【电子与信息科学 / Electronics and Information Science】

基于机器学习的集群双向 DAS 能效技术

何春龙, 周月华, 钱恭斌, 丁 雪

深圳大学电子与信息工程学院, 广东深圳 518060

摘 要: 双向分布式天线系统(bidirectional distributed antenna system, BDAS)是构建未来绿色通信的重要技术之一, 但其在提高能效和降低能耗的同时, 也给小区中的远程接入单元(remote access unit, RAU)与用户带来严重干扰. 本研究提出在 BDAS 中构建基于机器学习的集群通信模式. 首先, 通过对 BDAS 中的用户进行聚类分析, 将每个用户都归属一个集群; 然后, 根据集群中心与每个 RAU 的距离为集群中的用户选择通信所需的唯一服务基站; 最后, 在加入机器学习的集群 BDAS 中, 对以最大化系统能量效率(energy efficiency, EE)和频谱效率(spectral efficiency, SE)为优化目标的功率分配方案进行求解. 仿真结果表明, 相比单一的 BDAS, 基于机器学习的集群双向分布式天线系统(BDAS based on machine learning generated clusters, BDAS-MLGC)能更有效地提高系统的 SE 和 EE.

关键词: 无线通信技术; 机器学习; 频谱效率; 能量效率; k 均值; 高斯混合模型; 分布式天线系统
中图分类号: TN92 **文献标志码:** A **doi:** 10.3724/SP.J.1249.2020.06567

Machine learning generated clusters-based energy efficient power allocation for bidirectional DAS

HE Chunlong, ZHOU Yuehua, QIAN Gongbin, and DING Xue

College of Electronics and Information Engineering, Shenzhen University, Shenzhen 518060, Guangdong Province, P. R. China

Abstract: The bidirectional distributed antenna system (BDAS) is one of the important technologies for building the green communication system in the future. However, BDAS also causes serious interference to the remote access units (RAU) and users while bringing the improvement of energy efficiency and reduction of energy consumption. In this paper, we propose a communication mode based on machine learning generated clusters. Firstly, each user belongs to a cluster through clustering analysis of users in BDAS. And then, according to the distance between the cluster center and each RAU, we chose the only serving base station for users in the cluster, and finally we solve out the power allocation scheme with the optimization objectives of maximizing the system energy efficiency (EE) and spectral efficiency (SE). Simulation results show that the BDAS-MLGC, which is called the bidirectional distributed antenna system based on machine learning generated clusters, can improve SE and EE of the system more effectively than single BDAS.

Key words: wireless communication technology; machine learning; spectral efficiency; energy efficiency; k -means; Gaussian mixture model; distributed antenna system

随着移动通信技术的快速发展, 通信数据流量的爆炸式增长给未来通信网络带来了极大的挑战.

为有效面对这些挑战, 研究人员提出了虚拟小区、分布式天线系统(distributed antenna system, DAS)和设

Received: 2020-05-07; **Accepted:** 2020-08-07

Foundation: National Natural Science Foundation of China (61601300); Shenzhen Overseas High-level Talents Innovation and Entrepreneurship Project (KQJSCX20180328093835762); Shenzhen Basic Research Project (JCYJ20190808122409660)

Corresponding author: Associate professor QIAN Gongbin. E-mail: qiangb@szu.edu.cn

Citation: HE Chunlong, ZHOU Yuehua, QIAN Gongbin, et al. Machine learning generated clusters-based energy efficient power allocation for bidirectional DAS [J]. Journal of Shenzhen University Science and Engineering, 2020, 37(6): 567-575. (in Chinese)

备间通信技术等颇具创新性的方案. DAS 可显著提高系统的频谱效率(spectral efficiency, SE)和能量效率(energy efficiency, EE)^[1]. 然而, DAS 在提高通信系统能效、降低通信能耗的同时,也给小区中的远程接入单元(remote access unit, RAU)与用户带来严重干扰^[2]. 另外,在实际通信系统中,往往会同时存在上行链路和下行链路用户的双向通信^[3]. 在大规模多输入多输出(multiple-input multiple-output, MIMO)系统中,文献[4]提出可动态分配上行链路和下行链路远程无线电头数量的双向动态网络(bidirectional dynamic network, BDN)传输策略. 文献[5]分析了具有大规模 MIMO 的 BDN 中的系统性能,且通过数值仿真证明相较于动态时分双工系统,双向动态网络在频谱效率以及能量效率方面的优越性. 在大型 DAS 中,为最大化双向动态网络中的频谱效率,文献[6]为用户选择最有效的天线,提出一种新颖的分布式柔性天线集群(flexible antenna clustering, FAC)策略,并通获得比时分双工系统更好的 SE 性能. 为使 DAS 更加符合实际情况,本研究将研究对象定为双向分布式天线系统(bidirectional distributed antenna system, BDAS).

对于更复杂的第 5 代(5th generation, 5G)无线网络,研究人员已开始将机器学习应用到无线通信领域. 从系统设计和优化的角度,尤肖虎等^[7]阐述了机器学习算法在 5G 无线通信系统中主要涉及 3 类技术问题:组合优化(如资源分配)、检测(如最优接收机设计)和估计(如信道估计),并指出人工智能(artificial intelligence, AI)技术是优于传统通信技术的替代解决方案. 其中,聚类算法是无监督学习算法中的代表算法,被研究得最多,同时也是应用最广泛的算法. 为有效降低通信时延并保证通信系统的稳定性, k 均值(k -means)聚类算法被用于优化异构网络中的虚拟信道^[8]. LIANG 等^[9]运用 k -means 算法设计了 MIMO 通信系统的收发机. AL-BAKAY 等^[10]在无线通信系统中运用 k -means 算法估计频率偏移,在确保低复杂度的情况下可显著提高通信系统的效率. 在发射机信道估计中采用高斯混合模型(Gaussian mixture model, GMM)聚类算法可保证交换消息的完整性和真实性^[11]. LI 等^[12]应用 GMM 获取无线信道参数的分布,使用卷积神经网络(convolutional neural network, CNN)自动区分不同的无线信道. 本研究提出基于机器学习的集群双向分布式天线系统(bidirectional distributed antenna system based on machine learning generated clusters,

BDAS-MLGC)的功率分配方案,运用 k -means 和 GMM 集群算法构建基于机器学习的集群通信模式,并在此通信场景下探究最大化系统 SE 和 EE 的功率分配方案.

