

控制与决策

Control and Decision

基于深度无监督学习的多小区蜂窝网资源分配方法

孙明, 王淑梅, 郭媛, 曹伟, 徐耀群

引用本文:

孙明, 王淑梅, 郭媛, 曹伟, 徐耀群. 基于深度无监督学习的多小区蜂窝网资源分配方法[J]. *控制与决策*, 2022, 37(9): 2333–2342.

在线阅读 View online: <https://doi.org/10.13195/j.kzyjc.2020.1579>

您可能感兴趣的其他文章

Articles you may be interested in

[基于深度强化学习的微电网在线优化调度](#)

Online optimal scheduling of a microgrid based on deep reinforcement learning

控制与决策. 2022, 37(7): 1675–1684 <https://doi.org/10.13195/j.kzyjc.2021.0835>

[基于李雅普诺夫随机优化的车辆边缘计算资源管理](#)

Resource management of vehicle edge computing based on Lyapunov stochastic optimization

控制与决策. 2022, 37(3): 721–728 <https://doi.org/10.13195/j.kzyjc.2020.1211>

[认知智能电网中基于能效优化的频谱分配策略](#)

Spectrum allocation strategy based on energy efficiency optimization in cognitive smart grid

控制与决策. 2021, 36(8): 1901–1910 <https://doi.org/10.13195/j.kzyjc.2019.1448>

[异构网络中基于鸽群优化算法的D2D资源分配机制](#)

Resource allocation for D2D based on pigeon-inspired optimization algorithm in heterogeneous networks

控制与决策. 2020, 35(12): 2959–2967 <https://doi.org/10.13195/j.kzyjc.2019.0526>

[基于改进堆叠自动编码器的循环冷却水系统工艺介质温度预测控制方法](#)

Predictive control method of process medium temperature in circulating cooling water system based on improved stacked auto encoders

控制与决策. 2020, 35(12): 2835–2844 <https://doi.org/10.13195/j.kzyjc.2019.0694>

基于深度无监督学习的多小区蜂窝网资源分配方法

孙明^{1†}, 王淑梅^{1,2}, 郭媛¹, 曹伟¹, 徐耀群²

(1. 齐齐哈尔大学 计算机与控制工程学院, 黑龙江 齐齐哈尔 161006;

2. 哈尔滨商业大学 计算机与信息工程学院, 哈尔滨 150028)

摘要: 针对多小区蜂窝网络资源分配所要求的低能耗、高速率和低延时问题, 提出一种基于深度无监督学习的多小区蜂窝网络资源分配方法. 首先, 构建基于无监督学习的深度功率控制神经网络, 通过约束处理输出优化的信道功率控制方案以最大化能量效率的期望; 然后, 构建基于无监督学习的深度信道分配神经网络, 通过约束处理输出优化的信道分配方案, 并联合前期训练好的深度功率控制神经网络拟合输出优化的信道功率, 进一步优化能量效率的期望. 仿真结果表明, 所提出的方法在保证低计算时延的同时可获得优于其他算法的能量效率和传输速率.

关键词: 蜂窝网; 资源分配; 无监督学习; 深度神经网络; 能量效率

中图分类号: TP273

文献标志码: A

DOI: 10.13195/j.kzyjc.2020.1579

引用格式: 孙明, 王淑梅, 郭媛, 等. 基于深度无监督学习的多小区蜂窝网资源分配方法[J]. 控制与决策, 2022, 37(9): 2333-2342.

Deep unsupervised learning based resource allocation method for multi-cell cellular networks

SUN Ming^{1†}, WANG Shu-mei^{1,2}, GUO Yuan¹, CAO Wei¹, XU Yao-qun²

(1. College of Computer and Control Engineering, Qiqihar University, Qiqihar 161006, China; 2. School of Computer and Information Engineering, Harbin University of Commerce, Harbin 150028, China)

Abstract: Aiming at the problem of low energy consumption, high speed, and low latency required for resource allocation for multi-cell cellular networks, a deep unsupervised learning based resource allocation method is proposed. Firstly, an unsupervised learning based deep power control neural network is constructed to output an optimized channel power control scheme by constraint handling, so as to maximize the expectation of energy efficiency. Then, an unsupervised learning based deep channel allocation neural network is constructed to output an optimized channel allocation scheme by constraint handling, and the unsupervised learning based deep power control neural network trained well previously is combined to fit and output the optimized channel power control scheme to further optimize the expectation of energy efficiency. The simulation results show that the proposed method can obtain better transmission rate and energy efficiency than other algorithms while ensuring low computational latency.

Keywords: cellular networks; resource allocation; unsupervised learning; deep neural networks; energy efficiency

0 引言

无线频谱资源一直是移动通信的紧缺资源^[1-2]. 新一代无线通信网络促使频谱资源优化目标日益趋向多样化、综合化, 并对高速率、低能耗和低延时提出了越来越高的要求^[1,3-8]. 如何高效地管理新一代无线通信的频谱资源, 仍然是需要研究的重点问题.

在蜂窝网络的资源分配上, 目前已提出了迭代优

化法、启发式算法等多种方法和理论. 尽管这些方法在蜂窝网的资源分配上具有各自的优越性, 但它们均属于计算密集型算法^[8], CPU 占用率高、计算耗时长; 同时, 它们还缺乏应对动态变化环境的能力, 需要频繁地计算来应对信道的衰落变化, 既影响通信系统的实时性, 又增加通信系统的计算负担, 无法适用于新一代的无线通信网络.

深度学习在线推断具有计算复杂度低及延时低

收稿日期: 2020-11-16; 录用日期: 2021-07-06.

基金项目: 国家自然科学基金项目(61872204, 71803095); 全国统计科学研究项目(2020LY074); 黑龙江省自然科学基金联合引导项目(LH2019F038); 黑龙江省省属高校青年创新人才基本科研业务项目(135309340).

†通讯作者. E-mail: snogisunming@163.com.

等优点^[4,9-10],已被用于解决无线移动通信的资源分配问题.例如,在深度强化学习方面,文献[4]根据增广拉格朗日乘法构建深度强化学习网络,提出了多小区蜂窝网络的频谱效率、能量效率的多目标优化方法.在深度监督学习方面,文献[8]将连续的功率域离散为多个功率等级,采用遗传算法为最大化系统吞吐量问题提供信道分配和功率等级的标签数据,然后训练堆叠自动编码器逼近标签数据,但由于该方法对功率进行了离散化处理,由其训练所得的深度学习模型的优化性能也将受到限制.在深度无监督学习方面,文献[11]提出了基于卷积神经网络的无监督学习深度功率控制方法,最大化无线网络多收发对的频谱效率或能量效率;文献[12]根据集成学习的思想利用多个独立训练的网络构建集成网络,并利用无监督学习对功率进行控制以最大化无线网络多收发对的速率和.文献[13-14]利用拉格朗日乘法对干扰约束进行处理,并基于无监督学习研究了认知无线网络和D2D(device-to-device)网络干扰受限的功率控制问题;文献[15-17]主张采用拉格朗日对偶法处理约束,在无监督学习的同时对拉格朗日乘子进行迭代训练,以避免对拉格朗日乘子进行大量的反复试验选择.然而,上述深度无监督学习方法只用于信道的功率控制,而未用于信道分配.

