

DOI:10.19651/j.cnki.emt.2416044

基于 Dueling-DQN 的异构无线网络 垂直切换算法研究 *

李敏之 李转怀

(兰州交通大学电子与信息工程学院 兰州 730070)

摘要: 针对当前异构无线网络中切换算法考虑的服务质量指标较少,用户频繁切换愈加严重的问题,提出了基于主观加权与改进的深度强化学习相结合的异构无线网络垂直切换方法。首先,提出了一种支持异构无线网络的软件定义网络架构;其次,提出了主客观加权相结合的属性加权算法;最后,将网络选择问题利用改进的 Dueling-DQN 方法解决。仿真结果表明,本文所提算法在不同用户类型网络下切换次数分别减少了 11.25%, 13.34%, 18.76%, 13.75%, 吞吐量提升了 16.64%。因此本文所提算法有效避免了乒乓切换,减少切换次数并且提升了吞吐量。

关键词: 异构无线网络;软件定义网络;层次分析法;熵值法;深度 Q 学习

中图分类号: TN929.5 文献标识码: A 国家标准学科分类代码: 510.5015

Research on vertical handover algorithm of heterogeneous wireless network based on Dueling-DQN

Li Minzhi Li Zhuanhuai

(School of Electronic and Information Engineering, Lanzhou Jiaotong University, Lanzhou 730070, China)

Abstract: In view of the fact that the network selection algorithm in the heterogeneous wireless network has few quality of service indicators, and the frequent switching of users is becoming more and more serious, In this paper, a vertical handover method for heterogeneous wireless networks based on subjective and objective weighting combined with improved deep reinforcement learning is proposed. Firstly, a software-defined network architecture supporting heterogeneous wireless networks was proposed; secondly, an attribute weighting algorithm combining subjective and objective weighting was proposed; finally, the network selection problem is solved by using Dueling-DQN. The simulation results show that the proposed algorithm reduces the number of switching times by 11.25%, 13.34%, 18.76% and 13.75% respectively under different user types of networks, and increases the throughput by 16.64%. Therefore, the algorithm proposed in this paper effectively avoids ping-pong switching, reduces the number of switching times and improves the throughput.

Keywords: heterogeneous wireless network; software defined network; analytic hierarchy process; entropy method; deep Q-learning

0 引言

在新一代移动通信系统中,多网融合是未来移动通信发展的一个重要方向,而异构网络的选择是新一代移动通信系统的核心技术,具有重要的理论和现实意义。如何在不同业务类型下准确地选取合适的候选网络,减少切换次数,是值得深入研究的问题。

现阶段对垂直切换技术的研究仍然面临着诸多挑战。

垂直切换是指用户可以在不同的无线网络之间切换以满足其需求,为终端提供无缝移动性和会话连续性的过程。基于异构无线网络垂直切换主要分为以下几种:基于接收信号强度(really simple syndication, RSS)的网络选择算法,提出基于用户通过测量各个候选网络的 RSS,如果候选网络中的 RSS 大于当前所在网络,则进行切换,但是所利用的数学模型相对简单,并且复杂度^[1];基于多属性决策(multiple attribute decision making, MADM)的网络选择算法,提出了

一种 MADM 和灰色关联度(grey relation analysis, GRA)相结合的垂直切换算法,利用用户的偏好程度和网络排序对不同业务下的流量进行分析,最终有效减少了切换次数和排序异常,但计算复杂度较大且未对乒乓效应进行分析^[2];基于模糊逻辑的网络选择算法,提出了模糊逻辑和逼近理想解排序(technique for order preference by similarity to an ideal solution, TOPSIS)相结合的垂直切换算法^[3],引入信噪比(signal to interference plus noise ratio, SINR)和用户停留时间^[4]提高了决策的准确化,但没有考虑到用户的满意程度;基于马尔科夫决策过程(markov decision process, MDP)的网络选择算法,目标是最大化每个连接网络的期望回报值,采用层次分析法(the analytic hierarchy process, AHP)为每个网络属性分配权重,但是算法并没有区分业务类型^[5];基于智能优化决策的网络选择算法,提出了一种基于遗传禁忌算法优化的模糊神经网络选择算法^[6],对带宽,RSS,用户的移动速度和负载进行模糊化处理并且加入了预判机制,能够提供相对较优的服务,提出了基于神经网络的垂直切换算法^[7],结合 MADM 选择具有最高网络下载速率的无线网络作为预切换网络,再结合神经网络进行切换判决,该算法提高了切换成功率,但收敛程度较高。提出了深度 Q 学习(deep Q-learning, DQN)的垂直切换算法^[8],结合进化策略使得切换变得简单有效且提高了参数学习的收敛速度^[9],融合了参考信号接收功率和切换次数两项指标作为联合奖励函数^[10],减少了无人机在蜂窝基站之间的无效切换次数。基于上述问题,本文提出了一种在(software defined network, SDN)架构下基于 MADM 和改进的 Dueling-DQN 垂直切换方法,使得多个终端可以根据自身的偏好,自主地进行最优切换。

1 SDN 架构

SDN 是一种新兴的网络架构,其核心是通过 OpenFlow 协议实现控制和转发的分离并对网络资源进行合理调配,实现对网络性能的优化^[11]。网络设备只需关注分组的转发与分离,无需耗费任何资源进行分组转发路径的计算或更新,从而加快了分组的处理和转发速度。