1 系统配置

BDAS-MLGC 系统模型如图 1. 该系统模型考虑了上行用户和下行用户同时存在的场景, N 个装有 1 个收发天线的服务基站(base station, BS)均匀分布在小区中. K 个装有 1 个收发天线的蜂窝用户设备随机分布在小区中. 其中,下行链路蜂窝用户的数为 K_d ,上行链路蜂窝用户的数为 K_u . 为便于分析,假设:① 服务基站不受来自它所发送和接收的信息之间的干扰;② 信道状态信息对于接收端和发送端均已知;③ 蜂窝用户共享相同的频谱资源;④ 将系统带宽归一化为 1 Hz.

图 1 BDAS-MLGC 系统模型

Fig. 1 BDAS-MLGC system model

1.1 信道模型

在 BDAS-MLGC 系统中,第 n 个服务基站和第 k 个通信用户之间的信道状态 $h_{n,k}$ 可建模为复合衰落信道^[13],数学表达式为

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

其中, $g_{n,k}$ 和 $w_{n,k}$ 分别为第 n ($n = 1, 2, \dots, N$) 个服务基站和第 k ($k = 1, 2, \dots, K$) 个蜂窝用户之间的大规模衰落系数和小规模衰落系数. $g_{n,k}$ 是一个独立同分布的零均值单位方差的循环对称复高斯随机变量. $w_{n,k}$ 可表示^[14]为

$$w_{n,k} = \sqrt{\frac{cs_{n,k}}{d_{n,k}^\alpha}} \quad (2)$$

其中, c 为以参考距离为 1 km 的中值平均路径; $d_{n,k}$ 为蜂窝用户 k 到服务基站的距离, α 为路径衰落系数,取值范围一般为 $[3, 5]$; $s_{n,k}$ 是阴影衰落系数,服从对数正态分布,即 $10\lg s_{n,k}$ 是一个高斯随机变量,均值为 0,方差为 σ^2 .

1.2 构建机器学习集群通信模式

运用集群算法构建 BDAS-MLGC 系统.

1.2.1 采用 k -means 集群算法构建机器学习集群

d 维蜂窝用户的位置坐标分布数据集 $U = \{u_i | u_i \in \mathbf{R}^d, i = 1, 2, \dots, K, d = 2\}$. 其中, u_i 为第 i 个蜂窝用户的位置坐标; K 为小区中蜂窝用户的总数; d 是用户位置坐标的维度, 本研究设 $d = 2$. k -means 集群算法^[15]的目标是找到 S 个经集群分析后的蜂窝用户集群 $C = \{c_1, c_2, \dots, c_S\}$ 的集群中心 $M = \{\mu_1, \mu_2, \dots, \mu_S\}$, 从而使每个蜂窝用户 $u_j (j = 1, 2, \dots, K)$ 与其所对应的最近的集群中心 $\mu_i (i = 1, 2, \dots, S)$ 的距离和最小, 该距离平方和为

$$E = \sum_{i=1}^S \sum_{u_j \in c_i} \|u_j - \mu_i\|_2^2 \quad (3)$$

其中, $\mu_i = \frac{1}{|c_i|} \sum_{u_j \in c_i} u_j$ 是蜂窝用户集群 c_i 的集群中心, $|c_i|$ 是集群 c_i 中的蜂窝用户数^[16]. 在一定程度上, 式(3)描述了集群内蜂窝用户围绕集群中心的紧密关系程度, E 值越小表示集群内蜂窝用户间的相似程度越高. k -means 聚类算法通过最小化式(3), 将蜂窝用户的位置坐标数据集 U 划分为预定的 S 类. 具体来说, 预先从包含 K 个蜂窝用户样本的用户位置集合 U 中选择 S 个用户作为初始用户集群的集群中心. 对于其余的蜂窝用户则根据其与初始集群中心之间的相似性(距离), 分别分配到与之最相似的蜂窝用户集群中, 然后再次计算并更新每个新获得的用户集群的集群中心. 迭代此过程至每个蜂窝用户集群的集群中心不再改变. 在 BDAS-MLGC 中使用 k -means 集群算法对蜂窝用户进行集群分析的算法伪代码请扫描论文末页右下角二维码查看.

1.2.2 高斯集群算法构建机器学习集群

与使用蜂窝用户数据原型向量来描述聚类结构的 k -means 聚类算法不同, GMM 聚类算法采用概率模式(高斯分布)描述聚类原型. 根据文献[17]描述, 可将高斯分布分量组合在一起用以构建整体概率模型, 这意味着若它足够灵活, 则可基于高斯混合分布来近似任何分布. 因此, 可使用 GMM 聚类算法, 并通过高斯混合分量逼近所描述数据的分布, 从而将 BDAS-MLGC 中的蜂窝用户数据的聚类原型描述为高斯分布的概率模式. 具体来说, 描述蜂窝用户聚类原型的高斯混合分布由 S 个混合分布分量组成, 每个分量都是一个高斯分布, 这 S 个高斯混合分布分量可表示为

$$P_M(u_i) = \sum_{i=1}^S \alpha_i P(u_i | \mu_i, \varepsilon_i) \quad (4)$$

其中, μ_i 和 ε_i 分别为蜂窝用户聚类原型的高斯混合分布的第 i 个高斯混合分量的均值和协方差矩阵; 混合系数 α_i 为选择描述蜂窝用户聚类原型的高斯混合分布中的第 i 个高斯混合分量的概率, $\alpha_i > 0$, 且 $\sum_{i=1}^S \alpha_i = 1$;

$$P(u_j | \mu_i, \varepsilon_i) = \frac{1}{(2\pi)^{\frac{d}{2}} |\varepsilon_i|^{\frac{1}{2}}} e^{-\frac{1}{2}(u_j - \mu_i)^T \varepsilon_i^{-1} (u_j - \mu_i)} \quad (5)$$

