



专题：无人机空地无线通信技术

基于 5G 无人机通信的多智能体异构网络选择方法

丁雨¹, 李晨凯¹, 韩会梅¹, 卢为党¹, 任元红², 高原³, 曹江³

(1. 浙江工业大学信息工程学院, 浙江 杭州 310023;

2. 北方自动控制技术研究所, 山西 太原 030006;

3. 中国人民解放军军事科学院, 北京 100091)

摘要: 5G 无人机通信网络和各种不同无线接入技术的结合使无线异构网络呈现多样化的发展趋势。然而, 用户繁多且不同的业务请求对网络要求也不同, 造成网络接入选择问题。提出了一种基于 5G 无人机通信的多智能体异构网络选择方法, 将用户分为多个智能体, 从用户端和网络端两个方面出发, 将用户侧的时延和传输速率需求与网络侧的负载均衡需求综合考虑作为即时回报的相关参数, 通过基于 Nash Q -Learning 的算法进行学习, 得到异构网络环境下的网络选择决策模型。仿真结果表明, 所提异构网络选择方法针对不同业务类型用户的需求均能选择合适的网络, 同时均衡网络的负载, 充分利用异构无线网络的资源。

关键词: 无人机通信; 异构网络选择; Nash Q -Learning; 负载均衡

中图分类号: TN925

文献标志码: A

doi: 10.11959/j.issn.1000-0801.2022239

Multi-agent heterogeneous network selection method based on 5G UAV communication

DING Yu¹, LI Chenkai¹, HAN Huimei¹, LU Weidang¹, REN Yuanhong², GAO Yuan³, CAO Jiang³

1. College of Information Engineering, Zhejiang University of Technology, Hangzhou 310023, China

2. North Automatic Control Technology Institute, Taiyuan 030006, China

3. Academy of Military Science of the PLA, Beijing 100091, China.

Abstract: The combination of 5G UAV communication and different wireless access technologies makes wireless heterogeneous networks present a diversified development trend. However, various types of service requests by users have different requirements on the heterogeneous networks, which leads to the problem of network selection. A multi-agent heterogeneous network selection method based on 5G UAV communication was proposed. Specifically, the users were divided into multiple agents. From the two aspects of the user side and the network side, the delay and transmission rate requirements of the user side and the load balancing requirement of the network side were comprehensively considered as the relevant parameters of immediate reward. The simulation results show that the proposed heterogeneous network selection method can select the appropriate network according to the

收稿日期: 2022-06-02; 修回日期: 2022-08-11

通信作者: 卢为党, luweid@zjut.edu.cn

基金项目: 国家自然科学基金资助项目 (No.61871348, No.62222121)

Foundation Items: The National Natural Science Foundation of China (No. 61871348, No.62222121)

needs of users of different service types, balance the load of the network, and make full use of the resources of the heterogeneous wireless network.

Key words: UAV communication, heterogeneous network selection, Nash Q -Learning, load balancing

0 引言

无人机因高机动性和可控性被广泛应用于 5G 网络,不仅可以扩展通信覆盖范围,还可以通过高效移动改善链路通信质量,提高传输速率^[1-2]。随着 5G 无人机通信技术的快速发展,用户对请求的业务量显著增加,同时请求的业务类型也多种多样。

各种不同无线接入技术应用于 5G 无人机通信网络使无线网络具有独特的网络特性,适用于不同的通信场景。异构网络作为 5G 无人机通信网络的关键技术之一,要充分发挥其功效,首先要解决的是不同类型异构网络的融合问题。网络选择是异构网络融合的关键技术^[3],在异构网络环境中,用户需要根据其需求和偏好选择连接最佳网络,这就依靠一个灵活高效的异构网络选择算法。选择异构网络需要考虑不同接入网络之间的差异,因此需要考虑很多因素,从用户角度考虑,将涉及接收信号的强度、数据传输速率、覆盖范围、实时性和用户的移动性等。从网络角度考虑,将涉及系统的容量、网络的负载以及阻塞率等^[4-6]。

因此,一个合理高效的异构网络选择算法需要综合考虑各个方面的因素。在目前提出的异构网络选择算法中,由于研究和优化目标的不同,有多种针对性不同的算法。文献[7]从用户端考虑,选择用户吞吐量最大化的网络接入,或者根据接收信号强度进行异构网络的选择。从网络端考虑,有均衡网络负载、避免网络拥塞以及最小化通信成本等网络选择算法^[8-10]。这些算法通常只考虑了其中一两个因素,存在网络属性参数单一的问题,因此适用性不高。多属性决策算法对影响网络选择的多个属性参数进行分析,设计出一套对网络进行评价的策略来选