如图 1 所示,基于控制层提供的北向 API 可根据需求开发业务应用,而控制层和数据层之间的通信主要通过 OpenFlow 协议实现^[12]。因为深度强化学习模型可以适应不同类型的输入数据,因此,基于 SDN 的架构可以扩展到异构无线网络的切换场景。在 SDN 体系结构下,数据层将收集的信息进行归一化处理,生成候选网络,在控制层对候选网络的优先级进行优化,这也是本文的重点。根据优化结果作出切换决策,从而完成整个切换的过程。

2 主客观加权

2.1 建立 MADM 矩阵

异构无线网络的垂直切换是一个比较复杂的问题,它

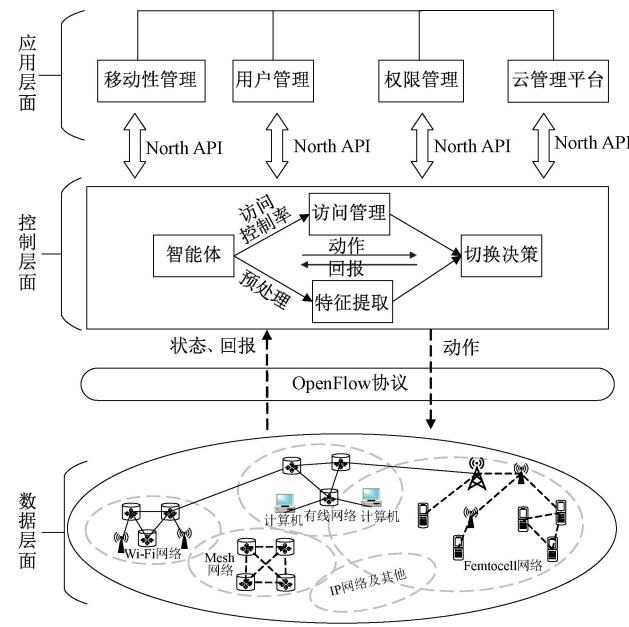


图 1 SDN 体系结构

Fig. 1 SDN architecture

既包含了网络端的性能指标,也包含了终端的特性和用户的偏好。垂直切换 MADM 的第一步是构建决策矩阵,决策矩阵 D 由式(1)表示, M 表示网络的数量, N 表示属性的数量。

$$D = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \quad (1)$$

2.2 主观权重的计算

AHP 是一种强大而有效的主观加权方法,由于其计算复杂度低、计算能力强的优点,在 MADM 中被广泛应用^[13]。AHP 被用于计算网络属性的主观权重^[14],使用 AHP 算法的权重向量计算过程如下所示:

1) 建立层次结构模型:如图 2 所示,该层次结构模型由 4 层组成。

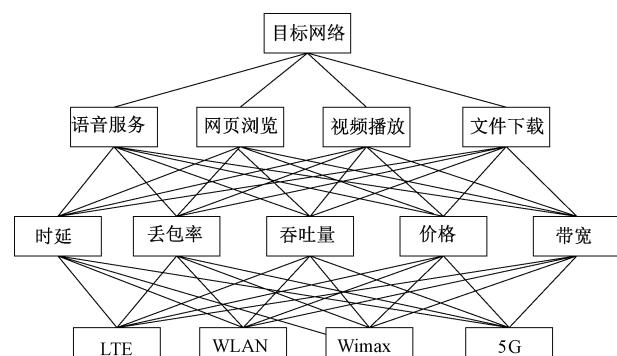


图 2 层次结构模型

Fig. 2 Hierarchical model

2) 构造成对比较矩阵:

成对比较是指本层因素相对于上一层不同业务类型的重要性。如图表 1 中 1~9 标度规则并且结合业务类型特点及用户的主观偏好, 式(2)~(5)分别构建了属性指示层相对于上层业务的成对比较矩阵。

表 1 标度法规则表^[15]

Table 1 Scaling rules

标度	重要性程度
1	i 与 j 同等重要
3	i 略比 j 重要
5	i 明显重要
7	i 很重要
9	i 非常重要
2,4,6,8	上述程度的中间值

$$\mathbf{P}^1 = \begin{bmatrix} 1 & 2 & 5 & 4 & 1 \\ 1/2 & 1 & 3 & 3 & 2 \\ 1/5 & 1/3 & 1 & 3 & 1/6 \\ 1/4 & 1/3 & 1/3 & 1 & 1/4 \\ 1 & 1/2 & 6 & 4 & 1 \end{bmatrix} \quad (2)$$

$$\mathbf{P}^2 = \begin{bmatrix} 1 & 1/3 & 5 & 1/5 & 1/3 \\ 3 & 1 & 7 & 1/4 & 1/2 \\ 1/5 & 1/7 & 1 & 1/7 & 1/5 \\ 5 & 4 & 7 & 1 & 5 \\ 3 & 1/2 & 5 & 1/5 & 1 \end{bmatrix} \quad (3)$$

$$\mathbf{P}^3 = \begin{bmatrix} 1 & 5 & 7 & 4 & 6 \\ 1/5 & 1 & 2 & 3 & 4 \\ 1/7 & 1/2 & 1 & 2 & 3 \\ 1/4 & 1/3 & 1/2 & 1 & 1/2 \\ 1/6 & 1/4 & 1/3 & 2 & 1 \end{bmatrix} \quad (4)$$

$$\mathbf{P}^4 = \begin{bmatrix} 1 & 4 & 1/4 & 1/3 & 1/5 \\ 1/4 & 1 & 1/7 & 1/6 & 1/8 \\ 4 & 7 & 1 & 1 & 1/2 \\ 3 & 6 & 1 & 1 & 2 \\ 5 & 8 & 2 & 1/2 & 1 \end{bmatrix} \quad (5)$$

