

Received January 21, 2020, accepted February 7, 2020, date of publication February 11, 2020, date of current version February 20, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2973253

Power Allocation Schemes Based on Deep Learning for Distributed Antenna Systems

GONGBIN QIAN^{ID}, ZHUKUN LI^{ID}, CHUNLONG HE^{ID}, (Member, IEEE),
XINGQUAN LI^{ID}, AND XUE DING^{ID}

Shenzhen Key Laboratory of Advanced Machine Learning and Applications, Shenzhen University, Shenzhen 518060, China

Corresponding author: Chunlong He (hclong@szu.edu.cn)

This work was supported in part by the Natural Science Funding of Guangdong Province under Grant 2017A030313336, in part by the Shenzhen Overseas High-level Talents Innovation and Entrepreneurship under Grant KQJSCX20180328093835762, and in part by the Tencent Rhinoceros Birds Scientific Research Foundation for Young Teachers of Shenzhen University.

ABSTRACT In recent years, a lot of power allocation algorithms have been proposed to maximize spectral efficiency (SE) and energy efficiency (EE) for the distributed antenna systems (DAS). However, the traditional iterative power allocation algorithms are difficult to be implemented in reality because of their high computational complexity. With the development of machine learning algorithms, it has been proved that the machine learning method has excellent learning ability and low computational complexity, which can approximate the traditional iterative power allocation well and be easily to be implemented in reality. In this paper, we propose a new deep neural network (DNN) model for DAS. From the perspective of machine learning, traditional iterative algorithms can be regarded as a nonlinear mapping between user channel realizations and optimal power allocation schemes. Therefore, we train the DNN to learn the nonlinear mapping between the user channel realizations and the corresponding power allocation schemes based on the traditional iterative algorithm. Then, a power allocation schemes based on DNN method is developed to maximize SE and EE for DAS. The simulation results show that the proposed scheme can not only obtain the almost similar performance as the traditional iterative algorithm, but also reduce much online computational time.

INDEX TERMS Distributed antenna systems, deep learning, power allocation algorithm, sub-gradient algorithm, spectral efficiency, energy efficiency.

I. INTRODUCTION

In recent years, with the improvement of the communication technology, there is a rapid growth in data transmission of cellular networks. The explosive growth data transmission demand has brought tremendous pressure to the fifth-generation (5G) wireless system [1], [2]. In order to provide high rate of data transmission, many researchers have proposed a DAS, which can effectively reduce the access distance between the remote access unit (RAU) and user equipment (UE) [3], [4]. Existing researchs showed that there are a lot of advantages of DAS, such as increasing SE [4]–[7] and improving EE [8]–[12].

In order to provide high data rate transmission, it is also significant for us to optimize SE and EE power allocation in DAS. A lot of algorithms for SE and EE power allocation

have been proposed in the past years. In [13], the authors have exploited the water-filling algorithm to maximum EE of MIMO systems. Moreover, a sub-gradient algorithm and fractional programming have been proposed in [14] to optimize three different objectives for DAS, including maximum SE optimization, minimum transmit power optimization and maximum EE optimization. However, most of works mainly based on the traditional methods to get the optimal power allocation algorithm, which usually sacrifice a large amount of online computational time and complex calculation. It is unrealistic to use the traditional iterative algorithms in actual systems whose user channel state information is constantly changing. Therefore, we should find better schemes to reduce the system's online computational complexity.

Deep learning has been successfully applied in many areas, including computer vision, natural language process and so on. Due to its natural advantages, many researchers try using deep learning algorithms to solve the problem of wireless

The associate editor coordinating the review of this manuscript and approving it for publication was Muhammad Khandaker^{ID}.

communication. A lot of studies found that deep learning algorithms can not only get excellent performance in the wireless communication, but also reduce the computational complexity. Deep learning algorithms make it possible to realize the real-time power allocation in actual systems. In [15], the authors used deep learning for channel estimation and signal detection in orthogonal frequency division multiplexing (OFDM) system, which showed that the performance of deep learning algorithm is better than traditional algorithm in reducing time and computational complexity. In [16], the authors regarded the weighted mean square error algorithm (WMMSE) as traditional iterative algorithm and trained the deep neural network to learn the nonlinear mapping between the channel realizations and the resource allocation of the WMMSE algorithm. The experimental results showed that the deep neural network can not only get the almost similar performance as the WMMSE algorithm, but also reduce much online computational time. The authors in [17] exploited deep learning method to solve the EE optimization of system through power control in wireless interference networks. The results showed that the neural network solution can satisfy the demands with less online complexity. The authors in [18] proposed the ensemble deep neural networks to solve the optimal power control problem and trained the deep learning model by adding noise as one of the input neuron, which can better cope with different noises in reality. In [19], the authors used convolutional neural network (CNN) to solve the wireless scheduling problem of device-to-device (D2D) by using user geographic location instead of user channel state information (CSI). The experimental results showed that the performance of CNN method is very close to the traditional iterative method which used CSI, and can save the large cost of measuring the user CSI. However, to the best of the authors' knowledge, there is little work that has studied the power allocation schemes based on DNN method for DAS. We hope to use deep learning algorithms to realize real-time power allocation in DAS so that the it can be applied in actual wireless communication system.

