IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received April 9, 2019, accepted April 24, 2019, date of publication April 30, 2019, date of current version May 20, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2914159

Energy Efficient Power Allocation Based on Machine Learning Generated Clusters for Distributed Antenna Systems

CHUNLONG HE, (Member, IEEE), YUEHUA ZHOU, GONGBIN QIAN, XINGQUAN LI, AND DAQUAN FENG, (Member, IEEE)
College of Information Engineering, Shenzhen University, Shenzhen 518060, China

Corresponding author: Yuehua Zhou (yuehzhou@163.com)

This work was supported in part by the Natural Science Foundation of China under Grant 61601300, 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 this paper, we consider the combination of machine learning (ML) and wireless communication. We design a machine learning generated clusters model in a distributed antenna system (DAS), which is constructed by two different ML clustering algorithms, i.e., k-means algorithm and Gaussian mixture model-based (GMM) algorithm. Under the communication scenario of DAS with ML generated clusters model, we investigate two different power allocation optimization problems with the interference of maximizing spectral efficiency (SE) and energy efficiency (EE) in DAS, respectively. We compare the SE and EE of DAS with ML generated clusters model and the conventional model. The simulation results verify the effectiveness of DAS with ML generated clusters model, which can obtain the much better performance of SE and EE compared with the conventional communication model in DAS.

INDEX TERMS Machine learning, spectral efficiency, energy efficiency, k-means, mixture of Gaussian model, distributed antenna system.

I. INTRODUCTION

The update and exchange of wireless data impose a huge burden on the existing cellular networks, and at the same time pose many serious challenges for the design, operation, and maintenance of the fifth generation (5G) wireless networks. In order to reduce the pressure on the communication networks caused by the explosive growth of data traffic, many new proposals for reforming and updating traditional communication technologies to fit the 5G wireless networks have emerged in the existing research work [1], and the effectiveness of many innovative approaches has been confirmed, which can be seen from the changes of the single-input single-output (SISO) techniques to the multiple-input multiple-output (MIMO) techniques [2]–[4], the traditional cellular communication modes to the communication mode with device to device [5]–[7], the centralized antenna systems (CAS) to the distributed antenna systems (DAS) [8]–[10], and the appearance of the ultra-dense networks [11].

The research of designing the efficient power allocation scheme to improve the energy efficiency (EE) [12], [13] and spectral efficiency (SE) [14]–[16] has always been a hot spot. In [17], Singh *et al.* discussed two transmit beamforming design problems including sum-power minimization problem and EE maximization problem in MIMO system. In [18], Cirik *et al.* investigated a resource allocation problem of joint power and subcarrier allocation to maximize the EE in multiuser relay networks.

As discussed in the existing research works [19], [20], DAS can significantly improve the EE and SE of the communication systems compared to CAS. However, it is inevitable that there is a serious interference problem among users and a problem of high computational complexity of the center unit, which is connected with the other remote access units (RAUs) by fibre [21], in the DAS. In order to solve the above two

The associate editor coordinating the review of this manuscript and approving it for publication was Yuan Gao.

problems and ensure the good performance of the communication systems, we introduce the machine learning (ML) algorithms into wireless communication systems. For more complex and dynamic of 5G wireless networks, artificial intelligence (AI) techniques provides us with a superior alternative option over traditional communication technologies in terms of joint optimization, detection, estimation and other communication behaviors [22]. Based on the channel state information (CSI) as a training sample set, the antenna selection issue in MIMO is considered as a ML behavior by constructing a multi-class classifier, which is achieved by k-nearest neighbors (KNN), support vector machine (SVM), naive-Bayes (NB) [23], [24], for a good communication performance. In [25], a framework for resource allocation based on ML elaborates on the process of applying ML algorithms to resource allocation in wireless communication networks, and a beam allocation problem has been solved by the KNN algorithm under low complexity conditions in a single cell MIMO system. The k-means algorithm has been exploited to design the next generation 5G wireless networks [26]. The gaussian mixture model-based (GMM) algorithm contributes to clustering the users for interest estimate in social networks [27]. The authors applied the k-NN algorithm to classify the cellular users for getting the power allocation schemes of DAS according to the historical data constructing by the traditional method [28]. The Q-learning was introduced to offer an alternative option to solve the high complexity problem of the traditional resource allocation solution in multi-cell, multi-user system, which is possibly applied to DAS [22]. The authors applied the DNN (deep neural network) to obtain the power allocation of maximizing EE in a multi-cell interference network, which also provides insights into the combination of ML and power allocation in DAS [29]. From the previous works, it has become an inevitable trend for ML to be applied to 5G wireless networks to provide communication systems with better communication performance than conventional communication technologies.

In this paper, firstly, we propose to use k-means algorithm and GMM algorithm to form a ML generated clusters communication model in DAS. By using the ML clustering algorithms, the users in each cluster are served by the only one base station selected by the clustering center, which significantly reduces interference from the other RAUs and reduces computational complexity of the center unit. Then, under the communication scenario of DAS with ML generated clusters model, we firstly discuss the problems of maximizing SE and EE under the requirements of each users' minimum SE and the maximum transmit power of RAUs in DAS with ML generated clusters model, respectively. Last, we compare the computational complexity of the proposed algorithm and the conventional algorithm to gain insight into the advantages of DAS with the ML generated clusters mode. And we consider the existing power allocation scheme appeared in the literature [32] as the comparison algorithm i.e., the conventional method in this paper, and the conventional communication model in DAS is typically used by each RAU, which serves all the cellular users. However, when all RAUs serve the kth cellular user in this model, it will bring severe interference from the communication of RAUs to the other cellular users [32]. Simulation results show that compared with the conventional algorithm, the ML generated clusters model can achieve the much better communication performance in DAS. When we consider the problem of maximizing SE in DAS, using ML generated clusters model in DAS can improve the SE of DAS, especially when the maximum transmit power is large. When we investigate the problem of maximizing EE in DAS, the communication model of ML generated clusters also can obtain a better performance. For example, when the maximum transmit power is 20 dBm, EE reaches maximum and EE of using ML generated clusters model in DAS constructed by k-means clustering algorithm and GMM clustering algorithm is about 230% and 180% higher than EE of conventional communication model in DAS.

The remainder of this paper is organized as follows. Section II introduces the system configure. Specifically, there are the system model, total power consumption. In section III, we investigate the ML generated clusters model achieved by the k-means clustering algorithm and GMM clustering algorithm, respectively. In section IV, we formulate the maximum SE optimization problem and the maximum EE optimization problem of ML generated clusters model in DAS, and develop the corresponding optimal power allocation algorithms to obtain the solutions, respectively. section V presents the simulation results to demonstrate the effectiveness and the validity of using ML generated clusters model in DAS. We draw the conclusion in section VI.

