

## D2D 中继辅助通信的能效优化算法研究

王雪<sup>1</sup>, 金涛<sup>1</sup>, 钱志鸿<sup>1</sup>, 胡良帅<sup>1</sup>, 王鑫<sup>1,2</sup>

(1. 吉林大学通信工程学院, 吉林 长春 130012; 2. 吉林农业大学信息技术学院, 吉林 长春 130118)

**摘 要:** 针对蜂窝系统下 D2D 中继辅助通信的能量效率优化问题, 在保证蜂窝用户与 D2D 用户服务质量的条件下, 提出了一种能效优化算法。首先, 以最大化系统中总 D2D 用户能效为目标, 将优化问题建模为一个混合整数非线性规划问题。然后, 将能效优化问题转化为功率控制、中继选择和信道分配 3 个子问题分别进行求解。最后, 利用 Dinkelbach 方法和拉格朗日乘子方法解决功率控制问题; 利用所提出的基于 Q 学习的中继选择算法完成中继的选择; 基于功率控制和中继选择, 利用匈牙利算法完成信道的分配。仿真结果表明, 与现有算法相比, 所提算法可有效提高系统 D2D 用户的总能效。

**关键词:** D2D 通信; 中继选择; 功率控制; 信道分配; 能量效率

**中图分类号:** TP393

**文献标识码:** A

**doi:** 10.11959/j.issn.1000-436x.2020048

## Research on maximizing energy efficiency for relay-aided D2D communication

WANG Xue<sup>1</sup>, JIN Tao<sup>1</sup>, QIAN Zhihong<sup>1</sup>, HU Liangshuai<sup>1</sup>, WANG Xin<sup>1,2</sup>

1. College of Communication Engineering, Jilin University, Changchun 130012, China

2. College of Information & Technology, Jilin Agricultural University, Changchun 130118, China

**Abstract:** In order to optimize the energy efficiency for relay-aided D2D communication underlying cellular network, an algorithm was proposed while guaranteeing the quality of service of D2D users and cellular users. Firstly, the problem was formulated as a mixed integer nonlinear fractional programming problem, with the goal of maximizing the energy efficiency of all D2D users. Then, the maximizing energy efficiency problem was divided into three sub-problems, such as power allocation, relay selection and subcarrier assignment. Finally, the power allocation problem, the relay selection problem and the subcarrier assignment problem were addressed by Dinkelbach method and Lagrange multiplier method, Q-learning and Hungarian algorithm, respectively. The simulation results show that the proposed algorithm can improve the total energy efficiency of D2D users compared with existing algorithms.

**Key words:** D2D communication, relay selection, power control, subcarrier assignment, energy efficiency

### 1 引言

作为 5G 的关键技术之一, D2D (device-to-device) 通信技术是指通信网络中邻近设备不通过基站而直接交换信息的技术<sup>[1]</sup>。在无线通信系统中加入 D2D 通信技术, 不仅可以提高系统频谱效率, 降低时延, 还可以减轻核心网络的负载<sup>[2]</sup>。但是,

当 D2D 通信中用户间距离较远或者链路质量较差时, 需要引入中继来保证信息的可靠、有效传输。并且, 引入中继可以扩大蜂窝系统的覆盖范围<sup>[3-4]</sup>。随着用户设备的大量增长以及终端设备电池容量的限制, 如何有效地提高能量效率, 实现绿色通信, 是未来无线通信发展的关键。

现有的文献大多研究的是蜂窝通信与 D2D 直

收稿日期: 2020-01-20; 修回日期: 2020-02-20

基金项目: 国家自然科学基金资助项目 (No.61771219); 吉林大学基础科研基金资助项目 (No.SXGJQY2017-9, No.2017TD-19)

**Foundation Items:** The National Natural Science Foundation of China (No.61771219), Fundamental Research Funds of Jilin University (No.SXGJQY2017-9, No.2017TD-19)

连通信并存场景下的能效优化问题。文献[5]通过使用联合迭代拍卖算法,优化了包括蜂窝用户在内的所有终端设备的能效。文献[6]考虑 D2D 通信复用蜂窝链路下行频谱资源,在满足蜂窝用户 QoS (quality of service) 的条件下,最大化系统总 D2D 链路的能量效率。文献[7]则考虑的是 D2D 通信复用蜂窝链路上行频谱资源,基于能效构造 D2D 链路 with 蜂窝链路之间的二分图,利用盖尔-沙普利算法解决 D2D 用户与蜂窝用户之间的匹配问题。文献[8]通过联合解决 D2D 直连通信下的功率控制和信道分配问题,最大化系统 D2D 用户对能效。文献[9]将非凸形式的能效优化问题转换为凸函数进行求解,提出了一种联合信道分配和功率控制的两层优化算法。文献[10]在 D2D 直连通信中加入了能量采集 (EH, energy harvesting) 技术,研究了包含能量采集时隙、功率和信道分配的能效优化问题。

在蜂窝通信与 D2D 中继通信并存场景下,可以从模式选择、功率控制、中继选择/信道分配等多个方面去优化能效。文献[11-12]考虑从功率控制角度优化能效,不同的是,前者利用非线性整数规划方法解决了单个 D2D 中继链路的能效最大化问题,后者则利用在线学习策略从全局角度最大化系统的能量效率。考虑功率控制和信道分配 2 个方面,文献[13]首先最大化 D2D 中继通信的第一跳链路和第二跳链路的能效,然后基于两跳链路的能效值,采用匹配算法选择合适的中继,最大化系统中 D2D 链路的总能效。文献[14]假设中继具有从射频信号中获得能量的 EH 技术功能,研究 D2D 通信中的功率控制和中继选择问题。文献[15]提出了一种基于能效最优的模式选择策略,在确定最优中继数量的条件下,通过选择使用中继通信或者协同通信的方式来优化系统的能效。文献[16]考虑了 D2D 通信的直连、中继及协同 3 种通信方式共存的场景,通过将优化问题拆分成联合模式选择、功率控制和信道分配 3 个子问题分别求解,提高了系统的 D2D 链路能效。

文献[11-16]大多只是从功率控制和中继选择,或者功率控制和信道分配两方面考虑优化 D2D 中继通信的能效,缺乏联合三者的研究,因此,本文在 D2D 中继辅助通信系统中,通过联合考虑功率控制、中继选择以及信道分配以最大化系统 D2D 用户的总能效。本文首先建模能效优化问题,并将

