

引用格式:丁雨,李晨凯,卢为党,等.基于强化学习的异构网络垂直切换方法[J].信息对抗技术,2023,2(3):35-43.[DING Yu, LI Chenkai, LU Weidang, et al. Vertical handover for heterogeneous networks based on deep reinforcement learning[J]. Information Countermeasure Technology, 2023, 2(3):35-43. (in Chinese)]

基于强化学习的异构网络垂直切换方法

丁雨¹,李晨凯¹,卢为党^{1*},张清清¹,任元红²,高原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 网络切换技术不仅可以保证用户的网络连接,同时能够以较强的信号传输网络数据。网络切换技术的性能对网络服务质量(QoS)的影响至关重要。然而,现有的切换算法多数存在严重的乒乓效应,这不仅造成网络资源的严重浪费,还会损害用户的QoS。为此,提出了一种基于强化学习的异构网络垂直切换方案,主要从触发切换、网络选择及判决切换等方面进行优化。在触发切换时将垂直切换考虑成必要切换和择优切换,通过Q-Learning(QL)算法在选择网络时优化垂直切换;然后以QoS为条件,在判决切换时加入驻留定时器,从多角度减少用户切换次数,降低乒乓效应对异构网络垂直切换的影响。仿真结果表明,基于强化学习的异构网络垂直切换方法可以在保证QoS的条件下有效减少用户切换网络的次数,短时间内频繁切换的情况也有所改善,降低了乒乓效应的影响。

关键词 垂直切换;网络选择;判决切换;服务质量;异构网络

中图分类号 TN 925

文章编号 2097-163X(2023)03-0035-09

文献标志码 A

DOI 10.12399/j.issn.2097-163x.2023.03.003

Vertical handover for heterogeneous networks based on deep reinforcement learning

DING Yu¹, LI Chenkai¹, LU Weidang^{1*}, ZHANG Qingqing¹, REN Yuanhong², GAO Yuan³

(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 Sciences of the PLA, Beijing 100091, China)

Abstract Network handover technology can not only ensure the network connection of users, but also transmit network data with strong signals. The performance of network handover has a critical impact on the quality of service (QoS). However, most of the existing handover algorithms have serious ping-pong effect, which may cause a waste of network resources and damage the QoS. Therefore, a vertical handover scheme for heterogeneous networks based on reinforcement learning was proposed, which was mainly optimized from the aspects of triggering handoff, network selection and decision handover. Specifically, the necessary and the preferred handover were considered for the vertical handover when the handover was triggered. Then, the vertical handover was optimized when the network through Q-Learning(QL) was selected. Based on QoS, the resident timer was also added when deci-

ding the handover to reduce the number of user handover from multiple angles and reduced the impact of ping-pong effect on the proposed vertical handover of heterogeneous networks. Simulation results have shown that the proposed vertical handover scheme for heterogeneous networks could effectively reduce the number of times for handover, improve the situation of frequent handover in a short time, and reduce the impact of the ping-pong effect while ensuring the quality of service.

Keywords vertical handover; network selection; decision handover; QoS; heterogeneous network

0 引言

无线接入为传统无线通信网络带来了蓬勃发展的机会,出现具有不同特性的网络,用来服务于特定的应用场景^[1]。用户对于无线通信网络的服务质量(quality of service, QoS)有个性化和多样化的需求,这对目前仅提供有限服务的单一无线网络产生了较大挑战^[2-3]。

异构网络由多种类型的网络系统组成,在异构网络中,各种通信技术互相融合、补充和促进,因而在提供更多网络连接的同时,有效满足多样化的 QoS^[4-6]。然而在一些特殊场景,例如山区、地震及洪水发生地周围,由于铺设固定基站难度大、成本高,因此需要采取更加灵活、操控简单的方法来提供网络服务。移动式无人机可以作为网络基站载体,具有快速部署的优势,能够在这些特殊场景下提供网络连接^[7-10]。文献[8]指出无人机具有视距传输效率高以及高度等可控的优势,为无线网络展现新的设计自由度,在其研究中考虑了信道状态信息的过时性,探讨了无人机通信异构网络的切换问题。文献[9]指出无人机网络融入异构网络中为未来天地一体化通信提供了重要研究方向。文献[10]引入平均服务成功概率来衡量无人机通信异构网络传输性能。

在异构网络中,多种类型网络并存,用户有较大可能处于同时被多种类型网络覆盖且可使用的情况^[11]。用户可能因为移动而需要切换网络,也有可能期望获得更好的 QoS 或更低成本等目的而切换网络^[12-13]。网络切换中不同网络之间的切换为垂直切换^[14-18]。文献[15]提出了一种改善用户体验质量的垂直切换算法来解决由异构网络的超高动态性引起掉话率增长的问题。文献[16]提出了一种考虑负载均衡和用户体验的垂直切换算法来为服务用户选出当前状态满

