

DOI:10.19651/j.cnki.emt.2517809

非完美 SIC D2D-NOMA 系统吞吐量最大化功率分配*

张东方¹ 叶成荫²

(1. 辽宁石油化工大学人工智能与软件学院 抚顺 113001; 2. 辽宁石油化工大学信息与控制工程学院 抚顺 113001)

摘要: 针对 D2D-NOMA 系统在非完美串行干扰消除条件下的吞吐量最大化问题,提出了一种基于非完美 SIC 的 D2D-NOMA 系统吞吐量最大化算法。首先在蜂窝用户和 D2D 用户的服务质量、最大发射功率下,考虑非完美 SIC 条件,建立一个以 D2D-NOMA 系统的吞吐量最大化为目标的优化模型。然后将模型转换为以吞吐量为奖励的马尔可夫决策过程。利用 TD3 算法设计了一个功率分配算法,实现了系统吞吐量最大化。经仿真验证,所提出的算法与 DDPG 算法、遗传算法和随机功率分配算法相比,非完美 SIC 下 D2D-NOMA 系统的吞吐量分别提高了约 14%、25% 和 40%,也展现出更优的收敛性和公平性。

关键词: 非完美 SIC; D2D-NOMA 系统; 吞吐量最大化; TD3 算法

中图分类号: TN929.5 **文献标识码:** A **国家标准学科分类代码:** 510

Throughput-maximizing power allocation for D2D-NOMA system under imperfect SIC

Zhang Dongfang¹ Ye Chengyin²

(1. School of Artificial Intelligence and Software, Liaoning Petrochemical University, Fushun 113001, China;

2. School of Information and Control Engineering, Liaoning Petrochemical University, Fushun 113001, China)

Abstract: Aiming at the throughput maximization problem of D2D-NOMA system under the condition of imperfect successive interference cancellation, a D2D-NOMA system throughput maximization algorithm based on imperfect SIC is proposed. Firstly, under the constraints of service quality and maximum transmission power of cellular users and D2D users, an optimization model with the goal of maximizing the throughput of D2D-NOMA system is established by considering the imperfect SIC condition. Then the model is converted into a Markov decision process with throughput as reward. A power allocation algorithm is designed using the TD3 algorithm to maximize the system throughput. The simulation results show that compared with the DDPG algorithm, the genetic algorithm and the random power allocation algorithm, the proposed algorithm improves the throughput of D2D-NOMA system under imperfect SIC by 14%, 25% and 40%, respectively, and also shows better convergence and fairness.

Keywords: imperfect SIC; D2D-NOMA system; throughput maximization; TD3 algorithm

0 引言

随着数据传输需求的不断增加和用户对网络性能要求的提升,现有网络越来越难以满足对更高吞吐量需求^[1-3]。越来越多研究将设备到设备(device-to-device, D2D)技术与非正交多址接入(non-orthogonal multiple access, NOMA)技术结合起来,以提高通信系统吞吐量^[4]。但也增加了用户间干扰^[5]。因此,如何设计有效的功率分配算法对提升 D2D-NOMA 系统吞吐量至关重要。

目前,国内外学者基于不同方法对 D2D-NOMA 系统

吞吐量最大化的功率分配问题开展了大量的研究工作,取得了显著成果。Alemaishat 等^[6]利用 Karush-Kuhn-Tucker 条件设计了最优的功率分配方案,显著提升了能量效率和系统吞吐量。Li 等^[7]提出一种动态功率分配方法。该方法优先考虑弱用户和远用户,选择系数使得远用户的目标速率得到满足,随后将剩余功率分配给近用户,使得产生了相对更好的吞吐量;宋勇春等^[8]要优化 D2D 组内用户功率分配系数,提出了基于爬山算法的自适应遗传算法。实现了系统吞吐量最大化。然而,这些研究都是基于完美的串行干扰消除(successive interference cancellation, SIC)

收稿日期:2025-01-03

* 基金项目:国家自然科学基金(62203202)项目资助

假设,即接收端能够完全消除干扰。受限于硬件性能、信号处理不当及各种环境噪声等因素,SIC 常常无法完美工作^[9-11],导致残余干扰累积,从而对系统性能造成显著负面影响。因此,针对 D2D-NOMA 中的非完美 SIC 的问题,Wu 等^[12]提出考虑基于 NOMA 的 D2D 通信系统的能量效率和用户公平性的功率分配方案。Yu 等^[13]针对 D2D 接收机译码的非完美 SIC 问题,提出了在单载波底层 D2D 通信场景下,考虑蜂窝用户的服务质量(quality of service, QoS)和 D2D 发射功率的约束,将功率优化问题转化为对偶问题,然后用次梯度法更新对偶变量,显著提升了 D2D 通信的数据速率。Shi 等^[14]同时考虑与传统的按照接收信号强度从强到弱的顺序进行解码顺序和用户全接入策略不同,在非完美 SIC 或用户的 QoS 要求不同的情况下,研究 NOMA 系统中的动态接入和动态解码顺序,提出了一种基于 DRL 的智能分层用户接入和解码顺序选择算法,其明显优于传统的接入和解码方案。因此,在实际环境中分析非完美 SIC 的 D2D-NOMA 系统的吞吐量最大化功率分配问题具有重要意义。

本文针对 D2D-NOMA 系统中 D2D 发射机通过 NOMA 协议与两个 D2D 接收机进行通信时 SIC 非完美的实际情况,提出了基于非完美 SIC 的 D2D-NOMA 系统吞吐量最大化功率分配算法。算法通过以 D2D 用户吞吐量最大化为目标,约束最大发射功率、系统所有用户的服务质量要求,构建一个优化问题模型,对功率分配进行优化。仿真验证了所提算法优于遗传算法、深度确定性策略梯度(deep deterministic policy gradient, DDPG)算法和基于双延迟深度确定性策略梯度算法(twin delayed deep deterministic policy gradient, TD3)的随机功率分配方法。