能效优化问题拆分为功率控制、中继选择和信道分配 3 个子问题。然后,利用非线性整数规划及凸优化理论解决功率控制问题,利用强化学习中的 Q 学习算法解决中继选择问题。最后,基于功率控制和中继选择,结合匹配理论,使用匈牙利算法求解信道分配问题。仿真结果表明,与现有文献对比,本文所提的算法在保证蜂窝用户和 D2D 用户的传输速率的条件下,有效地增加 D2D 用户的能量效率。

## 2 系统模型及问题描述

### 2.1 系统模型

假设在单小区蜂窝系统中,存在  $N$  对 D2D 用户、 $L$  个理想中继以及  $K$  个蜂窝用户,其中,一对 D2D 用户包含一个 D2D 发射端和一个 D2D 接收端。D2D 用户对索引集合表示为  $M=\{1,2,\dots,N\}$ , D2D 发射端的索引集合表示为  $S=\{1,2,\dots,N\}$ , D2D 接收端的索引集合表示为  $D=\{1,2,\dots,N\}$ , 中继的索引集合表示为  $R=\{1,2,\dots,L\}$ , 蜂窝用户的索引集合表示为  $C=\{1,2,\dots,K\}$ 。假设 D2D 用户的发射端与接收端之间由于链路质量较差,没有直连通路,只能通过中继进行通信,且中继均采用放大转发协议,同时每个蜂窝用户都已经预先分配了正交信道。系统模型如图 1 所示,假设某一 D2D 用户对  $m=(s,d)(m\in M,s\in S,d\in D)$  复用蜂窝用户  $c\in C$  的上行链路频谱资源,通过中继  $r\in R$  进行通信。在 D2D 中继通信第一跳链路中,中继  $r$  和基站的信干噪比 (SINR, signal to interference plus noise ratio) 分别为

$$\text{SINR}_D^{(r)} = \frac{P_s G_{sr}}{P_c G_{cr} + N_0} \tag{1}$$

$$\text{SINR}_{BS}^{(r)} = \frac{P_c G_{cb}}{P_s G_{sb} + N_0} \tag{2}$$

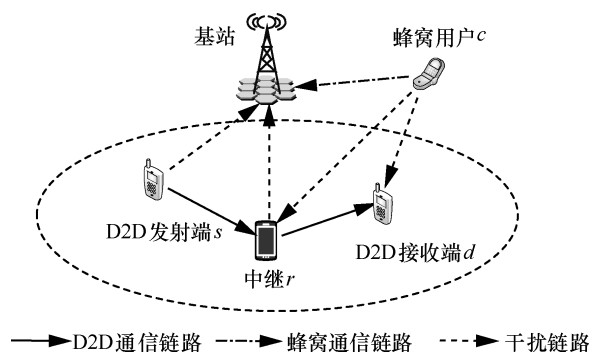


图 1 系统模型

其中,  $P_s$  和  $P_c$  分别表示 D2D 发射端和蜂窝用户的传输功率,  $G_{sr}$ 、 $G_{sb}$ 、 $G_{cr}$  和  $G_{cb}$  分别表示 D2D 发射端到中继、D2D 发射端到基站、蜂窝用户到中继和蜂窝用户到基站的信道增益,  $N_0$  表示加性高斯白噪声。

在 D2D 中继通信第二跳链路中, D2D 接收端  $d$  和基站的信干噪比分别为

$$\text{SINR}_D^{(r_2)} = \frac{P_r G_{rd}}{P_c G_{cd} + N_0} \quad (3)$$

$$\text{SINR}_{\text{BS}}^{(r_2)} = \frac{P_c G_{cb}}{P_r G_{rb} + N_0} \quad (4)$$

其中,  $P_r$  表示中继的传输功率,  $G_{rd}$ 、 $G_{rb}$  和  $G_{cd}$  分别表示中继到 D2D 接收端、中继到基站和蜂窝用户到 D2D 接收端的信道增益。

## 2.2 问题描述

根据式(1)和式(3), D2D 发射端  $s$  到 D2D 接收端  $d$  的能效表达式为

$$\text{EE}_{s,d} = x_{mr} \frac{W \text{lb}[1 + \min(\text{SINR}_D^{(r_1)}, \text{SINR}_D^{(r_2)})]}{\frac{1}{\eta}(P_s + P_r) + 2P_{\text{cir}}} \quad (5)$$

其中,  $x_{mr}$  表示中继选择因子, 当 D2D 用户对  $m$  通过中继  $r$  完成通信时  $x_{mr}=1$ , 否则  $x_{mr}=0$ ;  $P_{\text{cir}}$ 、 $W$  和  $\eta$  分别表示电路功率、信道带宽和功率放大系数。

本文旨在最大化整个系统的 D2D 用户对的能效, 具体表达式为

$$\max_{X, Y, P_s, P_r, P_c} \sum_{m \in M} \sum_{c \in C} \sum_{r \in R} y_{mc} \text{EE}_{s,d} \quad (6)$$

$$\text{s.t.} \sum_{r \in R} x_{mr} \leq 1, \forall m \in M, \sum_{m \in M} x_{mr} \leq 1, \forall r \in R \quad (7)$$

$$x_{mr} \in \{0, 1\}, \forall m \in M, \forall r \in R \quad (8)$$

$$\sum_{c \in C} y_{mc} \leq 1, \forall m \in M, \sum_{m \in M} y_{mc} \leq 1, \forall c \in C \quad (9)$$

$$y_{mc} \in \{0, 1\}, \forall m \in M, \forall c \in C \quad (10)$$

$$0 \leq P_s \leq P_{\text{max}}^D, \forall s \in S \quad (11)$$

$$0 \leq P_r \leq P_{\text{max}}^D, \forall r \in R \quad (12)$$

$$0 \leq P_c \leq P_{\text{max}}^C, \forall c \in C \quad (13)$$

$$W \text{lb}(1 + \text{SINR}_D^{(r_1)}) \geq R_{\text{min}}^D \quad (14)$$

$$W \text{lb}(1 + \text{SINR}_D^{(r_2)}) \geq R_{\text{min}}^D \quad (15)$$

$$\frac{W}{2} [\text{lb}(1 + \text{SINR}_{\text{BS}}^{(r_1)}) + \text{lb}(1 + \text{SINR}_{\text{BS}}^{(r_2)})] \geq R_{\text{min}}^C \quad (16)$$