In this paper, we therefore investigate the deep learning algorithm for the downlink DAS. We define those iterative algorithms as traditional optimization algorithms. One of the most out-standing algorithms for the DAS is sub-gradient algorithm which is mentioned in [15]. Therefore, we define the sub-gradient algorithm as the traditional algorithm and design a DNN to approximate it. We randomly generate a large number of channel realizations and use the traditional sub-gradient algorithm to get the corresponding optimal power allocation schemes. Then we define those channels realizations and the corresponding optimal power allocation schemes as the inputs and labels of the DNN, respectively. We use the channel realizations and labels to train the deep neural network so that the DNN can learn the nonlinear mapping between the channel realizations and the results of the traditional sub-gradient power allocation algorithm. We focus on designing a suitable structure and appropriate parameters of the DNN so that the performance of the DNN method is

close enough to the traditional sub-gradient algorithm and reduce a large amount of online computation time.

The remainder of this paper can be listed as follow. In Section II, the DAS model is presented including the system configure and the channel model. In addition, we formulate maximum SE optimization and maximum EE optimization problem of DAS. In Section III, we present the system architecture, the network structure, the process of data generation and stage of training the DNN. In Section IV, we present the parameters selection of the deep neural network and the simulation results presented to demonstrate our assumption. We conclude this paper in Section V.

II. SYSTEM MODEL AND PROBLEM FORMULATION

A. SYSTEM MODEL

In this section, we consider a downlink scenario of DAS. There are N RAUs with one antenna and K cellular UEs with single-antenna in DAS. N RAUs are uniformly deployed in the cell and connected to the central base station. K UEs are randomly distributed in the cell, which showed in Fig.1. We use $h_{n,k}$ to denote the channel frequency response between the n th RAU and the k th UE, which consists of a small and large scale fading [9] and can be expressed as

$$h_{n,k} = g_{n,k}w_{n,k}, \quad (1)$$

where $g_{n,k}$ represents the small-scale fading between the n th RAU and the k th UE, $w_{n,k}$ represents the large-scale fading which is independent of $g_{n,k}$ [20].

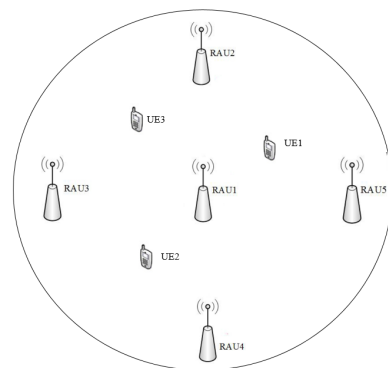


FIGURE 1. System model.

B. MAXIMUM SE OPTIMIZATION

In this part, the maximum SE optimization problem of DAS is shown. We assume that the perfect channel state information (CSI) is available at both transmitter and receiver side. In order to reduce the computational time of traditional algorithm so that we can generate enough train data set for DNN, we assume that the channels are orthogonal so there is no interference to each other. The problem of maximizing SE

for the downlink DAS can be modeled as [14]

$$\max_{P_{n,k}} \sum_{k=1}^K \log_2 \left(1 + \frac{\sum_{n=1}^N p_{n,k} |h_{n,k}|^2}{\sigma^2} \right) \quad (2)$$

$$s.t. \sum_{k=1}^K p_{n,k} \leq P_{max}^n, \quad (2a)$$

$$p_{n,k} \in [0, P_{max}^n], \quad (2b)$$

where P_{max}^n is the maximum transmit power of the n th RAU. $p_{n,k}$ is the transmit power from the n th RAU to the k th cellular UE. σ^2 represents the power of the complex additive white Gaussian noise (AWGN) of the UE [14].

C. MAXIMUM EE OPTIMIZATION

From [11], the problem of maximizing EE can be modeled as

$$\max_{P_{n,k}} \frac{\sum_{k=1}^K \log_2 \left(1 + \frac{\sum_{n=1}^N p_{n,k} |h_{n,k}|^2}{\sigma^2} \right)}{\frac{1}{\tau} \sum_{k=1}^K \sum_{n=1}^N p_{n,k} + NP_d + P_c + P_o} \quad (3)$$

$$s.t. \sum_{k=1}^K p_{n,k} \leq P_{max}^n, \quad (3a)$$

$$p_{n,k} \in [0, P_{max}^n], \quad (3b)$$

where P_d is the constant circuit power consumption per RAU and P_c denotes the constant basic power consumption and P_o denotes the disseminated of the optical fiber transmission. τ represents the power amplifier efficiency [14].

D. TRADITIONAL SUB-GRADIENT ALGORITHM

According to [11], the maximum SE optimization can find the optimal power allocation schemes by using the sub-gradient algorithm. When the objective problem is maximizing EE, we should transform the non-convex objective function into an equivalent objective function with subtractive form by using fractional programming. However, the time complexity of using sub-gradient algorithm is still high without considering the interference between users, and the time complexity is extremely high after combining fractional programming for maximizing the EE optimization. Therefore, we cannot use algorithms directly in the actual systems due to the high computational complexity.

III. DEEP NEURAL NETWORK BASED METHOD

In this part, we therefore consider the deep learning algorithm. We regard the DNN as a ‘‘black box’’ and use DNN to approximate the sub-gradient algorithm in an end-to-end fashion. In the proposed model, the sub-gradient algorithm can be treated as unknown nonlinear mapping between the channel realizations and corresponding power allocation schemes. The deep neural network can deal with this nonlinear mapping problem. Therefore, we can design a DNN model to learn the nonlinear mapping between the channel realizations and corresponding power allocation schemes

from a large number of data, which aims to design an appropriate DNN to realize a real-time power allocation in the DAS.