在 BDAS-MLGC 中, 给定 d 维蜂窝用户的位置坐标分布数据集 $U = \{u_j | u_j \in \mathbf{R}^d, j = 1, 2, \dots, K, d = 2\}$. 令随机变量 $Z_j \in \{1, 2, \dots, S\}$ 表示用户 u_j 的高斯混合分量. 显然, Z_j 的先验概率 $P(Z_j = i)$ 对应着 $\alpha_i, i = 1, 2, \dots, S$. 根据贝叶斯定理, Z_j 的后验分布为

$$P_M(Z_j = i | u_j) = \frac{P(Z_j = i) P_M(u_j | Z_j = i)}{P_M(u_j)} = \frac{\alpha_i P(u_j | \mu_i, \varepsilon_i)}{\sum_{l=1}^S \alpha_l P(u_j | \mu_l, \varepsilon_l)} \quad (6)$$

换句话说, $P_M(Z_j = i | u_j)$ 表示由第 i 个高斯混合分量产生蜂窝用户样本 u_j 的后验概率, 为便于描述, 记为 $r_{j,i} (i = 1, 2, \dots, S)$. 当已知高斯混合分布时, GMM 聚类算法会将蜂窝用户样本集 U 划分为 S 个集群 $C = \{c_1, c_2, \dots, c_S\}$, 并且每个蜂窝用户样本 u_j 的集群标签 λ_j 为

$$\lambda_j = \arg \max_{i \in \{1, 2, \dots, S\}} r_{j,i} \quad (7)$$

对于给定的蜂窝用户数据集 U , 最大似然估计可表示为

$$L(U) = \ln \left[\prod_{j=1}^K P_M(u_j) \right] = \sum_{j=1}^K \ln \left[\sum_{i=1}^S \alpha_i P(u_j | \mu_i, \varepsilon_i) \right] \quad (8)$$

式(8)的模型参数使用期望最大化(expectation maximize, EM)^[18]算法通过迭代优化获得的, 从而使 GMM 集群算法适应给定的蜂窝用户样本集合 U . 在 EM 算法的第 i 次迭代过程中, 每个高斯分量的混合系数为

$$\alpha_i = \frac{1}{K} \sum_{j=1}^K r_{j,i} \quad (9)$$

式(9)表示每个高斯分量的混合系数由属于该分量的用户样本的平均后验概率确定. 高斯混合分

量的均值和协方差矩阵分别为

$$\mu_i = \frac{\sum_{j=1}^K r_{j,i} u_j}{\sum_{j=1}^K r_{j,i}} \quad (10)$$

$$\varepsilon_i = \frac{\sum_{j=1}^K r_{j,i} (u_j - \mu_i)(u_j - \mu_i)^T}{\sum_{j=1}^K r_{j,i}} \quad (11)$$

由式(10)可见,描述蜂窝用户聚类原型的高斯混合分布的每个混合成分的平均值可通过样本加权平均值估算,该加权平均值是每个用户样本属于该成分的后验概率.将 GMM 集群算法运用到 BDAS-MLGC 中进行用户集群分析的过程如图 2.

```

1 初始化: 用户集合  $U$ , 其中用户数为  $K$ , 且  $S = N$ , 模型参数  $\{\alpha_i, \mu_i, \varepsilon_i, |i = 1, 2, \dots, S\}$ ;
2 repeat
3   for  $j = 1, 2, \dots, K$ 
4     根据式(7)计算  $u_j$  由各个混合成分生成的后验概率  $r_{j,i} = P_M(Z_j = i | u_j), i = 1, 2, \dots, S$ ;
5   end for
6   for  $j = 1, 2, \dots, S$ 
7     根据式(9)计算新混合系数  $\alpha'_i$ ;
8     根据式(10)计算新平均值  $\mu'_i$ ;
9     根据式(11)计算新协方差矩阵  $\varepsilon'_i$ ;
10  end for
11 将模型参数由  $\{\alpha_i, \mu_i, \varepsilon_i, |i = 1, 2, \dots, S\}$  更新为  $\{\alpha'_i, \mu'_i, \varepsilon'_i, |i = 1, 2, \dots, S\}$ ;
12 until 根据式(8)的数值收敛;
13 return  $\mu_i, i = 1, 2, \dots, S$ .
```

图 2 BDAS 中的 GMM 集群算法伪代码

Fig. 2 Pseudocode of GMM in BDAS

通过将 k -means 集群算法和 GMM 集群算法运用到 BDAS, 对用户进行集群分析后可得到每个蜂窝用户所在的集群和每个集群所在的集群中心. 然后通过集群中心选择服务基站算法得到集群唯一的服务基站, 可实现在 BDAS 中形成机器学习集群通信模式. 此过程的具体步骤为:

1) 通过机器学习中的非监督学习算法对蜂窝用户进行聚类分析. 通过调用 k -means 集群算法(算法伪代码请扫描论文末页右下角补充材料图 S1) 与高斯集群算法(图 2)对 BDAS 中的蜂窝用户进行聚类分析, 令每个上行链路或下行链路用户都找到所属集群以及对应集群的集群中心.

2) 通过蜂窝用户所在集群的中心为集群用户选择对应的服务基站. 在执行步骤 1) 得到用户集

群以及每个集群所对应的集群中心的基础上, 运用集群中心选择服务基站的算法为集群用户选择唯一的服务基站, 具体步骤见图 3.

```

1 初始化:  $S = N, \mu_i, i = 1, 2, \dots, S$ ;
2 for  $i = 1, 2, \dots, S$ 
3   计算每个聚类中心  $\mu_i$  到  $BS_n, n = 1, 2, \dots, N$  之间的距离  $d_{i,n} = \|BS_n - \mu_i\|^2$ ;
4   根据最近的基站确定集群  $c_i$  的服务基站标签  $R_i$ 
      $R_i = \arg \min_{i \in \{1, 2, \dots, S\}} d_{i,n}$ ;
5 end for
6 return  $R_i, i = 1, 2, \dots, N$ .
```