由于深度无监督学习既不需要满足马尔科夫决策过程,也不需要大量的标签数据,而且能够避免监督信号对网络优化水平的局限^[8,12],因而与深度强化学习和深度监督学习相比具有一定的优势.但通过对上述深度无监督学习在无线通信研究上的文献梳理发现,利用深度无监督学习进行多小区蜂窝网络全频复用无线资源分配的研究极少;另外,由于多小区蜂窝网络全频复用无线资源分配包括信道分配和功率控制两个方面,前期的深度无监督学习方法^[11-17]无法用于既需信道分配又需功率控制的全频复用无线资源分配.

本文提出一种全新的基于深度无监督学习的多小区蜂窝网全频复用资源分配方法,研究多小区蜂窝网络全频复用的信道分配和功率控制,并通过最大化能量效率的期望来满足网络对低能耗、低延时和高速率的要求.本文所提出的方法包括基于无监督学习的深度功率控制神经网络(unsupervised learning based deep neural network for power control, ULDNN-PC)和基于无监督学习的深度信道分配神经网络(unsupervised learning based deep neural network for channel allocation, ULDNN-CA).其中:ULDNN-PC可

根据信道增益和随机信道分配方案输出优化的信道功率,实现能量效率期望的最大化;ULDNN-CA可根据信道增益信息输出优化的信道分配方案,并针对获得的信道分配方案,借助前期训练好的ULDNN-PC拟合得到优化的信道功率,进一步最大化能量效率的期望.仿真结果表明,本文所提出的基于深度无监督学习的资源分配方法在保证低计算时延的同时可获得优于其他算法的能量效率和传输速率.

1 系统模型

采用集中式控制的某多小区蜂窝网络,其下行链路系统如图1所示.

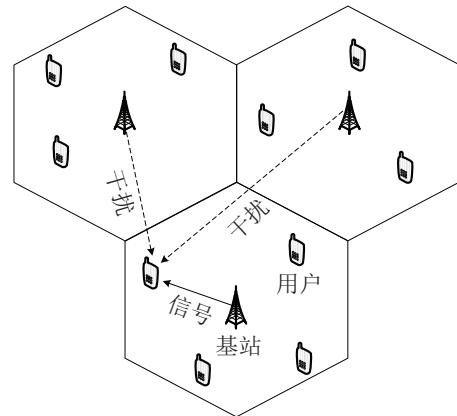


图1 集中式多小区全频复用下行链路系统模型

假定该蜂窝网络包含 M 个小区、 K 个用户并有 N 个正交的信道资源,每个小区中心有1个基站,用户随机分布在小区内,基站与用户均为单天线系统,信道资源在小区中进行全频复用,小区中的每个信道只分配给一个用户,且已知每个信道的带宽为 B 、噪声谱密度为 N_0 ;每个用户仅将位置信息通过导频信号传输给中心控制节点,中心控制节点可采用下式对信道增益进行建模和估计^[8]:

$$H_{m,k}^n = 10^{-(PL_{m,k} + X_\alpha)/10} |h_{m,k}^n|^2. \quad (1)$$

其中: $H_{m,k}^n$ 表示基站 $m(1 \leq m \leq M)$ 使用信道 $n(1 \leq n \leq N)$ 与用户 $k(1 \leq k \leq K)$ 通信时的信道增益; $PL_{m,k}$ 表示由基站 m 到用户 k 的路径损耗模型; X_α 表示阴影衰落,取均值为0、标准差为 α 的正态随机变量; $h_{m,k}^n$ 表示为基站 m 使用信道 n 与用户 k 通信的快衰落,取循环对称的独立同分布的均值为0、标准差为1的复高斯随机变量.

系统总的传输速率和能量效率^[4]可表示为

$$R = \sum_{m=1}^M \sum_{n=1}^N \sum_{k=1}^K \log_2 \left(1 + \frac{D_{m,k}^n P_{m,k}^n H_{m,k}^n}{(N_0 B + I_{m,k}^n) \Gamma} \right), \quad (2)$$

$$E = \sum_{m=1}^M \sum_{n=1}^N \sum_{k=1}^K \frac{B \log_2 \left(1 + \frac{D_{m,k}^n P_{m,k}^n H_{m,k}^n}{(N_0 B + I_{m,k}^n) \Gamma} \right)}{10^6 P_{m,k}^n}. \quad (3)$$

其中: R 为传输速率,单位为bit/(s·Hz); E 为系统的能量效率,单位为Mbit/(s·W); $D_{m,k}^n = 1$ 表示基站 m 将信道 n 分配给小区中的用户 k ,否则 $D_{m,k}^n = 0$; $p_{m,k}^n$ 表示基站 m 使用信道 n 与用户 k 通信时的功率; $\Gamma = -\ln(5 \cdot \text{BER})/1.6$, BER 为误比特率; $I_{m,k}^n$ 表示基站 m 使用信道 n 与用户 k 通信时的干扰,可表示为

$$I_{m,k}^n = \sum_{i=1, i \neq m}^M \sum_{j=1}^K D_{i,j}^n p_{i,j}^n H_{i,k}^n. \quad (4)$$

本文将优化目标表示为能量效率的期望以有效降低能耗. 与其他约束优化模型不同,为了保证全频复用的有效性,本文对信道传输所需的最低功率进行了约束,并通过考虑基站的总传输功率约束,将多小区蜂窝网全频复用的资源分配问题建模成如下的约束优化问题:

$$\begin{aligned} \text{(P1): } & \max_{\{D, p\}} E_{\mathbf{H}}[E]. \\ \text{s.t. } & C_1: \sum_{k=1}^K D_{m,k}^n = 1, D_{m,k}^n \in \{0, 1\}; \\ & C_2: p_{m,k}^n \geq 0; \\ & C_3: p_m^n \geq p_{\min}; \\ & C_4: N p_{\min} \leq p_{\text{tot},m} \leq p_{\max}. \end{aligned} \quad (5)$$

其中: $\mathbf{D} = \{D_{m,k}^n\}$, $\mathbf{p} = \{p_{m,k}^n\}$, $\mathbf{H} = \{H_{m,k}^n\}$; C_1 约束表示基站 m 只能将信道 n 分配给小区内的一个用户; C_2 约束表示基站 m 使用信道 n 与用户 k 通信的功率为非负值; C_3 约束表示信道传输所需的最低功率约束,其中 p_m^n 表示基站 m 在信道 n 上的传输功率, $p_m^n = \sum_{k=1}^K D_{m,k}^n p_{m,k}^n$; C_4 约束表示基站的总传输功率约束,其中 $p_{\text{tot},m}$ 表示基站 m 的总传输功率, $p_{\text{tot},m} = \sum_{n=1}^N p_m^n$.

2 基于深度无监督学习的资源分配方法

2.1 所提出的资源分配基本框架

一方面,约束优化问题P1是建立在信道增益期望上的,即只需要信道增益的统计信息,降低了对信道增益精确已知信息的依赖,这对优化算法的动态性能提出了挑战. 另一方面,若第 m 个小区有 k_m 个用户,则约束优化问题P1共有 $\prod_{m=1}^M k_m^N$ 种信道分配方案,即信道方案数量随着信道数量指数增长;多个小区间的同频信道干扰也使得功率的控制优化成为一个复杂的非凸优化问题. 基于上述原因,本文将约束优化问题P1的信道分配和功率控制进行分步求解,提出了基于深度无监督学习的资源分配基本框架,如图2所示.