A. SYSTEM ARCHITECTURE

In this paper, in order to prove that the power allocation based on the DNN can be applied to different maximum transmission power of RAU, we will compare the performance of the power allocation based on the DNN with the traditional sub-gradient algorithm under different maximum transmission power. The different maximum transmission power can be respectively represent as $P_{max1}, P_{max2}, \dots, P_{maxi}$. Because there are different optimal power allocation under different maximum transmission power of RAUs, the same channel realizations are put into the traditional sub-gradient algorithm and we will get different power allocation schemes. Therefore, for the DNN method, we fed the same channel realizations into the DNN and we will get different results of power allocation schemes. We train different DNNs by feeding the same channel realizations as inputs and the different results of power allocation by traditional algorithm as labels. The power allocation of traditional sub-gradient algorithm under different maximum transmission power which are the labels for different DNNs and can represented as $P(P_{max1}), P(P_{max2}), \dots, P(P_{maxi})$. Therefore, the system architecture of the DNN method for DAS can be illustrated in Fig.2.

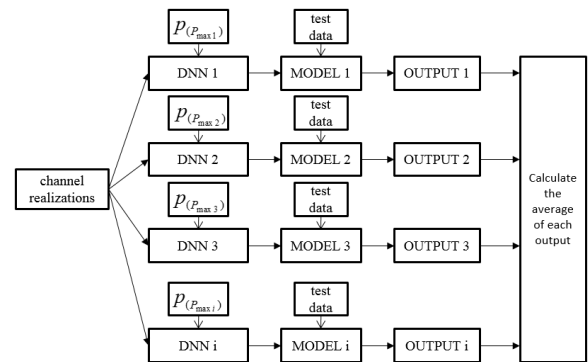


FIGURE 2. The structure of the deep neural network.

B. DATA GENERATION

Because the deep neural network algorithm is data-dependent method and it is very important for us to train the DNNs that prepare a large number of the training data and corresponding labels. Firstly, the channel model mentioned above is used to generate a lot of channel matrices $H^{(t)}$, where t is the index of training samples. Then it can be exploited in the traditional sub-gradient algorithm to get the optimal power allocation matrix $P_{P_{max1}}^{(t)}, P_{P_{max2}}^{(t)}, \dots, P_{P_{maxi}}^{(t)}$, which are labels for correspondent $H^{(t)}$, respectively. For convenience, we use $P^{(t)}$ to represent labels under different maximum transmission power. Therefore, $[H^{(t)}, P^{(t)}]$ denotes as the t th sample. By repeating the above process, we generate a large number

of samples as training samples. In addition, we use the cross validation method during training process and we randomly split 99% of the training samples into training dataset and 1% of the training samples into validation dataset. The validation dataset plays an important part in avoiding the over-fitting during the training process. Finally, in order to test the performance of the DNN method, we need to generate a large amount of samples as testing dataset.

C. NETWORK ARCHITECTURE

After training and testing dataset prepared, we need to design a deep neural network to approximate the sub-gradient algorithm to maximize SE and EE power allocation in DAS. A fully connected neural network is proposed, which includes one input layer, three hidden layers, and one output layers as shown in Fig.3.

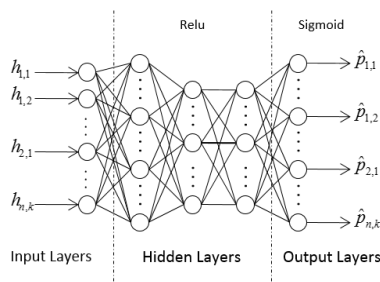


FIGURE 3. The structure of the deep neural network.

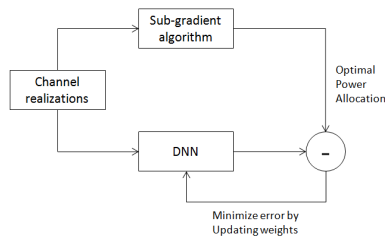


FIGURE 4. Data generation and training process.

The inputs of the DNN are the channel realizations $h_{n,k}$ and the outputs of the DNN are power allocation schemes $\hat{p}_{n,k}$. In the DAS model, we train the DNN to learn the nonlinear mapping between the channel realizations $h_{n,k}$ and the power allocation of traditional sub-gradient algorithm $p_{n,k}$. The outputs of the DNN should be a continuous value. Therefore, this problem is a nonlinear regression problem. In order to enhance the nonlinear fitting ability of the deep neural network, ReLU function is exploited as the activation function for the three hidden layers. ReLU function can be represent as

$$ReLU(x) = \max(0, x) \tag{4}$$

Since we first normalized the training data and labels, the output of the neural network should be between 0 and 1. We use

the sigmoid function as the activation function of the output layer.

$$sigmoid(x) = \frac{1}{1 + e^{-x}} \tag{5}$$

D. TRAINING STAGE

Training stage of the DNN mainly includes two processes, feed-forward operation and back propagation. The feed-forward operation is to calculate the loss value of the DNN. Otherwise, the back propagation is to update the weights and bias of the DNN by minimizing loss value, which depends on the important part of deep neural network, e.t. loss function. The loss function can be expressed as

$$\begin{aligned} loss &= \mathbb{E}[loss_{mse} + loss_{const}] \\ &= \mathbb{E}[\lambda_1 \sum_{n=1}^N \sum_{k=1}^K (\hat{p}_{n,k} - p_{n,k})^2 \\ &\quad + \lambda_2 \sum_{n=1}^N ReLU(\sum_{k=1}^K \hat{p}_{n,k} - P_{max}^n)], \end{aligned} \tag{6}$$

where $p_{n,k}$ represents the optimal power allocation of the traditional sub-gradient algorithm and $\hat{p}_{n,k}$ represents the output of the DNNs.