意度最高的目标网络,从而有效缓解网络拥塞。文献[17]通过提出的基于信道预留以及剩余带宽的切换网络方法实现网络接入概率的目的,提出了一种基于网络剩余带宽和信道预留的切换策略,其仿真结果表明该切换网络的策略无论是对于实时的还是非实时的业务,网络接入概率都可以获得提升。文献[18]指出用户对网络的偏好和对网络性能的需求日益增长,提出根据用户对网络的偏好和对网络性能需求的异构网络切换方法,其可以为用户在众多不同网络中切换并连接到适合它的网络。

在异构网络切换判决中,如果所用切换判决机制对网络环境的微小变化太过敏感,容易导致乒乓效应。乒乓效应是指在网络中,2个或多个通信节点之间的信息传输在较短时间内来回反弹,形成一种来回的“乒乓”现象。这样的过程会占用大量的网络带宽和处理能力,导致网络拥塞和性能下降,同时也会导致网络中出现过多的控制信息、传输延迟增加、带宽利用率低下等问题,从而降低用户的 QoS 体验,浪费网络资源。多数研究为减少乒乓效应的影响,在目标网络已确定的情况下判决是否继续切换以及切换时间,从而达到异构网络切换的无缝连接^[19-20]。文献[19]使用层次分析法来确定标准权重大小,通过多属性对网络进行决策。文献[20]根据移动用户状态,动态调整驻留定时器时间长短来避免不必要的切换网络动作。强化学习通过与环境交互学习来实现最大化回报的目标,广泛应用工业控制、无线通信、智慧交通以及智慧医疗等各方面,在网络切换中受到极大关注^[21]。文献[21]在选择网络时采用强化学习方法,基于性价比和博弈论的价格策论,通过神经网络学习和探寻满足纳什均衡条件的策略来确定选择网络。

目前的研究大多仅考虑信号接收强度为触

发切换和判决切换的标准,并只在判决切换方面进行优化,忽略了 QoS 对切换网络的作用和影响。此外,乒乓效应也可能导致严重后果。为此,本文提出一种基于强化学习的异构网络垂直切换方案,主要从触发切换、网络选择及判决切换等方面进行优化。

1 系统模型

考虑一个如图 1 所示的异构网络切换系统模型,其中包括 Wi-Fi、5G 无人机网络、5G CPE 网络以及多个移动的地面用户。5G CPE 网络由其中一个地面用户携带,而 5G 无人机网络和 Wi-Fi 同时搭载在无人机身上。在该异构网络切换系统中, M 个用户由于 N 种类型的业务请求需与网络连接通信。

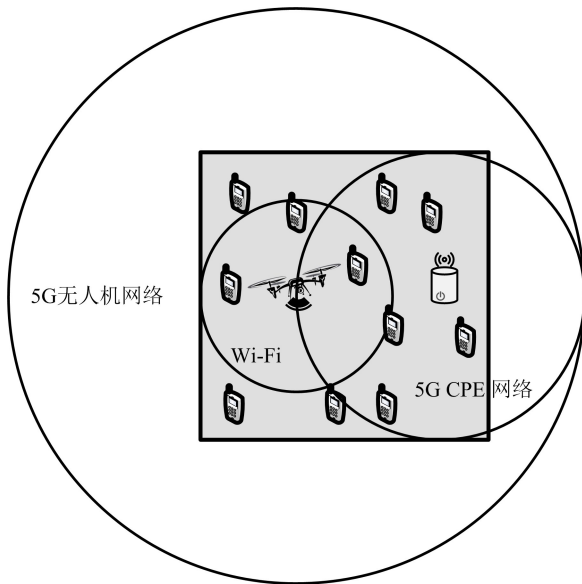


图 1 异构网络切换系统

Fig. 1 Heterogeneous network handover system

在该异构网络切换系统中,用户由于移动,其位置不断变化。用户在初始时随机分布在异构网络切换系统的阴影部分内,随后用户将随机发生移动,他们可能会频繁跨越网络边界,导致触发垂直切换。

地面用户由于 N 种类型业务请求需连接网络。假设用户当前被 r ($r \leq 3$) 个网络覆盖,当用户需要请求网络服务时,它可以从这 r 个网络中选择一个进行连接,来享受该网络提供的服务。由于用户处于移动的状态,当用户位置发生变化,使其不再被原来服务的网络所覆盖或是其追求更好的 QoS 时,此时用户可以进行异构网络的

垂直切换,与当前网络断开连接,与覆盖它的其他网络连接通信。在此异构网络切换系统中,用户同一时间不被允许与多个网络进行连接,只能与某一个网络连接。同时,考虑这 3 个网络的负载上限 $\{T_{TL_1}, T_{TL_2}, T_{TL_3}\}$,当某一网络已经接入的用户数量达到负载上限时,表示该网络无法再接受新的连接请求。此时,用户只能尝试连接其他网络。