3) 层次单排序及其一致性检验: 如式(6)所示, λ 为判决策矩阵的特征向量对应的最大特征值。

$$CI = \frac{\lambda - n}{n - 1} \quad (6)$$

4) 层次总排序及其一致性检验: 如式(7)所示, RI 均为一致性检验指标, 本文 RI 为 1.11^[16]。

$$CR = \frac{CI}{RI} \quad (7)$$

如果 $CR >= 0.1$, 则表示验证未通过, 应重新调整矩阵 \mathbf{P}^k , 直到 $CR < 0.1$ 。如果 $CR < 0.1$, 则表明一致性验证已通过。

5) 计算属性权重, 如式(8)所示。

$$W^k = (\omega_1^k, \omega_2^k, \dots, \omega_n^k)^T \quad (8)$$

2.3 客观权重的计算

熵值法可以用来确定网络属性的客观权重。该方法计算的属性权重被视为目标权重^[17]。使用熵值法计算每个属性的权重的步骤如下:

1) 归一化决策矩阵

针对各种指标的度量单位不一致, 需要对其进行标准化处理, 才能确定目标权重, 这样就能很好地解决各个指标的归一化问题, 按照属性对网络性能的不同影响, 可以分为效益型指标和成本型指标^[18], 收益型指标值越大越好, 例如带宽和吞吐量, 支出型指标值越小越好, 例如丢包率、时延等, 分别用式(9)、(10)进行归一化。

$$d'_{ij} = \frac{d_{ij} - \min_j(d_{ij})}{\max_j(d_{ij}) - \min_j(d_{ij})}, i, j \in [1, 2, \dots, n] \quad (9)$$

$$d'_{ij} = \frac{\min_j(d_{ij}) - d_{ij}}{\max_j(d_{ij}) - \min_j(d_{ij})}, i, j \in [1, 2, \dots, n] \quad (10)$$

2) 计算属性所占的比重, 如式(11)所示。

$$p_{ij} = d'_{ij} / \sum_{i=1}^m d'_{ij}, i, j = [1, 2, \dots, n] \quad (11)$$

3) 计算属性的熵值, 如式(12)所示。

$$e_j = -k \times \sum_{i=1}^n p_{ij} \ln(p_{ij}), i, j = [1, 2, \dots, n] \quad (12)$$

4) 计算信息熵冗余度, 如式(13)所示。

$$g_j = 1 - e_j, 1 \leq j \leq n \quad (13)$$

5) 计算各项客观属性的权重, 如式(14)所示。

$$w_j = \frac{g_j}{\sum_{j=1}^n g_j}, j = [1, 2, \dots, n] \quad (14)$$

2.4 综合权重的计算

本文结合了 AHP 和熵值法的优点, 综合权重可以用式(15)计算, ρ 和 θ 是权重系数, $\rho \geq 0, \theta \geq 0, \rho + \theta = 1$ 。取 $\rho = \theta = 0.5$ 。

$$W_j = \rho W_j + \theta W^k \quad (15)$$

3 算法描述

3.1 基本要素

1) 状态空间: 在决策时, 用户只能选择一个网络接入, 并且由各网络提供的服务质量(quality of service, QoS)参数来决定其网络状态。根据实时业务和非实时业务的特性, 选取 5 种 QoS 参数作为网络状态的表征, 定义为式(16)。

$$S = [T_1, E_1, B_1, PL_1, D_1, \dots, T_m, E_m, B_m, PL_m, D_m, m, k] \quad (16)$$

2) 动作空间: 切换动作用 a 表示, 则所有可选择的切换动作集可定义为式(17)、(18)表示的下一时刻网络的转移概率, g 和 g' 分别表示用户的当前状态和下一个决策时刻连接的网络。

$$a = \{a \mid a \in \{1, 2, \dots, m\}\} \quad (17)$$

$$P[s' \mid s, a] = \begin{cases} \prod_{m=1}^M P[t'_m, e'_m, b'_m, p t'_m, d'_m] & g' = a \\ 0 & g' \neq a \end{cases} \quad (18)$$

3) 奖励函数: 在每次决策时, 根据目前的网络状态, 选取目标网络, 并取得一个奖励值, 定义为式(19), 以体现不同 QoS 参数在切换决策中的重要程度。

$$r(s, a) = \sum_{j=1}^n W_j r_j(s, a) \quad (19)$$

当奖励函数为收益参数时, 如可用带宽, 奖励函数如式(20)所示。

$$r_n(s, a) = \begin{cases} 1 & n_a \geq U_n \\ (n_a - L_n)/(U_n - L_n) & L_n < n_a < U_n \\ 0 & n_a \leq L_n \end{cases} \quad (20)$$