图 3 集群中心选择服务基站算法

Fig. 3 Pseudocode of the algorithm of clustering center selecting the BS

3) 为每个上行链路或下行链路蜂窝用户找到所属集群. 对于每个上行用户, 执行步骤 1) 和 2) 后找到其对应的服务基站, 再把该用户加入到其所在的机器学习用户集群 B_n^u ; 对于每个下行用户, 同样经过步骤 1) 和步骤 2) 后得到其所对应的服务基站, 再将该用户加入到其所在的机器学习用户集群 B_n^d . BDAS-MLGC 中用户选择所属服务集群从而构建基于机器学习集群的通信模式的整体算法伪代码见图 4.

```

1 初始化: 设置上行用户数  $K_u$ , 下行用户数  $K_d$ , 基站数目  $N$ , 上行和下行用户初始集群  $B_n^u = \emptyset, B_n^d = \emptyset, n = 1, 2, \dots, N$ ;
2 while  $n \leq N$ 
3   while  $k \leq K_u$ 
4     if 采用  $k$ -means 或 GMM 集群算法求得第  $k$  个上行用户所在集群, 由图 3 的集群中心选择服务器算法求得该集群用户的服务基站数为  $n$ ;
5     将第  $k$  个上行的蜂窝用户加入到其所在的机器学习集群  $B_n^u$ ;
6   end
7 end
8 while  $l \leq K_d$ 
9   if 采用  $k$ -meansGMM 聚群算法求得第  $l$  个下行用户所在集群, 由图 3 的集群中心选择服务器算法求得该集群用户的服务基站为  $n$ ;
10  将第  $l$  个下行的蜂窝用户加入到其所在的机器学习集群  $B_n^d$ ;
11 end
12 end
13 end
14 return  $B_n^u, B_n^d, n = 1, 2, \dots, N$ .
```

图 4 BDAS 中基于机器学习集群通信模式构建算法伪代码

Fig. 4 Pseudocode of constructing machine learning generated clusters in BDAS

通过将机器学习集群算法应用到 BDAS 中构建 BDAS-MLGC, 小区中的上行链路或者下行链路蜂窝用户都会有一个自己所在的集群, 集群中心通过选择与之最近的基站作为该集群用户的服务基站. 此时, 在机器学习集群中的上行用户通信时会隔离同一机器学习集群之中的其他下行用户的干扰, 而下行用户通信时会隔离同一机器学习集群之中的其他上行用户的干扰. 在 BDAS-MLGC 中, 第 n 个机器学习集群服务的上行蜂窝用户数为 b_n^u , 下行蜂窝用户数为 b_n^d , $n = 1, 2, \dots, N$. 当在 BDAS 中形成基于机器学习集群的通信模式后, 每个被选择的服务基站都有其固定需要服务的上行链路和下行链路用户. 因此, 第 n 个服务基站 BS_n 服务第 k 个上行蜂窝用户的 SE 为

$$R_k^u = \text{lb} \left(1 + \frac{p_k^n |h_{k,n}|^2}{\gamma_u^M} \right) \quad (12)$$

其中, p_k^n 为第 n 个机器学习集群中第 k 个上行蜂窝用户向该集群的服务基站发射消息时的发射功率. 分母 γ_u^M 为

$$\gamma_u^M = \sum_{\substack{i=1, \\ i \neq k}}^{b_n^u} p_i^M |h_{k,n}|^2 + \sum_{\substack{m=1, \\ m \neq n}}^N \left(\sum_{j=1}^{b_m^u} p_j^M |h_{j,n}|^2 + \sum_{t=1}^{b_m^d} p_{m,t}^M |h_{m,n}|^2 \right) + \sigma^2 \quad (13)$$

这里, $p_{m,t}^M$ 为第 m 个机器学习集群中服务基站向该机器学习集群中第 t 个下行蜂窝用户发射功率; $h_{m,n}$ 为第 m 个用户与第 n 个服务基站之间的信道状态. 因此, 第 n 个机器学习集群中第 l 个下行蜂窝用户的 SE 为

$$R_l^d = \text{lb} \left(1 + \frac{p_{n,l}^M |h_{l,n}|^2}{\gamma_d^M} \right) \quad (14)$$

$$\gamma_d^M = \sum_{i=1, i \neq l}^{b_n^d} p_{n,i}^M |h_{l,n}|^2 + \sum_{m=1, m \neq n}^N \sum_{j=1}^{b_m^d} p_{m,j}^M |h_{l,m}|^2 + \sum_{n=1}^N \sum_{t=1}^{b_n^u} p_t^M |h_{l,t}|^2 + \sigma^2 \quad (15)$$

其中, $h_{l,t}$ 为第 t 个上行蜂窝用户与第 l 个下行蜂窝用户之间的信道状态.

因此, BDAS-MLGC 的总 SE 为

$$R_{\text{total}}^M = \sum_{k=1}^{K_u} R_k^u + \sum_{l=1}^{K_d} R_l^d \quad (16)$$

1.2.3 总功率消耗模型

由文献[19]可得, BDAS-MLGC 中总功率消耗 P_{total} 包含发射功率消耗和额外的电路功耗两部分.

令 P_{trans} 表示在 BDAS 中采用机器学习集群通信模式下系统的总发射功率, 则

$$P_{\text{trans}} = \sum_{k=1}^{K_u} p_k^M + \sum_{n=1}^N \sum_{l=1}^{b_n^d} p_{n,l}^M \quad (17)$$

额外的电路功耗 P_{circuit} 包括与实际发射功率无关的动态功耗 P_{dy} 、与处理单元以及电池有关的恒定静态功耗 P_{st} 和远程接入单元间连接光纤电路传输时的功耗 P_0 . 因此, 在 BDAS-MLGC 中总功率消耗模 P_{total} 可被改写^[20]成

$$P_{\text{total}} = \frac{P_{\text{trans}}}{\tau} + P_{\text{circuit}} = \frac{P_{\text{trans}}}{\tau} + (N + K_u)P_{\text{dy}} + P_{\text{st}} + P_0 \quad (18)$$

其中, τ 是射频功率放大器的效率.

2 功率分配

2.1 最大化基于 BDAS-MLGC 的 SE

BDAS-MLGC 即满足下述 3 个约束条件的最大化系统 SE 的功率分配优化问题: ① 上下行链路蜂窝用户的最小 SE 限制; ② 服务基站的功率要求; ③ 上行链路蜂窝用户最大发射功率要求. 结合式(17), 可把机器学习集群通信模式下最大化系统 SE 的优化问题建模为

$$\max_{P_{M_u}, P_{M_d}} R_{\text{total}}^M \quad (19)$$

$$\text{s. t. } R_k^u \geq R_{\min}^u, \forall k \in [1, K_u] \quad (19a)$$

$$p_k^M \leq p_{\max}^u, \forall k \in [1, K_u] \quad (19b)$$

$$R_l^d \geq R_{\min}^d, \forall l \in [1, K_d] \quad (19c)$$

$$\sum_{l=1}^{b_n^d} p_{n,l}^M \geq p_{\max}^d, \forall n \in [1, N] \quad (19d)$$

其中, $P_{M_d} = \{p_{n,l}^M, n = 1, 2, \dots, N, l = 1, 2, \dots, K_d\}$, $P_{M_u} = \{p_k^M, k = 1, 2, \dots, K_u\}$; $p_{\max}^u$ 为上行链路用户的最大发射功率; $p_{\max}^d$ 为服务基站的功率要求; $R_{\min}^u$ 和 $R_{\min}^d$ 分别为上行用户和下行用户的最小 SE 限制要求.