2 异构网络垂直切换算法

在基于强化学习的异构网络垂直切换算法中,分别从触发切换、选择网络和判决切换进行优化,从多角度减少用户切换次数,减少乒乓效应对提出异构网络的垂直切换影响。

2.1 触发切换

用户在移动过程中触发切换大致有以下原因:由于用户移动导致网络覆盖问题;网络在均衡负载时而使使得用户切换网络;当用户期望获得更好的 QoS,如期望更低能耗、更低资费等。为了使用户在接受网络服务时减少非必要切换的次数,将切换分成择优切换和必要切换。其中,择优切换表示此时正在连接的网络服务稳定,但用户由于自身原因为了实现更好的 QoS 而切换网络,其条件描述为:此时正在连接的网络供应的 QoS 不是该用户所有可以接入网络的最佳 QoS。必要切换表示此时正在连接的网络无法继续使用,用户为了保持正常的网络通信与数据传输而切换网络,其条件描述为:此时连接的网络接收信号强度 RSS 无法达到最小门限值。

触发必要切换时, RSS 表示为:

$$R_{RSS} [\text{dB}] = G_r + G_t + P_t - L_{\text{shadow}} - L \quad (1)$$

式中, G_r 和 G_t 分别是基站发射和用户接收的传播增益, P_t 是基站发射功率, L_{shadow} 和 L 分别是阴影衰弱和路径损耗。

Wi-Fi 的路径损耗表示为:

$$L [\text{dB}] = 20 \lg f + 26 \lg d + 42.6 \quad (2)$$

式中, f 和 d 分别为传输频率 (MHz) 和基站与用户的距离 (km)。

5G CPE 和 5G 无人机网络的路径损耗表示为:

$$L [\text{dB}] = 33.9 \lg f + (44.9 - 6.55 \lg h_B) \lg d + 46.3 + C_m - 13.82 \lg h_B - a(h_R, f) \quad (3)$$

$$a(h_R, f) = h_R(1.1 \lg f - 0.7) - (1.56 \lg f - 0.8) \quad (4)$$

式中, h_B 和 h_R 分别为基站和移动台的有效天线高度(m), C_m 为恒定偏移量:

$$C_m = \begin{cases} 3 \text{ dB}, & \text{大都市} \\ 0 \text{ dB}, & \text{郊区及中等城市} \end{cases} \quad (5)$$

为了使得网络可以达到基础的连接要求, 用户需要网络 RSS 达到基本的灵敏度要求。由于不同网络基础连接要求不同, 因此用户对各异构网络的灵敏度需求也各不相同, 假设用户对 Wi-Fi、5G CPE 和 5G 无人机的 RSS 要求灵敏度分别为 -80 、 -97 和 -97 dB。

触发择优切换时, 需衡量各个网络供应的 QoS。将传输速率以及时延纳入网络供应 QoS 的衡量因素。因此, 网络 QoS 表示为:

$$Q_{\text{QoS}} = \omega_{\text{rate}} U_{\text{RATE}} + \omega_d U_{\text{Delay}} \quad (6)$$

式中, U_{RATE} 和 U_{Delay} 分别是当前网络的传输速率和时延, ω_{rate} 和 ω_d 是 QoS 分别是对传输速率以及网络时延的偏好权重, 表示各个 QoS 的衡量因素在 QoS 中体现的重要程度是不一致的。

2.2 网络选择

在网络选择中, 考虑将用户当前正在连接的网络的信息加入到状态中去, 从而分析用户当前正在连接的网络对用户服务的好坏。在不同状态下使用 Q-Learning (QL) 算法选择其中的异构网络获得即时回报学习, 具体如下:

状态 s : 网络选择的算法需要为用户选出合适的网络, 优化均衡网络的负载, 同时还要避免乒乓效应对异构网络垂直切换的影响。首先考虑当前某网络 j 的负载为 L_j , 则其负载占用率为 $\frac{L_j}{T_{\text{TL}_j}}$ 。为了便于 QL 处理, 首先将连续的网络负载占用率量化成 4 个等级^[6], 如表 1 所列。

表 1 负载占用率量化区间

Tab. 1 Load occupancy rate quantification interval

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

因此, 本算法的状态为:

$$s = [l_1, l_2, l_3, a_p] \quad (7)$$

式中, l_1 、 l_2 、 l_3 分别为 5G CPE、5G 无人机以及 Wi-Fi 网络的负载占用率量化等级, a_p 为用户此

时正在连接与通信的网络种类, $a_p = \{0, 1, 2, 3\}$ 分别代表用户没有连接网络、用户此时正在连接的可用网络是 5G CPE、用户此时正在连接的可用网络是 5G 无人机以及用户此时正在连接的可用网络是 Wi-Fi。