择最佳网络。文献[11]通过对多个属性进行加权,对网络进行排序。文献[12]利用层次分析法和熵权法计算候选网络的属性权重。文献[13]根据基于优劣解距离法进行网络选择。基于多属性决策的网络选择算法通过多种网络参数综合评价网络性能,但对如何准确描述属性参数,没有确切的标准,存在的问题^[14]。

基于强化学习的网络选择算法能够在一定程度弥补多属性决策算法的缺陷,得到较为精准的判决结果,将用户的历史数据作为输入,根据用户的反馈进行不断学习训练,得到一个满足用户需求的网络选择模型。并且,在训练过程中可以根据反馈实时优化网络选择模型,能更好地适应未来网络复杂多变的环境,从而帮助用户更加科学合理地选择网络接入^[15]。文献[16]提出了一种基于马尔可夫决策模型的网络选择优化算法,将异构无线网络的选择问题转化为优化问题,使得用户端能够在异构网络环境中始终连接到最佳网络。文献[17]在异构网络中提出了一种基于 Q -Learning 的选择策略,能够有效地缓解异构网络中的负载不平衡和能量不平衡问题。文献[18]提出基于 Q -Learning 的单智能体网络选择算法和基于 Nash Q -Learning 的多智能体网络选择算法,提升了网络负载均衡性能。

然而,现有的多智能体异构网络选择算法只将网络的可用容量比例作为学习的状态和回报的相关参数,限制了用户的性能。为了进一步提升用户的性能和异构网络资源的利用,本文提出了一种兼顾用户端和网络端性能的基于 5G 无人机通信的多智能体异构网络选择方法。本文的主要工作如下。

- 提出了一个由 5G 客户终端设备 (5G customer premise equipment, 5G CPE)、



5G 无人机和 Wi-Fi 3 种网络组成的异构通信网络, 其中, Wi-Fi 和 5G 网络部署在同一架无人机上, 5G CPE 部署在地面用户上, 随机分布的 M 个用户因业务请求需要进行网络选择。

- 提出了一种基于 5G 无人机通信的多智能体异构网络选择方法。考虑了用户和网络的移动性, 随着用户和网络的移动, 用户所处位置的网络覆盖情况也随之发生变化。将用户分为多个智能体, 从用户端和网络端两个方面出发, 将用户侧的时延和传输速率需求与网络侧的负载均衡需求综合考虑作为即时回报的相关参数。
- 通过基于 Nash Q -Learning 的算法进行学习, 训练分析得到异构网络环境的网络选择决策模型。仿真结果表明, 算法针对不同业务类型用户的需求均能选择合适的网络, 同时均衡网络的负载, 充分利用了异构无线网络的资源。

1 系统模型

异构网络系统模型如图 1 所示, 本文考虑了一个由 5G CPE、5G 无人机和 Wi-Fi 3 种网络组成的异构网络, 其中, Wi-Fi 和 5G 网络部署在同一架无人机上, 5G CPE 部署在地面用户身上。地面用户部署异构多模终端, 能够根据 5G CPE、5G 无人机和 Wi-Fi 3 种网络性能好坏, 针对不同业务类型需求选择合适的网络接入。网络模型中共有 M 个用户因业务请求需要进行网络选择, 请求的业务类型共有 N 种。用户可以分别和 5G CPE、5G 无人机和 Wi-Fi 接入点连接, 每个用户一次最多只能接入一个网络。5G CPE 需要与 5G 基站进行连接, 它将 5G 基站发出的 5G 信号转换成宽带/Wi-Fi 信号, 进一步提升地面用户用网体验。 M 个用户在异构网络中随机分布, 以一定的速度和方向朝着指定的目标区域移动。

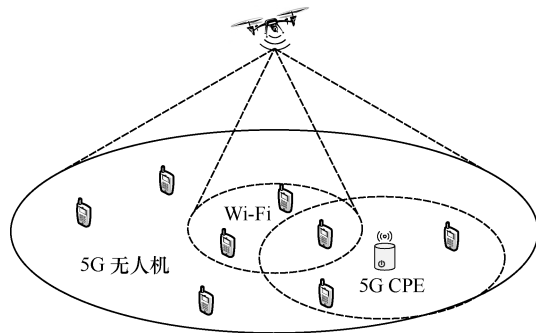


图 1 异构网络系统模型

各网络采用不同的无线接入技术, 其覆盖范围有限。由于受网络移动性的影响, 其覆盖区域时刻发生变化, 各个用户能被覆盖到的网络也在实时发生变化。用户因为业务请求需要进行网络连接, 根据当前时刻所在位置的网络覆盖情况选择网络接入。若用户不在任何网络的覆盖范围内, 则其不进行网络选择, 每个用户一次最多只能接入一个网络。此外, 每个网络设有一个能接入的最大负载值, 分别记为 $\{TL_1, TL_2, TL_3\}$, 当网络已接入的用户数达到其负载上限时, 其他用户将无法再接入该网络。

