



研究与开发

联合功率控制和信道分配的蜂窝网络能效优化算法

徐会彬

(湖州师范学院信息工程学院, 浙江 湖州 313000)

摘要: 为了提高混合设备到设备 (D2D) 蜂窝网络中 D2D 干扰导致系统能效下降的问题, 提出了联合功率控制和信道分配的能效优化 (EEPC) 算法, 进而提升系统能效。以 D2D 用户和蜂窝用户最小速率为约束条件, 建立最大化能效的优化问题; 利用块坐标下降法将优化问题转化为信道分配和功率控制两个子问题, 再分别利用 Q 学习算法、Dinkelbach 算法和优化最小 (MM) 算法求解。并对 Q 学习算法中贪婪搜索因子进行改进, 采用动态的搜索因子, 平衡探索与利用间的关系。性能分析表明, 提出的 EEPC 算法提升了系统能效。

关键词: 混合 D2D 蜂窝网络; 能效; 信道分配; Q 学习; Dinkelbach 算法

中图分类号: TN929.5

文献标志码: A

doi: 10.11959/j.issn.1000-0801.2024004

Energy efficiency optimization under joint transmission power and channel allocation for cellular network

XU Huibin

School of Information Engineering, Huzhou University, Huzhou 313000, China

Abstract: In order to solve problem of low-energy efficiency caused by device-to-device (D2D) interference in hybrid D2D cellular network, energy efficiency improvement based on joint power control and channel allocation (EEPC) algorithm was proposed to improve the energy efficiency. The optimization problem of maximizing energy efficiency was established with the constraint of ensuring the minimum rate of D2D users and cellular users. By using block coordinate descent method, the optimization problem was transformed into two sub-problems of channel allocation and power control, which were solved by Q-learning algorithm, Dinkelbach algorithm and majorization-minimization (MM) respectively. The greedy search factor in Q-learning algorithm was improved, and a dynamic search factor was used to balance the relationship between exploration and exploitation. Performance analysis shows that the proposed EEPC algorithm improves the energy efficiency of the system.

Key words: hybrid D2D cellular network, energy efficiency, channel allocation, Q-learning, Dinkelbach algorithm

收稿日期: 2023-09-23; 修回日期: 2024-01-08

基金项目: 浙江省软科学研究计划项目 (No.2021C35129); 湖州市自然科学基金资助项目 (No.2021YZ20)

Foundation Items: Zhejiang Soft Science Research Program (No.2021C35129), Huzhou Natural Science Foundation Project (No.2021YZ20)

0 引言

第三代合作伙伴计划 (3rd Generation Partnership Project, 3GPP) 提出设备到设备 (device-to-device, D2D)^[1] 技术, 有效地缓解蜂窝核心网络负载, 提升了本地服务的数据速率。D2D 技术允许通信网中邻近设备间进行直接通信, 无须通过基站转发, 从而缓解核心网络的数据压力^[2]。因此, 将 D2D 技术应用于传统蜂窝网, 形成混合 D2D 蜂窝网可极大地提升网络吞吐量和频谱资源利用率^[3]。

D2D 用户通过复用蜂窝用户的频谱资源传输数据, 虽然提高了频谱利用率和吞吐量, 但也对现有蜂窝网络带来一定的干扰, 降低蜂窝用户的速率。因此, 解决 D2D 用户与蜂窝用户间干扰成为推进 D2D 技术亟待解决的问题。

由于 D2D 用户复用蜂窝用户的信道, 合理地选择复用信道以及传输功率是缓解 D2D 用户与蜂窝用户间干扰的有效途径。选择复用信道属于信道分配问题, 若能合理分配信道, 可有效地降低 D2D 用户对蜂窝用户的干扰。D2D 用户的传输功率、基站传输功率与干扰息息相关, 传输功率低, 则形成干扰的概率小, 但通信质量差, 用户端所获取的速率低。反之, 高的传输功率则会加剧相互干扰^[4]。

文献[5]利用博弈论权衡蜂窝用户和 D2D 用户的传输功率, 最大化系统能效。类似地, 文献[6-7]也采用博弈论方法, 求解最大化多用户速率的 D2D 传输功率 (解)。尽管采用博弈论能够以分布式方式决策传输功率, 但是利用博弈论所产生的解并非全局最优解。此外, 针对信道分配问题, 文献[8]采用反向迭代组合拍卖理论实现 D2D 用户复用蜂窝用户的信道。文献[9]采用 Coalition (联盟) 博弈法给 D2D 用户分配信道。

目前, 由于机器学习在无线通信领域应用的

拓展, 基于机器学习的传输功率控制方法受到广泛关注和研究。文献[10]提出基于深度 Q 网络的功率控制算法。通过学习环境, 优化传输功率。但是该算法需要为深度神经网络输入大量状态信息。并且该算法在控制功率的同时没有考虑到 D2D 用户与蜂窝用户间的信道分配问题。即让 D2D 用户复用蜂窝用户的哪个信道也是降低干扰、提升速率的关键。

据此, 针对混合 D2D 蜂窝网, 提出联合功率控制和信道分配的能效优化 (energy efficiency improvement based on joint power control and channel allocation, EEPC) 算法。该算法以满足 D2D 用户和蜂窝用户的最小速率为约束条件, 构建最大化能效的优化问题。利用块坐标下降法将优化问题分解成两个子问题。先利用 Q 学习算法为每个 D2D 用户对分配信道, 再利用 Dinkelbach 算法求解 D2D 用户和基站的功率分配方案。仿真结果表明, 在保证 D2D 用户和蜂窝用户的最小速率的基础上, EEPC 算法的能效优于同类算法。