1 系统模型和问题表述

1.1 系统模型

考虑如图 1 所示的单小区上行蜂窝网络。基站(base station, BS)位于小区的**中心。小区包含 M 个蜂窝用户(cellular users, CU), $\mathbf{C} = \{C_1, C_2, \dots, C_m, \dots, C_M\}$; N 个 D2D 用户(D2D users, DU)组, $\mathbf{D} = \{D_1, D_2, \dots, D_n, \dots, D_N\}$ 。其中,每组包含一个 D2D 发射机(D2D transmitter, DT)和两个 D2D 接收机(D2D receiver, DR),其中对 $\forall n \in N$,有 $\mathbf{D}_n = \{DR_{n1}, DR_{n2}\}$ 。CU 和 DT 随机均匀分布在以基站为中心的小区内,DR 随机均匀分布在以 DT 为中心的小区内。基站通过传统的正交多址接入技术为 CU 提供服务,DT 通过 NOMA 技术为 D2D 接收机提供服务,DR 间通过建立直接链接轮流发送和接收信号避免自干扰。每个蜂窝用户的发射功率均为 P_C 。在系统通信模型中,给每一个蜂窝用户分配一个信道 $SC_m \in \mathbf{SC}$, $\mathbf{SC} = \{SC_1, SC_2, \dots, SC_M\}$,各信道相互正交。为了方便研究,假设 BS 知道系统信道状态信息和通信用户的 QoS 需求。信道默认遵循准静态瑞利衰落信道模型^[15],公式为:

$$|h|^2 = |\hat{h}|^2 d^{-\mu} \quad (1)$$

式中: \hat{h} 服从小尺度衰落(0,1)分布, d 为通信距离, μ 为路径损耗指数。

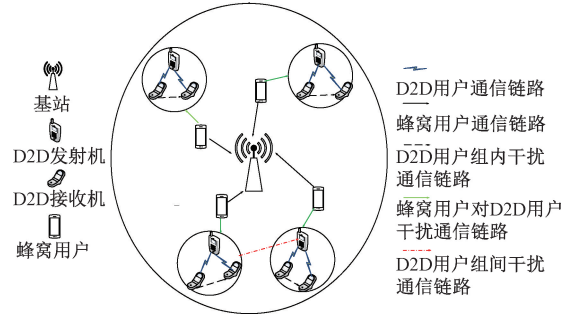


图 1 系统模型

Fig. 1 System model

1.2 非完美 SIC 下 D2D-NOMA 通信系统模型分析

1) 系统干扰分析

由于蜂窝上行链路可用信道资源有限^[16],为了提高系统的频谱利用率,采用多个 D2D 用户对复用**一个蜂窝用户信道的多对一模式,但是也增加了系统用户间干扰。第 n 个 D2D 组的接收机接收的干扰如下:

(1)来自叠加信号对同一 D2D 组内其他接收机非完美 SIC 解码的组内干扰,表示为:

$$IF_{DR_{nj}}^{co} = P_{D_n} \alpha_{nj} |h_{DR_{ni}-DR_{nj}}|^2, i, j \in \{1, 2\} \text{ 且 } i \neq j \quad (2)$$

式中: P_{D_n} 为第 n 个 D2D 组内 D2D 发射机 DT_n 的发射功率, α_{nj} 为第 n 个 D2D 组接收机 DR_{nj} 的功率分配系数, $h_{DR_{ni}-DR_{nj}}$ 为同一 D2D 组内,一个接收机到另一个接收机之间的信道增益。

(2)来自复用相同信道的其他 D2D 组的 DT 的组间干扰,表示为:

$$IF_{DR_{nj}}^{cr} = \sum_{n' \neq n} \eta_{n',n} P_{D_{n'}} |h_{DT_{n'}-DR_{nj}}|^2, j \in \{1, 2\} \quad (3)$$

式中: $\eta_{n',n}$ 为信道复用系数,不同 DT 复用相同信道时取 1,否则取 0。 $|h_{DT_{n'}-DR_{nj}}|^2$ 为 $DT_{n'}$ 和 DR_{nj} 之间的信道增益。

(3)来自复用相同信道 SC_m 的 CU 的蜂窝干扰,表示为:

$$IF_{C_m-DR_{nj}} = P_C |h_{C_m-DR_{nj}}|^2, j \in \{1, 2\} \quad (4)$$

式中: $h_{C_m-DR_{nj}}$ 为 C_m 和 DR_{nj} 之间的信道增益。

在 BS 收到与第 m 个蜂窝用户 C_m 同频段 DT 的同频干扰为:

$$IF_{DT_n-C_m} = \sum_{n=1}^N \eta_{n,m} P_{D_n} |h_{DT_n-BS}|^2 \quad (5)$$

式中: $\eta_{n,m}$ 当 C_m 与 DT_n 共用一个信道资源时取 1,否则取 0。 h_{DT_n-BS} 为 DT_n 和 BS 之间的信道增益。

2) 蜂窝用户通信链路信号传输

在上行传输时,CU 以发射功率 P_C 向 BS 传输数据,则

P_C 需要满足: $0 < P_C < P_C^{\max}$, 其中, P_C^{\max} 为 CU 发送数据的最大功率限制。BS 从 SC_m 接收的信号为:

$$y_{C_m-BS} = h_{C_m-BS} \sqrt{P_C} x_{C_m} + \sum_n h_{DT_n-BS} \sqrt{P_{D_n}} \eta_{n,m} x_{DT_n} + \zeta_m \quad (6)$$

式中: h_{C_m-BS} 为 C_m 和 BS 之间的信道增益, x_{C_m} 、 x_{DT_n} 分别为 C_m 发送给 BS 的信号、 DT_n 发送的信号, ζ_m 表示加性高斯白噪声, 它的均值为 0, 方差为 1。

