

基于图卷积神经网络的超密集物联网资源分配策略

黄杰, 李幸星, 杨凡, 丁睿杰, 蔡杰良, 姚凤航, 张鑫

(重庆理工大学电气与电子工程学院, 重庆 400054)

摘要: 针对超密集物联网 (UD-IoT) 中存在大量隐藏终端干扰严重影响资源管理问题, 提出了一种基于图卷积神经网络的深度确定性梯度的超密集物联网资源分配策略。通过矩阵变换构建冲突图模型, 采用极大团和超图理论将冲突图模型转化为冲突超图模型, 进而将无冲突资源分配问题转化为超图顶点着色问题, 并提出了一种基于图卷积神经网络的深度确定性梯度的超密集物联网资源分配算法, 采用图卷积强化学习实现无冲突资源分配和资源复用率最大化。仿真实验表明, 所提算法具有更高的资源复用率和吞吐量, 可以在超密集物联网中提供更好的性能。

关键词: 超密集物联网; 资源分配; 深度强化学习; 图卷积神经网络

中图分类号: TN929.5

文献标志码: A

DOI: 10.11959/j.issn.1000-436x.2024178

Resource allocation strategy for ultra-dense Internet of things based on graph convolutional neural network

HUANG Jie, LI Xingxing, YANG Fan, DING Ruijie, CAI Jieliang, YAO Fenghang, ZHANG Xin

School of Electrical and Electronic Engineering, Chongqing University of Technology, Chongqing 400054, China

Abstract: To address the significant issue of hidden terminal interference that severely impacted resource management in ultra-dense Internet of things (UD-IoT) environments, a deep deterministic gradient-based conflict-free resource allocation strategy using graph convolution neural network was proposed. The conflict graph model was constructed by employing matrix transformations to represent potential hidden terminal interference among devices. Then, using the concepts of maximal cliques and hypergraph theory, the conflict graph model was transformed into a conflict hypergraph model. This transformation allowed the conflict-free resource allocation problem to be formulated as a hypergraph vertex coloring problem. A deep deterministic gradient-based conflict-free resource allocation algorithm, leveraging graph convolutional neural network reinforcement learning, was developed to achieve conflict-free resource allocation and maximize resource reuse. Simulation results demonstrated that the proposed algorithm achieved higher resource reuse rates and throughput compared to existing methods, providing superior performance in ultra-dense IoT.

Keywords: ultra-dense Internet of things, resource allocation, deep reinforcement learning, graph convolutional neural network

0 引言

随着无线通信技术的发展, 物联网 (IoT, Inter-

net of things) 迅速发展, 物联网设备 (IoT device, Internet of things device) 数量激增^[1]。2022年, 无线基

收稿日期: 2024-06-15; 修回日期: 2024-09-10

通信作者: 杨凡, 34179861@qq.com

基金项目: 国家自然科学基金资助项目 (No.62301094); 重庆市教育委员会科技研究基金资助项目 (No.KJQN202201157, No. KJQN202301135); 重庆理工大学科研创新团队培育计划基金资助项目 (No.2023TDZ003)

Foundation Items: The National Natural Science Foundation of China (No.62301094), Science and Technology Research Program of Chongqing Education Commission (No.KJQN202201157, No.KJQN202301135), Cultivation Program of Scientific Research and Innovation Team of Chongqing University of Technology (No.2023TDZ003)

基础设施中存在 132 亿 IoTD 连接, 并且以每年 18% 的增长率增长, 预计在 2028 年达到 347 亿。这意味着物联网在未来会演变为超密集物联网 (UD-IoT, ultra-dense Internet of things), 并会面临网络吞吐量和频谱资源利用率有限等挑战^[2]。终端直通 (D2D, device-to-device) 方式可用于缓解 UD-IoT 中网络吞吐量和资源利用率有限等挑战, D2D 通信支持 2 个 IoTD 之间直接通信, 并在 2 个设备之间共享资源^[3]。然而, 当 IoTD 间通信范围交叠且采用同一信道进行传输时, 将会出现严重的干扰现象。因此, 在 D2D 辅助的 UD-IoT 中, 必须依靠有效的资源管理方法, 既保障网络性能, 又避免严重干扰。

针对 D2D 辅助的 UD-IoT 资源管理, 现有研究多关注网络频谱效率和吞吐量等性能, 且多采用优化方法解决资源分配问题^[4-9]。针对毫米波频段中 D2D 通信的资源分配问题, 文献[4]提出了部分资源复用方法, 该方法将信道分配给设备使小区间干扰最小化, 进而提升网络吞吐量。针对大规模 5G 窄带物联网 D2D 通信中干扰控制和频谱利用问题, 文献[5]基于匈牙利算法提出避免干扰的资源分配方案, 考虑了小区边缘窄带用户设备, 提高了数据传输速率。文献[6]建立了一个基于功率分配、无线电资源分配和传输时间比的联合优化模型, 有效地提高整体吞吐量和频谱效率。为保证用户的服务质量, 文献[7]设计了一种双时间尺度资源分配方案, 减少了资源分配开销, 提高了用户体验。文献[8]提出了一种基于顺序协商的低复杂度算法, 实现集体信道的 D2D 应用, 从而增强设备的总容量。针对 D2D 通信的蜂窝网络在无线工业应用中的资源管理, 文献[9]基于匹配博弈提出了一种资源管理方案, 通过激励机制实现 D2D 用户与蜂窝资源的匹配, 实现最大化整体系统吞吐量。然而, 上述方法高度依赖于优化模型的准确性且计算量随网络规模极速增长, 难以适用于拥有海量 IoTD 的超密集环境。

随着人工智能 (AI, artificial intelligence) 的发展, 机器学习已经成为处理大量数据、高计算任务和数学上复杂的非线性非凸问题的非常有效的技术^[10]。针对网络规模极速增长, 越来越多的研究人员将强化学习应用于无线通信系统的资源管理和分配^[11-17]。在超密集小区网络中, 为应对多目标优化和跨小区干扰带来的复杂性, 文献[12]提出了一种基于多智能体强化学习的解决方案, 解决了具有