1 系统模型

1.1 网络模型

考虑混合 D2D 蜂窝网场景, 其系统模型如图 1 所示。基站位于小区中心, 且小区半径为 R 。小区有 K 个蜂窝用户和 M 对 D2D 通信对 (D2D pair, D2D-P)。 K 个蜂窝用户和 M 对 D2D 通信对分别形成两个集: $\mathcal{K} = \{1, 2, \dots, K\}$ 和 $\mathcal{M} = \{1, 2, \dots, M\}$ 。D2D-P 复用蜂窝用户的信道。每个 D2D-P 中 D2D 发射设备 (D2D-T) 离 D2D 接收设备 (D2D-R) 间的最大距离为 d 。假定系统内有 N 个信道, 这些信道形成信道集 $\mathcal{N} = \{1, 2, \dots, N\}$ 。

假定小区基站能通过公共控制信道获取蜂窝用户的信道信息, 并由基站控制实施信道分配策略。

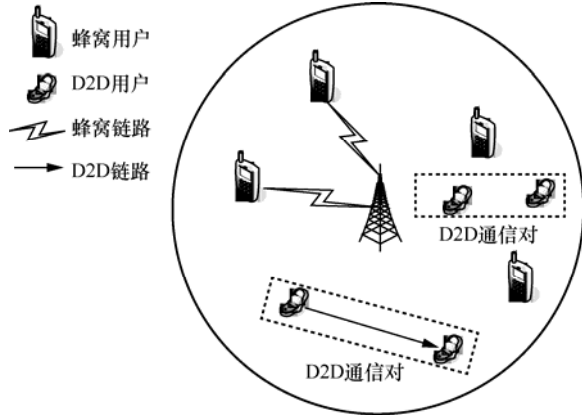


图1 系统模型

1.2 能效

将蜂窝用户和D2D用户速率之和与它们的总功耗的比值定义为系统能效 (energy efficiency, EE) [11]:

$$EE = \frac{R_{\text{tot}}}{P_{\text{tot}}} = \frac{\sum_{m=1}^M r_m + \sum_{k=1}^K r_k}{P_{\text{tot-D2D}} + P_{\text{tot-BS}}} \quad (1)$$

其中, $P_{\text{tot-D2D}}$ 、 $P_{\text{tot-BS}}$ 分别表示 D2D-P、基站总功耗; r_m 、 r_k 分别表示 D2D-P、蜂窝用户的速率和, 它们的定义分别如式 (2)、式 (3) 所示。

$$r_m = \sum_{n=1}^N \rho_{m,k}^{(n)} \log[1 + \Gamma_m^{(n)}] \quad (2)$$

$$r_k = \sum_{n=1}^N \rho_{m,k}^{(n)} \log[1 + \Gamma_k^{(n)}] \quad (3)$$

其中, $\rho_{m,k}^{(n)} \in \{0,1\}$ 为布尔变量, $\rho_{m,k}^{(n)} = 1$ 表示第 k 个蜂窝用户的信道 n 被第 m 个 D2D 用户复用; $\rho_{m,k}^{(n)} = 0$ 表示未复用此信道 n 。 $\Gamma_m^{(n)}$ 表示第 m 个 D2D 用户从子信道 n 上所接收信号的信号与干扰加噪声比 (signal to interference plus noise ratio, SINR) [12], 而 $\Gamma_k^{(n)}$ 表示第 k 个蜂窝用户在子信道 n 上所接收信号的 SINR。 $\Gamma_m^{(n)}$ 和 $\Gamma_k^{(n)}$ 的定义分别如式 (4) 和式 (5) 所示。

$$\Gamma_m^{(n)} = \frac{(p_{m,k}^{(n)} |h_m^{(n)}|^2)}{\sum_{k=1}^K \rho_{m,k}^{(n)} \alpha_k p_k^{(n)} |h_{m,k}^{(n)}|^2 + N_0} \quad (4)$$

$$\Gamma_k^{(n)} = \frac{\alpha_k p_k^{(n)} |h_k^{(n)}|^2}{\alpha_k \sum_{m=1}^M \rho_{m,k}^{(n)} (p_{m,k}^{(n)}) |h_{m,k}^{(n)}|^2 + N_0} \quad (5)$$

其中, $p_{m,k}^{(n)}$ 表示第 m 个 D2D 用户的传输功率, 下标 “ k ” 表示第 m 个 D2D 用户复用了第 k 个蜂窝用户的信道 n ; α_k 表示功率分配因子; $p_k^{(n)}$ 表示基站向第 k 个蜂窝用户的传输功率; N_0 表示 D2D-T 端的高斯白噪声; $h_k^{(n)}$ 表示第 k 个蜂窝用户的发送端与基站间的信道增益; $h_m^{(n)}$ 表示第 m 个 D2D 用户对间信道增益; $h_{m,k}^{(n)}$ 表示 D2D 用户对蜂窝用户所造成的干扰链路的信道增益。

D2D 用户复用信道如图 2 所示, 图 2 给出上述变量间关系。

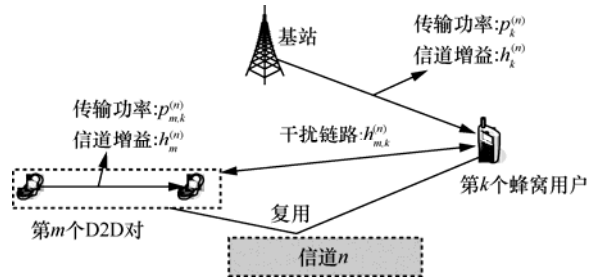


图2 D2D 用户复用信道

此外, 利用式 (6) 和式 (7) 分别计算式 (1) 中 $P_{\text{tot-D2D}}$ 和 $P_{\text{tot-BS}}$:

$$P_{\text{tot-D2D}} = \sum_{m=1}^M \sum_{n=1}^N \tau_i p_{m,k}^{(n)} + \sum_{m=1}^M p_m^{(\text{circuit})} \quad (6)$$

$$P_{\text{tot-BS}} = \sum_{k=1}^K \sum_{n=1}^N p_k^{(n)} + p_{\text{BS}}^{(\text{circuit})} \quad (7)$$

