

基于Q-Learning算法的毫微微小区功率控制算法

李云 唐英* 刘涵霄

(重庆邮电大学移动通信技术重庆市重点实验室 重庆 400065)

摘要: 该文研究macro-femto异构蜂窝网络中移动用户的功率控制问题, 首先建立了以最小接收信号信干噪比为约束条件, 最大化毫微微小区的总能效为目标的优化模型; 然后提出了基于Q-Learning算法的毫微微小区集中式功率控制(PCQL)算法, 该算法基于强化学习, 能在没有准确信道状态信息的情况下, 实现对小区内所有用户终端的发射功率统一调整。仿真结果表明该算法能实现对用户终端的功率有效控制, 提升系统能效。

关键词: 集中式功率控制; Q-Learning算法; 能效优化

中图分类号: TN92

文献标识码: A

文章编号: 1009-5896(2019)11-2557-08

DOI: 10.11999/JEIT181191

Power Control Algorithm Based on Q-Learning in Femtocell

LI Yun TANG Ying LIU Hanxiao

(Chongqing Key Laboratory of Mobile Communication Technology, The Chongqing University of Posts and Telecommunications, Chongqing 400065, China)

Abstract: The power control problem of mobile users in macro-femto heterogeneous cellular networks is studied. Firstly, an optimization model that maximizes the total energy efficiency of femtocells with the minimum received signal-to-noise ratio as the constraint is established. Then, a femtocell centralized Power Control algorithm based on Q-Learning (PCQL) is proposed. Based on reinforcement learning, the algorithm can adjust the transmit power of the user terminal without accurate channel state information simultaneously. The simulation results show that the algorithm can effectively control the power of the user terminal and improve system energy efficient.

Key words: Centralized power control; Q-Learning algorithm; Energy-efficient optimization

1 引言

现有的移动业务主要来自室内移动用户, 而室外的蜂窝网络需要面对严重的干扰和穿透损耗等, 难以满足大容量的室内覆盖, 且能量消耗大。毫微微基站(Femtocell Base Station, FBS)是一种部署在室内, 传输距离很短(通常小于100 m)的小型基

站, 其部署的主要目的是增加室内的覆盖率, 为用户提供高速率的移动服务, 从而提升室内移动用户的服务质量。基于此, 室外部署宏基站与室内部署毫微微基站相结合的异构蜂窝网络结构是支持移动通信业务快速增长的有效方式^[1,2]。

在异构蜂窝网络中, 由于远近效应、用户和基站间的小区内干扰和小区间干扰的存在, 要求基站和用户终端选择合适的发射功率。功率控制对于提升异构蜂窝网络的性能至关重要, 合理地为用户分配发射功率, 能够降低能耗, 增加系统容量和能效。与传统蜂窝网络相比, 异构蜂窝网络中存在大量毫微微基站; 另外, 在网络结构、管理方式等方面异构蜂窝网络存在明显差异; 因此, 不能直接采用传统蜂窝网络的功率控制技术, 需要研究适合异构蜂窝网络特点的功率控制理论和方法。

已有的异构蜂窝网络功率控制技术主要通过集中式和分布式两种方式来实现。文献[3]提出了一种联合空本子帧和功率控制的集中式算法, FBS根据宏小区用户遭受的干扰水平, 调节其发射功率, 以

收稿日期: 2018-12-28; 改回日期: 2019-04-10; 网络出版: 2019-05-21

*通信作者: 唐英 17749963914@163.com

基金项目: 国家自然科学基金(61671096), 重庆市研究生科研创新项目(CYS17220), 重庆市“科技创新领军人才支持计划”(CST-CCXLJRC201710), 重庆市基础科学与前沿技术研究项目(cstc2017jcyjBX0005), 重庆市留学人员创业创新支持计划

Foundation Items: The National Natural Science Foundation of China (61671096), The Chongqing Research and Innovation Program of Graduated Students (CYS17220), The Chongqing Science and Technology Innovation Leadership Talent Support Program (CSTCCXLJRC201710), The Chongqing Research Program of Basic Science and Frontier Technology (cstc2017jcyj-BX0005), The Chongqing Overseas Students Entrepreneurship and Innovation Support Plan

提升宏小区用户的性能。基于Voronoi图,文献[4]提出根据估计的业务量来控制FBS的发送功率,以改善室内覆盖并进行负载均衡。文献[5]通过引入中断概率,建立数学模型分析能效优化问题,提出了一种不依赖多余信令交换的功率控制策略。文献[6]通过调整FBS的功率,动态改变FBS的覆盖范围,以使宏基站(Macrocell Base Station, MBS)空闲休眠来减小能耗。为了减轻网络干扰并降低能耗,文献[7]联合用户关联和功率控制技术,提出一种双层迭代的功率控制算法,在已知FBS发射功率的情况下更新用户关联索引,在固定用户关联索引的情况下利用拉格朗日算法调节FBS的发射功率。针对异构蜂窝网络上行链路的干扰抑制、能效优化等问题。文献[8]针对macro-femto网络的干扰问题,提出利用随机几何学分析上行链路的功率控制机制,以抑制小区间的干扰。文献[9]基于随机几何学联合考虑上行链路传输速率、用户中断概率和小区间干扰,提出一种调整功率控制系数的上行链路功率控制方案,以最大化系统能效。文献[10]综合考虑联合基站选择、信道分配和模式选择,利用凸优化知识,提出一种基于参数Dinkelbach算法的迭代功率控制算法,以优化上行链路的能效,提升系统性能。文献[11]引入不同路径损耗模型以及不同的接入区域,采用分数式功率控制方法控制用户终端的发送功率,旨在提升上行链路能效。