2 多智能体异构网络选择

2.1 多智能体异构网络选择模型

以用户请求的业务类型为标准, 将网络中具有相同业务类型的用户视为一个智能体, 即请求的业务类型有 N 种, 就有 N 个智能体。 N 个智能体中的用户随机排成一个队列, 在每一次学习过程中, 选择队列的第一个用户执行动作, 完成后即从队列中移除。当队列为空时, 表示完成了一次学习。多智能体异构网络选择时使用的 Nash Q -Learning 算法的状态、动作和回报函数设置如下。

(1) 状态 s

用网络负载表示网络的实时状态, 假设 5G CPE 网络、5G 无人机网络和 Wi-Fi 网络的负载分别为 L_1 、 L_2 和 L_3 , 则 3 个网络的负载占用率分别为 $\frac{L_1}{TL_1}$ 、 $\frac{L_2}{TL_2}$ 和 $\frac{L_3}{TL_3}$ 。由于网络的负载占用率是连续的, 将其离散化使其能够作为 Nash

Q -Learning 的状态。为将状态空间减少到有限集合, 将网络负载占用率映射到一组量化值, 分为 4 个等级, 异构网络系统模型见表 1。可以看出, 负载占用率越高, 区间划分越细致, 这样可以使算法在负载占用率趋于饱和时做出更为细致的判断。

表 1 异构网络系统模型

负载占用率	量化等级
[0,0.5)	1
[0.5,0.75)	2
[0.75,0.875)	3
[0.875,1]	4

将状态空间 s 定义为:

$$s = \{l_1, l_2, l_3\}, l_1, l_2, l_3 \in \{1, 2, 3, 4\} \quad (1)$$

其中, l_1 、 l_2 和 l_3 分别表示 5G CPE 网络、5G 无人机网络和 Wi-Fi 网络负载占用率量化后的数值。

(2) 动作 a

定义动作空间 A 表示动作集合

$$A = \{a_i, a_i \in \{1, 2, 3\}, i \in \{1, 2, \dots, N\}\} \quad (2)$$

其中, a_i 表示智能体 i 选择的动作, $a_i = 1$ 表示智能体 i 选择接入 5G CPE 网络, $a_i = 2$ 表示智能体 i 选择接入 5G 无人机网络, $a_i = 3$ 表示智能体 i 选择接入 Wi-Fi 网络。

(3) 回报函数 r

为了判断并学习所选取动作的优劣, 设置回报函数计算执行某动作后得到的即时回报, 并将其作为历史经验进行学习。本文考虑的评价指标包括网络的时延和传输速率, 以及网络的均衡性。为了使不同业务的回报函数能体现网络负载情况对动作选择的影响, 实现网络负载的均衡, 本文在回报函数中加入了一项负载占用率方差 F , 衡量 3 个网络之间的负载均衡性。由于时延、速率和负载占用率方差之间的数量级不同, 且这些属性中有些属于效益型属性, 即它们的值越大表示性能越好, 而有些属于成本型属性, 即它们的值越小表示性能越好。因此, 本文对时延 D 、传输速率 E 和负载占用率方差 F 进行规范化处理, 且效益型属性和成本型属性

的规范化公式不同, 分别表示为:

$$y_i^{\text{benefit}} = \frac{x_i - \min(x)}{\max(x) - \min(x)} \quad (3)$$

$$y_i^{\text{cost}} = \frac{\max(x) - x_i}{\max(x) - \min(x)} \quad (4)$$

在这 3 个属性中, 传输速率 E 是效益型属性, 时延 D 和负载占用率方差 F 是成本型属性。 d 、 e 、 f 分别表示属性规范化处理后的时延、速率和负载占用率方差。因此, 回报函数定义为:

$$R_i = \omega_i d + \theta_i e + \mu f, i \in \{1, 2, \dots, N\} \quad (5)$$

其中, ω_i 和 θ_i 分别表示业务类型 i 对时延和传输速率的偏好, μ 表示系统对网络负载均衡的偏好。

强化学习算法通过智能体与环境的交互进行策略学习和优化, 在 Nash Q -Learning 算法中, 与环境进行交互的是多个智能体的联合动作。在该异构网络系统中, 所有用户按照业务类型划分为 N 个智能体, 单一时刻中, 每个智能体中均有一个用户执行网络选择动作, N 个智能体不分先后在同一时刻做出各自的动作, 这一时刻的联合动作定义为 $(a_1, a_2, \dots, a_N)$, 在多智能体场景下的最优策略是实现智能体整体策略的 Nash 均衡。Nash 均衡策略可表示为 $(\pi_1^*, \pi_2^*, \dots, \pi_N^*)$, 满足:

$$V_i(s', \pi_1^*, \pi_2^*, \dots, \pi_N^*) \geq V_i(s', \pi_1, \pi_2, \dots, \pi_N) \quad (6)$$