其中,  $X$  表示中继选择矩阵, 矩阵  $X$  第  $m$  行、第  $r$

列的元素是中继选择因子  $x_{mr}$ ;  $Y$  表示信道选择矩阵, 矩阵  $Y$  第  $m$  行、第  $c$  列的元素是信道复用因子  $y_{mc}$ , 当 D2D 对  $m$  复用蜂窝用户  $c$  的信道通信时  $y_{mc}=1$ , 否则  $y_{mc}=0$ ; 向量  $P_s = \{P_s\}_{s \in S}$ ,  $P_r = \{P_r\}_{r \in R}$  和  $P_c = \{P_c\}_{c \in C}$  分别表示 D2D 发射端、中继和蜂窝用户的传输功率向量;  $P_{\text{max}}^D$  表示 D2D 发射端和中继的最大传输功率;  $P_{\text{max}}^C$  表示蜂窝用户的最大传输功率;  $R_{\text{min}}^D$  表示 D2D 通信链路必须满足的最小传输速率;  $R_{\text{min}}^C$  表示蜂窝用户必须满足的最小传输速率。

约束式(7)和式(8)保证一个中继只能服务一个 D2D 用户对, 并且一个 D2D 用户对只能复用一个中继。类似地, 约束式(9)和式(10)保证一条信道只能被一个 D2D 用户对复用, 并且一个 D2D 用户对只能复用一条蜂窝用户的信道。约束式(11)~式(13)分别是 D2D 发射端、中继和蜂窝用户的传输功率限制。约束式(14)~式(16)分别是 D2D 第一跳链路、D2D 第二跳链路和蜂窝用户链路的传输速率限制。

由多个约束条件和变量可以看出, 本文的能效优化问题是一个混合整数非线性规划 (MINLP, mixed integer non-linear programming) 的 NP-hard 问题, 而 NP-hard 问题没有一个有效的直接求解方法。为了求解这个优化问题, 本文将其分为功率控制、中继选择和信道分配 3 个子问题进行求解。

## 3 功率控制、中继选择和信道分配算法

### 3.1 基于 Dinkelbach 方法和拉格朗日乘子法的功率控制算法

假设 D2D 用户对  $m$  通过中继  $r$  进行通信, 并且复用蜂窝用户  $c$  的信道资源, 则  $x_{mr}=1$  且  $y_{mc}=1$ , D2D 发送端  $s$  到接收端  $d$  的能效优化问题为

$$\max_{P_s, P_r, P_c} \text{EE}_{s,d} \quad (17)$$

s.t. 式 (11)~式(16)

由式(5)可得出, D2D 发射端到接收端的能效是关于蜂窝用户功率  $P_c$  的减函数, 所以为了最大化 D2D 发射端到接收端的能效, 功率  $P_c$  需要取得最小值。根据约束式(13)、式(16)和文献[11], 得出蜂窝用户最小传输功率为

$$P_c^{\text{min}} = \min \left\{ P_{\text{max}}^C, \frac{\left( 2^{\frac{R_{\text{min}}^D}{W}} - 1 \right)}{G_{cb}} \left[ \max(P_{\text{max}}^D G_{sb}, P_{\text{max}}^D G_{rb}) + N_0 \right] \right\} \quad (18)$$

在计算出蜂窝用户功率最小传输功率  $P_c^{\text{min}}$  之

后, 优化式(17)中仍然有 2 个变量  $P_s$  和  $P_r$ , 并且式(17)是一个分数形式, 其分子不是一个确定的表达式, 不能通过已有的数学方法直接求解。在文献[11]中已经证明, 只有满足  $\text{SINR}_D^{(n)} = \text{SINR}_D^{(s)}$  的条件, 才能得到 D2D 发射端到 D2D 接收端的最优能效值。因此, 本文在  $\text{SINR}_D^{(n)} = \text{SINR}_D^{(s)}$  条件下, 将得到的蜂窝用户最小传输功率  $P_c^{\min}$  代入式(17)所示的优化问题中, 优化问题可以改写为

$$\max_{P_s} \frac{W\text{lb}(1 + A_0 P_s)}{\frac{1}{\eta} \left( 1 + \frac{A_0}{B_0} \right) P_s + 2P_{\text{cir}}} \quad (19)$$

s.t. 式(14) 和式(20)

$$0 \leq P_s \leq P_{\text{th}}^D \quad (20)$$

其中,  $A_0 = \frac{G_{sr}}{P_c^{\min} G_{cr} + N_0}$ ,  $B_0 = \frac{G_{rd}}{P_c^{\min} G_{cd} + N_0}$ ,

$P_{\text{th}}^D = \min \left\{ P_{\text{max}}^D, \frac{P_{\text{max}}^D B_0}{A_0} \right\}$  为 D2D 发射端的功率上界。

式(19)是一个分数形式的非凸函数, 所以利用 Dinkelbach 方法将其转换成等价的凸函数形式。首先, 用变量  $q_1^*$  表示 D2D 发射端到接收端的能效值, 并且定义式(19)所示的最大能效为

$$q_1^* = \max_{P_s} \frac{W\text{lb}(1 + A_0 P_s)}{\frac{1}{\eta} \left( 1 + \frac{A_0}{B_0} \right) P_s + 2P_{\text{cir}}} = \frac{W\text{lb}(1 + A_0 P_s^*)}{\frac{1}{\eta} \left( 1 + \frac{A_0}{B_0} \right) P_s^* + 2P_{\text{cir}}} \quad (21)$$

其中,  $P_s^*$  是能效取最大值时 D2D 发射端的传输功率。由文献[17]有引理 1 成立。

**引理 1** 当且仅当以下条件成立时得到  $q_1^*$

$$\max_{P_s} W\text{lb}(1 + A_0 P_s) - q_1^* \left[ \frac{1}{\eta} \left( 1 + \frac{A_0}{B_0} \right) P_s + 2P_{\text{cir}} \right] = 0 \quad (22)$$