出于简化功率控制算法的复杂度,文献[12-20]提出基于分布式的功率控制算法。针对下行链路,文献[12]借助博弈论,以最大化异构蜂窝网络容量为目标,讨论FBS如何根据来自MBS反馈的干扰信息,通过自适应调整自身发送功率,降低小区间干扰。文献[13]提出一种联合动态信道和自功率控制方案,以抑制用户自身受到的干扰,从而满足移动业务的服务质量要求。文献[14]提出FBS根据启发式算法获得的最优功率传播因子,为自身选择合适的发射功率,以优化整个网络的能效。文献[15]以最大化MBS和FBS的容量为目标,将MBS和FBS的功率控制问题抽象为斯坦克尔伯格博弈模型,其中MBS为领导者,FBS为跟随者,并进一步讨论了增加MBS和FBS的容量的博弈策略。基于非合作式博弈模型,文献[16]将MBS和FBS视作博弈的参与者,二者利用迭代式注水算法和随机统计近似算法调节各自的发射功率,以优化系统的性能。文献[17,18]根据凸优化和博弈理论,分别提出了梯度分布式功率控制方法以及基于博弈的能效控制方法,以提升系统的能效。文献[19,20]提出分布式的上行链路功率控制算法。文献[19]联合考虑功率控

制和接入控制,将异构网络频谱效率和能效优化问题转化为NP难问题,以提升上行链路系统性能。基于Dinkelbach算法,文献[20]针对上行链路的能效优化问题,提出一种联合考虑中继选择、子载波选择的功率控制算法。

已有功率控制技术的相关研究主要针对异构蜂窝网络的干扰抑制、能效优化的功率控制问题。且在现有的能效优化功率控制问题中,主要研究点在于下行链路的功率控制,而关注上行链路的功率控制的研究相对较少。在已有的功率控制算法中,分布式的功率控制算法无需全局信息,复杂度较低,但是分布式的功率控制方式获得优化结果是次优的。基于集中式控制理念的功率控制算法,通常需要获得一个全局的归一化信道增益矩阵来得到最优解,其信令开销大,运算复杂度较高,获取实时的信道增益困难。

针对现有分布式和集中式功率控制算法的上述问题,本文利用强化学习具有感知外部环境,能够在不确定的环境下进行智能决策的特性^[21],提出了基于Q-Learning算法的移动用户集中式功率控制(PCQL)算法,该算法能够在不确定信道状态信息的情形下,通过智能学习,对毫微微小区内所有用户发送功率进行集中式地自适应调整,以实现系统能效的优化。

2 系统模型

2.1 网络模型

不失一般性,假设异构蜂窝网络由1个宏小区和多个毫微微小区组成,在这里用 $B = \{B_0, B_1, \dots, B_M\}$ 表示基站集合,其中, B_0 表示宏基站, B_1, B_2, \dots, B_M 表示 M 个FBS, B_b 表示第 b 个基站;用 $U = \{U_0, U_1, \dots, U_M\}$ 表示各个基站用户集合构成的集合,其中 U_b 表示第 b 个基站的用户集合,具体地, $U_b = \{u_{b,1}, u_{b,2}, \dots, u_{b,N_b}\}$,其中 $u_{b,\mu}$ 表示第 b 个小区中第 μ 个用户终端, N_b 表示第 b 个小区的用户个数;用户间共享相同的频谱带宽 W ,当 $u_{b,\mu}$ 与 B_b 进行通信时,既会遭受到同小区内用户的干扰,又会遭受到其它小区内用户的干扰,具体的网络模型如图1所示。

用 $p_{b,\mu}$ 表示 $u_{b,\mu}$ 的发射功率;用 $P_{b,\mu}^c$ 表示 $u_{b,\mu}$ 的固有功率消耗;用 $\text{SINR}_{b,\mu}$ 表示 $u_{b,\mu}$ 与 B_b 通信时, B_b 接收到的信干噪比,如式(1)所示;其中 $I_{b,\mu}$ 表示 $u_{b,\mu}$ 与 B_b 通信时, B_b 接收到的干扰信号,如式(2)所示。用 $R_{b,\mu}$ 表示 $u_{b,\mu}$ 的吞吐量,如式(3)所示

$$\text{SINR}_{b,\mu} = \frac{p_{b,\mu} g_{b,\mu}^b}{I_{b,\mu} + Wn_0} \quad (1)$$

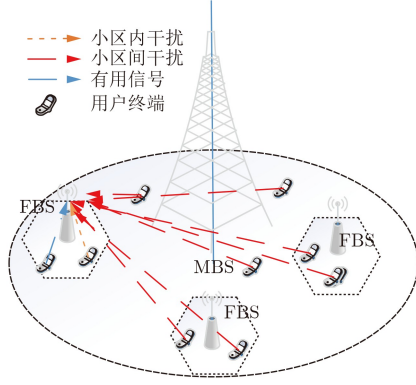


图1 异构蜂窝网络模型

$$I_{b,\mu} = \sum_{\substack{i \in U_b \\ i \neq \mu}} p_{b,i} g_{b,i}^b + \sum_{\substack{m=0 \\ m \neq b}}^M \sum_{n \in U_m} p_{m,n} g_{m,n}^b \quad (2)$$

$$R_{b,\mu} = W \log_2 \left(1 + \frac{p_{b,\mu} g_{b,\mu}^b}{I_{b,\mu} + W n_0} \right) \quad (3)$$