其中, V 表示状态—动作对的价值函数, $(\pi_1, \pi_2, \dots, \pi_N)$ 表示所有可能的联合动作。

2.2 多智能体异构网络选择算法

本节提出了一种多智能体异构网络选择算法。

首先, M 个用户根据业务类型的不同, 组成 N 个智能体, 每个智能体中的用户随机排成队列。从初始状态 s_0 开始, N 个智能体根据贪婪算法执行各自的动作组成联合动作 $(a_1, a_2, \dots, a_N)$ 。联合动作执行后, 每个智能体得到环境反馈获得各自的收益, 计算网络的下一个状态, 更新 Q 表及环境到达转移状态 s' , 随机博弈转移到下一个博弈阶段。在状态 s 下, 智能体重复执行联合动作的过程, 继续博弈。在开始下一次学习之前, 重置网络状态和服务队列。



由于多智能体系统中，代表收益的状态动作值函数不仅与自身所采取的动作有关，还与同一时刻同一状态下其他智能体的动作有关，因此将状态动作值定义为 $Q(s, a_1, a_2, \dots, a_N)$ ，表示智能体在状态 s 下执行动作 $(a_1, a_2, \dots, a_N)$ 后的即时回报与所有可能转移状态 s' 下期望收益的和。因此，智能体 i 的 Nash Q 函数可以表示为：

$$Q_i^*(s, a_1, a_2, \dots, a_N) = R_i(s, a_1, a_2, \dots, a_N) + \beta \sum_{s' \in S} p(s' | s, a_1, a_2, \dots, a_N) V_i(s', \pi_1^*, \pi_2^*, \dots, \pi_N^*) \quad (7)$$

其中， $R_i(s, a_1, a_2, \dots, a_N)$ 表示智能体 i 在状态 s 下执行动作 $(a_1, a_2, \dots, a_N)$ 后获得的即时回报值， β 表示折扣因子， s' 表示智能体 i 执行动作 $(a_1, a_2, \dots, a_N)$ 后可能的转移状态， p 表示对应的状态转移概率， $(\pi_1^*, \pi_2^*, \dots, \pi_N^*)$ 表示状态 s' 下满足 Nash 均衡的多智能体联合动作， $V_i(s', \pi_1^*, \pi_2^*, \dots, \pi_N^*)$ 表示对应的智能体 i 在状态 s' 下的期望收益。因此，式 (7) 表示智能体 i 在状态 s 下执行动作 $(a_1, a_2, \dots, a_N)$ 后的即时回报与其在执行动作后对应的所有可能转移状态 s' 下执行满足 Nash 均衡的联合动作的期望收益和。

Nash Q -Learning 算法通过多智能体与环境的不断交互，根据环境的反馈不断更新对应智能体的状态动作值，即 Q 值。在时刻 t ，各个智能体在观察当前环境的状态 s 后，根据状态执行动作 $(a_1, a_2, \dots, a_N)$ 。当所有的智能体都执行完动作后，环境的状态 s 变为其转移状态 s' 。为了考虑 Nash 均衡回报，联合动作执行后，各个智能体需要观察其他智能体采取的行动。因此，每个智能体都存有 N 个 Q 值表，每个 Q 值表对应当前系统中的一个智能体。根据观察到的联合动作，回报以及环境的转移状态 s' ，各个智能体根据式 (8) 更新 Q 值。

$$Q_{i,t+1}^j(s, a_1, \dots, a_N) = (1 - \alpha) Q_{i,t}^j(s, a_1, \dots, a_N) + \alpha [R_i^j(s, a_1, \dots, a_N) + \beta \text{Nash} Q_{i,t}^j(s')] \quad (8)$$

其中，

$$\text{Nash} Q_{i,t}^j(s') = \pi_1(s') \cdots \pi_N(s') \cdot Q_{i,t}^j(s') \quad (9)$$

智能体 i 在状态 s' 下执行动作 $\pi_i, i=1, 2, \dots, N$ ， N 个智能体的动作 $\pi_1(s') \cdots \pi_N(s')$ 为在状态 s' 下满足 Nash 均衡的联合动作。 $\text{Nash} Q_{i,t}^j(s')$ 为 t 时刻智能体 i 在转移状态 s' 下执行满足 Nash 均衡的联合动作后的收益。基于 Nash Q -Learning 的多智能体网络选择见算法 1。