**证明** 证明过程请参考文献[17]。

证毕。

由引理 1 可将优化问题式(19)从分数形式转化为等价的减式形式, 即

$$\max_{P_s} \left\{ W\text{lb}(1 + A_0 P_s) - q_1^* \left[ \frac{1}{\eta} \left( 1 + \frac{A_0}{B_0} \right) P_s + 2P_{\text{cir}} \right] \right\} \quad (23)$$

s.t. 式(14)和式(20)

式(23)为一个带约束条件的凸优化问题, 可使用拉格朗日乘法求解, 但是所得结果与  $q_1^*$  有关, 为了获得  $q_1^*$  的值, 本文通过迭代进行计算。假设在第  $n$  次迭代中功率变量为  $P_s(n)$  以及能效值为  $q_1(n-1)$ , 则式(23)的增广拉格朗日式为

$$L(P_s, \delta, \theta) = W\text{lb}(1 + A_0 P_s(n)) - q_1(n-1) \left[ \frac{1}{\eta} \left( 1 + \frac{A_0}{B_0} \right) P_s(n) + 2P_{\text{cir}} \right] - \delta(n) [P_s(n) - P_{\text{th}}^D] + \theta(n) \{ W\text{lb}(1 + A_0 P_s(n)) - R_{\text{min}}^D \} \quad (24)$$

其中,  $\delta(n)$  和  $\theta(n)$  分别是约束式(14)和式(20)的拉格朗日算子。根据拉格朗日对偶分解, 式(19)的能效优化问题等价于

$$\min_{(\delta, \theta \geq 0)} \max_{(P_s)} L(P_s, \delta, \theta) \quad (25)$$

由 KKT(Karush-Kuhn-Tucker)条件可求得

$$\hat{P}_s(n) = \left\{ \frac{W\eta[1 + \theta(n)]\text{lbe}}{q_1(n-1) + \eta\delta(n)} - \frac{N_0 + P_c^{\min} G_{cr}}{G_{sr}} \right\}^+ \quad (26)$$

其中,  $\{x\}^+ = \max\{0, x\}$ 。拉格朗日算子  $\delta(n)$  和  $\theta(n)$  的值可由梯度下降法迭代更新, 计算式为

$$\delta(n, \tau + 1) = \{ \delta(n, \tau) + \alpha [\hat{P}_s(n, \tau) - P_{\text{th}}^D] \}^+ \quad (27)$$

$$\theta(n, \tau + 1) = \{ \theta(n, \tau) - \alpha [W\text{lb}(1 + A_0 \hat{P}_s(n, \tau)) - U_{\text{min}}] \}^+ \quad (28)$$

其中,  $\tau$  表示拉格朗日算子的第  $\tau$  次迭代,  $\alpha$  为步长。中继的最优功率  $P_r^*$  的计算式为

$$P_r^* = \min \left\{ P_{\text{max}}^D, \frac{B_0}{A_0} P_s^* \right\} \quad (29)$$

在功率控制问题的求解过程中包含两层迭代。其中, 外层迭代是为了求解  $q_1$ , 迭代次数用  $n$  表示, 最大迭代次数为  $I_{\text{outer}}$ , 迭代误差为  $\Delta_1$ ; 内层迭代为了更新拉格朗日算子  $\delta$  和  $\theta$ , 迭代次数用  $\tau$  表示, 最大迭代次数为  $I_{\text{inner}}$ , 迭代误差为  $\Delta_2$ 。功率控制的具体步骤如算法 1 所示。

**算法 1** 基于 Dinkelbach 方法和拉格朗日乘子法的功率控制算法

**输入**  $G_{sr}$ 、 $G_{sb}$ 、 $G_{cr}$ 、 $G_{cb}$ 、 $G_{rd}$ 、 $G_{rb}$ 、 $G_{cd}$ 、 $N_0$ 、 $W$ 、 $P_{\text{max}}^D$ 、 $P_{\text{max}}^C$ 、 $R_{\text{min}}^D$  和  $R_{\text{min}}^C$

**输出**  $q_1^*$ 、 $P_s^*$  和  $P_r^*$

1) for  $\forall m \in M, \forall r \in R$

2) 初始化:  $n=1, \tau=0, q_1(0), \delta(n, \tau), \theta(n, \tau), I_{\text{outer}}, I_{\text{inner}}, \Delta_1$  和  $\Delta_2$ , 根据式(18)计算  $P_c^{\min}$

3) while  $n < I_{\text{outer}}$  do

4) 根据式(26)计算  $\hat{P}_s(n)$

5) if  $\left| \text{Wlb}(1 + A_0 \hat{P}_s(n)) - \right.$

$$q_1(n-1) \left[ \frac{1}{\eta} \left( 1 + \frac{A_0}{B_0} \right) \hat{P}_s(n) + 2P_{\text{cir}} \right] < \Delta_1$$

$$6) \quad q_1^* = \frac{\text{Wlb}(1 + A_0 \hat{P}_s(n))}{\frac{1}{\eta} \left( 1 + \frac{A_0}{B_0} \right) \hat{P}_s(n) + 2P_{\text{cir}}}$$

7)  $P_s^* = \hat{P}_s(n)$

8) break

9) else

10)  $n \rightarrow n+1$

11) while  $\tau < I_{\text{inner}}$  do

12) 设置  $\delta(n) = \delta(n-1), \theta(n) = \theta(n-1)$

和  $\hat{P}_s(n) = \hat{P}_s(n-1)$

13) if  $|\delta(n, \tau+1) - \delta(n, \tau)| < \Delta_2$

且  $|\theta(n, \tau+1) - \theta(n, \tau)| < \Delta_2$

14) 更新拉格朗日算子  $\delta(n, \tau), \theta(n, \tau)$

以及  $\hat{P}_s(n, \tau)$

15) 计算

$$q_1(n-1) = \frac{\text{Wlb}(1 + A_0 \hat{P}_s(n))}{\frac{1}{\eta} \left( 1 + \frac{A_0}{B_0} \right) \hat{P}_s(n) + 2P_{\text{cir}}}$$

16) break

17) end if

18)  $\tau \rightarrow \tau+1$

19) end while

20) end if

21) end while

22) end for

### 3.2 基于 Q 学习的中继选择算法

在为每个 D2D 用户对选择合适的中继之前, 为了减小计算复杂度, 本文为每个 D2D 用户对划分了备选的中继集合。对于 D2D 用户对  $m$ , 可选用中继的区域划分如图 2 所示, 备选中继所在区域处于以 D2D 发射端  $s$  和 D2D 接收端  $d$  的距离  $D_{sd}$  为半径, 分别以 D2D 发射端  $s$  和 D2D 接收端  $d$  为圆