当奖励函数是成本参数时, 如延迟抖动, 奖励函数如式(21)所示。

$$r_n(s, a) = \begin{cases} 1 & 0 < n_a \leq U_n \\ (U_n - n_a)/(U_n - L_n) & L_n < n_a < U_n \\ 0 & n_a \geq U_n \end{cases} \quad (21)$$

4) 期望回报值: 当网络状态满足切换条件时, 用户将执行网络选择。基于 MDP 模型的网络选择是最大化每个网络的期望回报值, 由式(22)计算出切换网络的期望回报值, 期望回报值最大的网络即为最佳接入网络。

$$v(s) = \max_{a \in A} \left\{ r(s, a) + \sum_{s' \in S} \gamma P[s' \mid s, a] v(s') \right\} \quad (22)$$

3.2 改进经验回放机制

传统的 DQN 方法是通过随机抽样的方式, 从经验池中随机选取样本且不能充分考虑样本的重要程度, 从而造成了高信息量的样本在网络训练中的使用率较低或出现没有被使用的情况, 降低了训练的效率^[19]。针对这一问题, 人们提出一种优先经验回放机制, 智能体会首先选择训练最有价值的样本, 为了防止过拟合, 低价值的样本也有被选择的概率, 但采样时间会比较长, 容易产生局部最优解^[20]。基于上述问题, 本文采用随机分层均匀的抽样方法, 对抽取的样本进行分层, 层数最多为抽样总数的 1/2, 以确保抽样方法不再是传统的抽样方式。该方法在保证抽样多样性的前提下, 也保证了对样本进行合理的选取, 并且可以加快算法的收敛速率。流程如图 3 所示。

在经验回放机制中采用时序误差 δ , 可以增加高价值样本的重要程度并且可以降低错误发生的概率。计算公式如式(23)所示, 从而可以避免网络过拟合的问题。

$$\delta = r + \gamma \max_{a'} Q(s', a', \theta) - Q(s, a, \theta) \quad (23)$$

3.3 Dueling-DQN 算法

DQN 算法是一种将 Q-learning 通过神经网络近似值

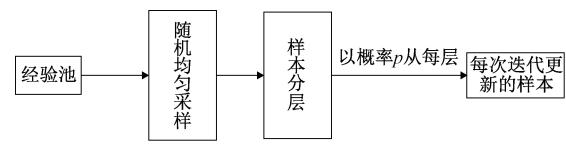


图 3 采样流程图

Fig. 3 Sampling flowchart

函数的一种方法, 采用神经元网络, 对各状态下各行为的 Q 值进行估计。其框架如图 4 所示。

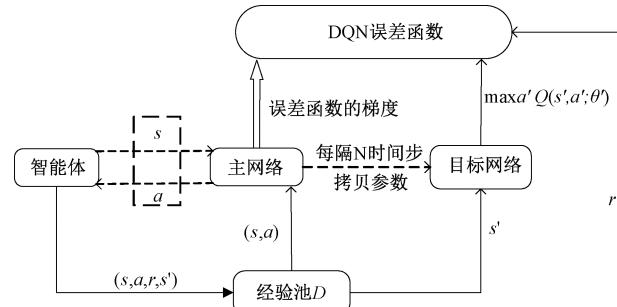


图 4 DQN 框架图

Fig. 4 DQN framework diagram

Dueling-DQN 属于 DQN 算法的改进, 该算法采用了分解 Q 值的方法, 分为状态值函数和优势函数。这种架构允许神经网络同时估计状态价值和不同行动的优势, 从而提高学习效率和稳定性。网络架构的关键组成部分如图 5 所示。

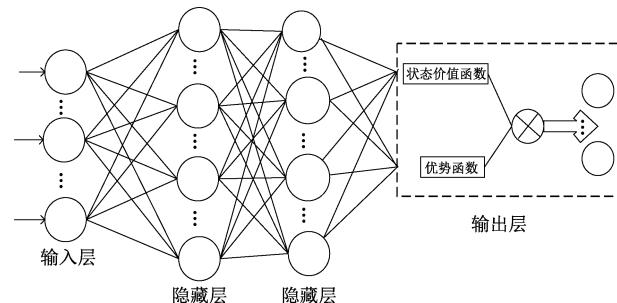


图 5 主网络结构图

Fig. 5 Main network structure

最终得出在每一种状态下采取动作所得到的 Q 值, 如式(24)所示。

$$Q(s, a; \alpha, \theta', \beta) = V(s; \alpha, \beta) + (A(s, a; \alpha, \theta') - \frac{1}{|A|} \sum_{a'} A(s, a'; \alpha, \theta')) \quad (24)$$

Dueling-DQN 模型在学习过程中, 通过对 Q 网络的参数进行调节, 使得网络的参数趋于最优。周期性地从经验池中选取状态和动作输入到 Q 网络中, 然后计算下一状态下的最优动作, 直至网络收敛。其中 Dueling-DQN 计算 Q 值的方式遵循贝尔曼方程^[21]可以得到主网络中 Q 值更新, 如式(25)、(26)所示。

$$Q(s, a) = E[R + \gamma \max_{a'} Q(s', a') \mid s, a] \quad (25)$$

$$Q_M(s, a, \alpha) = Q_m(s, a, \alpha) + \lambda(r + \gamma \max_a Q_M(s', a, \alpha)) - Q_M(s, a, \alpha) \quad (26)$$