动作 a : 用户根据环境给出的状态, 通过强化学习选择切换到哪个网络, 假设 A 表示在该算法中的可选动作组合, 表示为:

$$A = \{a, a \in \{1, 2, 3\}\} \quad (8)$$

式中, $a = 1, 2, 3$ 分别代表此时用户切换到 5G CPE、5G 无人机和 Wi-Fi 网络。

回报 r : 根据异构网络切换系统给出的环境状态, 该算法选择动作应用至环境中获得效用函数和即时回报。效用函数用来奖励 QoS 更好同时可以均衡负载的切换网络动作, 与传输速率、时延以及负载占用率有关系, 表示为:

$$G(s, a) = \omega_{\text{rate}} U_{\text{RATE}}^a + \omega_d U_{\text{Delay}}^a + \omega_{\text{lb}} U_{\text{LB}}^a \quad (9)$$

式中, U_{RATE}^a 、 U_{Delay}^a 以及 U_{LB}^a 分别为动作 a 可以得出的传输速率、时延以及网络负载占用率方差, ω_{rate} 、 ω_d 以及 ω_{lb} 分别代表效用函数对传输速率、时延以及负载均衡的权重, 不同的业务对连接网络的性能需求不一样, 其权重对应的值也不相同。

回报 $r(s, a)$ 表示为:

$$r(s, a) = \rho G(s, a) l\{a = a_p\} + (1 - \rho) G(s, a) l\{a \neq a_p\} \quad (10)$$

式中, $l\{x\}$ 主要起到指示作用, 当满足对应条件 x 时, $l\{x\}$ 为 1, 否则为 0; ρ ($\rho \in (0.5, 1]$) 为回报切换因子, 其主要为了减少地面用户切换网络的总次数。若用户选择执行 a 与在这之前连接的网络 a_p 一致, 则回报 $r(s, a)$ 为 $\rho G(s, a)$ 。若用户选择执行 a 与在这之前连接的网络 a_p 不一致, 回报 $r(s, a)$ 则为 $(1 - \rho) G(s, a)$ 。由于 $\rho \in (0.5, 1]$, 因此有 $\rho > 1 - \rho$, 这表明当用户没有产生网络切换时, 可得到较高的回报值, 符合实际。

在 QL 中, 每当学习过程开始之前, M 个用户以随机的方式组成初始队列, 依次根据异构网络垂直切换系统给出的状态开始选择网络。当用户执行动作后, 此时智能体会获得回报, 同时环境得到下一时刻的不同状态, 智能体也将根据得到的信息去更新 $Q(s, a)$ 。接下来, 从初始队列中将已经实施过动作的用户移除。新队列的

第一个用户根据系统状态选择网络,并重复此类操作,直至队列中所有用户均被处理。这个过程表明智能体在一次迭代中完成了对所有用户的学习,并准备开始下一轮学习。在下次学习之前,智能体需要重置网络服务队列的状态,以确保每一次的学习环境都是相同的。

在动作的选择中,该算法以 ϵ 概率使用贪婪策略探索更多可能情况来学习。智能体以 $1-\epsilon$ 概率进行利用,该算法的“利用”是指在当前环境状态 s 下选择最大 $Q(s, a)$ 对应的动作来得到最好收益。该算法网络选择表示为:

$$a = \begin{cases} \operatorname{argmax}_{a \in A} Q(s, a), & \epsilon < \operatorname{rand}(0, 1) < 1 \\ \text{随机连接网络}, & 0 < \operatorname{rand}(0, 1) \leq \epsilon \end{cases} \quad (11)$$

完成动作得到回报后,需对当前环境状态和动作组合对应的 Q 值进行更新,即:

$$Q^{\text{new}}(s, a) = \alpha \times (r + \beta \cdot \max_a Q(s', a)) + (1 - \alpha)Q(s, a) \quad (12)$$

上述算法的具体过程如算法 1 所示。

算法 1 基于强化学习的异构网络垂直切换算法
Initialize: 贪婪策略探索概率 ϵ 、网络初始状态 s_0 、 Q 值表,初始值为 0、学习率 α 、折扣因子 β 、回报切换因子 ρ

1. for $a = 1$: episode do
2. 系统中的用户以随机的方式组成初始队列,并计算初始状态 s_0 。
3. while 智能体服务请求队列不为空 do
4. 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,根据贪婪算法选择一个动作执行
5. 智能体得到回报,计算下一状态 s'
6. 智能体根据式(12)更新 Q 值表
7. 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
8. end while
9. 随机重置队列
10. end for
11. 输出:智能体的 Q 值表

2.3 判决切换

由于地面用户是移动的,因此,网络状态以及用户的位置在完成切换后也会不同,所选择的网络可能会出现不可用或不是最优的情况。在短时间内出现该情况将会触发切换,增加垂直切换的次数,提高了乒乓效应出现的概率。因此,