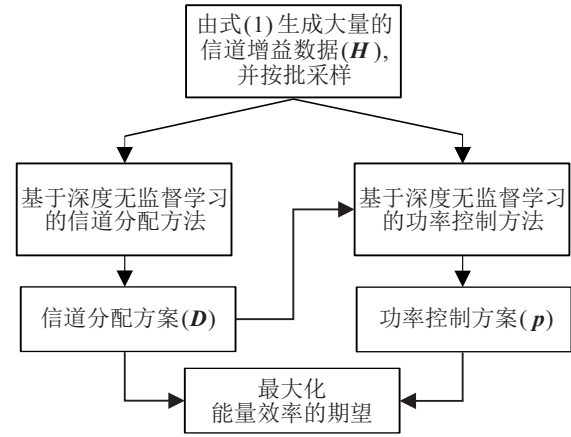


图2 基于无监督学习的资源分配基本框架

如图2所示,与传统方法^[2-3]先进行信道分配、后进行功率控制不同,本文方法先进行功率控制无监督学习、后进行信道分配无监督学习,分为基于深度无监督学习的功率控制方法和基于深度无监督学习的信道分配方法.

1) 基于深度无监督学习的功率控制方法根据采样得到的随机信道分配方案和信道增益信息,利用基于深度无监督学习的功率控制神经网络(ULDNN-PC)输出优化的信道功率,最大化该随机信道分配方案的能量效率期望.

2) 基于深度无监督学习的信道分配方法根据采样得到的信道增益信息,利用基于深度无监督学习的信道分配神经网络(ULDNN-CA)输出优化的信道分配方案,并针对获得的信道分配方案,调用前期训练好的ULDNN-PC拟合输出优化的信道功率,最终实现能量效率期望的最大化.

约束优化问题P1被ULDNN-PC和ULDNN-CA进行了分解. 其中,ULDNN-PC要解决的约束优化问题可描述为

$$\begin{aligned} \text{(P2): } & \max_{\mathbf{p}} E_{\{\mathbf{H}, \mathbf{D}\}}[E]. \\ \text{s.t. } & C_1: p_{m,k}^n \geq 0; \\ & C_2: p_m^n \geq p_{\min}; \\ & C_3: N \cdot p_{\min} \leq p_{\text{tot},m} \leq p_{\max}. \end{aligned} \quad (6)$$

而ULDNN-CA要解决的约束优化问题可描述为

$$\begin{aligned} \text{(P3): } & \max_{\mathbf{D}} E_{\mathbf{H}}[E]; \\ \text{s.t. } & C_1: \sum_{k=1}^K D_{m,k}^n = 1, D_{m,k}^n \in \{0, 1\}. \end{aligned} \quad (7)$$

2.2 基于无监督学习的深度功率控制神经网络(ULDNN-PC)

本文由约束优化问题P2来构建ULDNN-PC. 该网络为全连接的深度前向传播网络,其结构如图3所示.

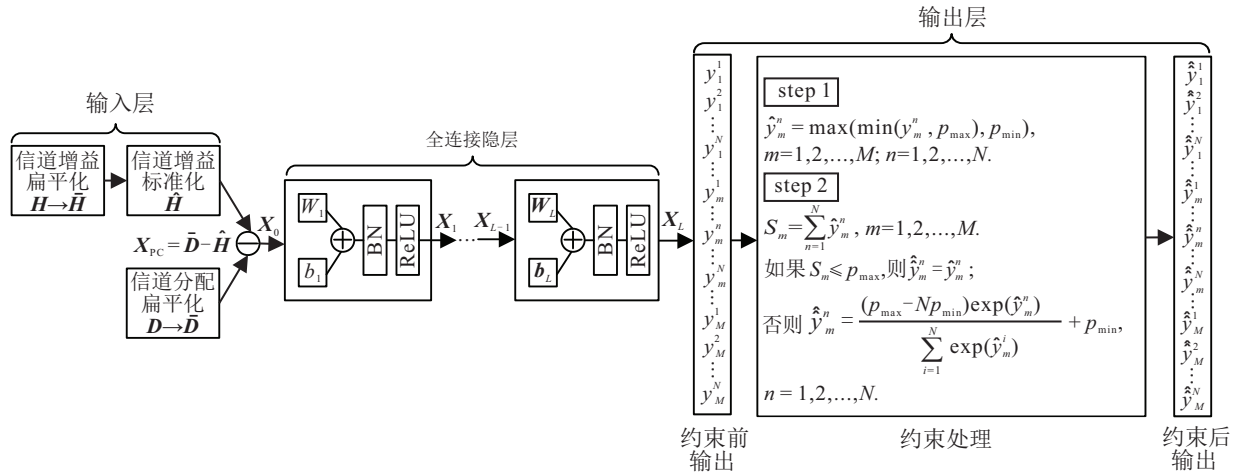


图3 ULDNN-PC深度网络结构

1) 输入层: 根据约束优化问题P2, ULDNN-PC的输入应由信道增益 H 和信道分配 D 组成. 为了降低输入层神经元的数量, 提高ULDNN-PC的优化性能, 本文对 H 和 D 进行了以下处理:

$$\hat{H} = \frac{\log_{10}(\bar{H}) - \min[\log_{10}(\bar{H})]}{\max[\log_{10}(\bar{H})] - \min[\log_{10}(\bar{H})]}, \quad (8)$$

$$X_{PC} = \bar{D} - \hat{H}. \quad (9)$$

其中: \bar{H} 、 \hat{H} 、 \bar{D} 、 X_{PC} 均是 $M \times N \times K$ 维的向量; \bar{H} 是信道增益 H 的扁平化向量; \hat{H} 是 \bar{H} 的标准化向量, 取值范围在 $[0, 1]$ 内; \bar{D} 是信道分配 D 的扁平化向量; X_{PC} 是ULDNN-PC的输入, 取值范围在 $[-1, 1]$ 内.

2) 全连接隐层: 每个全连接隐层均依照次序由1个全连接层、1个批归一化层以及1个ReLU激活层组成. 假设共有 L 个全连接隐层, 则第 l ($1 \leq l \leq L$) 个全连接隐层的输出可表示为

$$X_l = \text{ReLU}(\text{BN}(W_l X_{l-1} + b_l)). \quad (10)$$

其中: X_{l-1} 是第 $l-1$ 层的输出, $X_0 = X_{PC}$, W_l 、 b_l 为第 l 层的权值矩阵和偏置向量, $\text{ReLU}(\cdot)$ 为修正线性单元激活函数, $\text{BN}(\cdot)$ 为批标准化函数.

3) 约束前输出: 约束前输出 $Y = X_L$. 为了使输出层神经元与约束优化问题P2的约束变量 p_m^n 相对应, 本文将ULDNN-PC输出层神经元的个数定义为 $M \times N$, 即 Y 为 $M \times N$ 维向量. 为了方便, 记

$$Y = [y_1^1, \dots, y_1^N, \dots, y_m^1, \dots, y_m^N, \dots, y_M^1, \dots, y_M^N]^T. \quad (11)$$

其中 y_m^n 为约束前输出神经元的值, 与 p_m^n 相对应, 表示神经网络输出的基站 m 在信道 n 上的传输功率.