其中, $p_m^{(\text{circuit})}$ 和 $p_{\text{BS}}^{(\text{circuit})}$ 分别表示第 m 个 D2D 用户和蜂窝基站的电路功率; τ_i 表示传输信息数据的时间, 下标 i 表示信息数据。

1.3 Dinkelbach 算法概述

Dinkelbach 算法可求解分式规划问题:

$$\max_x \frac{f(x)}{g(x)}, \quad x \in \mathcal{X} \quad (8)$$

其中, $\mathcal{X}$ 表示变量 x 的值域。

通过引入变量 λ ，将原来分式规划问题转换成相减形式^[10]：

$$F(\lambda) = \max_x \{f(x) - \lambda g(x)\} \quad (9)$$

令 $\mathbf{x}^*$ 表示式 (8) 的最优解，其可表述为 $\mathbf{x}^* = \arg \max_x \{f(x) - \lambda^* g(x)\}$ 。其中， λ^* 表示函数 $F(\lambda)$ 的唯一零点^[13]。Dinkelbach 算法通过不断迭代，得到最优解 $\mathbf{x}^*$ 。

2 联合信道和功率分配算法

2.1 优化问题的构建

本文旨在通过优化 D2D 复用的信道和 D2D 用户传输功率以及基站传输功率，满足 D2D 用户和蜂窝用户的最低速率要求，最大化通信系统能效。因此，建立最大化能效的优化问题：

$$\max_{\rho_{m,k}^{(n)}, p_{m,k}, p_k} \text{EE} \quad (10)$$

$$\text{s.t. } r_m \geq R_{\min}^{\text{D2D}}, \forall m \in \mathcal{M} \quad (10a)$$

$$r_k \geq R_{\min}^{\text{CU}}, \forall k \in \mathcal{K} \quad (10b)$$

$$\rho_{m,k}^{(n)} \in \{0, 1\} \quad (10c)$$

$$\sum_{n=1}^N \sum_{k=1}^K \rho_{m,k}^{(n)} p_k^{(n)} \leq p_{\max}^{(\text{BS})}, \forall m \in \mathcal{M} \quad (10d)$$

$$\sum_{n=1}^N \sum_{m=1}^M \rho_{m,k}^{(n)} p_{m,k}^{(n)} \leq p_{\max}^{(\text{D2D})}, \forall k \in \mathcal{K} \quad (10e)$$

其中， $R_{\min}^{\text{D2D}}$ 、 $R_{\min}^{\text{CU}}$ 分别表示 D2D 用户、蜂窝用户的最小速率；约束项 (10a) 和 (10b) 满足每个 D2D 用户和蜂窝用户的最低速率要求； $p_{\max}^{(\text{BS})}$ 、 $p_{\max}^{(\text{D2D})}$ 分别表示基站、D2D 发射器的最大传输功率。

由于式 (10) 包含布尔变量 $\rho_{m,k}^{(n)}$ ，该优化问题为混合整数非线性问题。为此，采用块坐标下降法，将优化问题分解为两个子问题：信道分配和功率控制。首先，给定 D2D 用户和蜂窝用户的传输功率，利用 Q 学习算法求解信道分配问题，即选择最优的复用信道；其次，将信道分配的结果作为已知条件，对 D2D 用户传输功率 $p_{m,k}^{(n)}$ 和蜂窝用户的传输功率 $p_k^{(n)}$ 进行联合优化，进而最大化系统能效。

2.2 基于 Q 学习算法的信道分配

建立 Q 值表是 Q 学习算法的基础。Q 学习算法通过不断试错，向环境学习获取最优的奖励，并不断地更新 Q 表，最终得到一个收敛的 Q 表。利用此 Q 表指导智能体选择动作^[14]。

(1) 智能体：D2D 用户为智能体，最初随机地选择复用信道，并计算所获取的回报。

(2) 状态：将 D2D 用户的发射设备与接收设备间的信道增益、基站与蜂窝用户间的信道增益视为状态。

(3) 动作：将在每个状态为 D2D 用户选择复用哪个信道视为动作。

(4) 奖励：以式 (1) 为奖励函数。若智能体选择的动作满足式 (10a)、式 (10b)、式 (10c) 所示的约束条件，则以式 (1) 所计算的能效值为奖励，否则奖励值为零。

利用 Q 学习算法实现信道分配的步骤如下。

步骤 1 先对学习率、折扣率以及贪婪值等参数进行初始化，并随机地选择初始状态 s 。

步骤 2 采用 ε 贪婪策略选择动作。 ε 贪婪策略是指随机生成一个随机数，若随机数大于 ε ，就从当前 Q 表中选择具有最大 $Q(s, a)$ 所对应的动作；反之，就随机选择一个动作。

步骤 3 计算奖励值，并利用式 (11) 更新 Q 值：

$$Q(s, a) \leftarrow Q(s, a) + \alpha(r + \gamma \times \max_{a'} (Q(s', a')) - Q(s, a)) \quad (11)$$

其中， $Q(s, a)$ 表示在状态 s 下选择动作 a 下的 Q 值； α 为学习率； γ 为折扣率； $\max_{a'} (Q(s', a'))$ 表示在状态 s' 下最大的 Q 值； r 代表奖励。

步骤 4 判断是否达到迭代终止条件，即是否达到最大迭代次数。若达到就停止迭代，反之就跳转到步骤 2。

此外，传统 Q 学习算法采用固定的 ε 值^[15]。若 ε 值固定，则总是以相同概率随机地选择动作，并没有对选择行为进行约束，这降低了搜索效率^[16]。



实质上, 应在迭代初期增加 ε 值, 增强算法探索未知区域的能力, 即发挥探索能力; 在迭代后期应降低 ε 值, 提升算法利用已知信息的能力, 即发挥利用能力。为此, 本文不再采用固定 ε 值的机制, 而是动态搜索因子, 即 ε 值随迭代次数逐步减少。

$$\varepsilon = e^{\beta \times i} \quad (12)$$

其中, i 表示迭代次数; 参数 β 为衰减因子。文献[15]中取 $\beta = -0.90$ 。本文通过仿真对比分析发现, 当 $\beta = -0.95$ 时, 在本文实验环境下的收敛性能最好。因此, 在仿真中 β 取 -0.95 。