心所画的 2 个圆的重合区域。D2D 用户对  $m$  的备选中继集合可表示为

$$G_m \triangleq \{r : D_{sr} \leq D_{sd}, D_{rd} \leq D_{sd}, \forall r \in R, \forall m \in M\} \quad (30)$$

其中,  $D_{sr}$  和  $D_{rd}$  分别表示 D2D 发射端到中继和中继到 D2D 接收端的距离。

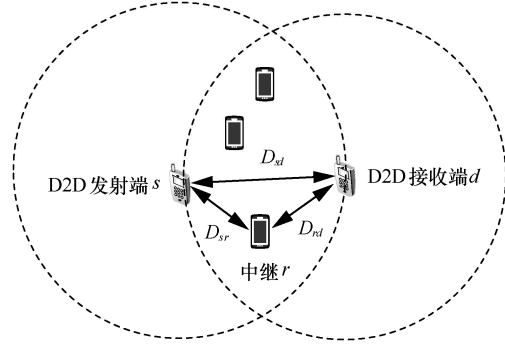


图 2 备选中继区域示意

在给每个 D2D 用户对划分备选中继区域之后, 为了保证 D2D 用户对的正常通信, 本文设计优先完成备选中继数较少的 D2D 用户对的中继选择。具体地, 用  $|G_m|$  表示 D2D 用户对  $m$  的备选中继数, 根据  $|G_m|$  的大小升序排列, 得到新的 D2D 用户对集合  $M^*$ 。然后, 利用本文提出的基于 Q 学习的中继选择算法完成每个 D2D 用户对的中继选择。

Q 学习是一种典型的强化学习算法, 可以在任意给定的马尔可夫决策过程中通过学习过程找到最优的动作<sup>[18]</sup>。在一次学习过程中, 学习的主体称为智能体 (agent), 智能体通过感知周围的环境, 判断当前所处的状态  $s_t$ , 其中  $t$  表示在一次迭代学习过程的时刻。之后根据一定的策略 (例如  $\epsilon$ -贪婪算法) 从动作集合  $A_t$  选择某个动作  $a_t$ 。在执行某个动作之后, 智能体会收到环境的反馈, 称为奖赏  $R_t$ ,  $R_t$  反映了所执行动作的好坏。在每次选择动作时, 智能体通常会大概率地选择当前已知的最优动作, 但是也会小概率地去尝试新的动作, 最终达到最大化总奖赏值的目的。Q 学习通过更新学习动作状态值函数  $Q(s_t, a_t)$  来获得最优策略, 其中  $Q(s_t, a_t)$  表示在给定状态  $s_t$  下执行动作  $a_t$  的收益。动作状态值函数  $Q(s_t, a_t)$  在时刻  $t$  的更新式为

$$Q_t(s_t, a_t) = Q_{t-1}(s_t, a_t) + \alpha [R_t + \gamma \max_{a \in A_t} Q(s_{t+1}, a) - Q_t(s_t, a_t)] \quad (31)$$

其中,  $\alpha \in (0, 1]$  表示学习因子,  $R_t$  表示奖赏值,  $\gamma$  表示折扣因子。从式(31)可以看出, 从状态  $s_t$  出发,

依照某种策略选择动作  $a_t$ ，随后更新 Q 值函数。当所有的动作状态值都被频繁地执行后，Q 估计值最终会收敛到最优 Q 值上，从而获得最优策略。

本文将中继选择的过程建模为 D2D 用户对学习寻找最优中继的过程，分别定义 Q 学习中的智能体，动作集合  $A_t$ 、状态集合  $S_t$  和奖赏函数  $R_t$  如式(32)~式(34)所示。

智能体：所有 D2D 用户对。

$$A_t = (a_1, a_2, \dots, a_l) \quad (32)$$

其中， $a_l$  表示选择中继  $l$  完成通信。在  $t$  时刻的动作  $a_t$  可以是集合  $A_t$  中任一元素。

$$S_t = \begin{cases} s_1, EE_{s,d} < EE^{th} \\ s_2, EE_{s,d} \geq EE^{th} \end{cases} \quad (33)$$

其中， $EE_{s,d}$  表示 D2D 发射端到 D2D 接收端的能效值，可通过功率控制进行计算； $EE^{th}$  表示 D2D 链路的能效阈值。

$$R_t = \begin{cases} \frac{1}{\lambda} EE_{s,d}, & EE_{s,d} \geq EE^{th} \\ -1, & \text{其他} \end{cases} \quad (34)$$

其中， $\lambda$  表示放大系数。奖赏函数反映了中继选择的好坏，若选择的中继满足 D2D 链路的最小能效时，返回一个正值；若不满足，返回一个负值。当每对 D2D 用户完成 Q 学习过程后，会生成对应于中继的 Q 表，根据 Q 表，便可以选择最优中继。具体的基于 Q 学习的中继选择过程如算法 2 所示。

**算法 2** 基于 Q 学习的中继选择算法

- 1) 对于  $\forall m \in M$ ，计算  $|G_m|$ ，根据  $|G_m|$  升序排列得到新的 D2D 用户对集合  $M^*$
- 2) 找到集合  $M^*$  中第一个元素  $m^*$ ，对于  $m^*$ ，设置  $Q(s_t, a_t) = 0$
- 3) 由式(33)得到  $m^*$  的初始状态  $s_t$
- 4) while Q 值没有收敛
- 5) 根据  $\epsilon$ -贪婪算法选择动作  $a_t$
- 6) 计算此时 D2D 用户对的能效  $EE_{s,d}$
- 7) 根据式(34)计算  $R_t$
- 8) 观测新的状态  $s_{t+1}$
- 9) 根据式(31)计算 Q 值并更新 Q 表
- 10) 更新智能体的状态为  $s_{t+1}$
- 11) end while
- 12) 根据 Q 表计算出  $m^*$  的最佳中继  $r^*$
- 13) 从  $M^*$  中删除  $m^*$  并更新  $|G_m|$  和  $M^*$
- 14) for  $m \in M^* \setminus \{m^*\}$

15) 重复步骤 2)~步骤 14)