目标网络根据动作计算的 Q 值可表示如式(27)所示。

$$Q_T(s', a', \alpha) = r + \gamma Q_T[s', a^{\max}(s, \alpha), \alpha] \quad (27)$$

在训练过程中, 通过运用均方误差 (mean squared error, MSE) 的方法计算损失函数, 并利用梯度下降法 (gradient descent method, GDM)^[22] 和反向传播过程完成参数的更新, 计算如式(28)所示。

$$\text{Loss} = E[(Q_T - Q_M)^2 + \lambda \times (A(s, a; \alpha, \theta') - V(s; \alpha, \beta))^2] \quad (28)$$

并且采用了 ϵ -greedy 策略加快算法的收敛速度, 具体公式如式(29)所示, 在训练阶段, 智能体按照 ϵ 的概率随机选择策略, 并将其逐渐降低直至 0。

$$\pi(s) = \begin{cases} \underset{a}{\operatorname{argmax}} Q(s, a), & p = 1 - \epsilon \\ \text{random}(\pi), & p = \epsilon \end{cases} \quad (29)$$

4 仿真结果及分析

4.1 参数设置

利用 MATLAB 仿真工具, 仿真环境由 5G、Wimax、WLAN、LTE 四种网络构成。终端用户以 5 m/s 的速度进行运动, 服从泊松分布。仿真实验将本文所提算法与文献中 MDP、DQN、ES-DQN、MDP-GA-SA^[23] 算法进行比较, 验证改进算法的优势。仿真参数的设置如表 2 和 3 所示。

表 2 系统仿真参数设置

Table 2 System simulation parameter settings

	5G	Wimax	LTE	WLAN
数量/个	2	3	3	5
范围/m	1 000	300	200	100
发射功率/dBm	35	25	20	15
噪声功/(dBm·Hz ⁻¹)	-175	-175	-175	-175
容纳最大终端数	120	30	20	15

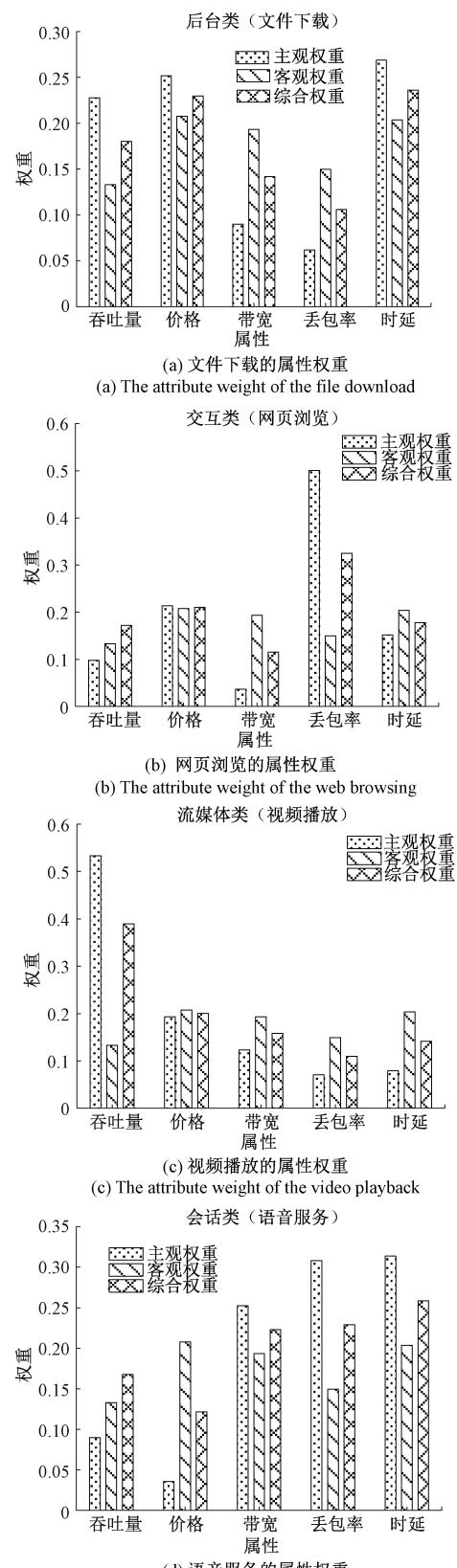
表 3 算法仿真参数设置

Table 3 Algorithm simulation parameter settings

参数名称	参数值
探索率 ϵ	0.25
折扣因子 γ	0.6~0.72
学习率 λ	0.000 2
经验回放池容量 D	2 500
网络参数更新频率 F	100
训练回合数	10 000

4.2 综合权重分析

图 6(a)~(d) 分别显示不同业务类型下的主观权重、客观权重和综合权重。由图 6(a)~(d) 知后台类以价格和时延为主, 交互类主要以丢包率为主, 流媒体类主要以吞吐量为主, 会话类主要以时延和丢包率为主。