其中, $g_{b,\mu}^b$ 表示 $u_{b,\mu}$ 与 B_b 之间的信道增益; 信道是功率谱密度为 n_0 的高斯白噪声信道。

2.2 优化模型

本文以最大化上行能效为优化目标。上行链路用户能效定义为单位功率用户的吞吐量, 用 $\eta_{b,\mu}$ 表示 $u_{b,\mu}$ 的能效, 如式(4)所示

$$\eta_{b,\mu} = \frac{R_{b,\mu}}{P_{b,\mu}} = \frac{W \log_2 \left(1 + \frac{p_{b,\mu} g_{b,\mu}^b}{I_{b,\mu} + W n_0} \right)}{(p_{b,\mu} + P_{b,\mu}^c)} \quad (4)$$

其中, $P_{b,\mu}$ 表示 $u_{b,\mu}$ 的总能耗。上行链路毫微微小区 b 的能效定义为小区 b 内用户能效之和, 表示为

$$\eta_b = \sum_{\mu \in U_b} \eta_{b,\mu} = \sum_{\mu \in U_b} \frac{W \log_2 \left(1 + \frac{p_{b,\mu} g_{b,\mu}^b}{I_{b,\mu} + W n_0} \right)}{(p_{b,\mu} + P_{b,\mu}^c)} \quad (5)$$

则毫微微小区上行能效优化的数学模型为

$$\left. \begin{aligned} \max_{p_{b,\mu}} \quad & \sum_{\mu \in U_b} \frac{W \log_2 \left(1 + \frac{p_{b,\mu} g_{b,\mu}^b}{I_{b,\mu} + W n_0} \right)}{(p_{b,\mu} + P_{b,\mu}^c)} \\ \text{s.t.} \quad & \text{SINR}_{b,\mu} \geq \text{SINR}_{b,\mu}^{\min}, \quad \forall b \in B, \forall \mu \in U_b \\ & p_{b,\mu} \leq p_{b,\mu}^{\max}, \quad \forall b \in B, \forall \mu \in U_b \end{aligned} \right\} \quad (6)$$

其中, $\text{SINR}_{b,\mu}^{\min}$ 为 B_b 解码来自 $u_{b,\mu}$ 的接收信号所需要的最小信干噪比, $p_{b,\mu}^{\max}$ 为 $u_{b,\mu}$ 最大的发射功率, 二者构成系统的约束条件。

3 基于Q-Learning算法的毫微微小区功率控制算法

在利用Q-Learning算法进行功率控制时, 需要

定义Q-Learning算法的3大要素, 即动作空间(A)、奖赏函数(\mathfrak{R})、状态空间(S)。其中, 动作空间被定义为控制实体(为实现在FBS的一个功率控制实体, 以下简称代理)在某种状态 s 下可以执行的动作 a 的所有可能, 奖赏函数被定义为代理在状态 s 下执行动作 a 所带来的回报值, 状态空间被定义为代理可能处于的状态集合。本文将“状态-动作值函数”表示为 $Q(s, a)$ ^[22]。结合功率控制问题, 定义功率控制动作空间、状态空间、能效奖赏函数如下。

3.1 功率控制动作空间

定义发射功率向量 $\mathbf{P}_b = [p_{b,1} p_{b,2} \cdots p_{b,N_b}]$ 为毫微微小区 b 内代理为各用户终端分配的发射功率, 也即代理采用的动作 $\mathbf{a}_b = \mathbf{P}_b$ 。利用式(1)和 $\text{SINR}_{b,\mu}^{\min}$, 得到用户 $u_{b,\mu}$ 的发射功率的取值范围为 $p_{b,\mu}^{\min} \leq p_{b,\mu} \leq p_{b,\mu}^{\max}$, 在该范围包括 $d_{b,\mu}$ 个发送功率等级, $u_{b,\mu}$ 的发射功率等级集合表示为 $\rho_{b,\mu} = \{p_{b,\mu}^1, p_{b,\mu}^2, \cdots, p_{b,\mu}^{d_{b,\mu}}\}$ 。因此, 代理的功率控制动作空间可以表示为 $A_b (A_b = (\rho_{b,1} \times \rho_{b,2} \times \cdots \times \rho_{b,N_b}))$ 。

3.2 状态空间

本文定义代理的状态为 $\mathbf{s}_b = [I_b \mathbf{l}_b]$, \mathbf{s}_b 的所有可能构成了代理的状态空间 S , 即 $\forall \mathbf{s}_b \in S$ 。其中 $I_b = [I_{b,1} I_{b,2} \cdots I_{b,N_b}]$ 表示干扰向量; $\mathbf{l}_b = [\lambda_{b,1} \lambda_{b,2} \cdots \lambda_{b,N_b}]$ 是一个系数向量, $\lambda_{b,\mu}$ 用于判断实际的 $\text{SINR}_{b,\mu}$ 与 $\text{SINR}_{b,\mu}^{\min}$ 之间的大小关系, 其中 $\lambda_{b,\mu} = \begin{cases} 1, & \text{SINR}_{b,\mu} \geq \text{SINR}_{b,\mu}^{\min} \\ 0, & \text{其它} \end{cases}$ 。此处的上行干扰和信干噪比可以通过基站端测量得到。

3.3 能效奖赏函数

本文的优化目标是最大化毫微微小区的能效, 因而它定义代理的奖赏函数为毫微微小区 b 的能效, 则代理在状态 \mathbf{s}_b 下采取动作 \mathbf{a}_b 的奖赏函数可以表示为

$$\mathfrak{R}_b(\mathbf{s}_b, \mathbf{a}_b) = \sum_{\mu \in U_b} \frac{\lambda_{b,\mu} W \log_2(1 + \text{SINR}_{b,\mu})}{(p_{b,\mu} + P_{b,\mu}^c)} \quad (7)$$