小区间干扰和多个优化目标的资源分配问题。考虑强化学习收敛速度较慢, 文献[13]引入了斯塔克尔伯格 Q 值来指导学习方向更快地收敛。针对移动物联网中动态资源分配问题, 文献[14]提出了一种在线深度强化学习 (DRL, deep reinforcement learning) 算法, 根据系统的实时反馈数据动态地调整资源分配方法。针对 D2D 辅助蜂窝网络的频谱接入问题, 文献[15]建立了基于 DRL 的频谱接入方案, 在设备无先验知识情况下, 自主学习最优策略以实现高效频谱访问, 从而使总吞吐量最大。针对物联网中设备面临的邻近增益、干扰和公平性等问题, 文献[16]提出了一种基于 DRL 的联合功率控制和资源块调度方案, 提高了网络数据传输速率。针对密集智能终端部署带来的严重小区干扰, 文献[17]提出了一种面向干扰的无线电资源分配框架和基于双深度 Q 网络的高性能资源分配算法, 实现了精确干扰识别和有效跨小区干扰协调。现有研究已证明, 基于强化学习的资源分配算法可以匹配实际网络的动态场景参数, 且适用于大规模网络。虽然当前研究考虑了小区干扰^[17]、同频干扰^[18]、交叠干扰^[19]等, 但是并未考虑超密集物联网密集部署场景下的海量隐藏终端干扰的问题。在超密集物联网中, 海量 IoTD 密集部署使得终端间通信范围密集交叠覆盖, 导致大量潜在隐藏终端干扰, 极大提高了资源复用冲突可能性, 严重增加了超密集物联网中资源管理的难度。

针对超密集物联网密集部署的海量隐藏终端干扰下的资源分配问题。本文提出了一种基于图卷积神经网络的深度确定性策略梯度 (GCNN-DDPG, graph convolutional neural network-deep deterministic policy gradient) 算法的超密集物联网资源分配策略。首先通过构建冲突图模型分析资源复用冲突关系, 然后将冲突图模型转化为冲突超图模型同时分析多个传输链路 (TL, transmission link) 之间的冲突关系, 并将资源分配问题转化为超图的顶点强着色问题。最后通过 GCNN-DDPG 算法实现无冲突资源分配。本文主要贡献包括以下 3 个方面。

1) 针对 D2D 辅助的 UD-IoT 中的隐藏终端干扰问题, 本文分析了 TL 间的资源复用冲突关系并对 TL 间的冲突类型进行分类。然后, 基于 UD-IoT 层之间的资源冲突关系, 建立了冲突图模型, 并设计了一种矩阵变换的方法构造冲突图, 直观地反映了

UD-IoT层之间的资源冲突关系。

2)针对冲突图模型不能同时分析多TL间的冲突关系,本文通过极大团和超图理论将冲突图模型转化为冲突超图,将无冲突资源分配问题转化为超图顶点着色问题。

3)针对超图顶点着色问题,本文提出了一种基于GCNN-DDPG的超密集物联网资源分配策略。GCNN-DDPG算法采用了Actor-critic网络学习D2D辅助的UD-IoT中资源分配过程,并根据经验回放池中的样本数据动态调整资源分配方案。仿真结果表明,在保证UD-IoT无冲突的情况下,实现了资源复用率的提高,并提高了UD-IoT的网络吞吐量。

1 系统模型

本文场景为超密集物联网,包括基带单元(BBU, building base band unit)池和D2D辅助的UD-IoT层,如图1所示。UD-IoT包含 $N = \{1, 2, \dots, n\}$ 个IoTD,在IoTD之间存在 $M = \{1, 2, \dots, m\}$ 个TL,以及多个射频拉远头(RRH, remote radio head)。RRH通过高速传输链路连接BBU池,RRH负责基本覆盖和辅助接入。IoTD采用D2D通信方式,并且根据多跳TL进一步扩大通信范围,最后通过UD-IoT层的RRH连接BBU池。BBU池由BBU和服务器组成,负责收集网络中所有环境信息,并将资源分配给UD-IoT中的TL。在超密集物联网中,海量IoTD密集部署使得终端间通信范围密集交叠覆盖,导致大量潜在隐藏终端干扰。因此,本文主要考虑UD-IoT中海量隐藏终端

干扰问题。

在本文中令 D_{ij} 表示任意2个IoTD i 和IoTD j 之间的欧氏距离,假设所有的IoTD具有相同的传输功率,并且采用瑞利衰落模型。 S_{ij} 为接收端IoTD i 收到发送端IoTD j 发送的信号。IoTD i 的信干噪比(SINR, signal-to-interference-plus-noise ratio)可表示为

$$\gamma_{ij} = \frac{S_{ij}}{Z + \sum_{k \neq i} S_{ik}} \quad (1)$$

其中, Z 为环境噪声, k 表示其他IoTD发送的干扰。

当SINR达到特定的阈值 s_{th} 时,接收端 j 可正常解码,根据式(1),传输速率可表示为

$$R_{ij} = B \ln(1 + \gamma_{ij}) \quad (2)$$

其中, B 为信道带宽, γ_{ij} 表示IoTD i 的信干噪比。

假设所有的IoTD具有相同通信的半径表示为 $O_T = \left(\frac{P}{s_{th}Z}\right)^{\frac{1}{\alpha}}$,其中 P 表示IoTD的传输功率, α 表示衰落因子。

根据所有IoTD之间的欧氏距离 D_{ij} 和通信半径 O_T ,本文建立了基于图论的通信模型。通信模型可以用 $G = \{V, E\}$ 表示,其中 $V = \{v_1, v_2, \dots, v_n\}$ 为顶点集,表示所有的IoTD; $E = \{e_1, e_2, \dots, e_m\}$ 为边集,表示TL。顶点与边的关系表示为

$$(e_m, v_n) = \begin{cases} 0, v_n \notin e_m \\ 1, v_n \in e_m \end{cases} \quad (3)$$

