b,m}(t) - \tau_b(t) \sum_{n=1}^N \sum_{b \in B, b \neq j} g^u_{h,b,m} p^u_{b,m}(t)}) .
\end{align*}
\]

The decision reference value of user association of iteration \( t \) is given by

\[
\begin{align*}
j_{b,m} &= \log \left( \frac{1 + \gamma^u_{h,b}}{\lambda_b(t) s^u_{h,b,m} n^u_{h,b,m} - \nu_c(t) \sum_{n=1}^N \sum_{b \in B, b \neq j} g^u_{h,b,m} p^u_{b,m}(t) - \tau_b(t) \sum_{n=1}^N \sum_{b \in B, b \neq j} g^u_{h,b,m} p^u_{b,m}(t)}{\lambda_b(t) s^u_{h,b,m} n^u_{h,b,m}} \right) \\
&= \frac{\log (1 + \gamma^u_{h,b}) - \lambda_b(t) s^u_{h,b,m} n^u_{h,b,m}}{\lambda_b(t) s^u_{h,b,m} n^u_{h,b,m}} .
\end{align*}
\]

B. Subchannel Allocation

In this subsection, a novel co-training semi-supervised learning method for subchannel allocation is proposed. In order to solve the channel allocation problem, we transform problem (12) into a loss function in the deep learning model, and get the optimal power allocation by minimizing the loss function. The loss function can be written as

\[
\begin{align*}
\hat{f}_{X,S,P}(\mu, \lambda, \nu, \tau) &= \log \left( \frac{1 + \gamma^u_{h,b}}{\lambda_b(t) s^u_{h,b,m} n^u_{h,b,m}} \right) - \lambda_b(t) s^u_{h,b,m} n^u_{h,b,m}(t) - \mu_b(t) + \nu_c(t) \sum_{n=1}^N \sum_{b \in B, b \neq j} g^u_{h,b,m} p^u_{b,m}(t) - \tau_b(t) \sum_{n=1}^N \sum_{b \in B, b \neq j} g^u_{h,b,m} p^u_{b,m}(t) \\
\end{align*}
\]

Algorithm 1: Two-side Matching Algorithm.

1: Initialize allocation strategy \( S \);
2: for \( s \) in \( B \) do
3: Initialize the set \( V_s(n) \) as users allocated to subchannel \( n \), and set \( V_s^c \) as users of unassigned channels;
4: while \( V_s^c \neq \emptyset \) do
5: for \( m \) in \( A \) do
6: if \( x_{b,m} = 1 \) then
7: Select the subchannel \( n^* \) with the best channel conditions.
8: if \( V_s(n^*) \neq 2 \) then
9: Let \( S_{b,m}^c = 1 \); 
10: Update \( V_s(n^*) \) and \( V_s^c \);
11: end if
12: if \( V_s(n^*) = 2 \) then
13: Calculate the utility function of any two of user \( m \), user \( m_1 \), and \( m_2 \) who occupy subchannel \( n^* \);
14: Select two users which maximize the EE of the subchannel \( n^* \), and update \( V_s(n^*) \) and \( V_s^c \);
15: end if
16: end if
17: end for
18: end while
19: end for

\[
\min_S \left| \hat{S} - \arg \max_S \sum_{b=1}^B \sum_{m=1}^M \sum_{n=1}^N \frac{R_n}{U} \right|^2 \\
\text{s.t.} \quad C2, C4 - C6
\]

where \( \hat{S} \) means the predicted subchannel allocation strategy.

In this paper, the subchannel allocation scheme is initially generated by two-sided matching scheme as shown in Algorithm 1. Users and subchannels are regarded as two players pursuing the maximization of their respective utility.