4) 约束处理与约束后输出: 为了使输出层神经元满足P2中 C_2 与 C_3 约束, 本文对约束前输出神经元进行了以下两步的约束处理:

step 1: 将约束前输出神经元的值 y_m^n 调整为 \hat{y}_m^n ,

即

$$\hat{y}_m^n = \max(\min(y_m^n, p_{\max}), p_{\min}). \quad (12)$$

step 2: 对于每一小区 m , 计算该小区基站的总功率 S_m , 有

$$S_m = \sum_{n=1}^N \hat{y}_m^n. \quad (13)$$

约束后输出神经元的值, 以使神经网络的输出满足P2中 C_2 与 C_3 约束, 即

$$\hat{y}_m^n = \begin{cases} \hat{y}_m^n, & S_m \leq p_{\max}; \\ \frac{\Delta p \cdot \exp(\hat{y}_m^n)}{\sum_{i=1}^N \exp(\hat{y}_m^i)} + p_{\min}, & \text{otherwise.} \end{cases} \quad (14)$$

其中 $\Delta p = p_{\max} - N \cdot p_{\min}$.

由于一个信道只能被基站分配给小区中的一个用户, 本文根据下式将约束后输出 \hat{y}_m^n 转化为 $p_{m,k}^n$, 以满足P2中 C_1 约束:

$$p_{m,k}^n = \begin{cases} \hat{y}_m^n, & D_{m,k}^n = 1; \\ 0, & D_{m,k}^n = 0. \end{cases} \quad (15)$$

5) 损失函数: 无监督学习训练ULDNN-PC时所采用的损失函数可定义为

$$L_{PC} = -E_{\{H,D\}}[E]. \quad (16)$$

在损失函数关于信道增益 H 和信道分配 D 期望

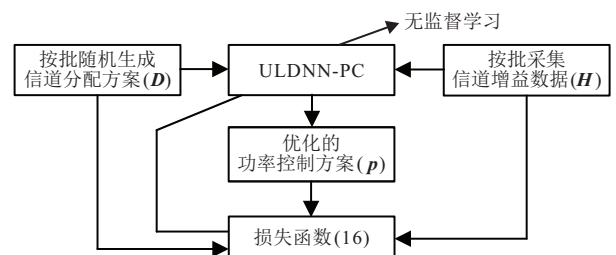


图4 ULDNN-PC训练示意图

的实现上, 本文通过在无监督学习中调用 mini-batch 梯度下降法来对 ULDDNN-PC 的权值进行训练, 如图 4 所示. 当损失函数值无法再下降时就可停止训练, 此时能量效率的负期望就达到了局部极小, 表明 ULDDNN-PC 已完成训练.

2.3 基于无监督学习的深度信道分配神经网络 (ULDDNN-CA)

本文由约束优化问题 P3 来构建 ULDDNN-CA. 该网络为全连接的深度前向传播网络, 其结构如图 5 所示.

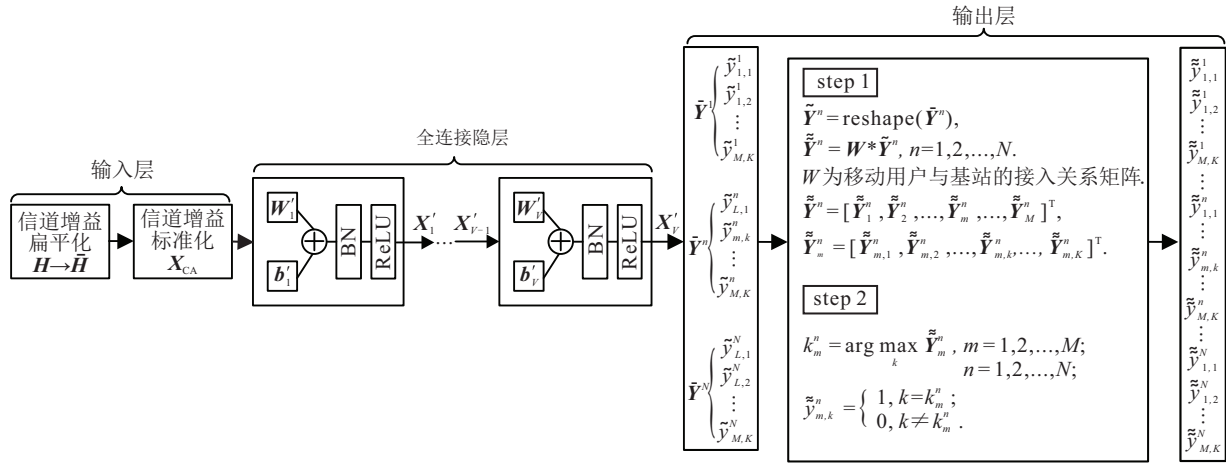


图 5 ULDDNN-CA 深度网络结构

1) 输入层: 根据约束优化问题 P3, ULDDNN-CA 的输入应为信道增益. 为了提高 ULDDNN-CA 的优化性能, 本文对信道增益进行了处理, 并将处理后的结果作为 ULDDNN-CA 的输入. 处理方法如下所示:

$$X_{CA} = \frac{\log_{10}(\bar{\mathbf{H}}) - E[\log_{10}(\bar{\mathbf{H}})]}{\sqrt{E\{(\log_{10}(\bar{\mathbf{H}}) - E[\log_{10}(\bar{\mathbf{H}})])^2\}}}, \quad (17)$$

其中 X_{CA} 是 $M \times N \times K$ 维的 ULDDNN-CA 输入向量, 取值范围在 $[-1, 1]$ 内.

2) 全连接隐层: 每个全连接隐层均依照次序由 1 个全连接层、1 个批归一化层和 1 个 ELU 激活层组成. 假设共有 V 个全连接隐层, 则第 v ($1 \leq v \leq V$) 个全连接隐层的输出 X'_v 可表示为

$$X'_v = \text{ELU}(\text{BN}(\mathbf{W}'_v X'_{v-1} + \mathbf{b}'_v)). \quad (18)$$

其中: X'_{v-1} 是第 $v-1$ 层的输出, $X'_0 = X_{PC}$, \mathbf{W}'_v 、 \mathbf{b}'_v 为第 v 层的权值矩阵和偏置向量, $\text{ELU}(\cdot)$ 为指数线性单元激活函数.

3) 约束前输出: 约束前输出 $\mathbf{Y}' = X'_v$. 为了使输出层神经元与 P3 的约束变量相对应, 本文将 ULDDNN-CA 输出层神经元的个数定义为 $M \times N \times K$, 由此 \mathbf{Y}' 的维数为 $M \times N \times K$. 为了方便, 将 \mathbf{Y}' 记为

$$\mathbf{Y}' = [\bar{\mathbf{Y}}^1, \bar{\mathbf{Y}}^2, \dots, \bar{\mathbf{Y}}^n, \dots, \bar{\mathbf{Y}}^N]^T. \quad (19)$$

其中 $\bar{\mathbf{Y}}^n$ 为 $M \times N$ 维的行向量, 可表示为

$$\bar{\mathbf{Y}}^n = [\tilde{y}_{1,1}^n, \dots, \tilde{y}_{1,K}^n, \tilde{y}_{2,1}^n, \dots, \tilde{y}_{2,K}^n, \dots, \tilde{y}_{M,1}^n, \dots, \tilde{y}_{M,k}^n, \dots, \tilde{y}_{M,K}^n]. \quad (20)$$

$\tilde{y}_{m,k}^n$ 为约束前输出神经元的值, 与 $D_{m,k}^n$ 相对应.