II. SYSTEM CONFIGURE

In DAS, N RAUs with one antenna each are uniformly distributed in the cell, and K singer-antenna cellular users are randomly distributed in the cellular cell, respectively. In this paper, we make the following assumptions: i) the users share the same spectrum. ii) the total system bandwidth is normalized to unit. iii) the cell is a circle with the radius R.

A. CHANNEL MODEL

As the work discussed in [30], the channel $h_{n,k}$ between kth user and nth RAU can be modeled as the composite fading channel, which can be expressed as the following

$$h_{n,k} = g_{n,k} w_{n,k}, \tag{1}$$

where $g_{n,k}$ is the small scale fading, and can be expressed as independent and identically distributed complex Gaussian random variables with zero mean and unit variance. The large scale fading is denoted by $w_{n,k}$, which is independent of the small scale fading $g_{n,k}$, and can be written as [12]

$$w_{n,k} = \sqrt{\frac{c s_{n,k}}{d_{n,k}^{\alpha}}}, \tag{2}$$

where c is the median of the mean path gain at a reference distance of $d_{n,k} = 1\ km$, $d_{n,k}$ is the distance between kth user and nth RAU, α is the path loss factor and is typically between 3 and 5, and $s_{n,k}$ is log-normal shadow fading variable, which means $10log_{10}s_{n,k}$ is a zero mean Gaussian random variable with standard deviation σ_{sh} [12], [31].

B. ML ALGORITHM

In this part, we focus particularly on the combination of ML and DAS. In DAS, the conventional communication model is typically used by each RAU, which serves all cellular users. However, when all RAUs serve the kth cellular user in this model, it will cause severe interference from the communication of RAUs to the other cellular users [32]. Different from the conventional communication of DAS, we propose a ML generated clusters model in DAS that is based on the k-means algorithm and GMM algorithm. With the ML generated clusters model in DAS, unsupervised learning algorithm divides K cellular users into several clusters with a certain number of users. Meanwhile, a corresponding clustering center will be generated in each cluster. The mth clustering center automatically select the closest RAU to the clustering center for all users in the mth cluster, before communicating with RAUs. When each cluster is associated with the corresponding RAU, the nth RAU serves only these users that select nth RAU as their unique serving base station. Finally, we discuss the problems of maximizing SE and EE under the requirements of each users' minimum SE and the maximum transmit power of RAUs in DAS with ML generated clusters model, respectively.

III. ML CLUSTERS MODEL IN DAS

In this section, we will introduce two unsupervised learning algorithm to form the ML generated clusters model in DAS, which can significantly decrease the interference of the communication systems.

A. K-MEANS CLUSTERING ALGORITHM

Given the d-dimension user data set $\mathbf{X} = \{\mathbf{x_i}|\mathbf{x_i} \in R^d, i = 1, 2, \ldots, K, d = 2\}$, which illustrates the location distribution information of the K users, where $\mathbf{x_i}$ denotes the position coordinates of the ith cellular user, and K is the number of the cellular users, and d is the dimension of the position coordinates (i.e., d = 2). The k-means clustering algorithm [33] minimizes the squared error, which is expressed as the following, based on the clustering results $\mathbf{C} = \{c_1, c_2, \ldots, c_S\}$,

$$E = \sum_{i=1}^{S}\sum_{\mathbf{x}\in c_i} \|\mathbf{x} - \mu_{\mathbf{i}}\|_2^2, \tag{3}$$

where $\mu_{\mathbf{i}} = \frac{1}{|c_i|}\sum_{\mathbf{x}\in c_i}\mathbf{x}$ denotes the clustering center (the mean of all user samples) in the cluster c_i, and $|c_i|$ is the number of the user samples in c_i [34].

Intuitively, the formulation (6) to a certain extent describes the degree of the closeness of users around the clustering center. The smaller the value of E is, the higher the degree of similarity the user samples. The k-means clustering algorithm divides the users of data set $\mathbf{X}$ into pre-determined S clusters by minimizing the squared error function (6). Specifically, S users are selected as the initial clustering center in advance from the user set containing K user samples. For the rest of the user samples, according to their similarity (distance) with these clustering centers, they are respectively assigned to the clusters most similar to them. Then calculate the clustering center of each newly obtained cluster. The process is repeated until each clustering no longer changes.

B. GMM CLUSTERING ALGORITHM

Different from the k-means clustering algorithm, which the sample data prototype vectors are used to describe the clustering structure, GMM clustering algorithm employed the probability mode (gaussian distribution) to describe the clustering prototype. According to [35], these Gaussian distributed components can be combined together to construct the overall probability model, which means it is flexible enough to approximate any distribution based on the Gaussian mixture distributions. So we can describe the clustering prototype of DAS users as the probability mode of Gaussian distributions by using the GMM algorithm to progressively approximate the distribution of the described data by Gaussian mixture components. The clusters are determined by the probability corresponding to the sample data prototype. Specifically, gaussian mixture distribution, which can be defined as following, consists of S mixture components, each of which is a gaussian distribution,

$$P_M(\mathbf{x}) = \sum_{i=1}^{S}\alpha_i P(\mathbf{x}|\mu_{\mathbf{i}}, \varepsilon_{\mathbf{i}}), \tag{4}$$

where $\mu_{\mathbf{i}}$ and $\varepsilon_{\mathbf{i}}$ are parameters i.e., the mean and covariance matrix of the ith gaussian mixture component, respectively. $P(\mathbf{x}|\mu_{\mathbf{i}}, \varepsilon_{\mathbf{i}})$ can be expressed as

$$P(\mathbf{x}|\mu_{\mathbf{i}}, \varepsilon_{\mathbf{i}}) = \frac{1}{(2\pi)^{\frac{d}{2}}|\varepsilon_{\mathbf{i}}|^{\frac{1}{2}}}e^{-\frac{1}{2}(\mathbf{x}-\mu_{\mathbf{i}})^T\varepsilon_{\mathbf{i}}^{-1}(\mathbf{x}-\mu_{\mathbf{i}})}, \tag{5}$$

and $\alpha_i > 0$ is the corresponding mixture coefficient, which is satisfied

$$\sum_{i=1}^{S}\alpha_i = 1, \tag{6}$$

where α_i represents the probability of selecting the ith gaussian mixture component.