2.3 基于 Dinkelbach 算法的功率分配

利用 Dinkelbach 算法^[17]解决目标问题式 (10)。具体而言, 先定义两个功率矢量:

$$\begin{cases} \mathbf{p}^* = (p_{1,K}^{*(n)}, p_{2,K}^{*(n)}, \dots, p_{M,K}^{*(n)}) \\ \tilde{\mathbf{p}}^* = (p_1^{*(n)}, p_2^{*(n)}, \dots, p_K^{*(n)}) \end{cases} \quad (13)$$

其中, $\mathbf{p}^*$ 、 $\tilde{\mathbf{p}}^*$ 分别表示 D2D-P、蜂窝用户的最优传输功率矢量。

再利用式 (1) ~ 式 (5) 计算总速率 R_{tot} :

$$\begin{aligned} R_{\text{tot}} = & \sum_{m=1}^M \sum_{n=1}^N \tau_i \text{lb} \left[\sum_{k=1}^K \alpha_k p_k^{(n)} |h_{k,m}^{(n)}|^2 + N_0 + p_{m,k}^{(n)} |h_m^{(n)}|^2 \right] + \\ & \sum_{k=1}^K \sum_{n=1}^N \text{lb} \left[\alpha_k \sum_{m=1}^M p_{m,k}^{(n)} |h_{k,m}^{(n)}|^2 + N_0 + p_{m,k}^{(n)} |h_m^{(n)}|^2 \right] - \\ & \sum_{m=1}^M \tau_i \text{lb} \left[\alpha_k \sum_{k=1}^K p_k^{(n)} |h_{k,m}^{(n)}|^2 + N_0 \right] - \\ & \sum_{k=1}^K \text{lb} \left[\alpha_k \sum_{m=1}^M p_{m,k}^{(n)} |h_{k,m}^{(n)}|^2 + N_0 \right] \end{aligned} \quad (14)$$

为了求解式 (10), 将目标问题转化成两个凹函数差的形式。

$$\begin{aligned} R_{\text{tot}}(\tilde{\mathbf{p}}, \mathbf{p}) = & \left[\sum_{m=1}^M f_{1,m}(\tilde{\mathbf{p}}, \mathbf{p}) + \sum_{k=1}^K f_{2,k}(\tilde{\mathbf{p}}, \mathbf{p}) \right] - \\ & \left[\sum_{m=1}^M g_{1,m}(\tilde{\mathbf{p}}) + \sum_{k=1}^K g_{2,k}(\mathbf{p}) \right] \end{aligned} \quad (15)$$

其中, $f_{1,m}$ 、 $f_{2,k}$ 、 $g_{1,m}$ 、 $g_{2,k}$ 的定义如下。

$$\begin{cases} f_{1,m} = \sum_{n=1}^N \tau_i \text{lb} \left[\sum_{k=1}^K \alpha_k p_k^{(n)} |h_{k,m}^{(n)}|^2 + N_0 + p_{m,k}^{(n)} |h_m^{(n)}|^2 \right] \\ f_{2,k} = \sum_{n=1}^N \text{lb} \left[\alpha_k \sum_{m=1}^M p_{m,k}^{(n)} |h_{k,m}^{(n)}|^2 + N_0 + p_{m,k}^{(n)} |h_m^{(n)}|^2 \right] \\ g_{1,m} = \tau_i \text{lb} \left[\alpha_k \sum_{k=1}^K p_k^{(n)} |h_{k,m}^{(n)}|^2 + N_0 \right] \\ g_{2,k} = \text{lb} \left[\alpha_k \sum_{m=1}^M p_{m,k}^{(n)} |h_{k,m}^{(n)}|^2 + N_0 \right] \end{cases}$$

然后, 将式 (10) 所示的目标问题重构为:

$$\max_{\tilde{\mathbf{p}}, \mathbf{p}} \frac{R_{\text{tot}}(\tilde{\mathbf{p}}, \mathbf{p})}{P_{\text{tot}}(\tilde{\mathbf{p}}, \mathbf{p})} \quad (16)$$

$$\text{s.t. } f_{1,m}(\tilde{\mathbf{p}}, \mathbf{p}) - g_{1,m}(\tilde{\mathbf{p}}) \geq R_{\min}^{\text{D2D}} \quad (16a)$$

$$f_{2,k}(\tilde{\mathbf{p}}, \mathbf{p}) - g_{2,k}(\mathbf{p}) \geq R_{\min}^{\text{CU}} \quad (16b)$$

$$(10a), (10b), (10c), (10d) \quad (16c)$$

其中, $P_{\text{tot}}(\tilde{\mathbf{p}}, \mathbf{p})$ 表示总功耗, 即通过优化 $\tilde{\mathbf{p}}$ 、 $\mathbf{p}$ 两个变量后所产生的总功耗。

尽管式 (15) 的每一项都是凹函数, 但两个或多个凹函数的差未必是凹函数。为此, 采用优化最小 (majorization-minimization, MM) 算法, 并结合一阶泰勒展开式构建序列的替代函数^[18], 使其在邻近点 ($\mathbf{p}^{t-1}$ 和 $\tilde{\mathbf{p}}^{t-1}$) 呈凸面性。

$$\begin{aligned} \hat{g}_{1,m}(\tilde{\mathbf{p}}) = & g_{1,m}(\tilde{\mathbf{p}})^{t-1} + \nabla g_{1,m} \\ & (\tilde{\mathbf{p}})^{t-1} (\tilde{\mathbf{p}} - (\tilde{\mathbf{p}})^{t-1}) \geq g_{1,m}(\tilde{\mathbf{p}}) \end{aligned} \quad (17)$$

$$\begin{aligned} \hat{g}_{2,k}(\mathbf{p}) = & g_{2,k}(\mathbf{p})^{t-1} + \nabla g_{2,k} \\ & (\mathbf{p})^{t-1} (\mathbf{p} - (\mathbf{p})^{t-1}) \geq g_{2,k}(\mathbf{p}) \end{aligned} \quad (18)$$