其中, $(e_m, v_n) = 1$ 表示顶点 v_n 与边 e_m 相关联,即

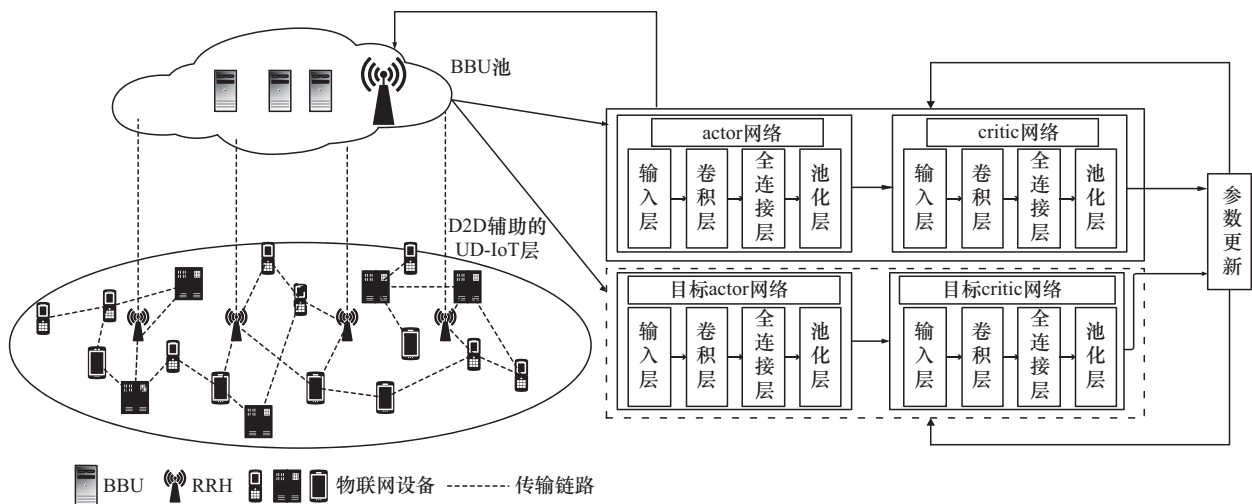


图1 超密集物联网场景

IoTD 可以通过 TL 与相连接的其他 IoTD 通信，反之则不能通信。故可以写出其关联矩阵 H 为

$$H = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \quad (4)$$

其中， a_{mn} 表示 (e_m, v_n) 的取值。

在通信模型中，存在 2 种无资源冲突情况，分别为：1) 顶点不与任何边相关联，即它是一个孤立的顶点；2) 只有 2 个顶点和一条边，这就意味着 TL 在作资源分配时不会与其他 TL 冲突。在这 2 种情况下，可以直接将通信资源分配给 TL。为了方便分析超密集物联网密集部署的海量隐藏终端干扰问题，图 2 给出一个 D2D 通信示例，其中包括 20 个 IoTD 和 32 个 TL。

为避免 D2D 辅助的 UD-IoTD 中 TL 之间的资源冲突，对潜在的冲突类型进行分类分析，分类如下。

直接冲突： TL₁ 和 TL₂ 如果同时使用同一个信道，则会发生直接冲突。例如，图 2 中 IoTD₁ 和 IoTD₃ 同时向 IoTD₂ 发送信息则直接冲突。

隐藏终端冲突： TL₁ 和 TL₃ 如果同时使用一个信道，则会发生隐藏终端冲突。例如，图 2 中 IoTD₁ 向 IoTD₂ 发送消息，同时 IoTD₃ 向 IoTD₄ 发送消息，则会发生隐藏终端冲突。

直接冲突问题本质上是典型的边着色问题，可以通过边着色算法解决。然而，隐藏终端冲突问题不适合利用典型的边着色算法来解决。因此，有必要进一步建立冲突模型并分析。

2 超密集物联网的冲突超图模型

本节分析了 TL 之间的资源冲突，根据资源冲突类型和 UD-IoT 通信网络拓扑，通过矩阵变换建立冲突图模型。为了同时分析多 TL 间冲突关系，基于极大团和超图理论将冲突图转换为冲突超图。冲突超图在保证冲突信息不变的前提下，能够同时分析多 TL 间的冲突关系。

2.1 冲突图模型

令 $G_c = \{V_c, E_c\}$ 表示冲突图模型，其中 $V_c = E = \{e_1, e_2, \dots, e_m\}$ 表示所有的 TL 集合， E_c 表示 TL 之间资源冲突关系的边集合。顶点与边之间邻接矩阵 G_e 表示为

$$G_e = \begin{pmatrix} b_{11} & \cdots & b_{1m} \\ \vdots & \ddots & \vdots \\ b_{n1} & \cdots & b_{nm} \end{pmatrix} \quad (5)$$

其中， $b_{nm} = (e_n, e_m)$ ，若 e_n 与 e_m 有冲突，则 $b_{nm} = 1$ ，反之则为 0。

在 D2D 辅助的 UD-IoT 中，TL_{*i*} 和 TL_{*j*} 的可达性表示 TL_{*i*} 到 TL_{*j*} 存在一条路径，路径长度为路径中顶点的数量。TL_{*i*} 和 TL_{*j*} 之间如果存在多条路径，但其中任何一条路径发生冲突，它们就不能共享资源。因此，在分析冲突关系中无须考虑造成 TL 之间冲突的路径数量。根据可达性，互为直接冲突的 TL 只需一步即可到达，所以直接冲突计算方法为

$$G_{e_1} = (G_g^T G_g) \odot 1 \quad (6)$$

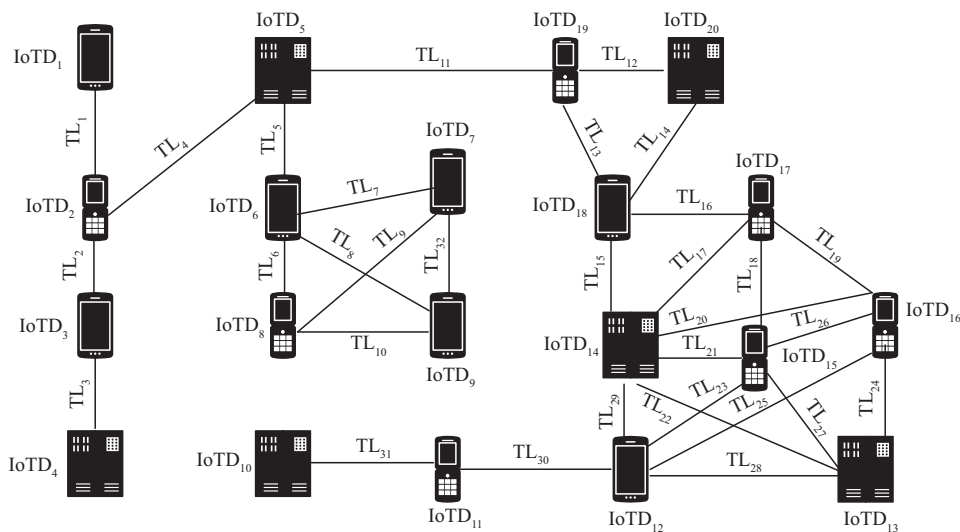


图 2 D2D 通信示例

$$c \odot d = \begin{cases} 1, c \geq d \\ 0, \text{其他} \end{cases} \quad (7)$$

其中, $G_{e_1} \in \mathbf{R}^{|E| \times |E|}$ 表示直接冲突的冲突图邻接矩阵, G_g 为图 2 的关联矩阵, G_g^T 是 G_g 的转置矩阵, $G_g \in \mathbf{R}^{|V| \times |E|}$, \odot 表示一种运算规则, c 和 d 表示任意 2 个数。