Given the d-dimension user data set $\mathbf{X} = \{\mathbf{x_j}|\mathbf{x_j} \in R^d, j = 1, 2, \ldots, K\}$, which illustrates the location distribution information of the K users, where $\mathbf{x_i}$ denotes the position coordinates of the ith cellular user, and K is the number of the cellular users, and d is the dimension of the position coordinates (i.e., d = 2).. Let the random variable $Z_j \in \{1, 2, \ldots, S\}$ present the gaussian mixture component of user sample $\mathbf{x_j}$. Obviously, the prior probability of variable Z_j, i.e., $P(Z_j = i)$, corresponds to $\alpha_i, (i = 1, 2, \ldots, S)$. According to Baye's theorem, the posterior distribution of Z_j can be expressed as

$$P_M(Z_j = i|\mathbf{x_j}) = \frac{P(Z_j = i)P_M(\mathbf{x_j}|Z_j = i)}{P_M(\mathbf{x_j})} = \frac{\alpha_i P(\mathbf{x_j}|\mu_\mathbf{i}, \varepsilon_\mathbf{i})}{\sum_{l=1}^{S} \alpha_l P(\mathbf{x_j}|\mu_\mathbf{i}, \varepsilon_\mathbf{i})}. \tag{7}$$

In other words, $P_M(Z_j = i|\mathbf{x_j})$ shows the posterior probability of user sample $\mathbf{x_j}$ generated by the ith gaussian mixture component, which is denoted as $r_{j,i}, (i = 1, 2, \ldots, S)$ for the convenience of description. When the gaussian mixture distribution is known, GMM clustering algorithm will divide the sample user set $\mathbf{X}$ into S clusters $C = \{c_1, c_2, \ldots, c_S\}$, and the cluster tag λ_j of each user sample $\mathbf{x_j}$ is determined as following

$$\lambda_j = \underset{i\in\{1,2,\ldots,S\}}{\operatorname{argmax}} \; r_{j,i}. \tag{8}$$

For the given sample user set $\mathbf{X}$, the maximum likelihood is expressed as follows

$$L(\mathbf{X}) = \ln\left(\prod_{j=1}^{K}(P_M(\mathbf{x_j})\right) = \sum_{j=1}^{K} \ln\left(\sum_{i=1}^{S} \alpha_i P(\mathbf{x_j}|\mu_\mathbf{i}, \varepsilon_\mathbf{i})\right). \tag{9}$$

The model parameters of (11) are obtained by iterative optimizations using the Expectation Maximization (EM) [36] algorithm to fit the GMM to the given user set $\mathbf{X}$. In the ith iteration of the EM algorithm, the mixture coefficient of each gaussian component are calculated as

$$\alpha_i = \frac{1}{K}\sum_{j=1}^{K} r_{j,i}, \tag{10}$$

which shows the mixture coefficient of each gaussian component is determined by the average posterior probability of the user sample belonging to the component. Whereas the mean value and the covariance value of each gaussian component are calculated as

$$\mu_\mathbf{i} = \frac{\sum_{j=1}^{K} r_{j,i}\mathbf{x_j}}{\sum_{j=1}^{K} r_{j,i}}, \tag{11}$$

$$\varepsilon_\mathbf{i} = \frac{\sum_{j=1}^{K} r_{j,i}(\mathbf{x_j} - \mu_\mathbf{i})(\mathbf{x_j} - \mu_\mathbf{i})^T}{\sum_{j=1}^{K} r_{j,i}}. \tag{12}$$

The formulation (14) presents that the mean of each mixed component can be estimated by a sample weighted average, which is the posterior probability of each user sample belongs to the component.

After clustering user samples with the k-means algorithm and GMM algorithm, we can get the the clustering results and the clustering centers corresponding to each cluster. Next, we will discuss how to select the appropriate unique serving base station for the users in each cluster by calculating the distance between each clustering center and each RAU. The specific process is as follows

Step1: Define the distance between the ith clustering center and nth RAU, which can be expressed as

$$d_{i,n} = \|RAU_n - \mu_\mathbf{i}\|^2, \tag{13}$$

where $\mu_\mathbf{i}$ is the ith clustering center and RAU_n is the nth RAU, $n \in [1, 2, \ldots, N]$, $i \in [1, 2, \ldots, S]$.

Step2: Determine the serving base station tag R_i of the cluster c_i according to the nearest RAU, which can be described as

$$R_i = \underset{i\in\{1,2,\ldots,S\}}{\operatorname{argmin}} \; d_{i,n}. \tag{14}$$

Step3: Put the all users of the cluster c_i into the corresponding set B_{R_i}, which can be described as

$$B_{R_i} = B_{R_i} \bigcup \{c_i\}. \tag{15}$$

After the discussions above, the detail steps of k-means clustering algorithm and GMM clustering algorithm in DAS are described in TABLE 1 and TABLE 2.

TABLE 1. K-means clustering algorithm in DAS.

Algorithm 1 k-means Clusterng Algorithm in DAS

1: Initializing the RAUs set $B_{i_1} = \varnothing, i_1 \in [1, N]$, and user data set $|\mathbf{X}| = K, S = N$.
2: randomly select S user samples from $\mathbf{X}$ as the initial clustering center, $\{\mu_\mathbf{1}, \mu_\mathbf{2}, \ldots, \mu_\mathbf{S}\}$.
3: **repeat**
  let $c_i = \varnothing, i \in [1, S]$.
4: **for** $j = 1, 2, \ldots, K$
    calculate the distance between the user sample $\mathbf{x_j}$ and each clustering center $\mu_\mathbf{i}, i \in [1, S]$: $d_{j,i} = \|\mathbf{x_j} - \mu_\mathbf{i}\|_2$.
    determine the clustering tag of $\mathbf{x_j}$ based on the nearest clustering center: $\lambda_j = \operatorname{argmin}_{i\in\{1,2,\ldots,S\}} d_{j,i}$.
    put sample user $\mathbf{x_j}$ into the corresponding cluster: $c_{\lambda_j} = c_{\lambda_j} \bigcup \{\mathbf{x_i}\}$.
  end for
5: **for** $i = 1, 2, \ldots, S$
    calculate the new clustering center: $\mu'_\mathbf{i} = \frac{1}{|c_i|}\sum_{\mathbf{x}\in c_i} \mathbf{x}$.
    if $\mu'_\mathbf{i} \neq \mu_\mathbf{i}$
      update the current clustering center $\mu_\mathbf{i}$ to $\mu'_\mathbf{i}$
    else
      keep the current clustering center $\mu_\mathbf{i}$.
    end if
  end for
6: **until** the current clustering center is no longer updated.
7: **for** $i = 1, 2, \ldots, S$
    calculate the distance between each clustering center $\mu_\mathbf{i}$ and $\text{RAU}_n, n \in [1, N]$: $d_{i,n} = \| \text{RAU}_n - \mu_\mathbf{i}\|_2$.
    determine the serving base station tag R_i of the cluster c_i according to the nearest RAU: $R_i = \operatorname{argmin}_{i\in\{1,2,\ldots,S\}} d_{i,n}$.
    put the all users of the cluster c_i into the corresponding set : $B_{R_i} = B_{R_i} \bigcup \{c_i\}$.
  end for
8: return $B_{i_1}, i_1 \in [1, N]$.

IV. POWER ALLOCATION SCHEME OF DAS WITH ML GENERATED CLUSTERS

When we use the clustering algorithm to construct the ML generated clusters model in DAS, the number of cellular users severed by the jth RAU can be obtained, i.e., $|B_j| = B_j, j \in [1, N]$, and the cellular users can use the above communication mode to form the ML generated clusters model in DAS.