算法 1 基于 Nash Q -Learning 的多智能体网络选择

初始化： 学习率 α ，折扣因子 β ，贪婪策略探索概率 ε 和网络初始状态 $s_0 = \{l_1, l_2, l_3\}$ 。

当前学习智能体的索引表示为 i ，在每个智能体中创建 N 个 Q 值表，记录系统中所有智能体的 Q 值，且各个 Q 值表中的初始值均为 0

for $a=1:\text{episode}$ do

智能体 i 中的用户随机排成一个队列，队列长度为 L_i ，并计算初始状态 s

while $(L_1 \neq 0 \& L_2 \neq 0 \& \cdots L_N \neq 0)$ do

for $i=1:N$ do

智能体 i 队列中的第一个用户判断自身所处位置的网络覆盖情况，根据贪婪策略选择一个动作执行，并判断该动作是否超出网络负载上限。若超出上限则重新选择

end for

for $i=1:N$ do

执行选出的联合策略 $a_1, a_2, \dots, a_N$ ，

智能体 i 得到回报

end for

计算下一个状态 s'

for $i=1:N$ do

智能体 i 根据式 (8) 更新 Q 值表

end for

系统从 s 转移到 s' 状态，同时移除所有智能体队列中的第一个用户

end while

随机重置队列

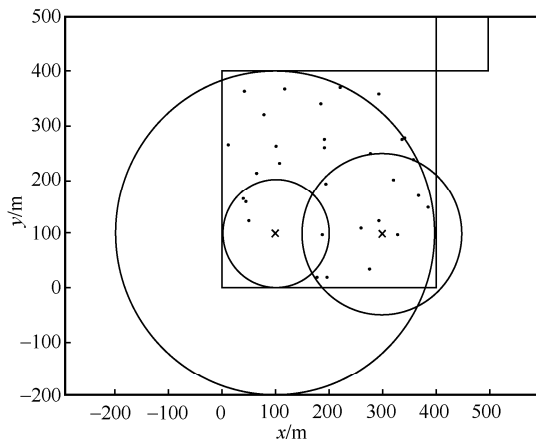
end for

输出： 智能体的 Q 值表

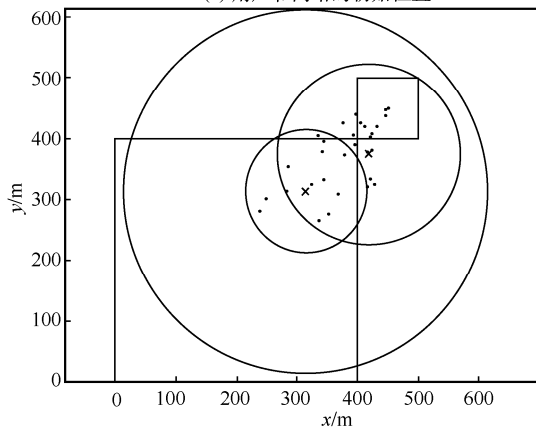
3 仿真分析

3.1 参数设置

用户和网络的位置分布如图2所示,图2(a)显示的是用户和网络的初始位置,用户随机均匀分布在 $400\text{ m} \times 400\text{ m}$ 的正方形区域内,该区域左下角顶点坐标为 $(0,0)$,目标区域为一个 $100\text{ m} \times 100\text{ m}$ 的正方形区域,该区域左下角顶点坐标为 $(400,400)$ 。 M 个用户朝着目标区域的方向以 $3 \sim 6\text{ m/s}$ 的速度移动,搭载5G网络和Wi-Fi网络的无人机的初始坐标为 $(100,100)$,携带5G CPE的用户初始坐标为 $(300,100)$ 。5G CPE、5G无人机和Wi-Fi 3个网络的覆盖半径分别为 150 m 、 300 m 和 100 m 。3个网络中心的移动速度均为 $3 \sim 6\text{ m/s}$ 。图2(b)显示了用户和网络移动后某一时刻的位置。



(a) 用户和网络的初始位置



(b) 用户和网络移动后某一时刻的位置

图2 用户和网络的位置分布

假设每个网络的负载上限 $\{TL_1, TL_2, TL_3\}$ 为 $\{25, 25, 25\}$ 。学习率 α 设为0.1,折扣因子 β 设为0.8,探索率 ε 设为0.4。用户请求的业务类型数 N 为3,分别是语音类业务、视频类业务和文本类业务。假设每种业务类型的用户数相等,均为 $M/3$ 。其中,资源需求包括负载占用率方差、时延及传输速率。3种业务对各个属性的偏好值如图3所示。资源需求与负载之间的关联是在每次连接时占用一个负载,同时每个网络可接入的负载上限一样,通过占上限的比例计算负载占用率。根据图3中的偏好值可算出对应的权重值,3种网络的参数范围见表2。