再构建两式相减形式的目标问题, 如式 (19) 所示。

$$\max_{\tilde{\mathbf{p}}, \mathbf{p}} \hat{R}_{\text{tot}}(\tilde{\mathbf{p}}, \mathbf{p}) - \lambda P_{\text{tot}}(\tilde{\mathbf{p}}, \mathbf{p}) \quad (19)$$

$$\text{s.t. } f_{1,m}(\tilde{\mathbf{p}}, \mathbf{p}) - g_{1,m}(\tilde{\mathbf{p}}) \geq R_{\min}^{\text{D2D}} \quad (19a)$$

$$f_{2,k}(\tilde{\mathbf{p}}, \mathbf{p}) - g_{2,k}(\mathbf{p}) \geq R_{\min}^{\text{CU}} \quad (19b)$$

$$(10a), (10b), (10c), (10d) \quad (19c)$$

$$\text{其中, } \hat{R}_{\text{tot}}(\tilde{\mathbf{p}}, \mathbf{p}) = \left[\sum_{m=1}^M f_{1,m}(\tilde{\mathbf{p}}, \mathbf{p}) + \sum_{k=1}^K f_{2,k}(\tilde{\mathbf{p}}, \mathbf{p}) \right] - \left[\sum_{m=1}^M \hat{g}_{1,m}(\tilde{\mathbf{p}}) + \sum_{k=1}^K \hat{g}_{2,k}(\mathbf{p}) \right].$$

依据第 1.3 节对 Dinkelbach 算法的表述, 构建一个辅助函数 $F(\lambda) = \max_{\mathbf{x}} \{f(\mathbf{x}) - \lambda g(\mathbf{x})\}$, $f(\mathbf{x}) = \hat{R}_{\text{tot}}(\mathbf{x})$, $g(\mathbf{x}) = P_{\text{tot}}(\mathbf{x})$, 其中 $\mathbf{x}$ 包含 $\mathbf{p}^*$ 、 $\tilde{\mathbf{p}}^*$ 两个矢量。将 $\mathbf{p}^*$ 、 $\tilde{\mathbf{p}}^*$ 融合成一个矢量 $\mathbf{x}^*$, 则 $\mathbf{x}^*$ 表示 $\mathbf{x}$ 的最优解, 求解过程如算法 1 所示。

算法 1 基于 Dinkelbach 算法求解式 (19)

步骤 1 初始化迭代次数 $\ell = 0$, $f(x^{(0)})$, $g(x^{(0)})$, $\lambda^{(0)} = 0$ 令精度 $\varsigma > 0$, 且保证 $F(\lambda^{(0)}) > \varsigma$ 。

步骤 2 求解 $x^{(\ell+1)} = \arg \max_{\mathbf{x}} \{f(\mathbf{x}) - \lambda g(\mathbf{x})\}$, 并代入 $F(\lambda^{(\ell)})$ 中。

步骤 3 更新 $\lambda^{(\ell+1)} = f(x^{(\ell)}) / g(x^{(\ell)})$ 。

步骤 4 更新迭代次数 $\ell = \ell + 1$ 。

步骤 5 若 $F(\lambda^{(\ell)}) < \varsigma$, 则输出最优解 $x^{(\ell-1)}$, 否则, 返回步骤 2, 继续执行。

求解目标问题的框架如图 3 所示, 给出了求解式 (10) 的框架。先利用块坐标下降法将式 (10) 的目标问题分别成两个子问题, 并分别用改进的 Q 学习算法、Dinkelbach 算法和 MM 算法求解。

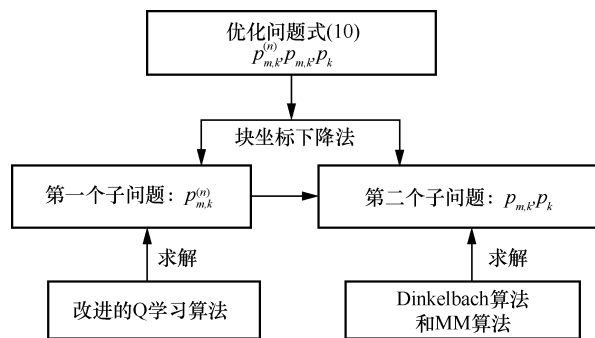


图 3 求解目标问题的框架

3 性能分析

3.1 仿真参数

利用 Python3.10 软件建立仿真平台, 在半径

为 200 m 蜂窝区内部署 M 对 D2D-P 和 K 个蜂窝用户。默认的仿真参数见表 1。

表 1 仿真参数

仿真参数	变量	值
蜂窝半径	R	200 m
D2D 对数	M	10~50
蜂窝用户数	K	10~50
D2D 传输距离	d	0~80 m
最大的 D2D 传输功率	$P_{\text{max}}^{\text{D2D}}$	23 dBm
最大的 BS 传输功率	$P_{\text{max}}^{\text{BS}}$	26 dBm
噪声功率	N_0	-100 dBm
D2D 用户的最小速率	$R_{\text{min}}^{\text{D2D}}$	2 bit·(Hz·s ⁻¹)
蜂窝用户的最小速率	$R_{\text{min}}^{\text{CU}}$	1 bit·(Hz·s ⁻¹)