16) end for

17) 输出中继选择矩阵 X

**3.3 基于匹配理论的信道分配算法**

在 3.1 节和 3.2 节中，通过功率控制和中继选择可以得到 D2D 用户对在复用某一蜂窝用户信道下的最大能效，本节讨论如何为每个 D2D 用户对分配合适的信道以最大化系统总 D2D 用户的能效。

由于蜂窝用户集合  $C$  和 D2D 用户集合  $M$  是 2 个互不相交的集合，因此本文构建二分图  $G = (V, E)$ ，其中二分图中顶点  $V$  表示蜂窝用户集合和 D2D 用户对集合，边  $E$  表示为 D2D 用户对和蜂窝用户之间的连线。在 3.1 节的功率控制中，通过约束式(16)已经满足蜂窝用户的 QoS，所以在信道分配时，只需要考虑如何最大化系统总 D2D 用户对的能效。因此，设定 D2D 用户与蜂窝用户之间边的权重表示的是该 D2D 用户复用该蜂窝用户信道时的能效值。引入  $N \times K$  的矩阵  $T$  表示权重矩阵，矩阵的行表示 D2D 用户对，矩阵的列表示蜂窝用户。

根据以上分析，信道分配问题可以转换为最大二部图匹配问题，并使用匈牙利算法进行求解。由匈牙利算法计算出一个包含信道分配信息的布尔矩阵  $Y$ ，根据矩阵  $Y$  便可为每个 D2D 用户对分配合适的信道。具体信道分配过程如算法 3 所示。

**算法 3** 基于匹配理论的信道分配算法

- 1) 对于  $\forall m \in M, \forall c \in C$ ，通过功率控制和中继选择得到  $N \times K$  的权值矩阵  $T$
- 2) 使用匈牙利算法得到信道分配矩阵  $Y$
- 3) for  $i = 1$  to  $N$
- 4) for  $j = 1$  to  $K$
- 5) if  $y_{ij} = 1$
- 6) 将蜂窝用户  $j$  的信道分配给 D2D 用户  $i$
- 7) end if
- 8) end for
- 9) end for

**3.4 复杂度分析**

本文所提出的能效优化算法由 3 个部分组成，分别是基于 Dinkelbach 方法和拉格朗日乘子法的功率控制算法、基于 Q 学习的中继选择算法和基于匹配理论的信道分配算法。在假设 D2D 用户获得确定辅助的中继以及复用的信道情况下，得出单一 D2D 用户能效最大时的功率分配解。之后，基于功

率分配的解，在假设复用信道确定的条件下，利用 Q 学习计算 D2D 用户在复用任意一条信道时的最佳中继。基于功率控制和中继选择，计算 D2D 用户复用每条可用信道时的能效矩阵，利用匈牙利算法得到信道分配解。虽然本文提出的能效算法复杂度较高，但是能在理论上获得系统总 D2D 用户的最大能效值。

本文所提能效优化算法的复杂度取决于功率控制、中继选择和信道分配算法。基于 Dinkelbach 方法和拉格朗日乘子法的功率控制算法复杂度为  $O(I_{outer}I_{inner})$ ，考虑  $N$  个 D2D 用户对， $K$  个蜂窝用户以及每个 D2D 用户对可复用的最大中继数  $L$ ，则功率控制算法的复杂度为  $O(MNRI_{outer}I_{inner})$ 。基于 Q 学习的中继选择算法的复杂度取决于 Q 学习中状态集合数  $|S_i|$  和动作集合数  $|A_i|$ ，所以中继选择算法的复杂度为  $O(|S_i||A_i|)$ 。基于匹配理论的信道分配算法复杂度为  $O(K^3)$ ，其中  $K$  为蜂窝用户数。综上所述，本文提出的联合功率控制、中继选择和信道分配的能效优化算法的复杂度为  $O(MNRI_{outer}I_{inner} + |S_i||A_i| + K^3)$ 。

#### 4 仿真结果及分析

本文使用 Matlab 仿真工具对算法进行仿真，重复执行 1 000 次蒙特卡洛实验，然后对结果取平均值。每一次算法执行过程中，中继、蜂窝用户和 D2D 用户均随机分布在系统中。假设系统内 D2D 与蜂窝链路的阴影衰落均服从均值为 0、标准差分别为 12 dB 与 10 dB 的对数正态分布，其余主要仿真参数如表 1 所示。

表 1 实验仿真参数

参数	取值
小区半径/m	500
高斯白噪声 $N_0$ /dBm	-114
系统带宽 $W$ /kHz	180
电路功率 $P_{cir}$ /mW	50
蜂窝链路路径损耗模型/dB	$128.1+37.6\ln d$
D2D 链路路径损耗模型/dB	$140+40\ln d$
蜂窝链路最小传输速率 $R_{min}^C$ /(Mbit·s <sup>-1</sup> )	0.9
D2D 链路最小传输速率 $R_{min}^D$ /(Mbit·s <sup>-1</sup> )	0.9
D2D 用户间最大传输距离 $D_{max}$ /m	50~150
蜂窝用户数 $K$ /个	10~20
D2D 用户数 $N$ /对	5~15
中继数 $L$ /个	50~150

仿真中，将本文所提算法与文献[13]的 BEEPER 算法、随机中继选择算法和随机信道分配算法进行对比。其中，文献[13]的 BEEPER 算法分别优化 D2D 中继通信的第一跳和第二跳链路的能效，然后再完成中继选择，但是并没有考虑信道的分配，所以，为了全面评价本文所提算法，与 BEEPER 算法进行对比时，在 BEEPER 算法中加入了本文提出的信道分配算法。随机中继选择算法是本文所提功率控制和信道分配算法的结合，而每个 D2D 用户的辅助中继是从备选中继区域中随机选择的。随机信道分配算法是本文所提功率控制和中继选择算法的结合，而信道的分配是随机的。

图 3 是在  $N=K=10$ 、 $L=50$  和  $D_{max}=100$  的条件下，系统 D2D 用户总能效的累计分布函数曲线。由图 3 可知，本文所提算法最大可达能效大于其余 3 种算法，且比 BEEPER 算法约高 2%，比随机中继选择算法高 13%，比随机信道分配算法高 41%。本文所提算法优于 BEEPER 算法的原因是本文算法直接对 D2D 发射端到接收端的能效优化问题进行求解，而 BEEPER 算法通过分别最大化 D2D 通信第一跳和第二跳链路的能效来解决问题。本文所提算法优于随机中继选择算法和随机信道分配的原因是前者考虑从功率、中继和信道 3 个方面优化系统中 D2D 用户的总能效，而后两者仅是从功率和中继或者功率和信道 2 个方面去优化能效。