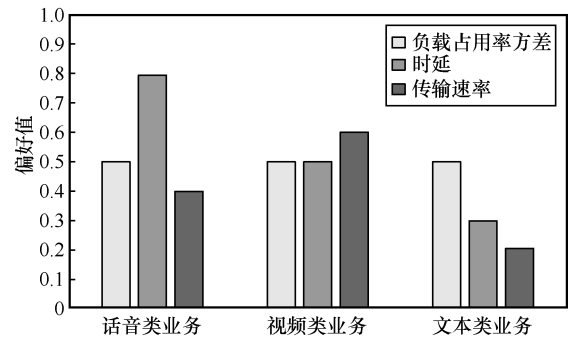


图3 3种业务对各个属性的偏好值

表2 3种网络的参数范围^[1,16,19]

网络	时延/ms	传输速率/(Gbit·s ⁻¹)
5G CPE	20~45	1.2~1.4
5G 无人机	10~30	1.4~1.6
Wi-Fi	40~70	1.7~1.9

3.2 仿真结果

当用户数为30时,学习过程中的负载占用率变化如图4所示。一开始,由于用户的分布较为分散,能同时被3个网络覆盖的用户比较少,3个网络接入用户数很大程度上取决于各个网络的覆盖范围。5G无人机网络的覆盖范围最大,其次是5G CPE,Wi-Fi的覆盖范围最小,因此在一开始,选择5G无人机的用户最多,选择Wi-Fi的用户最少。随着用户和网络的移动,它们的位



置逐渐聚合,直至最后大部分用户都可以同时被 3 个网络覆盖。经过学习,3 个网络的负载率逐渐收敛。5G 无人机的负载占用率逐渐减少,但最终仍为负载占用率最高的网络。覆盖范围最小的 Wi-Fi 的负载占用率逐渐升高,其负载占用率最终超过 5G CPE。

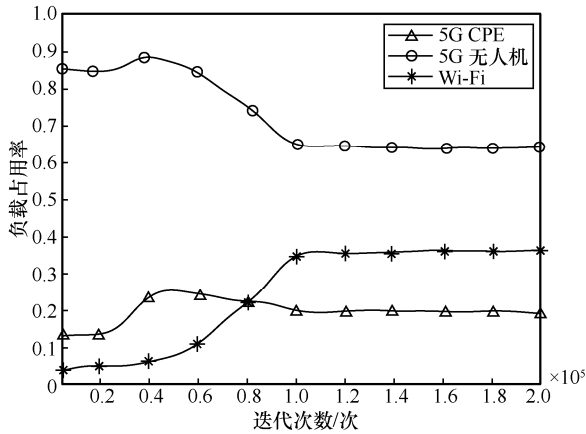
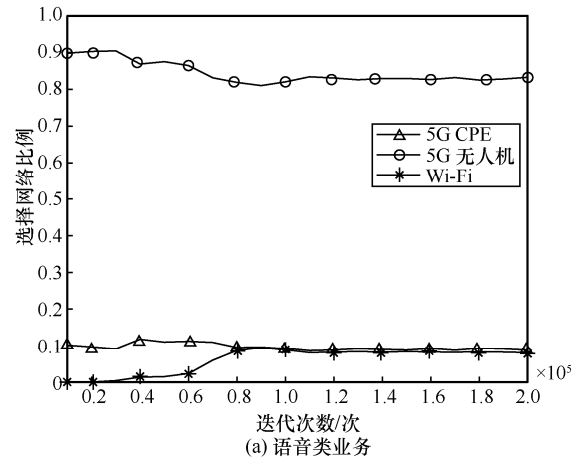


图4 学习过程中的负载占用率变化

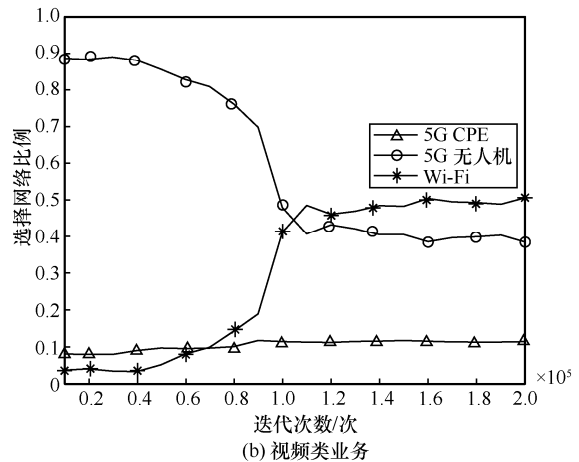
当用户数为 60 时,不同业务类型选择 3 种网络的比例变化如图 5 所示。随着用户和网络的移动,用户和网络逐渐聚合,最终大部分用户都可以被 3 个网络同时覆盖。此时不同业务的网络选择比例开始产生区别,如图 5 (a) 所示,对时延要求较高的语音类业务用户更多地选择接入时延最短的 5G 无人机网络,而从图 5 (b) 中可以看出对传输速率要求较高的视频类业务用户则更多地接入了传输速率最大的 Wi-Fi 网络。从图 5 (c) 中可以看出文本类业务对时延的偏好略高于传输速率,但都比语音类业务的偏好值低,因此文本类业务选择 5G 无人机网络的用户更多,但选择比例低于语音类业务相应的 5G 无人机网络。