此外, 通过蒙特卡洛仿真产生训练数据。在仿真中 D2D 用户和基站均以随机方式选择传输功率。同时, 也以随机方式给 D2D 分配复用信道。将仿真过程中收集的历史数据上传到基站, 这些信息包括状态、动作值、奖励值以及信道状况信息。并将这些信息存入经验池, 再从中随机地抽取一份批量数据, 用于训练模型。Q 学习算法的主要参数为: 训练的总局数为 6 000, 每局迭代次数为 100 次; 经验池容量为 2 000; 批量数据大小为 64; 学习率为 5×10^{-4} ; 折扣因子为 0.99。

3.2 动态搜索因子与固定搜索因子对 Q 学习收敛性能影响

首先分析采用动态搜索因子对 Q 学习算法的影响。图 4 给出了 $\varepsilon = 0.7$ 和 $\varepsilon = e^{-0.95i}$ 两种情况下的奖励值。不同轮数下的奖赏如图 4 所示, 从图 4 可知, 采用动态搜索因子 ($\varepsilon = e^{-0.95i}$) 收敛更快, 并且所收敛的奖励值也略优于 $\varepsilon = 0.7$ 下的奖励值。原因在于: 在基于动态搜索因子环境下, 最初探索能力强, 能尝试新的区域, 在搜索后期, 增强了利用能力, 能充分挖掘现有知识能力。

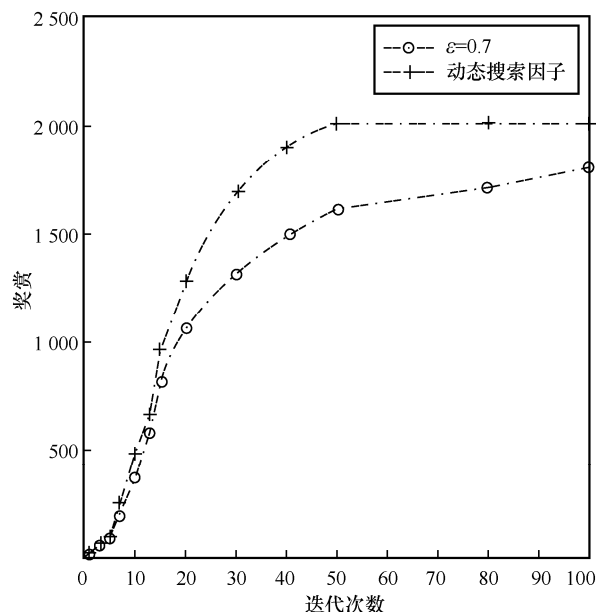


图4 不同轮数下的奖赏

3.3 Dinkelbach 算法的迭代次数对能效的影响

Dinkelbach 算法的迭代次数对 EEPC 算法能效的影响如图 5 所示, 其中 $d = 50 \text{ m}$ 。

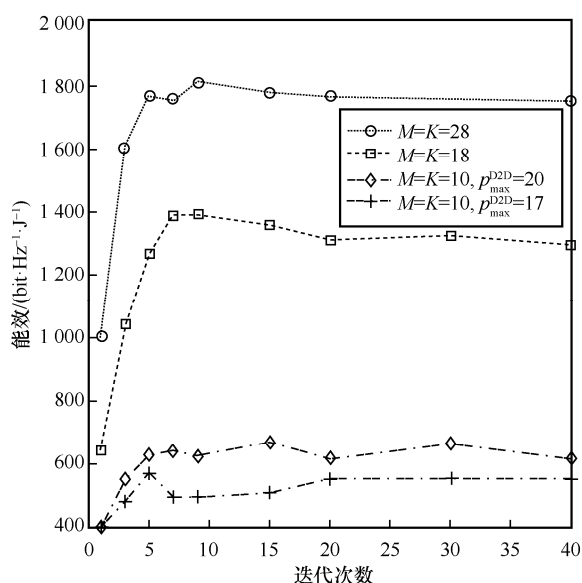


图5 Dinkelbach 算法的迭代次数对 EEPC 算法能效的影响

从图 5 可知, EEPC 算法的最初能效较低。经 4 次迭代后, EEPC 算法的能效值就得到收敛。这说明利用 Dinkelbach 算法求解功率分配的有效性。同时, D2D 用户数和蜂窝用户数的增加,

不影响 Dinkelbach 算法的收敛性能。此外, 观察图 5 不难发现, D2D 用户数和蜂窝用户数越多, 能效越高。当 D2D 用户数和蜂窝用户数固定时, 增加 $p_{\max}^{\text{D2D}}$ 的值降低了能效, 这符合逻辑。传输功率越高, 所消耗的能量就越大。

3.4 不同算法的能效对比分析

为了更好地分析 EEPC 算法的性能, 选择 3 个基准算法进行性能分析。

(1) 传输功率固定的最优子信道分配 (fixed-power optimization sub-channel allocation, FPOS) 算法。在 FPOS 算法中, D2D 用户的功率为 $p_{\max}^{\text{D2D}}$, 基站传输功率为 $p_{\max}^{\text{BS}}$, 采用最优信道分配策略。

(2) 传输功率固定的随机子信道分配 (fixed-power random sub-channel allocation, FPRS) 算法。在 FPRS 算法中, D2D 用户的功率为 $p_{\max}^{\text{D2D}}$, 基站传输功率为 $p_{\max}^{\text{BS}}$; 采用随机方式分配信道。EEPC、FPOS 和 FPRS 算法的异同之处见表 2。

表2 EEPC、FPOS 和 FPRS 算法的异同之处

	信道分配	功率分配
EEPC	改进的 Q 学习算法的信道分配	Dinkelbach 算法和 MM 算法
FPOS	改进的 Q 学习算法的信道分配	不进行功率分配, D2D 用户基站以最大功率传输信号
FPRS	以随机方式进行信道分配	不进行功率分配, D2D 用户和基站以最大功率传输信号