The proposed deep learning scheme for subchannel allocation is based on semi-supervised learning regression algorithm. In order to improve the accuracy of DNN, semi-supervised learning is used to improve the performance of NOMA heterogeneous network. Given a labeled data set \( L_s = \{(G_1, S_1), (G_2, S_2), \ldots, (G_L, S_L)\} \) generated by Algorithm 1 and an unlabeled data set \( U = \{G'_1, G'_2, \ldots, G'_U\} \), where \( G_i \) is matrix representing the i-th group of gains, which contains all \( g^u_{b,m} \) generated by the i-th initialization, \( S_i \) is the i-th group of channel allocation matrix, which contains all \( s^u_{h,b,m} \), \( |L| \) and \( |U| \) are the number of data samples \( L_s \) and \( U \), respectively. We aim to get a learner \( H_S : G \rightarrow S \) that can accurately predict the real label of an unlabeled input. The input of the learner is \( G \), and the output of the learner is \( S \).
D. J. Miller and H. S. Uyar analyzed the feasibility and rationality of using unlabeled data to improve learners performance based on data distribution estimation theory [31]. Assuming that all data follow a mixture of \( L \) Gaussian distributions, then the distribution is

\[
H_S(g) = \sum_{i=1}^{L} \alpha_i H_S(g|\theta_i),
\]

where \( \sum_{i=1}^{L} \alpha_i \) is a mixing coefficient and \( \theta = \{ \theta_i \} \) is the parameter. The random variable is a label that is determined by the selected mixture component \( m_i \) and feature vector \( g_i \) in the probability of \( P(c_i|g_i, m_i) \). According to the maximum posteriori probability hypothesis, the optimal classification formula can be written as

\[
H(g) = \arg \max_k \sum_j P(c_i = k|g_i, m_i = j) P(m_i = j|g_i),
\]

where \( P(m_i = j|g_i) = \frac{a_i H_S(g|\theta_i)}{\sum_{i=1}^{L} a_i H_S(g|\theta_i)} \).

According to the description above, the purpose of semi-supervised learning can be interpreted as using training data to estimate \( P(c_i = k|g_i, m_i = j) \), and \( P(m_i = j|g_i) \). Obviously, the former is related to labels while the latter is not related to labels. Therefore, if there is mass unlabeled data, the generalization ability of learners is improved. The analysis results of T. Zhang and F. J. Oles [32] showed that if a parameterized model can be decomposed into the form of \( P(x, y|\theta) = P(y|x, \theta) P(x|\theta) \), the simulated model parameters can be better supported by unlabeled data, thus improving the performance of the learning model.

Co-training semi-supervised learning algorithm is the most important semi-supervised learning method [30]. It initializes more than two learners based on clustering hypothesis or popular hypothesis. In the process of learning, the labeled data with the highest confidence after labeling is selected according to specific criteria. After labeling, the labeled data are put into each other’s labeled data set of the other’s learner to update the model. Co-training method is also based on the compatibility and complementarity of multi-view data [33]. Assuming that the data has two sufficient redundant and conditionally independent views. Sufficient means that each view contains enough information to produce the optimal learner, conditionally independent is that in given labels, each view is independent.

The co-training semi-supervised deep learning based subchannel allocation scheme is shown in Algorithm 2. The key of the co-training is to define the rule for selecting high confidence labeled data. To solve this problem, the criterion defined by Algorithm 2 is that the unlabeled data with the most consistent labeled training set is the data with the highest confidence. The rules for calculating confidence are as follows:

\[
\Delta_u = \frac{1}{|L|} \sum_{x_i \in L_S} (S_i - H(G_i))^2 - \frac{1}{|L|} \sum_{x_i \in L_S} (S_i - H'(G_i))^2,
\]

where \( H \) represents learning model, \( g_u \in U \), \( H' \) represents the retrained learning model using the original labeled data set adding labeled data generated by model \( H \).

According to Zhou’s research results [33], in order to start a co-training process with good effects, the two initialized learners must have relatively large differences. In extreme cases, if two learners are very close, the added labeled data will be very similar, and eventually it will evolve into a single learner, that is, self-training algorithm. When designing learners of Algorithm 2, we ensure the largest difference between learner \( H^1 \) and learner \( H^2 \). The differences of the models are mainly obtained by using different hidden layers, different network parameters or different number of nodes on the hidden layer.