(d) 语音服务的属性权重

(d) The attribute weight of the speech services

图 6 属性权重

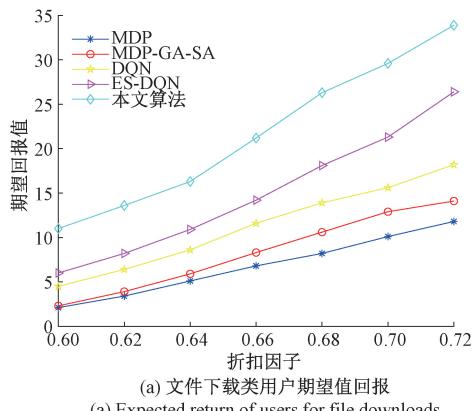
Fig. 6 Attribute weights

4.3 算法期望回报值分析对比

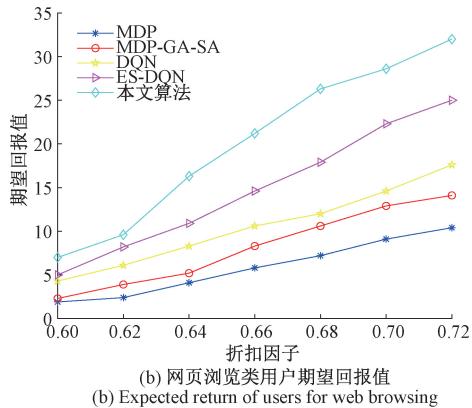
图 7(a)~(d) 分别为折扣因子与文件下载、网页浏览、视频播放和语音服务用户的期望回报值之间的关系。图中将本文提出的基于改进的 Dueling-DQN 算法与 MDP、MDP-GA-SA、DQN、ES-DQN 进行性能对比。由图 7(a)~(d) 可知, 随着折扣因子值 γ 从 0.6 增大到 0.72, 算法的期望回报值均呈递增趋势, 这是因为 γ 反映了算法对网络长期收益的关注程度, 每一次决策的回报值根据不同用户业务类型特点, 构建奖励函数及网络状态的属性权重, 从而使算法的期望回报差距也随之增大。

4.4 切换次数对比分析

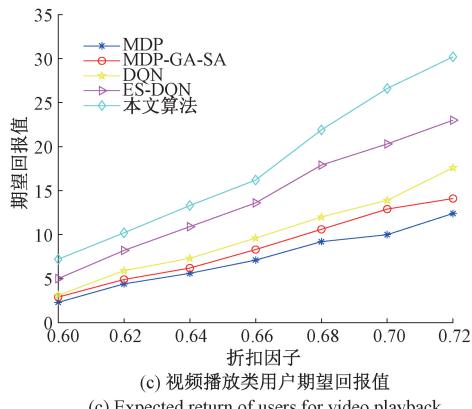
切换次数是评价垂直切换决策算法的一个非常重要的指标。图 8(a)~(d) 分别指折扣因子 0.6~0.72 时, 不同业



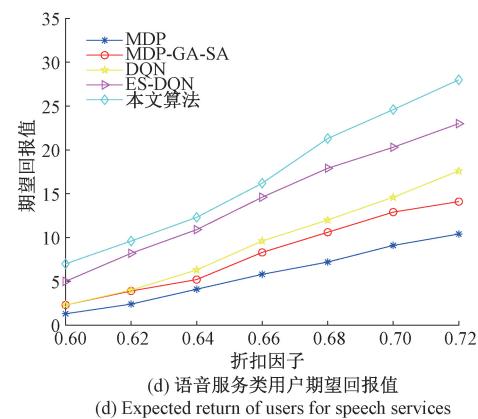
(a) 文件下载类用户期望回报
(a) Expected return of users for file downloads



(b) 网页浏览类用户期望回报值
(b) Expected return of users for web browsing



(c) 视频播放类用户期望回报值
(c) Expected return of users for video playback



(d) 语音服务类用户期望回报值
(d) Expected return of users for speech services

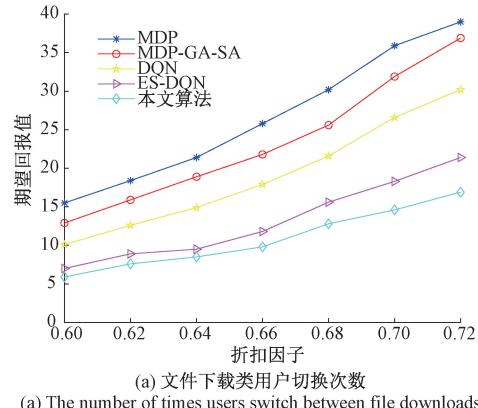
图 7 不同业务类型下用户的期望回报值

Fig. 7 Expected return of users for different service

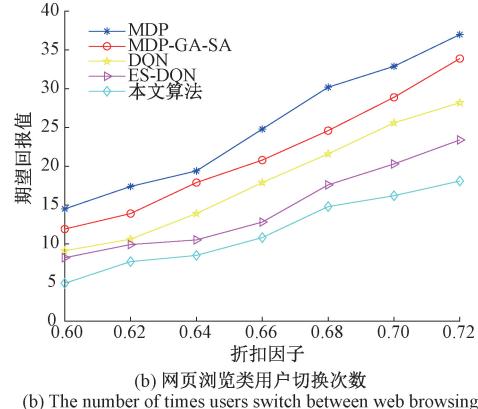
务类型下的切换次数对比。可看出, 随着 γ 的增大决策次数逐渐增多, 本文所提算法在不同类型下相较于文献[6]中所提 ES-DQN 算法分别减少了 11.25%, 13.34%, 18.76%, 13.75%, 这是因为使用随机分层抽样可以加快算法的收敛速度, 同时引入了 ϵ -greedy 策数并且避免了不必要的切换。避免了网络切换中发生频繁切换和“乒乓”现象的发生。实验结果表明, 该方法可以有效地减少切换次数。