4) 约束处理与约束后输出: 为了使输出层神经元满足 P3 中 C_1 约束, 本文对约束前输出神经元进行了以下两步的约束处理.

step 1: 首先, 对 $\bar{\mathbf{Y}}^n$ 进行 reshape 后记为 $\tilde{\mathbf{Y}}^n$, 即 $\tilde{\mathbf{Y}}^n$ 为约束前输出神经元在第 n 个信道上所形成的神经元值矩阵, 可表示为

$$\tilde{\mathbf{Y}}^n = \text{reshape}(\bar{\mathbf{Y}}^n) = \begin{bmatrix} \tilde{y}_{1,1}^n & \dots & \tilde{y}_{1,k}^n & \dots & \tilde{y}_{1,K}^n \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ \tilde{y}_{m,1}^n & \dots & \tilde{y}_{m,k}^n & \dots & \tilde{y}_{m,K}^n \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ \tilde{y}_{M,1}^n & \dots & \tilde{y}_{M,k}^n & \dots & \tilde{y}_{M,K}^n \end{bmatrix}. \quad (21)$$

然后根据移动用户与基站的接入关系矩阵 \mathbf{W} , 对神经元值矩阵 $\tilde{\mathbf{Y}}^n$ 进行如下计算:

$$\tilde{\tilde{\mathbf{Y}}}^n = \mathbf{W} * \tilde{\mathbf{Y}}^n. \quad (22)$$

其中

$$\tilde{\tilde{\mathbf{Y}}}^n = \begin{bmatrix} \tilde{\tilde{y}}_{1,1}^n & \dots & \tilde{\tilde{y}}_{1,k}^n & \dots & \tilde{\tilde{y}}_{1,K}^n \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ \tilde{\tilde{y}}_{m,1}^n & \dots & \tilde{\tilde{y}}_{m,k}^n & \dots & \tilde{\tilde{y}}_{m,K}^n \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ \tilde{\tilde{y}}_{M,1}^n & \dots & \tilde{\tilde{y}}_{M,k}^n & \dots & \tilde{\tilde{y}}_{M,K}^n \end{bmatrix}; \quad (23)$$

$\mathbf{W} = \{w_{m,k}\}$ 表示移动用户与基站的接入关系矩阵; $w_{m,k} = 1$ 表示移动用户 k 接入到基站 m , 否则 $w_{m,k} = 0$; 符号 “*” 表示 Hadamard 积.

step 2: 将矩阵中 $\tilde{\mathbf{Y}}^n$ 第 m 行所示的向量记为列向量 $\tilde{\mathbf{y}}_m^n$,表示第 n 个信道在第 m 个小区的 K 维神经元值向量,有

$$\tilde{\mathbf{Y}}_m^n = [\tilde{y}_{m,1}^n, \tilde{y}_{m,2}^n, \dots, \tilde{y}_{m,k}^n, \dots, \tilde{y}_{m,K}^n]^T. \quad (24)$$

通过下式计算得到第 n 个信道在第 m 个小区的神经元值向量 $\tilde{\mathbf{Y}}_m^n$ 的最大元素所对应的位置:

$$k_m^n = \arg \max_k \tilde{y}_{m,k}^n. \quad (25)$$

然后,根据下式对 $\tilde{y}_{m,k}^n$ 进行修正,使得输出层神经元的值满足P3中 C_1 约束:

$$\tilde{y}_{m,k}^n = \begin{cases} 1, & k = k_m^n; \\ 0, & k \neq k_m^n. \end{cases} \quad (26)$$

其中 $\tilde{y}_{m,k}^n$ 为约束后输出神经元的值.

5) 损失函数: 无监督学习训练ULDNN-CA时所采用的损失函数可定义为

$$L_{CA} = -E_{\mathbf{H}}[E]. \quad (27)$$

为了完成ULDNN-CA的训练,需要调用前期已经训练好的ULDNN-PC,如图6所示,即将ULDNN-CA输出的信道分配方案以及反映信道环境的信道增益按批输入到前期训练好的ULDNN-PC中,此时主要通过ULDNN-PC的拟合得到优化的信道功率,最小化能量效率的负期望.在损失函数关于信道增益 \mathbf{H} 期望的实现上,本文同样通过无监督学习中调用mini-batch梯度下降法来对ULDNN-CA的权值进行训练.当损失函数值无法再下降时即可停止训练,此时能量效率的负期望就达到了局部极小,意味着ULDNN-CA已完成训练.

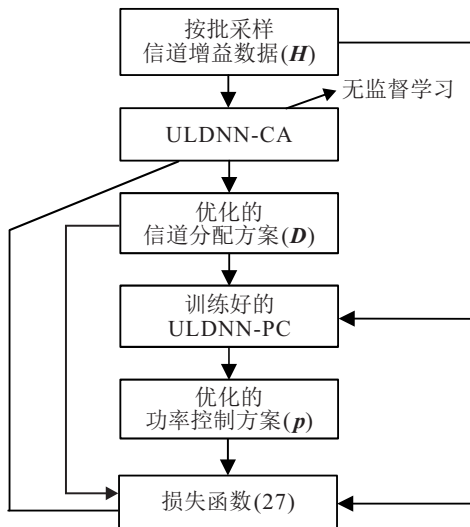


图6 ULDNN-CA训练示意图

2.4 基于深度无监督学习资源分配方法的在线推断

基于深度无监督学习资源分配方法的在线推断流程如下:

1) 将信道增益信息输入到训练好的ULDNN-CA;

2) 将ULDNN-CA输出的信道分配方案输入到训练好的ULDNN-PC,拟合输出优化的信道功率;

3) 根据信道增益信息、ULDNN-CA输出的信道分配方案、ULDNN-PC拟合输出的功率控制方案以及式(2)~(4)计算出传输速率和能量效率.

3 仿真与分析

本文仿真比较了不同算法在信号快衰落时变情况下的能量效率、传输速率和计算时延.为公平起见,所有算法均采用Pytorch 1.6.0进行仿真;所有深度学习算法的线下训练均采用NVIDIA Geforce RTX 2080 Ti 11G GPU、Intel(R) Core(TM) i9-9900K CPU 3.6 GHz和32G内存硬件环境;而所有深度学习算法的在线推断以及其他算法的性能比较均采用CPU进行.由于系统信道数量的增加会使优化难度指数增加,仿真中将小区和用户的数量固定不变,而使系统信道数量取4、6、8、10、12.具体的蜂窝网络仿真参数取值见表1.根据表1中的基站数(M)、移动用户数(K)和信道资源数(N),通过反复实验确定了如表2所示的ULDNN-PC和ULDNN-CA的神经网络结构.

表1 蜂窝网络仿真参数

参数	取值
基站/个	3
小区半径/m	200
基站的最大发射功率/dBm	38
信道的最低传输功率/dBm	26.99
载波频率/GHz	2.0
子信道带宽/kHz	180
噪声谱密度/(dBm/Hz)	-170
信道资源总数/个	4、6、8、10、12
移动用户数/个	10
路径损耗指数	3.2
参考距离/m	100
阴影衰落标准差 α	8
误比特率BER	10^{-3}

表2 ULDNN-PC和ULDNN-CA的网络结构

神经网络	网络层及其神经元数
ULDNN-PC	$M \times K \times N$ (input) \rightarrow 180(FC-BN-ReLU) \rightarrow 120(FC-BN-ReLU) \rightarrow 80(FC-BN-ReLU) $\rightarrow M \times N$ (output)
ULDNN-CA	$M \times K \times N$ (input) \rightarrow 300(FC-BN-ELU) $\rightarrow M \times K \times N$ (output)

注: input是输入层,FC是全连接层,BN是批归一化层,ReLU和ELU是激活层,output是输出层; $M=3, K=10, N=4, 6, 8, 10, 12$.

首先,分析ULDNN-PC和ULDNN-CA线下训练的优化特性.