为了获得不同 TL 间的直接冲突关系, 将主对角线元素设为 0, 即

$$\hat{G}_{e_1} = G_{e_1} - I \quad (8)$$

其中, I 表示单位矩阵。

与式(6)类似, 隐藏终端冲突的计算方法为

$$G_{e_2} = (\hat{G}_{e_1}^T \hat{G}_{e_1}) \odot 1 \quad (9)$$

与式(8)同理, G_{e_2} 中的主对角线元素设置为 0, 即

$$\hat{G}_{e_2} = G_{e_2} - I \quad (10)$$

根据式(8)和式(10), 得到冲突图的邻接矩阵 G_e 为

$$G_e = (\hat{G}_{e_1} + \hat{G}_{e_2}) \odot 1 \quad (11)$$

根据 G_e 可得冲突图, 如图 3 所示。

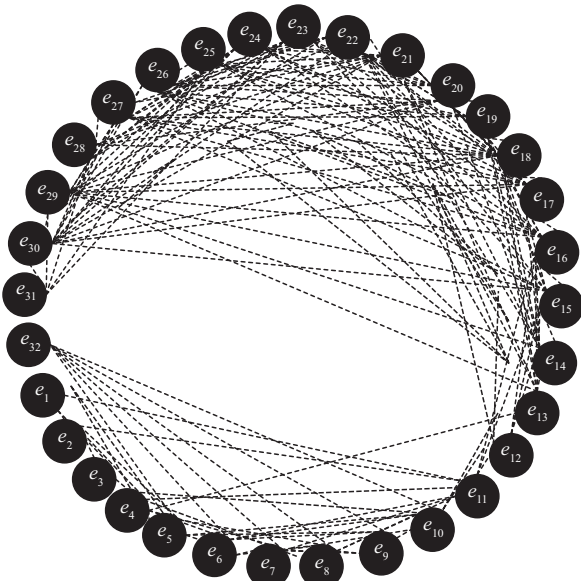


图 3 冲突图

图 3 表示 TL 间的冲突关系。由于冲突图中的边仅包括 2 个节点, 难以反映多 TL 的资源冲突关系。由于海量终端的密集部署, 多 TL 的资源冲突是造成隐藏终端干扰的主要原因。因此, 有必要进一步简化冲突图使其保持顶点之间冲突关系不变并

同时分析多个 TL 之间的冲突关系。

2.2 冲突超图模型

为了同时分析多个 TL 之间的冲突关系, 基于极大团和超图理论, 本文将冲突图简化为冲突超图。

令 $G_H = \{V_H, E_H\}$ 表示冲突超图, V_H 表示顶点集, E_H 表示超边集。超边中的所有顶点具有相同的关系。超图可以用一个关联矩阵 $H_g \in \mathbf{R}^{|V_H| \times |E_H|}$ 来表示, 其元素为

$$h(v, e) = \begin{cases} 1, v \in e \\ 0, \text{其他} \end{cases} \quad (12)$$

其中, $h(v, e)$ 是 H_g 中第 v 行 e 列的元素, $h(v, e) = 1$ 表示顶点 v 在超边 e 上, 即 v 和 e 是相关联的。

由于冲突图 $G_c = \{V_c, E_c\}$ 中顶点之间高密度连接的性质, 可生成许多完全子图 $G_s = \{V_s, E_s\}$ 。每个完全子图都为一个团, 其所有顶点彼此相邻, 可用一条超边表示。因此, 根据超边与极大团的关系, 在保持顶点相邻关系不变的情况下, 可以将冲突图简化为冲突超图。根据极大团生成定理^[20]以及贪心算法可求出冲突图中所有的极大团, 并将冲突图转换为冲突超图, 如图 4 所示。图 4 中, 顶点表示 TL, 超边表示它们之间的冲突关系, 在同一超边上的 TL 存在冲突关系, 不能共享同一种通信资源。无冲突资源分配问题的本质是有冲突关系的 TL 不能共享同一种通信资源。在超图的顶点着色问题中, 同一条超边上的节点不能着色相同, 这与无冲突资源分配问题的本质是一致的。因此, 无冲突资源分配可以转化为冲突超图的顶点着色问题。

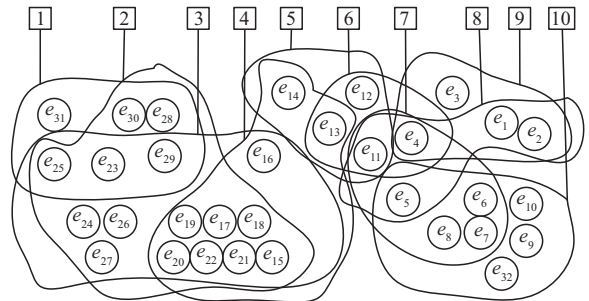


图 4 冲突超图

3 基于图卷积深度确定性梯度的强化学习模型

针对冲突超图中的顶点着色问题, 本文提出

GCNN-DDPG 算法首先构建考虑超密集物联网两类冲突的资源分配马尔可夫决策 (MDP, Markov decision process) 模型, 然后结合 GCNN-DDPG 算法求解最优资源分配策略 π^* 。

针对超图顶点着色问题, 首先通过马尔可夫决策过程将其转化为一个序列决策问题, 将 BBU 池作为智能体, 并结合 GCNN-DDPG 算法, 智能体通过训练不断寻找超图顶点着色方案。关于状态 S 、动作 A 、奖励 R 的具体定义如下。

1) 状态 S : $S = \{H, K_t\}$, 其中 H 是冲突超图的关联矩阵, K_t 表示所有 TL 的分配情况。

2) 动作 A : 智能体根据当前状态做出观察当前状态做出相应通信链路资源分配的集合, 即动作集。

3) 奖励 R : 智能体在当前状态下执行相应动作获得对应的回报。

如图 5 所示, 本文所提 GCNN-DDPG 算法结构由 actor 网络和 critic 网络组成, 前者用来生成动作, 后者用于评估生成动作的性能。本文算法创建 4 个网络用来提高稳定性和收敛性, 分别为 actor 网络、actor-target 网络、critic 网络、critic-target 网络, 并使用软更新对目标网络进行更新。

具体算法步骤如下。首先, 智能体将当前状态 s_t 输入 actor 网络中, 得到确定性策略 $a_t = \mu(s_t | \theta_\mu)$, 其中, $\mu(s_t)$ 表示 actor 网络, θ_μ 表示 actor 网络中的参数。然后, 智能体在环境中执行 a_t , 然后得到及