4.5 吞吐量分析

图 9 为不同用户数量下, 网络的平均吞吐量变化情况。当用户数超过 100 时, 由于网络可利用的带宽资源会随用



(a) 文件下载类用户切换次数
(a) The number of times users switch between file downloads



(b) 网页浏览类用户切换次数
(b) The number of times users switch between web browsing

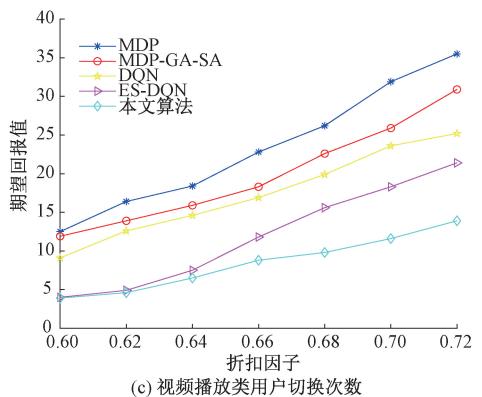
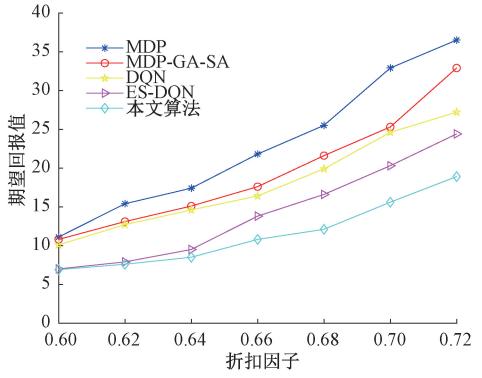
(c) 视频播放类用户切换次数
(c) The number of times users switch between video playback(d) 语音服务类用户切换次数
(d) The number of times users switch between speech services

图 8 不同业务类型下的切换次数

Fig. 8 The number of switchovers for different service types

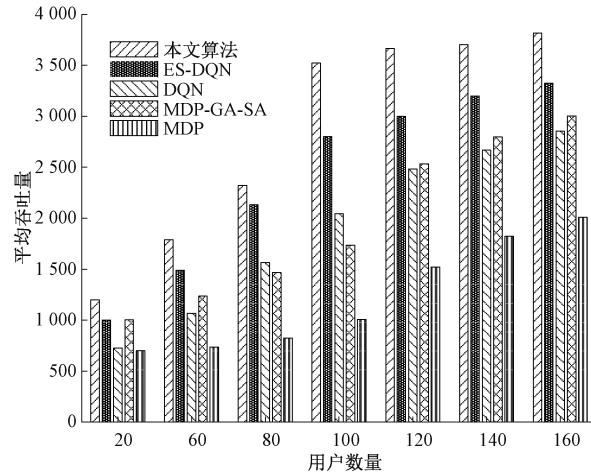


图 9 平均吞吐量

Fig. 9 Average throughput

户数的增加而减少,所以上升缓慢,当用户数为 160 时,本文所提算法较于所提 ES-DQN 算法提升了约 16.64%。主要原因在于,当供选择的网络数量较多时,不能有效选择接入的网络,使得网络总资源利用率较低,所提算法充分考虑用户在不同业务下的 QoS 需求,从而提升了网络的平均吞吐量。

5 结 论

本文分析了异构无线网络环境下垂直切换决策的过程,基于 SDN 垂直切换框架,提出了一种基于主观加权和改进的 Dueling-DQN 异构无线网络垂直切换决策方案,详细分析了客观加权法、主观加权法和综合加权法的过程。然后使用改进的 Dueling-DQN 网络选择算法,对不同的业务类型,设计不同的奖励函数,通过切换次数和吞吐量来判定选网策略的优劣。仿真结果表明,在四种业务类别下,可以减少总切换次数和提升吞吐量。因此,可以减少不必要的切换次数和能耗;同时可以提高频谱资源的利用率、网络服务质量以及用户体验质量。