则 C_m 的吞吐量为:

$$R_{C_m} = \log_2(1 + SINR_{C_m}) \quad (7)$$

其信噪比(signal-to-noise ratio, SINR)为:

$$SINR_{C_m} = \frac{P_C |h_{C_m-BS}|^2}{IF_{DT_n-C_m} + N_0} \quad (8)$$

3) D2D 用户通信链路信号传输

由于 DT 利用 NOMA 协议跟多个 DR 同时进行通信, 根据功率域 NOMA, 对于功率信号的复用, 在 DT 侧采用叠加编码技术, 在 DR 侧采用 SIC 技术。由 DT 向接收方发送带有功率分配系数的 2 个 DU 的叠加信号表示为:

$$x_{DT_n} = \sqrt{P_{D_{n1}}} x_{DT_{n1}} + \sqrt{P_{D_{n2}}} x_{DT_{n2}}, P_{D_{n1}} = \alpha_{n1} P_{D_n}, P_{D_{n2}} = \alpha_{n2} P_{D_n}, \alpha_{n1} + \alpha_{n2} = 1, \text{ 则 } DR_{n1} \text{ 接收的信号为:}$$

$$y_{DR_{n1}} = h_{C_m-DR_{n1}} \sqrt{P_C} x_{C_m} + h_{DT_n-DR_{n1}} x_{D_n} + \sum_{n' \neq n} h_{DT_{n'}-DR_{n1}} \sqrt{P_{D_{n'}}} \eta_{n',m} x_{D_{n'}} + s_{n1} \quad (9)$$

式中: $h_{C_m-DR_{n1}}$ 、 $h_{DT_n-DR_{n1}}$ 和 $h_{DT_{n'}-DR_{n1}}$ 分别为 C_m 和 DR_{n1} 、 DT_n 和 DR_{n1} 以及 $DT_{n'}$ 、 $DT_{n'}$ 和 DR_{n1} 之间的信道增益, $\eta_{n',m}$ 表示如果 DT_n 和 $DT_{n'}$ 共用相同的信道, 则为 1, 否则为 0。

同理, DR_{n2} 接收的信号为:

$$y_{DR_{n2}} = h_{C_m-DR_{n2}} \sqrt{P_C} x_{C_m} + h_{DT_n-DR_{n2}} x_{D_n} + \sum_{n' \neq n} h_{DT_{n'}-DR_{n2}} \sqrt{P_{D_{n'}}} \eta_{n',m} x_{D_{n'}} + s_{n2} \quad (10)$$

4) 非完美 SIC 解码

由于功率域 NOMA 以不同的功率电平为不同的用户提供服务, 所以不失一般性的假设 $h_{DT_n-DR_{n1}} > h_{DT_n-DR_{n2}}$, 则 DR_{n1} 对于 DR_{n2} 的信号的接收 SINR 大于 DR_{n2} 对其自身信号的接收 SINR, 则干扰消除成功。表示为:

$$\frac{|h_{DT_n-DR_{n1}}|^2 \alpha_{n2} P_{D_n}}{IF_{DR_{n1}}^{co} + IF_{DR_{n1}}^{cr} + IF_{C_m-DR_{n1}} + N_0} \geq \frac{|h_{DT_n-DR_{n2}}|^2 \alpha_{n2} P_{D_n}}{IF_{DR_{n2}}^{co} + IF_{DR_{n2}}^{cr} + IF_{C_m-DR_{n2}} + N_0} \quad (11)$$

式中: $|h_{DT_n-DR_{nj}}|^2 = |\hat{h}_{DT_n-DR_{nj}}|^2 d_{DT_n-DR_{nj}}^{-\mu}$, $j \in \{1, 2\}$, $|h_{DT_{n'}-DR_{nj}}|^2 = |\hat{h}_{DT_{n'}-DR_{nj}}|^2 d_{DT_{n'}-DR_{nj}}^{-\mu}$, $j \in \{1, 2\}$, $|h_{C_m-DR_{nj}}|^2 = |\hat{h}_{C_m-DR_{nj}}|^2 d_{C_m-DR_{nj}}^{-\mu}$, $j \in \{1, 2\}$ 。 $d_{DT_n-DR_{n1}}^{-\mu}$ 、 $d_{DT_n-DR_{n2}}^{-\mu}$ 、 $d_{DT_{n'}-DR_{n1}}^{-\mu}$ 、 $d_{DT_{n'}-DR_{n2}}^{-\mu}$ 、 $d_{C_m-DR_{n1}}^{-\mu}$ 、 $d_{C_m-DR_{n2}}^{-\mu}$ 分别表示 DT_n 和 DR_{n1} 、 DT_n 和 DR_{n2} 、 $DT_{n'}$ 和 DR_{n1} 、 $DT_{n'}$ 和 DR_{n2} 、 C_m 和 DR_{n1} 、 C_m 和 DR_{n2} 的距离。 $IF_{DR_{n1}}^{co}$ 、 $IF_{DR_{n2}}^{co}$ 分别表示来自

DR_{n1} 和 DR_{n2} 处的组内干扰, $IF_{DR_{n1}}^{cr}$ 、 $IF_{DR_{n2}}^{cr}$ 分别表示在 DR_{n1} 和 DR_{n2} 处接收到的组间干扰, $IF_{C_m-DR_{n1}}$ 、 $IF_{C_m-DR_{n2}}$ 分别表示在 DR_{n1} 和 DR_{n2} 处接收到蜂窝干扰。并由此化简上面的不等式为:

$$\Gamma(\eta) = |h_{DT_n-DR_{n1}}|^2 (IF_{DR_{n1}}^{cr} + IF_{C_m-DR_{n1}} + 1) - |h_{DT_n-DR_{n2}}|^2 (IF_{DR_{n2}}^{cr} + IF_{C_m-DR_{n2}} + 1) \geq 0 \quad (12)$$