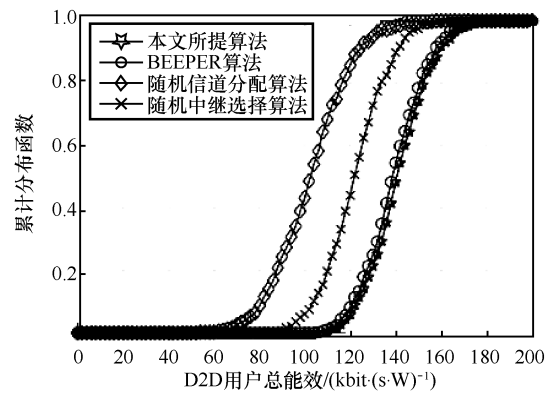


图 3 D2D 用户总能效的累计分布函数曲线

当  $L=50$  和  $D_{max}=100$  时，系统总 D2D 用户对能效随 D2D 用户对数目变化的曲线如图 4 所示。随着 D2D 用户数目的增加，4 种算法的 D2D 链路总能效都呈现了递增的趋势。在 D2D 用户数相同的情况下，本文所提算法的性能明显优于其他 3 种优化算法。这是因为本文所提算法联合考虑从功率控制、中继

选择和信道分配 3 个方面优化系统中 D2D 用户的总能效，并且在解决功率控制问题时，在假设  $SINR_D^{(1)}=SINR_D^{(2)}$  的条件下，直接求解出 D2D 发射端到接收端的最大能效。当  $N=K=15$  时，本文所提算法的能效值比 BEEPER 算法约高 2%，比随机中继选择算法高 13%，比随机信道分配算法高 41%。

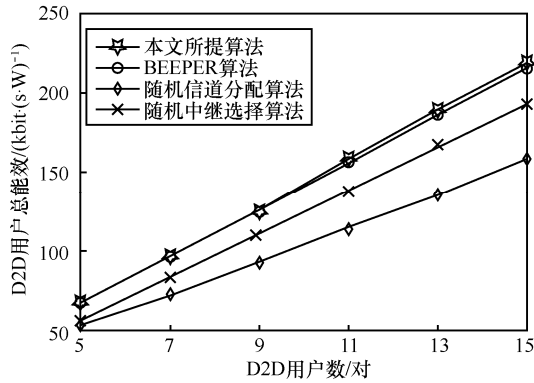


图 4 D2D 用户总能效随 D2D 用户数变化曲线

图 5 是当  $N=K=10$  和  $L=50$  时，系统总 D2D 用户对能效随 D2D 发射端与接收端最大传输距离  $D_{max}$  变化的曲线。从图 5 中可以看出，随着  $D_{max}$  的增大，总能效呈递减的趋势。这是因为当  $D_{max}$  增大时，D2D 发射端和中继需要增加发射功率以保证传输质量，从而降低了能效。但是从整体上来说，本文所提算法优于其余 3 种对比算法。当  $D_{max}=50$  时，本文算法的能效值比 BEEPER 算法约高 1%，比随机中继选择算法高 13%，比随机信道分配算法高 41%。

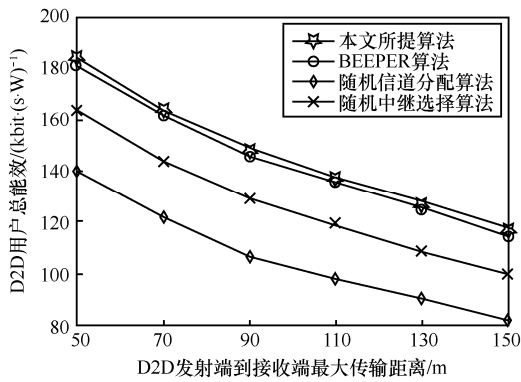


图 5 D2D 用户总能效随  $D_{max}$  变化曲线

图 6 给出了当  $K=N=10$  和  $D_{max}=100$  时，D2D 用户总能效随着中继数  $L$  变化的曲线。随着  $L$  的增加，本文所提算法、BEEPER 算法和随机信道分配算法的曲线呈现递增的趋势，而随机中继选择算法的曲线平缓，没有太大的波动。这是因为随着中继

数目增加，加入了中继选择过程的本文所提算法、BEEPER 算法和随机信道分配算法会选择更合适的中继进行通信，而随机中继选择算法在选择中继时是随机的，所以关于系统 D2D 用户总能效的曲线没有太大波动。当  $L=150$  时，本文所提算法的能效值高于 BEEPER 算法 2%，高于随机中继选择算法 26%，高于随机信道分配算法 34%。

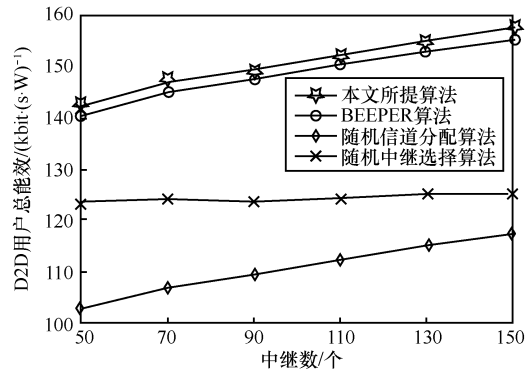


图 6 D2D 用户总能效随中继数变化曲线

图 3~图 6 的仿真中 D2D 用户数和蜂窝用户数是相等的，也就是说，D2D 用户数与可复用信道数是一致的，而图 7 给出了 D2D 用户数与信道数不一致的情况下，每种算法的能效曲线。在图 7 中，仿真参数分别为  $K=10$ 、 $L=50$  和  $D_{max}=100$ 。随着  $N$  增加，本文所提算法、BEEPER 算法和随机中继选择算法的曲线呈现递增并趋于平缓的趋势，而随机信道分配算法曲线平缓，没有太大的波动。因为信道资源的增加，加入了信道分配过程的 3 种算法会被分配到更合适的信道进行通信，而随机信道分配算法在信道分配时是随机的，所以能效曲线没有太大波动。当  $N=20$  时，本文所提算法的能效值高于 BEEPER 算法 2%，高于随机中继选择算法 10%，高于随机信道分配算法 49%。