The loss function contains two parts, including the mean square error between outputs of the DNN and labels $loss_{mse}$ and the constraint function error $loss_{const}$. The scaling factors λ_1 and λ_2 are used to balance the $loss_{mse}$ and $loss_{const}$ in order to ensure the DNN can be trained well enough. The first part of the loss function $loss_{mse}$ is used to reduce the error between $\hat{p}_{n,k}$ and $p_{n,k}$ so that the DNNs can reach the performance as the traditional sub-gradient algorithm. The second part of the loss function $loss_{const}$ ensure the output of the deep neural network strictly satisfying the constraint (2a)(3a). If $\sum_{k=1}^K \hat{p}_{n,k} \geq P_{max}^n$, the constraint (2a)(3a) are not satisfied and $ReLU(\sum_{k=1}^K \hat{p}_{n,k} - P_{max}^n) > 0$. Then the second part loss function $loss_{const}$ will force the network parameters to be updated to satisfy the constraint (2a)(3a). On the contrary, if $\sum_{k=1}^K \hat{p}_{n,k} \leq P_{max}^n$, the constraint (2a)(3a) are satisfied and $ReLU(\sum_{k=1}^K \hat{p}_{n,k} - P_{max}^n) = 0$, the second part loss function $loss_{const}$ will not influence the network training.

We adopt the RMSprop algorithm as the optimization algorithm, which is an efficient implementation of mini-batch gradient descent and we choose 0.9 as the decay rate [21]. In order to improve the performance of the DNN, we choose the Xavier initialization [22] to initialize the weights. When using a large learning rate, the training speed will be improved and will get high convergence error. Otherwise, when using a small learning rate, the train speed will be slow down and get the low convergence error. In addition, when using a big batch size, the convergence error will increase. otherwise, when using small batch size, the convergence error will decrease but unstable. Therefore, we will try different learning rate and batch size, then choose appropriate learning rate and batch size based on the validation error of previous 300 times

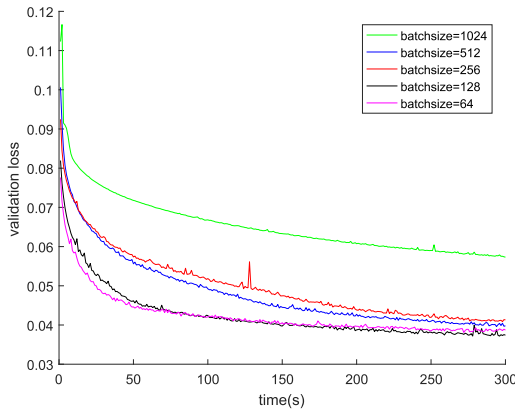


FIGURE 5. Batch size selection.

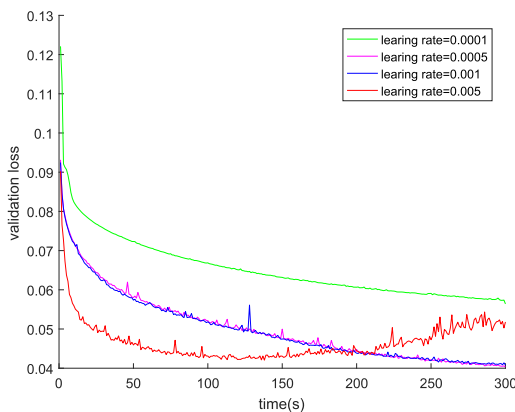


FIGURE 6. Learning rate selection.

of the DNN training, which are shown in Fig.5 and Fig.6, respectively.

IV. SIMULATION RESULTS

In this section, we generate 50000 samples as training dataset, of which 500 are randomly spilt into validation dataset and 49500 are randomly spilt into training dataset. Then, in order to reduce the influence of randomness on experimental results, we will generate 5000 samples as testing dataset and repeat the experiment 5000 times in two methods, finally get their average respectively.

A. SCENARIO I AND PARAMETERS SELECT

In this part, we consider the scenario I, which contains 5 RAUs and 3 UEs. A fully-connected deep neural network which contains five layers, one input layer with 15 nodes, three hidden layers with 50, 100, 50 nodes and one output layer with 15 nodes is proposed for DAS. We use ReLU as the activate function for both hidden layers and output layer. We set the batch size and the learning rate to be 512 and 0.001 respectively. In addition, we set the scale factors λ_1 and λ_2 to be 1 and 0.1, respectively.