TABLE 2. GMM clustering algorithm in DAS.

Algorithm 2 GMM Clustering Algorithm in DAS

1: Initializing the RAUs set $B_{i_1} = \varnothing, i_1 \in [1, N]$, user data set $|\mathbf{X}| = K, S = N$. and obtain the starting values $\{\alpha_\mathbf{i}, \mu_\mathbf{i}, \varepsilon_\mathbf{i}, |1 \leqslant i \leqslant S\}$.
3: **repeat**
4: **for** $j = 1, 2, ..., K$
calculate the posterior probability generated by each mixture gaussian component according to equation (5), i.e., $r_{j,i} = P_M(Z_j = i|\mathbf{x_j}), i \in [1, S]$.
end for
5: **for** $i = 1, 2, ..., S$
calculate the new clustering center according to equation (11).
calculate the new covariance matrix according to equation (12)
calculate the new mixture coefficient according to equation (10)
end for
7: update model parameter $\{\alpha_i, \mu_\mathbf{i}, \varepsilon_\mathbf{i}, |1 \leqslant i \leqslant S\}$ to $\{\alpha_i', \mu_\mathbf{i}', \varepsilon_\mathbf{i}', |1 \leqslant i \leqslant S\}$
8: **until** numerical convergence according to equation (9).
9: **for** $i = 1, 2, ..., S$
calculate the distance between each clustering center $\mu_\mathbf{i}$, $i \in [1, S]$ and RAU$_n$, $n \in [1, N]$: $d_{i,n} = \|$ RAU$_n - \mu_\mathbf{i}\|_2$.
determine the serving base station tag R_i of the cluster c_i according to the nearest RAU: $R_i = \mathrm{argmin}_{i \in \{1,2,...,S\}} d_{i,n}$.
put the all users of the cluster c_i into the corresponding set : $B_{R_i} = B_{R_i} \bigcup \{c_i\}$.
end for
10: return $B_{i_1}, i_1 \in [1, N]$.

When the jth RAU selected by a cluster, the SE of the ith user in this cluster can be expressed as

$$R_i = \log_2\left(1 + \frac{p_i|h_{i,j}|^2}{p_1 + p_2 + \sigma^2}\right), \tag{16}$$

where $p_1 = \sum_{k=1,k\neq i}^{B_j} p_k|h_{i,j}|^2$ and $p_2 = \sum_{n=1,n\neq j}^{N} \sum_{t=1}^{B_n} p_t|h_{i,n}|^2$, let p_i and p_t denote the transmit power of the jth RAU to the ith user and the nth RAU to the tth user in the nth cluster served by the nth RAU, respectively. $h_{i,j}$ and $h_{i,n}$ are the composite fading channels between the jth RAU and the ith user and between the ith user and the RAU of the nth cluster, respectively.

From the exiting work [37], the total power consumption P_{total} contains two parts: i) the transmit power consumption. ii) the extra circuit power consumption. Let the P_{trains} denote the transmit power consumption, which can be expressed as

$$P_{trains} = \sum_{n=1}^{N}\sum_{j=1}^{B_j} p_j, \tag{17}$$

The extra circuit power consumption includes the dynamic power consumption P_{dy}, which is independent of the actual transmit power [38], the constant basic power consumption P_{st}, and the consumption circuit of optical fiber transmission P_0. Then the extra circuit power consumption $P_{circuit}$ can be expressed as

$$P_{circuit} = NP_{dy} + P_{st} + P_0. \tag{18}$$

So, the P_{total} can be written as

$$\begin{aligned} P_{total} &= \frac{P_{trains}}{\tau} + P_{circuit} \\ &= \frac{P_{trains}}{\tau} + NP_{dy} + P_{st} + P_0, \end{aligned} \tag{19}$$

where τ is the drain efficiency of the radio frequency power amplifier.

A. MAXIMUM SE OF DAS WITH ML GENERATED CLUSTERS MODEL

Under the following two constraints: i) the requirements of users' minimum SE. ii) the maximum transmit power of RAUs. the optimization problem of maximizing SE of the downlink DAS with ML generated clusters can be modeled as

$$\max_{\mathbf{P_{SE}}} \ R_c = \sum_{i=1}^{K} R_i \tag{20}$$

$$s.t. \quad R_i \geq R_{min}, \quad \forall i \in [1, K], \tag{20a}$$

$$\sum_{j}^{B_j} P_j \leq P_{max}, \quad \forall j \in [1, N], \tag{20b}$$

where R_c is the total SE of DAS with ML generated clusters mode. $\mathbf{P_{SE}} = \{p_i, i \in [1, K]\}$, P_{max} is the maximum transmit power of RAUs. R_{min} denotes the minimum SE requirements of the each user.

Here, firstly, an efficient optimization algorithm based on difference of convex functions structure (D.C.) [39] programming is employed to transform (20) into following formulation

$$f_{SE}(\mathbf{P_{SE}}) = f_{cave}^{SE}(\mathbf{P_{SE}}) + f_{vex}^{SE}(\mathbf{P_{SE}}), \tag{21}$$

where the concave function and convex function can be denoted as the following

$$f_{cave}^{SE}(\mathbf{P_{SE}}) = \sum_{i=1}^{K} \log_2\left(\sum_{n=1}^{N}\sum_{i=1}^{B_n} p_i|h_{i,n}|^2 + \sigma^2\right), \tag{22}$$

$$\begin{aligned} f_{vex}^{SE}(\mathbf{P_{SE}}) = -\sum_{i=1}^{K} \log_2 &\left(\sum_{k=1,k\neq i}^{B_j} p_k|h_{i,j}|^2\right. \\ &\left. + \sum_{n=1,n\neq j}^{N}\sum_{t=1}^{B_n} p_t|h_{i,n}|^2 + \sigma^2\right). \end{aligned} \tag{23}$$

Then let C_1 denote the constraint set of (20), which is the nonlinear constraint condition, but (20a) can be transformed into the linear expression as following

$$\begin{aligned} P_i|h_{i,j}|^2 - (2^{R_{min}} - 1) &\left(\sum_{k=1,k\neq i}^{B_j} P_k|h_{i,j}|^2\right. \\ &\left. + \sum_{n=1,n\neq j}^{N}\sum_{t=1}^{B_n} p_t|h_{i,n}|^2 + \sigma^2\right). \end{aligned} \tag{24}$$

So the problem of maximizing SE of the downlink DAS with ML generated clusters model under the above two constraints can be transformed into a D.C. optimization problem with a convex constraint set, which is expressed as

$$\begin{aligned} \max_{\mathbf{P_{SE}} \in C_1} \quad & f_{SE}(\mathbf{P_{SE}}) \\ s.t. \quad & (20a), (20b). \end{aligned} \tag{25}$$

TABLE 3. Maximum SE of DAS with ML generated clusters model.