3.4 算法描述

在Q-Learning算法中, 代理根据各用户的基站信号受到的上行干扰和信干噪比, 动态调整用户终端的发送功率, 以实现上行能效(奖赏函数)的优化。代理与小区环境的互动过程如图2所示。代理与环境的互动过程可以建模为4元组的形式 $(s_b^t, \mathbf{a}_b^t, \mathfrak{R}_b^{t+1}, s_b^{t+1})$, 其中, t 是一个离散的时间计数器。该4元组表示代理在状态 s_b^t 时, 采取动作 \mathbf{a}_b^t 后, 获得值为 \mathfrak{R}_b^{t+1} 的奖赏, 然后状态转变 s_b^{t+1} 。即当用户的基站信号受到的上行干扰和信干噪比系数向量分别为 I_b^t, \mathbf{l}_b^t 时, 代理控制各用户终端的发

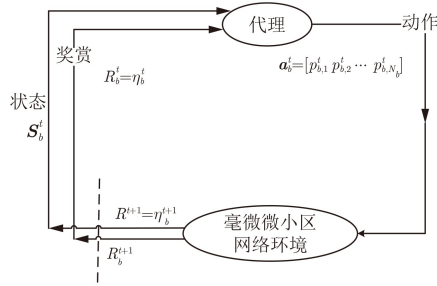


图2 代理自主学习过程

送功率为 $\mathbf{a}_b^t = [p_{b,1}^t, p_{b,2}^t, \dots, p_{b,N_b}^t]$ 后, 使得小区 b 的总能效为 $\eta_b = \mathfrak{R}_b^{t+1}(\mathbf{s}_b^t, \mathbf{a}_b^t)$ 。因为新的发射功率向量会导致干扰和信干噪比改变, 因此系统位于新的状态 $\mathbf{s}_b^{t+1} = [\mathbf{I}_b^{t+1}, \mathbf{l}_b^{t+1}]$ 。

代理的整个学习过程等价于学习状态-动作函数(也称作Q-函数), 其用作计算给定状态 \mathbf{s}_b 下采取动作 \mathbf{a}_b 的预期回报(即, 为各用户选择合适的发射功率后, 得到的能效值)。具体而言, 与状态-动作对和功率策略 π 相关的Q-函数, 可以表示为

$$Q_{\gamma}^{\pi}(\mathbf{s}_b, \mathbf{a}_b) = \mathbb{E}_{\pi_p} \left[\left(\sum_{t=0}^{+\infty} \gamma^t \mathfrak{R}_b^{t+1} \mid \mathbf{s}_b^0 = \mathbf{s}_b, \mathbf{a}_b^0 = \mathbf{a}_b \right) \right] \quad (8)$$

其中, γ 表示折扣奖赏因子。Q-Learning算法旨在获得最佳策略, 也即找到能够最大化Q-函数的策略。本文通过最大化Q-函数获得最佳的发射功率控制策略 π^* 。

考虑利用增量求和的方法可获得Q-函数的另一个表达式, 如式(9)所示

$$Q_{t+1}^{\pi}(\mathbf{s}_b, \mathbf{a}_b) = Q_t^{\pi}(\mathbf{s}_b, \mathbf{a}_b) + \alpha \left(\mathfrak{R}_b(\mathbf{s}_b, \mathbf{a}_b) + \gamma \max_{\mathbf{a}_b'} (Q_t^{\pi}(\mathbf{s}_b', \mathbf{a}_b')) - Q_t^{\pi}(\mathbf{s}_b, \mathbf{a}_b) \right) \quad (9)$$

其中, α 表示学习效率。结合式(7)和式(9), 更新Q-函数, 从而获得策略最佳的功率控制策略。详细的过程如下描述:

步骤1 设定信道模型(带宽, 增益等), 给定初始信息, 设置总的执行步数为 T , 确定代理的功率控制动作空间 A_b 以及状态空间 S ;

步骤2 给定代理的初始状态 $\mathbf{s}_b = [\mathbf{I}_b, \mathbf{l}_b]$;

步骤3 利用 ε -greedy 算法为小区内用户终端选择发射功率向量; 以概率 ε 对 A_b 中的动作进行探索, 即从整个动作空间中随机选择一个动作; 以概率 $1-\varepsilon$ 对 A_b 中的动作进行利用, 即选取使得当前 $Q^{\pi}(\mathbf{s}_b, \mathbf{a}_b)$ 最大的动作; 并且记录采取动作 \mathbf{a}_b 后, 代理处于的新状态 $\mathbf{s}_b^* = [\mathbf{I}_b^*, \mathbf{l}_b^*]$;

步骤4 计算采取动作 \mathbf{a}_b 后, 获得的奖赏值

$\mathfrak{R}_b(\mathbf{s}_b, \mathbf{a}_b)$ (也即, 采取发射功率向量 \mathbf{a}_b 后, 得到的系统能效值); 然后选择状态 \mathbf{s}_b^* 下最大的Q值, 结合式(9)更新 $Q^{\pi}(\mathbf{s}_b, \mathbf{a}_b)$;

步骤5 根据步骤4中得到的最新 $Q^{\pi}(\mathbf{s}_b, \mathbf{a}_b)$, 找出状态 \mathbf{s}_b 下使得Q值最大的动作 \mathbf{a}_b^{\otimes} , 然后更新功率控制策略 $\pi(\mathbf{s}_b) = \mathbf{a}_b^{\otimes}$ (即, 系统处于状态 $\mathbf{s}_b = [\mathbf{I}_b, \mathbf{l}_b]$ 时, 最佳的发射功率向量为 \mathbf{a}_b^{\otimes});