接下来, 对比分析 FPOS、FPRS 和 EEPC 算法的能效性能。

图 6 所示为 FPOS、FPRS 和 EEPC 算法的能效随 D2D 用户数和蜂窝用户数的变化情况, 其中 $d = 20 \text{ m}$, $M = K = 10 \sim 80$ 。从图 6 可知, 相比于 FPOS 和 FPRS 算法, EEPC 算法的能效性能更优。原因在于: EEPC 算法通过合理选择 D2D 用户选择复用信道, 缓解了干扰。同时, EEPC 算法优化 D2D 用户和蜂窝用户功率, 提升了系统能效。

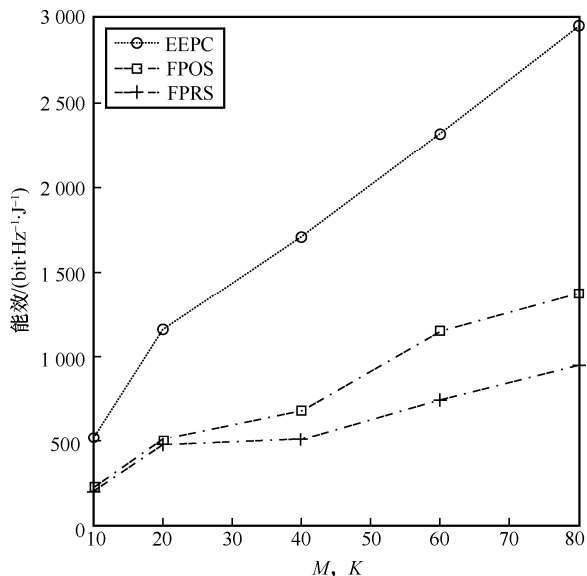


图6 FPOS、FPRS和EEPC算法的能效随D2D用户数和蜂窝用户数的变化情况

图7所示为传输距离 d 对FPOS、FPRS和EEPC算法能效的影响，其中 $M=K=30$ ， $d=10\sim 80$ 变化。

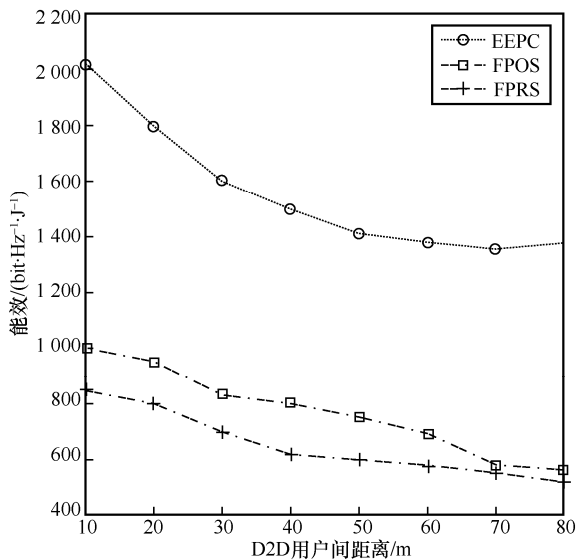


图7 传输距离 d 对FPOS、FPRS和EEPC算法能效的影响

从图7可知，能效随距离 d 的增加而下降。原因在于：当距离 d 增加，D2D用户需要更高的传输功率才能满足它们的最小速率的要求，这就增加了D2D用户的功耗。相比于FPOS和FPRS算法，EEPC算法提高了系统能效。例如，在

$d=10$ m时，FPOS和FPRS算法的能效分别约为1000和810，EEPC算法的能效约为2000。同FPOS、FPRS算法相比，EEPC算法的能效分别提升约100%、140%。尽管随着距离 d 的增加，EEPC算法的能效有所下降，但仍高于FPOS和FPRS算法的能效。

图8所示为能效随基站最大传输功率的变化情况。从图8可知，最初能效随基站最大传输功率的增加而上升。原因在于：最初基站传输功率的增加，提升蜂窝用户的通信质量，增加了蜂窝用户的速率。当基站传输功率增加至35 dBm时，能效达到峰值。随后，当基站传输功率继续增加，能效下降。原因在于：当基站持续增加传输功率，功率过大就容易造成干扰，降低了信噪比，进而无法再提升用户速率，能效下降。

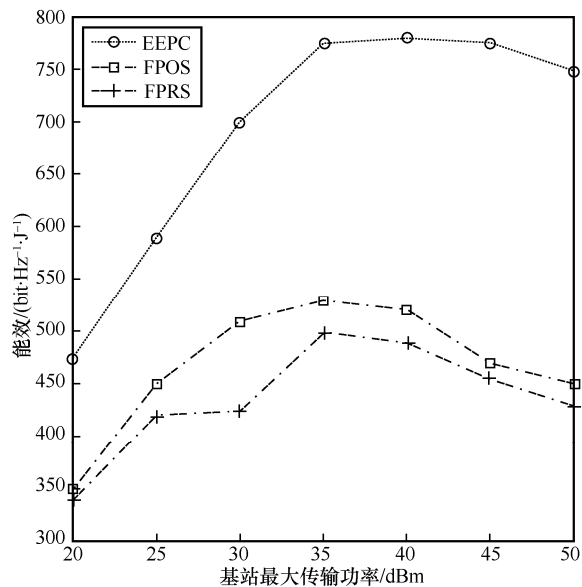


图8 能效随基站最大传输功率的变化情况

4 结束语

为了提升系统能效，提出EEPC算法。EEPC算法从优化D2D用户和蜂窝用户功率以及D2D用户所复用的信道，提升系统能效。利用强化学习的Q学习算法求解最优的信道分配策略，并联合Dinkelbach和MM算法对功率进行控制，降低



干扰,提升系统能效。性能分析表明,相比于 FPOS 和 FPRS 算法,EEPC 算法提升了系统能效。

本文只考虑单个小区的应用场景,后面将考虑多个小区的应用场景,并且相邻小区间也存在干扰。尽管考虑多个小区增加了算法的复杂性,但通过考虑多个小区能更进一步地优化功率分配策略,减少相邻小区间的干扰,进一步地提升系统能效。这将是后期的研究工作。