B. SCENARIO II AND PARAMETERS SELECT

In this part, in order to verify the scalability of the power allocation based on DNN method, we test the model in scenario II, which contains 5 RAUs and 10 UEs. The input layer is 50 nodes, the number of neurons in the hidden layer are 100, 150 and 100. The number of neurons in output layer is 50. We use the same batch size, learning rate and scale factors as the scenario I.

C. RESULTS

The simulation parameters are listed in Table.1. In the case of scenario I, when the objective problem is maximizing SE, the SE and EE performance of the sub-gradient algorithm and the power allocation based on DNN method is shown in Fig.7 and Fig.8. When the objective problem is maximizing EE, the SE and EE performance of the two methods is shown in Fig.11 and Fig.12.

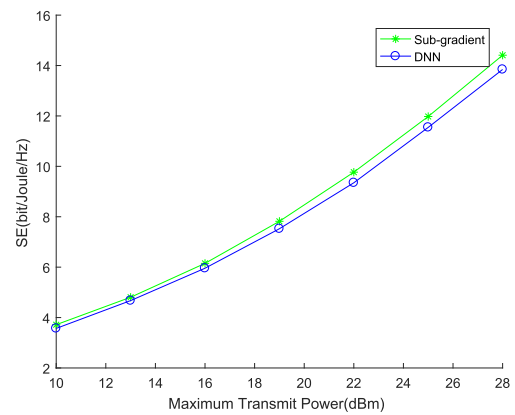


FIGURE 7. SE versus maximum transmit power in scenario I when the objective problem is maximizing SE.

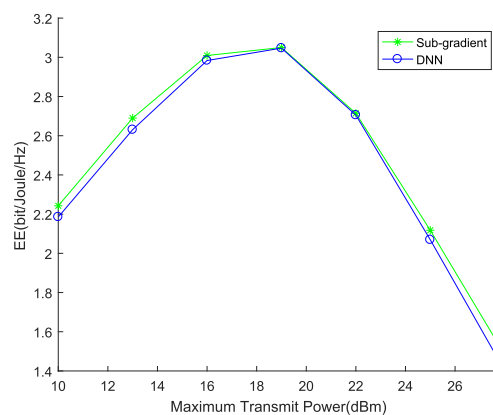


FIGURE 8. EE versus maximum transmit power in scenario I when the objective problem is maximizing SE.

In the case of scenario II, when the objective problem is maximizing SE, the SE and EE performance of the two methods is shown in Fig.9 and Fig.10. When the objective problem is maximizing EE, the SE and EE performance of the two methods is shown in Fig.13 and Fig.14.

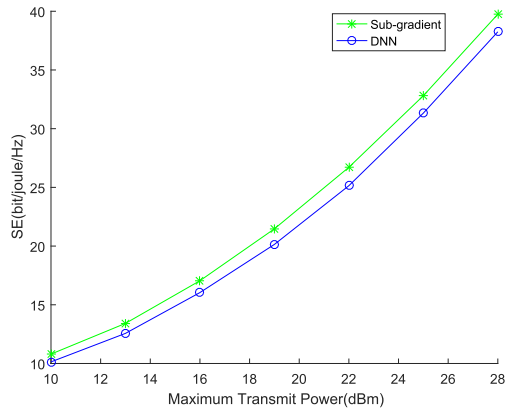


FIGURE 9. SE versus maximum transmit power in scenario II when the objective problem is maximizing SE.

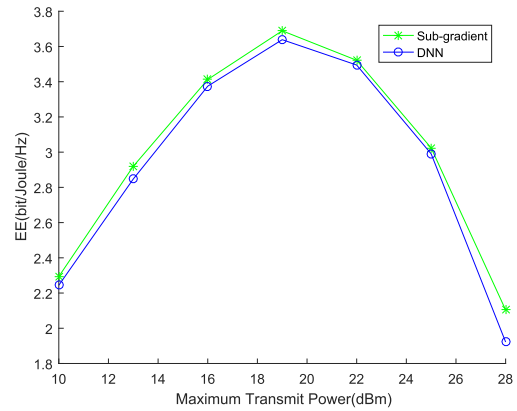


FIGURE 12. EE versus maximum transmit power in scenario I when the objective problem is maximizing EE.

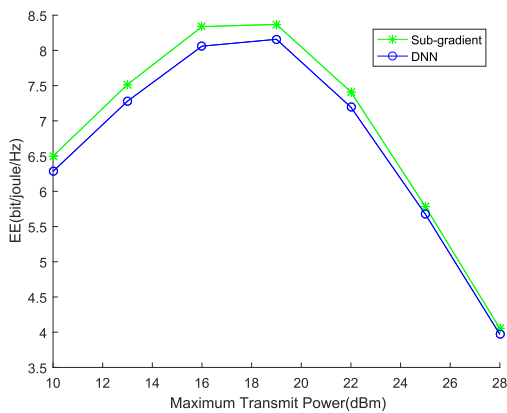


FIGURE 10. EE versus maximum transmit power in scenario I when the objective problem is maximizing SE.