步骤6 更新状态 $\mathbf{s}_b \leftarrow \mathbf{s}_b^*$, 再次执行步骤3—步骤5, 直到执行步数等于 T 。

本文所提算法的具体过程如表1所示。

表1 基于Q-Learning算法的毫微微小区功率控制算法(PCQL)

输入: $W, n_0, P_{b,\mu}^c, \text{SINR}_{b,\mu}^{\min}, p_{b,\mu}^{\max}, \gamma, \alpha, T, \varepsilon$, 动作空间 A_b ;

输出: $\pi^*, p_{b,\mu}^*(\mu \in U_b)$;

定义: \mathbf{k} 表示代理选取的动作; $\text{SINR}_{b,\mu}^{\text{real}}$ 表示 $u_{b,\mu}$ 与基站 b 通信时的实际信干噪比;

$$Q(\mathbf{s}_b, \mathbf{a}_b) = 0, \pi(\mathbf{s}_b, \mathbf{a}_b) = \frac{1}{|A_b(\mathbf{s}_b)|}, \mathbf{s}_b^t = \mathbf{s}_b^0;$$

for $t = 0, 1, \dots, T$ do

若 $\text{rand}() < \varepsilon$, 从 A_b 中随机选动作 \mathbf{k} ; 否则 $\mathbf{k} = \arg \max_{\mathbf{a}_b} Q(\mathbf{s}_b^t, \mathbf{a}_b^t)$;

根据式(1)确定 $\text{SINR}_{b,\mu}^{\text{real}}$;

for $\mu = 1, 2, \dots, N_b$ do

若 $\text{SINR}_{b,\mu}^{\text{real}} \geq \text{SINR}_{b,\mu}^{\min}$, 那么 $\lambda_{b,\mu} = 1$; 否则 $\lambda_{b,\mu} = 0$;

end for;

根据式(7)计算采取动作 $\mathbf{a}_b^t = \mathbf{k}$ 所带来的奖赏值 $\mathfrak{R}_b(\mathbf{s}_b^t, \mathbf{a}_b^t)$;

$$\mathbf{a}_b^{t+1} = \pi(\mathbf{s}_b^{t+1});$$

$$Q(\mathbf{s}_b^t, \mathbf{a}_b^t) \leftarrow Q(\mathbf{s}_b^t, \mathbf{a}_b^t) + \alpha \left(\mathfrak{R}_b(\mathbf{s}_b^t, \mathbf{a}_b^t) + \gamma \max_{\mathbf{a}_b'} (Q(\mathbf{s}_b^{t+1}, \mathbf{a}_b^{t+1})) - Q(\mathbf{s}_b^t, \mathbf{a}_b^t) \right);$$

$$\mathbf{s}_b^t \leftarrow \mathbf{s}_b^{t+1};$$

end for;

$$\pi^*(\mathbf{s}_b) = \arg \max_{\mathbf{a}_b} Q(\mathbf{s}_b, \mathbf{a}_b), \forall \mathbf{s}_b \in S.$$

4 仿真分析

4.1 仿真参数

本文使用Python仿真分析本文提出的PCQL算法的性能, 并同文献[10]进行比较。文献[10]是近两年提出的异构蜂窝网络中一种典型的能效优化功率控制算法(PCPD算法)。PCPD结合参数Dinkelbach算法和拉格朗日理论实现集中式的功率控制。PCPD在更新用户发射功率的过程中, 需要准确的信道状态信息。由于本文也是针对异构蜂窝网络的功率控制问题, 因此采用PCPD作为对比算法。将PCPD算法与本文所提的PCQL算法进行对比, 从

而验证PCQL算法的性能。本文的仿真参数设置如表2所示。

表 2 主要的仿真参数

参数名称	参数值
MBS/FBS	1个/4个
MUE/FUE最大的发射功率	37 dBm/30 dBm
MBS/FBS覆盖范围半径	250 m/50 m
$SINR_{b,\mu}^{\min}$	-9 dB
固定的电路功耗	100 mW
信道带宽	10 MHz
高斯白噪声的功率谱密度	10^{-11} W/Hz

设定信道增益为 $g_{b,u}^b = (r_{b,u}^b)^{-a_{b,u}^b} 10^{\varepsilon_{b,u}^b/10}$ ，其中， $(r_{b,u}^b)^{-a_{b,u}^b}$ 表示 $u_{b,\mu}$ 与基站 b 通信时的路径损耗， $a_{b,u}^b$ 为路径损耗因子； $10^{\varepsilon_{b,u}^b/10}$ 表示 $u_{b,\mu}$ 与基站 b 通信时的阴影衰落， $\varepsilon_{b,u}^b$ 表示均值为0、方差为 $(\sigma_{b,u}^b)^2$ 的正态随机变量。本文设定 $(\sigma_{b,u}^b)^2 = 10$ ， $a_{b,u}^b = 4$ ，设定PCQL算法的学习效率 $\alpha = 0.1$ ，折扣奖赏因子 $\gamma = 0.9$ 。

4.2 仿真结果

分别描述了毫微微小区内用户数为4时，PCQL算法和PCPD算法对用户终端进行发射功率控制得到的系统能效和系统吞吐量对比结果。从图3和图4可以知道，利用PCQL算法优化系统能效和吞吐量