参考文献:

- [1] 易鑫,裴二荣. D2D辅助NB-IoT系统的传输设备数优化算法[J]. 重庆邮电大学学报(自然科学版), 2021, 33(4): 606-613.
YI X, PEI E R. Transmission device number optimization algorithm for D2D-assisted NB-IoT system[J]. Journal of Chongqing University of Posts and Telecommunications (Natural Science Edition), 2021, 33(4): 606-613.
- [2] 陶静,朱琦. 基于 NOMA 的 D2D 通信联合子信道与功率分配算法[J]. 南京邮电大学学报(自然科学版), 2018, 38(3): 32-39.
TAO J, ZHU Q. Joint subchannel and power allocation algorithm in D2D communication based NOMA[J]. Journal of Nanjing University of Posts and Telecommunications (Natural Science Edition), 2018, 38(3): 32-39.
- [3] 余翔,张海波,杨路. 混合 D2D 蜂窝网络中基于模拟退火算法的资源调度策略[J]. 电信科学, 2017, 33(4): 71-77.
YU X, ZHANG H B, YANG L. Resource scheduling strategy based on simulated annealing algorithm in hybrid D2D cellular networks[J]. Telecommunications Science, 2017, 33(4): 71-77.
- [4] 张先超,赵耀,叶海军,等. 无线网络多用户干扰下智能发射功率控制算法[J]. 通信学报, 2022, 43(2): 15-21.
ZHANG X C, ZHAO Y, YE H J, et al. Intelligent transmit power control algorithm for the multi-user interference of wireless network[J]. Journal on Communications, 2022, 43(2): 15-21.
- [5] CHEN X, HU R Q, JEON J, et al. Energy efficient resource allocation for D2D communication underlaying cellular networks[C]//Proceedings of the 2015 IEEE International Conference on Communications (ICC). Piscataway: IEEE Press, 2015: 2943-2948.
- [6] ABRARDO A, MORETTI M. Distributed power allocation for D2D communications underlaying/overlaying OFDMA cellular networks[J]. IEEE Transactions on Wireless Communications, 2017, 16(3): 1466-1479.
- [7] LYU J B, CHEW Y H, WONG W C. A stackelberg game model for overlay D2D transmission with heterogeneous rate requirements[J]. IEEE Transactions on Vehicular Technology, 2016, 65(10): 8461-8475.
- [8] WANG F R, XU C, SONG L Y, et al. Energy-efficient radio resource and power allocation for device-to-device communication underlaying cellular networks[C]//Proceedings of the 2012 International Conference on Wireless Communications and Signal Processing (WCSP). Piscataway: IEEE Press, 2012: 1-6.
- [9] 杨佳颖,李汀,解培中. NOMA 增强型 D2D 通信的资源分配算法[J]. 信号处理, 2021, 37(8): 1441-1451.
YANG J Y, LI T, XIE P Z. NOMA enhanced D2D communication resource allocation algorithm[J]. Journal of Signal Processing, 2021, 37(8): 1441-1451.
- [10] NASIR Y S, GUO D N. Multi-agent deep reinforcement learning for dynamic power allocation in wireless networks[J]. IEEE Journal on Selected Areas in Communications, 2019, 37(10): 2239-2250.
- [11] 王鹏杰. 基于无线能量传输的无人机协作通信网络性能优化[D]. 北京: 北京邮电大学, 2021.
WANG P J. Performance optimization of UAV cooperative communication network based on wireless energy transmission [D]. Beijing: Beijing University of Posts and Telecommunications, 2021.
- [12] 肖宇,邓正宏. 信噪比约束下基于互信息的雷达波形设计[J]. 系统工程与电子技术, 2021, 43(7): 1775-1780.
XIAO Y, DENG Z H. MI-based radar waveform design under SINR constraint[J]. Systems Engineering and Electronics, 2021, 43(7): 1775-1780.
- [13] 廖雅婷. 智能超表面辅助的蜂窝网络多域资源优化技术研究[D]. 成都: 电子科技大学, 2022.
LIAO Y T. Research on multi-domain resource optimization technology of cellular network assisted by intelligent hypersurface[D]. Chengdu: University of Electronic Science and Technology of China, 2022.
- [14] 李威,张晓东,姜学峰,等. 基于改进强化学习的机器人路径规划研究[J]. 制造业自动化, 2023, 45(3): 148-151, 172.
LI W, ZHANG X D, JIANG X F, et al. Robot path planning research based on improved reinforcement learning[J]. Manufacturing Automation, 2023, 45(3): 148-151, 172.
- [15] 唐静远,魏文军. 离散多智能体有限时间 Q 学习协同输出调节[J]. 计算机应用研究, 2023, 40(1): 204-208.
TANG J Y, WEI W J. Cooperative output regulation of finite time Q-learning for discrete multi-agents[J]. Application Research of Computers, 2023, 40(1): 204-208.
- [16] 胡晓辉. 一种基于动态参数调整的强化学习动作选择机制[J]. 计算机工程与应用, 2008, 44(28): 29-31, 48.
HU X H. Action choice mechanism of reinforcement learning based on adjusted dynamic parameters[J]. Computer Engineering and Applications, 2008, 44(28): 29-31, 48.
- [17] DINKELBACH W. On nonlinear fractional programming[J]. Management Science, 1967, 13(7): 492-498.
- [18] KHA H H, TUAN H D, NGUYEN H H. Fast global optimal power allocation in wireless networks by local D.C. programming[J]. IEEE Transactions on Wireless Communications, 2012, 11(2): 510-515.

[作者简介]



徐会彬(1982-),男,博士,湖州师范学院信息工程学院讲师,主要研究方向为 VANET 安全、路由技术。