时奖励 r_t 并获得下一个状态 s_{t+1} 。最后, 将 (s_t, a_t, r_t, s_{t+1}) 存储在经验回放缓冲区用于进一步训练。当经验池存满, 随机选择 n 个数据来组成小批量数据。令 (s_n, a_n, r_n, s_{n+1}) 为小批量数据中一组数据。状态-动作 (s_n, a_n) 的估计值由 critic 网络得到, 用 $Q(s_n, a_n | \theta_Q)$ 表示, 其中 $Q(s, a)$ 表示 critic 网络, θ_Q 表示 critic 网络中的参数。状态-动作 (s_n, a_n) 的目标值可基于贝尔曼方程得到, 即

$$y_n = r_n + \gamma \bar{Q}(s_{n+1}, \bar{\mu}(s_{n+1} | \theta_{\bar{\mu}}) | \theta_{\bar{Q}}) \quad (13)$$

其中, γ 为折扣因子, $\bar{Q}(s, a)$ 表示 critic-target 网络, $\theta_{\bar{Q}}$ 表示 critic-target 网络参数, $\bar{\mu}(s)$ 表示 actor-target 网络, $\theta_{\bar{\mu}}$ 表示 actor-target 网络参数。

通过最小化均方误差 (MSE, mean squared error) 损失函数训练 critic 网络, 即

$$L(\theta_Q) = \frac{1}{N} \sum_{n=1}^N (y_n - Q(s_n, a_n | \theta_Q))^2 \quad (14)$$

critic 网络的梯度可通过对 $L(\theta_Q)$ 中 θ_Q 求微分得到, 即

$$\nabla_{\theta_Q} L(\theta_Q) = \frac{2}{N} \sum_{n=1}^N \nabla_{\theta_Q} Q(s, a | \theta_Q) |_{s=s_n, a=a_n} (y_n - Q(s_n, a_n | \theta_Q)) \quad (15)$$

actor 网络可通过最大化环境状态初始分布的预期收益进行训练, 即

$$J(\theta_\mu) = \frac{1}{N} \sum_{n=1}^N Q(s_n, \mu(s_n | \theta_\mu) | \theta_Q) \quad (16)$$

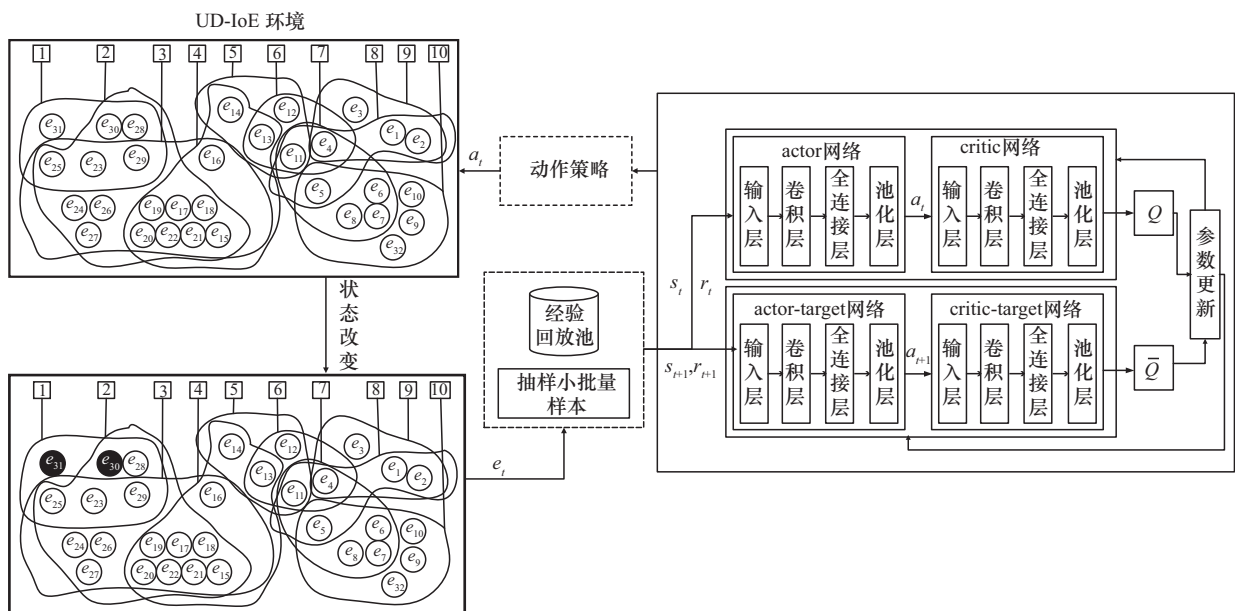


图 5 GCN-DDPG 算法

其梯度可通过微分的链式法则计算, 即

$$\nabla_{\theta_{\mu}} J(\theta_{\mu}) = \frac{1}{N} \sum_{n=1}^N \nabla_a Q(s, a | \theta_Q) |_{s=s_n, a=a_n} \nabla_{\theta_{\mu}} \mu(s | \theta_{\mu}) |_{s=s_n} \quad (17)$$

通过梯度对整个网络进行参数更新。

在 actor 网络和 critic 网络中建立两层图卷积模型来获取特征, 该模型可以表示为

$$f(\mathbf{X}, \mathbf{A}) = \sigma(\hat{\mathbf{A}} \text{ReLu}(\hat{\mathbf{A}} \mathbf{X} \mathbf{W}^0) \mathbf{W}^1) \quad (18)$$

其中, \mathbf{X} 为特征矩阵, 特征矩阵为资源分配矩阵, 即顶点染色矩阵; $\hat{\mathbf{A}} = \tilde{\mathbf{D}}^{-\frac{1}{2}} \tilde{\mathbf{A}} \tilde{\mathbf{D}}^{-\frac{1}{2}}$, $\tilde{\mathbf{A}} = \mathbf{A} + \mathbf{I}$, \mathbf{A} 表示一个图的邻接矩阵, \mathbf{I} 表示单位阵, $\tilde{\mathbf{D}}$ 表示一个对角阵, $\tilde{\mathbf{D}} = \sum_j \tilde{\mathbf{A}}_{ij}$; $\sigma(\cdot)$ 和 $\text{ReLu}(\cdot)$ 是激活函数; \mathbf{W}^k 表示第 k 层的权值矩阵, 其值会跟随网络更新。