首先, 定义 $P_1 = [P_{M_u}, P_{M_d}]$ 为优化变量, $f^{\text{SE}}(P_1)$ 为目标函数. 再由(16)可知, 优化问题(19)仍具有非凸性及非线性特征, 因此, 运用基于凸函数差分(difference of convex function, DC)^[21]优化理论对式(19)进行转化为两个凹凸部分的和的形式, 可以表述为

$$f^{\text{SE}}(P_1) = f_{\text{cave}}^{\text{SE}}(P_1) + f_{\text{vex}}^{\text{SE}}(P_1) \quad (20)$$

其中,

$$f_{\text{cave}}^{\text{SE}}(\mathbf{P}_1) = \text{lb} \left\{ \sum_{n=1}^N \sum_{i=1}^{b_n^u} p_{i,n}^M |h_{i,n}|^2 + \sum_{m=1, m \neq n}^N \sum_{t=1}^{b_m^d} p_{m,t}^M |h_{m,n}|^2 + \sigma^2 \right\} + \text{lb} \left\{ \sum_{n=1}^N \sum_{i=1}^{b_n^d} p_{n,i}^M |h_{i,n}|^2 + \sigma^2 + \sum_{m=1}^N \sum_{t=1}^{b_m^u} p_{t,m}^M |h_{t,l}|^2 \right\} \quad (21)$$

$$f_{\text{vex}}^{\text{SE}}(\mathbf{P}_1) = -\text{lb} \gamma_u^M - \text{lb} \gamma_d^M \quad (22)$$

用 C_1 表示式(19)中的约束集, 它表明满足最小 SE 限制的约束条件(19a)和(19c)是非线性, 因此可将(19a)和(19c)转换为如式(23)和式(24)的线性表达式.

$$p_k^M |h_{k,n}|^2 - (2^{R_{\min}^u} - 1) \gamma_u^M \geq 0 \quad (23)$$

$$p_{n,l}^M |h_{l,n}|^2 - (2^{R_{\min}^d} - 1) \gamma_d^M \geq 0 \quad (24)$$

在 BDAS-MLGC 中, 满足上述约束条件的最大化系统 SE 的优化问题, 可转化为具有凸约束集的 DC 优化问题, 即

$$\begin{aligned} \max_{\mathbf{P}_1 \in C_1} f^{\text{SE}}(\mathbf{P}_1) &= f_{\text{cave}}^{\text{SE}}(\mathbf{P}_1) + f_{\text{vex}}^{\text{SE}}(\mathbf{P}_1) \\ \text{s. t.} \quad &\text{式(19a)一式(19d)} \end{aligned} \quad (25)$$

当式(25)存在偏导数时, 可采用凹凸过程 (concave convex procedure, CCCP) 优化算法^[22] 解决该优化问题. 在算法的每次迭代中, 采用一阶泰勒展开将 $f_{\text{vex}}^{\text{SE}}(\mathbf{P}_1)$ 转换为线性表达式, 求解为

$$\mathbf{P}_1^{(j+1)} = \arg \max_{\mathbf{P}_1 \in C_1} (f_{\text{cave}}^{\text{SE}}(\mathbf{P}_1) + \nabla f_{\text{vex}}^{\text{SE}}(\mathbf{P}_1^{(j)}) \times \mathbf{P}_1^T) \quad (26)$$

其中, j 为迭代次数; $\nabla f_{\text{vex}}^{\text{SE}}(\mathbf{P}_1^{(j)})$ 为函数 $f_{\text{vex}}^{\text{SE}}(\mathbf{P}_1)$ 在 $\mathbf{P}_1 = [\mathbf{P}_{M_u}^{(j)}, \mathbf{P}_{M_d}^{(j)}]$ 处的梯度. BDAS-MLGC 中, 最大化系统 SE 的优化算法伪代码如图 5.

```

1 初始化: 设  $j = 0$ ,  $\mathbf{P}_1^{(0)} \in C_1$ , 容忍度  $\zeta > 0$ ;
2 do
3  $\mathbf{P}_1^{(j+1)} = \arg \max_{\mathbf{P}_1 \in C_1} (f_{\text{cave}}^{\text{SE}}(\mathbf{P}_1) + \nabla f_{\text{vex}}^{\text{SE}}(\mathbf{P}_1^{(j)}) \times \mathbf{P}_1^T)$ ;
4 选择内点法求解上述凸优化问题;
5 利用对数障碍函数实现问题转化, 拟牛顿法获取搜索方向, 用 Armijo 准则中线性回归算法获取最优步长;
6  $j = j + 1$ ;
7 while  $\|\mathbf{P}_1^{(j+1)} - \mathbf{P}_1^{(j)}\| > \zeta$ 
8 return  $\mathbf{P}_1^{(j+1)}$ .
```

图 5 BDAS-MLGC 中最大化 SE 算法伪代码

Fig. 5 Pseudocode of maximize SE in BDAS-MLGC

2.2 最大化 BDAS-MLGC 的 EE

文献[23]将能量效率定义为频谱效率与消耗总功率的比值. 因此, BDAS-MLGC 的能量效率为

$$\text{EE}_M = \frac{R_{\text{total}}^M}{\frac{P_{\text{trans}}}{\tau} + P_{\text{circuit}}} \quad (27)$$