可以考虑在网络选择完成后进行切换判决,判断是否需要再次进行切换,以减少不必要的切换,降低乒乓效应出现的概率。

切换在触发切换进程中可以被分为择优切换和必要切换。择优切换是指用户当前所处的网络无法提供最优服务质量而产生的切换。为了避免频繁的切换,需要将服务质量作为判定的标准,并引入驻留定时器辅助判决。其具体流程为:在网络选择后获得网络决策的结果,启动可变时长的驻留定时器,当其超时后通过计算 QoS 和网络决策结果判断该网络是否提供了最优服务。若是,则不再进行网络切换;反之,则继续完成其余切换。必要切换是指在网络接收信号强度抵御设定阈值的情况下产生的切换进程,此时连接的网络会出现不可用的情况,需要尽快选择新的网络进行切换以保证用户与网络之间的连通性。因此,在必要切换时可以直接判定继续进行网络切换。

在切换执行中,为实现最大程度上的无缝性,需要在通信从原网络切换至目标网络的过程中尽可能地降低切换时延。这需要根据前一步骤的判决结果进行相应的操作:当判决结果为执行切换时,将通过基于 QL 的网络选择算法将网络切换至选择的最优网络上;反之,将不再进行操作。

3 仿真分析

用户和网络的初始位置分布如图 2 所示,用户随机分布在 $300 \text{ m} \times 300 \text{ m}$ 的指定区域内。5G 无人机、Wi-Fi 以及 5G CPE 网络的覆盖区域各不相同,其中,5G 无人机可以覆盖的区域半径为 300 m ,Wi-Fi 和 5G CPE 覆盖的区域半径则分别为 100 、 150 m 。用户在移动中随机选择一个方向,并在该方向及其相反方向上以 $3 \sim 6 \text{ m/s}$ 的速度持续做来回运动。

假设 5G 无人机、Wi-Fi 以及 5G CPE 网络的负载上限 $\{T_{\text{TL}_1}, T_{\text{TL}_2}, T_{\text{TL}_3}\}$ 分别为 $\{25, 25, 25\}$ 。假设这 3 种异构网络的特性不同,5G 无人机、Wi-Fi 以及 5G CPE 网络对传输速率的参数范围分别为 $1.4 \sim 1.6$ 、 $1.7 \sim 1.9$ 、 $1.2 \sim 1.4 \text{ Gbit/s}$,对时延的参数范围分别为 $10 \sim 30$ 、 $40 \sim 70$ 、 $20 \sim 45 \text{ ms}$ 。QL 中,假设 $\epsilon = 0.4$, $\alpha = 0.1$, $\beta = 0.8$, $\rho = 0.7$ 。用户请求的业务类型包括语音类业务、视频类业务和文本类业务共 3 种,在计算即时回

报 r 时,3 种业务类型对效用函数参考属性(负载占用率方差、时延、传输速率)的权重偏好不同,具体的权重偏好如表 2 所列。

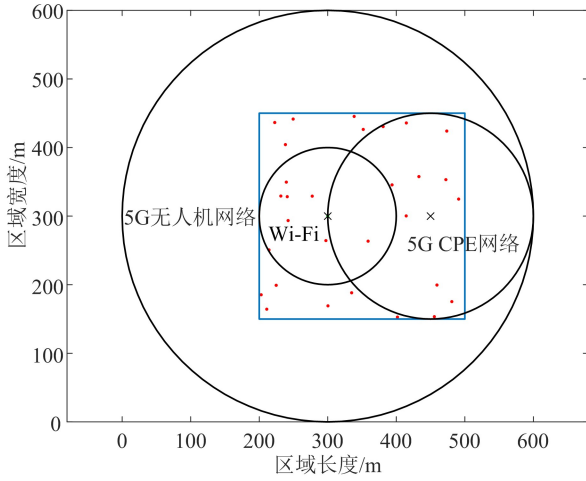


图 2 用户和网络初始位置分布

Fig. 2 The initial position of the users and heterogeneous networks

表 2 3 种业务类型对效用函数参考属性的偏好值

Tab. 2 Preference values of three business types for reference attributes of utility function

	负载占用率方差	时延/ms	传输速率/(Gbit · s ⁻¹)
语音类业务	0.5	0.8	0.4
视频类业务	0.5	0.5	0.6
文本类业务	0.5	0.3	0.2