Algorithm 3 Maximum SE power allocation algorithm in DAS with ML generated clusters model

1: Initializing $i = 0, \mathbf{P}_{\mathbf{SE}}^{(\mathbf{0})} \in C_1$ and the tolerance parameter $\zeta > 0$.
2: **do**
3: $\mathbf{P_{SE}}^{(i+1)} = \text{argmax}_{\mathbf{P_{SE}} \in C_1}(f_{cave}^{SE}(\mathbf{P_{SE}}) + \nabla f_{vex}^{SE}(\mathbf{P_{SE}}^{(i)}) * \mathbf{P_{SE}}^{T})$
4: Selecting the interior point method to deal with the convex problem above.
5: Applying the logarithm barrier function to transform the original constrained problem into unconstrained problem.
6: get the the research direction based on the Quasi-Newton method.
7: Using the backtracking linear research of Armijo rule to get the optimal step length.
8: $i = i + 1$.
9: **while** $\|\mathbf{P_{SE}}^{(i+1)} - \mathbf{P_{SE}}^{(i)}\| > \zeta$,
10: return $\mathbf{P}_{\mathbf{SE}}^{(\mathbf{i+1})}$.

When there is partial derivatives for the convex part of the D.C. object function, a concave convex procedure (CCCP) algorithm can be employed to solve the D.C. structure optimization problems [40]. In the every iteration of the algorithm, the first order taylor expansion is employed to transform the $f_{vex}^{SE}(\mathbf{P_{SE}})$ into linear expression, and we solve the formulation as following

$$\mathbf{P_{SE}}^{(i+1)} = \underset{\mathbf{P_{SE}} \in C_1}{\text{argmax}}\,(f_{cave}^{SE}(\mathbf{P_{SE}}) + \nabla f_{vex}^{SE}(\mathbf{P_{SE}}^{(i)}) * \mathbf{P_{SE}}^{T}), \quad (26)$$

where the $\nabla f_{vex}^{SE}(\mathbf{P_{SE}}^{(i)})$ denotes the gradient of $f_{vex}^{SE}(\mathbf{P_{SE}})$ at the point of $\mathbf{P_{SE}}^{(i)}$, and i is the iteration step. According to the discussion above, we convert the original problem (20) into a standard convex optimization problem (25), which can be solved by applying the interior point method to (26). We summarize the specific algorithm process in TABLE 3.

B. MAXIMUM EE OF DAS WITH ML GENERATED CLUSTERS MODEL

In this part, we will firstly discuss the EE model of DAS with the ML generated clusters model, and then investigate the objective problem of maximizing EE in the DAS with ML generated clusters model.

From the existing work [41], the EE of the DAS with ML generated clusters model can be expressed as

$$\eta_{EE} = \frac{R_{total}}{P_{total}}, \quad (27)$$

where R_{total} denotes the total SE of all users, and P_{total} denotes the total power consumption.

Next, maximizing EE of downlink DAS with the ML generated clusters model with requirements of the minimum SE and the constraints of maximum transmit power of each RAU can be expressed as

$$\max_{\mathbf{P_{EE}}} \quad EE = \frac{R_c}{\frac{P_{trains}}{\tau} + NP_{dy} + P_{st} + P_0} \quad (28)$$

$$s.t. \quad R_i \geq R_{min}, \quad \forall i \in [1, K], \quad (28a)$$

$$\sum_{j}^{B_j} p_j \leq p_{max}, \quad \forall j \in [1, N], \quad (28b)$$

where $\mathbf{P_{EE}} = \{p_i, i \in [1, K]\}$ is the optimization variables.

From the exiting work [42], we can exploit the fractional programming theory to rewrite the objective function as the following subtractive optimization problem

$$\underset{\mathbf{P_{EE}}}{\text{argmax}} \quad h_1(\mathbf{P_{EE}}, \omega_1) = R_c - \frac{\omega_1}{\tau}P_{trains} - \omega_1 NP_{dy} - \omega_1 P_{st} - \omega_1 P_0$$

$$s.t. \quad (25a), (25b). \quad (29)$$

In **Theorem 1**, which has been proved in the exiting work [31], the relationship between problem (28) and (29) can be explained.

Theorem 1: Let $G_1(\omega_1) = \max_{\mathbf{P_{EE}}} h_1(\mathbf{P_{EE}}, \omega_1)$ and $g_1(\omega_1) = \text{argmax}_{\mathbf{P_{EE}}} h_1(\mathbf{P_{EE}}, \omega_1)$. The optimal power allocation $\mathbf{P}_{\mathbf{EE}}^{*}$ reach the maximum EE in (28) if and only if $G_1(\omega_1^{*}) = \mathbf{P}_{\mathbf{EE}}^{*}$.

According to **Theorem 1**, we can know that the relationship between problem (28) and its transformation (29) is equivalent. In order to obtain the optimal solutions of (29), we use the similar method used in Section to IV.A to transform it into the following optimization problem

$$\max_{\mathbf{P_{EE}}} \quad \{f_{cave}^{EE}(\mathbf{P_{EE}}) + f_{vex}^{EE}(\mathbf{P_{EE}})\}$$

$$s.t. \quad (28a), (28b), \quad (30)$$

where

$$f_{cave}^{EE}(\mathbf{P_{EE}}) = \sum_{i=1}^{K} \log_2(\sum_{n=1}^{N}\sum_{i=1}^{B_n} p_i|h_{i,n}|^2 + \sigma^2), \quad (31)$$

$$f_{vex}^{EE}(\mathbf{P_{EE}}) = -\omega_1\left(\frac{P_{trains}}{\tau} + (N + U_{ac})P_{dy} + P_{st} + P_0\right) - \sum_{i=1}^{K}\log_2\left(\sum_{k=1,k\neq i}^{B_j} P_k|h_{i,j}|^2 + \sum_{n=1,n\neq j}^{N}\sum_{t=1}^{B_n} P_t|h_{i,n}|^2 + \sigma^2\right). \quad (32)$$

$f_{vex}^{EE}(\mathbf{P_{EE}})$ has the partial derivative, so the CCCP algorithm can be employed again to solve the objective problem (30), and the specific iteration step can be described as

$$\mathbf{P_{EE}}^{(i+1)} = \underset{\mathbf{P_{EE}} \in C_1}{\text{argmax}}\,(f_{cave}^{EE}(\mathbf{P_{EE}}) + \nabla f_{vex}^{EE}(\mathbf{P_{EE}}^{(i)}) * \mathbf{P_{EE}}^{T}), \quad (33)$$

where the $\nabla f_{vex}^{EE}(\mathbf{P_{EE}}^{(i)})$ denotes the gradient of $f_{vex}^{EE}(\mathbf{P_{EE}})$ at the point of $\mathbf{P_{EE}}^{(i)}$, i is the iteration step. According to the iterate step, we can obtain the optimal power optimal power allocation for maximizing EE in DAS with the ML generated clusters model, which is showed in TABLE 4.