Two network models are designed as 3 hidden layer network and 4 hidden layer network, the dropout coefficients are 0.8 and 1, batch-size are 200 and 500, respectively. We run the initial learners with different learning rates and choose the better learning rates of 0.01 and 0.05 respectively, as shown in Fig. 2. The Algorithm 2 runs under different number of hidden layer neurons from 600 to 1000 and from 60 to 10 in order to choose the optimal parameters, and the accuracy of the learner is evaluated as shown in the Fig. 3. By comparing different accuracy rates, the neurons number of hidden layer in the two networks is \( n_1 = 700 \) and \( n_2 = 80 \), respectively.

---

### Algorithm 2: Co-training Semi-Supervised Deep Learning based Subchannel Allocation.

**Input:** Labeled datasets of subchannel allocation \( L_S \), unlabeled dataset \( U \); maximum number of iterations \( T_{\text{max}} \); number of neurons in the hidden layers \( n_1, n_2 \);

1. Initialize and train two neural network models \( H^1 \) and \( H^2 \) using datasets \( L^1_S \) and \( L^2_S \) that are copied from dataset \( L_S \);
2. Randomly select an unlabeled datapool \( U' \) of size \( s \) from unlabeled dataset \( U \);
3. \textbf{repeat}
   4. \textbf{for} \( k = 1 \) and \( 2 \) \textbf{do}
      5. \textbf{for} each \( g_u \in U' \) \textbf{do}
         6. \( \tilde{S}_u \leftarrow H^k_u(G_u) \);
         7. Train network models \( H^k \) using datasets \( L^k_u \cup \{G_u, \tilde{S}_u\} \);
         8. \( \delta^k_{G_u} = \sum_{G \in L^k_u} (S_i - H^k(S_i))^2 - (S_i - H^k(G_i))^2 \);
      9. \textbf{end for}
     10. \textbf{if} there exists \( \delta^k_{G_u} > 0 \) \textbf{then}
         11. \( G^k_u \leftarrow \arg \max_{G \in L^k_u} \delta^k_{G_u} ; \tilde{S}_h = H^k(\tilde{S}_h) \);
         12. \( \omega_h = \{G^k_u, \tilde{S}_h\} ; U' = U' - \omega_h \);
     13. \textbf{else}
         14. \( \omega_h = \emptyset \);
     15. \textbf{end if}
   16. \( L^1_u \leftarrow L^1_u \cup \omega_1 ; L^2_u \leftarrow L^2_u \cup \omega_1 \);
   17. \textbf{if} one of \( L^1_u \) and \( L^2_u \) changes \textbf{then}
      18. Update neural network models \( H^1 \) and \( H^2 \);
      19. Replenish datapool \( U' \) to size \( s \) by randomly picking;
   20. \textbf{end if}
21. \textbf{until} Convergence or the number of iterations reaches \( T_{\text{max}} \).

**Output:** \( H^S(G) = \frac{1}{2}(H^1(G) + H^2(G)) \)
is the arg max $D$ is given by as the activation function, $K_j$ is shown as $+$ $+$ $n_0 = w$ and $s$ as the weight coefficient and $b_i$ as the bias of layer $j$. The output of the $i^{th}$ neural of layer $j$ is given by

$$y^j_i = \sum_{l=1}^{n_j} w^{j-1}_{ij} y^{j-1}_i + b^j_i.$$ (23)

Consider the loss function of power optimization problem, problem (13) can be rewritten as

$$\min_{\hat{P}} \left\| \hat{P} - \arg \max_P \sum_{b=1}^B \sum_{m=1}^M \sum_{n=1}^N P_{b,m} \right\|^2 \quad \text{s.t. C4 - C6}$$ (24)

where $\hat{P}$ shows maximization matrix of power optimization. We translate the maximization problem of non-deterministic polynomials into the problem of minimizing the error between the predicted value and the sample value, which is also directly expressed as the problem of the mean square error between the predicted value and the sample that represents the accuracy of the network model.