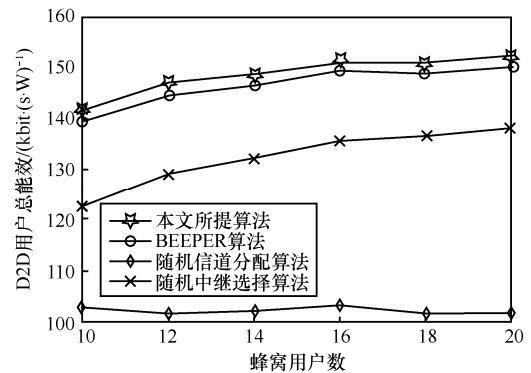


图 7 D2D 用户总能效随蜂窝用户数变化曲线

## 5 结束语

为了提高 D2D 中继通信系统中 D2D 用户的能效, 本文提出了一种联合功率控制、中继选择和信道分配的能效优化算法。首先, 在保证 D2D 用户和蜂窝用户的 QoS 条件下, 建模能效优化问题。其次, 使用 Dinkelbach 方法与拉格朗日乘子法, 求解能效的最优值。再次, 根据功率控制得到的能效值, 通过 Q 学习算法选择合适的中继。最后, 基于 D2D 用户集合与蜂窝用户集合构建二分图, 设置集合间边的权重值为能效值, 利用匈牙利算法完成信道分配。仿真结果表明, 本文所提算法有效地提高了系统 D2D 用户的总能效。

### 参考文献:

- [1] 钱志鸿, 王雪. 面向 5G 通信网的 D2D 技术综述[J]. 通信学报, 2016, 37(7): 1-14.  
QIAN Z H, WANG X. Reviews of D2D technology for 5G communication networks[J]. Journal on Communications, 2016, 37(7): 1-14.
- [2] LIU J, KATO N, MA J, et al. Device-to-device communication in LTE-advanced networks: a survey[J]. IEEE Communications Surveys & Tutorials, 2015, 17(4): 1923-1940.
- [3] HASAN M, HOSSAIN E. Distributed resource allocation for relay-aided device-to-device communication: a message passing approach[J]. IEEE Transactions on Wireless Communications, 2014, 13(11):6326-6341.
- [4] TIAN C, QIAN Z, WANG X, et al. Analysis of joint relay selection and resource allocation scheme for relay-aided d2d communication networks[J]. IEEE Access, 2019(7): 142715-142725.
- [5] WANG F, XU C, SONG L, et al. Energy-efficient resource allocation for device-to-device underlay communication[J]. IEEE Transactions on Wireless Communications, 2015, 14(4): 2082-2092.
- [6] HU J, HENG W, LI X, et al. Energy-efficient resource reuse scheme for D2D communications underlying cellular networks[J]. IEEE Communications Letters, 2017, 21(9): 2097-2100.
- [7] ZHOU Z, OTA K, DONG M, et al. Energy-efficient matching for resource allocation in D2D enabled cellular networks[J]. IEEE Transactions on Vehicular Technology, 2017, 66(6): 5256-5268.
- [8] LIU S, WU Y, LI L, et al. A two-stage energy-efficient approach for joint power control and channel allocation in D2D communication[J]. IEEE Access, 2019(7):16940-16951.
- [9] JIANG Y, LIU Q, ZHENG F, et al. Energy efficient joint resource allocation and power control for D2D communications[J]. IEEE Transactions on Vehicular Technology, 2016, 65(8): 6119-6127.
- [10] KUANG Z, LIU G, LI G, et al. Energy efficient resource allocation algorithm in energy harvesting-based D2D heterogeneous networks[J]. IEEE Internet of Things Journal, 2019, 6(1): 557-567.
- [11] WANG R, LIU J, ZHANG G, et al. Energy efficient power allocation for relay-aided D2D communications in 5G networks[J]. China Com-

munications, 2017, 6(6): 54-64.

- [12] ALQERM I, SHIHADA B. Energy efficient power allocation in multi-tier 5G networks using enhanced online learning[J]. IEEE Transactions on Vehicular Technology, 2017, 66(12): 11086-11097.
- [13] CHANG W, TENG J. Energy efficient relay matching with bottleneck effect elimination power adjusting for full-duplex relay assisted D2D networks using mm-wave technology[J]. IEEE Access, 2018, 6(6): 3300-3309.
- [14] LIU S, WU Y, LI L, et al. Optimal resource and power allocation with relay selection for RF/RE energy harvesting relay-aided D2D communication [J]. IEEE Access, 2019(7): 89670-89686.
- [15] KLAIQI B, CHU X, ZHANG J. Energy and spectral-efficient adaptive forwarding strategy for multi-hop device-to-device communications overlaying cellular networks[J]. IEEE Transactions on Wireless Communications, 2018, 17(9): 5684-5699.
- [16] ZHANG R, QI C, LI Y, et al. Towards energy-efficient underlaid device-to-device communications: a joint resource management approach[J]. IEEE Access, 2019(7): 31385-31396.
- [17] DINKELBACH W. On nonlinear fractional programming[J]. Management Science, 1967, 13(7):492-498.
- [18] WATKINS C, DAYAN P. Q-learning[J]. Machine Learning, 1992, 8(3-4): 279-292.

### [作者简介]



王雪(1984-), 女, 吉林白山人, 博士, 吉林大学副教授, 主要研究方向为物联网与 D2D 通信技术。



金涛(1994-), 男, 贵州贵阳人, 吉林大学硕士生, 主要研究方向为 D2D 通信技术。

钱志鸿(1957-), 男, 吉林长春人, 博士, 吉林大学教授、博士生导师, 主要研究方向为物联网、D2D、Wi-Fi、RFID 等无线网络与通信技术。

胡良帅(1995-), 男, 河南南阳人, 吉林大学硕士生, 主要研究方向为 D2D 通信技术。

王鑫(1982-), 女, 吉林长春人, 吉林大学博士生, 吉林农业大学副教授, 主要研究方向为 D2D 通信技术与无线传感器网络。