用户在各个时刻不使用驻留定时器和使用时长为 3 s 驻留定时器的结果如图 3~4 所示。网络值等于 0 表明没有网络连接,1 表明连接 5G

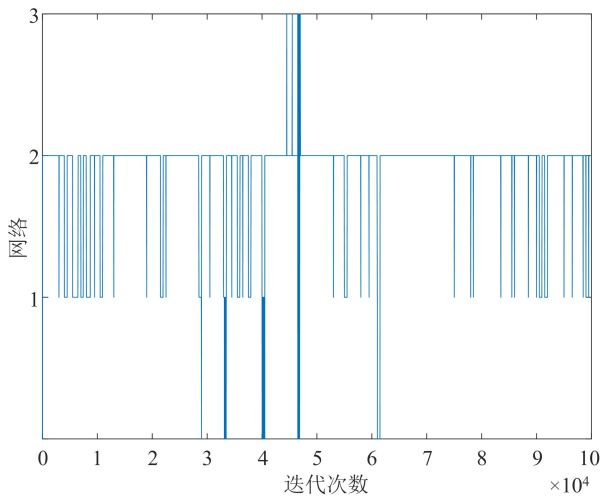


图 3 用户在无驻留定时器时各时刻连接网络

Fig. 3 The network that the user connects at each time without a dwell timer

CPE, 2 表明连接到 5G 无人机, 3 表明连接 Wi-Fi。可以看到,使用 3 s 的驻留定时器的网络切换次数相比不使用驻留定时器明显减少。例如,网络 1 和 2 间的切换,学习过程中切换次数减少。当不使用驻留定时器时,存在多处短时的、频繁的切换,而用 3 s 的驻留定时器后这样频繁的切换仅出现了 1 处。

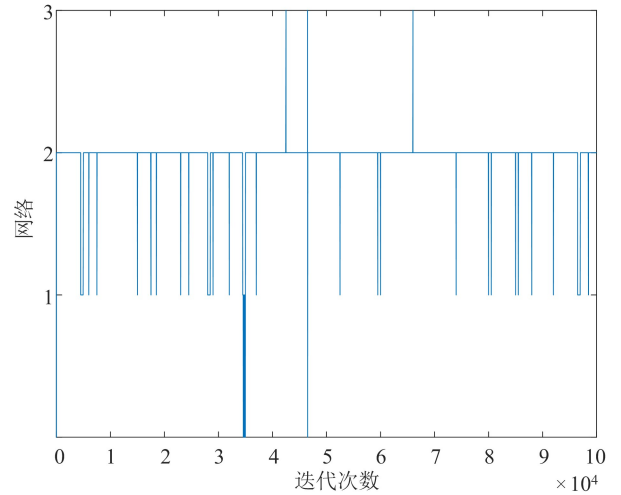


图 4 用户在驻留定时器为 3 s 时各时刻连接网络

Fig. 4 The network connected to the users at each time with a 3 s dwell timer

图 5~6 分别描述了不使用驻留定时器以及驻留定时器为 4 s 条件下用户在各时刻的 QoS 值。从图中可以看出,当没有使用驻留定时器时,用户的 QoS 变化不大。但当驻留定时器时长增大到 5 s 时, QoS 值大幅度下降的次数增多,且原先 QoS 较为平稳的时间段内也开始出现 QoS 频繁下降的情况。这是因为使用驻留定时器会牺牲一定程度的服务质量,这意味着降低切换次数和保证服务质量 QoS 值两者之间的权衡。

图 7 描述了语音、视频和文本 3 种业务类型中用户连接网络的比例。本文提出的异构网络垂直切换方法不仅减少了切换次数,避免网络选择功能受到影响,而且可以为具有不同业务需求的用户提供合适的网络选择。如图 7 所示,由于具有低时延的需求,语音类业务用户偏向选择 5G 无人机网络,因为其时延最短;而 Wi-Fi 由于其传输速率最快,可以满足视频类业务用户的需求;由于对网络的要求较低,文本类用户主要选择 5G 无人机网络和 Wi-Fi,这两种方式在该场景下差距不大,而此类用户受负载均衡