An iterative gradient algorithm is used to generate training data. Let $p_{b,m} = \sum_{n=1}^N s_{b,m}^n P_{b,m}^n$, and the power optimization function is given by:

$$f(p_{b,m}) = \sum_{b=1}^B \sum_{m=1}^M \sum_{n=1}^N x_{b,m} \ln \left( \log_2 \left( 1 + \gamma_{b,m}^n \right) \right) - \ln(K_b) + \sum_{b=1}^B \sum_{m=1}^M \sum_{n=1}^N p_{b,m} + \sum_{b=1}^B \sum_{m=1}^M \sum_{n=1}^N x_{b,m} s_{b,m} P_{b,m}^n + \sum_{b=1}^B \sum_{m=1}^M \sum_{n=1}^N x_{b,m} s_{b,m} P_{b,m}^n.$$ (25)

The complexity of user association algorithm.

In this subsection, a power optimization scheme based on DNN is proposed to maximize the EE. DNN scheme has been proved to have high approximation rate and computing performance for power control problem in [18]. In this scheme, the complex iteration process is replaced by the non-linear mapping. Its output operation is realized only by simple matrix multiplication and addition, which guarantees the real-time performance of the network. DNN is a multi-layer network model, which works mainly through the connection between neurons. Specifically, denote $\xi$ as the activation function, denote $n_j$ as the number of neural of layer $j$, denote $w^j_i$ as the weight coefficient and $b^j_i$ as the bias of layer $j$. The output of the $i^{th}$ neural of layer $j$ is given by

$$y^j_i = \sum_{l=1}^{n_j} w^{j-1}_{ij} y^{j-1}_i + b^j_i.$$ (23)

Consider the loss function of power optimization problem, problem (13) can be rewritten as

$$\min_{\hat{P}} \left\| \hat{P} - \arg \max_P \sum_{b=1}^B \sum_{m=1}^M \sum_{n=1}^N P_{b,m} \right\|^2 \quad \text{s.t. C4 - C6}$$ (24)

where $\hat{P}$ shows maximization matrix of power optimization. We translate the maximization problem of non-deterministic polynomials into the problem of minimizing the error between the predicted value and the sample value, which is also directly expressed as the problem of the mean square error between the predicted value and the sample that represents the accuracy of the network model.

An iterative gradient algorithm is used to generate training data. Let $p_{b,m} = \sum_{n=1}^N s_{b,m}^n P_{b,m}^n$, and the power optimization function is given by:

$$f(p_{b,m}) = \sum_{b=1}^B \sum_{m=1}^M \sum_{n=1}^N x_{b,m} \ln \left( \log_2 \left( 1 + \gamma_{b,m}^n \right) \right) - \ln(K_b) + \sum_{b=1}^B \sum_{m=1}^M \sum_{n=1}^N p_{b,m} + \sum_{b=1}^B \sum_{m=1}^M \sum_{n=1}^N x_{b,m} s_{b,m} P_{b,m}^n + \sum_{b=1}^B \sum_{m=1}^M \sum_{n=1}^N x_{b,m} s_{b,m} P_{b,m}^n.$$ (25)

The complexity of user association algorithm.

In this subsection, a power optimization scheme based on DNN is proposed to maximize the EE. DNN scheme has been proved to have high approximation rate and computing performance for power control problem in [18]. In this scheme, the complex iteration process is replaced by the non-linear mapping. Its output operation is realized only by simple matrix multiplication and addition, which guarantees the real-time performance of the network. DNN is a multi-layer network model, which works mainly through the connection between neurons. Specifically, denote $\xi$ as the activation function, denote $n_j$ as the number of neural of layer $j$, denote $w^j_i$ as the weight coefficient and $b^j_i$ as the bias of layer $j$. The output of the $i^{th}$ neural of layer $j$ is given by

$$y^j_i = \sum_{l=1}^{n_j} w^{j-1}_{ij} y^{j-1}_i + b^j_i.$$ (23)

Consider the loss function of power optimization problem, problem (13) can be rewritten as

$$\min_{\hat{P}} \left\| \hat{P} - \arg \max_P \sum_{b=1}^B \sum_{m=1}^M \sum_{n=1}^N P_{b,m} \right\|^2 \quad \text{s.t. C4 - C6}$$ (24)

where $\hat{P}$ shows maximization matrix of power optimization. We translate the maximization problem of non-deterministic polynomials into the problem of minimizing the error between the predicted value and the sample value, which is also directly expressed as the problem of the mean square error between the predicted value and the sample that represents the accuracy of the network model.