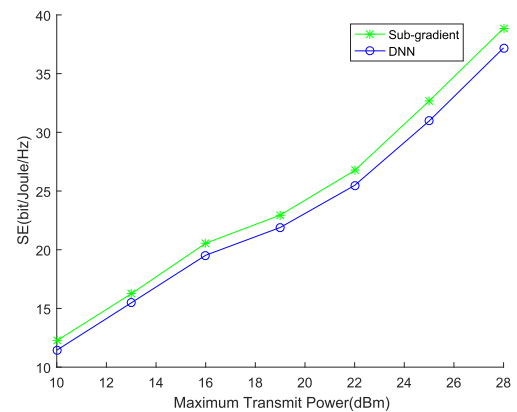


FIGURE 13. SE versus maximum transmit power in scenario II when the objective problem is maximizing EE.

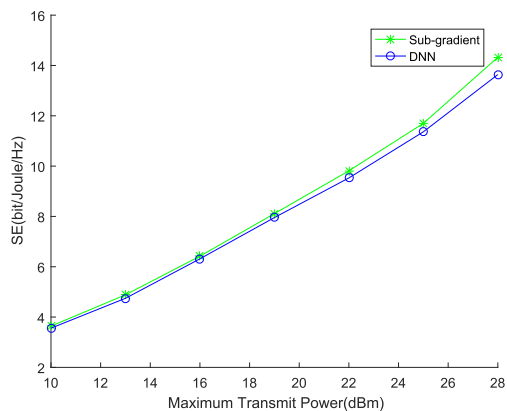


FIGURE 11. SE versus maximum transmit power in scenario I when the objective problem is maximizing EE.

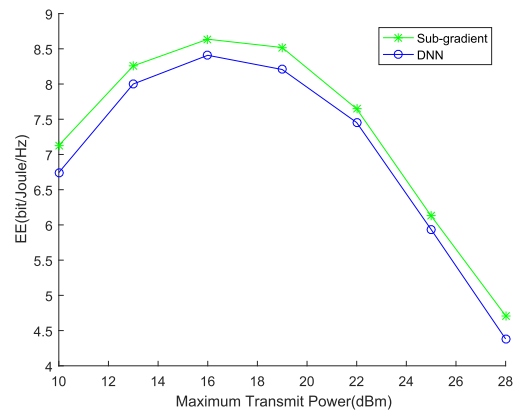


FIGURE 14. EE versus maximum transmit power in scenario II when the objective problem is maximizing EE.

In order to prove that the performance of power allocation based on DNN method is very close to the sub-gradient algorithm, the accuracy of the DNN method under two scenarios are shown in Table 1 and Table 2, respectively. In addition, to better compare the computational complexity between the two methods, the computational time of the two methods

under the two scenarios is shown in Table 3 and Table 4, respectively.

D. RESULTS ANALYZE

As shown in the simulation results above, we can see from Fig.6 to Fig.13, the performance of the DNN method is very close to the traditional sub-gradient algorithm. We can see

TABLE 1. Simulation parameters.

Parameters	Value
The cellular radius D	1000m
The number of UE K	3 or 10
The noise power	-104dBm
The maximum transmit power of UE	10~28dBm
The constant circuit power consumption	20dBm
The constant basic power consumption	30dBm
The optical fiber transmission	30dBm
Path loss exponent	4
Drain efficiency	38%

TABLE 2. Performance comparison for the two methods in scenario I.

PP	accuracy (DNN/Sub-gradient)			
	maxSE		maxEE	
	SE	EE	SE	EE
10	96.0%	97.4%	98.1%	97.9%
13	97.2%	97.9%	97.5%	97.1%
16	97.1%	99.1%	98.3%	98.1%
19	96.2%	99.8%	98.2%	98.7%
22	95.9%	99.6%	97.6%	99.1%
25	96.3%	97.6%	97.2%	99.0%
28	96.0%	97.5%	96.5%	93.1%

TABLE 3. Performance comparison for the two methods in scenario II.

PP	accuracy (DNN/Sub-gradient)			
	maxSE		maxEE	
	SE	EE	SE	EE
10	96.2%	96.7%	93.6%	95.1%
13	95.3%	97.1%	95.0%	97.2%
16	95.3%	97.2%	94.2%	97.5%
19	94.4%	97.3%	94.2%	95.5%
22	95.9%	97.7%	95.5%	97.5%
25	95.6%	98.9%	95.1%	97.1%
28	96.6%	97.9%	95.6%	93.2%

TABLE 4. Computational times for the two methods in scenario I.

PP	Computation Time(s)			
	maxSE		maxEE	
	DNN	Sub-gradient	DNN	Sub-gradient
10	0.000058	0.358288	0.000064	151.5821
13	0.000058	0.190786	0.000070	9.261512
16	0.000058	0.071224	0.000066	6.675778
19	0.000056	0.026411	0.000064	6.232736
22	0.000058	0.038264	0.000064	9.192754
25	0.000062	0.070213	0.000066	14.20951
28	0.000058	0.178401	0.000064	28.38203

form the Table 1 and Table 2. In the two scenarios described above, whether the objective problem is maximizing SE or maximizing EE, the accuracy of the DNN method can reach more than 92% of the traditional sub-gradient algorithm.

We can see from the table 3 and table 4. To begin with, for the DNN method, the online computational time is almost same at different maximum transmission power, because the structure of the DNN is same under the same scenario. For the traditional sub-gradient algorithm, the online computational time will be different under different maximum transmission power, because the number of iterations required to find the

TABLE 5. Computational times for the two methods in scenario II.