BDAS-MLGC 是在满足以下 3 个约束条件的基础上实现最大化系统 EE 的功率优化问题. 这 3 个约束条件为: ① 蜂窝用户的最小 SE 限制; ② 基站的最大发射功率限制; ③ 用户的最大发射功率限制. 结合式(18)和式(27)可把机器学习集群通信模式下最大化系统 EE 的优化问题建模为

$$\max_{\mathbf{P}_2} \text{EE}_M = \frac{R_{\text{total}}^M}{\frac{P_{\text{trans}}}{\tau} + (N + K_u)P_{\text{dy}} + P_{\text{st}} + P_0} \quad (28)$$

$$\text{s. t.} \quad R_k^M \geq R_{\min}^u, \forall k \in [1, K_u] \quad (28a)$$

$$p_k^M \leq p_{\max}^u, \forall k \in [1, K_u] \quad (28b)$$

$$R_l^M \geq R_{\min}^d, \forall l \in [1, K_d] \quad (28c)$$

$$\sum_{l=1}^{b_n^d} p_{n,l}^M \geq p_{\max}^d, \forall n \in [1, N] \quad (28d)$$

其中, $\mathbf{P}_2 = [P_{M_u}, P_{M_d}]$ 为优化变量. 由式(28)可得, BDAS-MLGC 中的最大化系统 EE 问题是一个具有分式形式的优化问题, 运用分式规划 (fractional programming, FP) 理论^[24], 可将具有分式形式的优化问题式(28)转化为含有减法结构的等价优化问题, 即

$$\begin{aligned} \arg \max_{\mathbf{P}_2} h_1(\mathbf{P}_2, \omega_1) &= R_{\text{total}}^M - \omega_1 \left\{ \frac{P_{\text{trans}}}{\tau} + (N + K_u)P_{\text{dy}} + P_{\text{st}} + P_0 \right\} \\ \text{s. t.} \quad &\text{式(28a)一式(28d)} \end{aligned} \quad (29)$$

其中, ω_1 是一个被引入的标量. 当定理 1 被满足时, 任何一个具有类似分式规划的函数都可被等价近似为一个具有减的形式非分式规划问题^[24].

定理 1 令 $G_2(\omega_1) = \max_{\mathbf{P}_2} h_2(\mathbf{P}_2, \omega_1)$ 和 $g_2(\omega_1) = \arg \max_{\mathbf{P}_2} h_1(\mathbf{P}_2, \omega_1)$. 当且仅当 $g_2(\omega_1^*) = \mathbf{P}_2^*$ 和 $G_2(\omega_1^*) = 0$ 时, 式(28)的最优功率分配 $\mathbf{P}_2^*$ 也是式(29)中的最优解.

由定理 1 可知, 式(28)和式(29)之间的等价性^[25]. 用 FP 理论调整后得到的等价优化问题(29)可再次被重写为如式(30)的带有 DC 结构的优化问题.

$$\begin{aligned} \max_{P_2} \quad & \{f_{\text{cave}}^{\text{EE}}(P_2) + f_{\text{vex}}^{\text{EE}}(P_2)\} \\ \text{s. t.} \quad & \text{式 (28a) — 式 (28d)} \end{aligned} \quad (30)$$

式(30)中的凹函数和凸函数部分可分别写成

$$f_{\text{cave}}^{\text{EE}}(P_2) = \text{lb}(p_k^M |h_{k,n}|^2 + \gamma_u^M) + \text{lb}(p_{n,l}^M |h_{l,n}|^2 + \gamma_d^M) \quad (31)$$

$$f_{\text{vex}}^{\text{EE}}(P_2) = -\omega_1 \left[\frac{P_{\text{trans}}}{\tau} + (N + K_u)P_{\text{dy}} + P_{\text{st}} + P_0 \right] - \text{lb} \gamma_u^M - \text{lb} \gamma_d^M \quad (32)$$

通过原始的最大化系统 EE 的优化问题转化为带 DC 优化结构的问题, 同时将原始问题的约束条件转化为凸集 C_2 后, 因 $f_{\text{vex}}^{\text{EE}}(P_2)$ 具有偏导数, 可同样使用 CCCP 算法解决目标问题(28), 描述特定的迭代步骤为

$$P_2^{(t+1)} = \arg \max_{P_2 \in C_2} (f_{\text{cave}}^{\text{EE}}(P_2) + \nabla f_{\text{vex}}^{\text{EE}}(P_2^{(t)}) \times P_2^T) \quad (33)$$

其中, t 为迭代次数; $\nabla f_{\text{vex}}^{\text{EE}}(P_2^{(t)})$ 为函数 $f_{\text{vex}}^{\text{EE}}(P_2)$ 在 $P_2^{(t)}$ 处的梯度. 根据迭代步骤, 可获得 BDAS-MLGC 中以最大化系统 EE 为优化目标最佳功率分配算法伪代码如图 6.

```

1 初始化: 设  $t = 0$ ,  $P_2^{(0)} \in C_2$ , 容忍度  $\zeta > 0$ ;  $\omega_1^{(0)} = 0.01$ ;
2 do
3    $P_2^{(t+1)} = \arg \max_{P_2 \in C_2} (f_{\text{cave}}^{\text{EE}}(P_2) + \nabla f_{\text{vex}}^{\text{EE}}(P_2^{(t)}) \times P_2^T)$ ;
4 选择内点法求解上述公式;
5 用对数障碍函数实现问题转化, 拟牛顿法获取搜索方向, 用 Armijo 准则中的线性回归算法获取最优步长;
6 根据求得的  $P_2^{(t+1)}$  更新  $\omega_1^{(t+1)} = \text{EE}|_{P_2^{(t+1)}}$ ;
7  $t = t + 1$ ;
8 until  $|h(P_2^{(t+1)}), \omega_1^{(t+1)}| < \zeta$ , 其中  $h(P_2^{(t+1)}, \omega_1^{(t+1)}) = \max_{P_2^{(t+1)} \in C_2} R_{\text{total}}^M - \omega_1^{(t+1)} \left\{ \frac{P_{\text{trans}}}{\tau} + (N + K_u)P_{\text{dy}} + P_{\text{st}} + P_0 \right\}$ ;
9 return  $P_{\text{EE}}^{(t+1)}, \omega_1^{(t+1)}$ .

```

图 6 BDAS-MLGC 中最大化系统 EE 算法伪代码

Fig. 6 Pseudocode of maximize EE in BDAS-MLGC

3 仿真结果

将文献[23]的单一 BDAS 下功率优化算法作为 BDAS-MLGC 算法的对比算法. 表 1 列出了仿真时系统的参数设置. 设小区中心基站为中央处理基站, 其他基站坐标为 $(d, 2n(n-1)/(N-1))$. 其中, $n = 1, 2, \dots, N-1$; $d = (3 - \sqrt{3}R/2)$.