An iterative gradient algorithm is used to generate training data. Let $p_{b,m} = \sum_{n=1}^N s_{b,m}^n P_{b,m}^n$, and the power optimization function is given by:

$$f(p_{b,m}) = \sum_{b=1}^B \sum_{m=1}^M \sum_{n=1}^N x_{b,m} \ln \left( \log_2 \left( 1 + \gamma_{b,m}^n \right) \right) - \ln(K_b) + \sum_{b=1}^B \sum_{m=1}^M \sum_{n=1}^N p_{b,m} + \sum_{b=1}^B \sum_{m=1}^M \sum_{n=1}^N x_{b,m} s_{b,m} P_{b,m}^n + \sum_{b=1}^B \sum_{m=1}^M \sum_{n=1}^N x_{b,m} s_{b,m} P_{b,m}^n.$$ (25)

The complexity of user association algorithm.

In this subsection, the asymptotic time complexity of proposed radio resource management is discussed.

- The complexity of user association algorithm. The proposed scheme is based on the Lagrange dual decomposition method. The complexity performance of the Lagrange dual decomposition method for user association is discussed in [2]. Thus, user association needs $O(MB + 3B + M)$ operations at each iteration.

- The complexity of subchannel allocation algorithm.

The proposed deep learning scheme for subchannel allocation is based on semi-supervised learning regression algorithm, which contains two learner $H_1^2$ and $H_2^2$ of different number of hidden layers and nodes per layer. The time complexity of DNN can be represented by floating-point operations (FLOPs). For each layer of neural network, the number of FLOPs can be expressed as

$$\text{FLOPs} = 2I_O i,$$ (28)

where $I_i$ is the input dimension of the $i^{th}$ layer and $O_i$ is the output dimension of the $i^{th}$ layer. Therefore, for our semi-supervised learning scheme, the number of FLOPs is:
Algorithm 3: Deep Neural Networks Based Power Optimization Scheme.

Data Generation:
1: for $n = 1$ to the size of data set do
2: repeat
3: for $b$ in $B$ do
4: for $m$ in $M$ do
5: (1) Calculate $\Delta p_{b,m} = \frac{\partial f}{\partial p_{b,m}}$
6: (2) $p_{b,m}(t + 1) = p_{b,m}(t) + \delta(t) \Delta p_{b,m}$
7: end for
8: end for
9: until Convergence or $t = T_{\text{max}}$
10: end for

Training Stage:
1: Initialize 3 layers DNN structure with $n_j$ neurons in each layer, the weight $w$ and bias $b$.
2: for $m = 1$ to training-epochs do
3: for $n = 1$ to batch-size do
4: Update the weight $w$ and bias $b$;
The activation function: RELU;
The optimization algorithm: Adam algorithm.
5: end for
6: end for

Testing Stage:
1: Generate the testing dataset.
2: Pass testing dataset through the trained power optimization model.
3: Evaluate the performance through DNN model.

$$FLOPs = 2 \left( \sum_{i=1}^{3} I_i O_i + \sum_{i=1}^{4} I_i O_i \right)$$

$$= 2B \left( \frac{M/B(N + 1)n_1 + n_1^2 + M/B(N + 1)n_2 + 2n_2^2}{\text{DNNmodel1}} \right) \left( \frac{O}{\text{DNNmodel2}} \right).$$  \hspace{1cm} (29)

While the time complexity of two-side matching algorithm would be $O(BN!2^{M/B})$ discussed in [5]. So the complexity comparison between learning-based subchannel allocation and matching subchannel allocation is as follows:

$$O(2B(M/B(N + 1)(n_1 + n_2) + n_1^2 + n_2^2)) < O(BN!2^{M/B}).$$  \hspace{1cm} (30)

Therefore, the proposed learning model for subchannel allocation has a lower complexity than matching algorithm.

- The complexity of power allocation algorithm.