时，其收敛速度略慢于PCPD算法，但其得到的优化结果优于PCPD算法的优化结果。

图5和图6分别描述了在准确估计信道状态信息时，系统能效和系统吞吐量与毫微微小区用户数的关系。从图5中可以知道系统的能效随着用户数的增加而增加，另外采用PCQL算法控制用户终端的发射功率得到的能效结果始终优于采用PCPD算法得到的优化结果。从图6中可以知道系统的吞吐量先增加，且增长的速度越来越慢，而后趋于一个稳定的值。这是因为用户数的增加，使得毫微微小区内吞吐量求和个数增加，所以吞吐量增加；另外用户数的增加，也导致干扰增加，所以吞吐量的增长速度越来越慢；当用户数增加到一定数量时，小区内干扰达到门限，使得系统吞吐量的不再随着用户数的增加而增加。从图6中还可以得到，采用PCQL算法得到的吞吐量结果始终优于采用PCPD算法得到的结果。

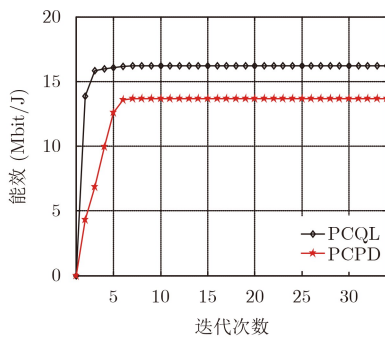


图 3 小区用户数为4时，系统能效对比

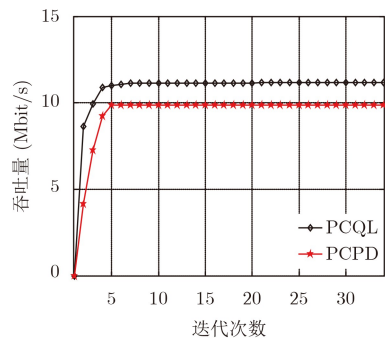


图 4 小区用户数为4时，系统吞吐量对比

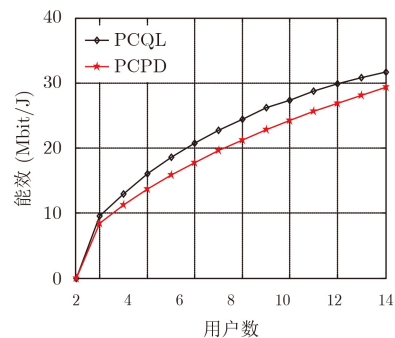


图 5 系统能效与用户数的关系

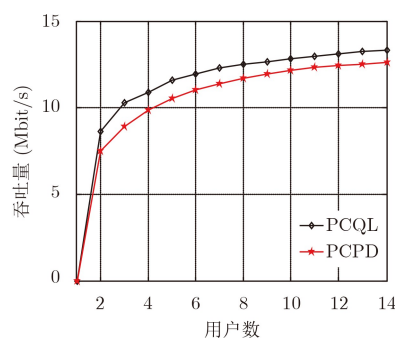


图 6 系统吞吐量与用户数的关系

图7和图8分别描述了信道状态信息估计误差为5%、10%、15%时，系统能效和系统吞吐量与毫微微小区内用户数的关系。对比图7和图5，可以知道，在两种场景下，采用PCQL算法获得的能效结果几乎相当，而采用PCPD算法得到的能效结果明显下降，当信道状态信息估计误差为15%时，能效结果下降了约10.7%。