对于一个图邻接矩阵 \mathbf{A} 和特征矩阵 \mathbf{X} , 在图卷积神经网络中, 根据以下传播规则更新每一层中的节点。

$$\mathbf{H}^{k+1} = \sigma(\tilde{\mathbf{D}}^{-\frac{1}{2}} \tilde{\mathbf{A}} \tilde{\mathbf{D}}^{-\frac{1}{2}} \mathbf{H}^k \mathbf{W}^k) \quad (19)$$

图 6 为图卷积神经网络结构, 在 actor 网络和 critic 网络中, 构建两层卷积层, 输入超图的邻接矩阵和特征矩阵, 然后输出节点特征, 根据节点特征进行资源分配和选取动作的评价。

D2D 辅助的 UD-IoT 基于 GCNN-DDPG 的超密集物联网资源分配算法流程如算法 1 所示。

算法 1 基于 GCNN-DDPG 的超密集物联网资源分配算法

输入 网络拓扑 $G(\mathbf{V}, \mathbf{E})$, 奖励折扣因子 γ , 经验回放大小 N_D , 目标网络更新频率 L , 学习率 α , 样本批采样大小 Γ

输出 超图点染色方案、网络吞吐量

1) 初始化训练次数 E , 初始化时间步 T 、经验池, 随机初始化 actor 网络参数 θ_{μ} 、critic 网

络参数 θ_Q

2) $\theta_Q = \theta_Q, \theta_{\mu} = \theta_{\mu}$

3) for $e = 1:E$ do

4) 初始化环境, 得到状态 s_1

5) for $t = 1:T$ do

6) 在 actor 网络中, 输入状态 s_t , 通过图卷积神经网络得到动作 $a_t = \mu(s_t | \theta_{\mu})$

7) 根据动作 a_t 获取下一个状态 s_{t+1} 和奖励值 r_t , 在 critic 网络中, 输入 s_{t+1} , 通过卷积神经网络得到动作 a_t 的评估值 Q

8) 将 (s_t, a_t, r_t, s_{t+1}) 存储在经验池中

9) 通过式(15)更新 critic 网络

10) 通过式(17)更新 actor 网络

11) 使用软更新更新目标网络

12) end for

13) end for

4 仿真结果与性能分析

实验采用的硬件平台为个人计算机, CPU 为 Intel(R) Xeon(R) Gold 6242R CPU @ 3.10GHz, GPU 为 NVIDIA RTX 3080Ti, 内存为 64 GB。本文在超密集物联网场景下进行仿真, 仿真实验参数如表 1 所示, 分别得到网络吞吐量和资源复用率性能的对比实验数据。将本文所提算法 (以下简称 GCNN-DDPG) 与其他 3 种算法进行网络吞吐量和资源复用率的比较, 3 种算法分别为基于随机匹配 (RM, random matching) 的资源分配算法^[21] (以下简称 RM)、基于贪婪匹配的资源分配算法^[21] (以下简称 GA) 和基于最大节点度 (MND, maximum node degree) 匹配的资源分配算法^[22] (以下简称 MND)。仿真实验时, 4 种资源分配算法都是在同一条件下完成的。仿真完成后, 通过网络吞吐量、资源复用率和信干噪比 3 项性能指标的对比分析, 反映 GCNN-DDPG 算法性能。

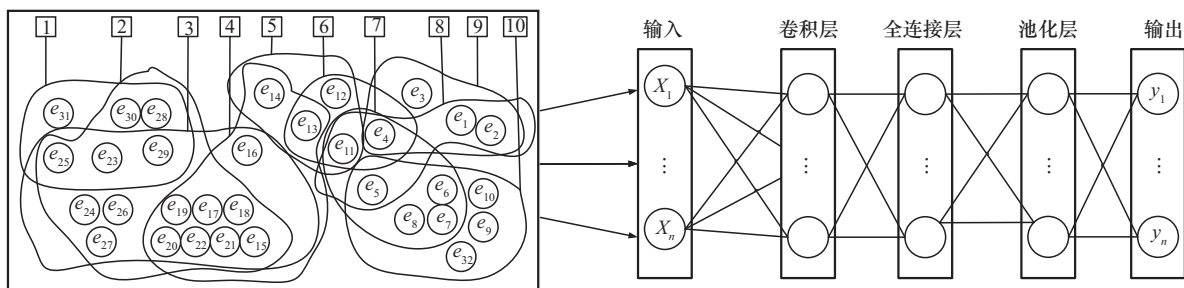


图 6 图卷积神经网络结构

表1 仿真实验参数

参数	数值
学习率	0.000 1
目标网络更新频率	20
经验回放采样数量	32
经验池容量	100
累积折扣因子	0.95
初始贪婪因子	0.9
最小贪婪因子	0.01
贪婪因子衰减率	99.5%
频谱带宽/MHz	1
最大可容忍误码率	70%
平均信干噪比/dB	20

4.1 UD-IoT 的性能评价指标

1) 网络吞吐量。该性能指标可评价资源分配算法分配完所有通信链路资源后 UD-IoT 的网络吞吐量^[23]，可表示为

$$\Gamma = Blb \left(1 + \frac{\bar{g}}{-\frac{2}{3} \ln \frac{P_b}{2}} \right) \quad (20)$$

其中， B 为通信链路的频谱带宽， \bar{g} 为平均信干噪比， P_b 为最大可容忍误码率。

2) 资源复用率。该性能指标可评价资源分配算法使 UD-IoT 所有通信链路无干扰后的通信资源复用率^[24]，可表示为

$$\lambda = \frac{M - \eta_*}{M} \quad (21)$$

其中， M 表示通信资源总数， η_* 为最终使用的通信资源数量。

3) 信噪比。系统中信号与干扰和噪声之和的比，如式(1)所示。

4.2 网络吞吐量性能对比

图7为UD-IoT最大网络吞吐量，其中最大吞吐量指系统在无干扰时的吞吐量。从图7可以看出，随着TL增大，网络吞吐量也增大。GCNN-DDPG相对于RM提高了27.2%的网络性能，相对于GA提高了37%的网络性能，相对于MND提高了19.2%的网络性能。这验证了本文所提算法可以有效地提高系统网络吞吐量。