For our learning scheme for power allocation, the number of FLOPs is:

$$FLOPs = 2 \left( \sum_{i=1}^{3} I_i O_i \right) = 2B \left( \frac{M/B(N + 1)n_j + n_j^2}{\text{DNNmodel1}} \right) \left( \frac{O}{\text{DNNmodel2}} \right).$$  \hspace{1cm} (31)

The complexity performance of the iterative gradient algorithm for power allocation is discussed in [2]. Thus, data generation of Algorithm 4 needs $O(MB + B)$ operations at each iteration. We suppose that the iterative gradient algorithm requires $t_p$ iterations to converge.

So the complexity comparison between learning-based power allocation and iterative gradient algorithm is as follows:

$$O \left( 2B \left( \frac{M/B(N + 1)n_j + n_j^2}{\text{DNNmodel1}} \right) \right) < O \left( t_p(BM + B) \right).$$  \hspace{1cm} (32)

Therefore, the proposed learning model for power allocation has a lower complexity than the iterative gradient algorithm.

IV. Simulation Results And Discussion

In this section, we simulated the performance of the proposed deep learning-based resource optimization algorithm. In our simulation, a NOMA-based mmWave network with all users distributed within the coverage of the macrocell is studied. The power of AWGN is $\sigma^2 = -134$ dBm; system bandwidth is 1200 MHz; the radius of the macrocell is 100 m; up to 2 users can be occupied on each subchannel; the maximum power of the macrocell is 9.5 dBm, and the maximum power of the small cells is 4.7 dBm. The parameters for deep learning model are as shown in Table I. And the simulation environment of the proposed scheme is as follows: Python 3.6 with TensorFlow 1.3.0 with NVIDIA GeForce GTX 1050. Fig. 4 shows the user association strategy obtained by Lagrange dual decomposition and Max-SINR algorithm. The network environment includes one macrocell BS, three small BSs and 60 users. The main purpose of user association is to solve the problem of load balancing and blind area coverage. By comparing users connected with different types of base stations, it can be seen that under max-SINR algorithm, most users are assigned to macrocell, indicating that the load balancing effect of this algorithm is the worst. Under the Lagrangian Dual method, the number of users connected to macrocell is reduced. Compared to max-SINR algorithm, Lagrange dual decomposition method based user association algorithm embodies better performance of load balancing.

As shown in the Fig. 5, the sum rate of the system is evaluated for different methods. Each line represents the results obtained from 5,000 random data test points. The distributions of sum rate under different algorithms are consistent approximately, ranging from $2 \times 10^9$ to $3.6 \times 10^9$. It is obviously that the results of deep learning schemes are close to the result of two-side matching algorithm. Furthermore, it is also obviously that the performance of semi-supervised learning scheme is much better than that of single DNN models, which proves that semi-supervised learning can address the problem of subchannel allocation well.

Fig. 6 shows the distribution of the EE of system by DNN approach and the gradient iteration algorithm over the whole test dataset. DNN approach gives a good behavior of approximation (about 98%) of the EE generated by gradient iteration algorithm. Fig. 7 shows the EE of systems versus number of training sequences. It is obviously that the results of DNN schemes converges to $5.31 \times 10^{12}$ bits/joule with the changes of epochs, during the training process. As shown in the Fig. 8, we evaluate the EE of system achieved by different schemes.
Each line represents the results obtained from 5,000 random data test points. The range of EE of the system under deep learning schemes is mostly distributed between $0.4 \times 10^{12}$ to $0.6 \times 10^{12}$. It is obviously that the results of DNN schemes are close to the results of gradient iteration algorithm, significantly superior to the other equal power scheme and random power scheme. Furthermore, Fig. 8 shows that semi-supervised learning for subchannel allocation and DNN model for power optimization perform well in EE optimization, and also proves that our proposed schemes can address the resource optimization problem well.

| Symbols                        | Values       |
|-------------------------------|--------------|
| The number of training sets   | 5000 | 5000 |
| The number of testing sets    | 1000 | 1000 |
| The number of layers          | 3 / 4 | 3 |
| The number of neurons per layer | 700 / 80 | 800 |
| Learn rate                    | 0.01 / 0.05 | 0.01 |
| Batch size                    | 200 / 500 | 200 |
| Epochs                        | 100 | 100 |
| Optimizer                     | RMSProp | Adam |

V. CONCLUSIONS

In this paper, a framework based on deep learning has been designed for dealing with user association, subchannel
allocation, and power control in NOMA mmWave networks. User association is solved by the Lagrange dual decomposition, the subchannel allocation is solved by semi-supervised learning and power allocation is solved by DNN scheme. In the proposed scheme, sample data are generated by iterative algorithm, and ML schemes are adopted in the decision-making stage. The proposed scheme improves the EE of the system, and reduces the dependence on the reference algorithm by using unlabeled data. Simulation results verify that in the NOMA-based mmWave network, the effectiveness of the proposed deep learning method in user association, subchannel allocation and power control.