控制的影响较大。异构网络垂直切换方法在考虑负载均衡和垂直切换次数的情况下,仍可以根据不同的通信和连接需求,为不同类型的业务选择合适的网络。

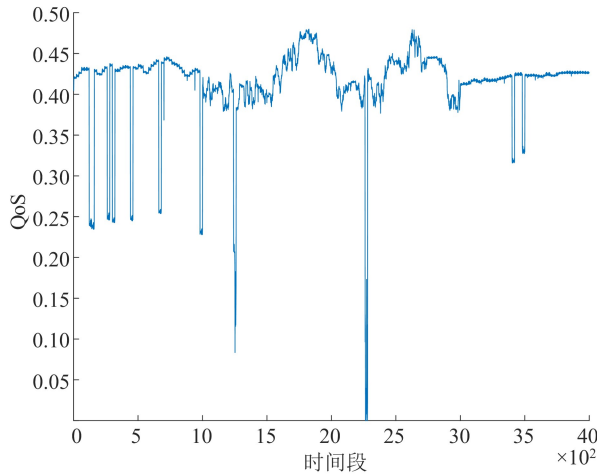


图 5 无驻留定时器时的用户各时刻 QoS

Fig. 5 QoS value at each time without dwell timer

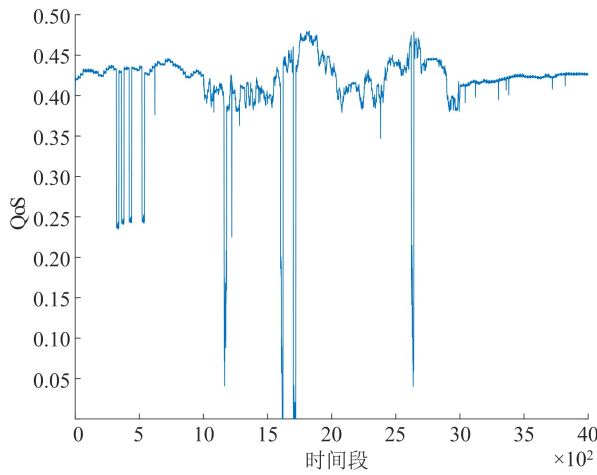


图 6 驻留定时器为 4 s 时的用户各时刻 QoS

Fig. 6 QoS value at each time with a 4 s dwell timer

图 8 表现了用户平均切换次数和驻留定时器时长之间的关系。其中 QL with ap 为在回报函数中考虑网络接入历史的网络选择算法,它考虑了当前接入网络,并基于 QL 的方法实现了网络选择。如图 8 所示,平均切换测试对着驻留定时器时长的增加而降低。相比于 QL 算法,在不同定时器时长的条件下,QL with ap 算法均能显著降低用户平均切换次数。这是由于用户在判决阶段,随着驻留定时器的时长增加,其移动的距离越长,导致原本选择的网络无法连接或无法提供最优服务质量的可能性增加,因此提高了取消

网络切换的概率。此外,在相同的驻留定时器时长下,QL with ap 相比 QL 算法可以更好地降低平均切换次数。

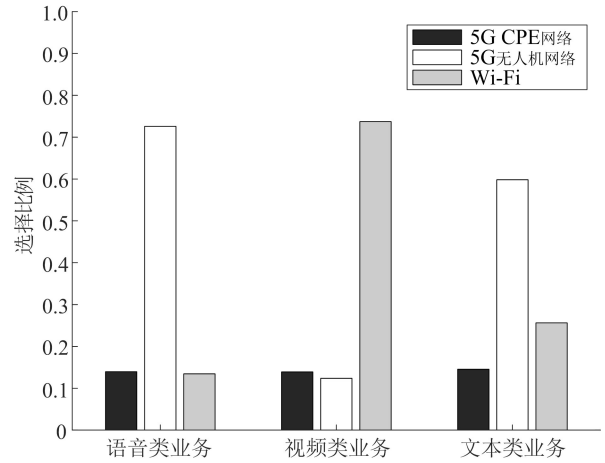


图 7 3 种业务类型用户选择网络比例

Fig. 7 Proportion of the selected networks for 3 business types

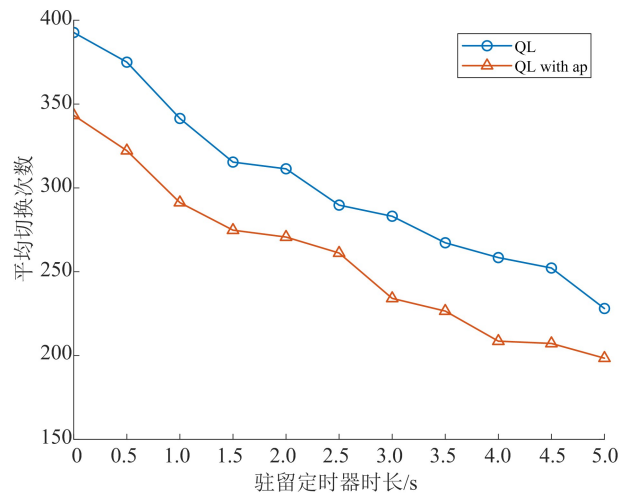


图 8 不同驻留定时器时长下的平均切换次数

Fig. 8 Average number of handovers with different dwell timers

驻留定时器时长为 3 s 的情况下,用户平均切换次数和用户数目关系如图 9 所示。用户平均切换次数随着用户数量的增加而递增。这是因为网络负载占用率由于用户数量的增加而增高,用户无法介入满负载的最优网络转而选择其他网络。这导致了垂直切换的再次出发,因此,切换次数会增加。或者说,网络满负载的概率随着用户数量的增多而提高,因此导致了高用户平均切换次数。此外,在使用/不使用驻留定时器的情况下,QL with ap 算法均优于 QL 算法。而在不同用户数目的情况下,有驻留定时器可以显著