当用户数为 30 时,不同负载占用率方差偏好值下的负载占用率如图 6 所示。对负载占用率方差的偏好值越大,说明对负载均衡的要求越高,其对应的权重越高。因此,在计算执行动作后的回报时,使得负载占用率变均衡的动作得到的回

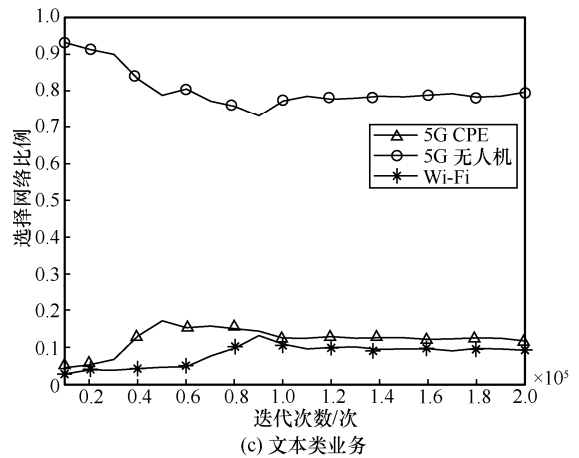
报值会更高,在之后的选择中该动作更有可能被选中。可以看出,随着负载占用率方差偏好值的增大,负载占用率最高的 5G 无人机网络的占用率逐渐下降,5G CPE 网络的负载占用率逐渐增加,3 个网络的负载占用率逐渐接近,负载均衡的效果逐渐提升。



(a) 语音类业务



(b) 视频类业务



(c) 文本类业务

图5 不同业务类型选择 3 种网络的比例变化

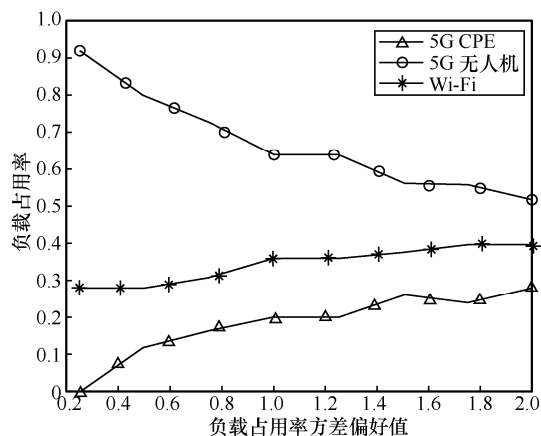


图6 不同负载占用率方差偏好值下的负载占用率

4 结束语

本文提出了一种兼顾用户端和网络端性能的基于 5G 无人机通信的多智能体异构网络选择算法,建立了用户和网络中心的移动模型和异构网络选择系统模型,以网络的负载占用率作为 Nash Q -Learning 的状态特征量,改进回报函数,将网络侧的负载均衡需求和用户侧的时延和传输速率需求综合考虑作为即时回报的相关参数。通过基于 Nash Q -Learning 的算法的学习,得到异构网络环境下的网络选择决策模型。仿真结果表明,所提异构网络选择方法针对不同业务类型用户的需求均能选择合适的网络,同时均衡网络的负载,充分利用异构无线网络的资源。

参考文献:

- [1] 贺子健, 艾元, 闫实, 等. 无人机通信网络的容量与覆盖性能[J]. 电信科学, 2017, 33(10): 65-70.
HE Z J, AI Y, YAN S, et al. Performance evaluation of capacity and coverage in UAV communication networks[J]. Telecommunications Science, 2017, 33(10): 65-70.
- [2] LU W D, DING Y, GAO Y, et al. Secure NOMA-based UAV-MEC network towards a flying eavesdropper[J]. IEEE Transactions on Communications, 2022, 70(5): 3364-3376.
- [3] 张祖凡, 张宇. 基于 3D-Markov 链的异构网络用户 QoE 分析方法[J]. 中国科学: 信息科学, 2016, 46(1): 85-99.
ZHANG Z F, ZHANG Y. QoE analysis method for heterogeneous network user based on 3D-Markov chain[J]. Scientia Sinica