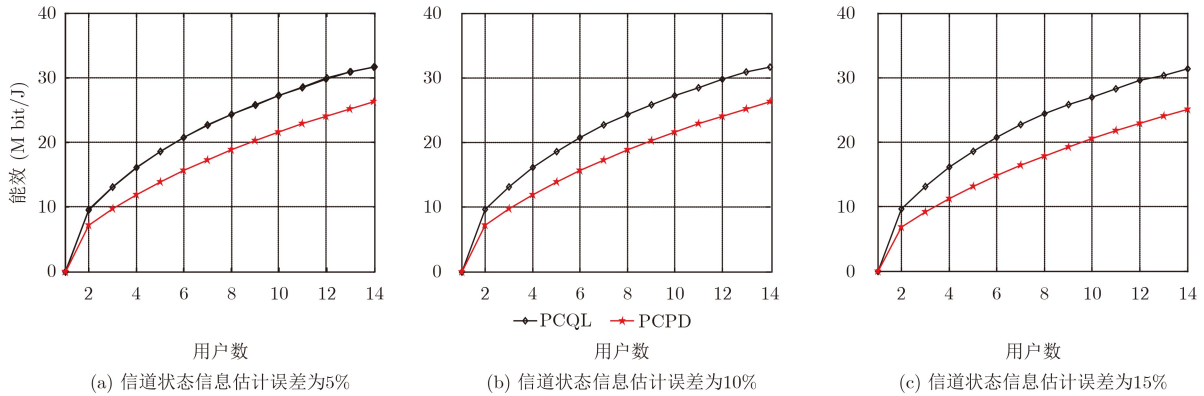


图7 信道状态信息存在估计误差时，系统能效与用户数的关系

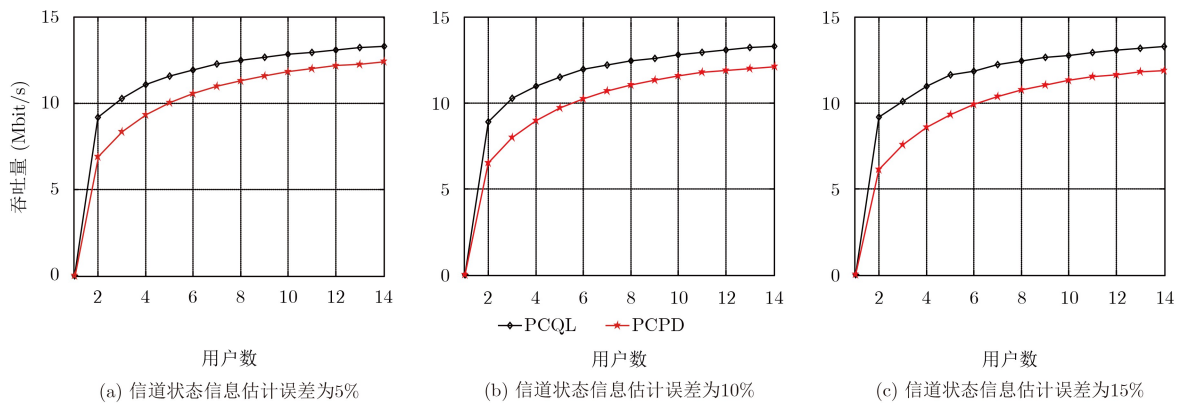


图8 信道状态信息存在估计误差时，系统吞吐量与用户数的关系

对比图8和图6，可以知道，采用PCQL算法在两种场景下获得的吞吐量优化结果几乎相当；而采用PCPD算法，当信道状态信息估计误差为15%时，得到的吞吐量结果下降了约4.76%。显然，在信道状态信息存在估计误差的情况下，PCQL算法能够实现对用户终端的功率有效控制。这是因为PCQL算法通过Q-Learning算法，根据基站端反馈的干扰和信干噪比，获得各状态下毫微微小区内用户的最佳发射功率，从而学习到最优的用户发射功率控制策略；当用户和基站进行通信时，给定用户任意的初始发射功率，代理会根据学习得到的最优发射功率控制策略，为用户终端调控最佳的发射功率，其无需测量信道状态信息，因而不受信道状态信息估计误差的影响。

图9和图10给出了两种算法在不同用户数下分别优化能效和吞吐量时的运行时间。由图9和图10可以知道，相较于PCPD算法，本文所提PCQL算法在用户数增加时其运行时间增加明显。这是因为随着用户的增加，功率控制动作空间的维度不断增加，因而代理的学习更长。因此，从运算复杂度看，PCQL不如PCPD。然而，由于PCQL不需要测量反馈信道状态信息，因此其信令开销会更小。

同时，当信道状态的估计不准确时，PCQL仍有比较好的性能。

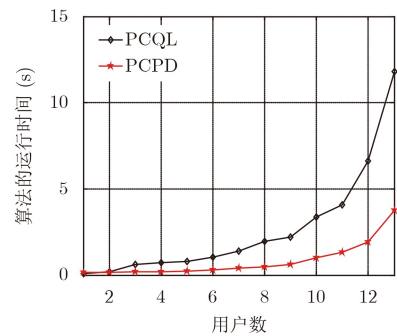


图9 能效优化的算法运行时间对比

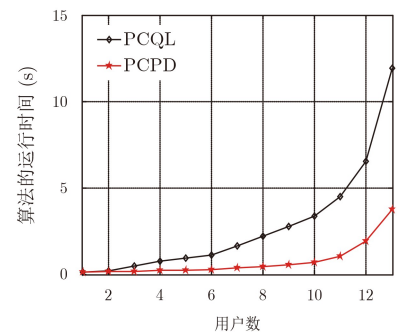


图10 吞吐量优化的算法运行时间对比

通过仿真数据并结合上述分析可以得到, 利用PCQL算法优化毫微微小区的能效, 能够有效地节约系统的能耗, 另外PCQL算法能够有效地应对信道状态信息估计误差的影响, 提升系统性能。

5 结论

本文研究毫微微小区移动用户的功率控制问题, 将单位功率的用户吞吐量定义为用户能效, 将毫微微小区用户能效之和定义为小区总能效; 并建立以用户最大发射功率和最小信干噪比为约束条件, 以最大化小区总能效为目标的最优化模型; 提出基于Q-Learning算法的毫微微小区功率控制算法。本文算法基于强化学习, 将基站端观测获得的上行干扰和信干噪对比系数定义为代理的状态; 将小区总能效映射为代理动作奖赏值; 将毫微微小区内各用户的发射功率构成的功率向量视作代理的动作, 在某个状态下为小区内用户终端选择合适的发射功率。仿真结果验证了本文算法设计的正确性和有效性。