在训练ULDNN-PC和ULDNN-CA时,本文总共

生成了150 000个信道增益数据以及相同数量的随机信道分配数据,其中135 000个数据作为训练集、15 000个数据作为验证集. 设定 ULDNN-PC 的学习速率为0.001、ULDNN-CA 的学习速率为0.0003. 本文采用早停法,当最优的验证损失在50个连续的训练周期中没有更新时停止训练过程. 基于提出的深度无监督学习方法,首先训练 ULDNN-PC,然后借助训练好的 ULDNN-PC 来训练 ULDNN-CA. 图7是当 $M = 3$ 、 $K = 10$ 、 $N = 4$ 时 ULDNN-PC 的训练损失和验证损失,图8是当 $M = 3$ 、 $K = 10$ 、 $N = 4$ 时 ULDNN-CA 联合 ULDNN-PC 的训练损失和验证损失.

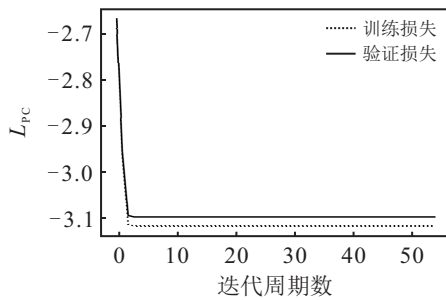


图7 ULDNN-PC 的训练损失和验证损失

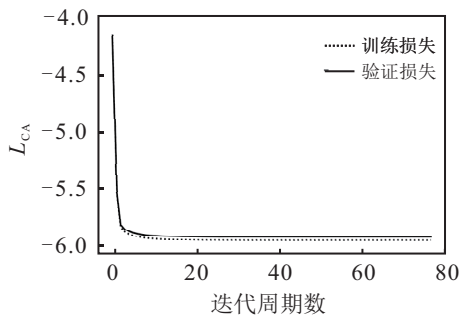


图8 ULDNN-CA 的训练损失和验证损失

比较图7和图8可以发现,ULDNN-CA 收敛时的训练损失和验证损失均小于 ULDNN-PC. 原因是,ULDNN-PC 在训练中采用的是随机生成的信道分配信息,而 ULDNN-CA 则是在训练中通过前期训练好的 ULDNN-PC 拟合输出优化的信道功率,进而通过 mini-batch 梯度下降得到优化的信道分配信息,因此能够更有效地优化能量效率的期望.

其次,为了检验 ULDNN-CA 联合 ULDNN-PC 的优化性能,将其与传统的集中式人工蜂群算法^[2,18]、集中式贪婪/随机信道分配及 WMMSE 信道功率控制算法^[19] 进行比较.

1) 在精确信道增益信息已知的情况下对各算法进行了比较,以检验本文方法的优化性能水平. 图9~图11 是各算法50次蒙特卡洛重复仿真的结果.

图9~图11 表明,各算法的能量效率、传输速率

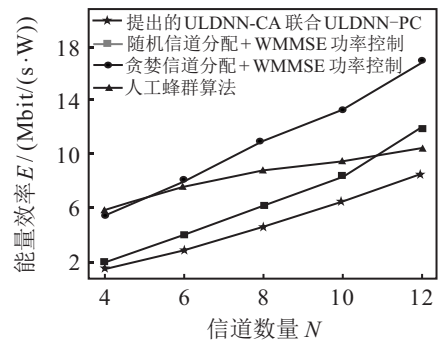


图9 精确信道增益信息已知时各算法的平均能量效率

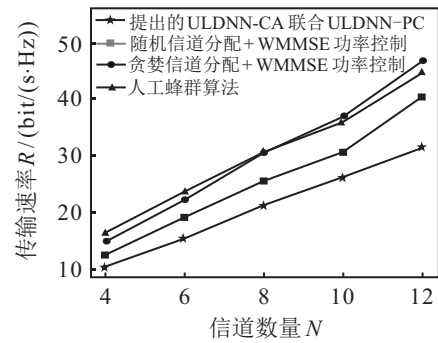


图10 精确信道增益信息已知时各算法的平均传输速率

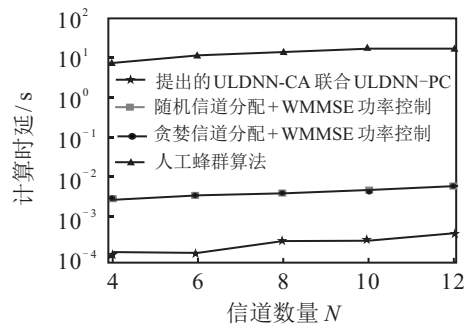


图11 精确信道增益信息已知时各算法的平均计算时延

以及计算时延均随着系统可用信道数量的增加而增加,并且所提出的 ULDNN-CA 联合 ULDNN-PC 能够在最大化能量效率的同时获得接近于或高于其他算法的传输速率. 尽管资源分配优化的难度随着系统可用信道数量的增加而增大(信道分配可选方案随信道数量的增加呈指数增加),但所提出的 ULDNN-CA 联合 ULDNN-PC 所耗费的计算时延仅在 10^{-4} s 数量级左右,远远小于贪婪/随机信道分配与 WMMSE 功率控制相结合的 10^{-3} s 数量级的计算时延、人工蜂群算法的 10^1 s 数量级的计算时延. 这主要与各算法的理论计算复杂度以及优化计算特点相关. 各算法的理论计算复杂度见表3.

由表3可知,迭代次数是影响人工蜂群算法、贪婪/随机信道分配和 WMMSE 功率控制算法计算复杂度的一个重要因素,因而也是增加算法计算时延的主要原因. 与人工蜂群算法、贪婪/随机信道分配和

表3 各算法的理论计算复杂度

算法	理论计算复杂度
本文方法	$O(K \times M \times N \times N_{CA-hidden1})$
人工蜂群算法	$O(Cycles \times M \times N \times N_{foods})^{[18]}$
贪婪信道分配+	信道分配 $O(K \times N)$
WMMSE 功率控制	功率控制 $O(Cycles \times N \times K^2)^{[19]}$
随机信道分配+	信道分配 $O(M \times N)$
WMMSE 功率控制	功率控制 $O(Cycles \times N \times K^2)^{[19]}$

注: $N_{CA-hidden1}$ 是 ULDNN-CA 第一个全连接隐层的神经元数量, N_{foods} 是人工蜂群的蜜源总数, Cycles 是人工蜂群或 WMMSE 的迭代次数.

WMMSE 功率控制算法不同, 所提出的 ULDNN-CA 联合 ULDNN-PC 能够按批进行在线推断, 其计算时延主要由矩阵相乘产生. 与人工蜂群算法、贪婪/随机信道分配与 WMMSE 功率控制算法的复杂迭代计算相比, 矩阵相乘调用了高度优化的 BLAS (basic linear algebra subroutine), 使得矩阵相乘具有远高于迭代运算的计算效率, 因而所提出的 ULDNN-CA 联合 ULDNN-PC 具有较快的计算速度.

如表 3 所示, 贪婪/随机信道分配与 WMMSE 功率控制算法的计算复杂度均由 WMMSE 功率控制的计算复杂度决定, 因而在图 11 中的计算时延表现几乎相同. 而人工蜂群算法需要优化的蜜源的维度随着信道数量的增加而增加, 加大了优化的难度^[20], 从而降低了优化效率. 因此, 如图 9 和图 11 所示, 随着系统信道数量的增加, 人工蜂群算法在能量效率上增加得越来越缓慢, 其计算时延也由 10^0 s 数量级增大到了 10^1 s 数量级.