- (Informationis), 2016, 46(1): 85-99.
- [4] 刘军, 李晓楠. 基于层次分析法的 WLAN/蜂窝网络切换判决算法[J]. 通信学报, 2013, 34(2): 65-72.
LIU J, LI X N. Handover algorithm for WLAN/cellular networks with analytic hierarchy process[J]. Journal on Communications, 2013, 34(2): 65-72.
- [5] 盛洁, 马冬. 异构无线网络业务接入多目标优化控制算法[J]. 电子学报, 2016, 44(2): 282-288.
SHENG J, MA D. Multi-objective optimal based service access control algorithm for heterogeneous wireless networks[J]. Acta Electronica Sinica, 2016, 44(2): 282-288.
- [6] ZHOU W A, XU Y J, LIANG T, et al. Optimal power allocation for heterogeneous network capacity maximization[J]. China Communications, 2013, 10(9): 97-103.
- [7] CHANG B J, CHEN J F. Cross-layer-based adaptive vertical handoff with predictive RSS in heterogeneous wireless networks[J]. IEEE Transactions on Vehicular Technology, 2008, 57(6): 3679-3692.
- [8] 石文孝, 范绍帅, 王楠, 等. 基于模糊神经网络的异构无线网络接入选择算法[J]. 通信学报, 2010, 31(9): 151-156.
SHI W X, FAN S S, WANG N, et al. Fuzzy neural network based access selection algorithm in heterogeneous wireless networks[J]. Journal on Communications, 2010, 31(9): 151-156.
- [9] MIAO J, HU Z, YANG K, et al. Joint power and bandwidth allocation algorithm with QoS support in heterogeneous wireless networks[J]. IEEE Communications Letters, 2012, 16(4): 479-481.
- [10] CHEN H N, WU H Y, KUMAR S, et al. Minimum-cost data delivery in heterogeneous wireless networks[J]. IEEE Transactions on Vehicular Technology, 2007, 56(6): 3511-3523.
- [11] LIU S M, PAN S, MI Z K, et al. A simple additive weighting vertical handoff algorithm based on SINR and AHP for heterogeneous wireless networks[J]. 2010 International Conference on Intelligent Computation Technology and Automation, 2010, 1: 347-350.
- [12] 魏淑芝, 朱琦. 基于网络选择的视频通信带宽博弈算法[J]. 通信学报, 2015, 36(2): 216-224.
WEI S Z, ZHU Q. Bandwidth allocation games based on network selection for video communication[J]. Journal on Communications, 2015, 36(2): 216-224.
- [13] CHAMODRAKAS I, MARTAKOS D. A utility-based fuzzy TOPSIS method for energy efficient network selection in heterogeneous wireless networks[J]. Applied Soft Computing, 2011, 11(4): 3734-3743.
- [14] GOYAL R K, KAUSHAL S, SANGAIAH A K. The utility based non-linear fuzzy AHP optimization model for network



selection in heterogeneous wireless networks[J]. Applied Soft Computing, 2018, 67: 800-811.

- [15] 赵彦清, 朱琦. 基于 Q 学习的异构网络选择新算法[J]. 计算机应用, 2011, 31(6): 1461-1464, 1479.

ZHAO Y Q, ZHU Q. New Q-learning based heterogeneous network selection algorithm[J]. Journal of Computer Applications, 2011, 31(6): 1461-1464, 1479.

- [16] XIE J L, GAO W J, LI C R. Heterogeneous network selection optimization algorithm based on a Markov decision model[J]. China Communications, 2020, 17(2): 40-53.

- [17] CHEN J L, YIN M X, DUAN X H, et al. Q-learning based selection strategies for load balance and energy balance in heterogeneous networks[C]//Proceedings of 2020 5th International Conference on Computer and Communication Systems (ICCCS). Piscataway: IEEE Press, 2020: 728-732.

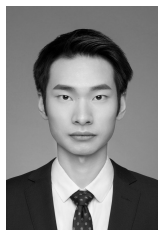
- [18] WANG X Q, SU X, LIU B. A novel network selection approach in 5G heterogeneous networks using Q-learning[C]//Proceedings of 2019 26th International Conference on Telecommunications (ICT). Piscataway: IEEE Press, 2019: 309-313.

- [19] YANG J, FENG L, JIN T. Enabling sector scheduling for 5G-CPE dense networks[J]. Security and Communication Networks, 2020(7): 1-18.

[作者简介]



丁雨(1998-), 女, 浙江工业大学信息工程学院硕士生, 主要研究方向为无人机通信、智能通信。



李晨凯(1996-), 男, 浙江工业大学信息工程学院硕士生, 主要研究方向为异构网络接入、智能通信。



韩会梅(1990-), 女, 浙江工业大学信息工程学院讲师, 主要研究方向为无人机通信、大规模接入。



卢为党(1984-), 男, 浙江工业大学信息工程学院教授, 主要研究方向为无人机通信、移动边缘计算。



任元红(1984-), 男, 北方自动控制技术研究所高级工程师, 主要研究方向为指控系统设计。



高原(1986-), 男, 中国人民解放军军事科学院副研究员, 主要研究方向为无人机通信。



曹江(1960-), 男, 中国人民解放军军事科学院研究员, 主要研究方向为指挥信息系统理论。