REFERENCES

[1] Q. Ye, B. Rong, Y. Chen, M. Al-Shalash, C. Caramanis, and J. G. Andrews, "User association for load balancing in heterogeneous cellular Veh. Technol. Trans. Wireless Commun., vol. 12, no. 6, pp. 2706–2716, Jan. 2013.

[2] H. Zhang, S. Huang, C. Jiang, K. Long, V. C. M. Leung, and H. V. Poor, "Energy efficient user association and power allocation in millimeter-wave-based ultra dense networks with energy harvesting base stations," IEEE J. Sel. Areas Commun., vol. 35, no. 9, pp. 1936–1947, Sep. 2017.

[3] Y. Lin, W. Bao, W. Yu, and B. Liang, "Optimizing user association and spectrum allocation in hetNets: A utility perspective," IEEE J. Sel. Areas Commun., vol. 33, no. 6, pp. 1025–1039, Jun. 2015.

[4] F. Wang, W. Chen, H. Tang, and Q. Wu, "Joint optimization of user association, subchannel allocation, and power allocation in multi-cell multi-association OFDMA heterogeneous networks," IEEE Trans. Commun., vol. 65, no. 6, pp. 2672–2684, Jun. 2017.

[5] F. Fang, H. Zhang, J. Cheng, S. Roy, and V. C. M. Leung, "Joint user scheduling and power allocation optimization for energy-efficient NOMA systems with imperfect CSI," IEEE J. Sel. Areas Commun., vol. 35, no. 12, pp. 2874–2885, Dec. 2017.

[6] H. Zhang, F. Fang, J. Cheng, K. Long, W. Wang, and V. C. M. Leung, "Energy-efficient resource allocation in NOMA heterogeneous networks," IEEE Wireless Commun., vol. 25, no. 2, pp. 48–53, Apr. 2018.

[7] H. Zhang, N. Yang, K. Long, M. Pan, G. K. Karagiannidis, and V. C. M. Leung, "Secure communications in NOMA system: Subcarrier assignment and power allocation," IEEE J. Sel. Areas Commun., vol. 36, no. 7, pp. 1441–1452, Jul. 2018.

[8] H. Zhang, M. Feng, K. Long, G. K. Karagiannidis, V. C. M. Leung, and V. Poor, "Energy efficient resource management in SWIPT enabled heterogeneous networks with NOMA," IEEE Trans. Wireless Commun., vol. 19, no. 2, pp. 835–845, Feb. 2020.

[9] S. Zhang, B. Di, L. Song, and Y. Li, "Sub-channel and power allocation for non-orthogonal multiple access relay networks with amplify-and-forward protocol," IEEE Trans. Wireless Commun., vol. 16, no. 4, pp. 2249–2261, Apr. 2017.

[10] W. Chen, P. Fan, and Z. Cao, "Water filling in cell: The optimal power allocation policy with channel and buffer state information," in Proc. IEEE Int. Conf. Commun., 2005, vol. 1, pp. 537–541.

[11] F. Fang, J. Cheng, and Z. Ding, "Joint energy efficient subchannel and power optimization for a downlink NOMA heterogeneous network," IEEE Trans. Veh. Technol., vol. 68, no. 2, pp. 1351–1364, Feb. 2019.

[12] C. Liang, F. R. Yu, H. Yao, and Z. Han, "Virtual resource allocation in information-centric wireless networks with virtualization," IEEE Trans. Veh. Technol., vol. 65, no. 12, pp. 9902–9914, Dec. 2016.