因此, DR_{n1} 解码时经由 SIC 技术移除了 DR_{n2} 的信号干扰, DR_{n2} 解码时则将 DR_{n1} 的信号当作噪声来处理。但由于解码能力有限, DR_{n1} 使用 SIC 技术并不能达到完美, 故 DR_{n2} 的吞吐量表示为:

$$R_{DR_{n2}} = \log_2(1 + SINR_{DR_{n2}}) \quad (13)$$

其中, $SINR_{DR_{n2}}$ 为:

$$SINR_{DR_{n2}} = \frac{|h_{DT_n-DR_{n2}}|^2 \alpha_{n2} P_{D_n}}{IF_{DR_{n2}}^{co} + IF_{DR_{n2}}^{cr} + IF_{C_m-DR_{n2}} + N_0} \quad (14)$$

同理, DR_{n1} 的吞吐量为:

$$R_{DR_{n1}} = \log_2(1 + SINR_{DR_{n1}}) \quad (15)$$

其中, $SINR_{DR_{n1}}$ 为:

$$SINR_{DR_{n1}} = \frac{|h_{DT_n-DR_{n1}}|^2 \alpha_{n1} P_{D_n}}{IF_{DT_n-DR_{n1}}^{cr} \beta + IF_{C_m-DR_{n1}} + N_0} = \frac{|h_{DT_n-DR_{n1}}|^2 \alpha_{n1} P_{D_n}}{IF_{DT_n-DR_{n1}}^{cr} \beta + IF_{C_m-DR_{n1}} + N_0 + \sum_{n', n' \neq n} \eta_{n',n} P_{D_n} |h_{DT_{n'}-DR_{n1}}|^2 \beta + P_C |h_{C_m-DR_{n1}}|^2 + N_0} \quad (16)$$

式中: β 表示非完美 SIC 参数^[17], $0 < \beta < 1$, 当 $\beta = 0$ 时, 表示完美 SIC, $\beta = 1$ 表示无 SIC。

综上所述, D2D 用户总吞吐量为:

$$R_{sum} = R_{DR_{n1}} + R_{DR_{n2}} = \log(1 + SINR_{DR_{n1}}) + \log(1 + SINR_{DR_{n2}}) \quad (17)$$

1.3 问题表述

本文的主要目的是在满足 CU 和 DU 的通信服务质量要求以及 NOMA 协议, 考虑非完美 SIC 引起的干扰, 寻求最优功率分配策略, 实现系统 D2D 用户总吞吐量最大化。从而, 优化问题可以表述为:

$$\begin{aligned} \max_{\{\alpha_{n1}, P_{D_n}\}} & \sum_{n=1}^N R_{sum} \\ C1: & 0 < \alpha_{n1} < 0.5, \forall D_n \in D \\ C2: & 0 < P_{D_n} < P_{D_n}^{\max}, \forall D_n \in D \\ C3: & SINR_{C_m} \geq \gamma_m^{\min}, \forall C_m \in C \\ C4: & SINR_{DR_{n1}} \geq \gamma_n^{\min}, \forall DR_{n1} \in D_n \\ C5: & SINR_{DR_{n2}} \geq \gamma_n^{\min}, \forall DR_{n2} \in D_n \end{aligned} \quad (18)$$

式中: C1 表示一组 D2D 组中接收机的功率分配系数, 约束为二者之和为 1, 且二者均大于 0。C2 表示 D2D 发射功率取值范围。C3 约束 CU 的服务质量要求。C4 和 C5 保证在 NOMA 机制下 DU 的服务质量需求, γ_m^{\min} 和 γ_n^{\min} 为 CU 和 D2D 组的两个接收机 DR_{n1} 和 DR_{n2} 达到服务质量需要的最低 SINR 值。

2 TD3 算法吞吐量最大化功率分配设计

2.1 马尔可夫决策

通过在时间上对网络环境进行建模,将功率分配问题规划为一个马尔可夫决策过程^[18],智能体大多数时间利用历史的信道状态信息作为训练信息,瞬时信道状态信息只在可用时作为补充来使用。所构建的网络模型作为环境来与智能体进行交互反馈。下面给出强化学习的要素:智能体定义为每个 D2D-NOMA 用户组中的发射机,它只具有感知能力的与其自身环境和状态相关的信息。状态空间包含网络中所有可能的状态,定义每个 D2D-NOMA 用户组及其先前的信道增益信息作为一个独立的状态:

$$\mathbf{S}_n^t = \begin{bmatrix} \alpha_{n1}^{t-1}, P_{D_n}^{t-1}, h_{C_m-BS}, h_{DT_n-BS}, h_{C_m-DR_{n1}}, h_{C_m-DR_{n2}}, \\ h_{DT_n-DR_{n1}}, h_{DT_n-DR_{n2}}, h_{DT_n-DR_{n1}}, h_{DT_n-DR_{n2}} \end{bmatrix}. \quad (19)$$

每个状态是经过归一化处理,以便输入到神经网络中。对于每个 D2D 发射机,在 t 时刻,D2D 发射机执行动作 \mathbf{a}_t ,根据策略 $\pi(\mathbf{s})$ 选择功率分配因子和 D2D 发射功率,每个智能体选择动作时必须满足蜂窝用户的通信服务质量。奖励函数 r 是基于系统 D2D 用户吞吐量的计算,目标是最大化总 D2D 用户吞吐量,同时满足蜂窝用户的最小 SINR 阈值。