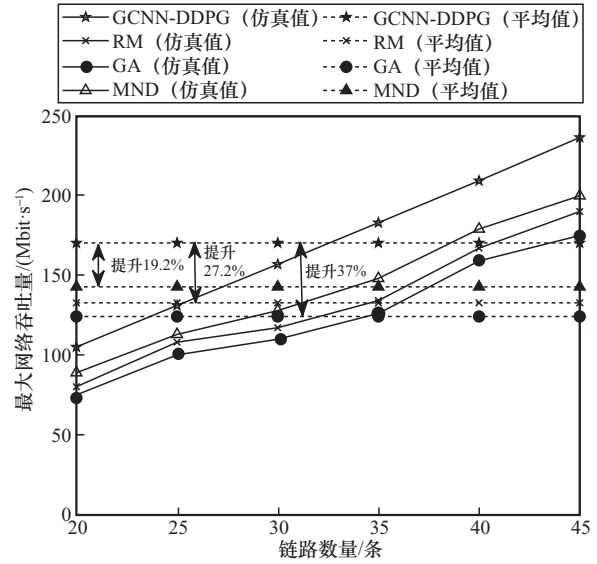


图7 UD-IoT最大网络吞吐量

图8为UD-IoT最小网络吞吐量，最小吞吐量指系统在遭受干扰时的网络吞吐量。从图8可以看出，随着TL的增加，GCNN-DDPG和3种对比算法得到的网络吞吐量呈现上升趋势。GCNN-DDPG得到的最小网络吞吐量明显高于RM和GA，相对于MND提高了9.8%的网络性能。这验证了本文所提算法可有效提高系统抗干扰能力。

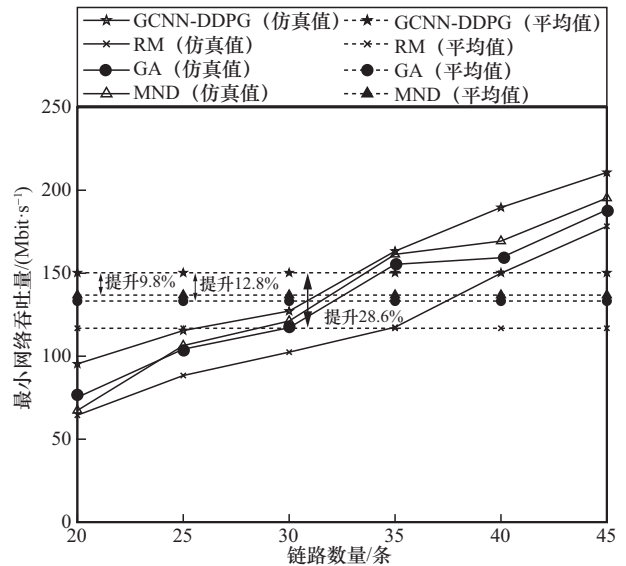


图8 UD-IoT最小网络吞吐量

4.3 资源复用率性能对比

图9为UD-IoT资源复用率，对于通信系统，资源复用率越高，系统的性能越好。从图9可以看出，GCNN-DDPG得到的资源复用率明显高于GA

和RM, 相对于MND提高了21%的资源复用率。这验证了本文所提算法可以有效提高资源复用率。

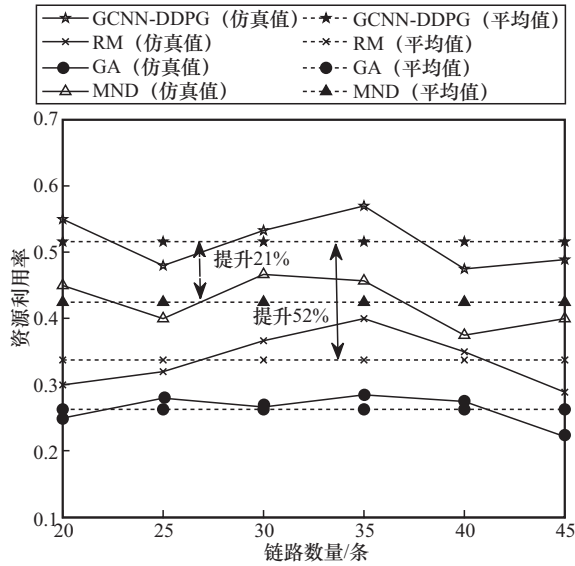


图9 UD-IoT资源复用率

4.4 信干噪比性能对比

图 10 为 UD-IoT 平均 SINR 对比, 对于通信系统来说, 具有较高 SINR 就意味着网络具有较低的干扰。从图 10 可以看出, GCNN-DDPG 比其他 3 种算法有更高的 SINR。这是因为本文所提算法通过优化资源分配, 提高了 SINR, 保证了传输速率的要求, 有效地避免了严重的隐藏终端干扰。

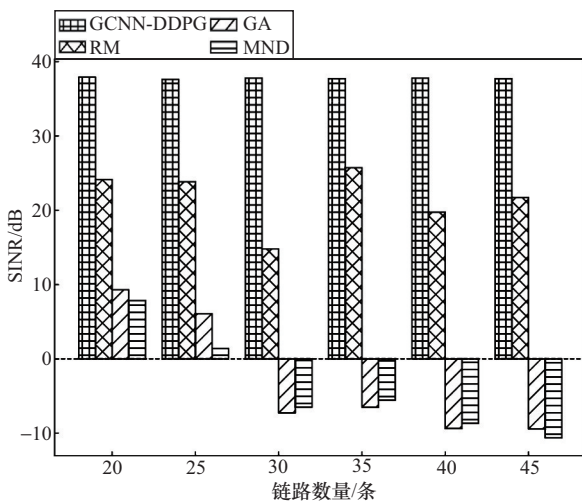


图10 平均SINR对比

5 结束语

针对超密集物联网部署的海量隐藏终端干扰下的资源分配问题, 本文提出了一种基于 GCNN-

DDPG的超密集物联网资源分配策略。为了同时分析多TL间的冲突关系, 本文首先提出了冲突超图模型, 用于分析在D2D辅助的UD-IoT的TL之间的资源冲突, 与冲突图模型相比, 冲突超图模型能够同时分析多TL间的冲突关系。然后, 将无冲突资源分配问题转化为超图的点强着色问题。最后, 提出了一种基于GCNN-DDPG的超密集物联网资源分配策略, GCNN-DDPG算法能够实现无冲突和最大化资源复用。实验结果表明, 与对比算法相比, 本文所提算法具有更高的网络资源重用率和网络吞吐量, 在D2D辅助的UD-IoT中具有更好的性能。

参考文献:

- [1] CHU Z, XIAO P, MI D, et al. IRS-assisted wireless powered IoT network with multiple resource blocks[J]. IEEE Transactions on Communications, 2023, 71(4): 2335-2350.
- [2] LI L X, CHENG Q Q, TANG X, et al. Resource allocation for NOMA-MEC systems in ultra-dense networks: a learning aided mean-field game approach[J]. IEEE Transactions on Wireless Communications, 2021, 20(3): 1487-1500.
- [3] AO S Y, NIU Y, HAN Z, et al. Resource allocation for RIS-assisted device-to-device communications in heterogeneous cellular networks[J]. IEEE Transactions on Vehicular Technology, 2023, 72(9): 11741-11755.
- [4] SARMA S S, HAZRA R, MUKHERJEE A. Symbiosis between D2D communication and industrial IoT for industry 5.0 in 5G mm-wave cellular network: an interference management approach[J]. IEEE Transactions on Industrial Informatics, 2022, 18(8): 5527-5536.
- [5] GBADAMOSI S A, HANCKE G P, ABU-MAHFOUZ A M. Interference avoidance resource allocation for D2D-enabled 5G narrowband Internet of things[J]. IEEE Internet of Things Journal, 2022, 9(22): 22752-22764.
- [6] KIM H M, NGUYEN H V, KANG G M, et al. SWIPT-aided device-to-device communications for massive IoT networks: a novel resource allocation with sparse code multiple access[J]. IEEE Internet of Things Journal, 2023, 10(22): 19617-19629.
- [7] YUAN Y L, YANG T, HU Y L, et al. Two-timescale resource allocation for cooperative D2D communication: a matching game approach[J]. IEEE Transactions on Vehicular Technology, 2021, 70(1): 543-557.
- [8] NAJLA M, BECVAR Z, MACH P. Reuse of multiple channels by multiple D2D pairs in dedicated mode: a game theoretic approach[J]. IEEE Transactions on Wireless Communications, 2021, 20(7): 4313-4327.
- [9] SANUSI I O, NASR K M, MOESSNER K. Radio resource management approaches for reliable device-to-device (D2D) communication in wireless industrial applications[J]. IEEE Transactions on Cognitive Communications and Networking, 2021, 7(3): 905-916.
- [10] YUAN Y Z, LI Z J, LIU Z X, et al. Double deep Q-network based distributed resource matching algorithm for D2D communication[J]. IEEE Transactions on Vehicular Technology, 2022, 71(1): 984-993.
- [11] SEID A M, ERBAD A, ABISHU H N, et al. Multiagent federated reinforcement learning for resource allocation in UAV-enabled Internet of medical things networks[J]. IEEE Internet of Things Journal, 2023, 10(22): 19695-19711.
- [12] SHI Z J, LIU J J. Massive access in 5G and beyond ultra-dense networks: an MARL-based NORA scheme[J]. IEEE Transactions on Communications, 2023, 71(4): 2170-2183.

[13] SHI D, LI L, OHTSUKI T, et al. Make smart decisions faster: deciding D2D resource allocation via stackelberg game guided multi-agent deep reinforcement learning[J]. IEEE Transactions on Mobile Computing, 2022, 21(12): 4426-4438.

[14] FAN Q, BAI J N, ZHANG H X, et al. Delay-aware resource allocation in fog-assisted IoT networks through reinforcement learning[J]. IEEE Internet of Things Journal, 2022, 9(7): 5189-5199.

[15] HUANG J F, YANG Y, GAO Z, et al. Dynamic spectrum access for D2D-enabled Internet of things: a deep reinforcement learning approach[J]. IEEE Internet of Things Journal, 2022, 9(18): 17793-17807.

[16] BUDHIRAJA I, KUMAR N, TYAGI S. Deep-reinforcement-learning-based proportional fair scheduling control scheme for underlay D2D communication[J]. IEEE Internet of Things Journal, 2021, 8(5): 3143-3156.

[17] PENG T, GUO Y C, WANG Y C, et al. An interference-oriented 5G radio resource allocation framework for ultradense networks[J]. IEEE Internet of Things Journal, 2022, 9(22): 22618-22630.

[18] SHAMAEI S, BAYAT S, HEMMATYAR A M A. Interference-aware resource allocation algorithm for D2D-enabled cellular networks using matching theory[J]. IEEE Transactions on Network and Service Management, 2024, 21(1): 759-772.

[19] HU J M, HENG W, ZHU Y P, et al. Overlapping coalition formation games for joint interference management and resource allocation in D2D communications[J]. IEEE Access, 2018, 6: 6341-6349.

[20] 汤小春, 周佳文, 田凯飞, 等. 大图中全部极大团的并行挖掘算法研究[J]. 计算机学报, 2019, 42(3): 513-531.
TANG X C, ZHOU J W, TIAN K F, et al. An efficient parallel algorithm for maximal clique enumeration in a large graph[J]. Chinese Journal of Computers, 2019, 42(3): 513-531.

[21] 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.

[22] YANG Y Y, XIE G. Mining maximum matchings of controllability of directed networks based on in-degree priority[C]//Proceedings of the 2016 35th Chinese Control Conference (CCC). Piscataway: IEEE Press, 2016: 1263-1267.

[23] FOSCHINI G J, SALZ J. Digital communications over fading radio channels[J]. The Bell System Technical Journal, 1983, 62(2): 429-456.

[24] 滑思忠, 丁爱玲, 郭达伟, 等. 基于二分图的D2D通信资源复用分配算法[J]. 计算机应用研究, 2017, 34(7): 2096-2098, 2103.
HUA S Z, DING A L, GUO D W, et al. Resource allocation algorithm in D2D cellular network based on bipartite graph[J]. Application Research of Computers, 2017, 34(7): 2096-2098, 2103.

[作者简介]



黄杰 (1988-), 男, 重庆人, 博士, 重庆理工大学副教授, 主要研究方向为无线通信理论、通信网络、下一代移动通信技术、认知无线电、无线通信资源分配。



李幸星 (2001-), 男, 重庆人, 重庆理工大学硕士生, 主要研究方向为无线通信理论、无线通信资源分配。



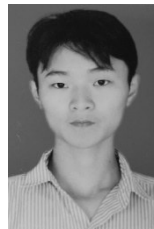
杨凡 (1983-), 男, 湖北广水人, 博士, 重庆理工大学副教授, 主要研究方向为无线带宽自适应传输、无线通信网络、下一代移动通信技术、无线通信中的编码技术。



丁睿杰 (2000-), 男, 重庆人, 重庆理工大学硕士生, 主要研究方向为无线带宽自适应传输、无线通信中的编码技术。



蔡杰良 (1998-), 男, 重庆人, 重庆理工大学硕士生, 主要研究方向为无线带宽自适应传输、无线通信中的编码技术。



姚凤航 (2000-), 男, 四川泸州人, 重庆理工大学硕士生, 主要研究方向为无线通信网络、下一代移动通信技术。



张鑫 (2000-), 男, 重庆人, 重庆理工大学硕士生, 主要研究方向为下一代移动通信技术、无线通信资源分配。