V. SIMULATION RESULTS

In this paper, we chose the power allocation of maximizing SE and EE with the proactive communication model that has appeared in the literature [32], i.e., conventional model

TABLE 4. Maximum EE of DAS with ML generated clusters model.

Algorithm 4 Maximum EE power allocation algorithm in DAS with ML generated clusters model

1: Initializing $t = 0, \mathbf{P}_{\mathbf{EE}}^{(\mathbf{0})} \in C_1$ and the tolerance parameter $\zeta > 0$. $\omega_1^{(0)} = 0.01$.
2: **repeat**
3: Converting the problem (28) into D.C. objective function as (29) and solving it by using CCCP algorithm.
4: Selecting the interior point method to deal with the convex problem by (33).
5: Applying the logarithm barrier function to transform the original constrained problem into unconstrained problem.
6: Getting the the research direction based on the Quasi-Newton method.
7: Using the backtracking linear research of Armijo rule to get the optimal step length.
8: $\omega_1^{(t+1)} = EE|_{\mathbf{P}_{\mathbf{EE}}^{(\mathbf{t+1})}}$.
9: $t = t + 1$.
10: **until** $|h(\mathbf{P}_{\mathbf{EE}}^{(\mathbf{t+1})}, \omega_1^{t+1})| < \zeta$, $h(\mathbf{P}_{\mathbf{EE}}^{(\mathbf{t+1})}, \omega_1^{t+1}) = \max_{\mathbf{P}_{\mathbf{EE}}^{(\mathbf{t+1})} \in C_1} \{R_c - \omega_1^{t+1} \frac{p_{trains}}{\tau} + NP_{dy} + P_{st} + P_0)\}$.
11: return $\mathbf{P}_{\mathbf{EE}}^{(\mathbf{t+1})}$ and ω_1^{t+1}.

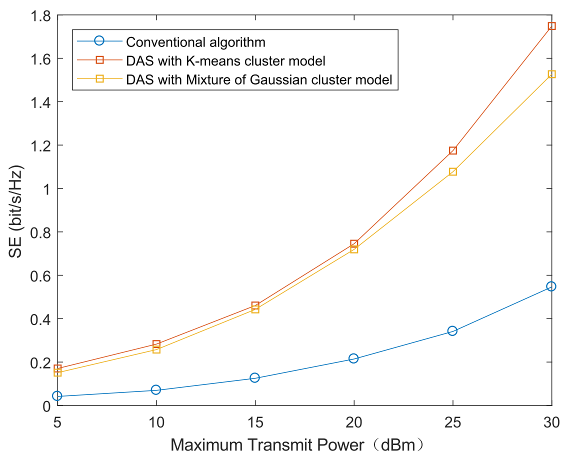


FIGURE 1. SE versus the maximum transmit power.

in DAS, as the comparison algorithm of the proposed algorithm. And in this section, we provide the numerical results to verify the validity and the effectiveness of applying ML to the wireless communication systems by comparing the performance of DAS using ML generated clusters mode and DAS with conventional model [32]. The system parameters in the simulations are listed in TABLE 5. We define the cell as a circle with radius R. N RAUs with one antenna each are uniformly distributed in the cell, and K singer-antenna cellular users are randomly distributed in the cellular cell.

Fig.1 compares the SE of DAS with ML generated clusters mode and conventional model versus the maximum transmit power from 5 dBm to 30 dBm. From Fig.1, we can easily know that the SE of DAS with ML generated clusters model is much better than the conventional communication model in DAS. While both the SE of DAS with and without ML generated clusters model increase with the growth of the maximum transmit power, the SE's growth trend of DAS

TABLE 5. Simulation parameters.

Parameters	Value
The cellular radius R	1000 m
The BS number N	5
The user number K	50
The noise power δ^2	-104 dBm
The minimum requirements of SE R_{min}	1 Mbps
The dynamic power consumption P_{dy}	30 dBm [32]
The static power consumption P_{st}	40 dBm [32]
The optical fiber dissipated power P_0	0.1484 dB/KM [43]
The maximum transmit power P_{max}	30 dBm
Path loss exponent α	4
Drain efficiency η	38% [44]

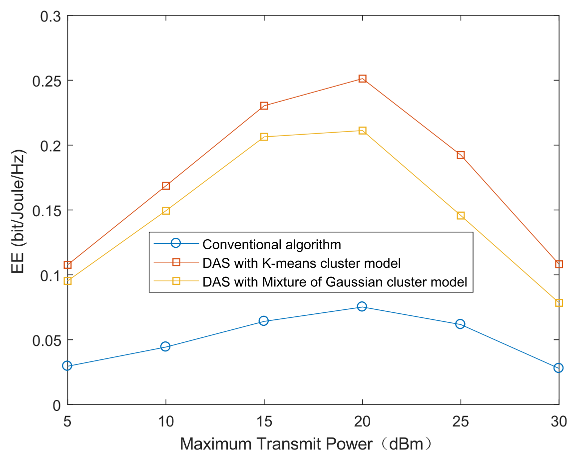


FIGURE 2. EE versus the maximum transmit power.

with ML generated clusters model grows much faster than conventional communication model in DAS, which is more obvious when the maximum transmit power is higher. And the SE of DAS with ML generated clusters model achieved by k-means clustering algorithm is better than the DAS with ML generated clusters model constructed by GMM clustering algorithm. Moreover, When the maximum transmit power is 30 dBm, the SE of DAS with ML generated clusters model achieved by k-means clustering algorithm and GMM clustering algorithm are approximately 220% and 170% higher than the conventional communication model in DAS, respectively. This represents using ML generated clusters model in DAS can improve the SE of DAS, especially when the maximum transmit power is large.

Fig.2 presents the change of EE in DAS with ML generated clusters mode and conventional model versus the maximum transmit power from 5 dBm to 30 dBm. It shows that the EE of DAS with ML generated clusters model is much better than DAS without ML generated clusters model. With the growth of the maximum transmit power, EE decreases but it is still higher in DAS with ML generated clusters mode than the conventional communication model in DAS, where the EE of DAS with ML generated generated clusters model obtained by k-means clustering algorithm is better than the DAS with ML generated clusters model built by GMM clustering algorithm. Moreover, when the maximum transmit power is 20 dBm, EE reaches maximum and EE of using ML generated clusters model in DAS constructed by k-means clustering algorithm and GMM clustering algorithm is about 230% and 180% higher than EE of conventional communication model in DAS. According to Fig.1 and Fig.2, we can conclude that SE and EE of DAS with ML generated clusters model are much better than the conventional communication model.