2.2 非完美 SIC D2D-NOMA 系统基于 TD3 算法的功率分配

非完美 SIC D2D-NOMA 系统基于 TD3 算法的功率分配结构如图 2 所示。

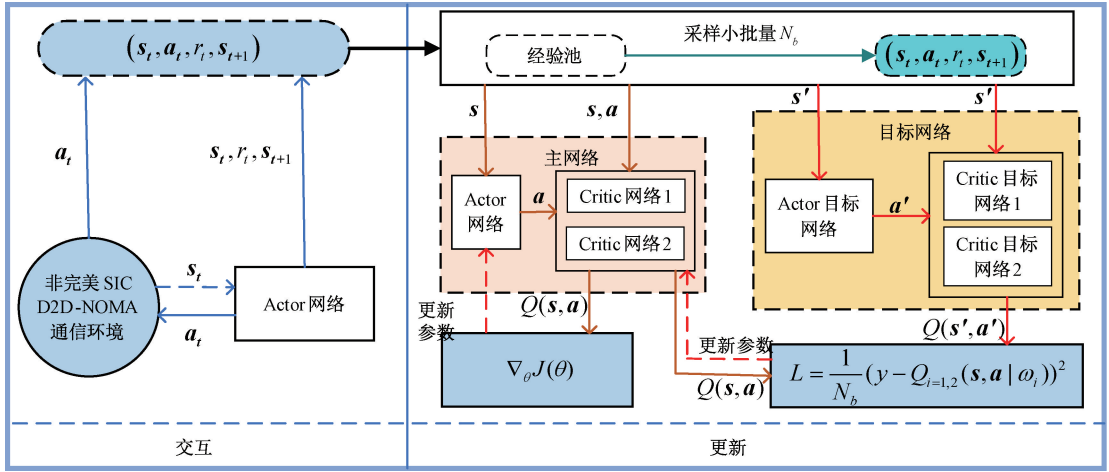


图 2 非完美 SIC D2D-NOMA 系统基于 TD3 算法的功率分配结构

Fig. 2 Power allocation structure of D2D-NOMA system based on TD3 algorithm under imperfect SIC

TD3 算法通过智能体与环境的交互,利用策略梯度算法来评估和改进参数化策略,从而最大化累积回报。在 D2D-NOMA 系统的功率分配场景中,TD3 算法包括 6 个网络组成,分别是 Actor 网络、Critic 网络 1、Critic 网络 2、Actor 目标网络、Critic 目标网络 1、Critic 目标网络 2,分别用于生成动作、评估动作价值和目标网络更新。整个算法流程可以分为 3 个步骤:Actor 过程、Critic 过程和策略和 Q 网络更新。

1) Actor 过程

在 Actor 过程中,TD3 使用策略梯度算法来改进策略网络,Actor 网络由两个深度神经网络组成,用于生成动作,同时通过 Actor 目标网络来保持稳定性。Actor 网络的参数设为 θ ,目标策略网络的参数设为 θ' 。在每个时刻 t ,Actor 网络从环境中接收状态向量 $\mathbf{s}(t)$,并通过确定性策略梯度方法生成动作 $\mathbf{a}(t)$,表达式为:

$$\mathbf{a}(t) = \pi_{\theta}(\mathbf{s}(t)) + N_t \quad (20)$$

式中: $N_t \sim \text{clip}(N(0, \sigma), -c, c)$ 表示平滑噪声,用于增强智能体的探索能力。Actor 网络的目标是最大化回报函数,通过策略梯度优化在线策略网络的参数 θ 。梯度的计算公式为:

$$\nabla_{\theta} J(\theta) = E_{\theta}[\nabla_{\theta} \pi_{\theta}(\mathbf{s}) \nabla_{\mathbf{a}} Q(\mathbf{s}, \mathbf{a}) |_{\mathbf{a}=\pi_{\theta}(\mathbf{s})}] \quad (21)$$

为减少训练数据间的相关性,引入经验回放机制,将每个时间步的经验 $(\mathbf{s}(t), \mathbf{a}(t), r(t), \mathbf{s}(t+1))$ 存储到回放缓冲区中,从中随机采样小批量数据,并使用蒙特卡罗方法估计梯度。

2) Critic 过程

Critic 网络用于评估 Actor 网络的性能,包含 Critic 网络 1 和 Critic 网络 2,以及对应的 Critic 目标网络 1 和 Critic 目标网络 2。这种双 Critic 架构用于减少 Q 值的过估计问题。对于每个采样的批量数据,状态 \mathbf{s} 和动作 \mathbf{a} 输入 Critic 网络 1 和 Critic 网络 2,分别计算出当前 Q 值 $Q_1(\mathbf{s}, \mathbf{a} | \omega_1)$ 和 $Q_2(\mathbf{s}, \mathbf{a} | \omega_2)$,其中 ω_1 和 ω_2 为 Critic 网络 1 和 Critic 网络 2 的参数。同时,目标 Q 网络生成目标 Q 值 y :

$$y = r + \gamma \min(Q_1(\mathbf{s}', \mathbf{a}' | \omega_1), Q_2(\mathbf{s}', \mathbf{a}' | \omega_2)) \quad (22)$$

式中: r 是奖励, γ 是范围在 $[0, 1]$ 的折扣因子, \mathbf{a}' 为 Actor 目标网络生成的下一步动作。通过最小化均方误差来更新 Critic 网络 1 和 Critic 网络 2 的参数,从而提高 Actor 网络的性能评估能力。具体公式为:

$$L = \frac{1}{N_b} (y - Q_1(\mathbf{s}, \mathbf{a} | \omega_1))^2 + (y - Q_2(\mathbf{s}, \mathbf{a} | \omega_2))^2 \quad (23)$$