PP	Computation Time(s)			
	maxSE		maxEE	
	DNN	Sub-gradient	DNN	Sub-gradient
10	0.000086	2.027886	0.000144	106.6814
13	0.000090	1.984795	0.000144	96.76731
16	0.000084	1.706011	0.000144	58.16991
19	0.000084	0.543031	0.000142	11.53632
22	0.000098	0.164041	0.000144	29.15942
25	0.000096	0.289052	0.000158	19.78721
28	0.000084	0.409261	0.000144	57.60124

optimal power allocation schemes may be different. In addition, for the sub-gradient algorithm, the computational time required to maximize EE is much higher than the maximum SE. What's more, the online computational time of the the DNN method is many times less than that of the traditional sub-gradient algorithm.

In order to further understand the advantages of the DNN method, we compare the time complexity of the DNN method and traditional sub-gradient algorithm. According to the [23], the online time complexity of the already trained well fully connected neural network is $O(n)$. When the objective problem is maximizing SE, the online time complexity of the traditional sub-gradient methods is $O(n^3)$. When the objective problem is maximizing EE, because the algorithm needs to combine fractional programming with sub-gradient algorithms, the online time complexity is higher than when maximizing SE.

In conclusion, the performance of the DNN method can achieve 92% of the traditional sub-gradient algorithm and provide with at least three orders of magnitude times speed up.

V. CONCLUSION

In this paper, we exploited the DNN algorithm to solve power allocation problem in DAS. Firstly, we introduced a system model and assumed that perfect channel state information is known at both transmitter and receiver side and the channels are orthogonal so there are no interference to each other. Secondly, we randomly generated a large number of the channel realizations and used the traditional sub-gradient algorithm to maximize SE and EE optimization and save the optimal power allocation schemes. Thirdly, we fed the channel realizations and the corresponding power allocation schemes into the DNN and used the loss function to train the DNN. Fourthly, we generated other channel realizations for testing dataset and fed into the DNN which had been trained well and got the outputs of DNN. Otherwise, the same testing dataset were fed into the traditional sub-gradient algorithm and calculate the optimal power allocation schemes. Finally, we showed the performance and computational time difference between the DNN method and the sub-gradient algorithm, respectively.

In the future, we will explore more excellent machine learning algorithms to solve the power allocation problem

in DAS. For example, we will try to use the CNN to realize power allocation by using user geographic location instead of user channel information. In addition, we will try to use migration learning to deal with the problem of insufficient training dataset in wireless communication systems. What's more, we will try to use the ensemble learning to make the generalization ability of the trained model stronger so that it can cope with various complicated situations in actual wireless communication systems. Finally, we hope to find more machine learning methods with excellent performance and low computational complexity to truly implement power allocation algorithms in actual wireless communication systems.