2) 在精确信道增益信息未知的情况下对各算法进行了比较, 以检验提出的 ULDNN-CA 联合 ULDNN-PC 的动态优化性能水平. 由于用户位置已知, 人工蜂群算法、贪婪/随机信道分配与 WMMSE 功率控制算法可通过信道增益的估计模型(1)获得先验的信道分配和功率控制方案, 并通过先验的信道分配和功率控制方案在精确信道增益信息未知的情况下对能量效率进行优化. 图 12 和图 13 是在 50 次蒙特卡洛重复仿真中各算法在不同信道数量时的能量效率

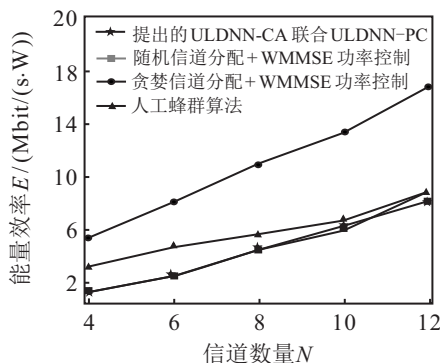


图 12 精确信道增益信息未知时各算法的平均能量效率

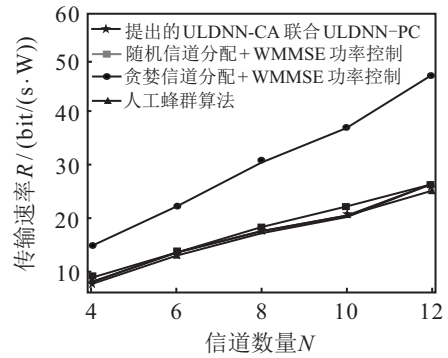


图 13 精确信道增益信息未知时各算法的平均传输速率和传输速率的平均值.

由图 12 和图 13 可知, 当精确信道增益信息未知时, 所提出的 ULDNN-CA 联合 ULDNN-PC 在能量效率和传输速率上远好于其他传统算法, 说明所提出的 ULDNN-CA 联合 ULDNN-PC 具有较好的动态优化性能.

最后, 将本文方法与卷积深度神经网络^[11]、深度强化学习算法^[4,21]等其他深度学习算法进行了比较. 由于卷积神经网络^[11]主要用于 D2D 功率控制, 本文将贪婪信道分配与其进行了结合. 仿真中, 卷积深度神经网络的结构为 5 个卷积层与 1 个全连接层, 并均含有批标准化 BN 层和 ReLU 激活层, 5 个卷积层的卷积核大小分别为 $(N, 16, 3, 3)$ 、 $(16, 8, 3, 3)$ 、 $(8, 8, 3, 3)$ 、 $(8, 8, 3, 3)$ 、 $(8, 8, 3, 3)$, 全连接层的大小为 $(240, M \times N)$. 深度强化学习^[21]的状态为功率, 动作为信道分配、信道功率以及功率修正, 将深度强化学习^[21]的网络结构设为 4 个全连接层, 均含有批标准化 BN 层和 ReLU 激活层, 全连接层的维度分别为 $M \times N$ 、180、120、80、 $N \times K + M \times N + d \times M$, 其中 $N \times K$ 、 $M \times N$ 、 $d \times M$ 分别为信道分配、信道功率和功率修正的动作数量, d 为修正向量的维度, 将修正向量设为 $[0, -0.001, 0.001, -0.003, 0.003]$, 即 $d = 5$. 深度强化学习^[4]的迭代结构设为 10, 状态为系统干扰, 动作为信道分配. 50 次蒙特卡洛重复仿真的结果如图 14 和图 15 所示.

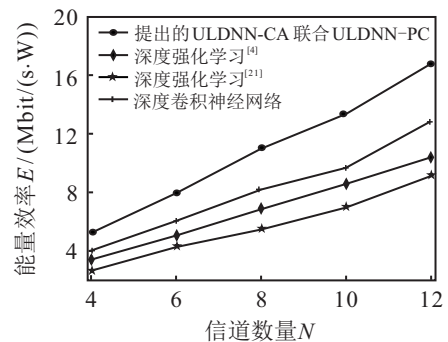


图 14 精确信道增益信息未知时各算法的平均能量效率

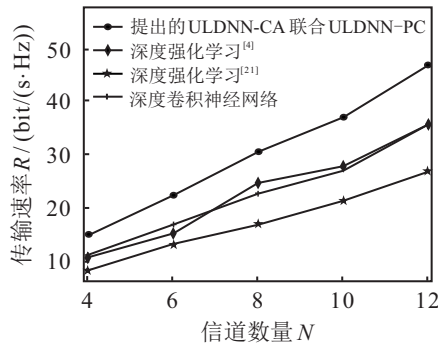


图 15 精确信道增益信息未知时各算法的平均传输速率

图 14 和图 15 表明, 本文所提出的方法在能量效率和传输速率上均好于卷积深度神经网络^[11]、深度强化学习算法^[4,21]。与卷积深度神经网络^[11] 的比较说明, 本文利用深度神经网络对信道分配与信道功率进行同时优化的方法是有效的。深度强化学习算

法^[21] 由于采用了离散化的功率动作, 限制了深度 Q 网络优化性能的发挥。深度强化学习算法^[4] 利用增广拉格朗日乘子约束优化进行信道分配和功率控制, 但不恰当的惩罚因子极易降低约束优化的性能^[15-17], 且在增广拉格朗日乘子约束优化的迭代中信道分配为连续的实数值而非离散的 0、1 值, 以上均导致所产生的功率控制方案削弱了深度 Q 网络的优化性能。

将本文所提出的方法与深度卷积神经网络^[11]、深度强化学习^[4,21]、人工蜂群算法^[18]、WMMSE 方法^[19] 的比较结果汇总于表 4。由表 4 可看出, 所提出的方法在信道数为 4 的能量效率上仅次于人工蜂群算法, 在信道数为 6、8、10 和 12 的能量效率上均优于其他算法; 在信道数为 4、6、8 的传输速率上仅次于人工蜂群算法, 在 10 和 12 的传输速率上均优于其他算法; 计算时延仅次于文献[21]的深度强化学习算法。

表 4 各算法的能量效率、传输速率和计算延时

算法	能量效率/(Mbit/(s·W))					传输速率/(bit/(s·Hz))					计算延时/s				
	4	6	8	10	12	4	6	8	10	12	4	6	8	10	12
本文方法	5.42	8.02	10.99	13.30	16.75	15.05	22.28	30.53	36.94	46.51	1e-4	1e-4	3e-4	3e-4	4e-4
深度卷积网络	4.07	6.13	8.17	9.74	12.78	11.30	17.02	22.70	27.05	35.51	1.3e-4	1.3e-4	2.5e-4	2.5e-4	2.5e-4
深度强化学习 ^[21]	2.75	4.38	5.55	7.00	9.16	8.46	13.35	16.93	21.36	26.73	2.5e-5	5e-5	7.5e-5	1e-4	1e-4
深度强化学习 ^[4]	3.41	5.11	6.89	8.62	10.43	10.84	15.32	24.70	27.65	35.49	9.3e-4	9.3e-4	9.3e-4	4.3e-3	5.3e-3
人工蜂群算法	5.79	7.62	8.80	9.40	10.44	16.44	23.59	30.77	35.91	44.37	8.20	11.05	14.06	16.53	18.91
贪婪信道分配+WMMSE 功率控制	1.99	3.89	6.22	8.38	11.98	12.77	19.21	25.67	30.55	40.10	2.9e-3	3.6e-3	4.2e-3	4.9e-3	5.78e-3
随机信道分配+WMMSE 功率控制	1.55	2.79	4.50	6.49	8.70	10.28	15.26	20.67	26.74	32.24	3e-3	3.7e-3	4.3e-3	5.2e-3	5.92e-3