[13] H. Zhang, N. Liu, K. Long, J. Cheng, V. C. M. Leung, and L. Hanzo, "Energy efficient subchannel and power allocation for the software defined heterogeneous VLC and RF networks," IEEE J. Sel. Areas Commun., vol. 36, no. 3, pp. 658–670, Mar. 2018.

[14] H. Zhang et al., "Energy efficient dynamic resource optimization in NOMA system," IEEE Trans. Wireless Commun., vol. 17, no. 9, pp. 5671–5683, Sep. 2018.

[15] C. Jiang, H. Zhang, Y. Ren, Z. Han, K. Chen, and L. Hanzo, "Machine learning paradigms for next-generation wireless networks," IEEE Wireless Commun., vol. 24, no. 2, pp. 98–105, Apr. 2017.

[16] Q. Mao, F. Hu, and Q. Hao, "Deep learning for intelligent wireless networks: A comprehensive survey," IEEE Commun. Surv. Tut., vol. 20, no. 4, pp. 2595–2621, Oct.-Dec. 2018.

Haijun Zhang (Senior Member, IEEE) is currently a Full Professor in the University of Science and Technology Beijing, China. He was a Postdoctoral Research Fellow with the Department of Electrical and Computer Engineering, the University of British Columbia (UBC), Canada. He serves as an Editor of IEEE TRANSACTIONS ON GREEN COMMUNICATIONS NETWORKING, and IEEE Communications Letters. He received the IEEE CSIM Technical Committee Best Journal Paper Award in 2018, IEEE ComSoc Young Author Best Paper Award in 2017, and IEEE ComSoc Asia-Pacific Best Young Researcher Award in 2019.
Haisen Zhang received the B.S. degree in the School of Computer and Communication Engineering from the University of Science and Technology of Beijing, Beijing, China, in 2018. He is currently pursuing the M.S. degree at the School of Computer and Communication Engineering from the University of Science and Technology of Beijing, Beijing, China. His research interests include resource allocation, power control, and energy efficiency in wireless communications.

Keping Long (Senior Member, IEEE) received his M.S. and Ph.D. degrees at UESTC in 1995 and 1998, Respectively. From July 2001 to November 2002, he was a research Fellow in the ARC Special Research Centre for Ultra Broadband Information Networks (CUBIN) at the University of Melbourne, Australia. He is now a Professor and Dean at the School of Computer and Communication Engineering (CCE), USTB. He is/ was a Member of the Editorial Committee of Sciences in China Series F and China Communications. He is/ was also a TPC and ISC member for COIN, IEEE IWCN, ICON, APOC, TPC Chair of COIN’05/’08, and TPC Co-Chair of COIN’08/’10. He was awarded the National Science Fund Award for Distinguished Young Scholars of China in 2007 and selected as the Chang Jiang Scholars Program Professor of China in 2008.

George K. Karagiannidis (Fellow, IEEE) was born in Pithagorion, Samos Island, Greece. He received the University Diploma (5 years) and Ph.D. degree, both in electrical and computer engineering from the University of Patras, in 1987 and 1999, respectively. From 2000 to 2004, he was a Senior Researcher at the Institute for Space Applications and Remote Sensing, National Observatory of Athens, Greece. In June 2004, he joined the faculty of Aristotle University of Thessaloniki, Greece where he is currently Professor in the Electrical & Computer Engineering Dept. and Head of Wireless Communications Systems Group (WCSG). He is also Honorary Professor at South West Jiaotong University, Chengdu, China. His research interests are in the broad area of Digital Communications Systems and Signal processing, with emphasis on Wireless Communications, Optical Wireless Communications, Wireless Power Transfer and Applications and Communications & Signal Processing for Biomedical Engineering. He has been involved as General Chair, Technical Program Chair and member of Technical Program Committees in several IEEE and non-IEEE conferences. In the past, he was Editor in several IEEE journals and from 2012 to 2015 he was the Editor-in-Chief of IEEE Communications Letters. Currently, he serves as Associate Editor-in-Chief of IEEE Open Journal of Communications Society. He is one of the highly-cited authors across all areas of Electrical Engineering, recognized from Clarivate Analytics as Web-of-Science Highly-Cited Researcher in the five consecutive years 2015-2019.