表 1 BDAS-MLGC 仿真参数

Table 1 Simulation parameters of BDAS-MLGC

参 数	取 值
蜂窝小区半径 R/km	1
路径损耗指数 α	4
功放效率 $\tau/\%$	38
噪声功率 σ^2/dBm	-104
基站最大发射功率 $P_{\text{max}}^d/\text{dBm}$	[10, 30]
动态消耗功率 P_{dy}/dBm	30
上行链路用户数 K_u	50
服务基站数 N	5
静态消耗功率 P_{st}/dBm	40
光纤消耗功率 P_0/km^{-1}	0.148 4
用户最小频谱效率 R_{min}^u 和 $R_{\text{min}}^d / (\text{Mbit} \cdot \text{s}^{-1})$	1
下行链路用户数 K_d	50
上行用户最大发射功率 P_{min}^u	[3.3, 17.5]

图 7 显示了系统的 SE 随最大发射功率从 10 dBm 增至 30 dBm 变化的情况. 由图 7 可见, 在 BDAS 中加入基于机器学习集群的通信模式后, 系统 SE (R_{total}^M) 比单一的 BDAS 的 SE 高得多. 随着最大发射功率的逐渐增加, 虽然两种通信模式下的系统 SE 皆呈现增长趋势, 但是单一的 BDAS 的 SE 的增长趋势远远低于加入基于机器学习集群通信模式后系统 SE 的增长趋势, 这在发射功率较大时尤为明显. 例如, 当最大发射功率为 30 dBm 时, 通过 k -means 聚类算法和 GMM 聚类算法构建的机器学习集群通信模式下系统的 SE 大约比单一的 BDAS 高近 148% 和 109%, 这表示在 BDAS 中使用机器学习

图 7 系统 SE 随最大发射功率增大变化曲线

Fig. 7 SE versus the maximum transmit power

集群通信模式可以提高系统的 SE, 尤其是在最大发射功率较大时. 综上可知, BDAS-MLGC 可有效提升系统的 SE.

图 8 显示了系统的 EE 随最大发射功率从 10 dBm 增至 30 dBm 变化的情况. 从整体上看, 两种通信模式下系统的 EE 都随服务基站的最大发射功率的增大呈先增而后降趋势. 但是在 BDAS 中采用基于机器学习的集群通信模式后系统的 EE 比于单一的 BDAS 的 EE 有了很大的提高. 由图 8 可知, 当服务基站的发射功率较小时 BDAS-MLGC 的 EE 增长速度非常快, 且当服务基站的发射功率为 20 dBm 时, 系统的 EE 达到最大值, 而单一的 BDAS 的 EE 的增长速率非常缓慢. 另外, 在 BDAS 中采用 k -means 集群算法和 GMM 集群算法后得到的 EE 皆高于单一的 BDAS. 例如, 当服务基站的最大发射功率为 20 dBm 时, 使用 k -means 聚类算法和 GMM 聚类算法构建的 BDAS-MLGC 实现的系统的 EE 比单一的 BDAS 的系统 EE 分别高出近 35% 和 34%. 可见, BDAS-MLGC 可有效提升系统的 EE.

图 8 系统 EE 随最大发射功率增大变化曲线

Fig. 8 EE versus the maximum transmit power

为更深入地了解在 BDAS 中使用机器学习算法的优势, 比较在单一的 DAS 中和基于机器学习的 DAS 中进行功率分配时算法的复杂度. 由文献[26]可知, 内点法在最坏情况下复杂度达到 $O(1/\zeta)$. 其中, ζ 为误差容忍度参数. 因此, 在单一的 BDAS 中进行功率分配时, 使用内点法的算法复杂度为 $O(tNKL)$. 其中, t 为迭代次数; $L = 1/\zeta$; N 为服务基站数目; K 为小区中蜂窝用户总数. 分别从相关文献[27-28]中, 可得出 k -means 聚类算法和 GMM 聚类算法的复杂度分别为 $O(t(S + K)d)$ 和

$O(tSKd^3)$, 其中, S 为集群数目; d 为用户坐标位置的维度. 由 k -means 和 GMM 实现的基于机器学习集群算法的计算复杂度为 $O(tKL + t(S + K)d)$ 和 $O(tKL + tSKd^3)$. 结果表明, 在 BDAS 中使用机器学习集群算法能够极大程度地提高系统的通信性能.

结 语

在 BDAS 中, 上行用户和下行用户的同时存在导致通信系统中存在更多的干扰, 同时复杂的干扰管理会给中心处理基站带来沉重的计算负担. 为此, 本研究将机器学习算法引入 BDAS 中构建 BDAS-MLGC. 首先, 引进集群算法对蜂窝用户进行集群分析, 从而得到相应的集群以及集群中心; 然后, 按照集群中心选择服务基站算法来确定每个集群的服务基站, 同时对应的上行或下行用户会被分配到相应的机器学习集群, 这样就构建了基于机器学习的集群通信模式. 在此通信模式下, 讨论在满足一定约束条件下的最大化系统 SE 和 EE 的功率分配方案. 仿真结果表明, 相比单一的 BDAS, BDAS-MLGC 显著提高了系统的通信性能.

基金项目: 国家自然科学基金资助项目 (61601300); 深圳市海外高层次人才创新创业专项资助项目 (KQJSCX20180328 093835762); 深圳市基础研究计划资助项目 (JCYJ2019 0808122409660)

作者简介: 何春龙 (1984—), 深圳大学副教授、博士. 研究方向: 分布式天线能效、机器学习及信道估计等. E-mail: hclong@szu.edu.cn

引 文: 何春龙, 周月华, 钱恭斌, 等. 基于机器学习的集群双向 DAS 能效技术[J]. 深圳大学学报理工版, 2020, 37 (6): 567-575.