以上的仿真和分析说明, 本文所提出的方法具有较低的计算时延, 且在最大化能量效率的同时能够获得接近于或优于其他算法的传输速率, 表明了本文方法的有效性。

4 结论

为了满足新一代蜂窝网络对低能耗、高速率和低延时的要求, 本文提出了一种基于深度无监督学习的多小区蜂窝网络全频复用资源分配方法, 通过约束处理方法构建了满足约束优化的 ULDDN-PC 和 ULDDN-CA 深度全连接神经网络, 并对信道分配和功率控制进行了分步求解, 即首先通过 ULDDN-CA 输出优化的信道分配方案, 然后借助前期训练好的 ULDDN-PC 拟合输出优化的信道功率, 最小化了能量效率的负期望。仿真结果表明, 所提出的 ULDDN-CA 联合 ULDDN-PC 的深度无监督学习资源分配方法具有较低的计算时延, 且能获得优于其他算法的能量效率和传输速率。

参考文献(References)

- [1] He X M, Wang K, Huang H W, et al. Green resource allocation based on deep reinforcement learning in content-centric IoT[J]. IEEE Transactions on Emerging Topics in Computing, 2020, 8(3): 781-796.
- [2] 孙明, 曹伟, 李大辉, 等. 保证公平的最大化 OFDMA 系统容量策略[J]. 控制与决策, 2020, 35(5): 1175-1182. (Sun M, Cao W, Li D H, et al. Strategy of maximizing capacity of OFDMA system for ensuring fairness[J]. Control and Decision, 2020, 35(5): 1175-1182.)
- [3] Sun M, Lee K Y, Xu Y Q, et al. Hysteretic noisy chaotic neural networks for resource allocation in OFDMA system[J]. IEEE Transactions on Neural Networks and Learning Systems, 2018, 29(2): 273-285.
- [4] 廖晓闽, 严少虎, 石嘉, 等. 基于深度强化学习的蜂窝网资源分配算法[J]. 通信学报, 2019, 40(2): 11-18. (Liao X M, Yan S H, Shi J, et al. Deep reinforcement learning based resource allocation algorithm in cellular networks[J]. Journal on Communications, 2019, 40(2):

- 11-18.)
- [5] Challita U, Dong L, Saad W. Proactive resource management for LTE in unlicensed spectrum: A deep learning perspective[J]. *IEEE Transactions on Wireless Communications*, 2018, 17(7): 4674-4689.
- [6] Li X J, Fang J, Cheng W, et al. Intelligent power control for spectrum sharing in cognitive radios: A deep reinforcement learning approach[J]. *IEEE Access*, 2018, 6: 25463-25473.
- [7] Ye H, Li G Y, Juang B H F. Deep reinforcement learning based resource allocation for V2V communications[J]. *IEEE Transactions on Vehicular Technology*, 2019, 68(4): 3163-3173.
- [8] Ahmed K I, Tabassum H, Hossain E. Deep learning for radio resource allocation in multi-cell networks[J]. *IEEE Network*, 2019, 33(6): 188-195.
- [9] Zappone A, Debbah M, Altman Z. Online energy-efficient power control in wireless networks by deep neural networks[C]. 2018 IEEE International Workshop on Signal Processing Advances in Wireless Communications (SPAWC). Kalamata, 2018: 1-5.
- [10] Sun H R, Chen X Y, Shi Q J, et al. Learning to optimize: Training deep neural networks for wireless resource management[C]. 2017 IEEE International Workshop on Signal Processing Advances in Wireless Communications (SPAWC). Sapporo, 2017: 1-6.
- [11] Lee W, Kim M, Cho D H. Deep power control: Transmit power control scheme based on convolutional neural network[J]. *IEEE Communications Letters*, 2018, 22(6): 1276-1279.
- [12] Liang F, Shen C, Yu W, et al. Power control for interference management via ensembling deep neural networks[C]. 2019 IEEE/CIC International Conference on Communications in China (ICCC). Changchun, 2019: 237-242.
- [13] Lee W. Resource allocation for multi-channel underlay cognitive radio network based on deep neural network[J]. *IEEE Communications Letters*, 2018, 22(9): 1942-1945.
- [14] Lee W, Kim M, Cho D H. Transmit power control using deep neural network for underlay device-to-device communication[J]. *IEEE Wireless Communications Letters*, 2019, 8(1): 141-144.
- [15] Lee H, Lee S H, Quek T Q S. Constrained deep learning for wireless resource management[C]. 2019 IEEE International Conference on Communications (ICC). Shanghai, 2019: 1-6.
- [16] Sun C J, Yang C Y. Unsupervised deep learning for ultra-reliable and low-latency communications[C]. 2019 IEEE Global Communications Conference (GLOBECOM). Waikoloa, 2019: 1-6.
- [17] Liu D, Sun C J, Yang C Y, et al. Optimizing wireless systems using unsupervised and reinforced-unsupervised deep learning[J]. *IEEE Network*, 2020, 34(4): 270-277.
- [18] Sharma N, Anpalagan A. Bee colony optimization aided adaptive resource allocation in OFDMA systems with proportional rate constraints[J]. *Wireless Networks*, 2014, 20(7): 1699-1713.
- [19] Shi Q J, Razaviyayn M, Luo Z Q, et al. An iteratively weighted MMSE approach to distributed sum-utility maximization for a MIMO interfering broadcast channel[J]. *IEEE Transactions on Signal Processing*, 2011, 59(9): 4331-4340.
- [20] 汪照, 李有明, 陈斌, 等. 基于鱼群算法的OFDMA自适应资源分配[J]. *物理学报*, 2013, 62(12): 509-515. (Wang Z, Li Y M, Chen B, et al. OFDMA adaptive resource allocation based on fish swarm algorithm[J]. *Acta Physica Sinica*, 2013, 62(12): 509-515.)
- [21] 周凡, 王鸿, 宋荣方. 密集异构蜂窝网络中基于深度强化学习的下行链路功率分配算法[J]. *南京邮电大学学报: 自然科学版*, 2021, 41(2): 17-24. (Zhou F, Wang H, Song R F. Deep reinforcement learning based downlink power allocation algorithm in dense heterogeneous cellular networks[J]. *Journal of Nanjing University of Posts and Telecommunications: Natural Science Edition*, 2021, 41(2): 17-24.)

作者简介

孙明(1979—), 男, 教授, 博士, 从事深度学习、智能优化算法等研究, E-mail: snogisunming@163.com;

王淑梅(1980—), 女, 讲师, 博士生, 从事深度学习、演化博弈的研究, E-mail: wangshumei_qiqihar@163.com;

郭媛(1974—), 女, 教授, 博士, 从事光学图像加密、深度学习等研究, E-mail: guoyuan171@126.com;

曹伟(1977—), 男, 教授, 博士, 从事智能控制、迭代学习控制等研究, E-mail: gjg1998@124.com;

徐耀群(1972—), 男, 教授, 博士生导师, 从事智能优化算法、深度学习等研究, E-mail: xuyq@hrbcu.edu.cn.

(责任编辑: 孙艺红)