3) 策略和 Q 网络的更新

TD3 算法通过策略评估和策略改进过程来优化策略。策略评估通过 Critic 网络 1 和 Critic 网络 2 估计 Q 值函数,而策略改进通过 Actor 网络的梯度上升方法来优化策略。在连续动作空间中,最大化动作值函数需要近似方法。Actor 网络的参数更新公式为:

$$\theta(t+1) = \theta(t) + \alpha_\theta \nabla_\theta J(\theta) \quad (24)$$

式中: α_θ 为学习率。每次更新后,噪声项 N_t 通过衰减因子减小,以增加策略的利用率。

TD3 还使用软更新模式来稳定目标网络。通过设置小步长 τ , 将旧的目标网络参数与当前网络参数进行加权平均,以此平滑更新,更新公式如下:

$$\theta' \leftarrow \tau\theta + (1-\tau)\theta' \quad (25)$$

$$\omega'_1 \leftarrow \tau\omega_1 + (1-\tau)\omega'_1 \quad (26)$$

$$\omega'_2 \leftarrow \tau\omega_2 + (1-\tau)\omega'_2 \quad (27)$$

具体算法伪代码如算法 1 所示。

算法 1 非完美 SIC D2D-NOMA 基于 TD3 的功率分配算法

输入: Actor 网络和 Critic 网络结构、环境参数

输出: 每个 D2D 组接收机的功率分配策略

初始化: Actor 网络, Critic 网络和回放缓冲区

- 1: For $episode=1$: 最大训练步长 do
- 2: 从环境中获取初始状态 \mathbf{s}
- 3: For $t=1$: 最大步数 do
- 4: 根据策略 $\pi(\mathbf{s})$ 从 Actor 网络中选择动作 \mathbf{a} , 加入高斯噪声 N_t , 得到 $\mathbf{a}' = \pi(\mathbf{s}) + N_t$
- 5: 执行动作 \mathbf{a}' , 观察奖励 r , 获取下一个状态 \mathbf{s}'
- 6: 将存储经验元组 $(\mathbf{s}, \mathbf{a}', r, \mathbf{s}')$ 到回放缓冲区中, 并从回放缓冲区采样进行训练
- 7: For $i=1$: 智能体最大数 do
- 8: 通过使用采样的小批量更新 Critic1 和 Critic2 网络参数来最小化损失
- 9: 使用采样的策略梯度更新 Actor 网络参数
- 10: 更新 Critic 和 Actor 网络的目标网络参数
- 11: End For
- 12: End For
- 13: End For

3 仿真与分析

3.1 仿真环境

D2D-NOMA 系统的仿真场景如图 3 所示。图 3 中显示了 10 个蜂窝用户, 20 个 D2D 组的情况, 并在 Pycharm 平台上, 利用 Keras 3.3.6 和 TensorFlow 2.10 对模型进

行评估, 仿真参数如表 1 所示。TD3 的深层神经网络架构如表 2 所示。

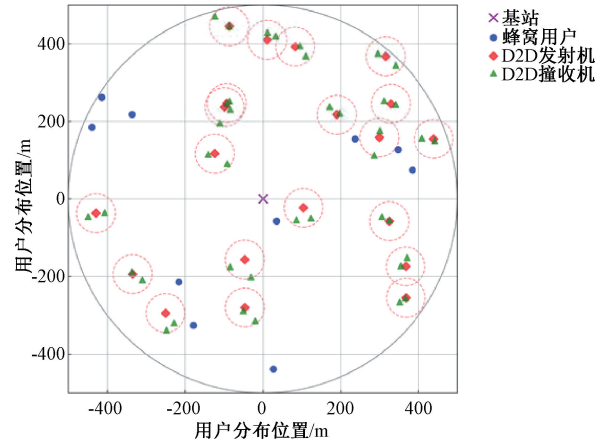


图 3 D2D-NOMA 系统仿真场景

Fig. 3 Simulated scenario of D2D-NOMA system

表 1 仿真参数

Table 1 Simulation parameter

参数	参数设置
高斯白噪声功率/dBm	-174
蜂窝用户最大发射功率/dBm	23
D2D 发射机最大发射功率/dBm	30
基站半径/m	500
D2D 最大传输距离/m	50
蜂窝用户个数	10
D2D 发射机个数	20
Actor 网络学习率	0.000 1
Critic 网络学习率	0.000 1
折扣因子	0.99
软更新因数	0.99
经验回放池容量	20 000
探索噪声	(0, 0.1)
策略平滑噪声	(0, 0.05)
路径损耗指数	3
非完美 SIC 参数	0.1, 0.2, 0.3, 0.4, 0.5

表 2 TD3 的深层神经网络架构

Table 2 TD3's deep neural network architecture

网络	层次	神经数量	激活函数
Actor	输入层	状态向量	无
Actor	全连接层	256	ReLU
Actor	全连接层	128	ReLU
Actor	输出层	1	Tanh
Critic	输入层	状态向量+动作向量	无
Critic	全连接层	512	ReLU
Critic	全连接层	256	ReLU
Critic	输出层	1	无

3.2 仿真结果分析

图 4 为不同 Actor 网络学习率的所提算法收敛性比较。从图中可以看出,在训练次数为 1 500 左右时,系统吞吐量都达到了稳定值,并且学习率为 0.000 1 最终的收敛值达到 190 Mbps 左右,表明所提算法可以在较短的时间内最大化 D2D 用户的吞吐量,体现出了优异的收敛性能。

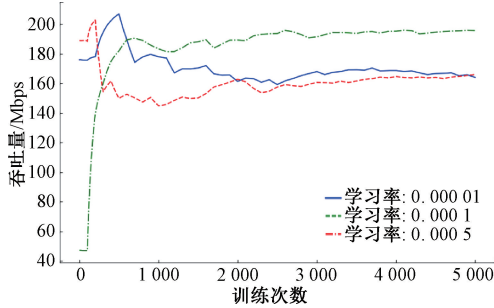


图 4 不同学习率的收敛性能

Fig. 4 Convergence performance of different learning rates

图 5 给出了不同功率分配算法下,不同 D2D 小组数量与系统吞吐量的关系图。从图 5 可以看出,随着 D2D 组数不断增加,系统 D2D 用户吞吐量均不断增加,且所提算法性能最优。