In order to gain insight into the advantages of DAS with the ML generated clusters mode, we compared the computational complexity of the conventional algorithm and the proposed algorithm based on k-means and GMM. From the existing work [16], the worst-case complexity of the interior point method reaches $O(1/\zeta)$, where ζ denotes the tolerance parameter. Therefore, the complexity of conventional algorithms using interior point method in DAS without ML generated clusters model is $O(tNKL)$, where t represents the number of iterations and L denotes $1/\zeta$. From the relevant literature [45], [46], we can get the complexity of k-means algorithm and GMM clustering algorithm are $O(t(S + K)d)$ and $O(tSKd^3)$, respectively, where S is the number of the clusters, and d is the dimension of the user samples. The computational complexity of the proposed algorithm achieved by k-means and GMM are $O(tKL + t(S + K)d)$ and $O(tKL + tSKd^3)$, respectively. According to the simulation results, the ML generated clusters mode in DAS performs a much better performance than the conventional communication model in DAS to solve the problem of maximizing SE and EE. Moreover, when the ML generated clusters mode is used in the future communication systems with more and more RAUs to provide users with high quality service, the N part will dominate the computation, which means the power distribution obtained by the ML generated clusters mode in DAS may be much better than the conventional power allocation scheme in the case of ensure low computational complexity.

VI. CONCLUSION

In this paper, we took ML generated clusters model into DAS by k-means clustering algorithm and GMM clustering algorithm, which was very different from the conventional communication model of DAS. Firstly, we divided the users into the clusters by using the ML clustering algorithms so that the clustering center of each cluster automatically selected the closest RAU to the clustering center for the users in the cluster as the users' only one serving base station. And after all the clusters selected their own serving base stations, the users can use the ML generated clusters model to communication in DAS. Under the communication scenario of DAS with ML generated clusters model, we investigated the maximum SE optimization problems and the maximum EE optimization problems under the requirements of each user's minimum SE and the constraints of the maximum transmit power of RAUs in DAS with the ML generated clusters model, respectively. Firstly, we converted the original optimization problem of maximizing SE into an equivalent problem with a D.C structure, and employed the CCCP algorithm to obtain the optimal power allocation scheme of the maximizing SE in DAS with the ML generated clusters model. Secondly, we discussed the process of obtaining the optimal solution of the maximum EE optimization problem, which included two steps. The first step was to finish the conversion from the maximum EE optimization problems to the equivalent optimization problem with subtraction structure by employing fractional programming theory, and then turned it into a solvable optimization problem with a D.C. structure. The second step was to use the CCCP algorithm to solve the optimization problem with D.C. form. Simulations results presented the SE and EE of DAS with ML generated clusters model are much better than the conventional communication model in DAS. ML algorithms provide unlimited possibilities for next generation 5G wireless networks. For the future study, the interesting area for exploration is to extend ML to more communication scenarios to solve more practical problems and achieve higher communication performance.

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. 2016.

[2] 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.

[3] X. Hong, Y. Jie, C.-X. Wang, J. Shi, and X. Ge, "Energy-spectral efficiency trade-off in virtual MIMO cellular systems," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 10, pp. 2128–2140, Oct. 2013.

[4] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "Energy and spectral efficiency of very large multiuser MIMO systems," *IEEE Trans. Commun.*, vol. 61, no. 4, pp. 1436–1449, Apr. 2013.

[5] G. I. Tsiropoulos, A. Yadav, M. Zeng, and O. A. Dobre, "Cooperation in 5G HetNets: Advanced spectrum access and D2D assisted communications," *IEEE Wireless Commun.*, vol. 24, no. 5, pp. 110–117, Oct. 2017.

[6] Q. Wu, G. Y. Li, W. Chen, and D. W. K. Ng, "Energy-efficient D2D overlaying communications with spectrum-power trading," *IEEE Trans. Wireless Commun.*, vol. 16, no. 7, pp. 4404–4419, Jul. 2017.

[7] X. Lin, J. Andrews, A. Ghosh, and R. Ratasuk, "An overview of 3GPP device-to-device proximity services," *IEEE Commun. Mag.*, vol. 52, no. 4, pp. 40–48, Apr. 2014.

[8] J. Yan, Y. Wang, G. Yang, Y. Guo, and W. Wu, "Energy efficient resource allocation in orthogonal frequency division multiple access-based distributed antenna systems," *IET Commun.*, vol. 10, no. 10, pp. 1214–1220, 2016.

[9] S.-R. Lee, S.-H. Moon, J.-S. Kim, and I. Lee, "Capacity analysis of distributed antenna systems in a composite fading channel," *IEEE Trans. Wireless Commun.*, vol. 11, no. 3, pp. 1076–1086, Mar. 2012.

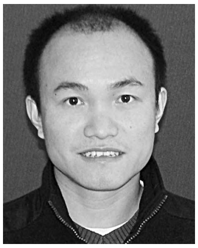
[10] C. Pan *et al.*, "Pricing-based distributed energy-efficient beamforming for MISO interference channels," *IEEE J. Sel. Areas Commu.*, vol. 34, no. 4, pp. 710–722, Apr. 2016.

[11] 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.

[12] C. He, B. Sheng, P. Zhu, X. You, and G. 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.

[13] 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.