参 考 文 献

- [1] LÓPEZ-PÉREZ D, DING M, CLAUSSEN H, *et al.* Towards 1 Gbps/UE in cellular systems: understanding ultra-dense small cell deployments[J]. *IEEE Communications Surveys & Tutorials*, 2015, 17(4): 2078–2101. doi: [10.1109/COMST.2015.2439636](https://doi.org/10.1109/COMST.2015.2439636).
- [2] YUNAS S F, VALKAMA M, and NIEMELÄ J. Spectral and energy efficiency of ultra-dense networks under different deployment strategies[J]. *IEEE Communications Magazine*, 2015, 53(1): 90–100. doi: [10.1109/MCOM.2015.7010521](https://doi.org/10.1109/MCOM.2015.7010521).
- [3] MARTOLIA D, SATHYA V, RANGISETTI A K, *et al.* Enhancing performance of victim macro users via joint ABSF and dynamic power control in LTE HetNets[C]. The Twenty-third National Conference on Communications, Chennai, India, 2017: 1–6.
- [4] SHIN D and CHOI S. Dynamic power control for balanced data traffic with coverage in femtocell networks[C]. The 8th International Wireless Communications and Mobile Computing Conference, Limassol, Cyprus, 2012: 648–653.
- [5] ZHANG Jinzhu, HONG Peilin, XUE Kaiping, *et al.* A novel power control scheme for femtocell in heterogeneous networks[C]. 2012 IEEE Consumer Communications and Networking Conference, Las Vegas, USA, 2012: 802–806.
- [6] PAN Zhenni, MEGUMI, SAITOU, *et al.* Neuron control-based power adjustment scheme for sleep two-tier cellular networks[C]. 2014 IEEE Wireless Communications and Networking Conference, Istanbul, Turkey, 2014: 3201–3206.
- [7] ZHOU Tianqing, LIU Zunxiong, ZHAO Junhui, *et al.* Joint user association and power control for load balancing in downlink heterogeneous cellular networks[J]. *IEEE Transactions on Vehicular Technology*, 2018, 67(3): 2582–2593. doi: [10.1109/TVT.2017.2768574](https://doi.org/10.1109/TVT.2017.2768574).
- [8] MARTIN-VEGA F J, GOMEZ G, AGUAYO-TORRES M C, *et al.* Analytical modeling of interference aware power control for the uplink of heterogeneous cellular networks[J]. *IEEE Transactions on Wireless Communications*, 2016, 15(10): 6742–6757. doi: [10.1109/TWC.2016.2588469](https://doi.org/10.1109/TWC.2016.2588469).
- [9] ZHANG Jing, LIAO Yan, and XIN Yili. Uplink power control for heterogeneous small cell networks[C]. 2016 IEEE 83rd Vehicular Technology Conference, Nanjing, China, 2016: 1–5.
- [10] WANG Min, GAO Hui, and LV Tiejun. Energy-efficient user association and power control in the heterogeneous network[J]. *IEEE Access*, 2017, 5: 5059–5068. doi: [10.1109/ACCESS.2017.2690305](https://doi.org/10.1109/ACCESS.2017.2690305).
- [11] ZHANG Jing, XIANG Lin, NG D W K, *et al.* Energy efficiency evaluation of multi-tier cellular uplink transmission under maximum power constraint[J]. *IEEE Transactions on Wireless Communications*, 2017, 16(11): 7092–7107. doi: [10.1109/TWC.2017.2739142](https://doi.org/10.1109/TWC.2017.2739142).
- [12] PAN Zhenni and SHIMAMOTO S. Cell sizing based energy optimization in joint macro-femto deployments via sleep activation[C]. 2013 IEEE Wireless Communications and Networking Conference, Shanghai, China, 2013: 4765–4770.
- [13] SHIFAT A S M Z, CHOWDHURY M Z, and JANG Y M. Game-based approach for QoS provisioning and interference management in heterogeneous networks[J]. *IEEE Access*, 2018, 6: 10208–10220. doi: [10.1109/ACCESS.2017.2704094](https://doi.org/10.1109/ACCESS.2017.2704094).
- [14] MISHRA S and MURTHY C S R. Increasing energy efficiency via transmit power spreading in dense femto cell networks[J]. *IEEE Systems Journal*, 2018, 12(1): 971–980. doi: [10.1109/JSYST.2016.2573845](https://doi.org/10.1109/JSYST.2016.2573845).
- [15] GURUACHARYA S, NIYATO D, KIM D I, *et al.* Hierarchical competition for downlink power allocation in OFDMA femtocell networks[J]. *IEEE Transactions on Wireless Communications*, 2013, 12(4): 1543–1553. doi: [10.1109/TWC.2013.022213.120016](https://doi.org/10.1109/TWC.2013.022213.120016).
- [16] WANG Haining, WANG Jiaheng, and DING Zhi. Distributed power control in a two-tier heterogeneous network[J]. *IEEE Transactions on Wireless Communications*, 2015, 14(12): 6509–6523. doi: [10.1109/TWC.2015.2456055](https://doi.org/10.1109/TWC.2015.2456055).
- [17] MAO Tingli, FENG Gang, LIANG Liang, *et al.* Energy-efficient power control for macro-femto networks[C]. The 22nd Wireless and Optical Communication Conference, Chongqing, China, 2013: 122–125.
- [18] MAO Tingli, FENG Gang, LIANG Liang, *et al.* Distributed energy-efficient power control for macro-femto networks[J].

- IEEE Transactions on Vehicular Technology*, 2016, 65(2): 718–731. doi: [10.1109/TVT.2015.2402618](https://doi.org/10.1109/TVT.2015.2402618).
- [19] LAI Weisheng, CHANG T H, and LEE T S. Joint power and admission control for spectral and energy efficiency maximization in heterogeneous OFDMA networks[J]. *IEEE Transactions on Wireless Communications*, 2016, 15(5): 3531–3547. doi: [10.1109/TWC.2016.2522958](https://doi.org/10.1109/TWC.2016.2522958).
- [20] LOODARICHEH R A, MALLICK S, BHARGAVA V K. Energy-efficient resource allocation for OFDMA cellular networks with user cooperation and QoS provisioning[J]. *IEEE Transactions on Wireless Communications*, 2014, 13(11): 6132–6146. doi: [10.1109/TWC.2014.2329877](https://doi.org/10.1109/TWC.2014.2329877).
- [21] GHADIMI E, CALABRESE F D, PETERS G, *et al.* A reinforcement learning approach to power control and rate adaptation in cellular networks[C]. 2017 IEEE International Conference on Communications, Paris, France, 2017: 1–7.
- [22] 周志华. 机器学习[M]. 北京: 清华大学出版社, 2016: 372–390.
- ZHOU Zhihua. Machine Learning[M]. Beijing: Tsinghua University Press, 2016: 372–390.
- 李 云: 男, 1974年生, 教授, 博士生导师, 主要研究领域为无线移动通信.
- 唐 英: 女, 1993年生, 硕士生, 研究方向为异构蜂窝无线网络.
- 刘涵霄: 男, 1994年生, 硕士生, 研究方向为异构蜂窝无线网络.