参考文献

- [1] 俞鹤伟,梁根.异构无线网络接入选择算法综述[J].哈尔滨工业大学学报,2017,49(11):178-188.
YU H W, LIANG G. A review of heterogeneous wireless network access selection algorithms [J]. Journal of Harbin Institute of Technology, 2017, 49(11):178-188.
- [2] ZHONG Y F, WANG H Q, LYU H W, et al. A vertical handoff decision scheme using subjective-objective weighting and grey relational analysis in cognitive heterogeneous networks [J]. Ad Hoc Networks, 2022, 134(7):102924.
- [3] 刘春玲,田玉琪,张琪珍,等.面向 5G 超密集异构网络的模糊逻辑切换算法[J].电讯技术,2023,63(11):1803-1810.
LIU CH L, TIAN Y Q, ZHANG Q ZH, et al. Fuzzy logic switching algorithm for 5G ultra-dense heterogeneous network [J]. Telecommunication Technology, 2023, 63(11):1803-1810.
- [4] MITHUN B P, PATIL R. Fuzzy based network controlled vertical handover mechanism for heterogeneous wireless network[J]. Materials Today: Proceedings, 2023, 80:2385-2389.
- [5] NING ZH L, SONG Q Y, LIU Y J, et al. Markov-based vertical handoff decision algorithms in heterogeneous wireless networks [J]. Computers & Electrical Engineering, 2014, 40(2):456-472.
- [6] 郭强,朱若函,张晓萌.基于遗传禁忌算法优化的模糊神经网络垂直切换算法[J].计算机应用研究,2016,33(3):840-842.
GUO Q, ZHU R H, ZHANG X M, et al. Fuzzy neural network vertical switching algorithm optimized based on genetic tabu algorithm [J]. Computer Application Research, 2016, 33(3):840-842.
- [7] TAN X N, CHEN G, SUN H Y, et al. Vertical handover algorithm based on multi-attribute and neural

- network in heterogeneous integrated network [J]. EURASIP Journal on Wireless Communications and Networking, 2020, 10(1):202.
- [8] SONG Y, LIM S H, JEON S W, et al. Handover decision making for dense HetNets: A reinforcement learning approach [J]. IEEE Access, 2023, 11(9): 24737-24751.
- [9] SUN J N, QIAN ZH H, WANG X, et al. ES-DQN-based vertical handoff algorithm for heterogeneous wireless networks[J]. IEEE Wireless Communications Letters, 2020, 9(8):1327-1330.
- [10] 段盈江,赵一帆,丁广恩,等.基于深度强化学习的无人机切换管理研究[J].无线电通信技术,2024,7:1-9.
DUAN Y J, ZHAO Y F, DING G EN, et al. Research on UAV switching management based on deep reinforcement learning [J]. Radiocommunication Technology, 2024, 7:1-9.
- [11] WU M T, HUANG W, SUN K, et al. A DQN-based handover management for SDN-enabled ultra-dense networks[J]. 2020 IEEE 92nd Vehicular Technology Conference, 2020, 18:1-6.
- [12] AMIN R, MARTIN R, NADIR S. Hybrid SDN networks: A survey of existing approaches[J]. IEEE Communications Surveys & Tutorials, 2018, 20(4): 3259-3306.
- [13] SARBAST M. A systematic review of analytic hierarchy process applications to solve transportation problems from 2003 to 2022[J]. IEEE Access, 2023, 11(7):11973-11990.
- [14] SANDEEPA P, KUMAR A, RAM A, et al. Consistency indices in analytic hierarchy process: A review[J]. Mathematics, 2022, 10(8):1206.
- [15] 钟玉峰. 基于多属性决策的认知网络接入选择模型及方法[D]. 哈尔滨: 哈尔滨工程大学, 2024.
ZHONG Y F. Access selection model and method of cognitive network based on multi-attribute decision-making [D]. Harbin: Harbin Engineering University, 2024.
- [16] SARBAST M. A novel parsimonious spherical fuzzy analytic hierarchy process for sustainable urban transport solutions [J]. Engineering Applications of Artificial Intelligence, 2024, 128:107447.
- [17] XIAO F Y. EFMCDM: Evidential fuzzy multicriteria decision making based on belief entropy [J]. IEEE Transactions on Fuzzy Systems, 2020, 28 (7): 1477-1491.
- [18] SANTHI J, PRABHA K. QOS aware vertical handover process in heterogeneous wireless network [J]. Measurement: Sensors, 2023, 26(8):100710.
- [19] 邓修朋,崔建明,李敏,等.深度强化学习在机器人路径规划中的应用[J].电子测量技术,2023,46(6):1-8.
DENG X P, CUI J M, LI M, et al. Application of deep reinforcement learning in robot path planning[J]. Electronic Measurement Technology, 2023, 46(6):1-8.
- [20] 许向阳,李京阳,彭文鑫.一种基于 Dueling DQN 改进的低轨卫星路由算法[J].长江信息通信,2023,36(7): 56-59.
XU X Y, LI J Y, PENG W X. An improved low-orbit satellite routing algorithm based on Dueling DQN[J]. Yangtze River Information and Communication, 2023, 36(7):56-59.
- [21] 谭俊杰,梁应敞.面向智能通信的深度强化学习方法[J].电子科技大学学报,2020,49(2):169-181.
TAN J J, LIANG Y CH. Deep reinforcement learning method for intelligent communication [J]. Journal of University of Electronic Science and Technology of China, 2020, 49(2):169-181.
- [22] 李兴怡,岳洋.梯度下降算法研究综述[J].软件工程, 2020,23(2):1-4.
LI X Y, YUE Y. Review of gradient descent algorithm[J]. Software Engineering, 2020, 23(2):1-4.
- [23] XIE J L, GAO W J. Heterogeneous network selection optimization algorithm based on a Markov decision model[J]. China Communications, 2020, 17(2):40-53.

作者简介

李敏之(通信作者),副教授,硕士研究生,主要研究方向为异构无线通信、智能信息处理等。

E-mail:1976209638@qq.com