[14] W. Choi and J. G. Andrews, "Downlink performance and capacity of distributed antenna systems in a multicell environment," *IEEE Trans. Wireless Commun.*, vol. 6, no. 1, pp. 89–92, Jan. 2007.
[15] X. Chen, X. Xu, and X. Tao, "Energy efficient power allocation in generalized distributed antenna system," *IEEE Commun. Lett.*, vol. 16, no. 7, pp. 1022–1025, Jul. 2012.
[16] 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.
[17] K. Singh, A. Gupta, and T. Ratnarajah, "Energy efficient resource allocation for multiuser relay networks," *IEEE Trans. Wireless Commun.*, vol. 16, no. 2, pp. 1218–1235, Feb. 2017.
[18] A. C. Cirik, S. Biswas, S. Vuppala, and T. Ratnarajah, "Beamforming design for full-duplex MIMO interference channels—QoS and energy-efficiency considerations," *IEEE Trans. Commun.*, vol. 64, no. 11, pp. 4635–4651, Nov. 2016.
[19] Y. Dong, H. Zhang, M. J. Hossain, J. Cheng, and V. C. M. Leung, "Energy efficient resource allocation for OFDMA full duplex distributed antenna systems with energy recycling," in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2015, pp. 1–6.
[20] J. Wang, W. Feng, Y. Chen, and S. Zhou, "Energy efficient power allocation for multicell distributed antenna systems," *IEEE Commun. Lett.*, vol. 20, no. 1, pp. 177–180, Jan. 2016.
[21] R. Heath, Jr., S. Peters, Y. Wang, and J. Zhang, "A current perspective on distributed antenna systems for the downlink of cellular systems," *IEEE Commun. Mag*, vol. 51, no. 4, pp. 161–167, Apr. 2013.
[22] X. You, C. Zhang, X. Tan, S. Jin, and H. Wu, "AI for 5G: Research directions and paradigms," *Sci. China Inf. Sci.*, vol. 62, no. 2, Feb. 2019, Art. no. 21301.
[23] D. He, C. Liu, T. Q. S. Quek, and H. Wang, "Transmit antenna selection in MIMO wiretap channels: A machine learning approach," *IEEE Wireless Commun.*, vol. 7, no. 4, pp. 634–637, Aug. 2018.
[24] J. Joung, "Machine learning-based antenna selection in wireless communications," *IEEE Commun. Lett.*, vol. 20, no. 11, pp. 2241–2244, Nov. 2016.
[25] J.-B. Wang *et al.*, "A machine learning framework for resource allocation assisted by cloud computing," *IEEE Netw.*, vol. 32, no. 2, pp. 144–151, Mar./Apr. 2018.
[26] H. Chen *et al.*, "Cost-minimized design for TWDM-PONbased 5G mobile backhaul networks," *IEEE/OSA J. Opt. Commun. Netw.*, vol. 8, no. 11, pp. B1–B11, Nov. 2016.
[27] D. An, X. Zheng, C. Rong, T. Kechadi, and C. Chen, "Gaussian mixture model based interest prediction in social networks," in *Proc. IEEE 7th Int. Conf. Cloud Comput. Technol. Sci. (CloudCom)*, Nov./Dec. 2015, pp. 196–201.
[28] Y. Liu, C. He, X. Li, C. Zhang, and C. Tian, "Power allocation schemes based on machine learning for distributed antenna systems," *IEEE Access*, vol. 7, pp. 20577–20584, 2019.
[29] S. Xu, P. Liu, R. Wang, and S. S. Panwar. (2018). "Realtime scheduling and power allocation using deep neural networks," [Online]. Available: https://arxiv.org/abs/1811.07416
[30] X.-H. You, D.-M. Wang, B. Sheng, X.-Q. Gao, X.-S. Zhao, and M. Chen, "Cooperative distributed antenna systems for mobile communications [coordinated and distributed MIMO]," *IEEE Wireless Commun.*, vol. 17, no. 3, pp. 35–43, Jun. 2010.
[31] 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.
[32] X. Li, C. He, and J. Zhang, "Spectral efficiency and energy efficiency of distributed antenna systems with virtual cells," *AEU-Int. J. Electron. Commun.*, vol. 96, pp. 130–137, Nov. 2018.
[33] D. Pollard, "Quantization and the method of k-means," *IEEE Trans. Inf. Theory*, vol. 28, no. 2, pp. 199–205, Mar. 2003.
[34] P. S. Bradley and U. M. Fayyad, "Refining initial points for k-means clustering," in *Proc. 15th Int. Conf. Mach. Learn.*, Jul. 1998, pp. 91–99.
[35] C. M. Bishop, *Pattern Recognition and Machine Learning* (Information Science and Statistics). New York, NY, USA: Springer, 2016.
[36] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *J. Roy. Statist. Soc., B, (Methodological)*, vol. 39, no. 1, pp. 1–22, 1977.
[37] D. W. K. Ng, E. S. Lo, and R. Schober, "Energy-efficient resource allocation in OFDMA systems with large numbers of base station antennas," *IEEE Trans. Wireless Commun.*, vol. 11, no. 9, pp. 3292–3304, Sep. 2012.
[38] O. Arnold, F. Richter, G. Fettweis, and O. Blume, "Power consumption modeling of different base station types in heterogeneous cellular networks," in *Proc. Future Netw. Mobile Summit*, Jun. 2011, pp. 1–8.
[39] L. T. H. An and P. D. Tao, "The DC (difference of convex functions) programming and DCA revisited with DC models of real world nonconvex optimization problems," *Ann. Oper. Res.*, vol. 133, nos. 1–4, pp. 23–46, 2005.
[40] R. G. Lanckriet and B. K. Sriperumbudur, "On the convergence of the concave-convex procedure," in *Proc. Adv. Neural Inf. Process. Syst.*, 2009, pp. 1759–1767.
[41] Y. Rui, Q. T. Zhang, L. Deng, P. Cheng, and M. Li, "Mode selection and power optimization for energy efficiency in uplink virtual MIMO systems," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 5, pp. 926–936, May 2013.
[42] W. Dinkelbach, "On nonlinear fractional programming," *Manage. Sci.*, vol. 13, no. 7, pp. 492–498, Mar. 1967.
[43] K. Nagayama, M. Kakui, M. Matsui, I. Saitoh, and Y. Chigusa, "Ultra-low-loss (0.1484 dB/km) pure silica core fibre and extension of transmission distance," *Electron. Lett.*, vol. 38, no. 20, pp. 1168–1169, Sep. 2002.
[44] S. Cui, A. J. Goldsmith, and A. Bahai, "Energy-constrained modulation optimization," *IEEE Trans. Wireless Commun.*, vol. 4, no. 5, pp. 2349–2360, Sep. 2005.
[45] X. He, D. Cai, Y. Shao, H. Bao, and J. Han, "Laplacian regularized Gaussian mixture model for data clustering," *IEEE Trans. Knowl. Data Eng.*, vol. 23, no. 9, pp. 1406–1418, Sep. 2011.
[46] G. Oliva, R. Setola, and C. N. Hadjicostis. (2013). "Distributed k-means algorithm." [Online]. Available: http://arxiv.org/abs/1312.4176

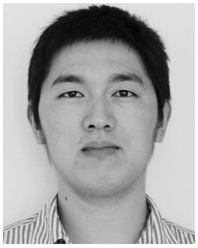


CHUNLONG HE received the M.S. degree in communication and information science from Southwest Jiaotong University, Chengdu, China, in 2010, and the Ph.D. degree from Southeast University, Nanjing, China, in 2014. From 2012 to 2014, he was a Visiting Student with the School of Electrical and Computer Engineering, 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 Associate Professor. His research interests include communication and signal processing, green communication systems, channel estimation algorithms, and limited feedback techniques. He is a member of the Institute of Electronics, Information, and Communication Engineering. He is currently an Associate Editor of the IEEE Access.

YUEHUA ZHOU received the B.S. degree in communications engineering from China West Normal University, Nanchong, China, in 2017. 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 and resource allocation.

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.

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.



DAQUAN FENG received the Ph.D. degree in information engineering from the University of Electronic Science and Technology of China, in 2015. He was a Research Staff Member with the State Radio Monitoring Center, Beijing, China, and then a Postdoctoral Research Fellow with the Singapore University of Technology and Design. From 2011 to 2014, he was a Visiting Student with the School of Electrical and Computer Engineering, Georgia Institute of Technology. Since 2016, he has been an Assistant Professor with the College of Electrical Engineering, Shenzhen University. His research interests include D2D communications, LTE-U, and the massive loT networks. He received the Excellent Doctoral Dissertation Award from the Chinese Institute of Electronics. He is currently an Associate Editor of the IEEE ACCESS.