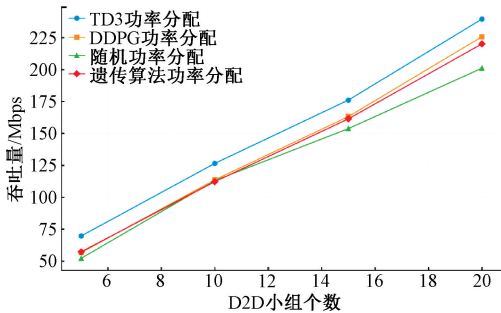


图 5 不同 D2D 组数对系统吞吐量的影响

Fig. 5 Impact of different D2D group numbers on system throughput

图 6 为所提算法在提高系统 D2D 用户吞吐量上表现出明显的优势。与遗传算法、DDPG 算法和随机功率分配方法相比,所提功率分配算法实现了更高的吞吐量。当 D2D 发射功率为 15 dBm 时,相较于 DDPG、遗传算法和随机分配方法分别提升了 14%、25% 和 40% 左右。

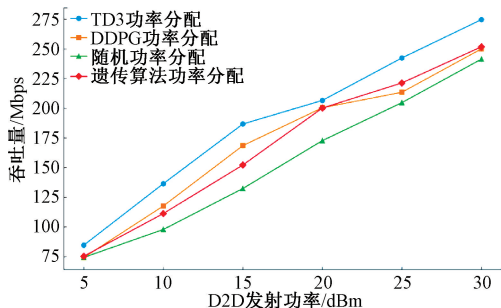


图 6 不同功率分配算法的性能比较

Fig. 6 Performance comparison of different power allocation algorithms

图 7 为不同算法的公平性指数与 D2D 发射功率之间的关系。采用的 Jain 公平性指数通常是用来衡量资源分配的公平性,这里用来衡量所提的功率分配算法的公平性。通过改变用户的平均吞吐量来测量所提出的算法的公平性,公式为:

$$F = \frac{(\sum_{n=1}^N R_{sum})^2}{N \sum_{n=1}^N (R_{sum})^2} \quad (28)$$

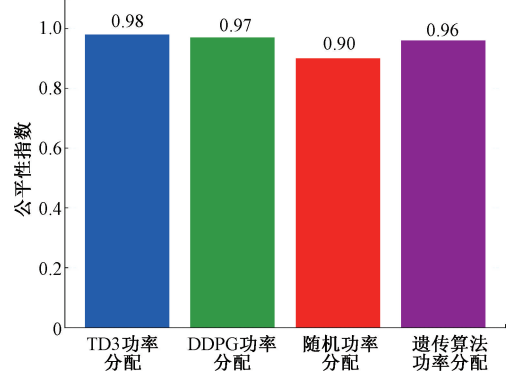


图 7 不同功率分配算法的公平性能比较

Fig. 7 Fairness comparison of different power allocation algorithms

如图 7 所示,所提功率分配算法公平性指数最高,表现优于其他算法,说明它在功率分配中能更有效地平衡各用户的公平性需求,更好地体现了所提算法的优势。

如图 8 所示,当考虑非完美 SIC 时,D2D-NOMA 通信系统的吞吐量受到明显的影响,尤其随着 D2D 发射功率的增加,影响逐渐显著,这是由于一方面随着 D2D 发射功率的增加,接收端的信噪比逐渐恶化,影响系统的吞吐量增益;另一方面随着 D2D 发射功率的增加,残留干扰信号的强度也随之增加,使得吞吐量的提升速度逐渐变缓。

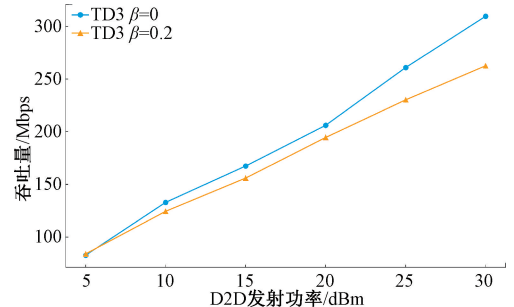


图 8 完美 SIC 和非完美 SIC($\beta=0.2$) 对系统 D2D 用户吞吐量的影响

Fig. 8 Impact of perfect SIC and imperfect SIC($\beta=0.2$) on system D2D user throughput

从图 9 可以看出,在固定发射功率为 20 dBm 时,当 β 从 0.1 增大到 0.5 时,吞吐量下降了约 28%。这是因为非完美 SIC 参数 β 值越大,NOMA 解码的误差越大,从而导致系统吞吐量下降显著。因此,非完美 SIC 的情况不容忽视

视,并且确保接收端的 SIC 解码精度对于提高系统的吞吐量至关重要。

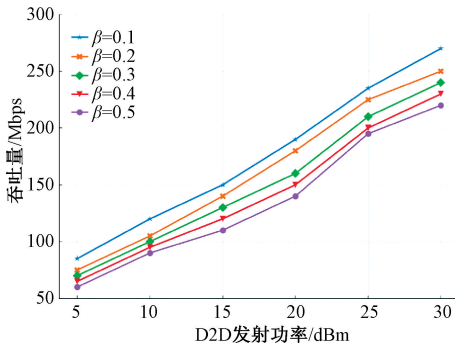


图 9 非完美 SIC 参数 β 对系统 D2D 用户吞吐量的影响

Fig. 9 Impact of imperfect SIC parameter β on system D2D user throughput

4 结 论

本文研究了 D2D-NOMA 通信系统中非完美 SIC 的吞吐量最大化问题。通过仿真实验与遗传算法、随机功率分配和 DDPG 功率分配算法相比较,验证了所提 TD3 算法在非完美 SIC 条件下展现出显著的性能优势,不仅提升了系统的吞吐量,还在动态环境中表现出更优的收敛性和公平性。本文考虑了 BS 已知完全信道状态信息,但实际上,完全的信道状态信息无法获取,因此,在非完美信道状态信息的条件下,如何在大规模用户和高干扰环境中有效提升系统吞吐量,将成为一个重要且具有挑战性的研究方向。