REFERENCES

- [1] Q. Wu, G. Y. Li, W. Chen, D. W. K. Ng, and R. Schober, "An overview of sustainable green 5G networks," *IEEE Wireless Commun.*, vol. 24, no. 4, pp. 72–80, Aug. 2017.
- [2] C. Pan, M. ElKashlan, J. Wang, J. Yuan, and L. Hanzo, "User-centric C-RAN architecture for ultra-dense 5G networks: Challenges and methodologies," *IEEE Commun. Mag.*, vol. 56, no. 6, pp. 14–20, Jun. 2018.
- [3] H. Ren, N. Liu, C. Pan, M. ElKashlan, A. Nallanathan, X. You, and L. Hanzo, "Power- and rate-adaptation improves the effective capacity of C-RAN for Nakagami- m fading channels," *IEEE Trans. Veh. Technol.*, vol. 67, no. 11, pp. 10841–10855, Nov. 2018.
- [4] T. Shen, S. Zhang, R. Chen, J. Wang, J. Hu, F. Shu, and J. Wang, "Two practical random-subcarrier-selection methods for secure precise wireless transmissions," *IEEE Trans. Veh. Technol.*, vol. 68, no. 9, pp. 9018–9028, Sep. 2019.
- [5] F. Shu, X. Wu, J. Hu, J. Li, R. Chen, and J. Wang, "Secure and precise wireless transmission for random-subcarrier-selection-based directional modulation transmit antenna array," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 4, pp. 890–904, Apr. 2018.
- [6] D. Wang, X. You, J. Wang, Y. Wang, and X. Hou, "Spectral efficiency of distributed MIMO cellular systems in a composite fading channel," in *Proc. IEEE Int. Conf. Commun. (ICC)*, Beijing, China, May 2008, pp. 1259–1264.
- [7] J. Wang, H. Zhu, and N. J. Gomes, "Distributed antenna systems for mobile communications in high speed trains," *IEEE J. Sel. Areas Commun.*, vol. 30, no. 4, pp. 675–683, May 2012.
- [8] C. Pan, H. Zhu, N. J. Gomes, and J. Wang, "Joint precoding and RRH selection for user-centric green MIMO C-RAN," *IEEE Trans. Wireless Commun.*, vol. 16, no. 5, pp. 2891–2906, May 2017.
- [9] C. He, B. Sheng, P. Zhu, X. You, and G. Y. Li, "Energy- and spectral-efficiency tradeoff for distributed antenna systems with proportional fairness," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 5, pp. 894–902, May 2013.
- [10] C. He, G. Y. Li, F.-C. Zheng, and X. You, "Power allocation criteria for distributed antenna systems," *IEEE Trans. Veh. Technol.*, vol. 64, no. 11, pp. 5083–5090, Nov. 2015.
- [11] H. Ren, N. Liu, and C. Pan, "Energy efficient transmission for multicast services in MISO distributed antenna systems," *IEEE Commun. Lett.*, vol. 20, no. 4, pp. 756–759, Apr. 2016.
- [12] C. Pan, H. Zhu, N. J. Gomes, and J. Wang, "Joint user selection and energy minimization for ultra-dense multi-channel C-RAN with incomplete CSI," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 8, pp. 1809–1824, Aug. 2017.
- [13] J. Xu and L. Qiu, "Energy Efficiency Optimization for MIMO broadcast channels," *IEEE Trans. Wireless Commun.*, vol. 12, no. 2, pp. 690–701, Feb. 2013.
- [14] C. He, "Comparison of three different optimization objectives for distributed antenna systems," *AEU-Int. J. Electron. Commun.*, vol. 70, no. 4, pp. 442–448, Apr. 2016.
- [15] H. Ye, G. Y. Li, and B.-H. Juang, "Power of deep learning for channel estimation and signal detection in OFDM systems," *IEEE Wireless Commun. Lett.*, vol. 7, no. 1, pp. 114–117, Feb. 2018.
- [16] H. Sun, X. Chen, Q. Shi, M. Hong, F. Xiao, and N. D. Sidiropoulos, "Learning to optimize: Training deep neural networks for wireless resource management," in *Proc. IEEE 18th Int. Workshop Signal Process. Adv. Wireless Commun. (SPAWC)*, Sapporo, Japan, Jul. 2017, pp. 1–6.
- [17] B. Matthiesen, A. Zappone, K.-L. Besser, E. A. Jorswieck, and M. Debbah, "A globally optimal energy-efficient power control framework and its efficient implementation in wireless interference networks," 2018, *arXiv:1812.06920*. [Online]. Available: <http://arxiv.org/abs/1812.06920>
- [18] F. Liang, C. Shen, W. Yu, and F. Wu, "Towards optimal power control via ensembling deep neural networks," 2018, *arXiv:1807.10025*. [Online]. Available: <http://arxiv.org/abs/1807.10025>
- [19] W. Cui, K. Shen, and W. Yu, "Spatial deep learning for wireless scheduling," 2018, *arXiv:1808.01486*. [Online]. Available: <http://arxiv.org/abs/1808.01486>
- [20] X. You, D. Wang, P. Zhu, and B. Sheng, "Cell edge performance of cellular mobile systems," *IEEE J. Sel. Areas Commun.*, vol. 29, no. 6, pp. 1139–1150, Jun. 2011.
- [21] G. Hinton, N. Srivastava, and K. Swersky. (2012). *Lecture 6A Overview of Mini-Batch Gradient Descent*. [Online]. Available: <https://class.coursera.org/neuralnets-2012-001/lecture>
- [22] X. Glorot and Y. Bengio, "Understanding the difficulty of training deep feedforward neural networks," in *Proc. 13th AISTATS*, Sardinia, Italy, May 2010, pp. 249–256.
- [23] M. Lee, Y. Xiong, G. Yu, and G. Y. Li, "Deep neural networks for linear sum assignment problems," *IEEE Wireless Commun. Lett.*, vol. 7, no. 6, pp. 962–965, Dec. 2018.



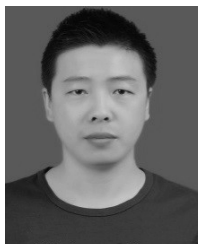
GONGBIN QIAN received the B.E. and M.Sc. degrees in communications and electronics system from the Harbin Institute of Technology, China, in 1990 and 1993, respectively. He joined the College of Information Engineering, Shenzhen University, China, in 1993. His research interests include signal processing, wireless communications, and communications systems.



ZHUKUN LI received the B.S. degree in electronic information engineering from Guangdong Ocean University, Zhanjiang, China, in 2018. He is currently pursuing the M.S. degree with the College of Information Engineering, Shenzhen University, Shenzhen, China. His research interests include wireless communication networks, resource allocation, and machine learning algorithms.



CHUNLONG HE (Member, IEEE) received the M.S. degree in communication and information science from Southwest Jiaotong University, Chengdu, China, in 2010, and the Ph.D. degree from Southeast University, Nanjing, China, in 2014. From September 2012 to September 2014, he was a Visiting Student with the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA, USA. Since 2015, he has been with the College of Information Engineering, Shenzhen University, where he is currently an Assistant Professor. His research interests include communication and signal processing, green communication systems, channel estimation algorithms, and limited feedback techniques. Dr. He is a member of the Institute of Electronics, Information, and Communication Engineering. He is also an Associate Editor of IEEE Access.



XINGQUAN LI received the master's degree in computer application technology from the Lanzhou University of Technology, Lanzhou, China, in 2015. He is currently pursuing the Ph.D. degree with the College of Information Engineering, Shenzhen University. His research interests include cooperative communications and green communications.



XUE DING received the B.S. degree in electronic and information engineering from the Beijing University of Chemical Technology, Beijing, China, in 2019. She is currently pursuing the M.S. degree with the College of Information Engineering, Shenzhen University, Shenzhen, China. Her research interests include wireless powered communication networks, resource allocation and UDNs with resource capacity, and user scheduling.

• • •