参考文献 / References:

- [1] WEI Zhongxiang, SUN Sumei, ZHU Xu, et al. Energy-efficient hybrid duplexing strategy for bidirectional distributed antenna systems [J]. IEEE Transactions on Vehicular Technology, 2018, 67(6): 5096-5110.
- [2] HE Anqi, WANG Lifeng, CHEN Yue, et al. Spectral and energy efficiency of uplink D2D underlaid massive MIMO cellular networks [J]. IEEE Transactions on Communications, 2017, 65(9): 3780-3793.
- [3] MARKO H. The bidirectional communication theory: a generalization of information theory [J]. IEEE Transactions on Communications, 1973, 21(12): 1345-1351.
- [4] XIN Yuanxue, WANG Dongming, ZHANG Rongqing, et al. Spectral efficiency of bidirectional dynamic networks with massive MIMO [C]// IEEE Global Communications Conference. Washington D C: IEEE, 2016. doi: 10.1109/GLOCOM.2016.7842288
- [5] XIN Yuanxue, YANG Liuqing, WANG Dongming, et al.

- Bidirectional dynamic networks with massive MIMO; performance analysis [J]. IET Communications, 2017, 11(4): 468-476.
- [6] XIN Yuanxue, ZHANG Rongqing, WANG Dongming, et al. Antenna clustering for bidirectional dynamic network with large-scale distributed antenna systems [J]. IEEE Access, 2017, 5: 4037-4047.
- [7] 尤肖虎, 张川, 谈晓思, 等. 基于 AI 的 5G 技术——研究方向与范例 [J]. 中国科学: 信息科学, 2018, 48(12): 1589-1602.
YOU Xiaohu, ZHANG Chuan, TAN Xiaosi, et al. AI for 5G: research directions and paradigms [J]. Scientia Sinica Informations, 2018, 48(12): 1589-1602. (in Chinese)
- [8] XIA Ming, OWADA Y, INOUE M, et al. Optical and wireless hybrid access networks: design and optimization [J]. Journal of Optical Communications and Networking, 2012, 4(10): 749-759.
- [9] LIANG Hanwen, CHUNG W H, KUO S Y. Coding-aided k -means clustering blind transceiver for space shift keying MIMO systems [J]. IEEE Transactions on Wireless Communications, 2016, 15(1): 103-115.
- [10] ALBAKAY N, HEMPEL M, ALAHMAD M, et al. Novel frequency offset estimation scheme for reliable wireless communication using modified k -means clustering [C]// The 15th International Wireless Communications & Mobile Computing Conference (IWCMC). Tangier, Morocco: IEEE, 2019: 826-829.
- [11] ANDREAS W, MICHAEL K, JI Lianghai, et al. Physical layer authentication for mission critical machine type communication using Gaussian mixture model based clustering [C]// Proceedings of IEEE 85th Vehicular Technology Conference (VTC-Spring). Sydney, Australia: IEEE, 2017: 1-5.
- [12] LI Haihan, LI Yunzhou, ZHOU Shidong, et al. Wireless channel feature extraction via GMM and CNN in the tomographic channel model [J]. Journal of Communications & Information Networks, 2017, 2(1): 41-51.
- [13] YOU Xiaohu, WANG Dongming, SHENG Bin, et al. Cooperative distributed antenna systems for mobile communications [J]. IEEE Wireless Communications, 2010, 17(3): 35-43.
- [14] HE Chunlong, SHENG Bin, ZHU Pengcheng, et al. Energy- and spectral-efficiency tradeoff for distributed antenna systems with proportional fairness [J]. IEEE Journal on Selected Areas in Communications, 2013, 31(5): 894-902.
- [15] POLLARD D. Quantization and the method of k -means [J]. IEEE Transaction Information Theory, 2003, 28(2): 199-205.
- [16] BRADLEY P S, FAYYAD U M. Defining initial points for k -means clustering [C]// Proceedings of the 15th International Conference on Machine Learning. San Francisco, USA: Morgan Kaufmann Publishers Inc, 1998: 91-99.
- [17] BISHOP C M. Pattern recognition and machine learning [M]. New York, USA: Springer-Verlag, 2016.
- [18] DEMPSTER A P, LAIRD N M, RUBIN D B. Maximum likelihood from incomplete data via the EM algorithm [J]. Journal of the Royal Statistical Society: Series B (Methodological), 1977, 39(1): 1-22.
- [19] NG D W K, LO E S, SCHÖBER R. Energy-efficient resource allocation in OFDMA systems with large numbers of base station antennas [J]. IEEE Transaction Wireless Communication, 2012, 11(9): 3292-3304.
- [20] WU Jiang, LIU Jianwei, LI Wanlin, et al. Low-complexity power allocation for energy efficiency maximization in DAS [J]. IEEE Communications Letters, 2015, 19(6): 925-928.
- [21] AN L T H, TAO P D. The DC (difference of convex functions) programming and DCA revisited with DC models of real world nonconvex optimization problems [J]. Annals of Operations Research, 2005, 133(4): 23-46.
- [22] LANCKRIET R G, SRIPERUMBUDUR B K. On the convergence of the concave-convex procedure [C]// Proceedings in of the 22nd International Conference on Neural Information Processing Systems. New York, USA: Curran Associates Inc, 2009: 1759-1767.
- [23] LI Xinquan, HE Chunlong, ZHANG Jihong. Spectral efficiency and energy efficiency of bidirectional distributed antenna systems with user centric virtual cells [J]. IEEE Access, 2018, 6: 49886-49895.
- [24] DINKELBACH W. On nonlinear fractional programming [J]. Management Sciences, 1967, 13(7): 492-498.
- [25] YOU Xiaohu, WANG Dongming, ZHU Pengcheng, et al. Cell edge performance of cellular mobile systems [J]. IEEE Journal on Selected Areas in Communications, 2011, 29(6): 1200-1213.
- [26] HE Chunlong, LI G Y, ZHENG Fuchun, et al. Power allocation criteria for distributed antenna systems [J]. IEEE Transactions on Vehicular Technology, 2015, 64(11): 5083-5090.
- [27] HE Xiaofei, CAI Deng, SHAO Yuanlong, et al. Laplacian regularized Gaussian mixture model for data clustering [J]. IEEE Transactions on Knowledge and Data Engineering, 2011, 23(9): 1406-1418.
- [28] OLIVA G, SETOLA R, HADJICOSTIS C N. Distributed k -means algorithm [EB/OL]. (2013-12-15) [2014-11-10]. <http://arxiv.org/abs/1312.4176>.

【中文责编: 英子; 英文责编: 木柯】

补充材料

<http://journal.szu.edu.cn>