参 考 文 献

- [1] MARAQA O, RAJASEKARAN A S, AL-AHMADI S, et al. A survey of rate-optimal power domain NOMA with enabling technologies of future wireless networks [J]. IEEE Communications Surveys & Tutorials, 2020, 22(4): 2192-2235.
- [2] 江超,张治中,胡正操,等. 基于改进的猎食者优化的 D2D 通信功率控制方法[J]. 电子测量技术, 2023, 46(6):31-36.
JIANG CH, ZHANG ZH ZH, HU ZH C, et al. D2D communication power control method based on improved predator optimization [J]. Electronic Measurement Technology, 2023, 46(6): 31-36.
- [3] 沈国丽,李君,李正权. D2D 通信中基于深度强化学习的资源分配[J]. 电子测量技术, 2022, 45(24):76-84.
SHEN G L, LI J, LI ZH Q. Resource allocation in D2D communication based on deep reinforcement learning [J]. Electronic Measurement Technology, 2022, 45(24):76-84.
- [4] ZHAO X, LIU F, ZHANG Y J, et al. Energy-

efficient power allocation for full-duplex device-to-device underlying cellular networks with NOMA[J]. Electronics, 2023, 12(16): 3433.

- [5] KHAN M A A, KAIDI H M, AHMAD N, et al. Sum throughput maximization scheme for NOMA-enabled D2D groups using deep reinforcement learning in 5G and beyond networks [J]. IEEE Sensors Journal, 2023, 23(13): 15046-15057.
- [6] ALEMAISHAT S, SARAEREH O A, KHAN I, et al. An efficient resource allocation algorithm for D2D communications based on NOMA[J]. IEEE Access, 2019, 7: 120238-120247.
- [7] LI Q, REN P Y, XU D Y. Security enhancement and QoS provisioning for NOMA-based cooperative D2D networks[J]. IEEE Access, 2019, 7: 129387-129401.
- [8] 宋勇春,王茜竹,高正念. 基于 HAGA 的 D2D-NOMA 资源分配优化算法[J]. 计算机工程, 2022, 48(2): 275-280.
SONG Y CH, WANG Q ZH, GAO ZH N. D2D-NOMA resource allocation optimization algorithm based on HAGA[J]. Computer Engineering, 2022, 48(2): 275-280.
- [9] TLEBALDIYEVA L, NAURYZBAYEV G, ARZYKULOV S, et al. Performance of NOMA-based mmWave D2D networks under practical system conditions [J]. IEEE Access, 2021, 9: 160958-160974.
- [10] 杨青青,肖春娟,彭艺. NOMA 残余干扰下 D2D 网络的能效资源分配算法[J]. 重庆邮电大学学报(自然科学版)2023, 35(4):606-614.
YANG Q Q, XIAO CH J, PENG Y. Energy efficiency resource allocation algorithm for D2D networks with NOMA residual interference [J]. Journal of Chongqing University of Posts & Telecommunications(Natural Science Edition), 2023, 35(4):606-614.
- [11] 周烁,仇润鹤. 不完全 SIC 下 CR-NOMA 系统的功率分配算法[J]. 计算机工程, 2021, 47(8):195-200, 209.
ZHOU SH, QIU R H. Power allocation algorithm of CR-NOMA system under imperfect SIC[J]. Computer Engineering, 2021, 47(8):195-200, 209.
- [12] WU G W, CHEN G F, CHEN G J. Energy-efficient utility function-based channel resource allocation and power control for D2D clusters with NOMA enablement in cellular networks [J]. IEEE Access, 2023, 11: 45001-45010.
- [13] YU SH SH, KHAN W U, ZHANG X Q, et al. Optimal power allocation for NOMA-enabled D2D

- communication with imperfect SIC decoding [J]. Physical Communication, 2021, 46: 101296.
- [14] SHI ZH Y, XIE X ZH, LU H B, et al. Deep reinforcement learning based dynamic user access and decode order selection for uplink NOMA system with imperfect SIC [J]. IEEE Wireless Communications Letters, 2020, 10(4): 710-714.
- [15] CHRYSOLOGOU A P, CHATZIDIAMANTIS N D, KARAGIANNIDIS G K. Cooperative uplink NOMA in D2D communications [J]. IEEE Communications Letters, 2022, 26(11): 2567-2571.
- [16] 江超,张治中,冯姣,等.蜂窝网络 D2D 通信资源分配研究[J].电子测量技术,2023,46(24):47-53.
JIANG CH, ZHANG ZH ZH, FENG J, et al. Study on D2D communication resource allocation in cellular network [J]. Electronic Measurement Technology, 2023, 46(24): 47-53.
- [17] LE CH B, DO D T. Joint evaluation of imperfect SIC and fixed power allocation scheme for wireless powered D2D-NOMA networks with multiple antennas at base station [J]. Wireless Networks, 2019, 25(8): 5069-5081.
- [18] WANG X, SHI H T, LI Y Q, et al. Energy efficiency resource management for D2D-NOMA enabled network: A dinkelbach combined twin delayed deterministic policy gradient approach [J]. IEEE Transactions on Vehicular Technology, 2023, 72(9): 11756-11771.

作者简介

张东方(通信作者),硕士研究生,主要研究方向为网络优化、深度强化学习。

E-mail:2479519352@qq.com

叶成荫,副教授,主要研究方向为 Internet 的控制与决策、5G 网络优化、机器学习等。

E-mail:chengyinye@foxmail.com