地降低平均切换次数。而基于 QL 算法改进的 QL with ap 算法可以在 QL 算法的基础上进一步提升避免乒乓效应的效果。

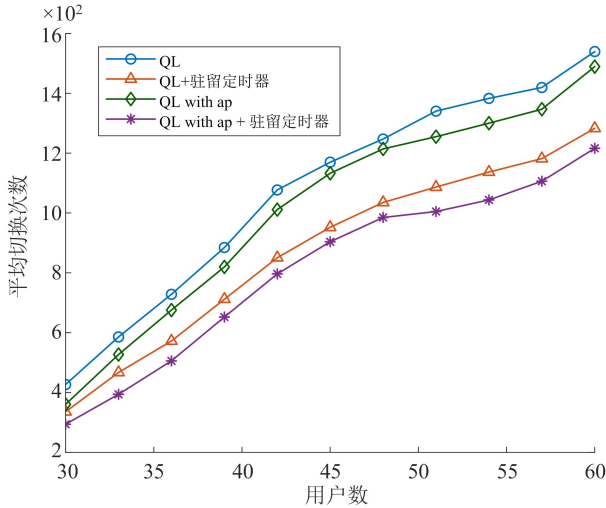


图9 不同用户数下的用户平均切换次数

Fig. 9 The average number of handovers with different numbers of users

4 结束语

本文对异构网络的切换问题,提出了一种基于强化学习的异构网络垂直切换方法。首先提出一个异构网络垂直切换系统,该系统中用户有较大的概率触发网络切换。接着从触发切换、网络选择和判决切换几个方面优化。在网络选择时提出基于 QL 的网络选择算法优化,更改状态和回报函数的设计,在切换判决时以 QoS 为条件,加入驻留定时器。仿真结果表明,本文提出的异构网络垂直切换方法在保证服务质量的条件下有效减少用户切换网络的次数,降低乒乓效应的影响。

参考文献

- [1] SU R, ZHANG D, VENKATESAN R, et al. Resource allocation for network slicing in 5G telecommunication networks: a survey of principles and models [J]. *IEEE Network*, 2019, 99: 1-8.
- [2] LI P, FANG Y G. On the throughput capacity of heterogeneous wireless networks[J]. *IEEE Transactions on Mobile Computing*, 2012, 11(12): 2073-2086.
- [3] DAMNJANOVIC A, MONTOJO J, WEI Y, et al. A survey on 3GPP heterogeneous networks [J]. *IEEE Wireless Communications*, 2011, 18(3): 10-21.
- [4] BEREZDIVIN R, BREINIG R, TOPP R. Next-generation wireless communications concepts and technologies[J]. *IEEE Communications Magazine*, 2002, 40(3): 108-116.
- [5] WANG C X, HAIDER F, GAO X Q, et al. Cellular architecture and key technologies for 5G wireless communication networks[J]. *IEEE Communications Magazine*, 2014, 52(2): 122-130.
- [6] 丁雨, 李晨凯, 韩会梅, 等. 基于 5G 无人机通信的多智能体异构网络选择方法[J]. *电信科学*, 2022, 38(8): 28-36.
DING Yu, LI Chenkai, HAN Huimei, et al. Multi-agent heterogeneous network selection method based on 5G UAV communication[J]. *Telecommunications Science*, 2022, 38(8): 28-36. (in Chinese)
- [7] MOZAFFARI M, SAAD W, BENNIS M, et al. A tutorial on UAVs for wireless networks: applications, challenges, and open problems[J]. *Communications Surveys & Tutorials*, 2019, 21(3): 2334-2360.
- [8] 尹博南, 彭木根, 刘晨熙. 无人机无线网络覆盖与切换性能分析[J]. *通信学报*, 2020(11): 22-29.
YIN Bonan, PENG Mugen, LIU Chenxi. Analysis of UAV wireless network coverage and handover performance[J]. *Journal on Communications*, 2020(11): 22-29. (in Chinese)
- [9] 张少伟, 朱雪田, 云翔. 5G 无人机异构网络的初始波束关联方案研究[J]. *电子技术应用*, 2020, 46(3): 33-37.
ZHANG Shaowei, ZHU Xuetian, YUN Xiang. Research on the initial beam correlation scheme for UAV-assisted 5G Het Net[J]. *Application of Electronic Technique*, 2020, 46(3): 33-37. (in Chinese)
- [10] LIN X S, XIA J J, WANG Z. Probabilistic caching placement in UAV-assisted heterogeneous wireless networks[J]. *Physical Communication*, 2019, 33: 54-61.
- [11] MARQUEZ-BARJA J M, AHMADI H, TORNELL S M, et al. Breaking the vehicular wireless communications barriers: vertical handover techniques for heterogeneous networks[J]. *IEEE Transactions on Vehicular Technology*, 2015, 64(12): 5878-5890.
- [12] SIDDIQUI F, ZEADALLY S. Mobility management across hybrid wireless networks: trends and challenges [J]. *Computer Communications*, 2006, 29(9): 1363-1385.
- [13] CHIN W H, FAN Z, HAINES R. Emerging technologies and research challenges for 5G wireless net-

- works[J]. IEEE Wireless Communications, 2014, 21(2): 106-112.
- [14] SONG W, JIANG H, ZHUANG W. Performance analysis of the WLAN-first scheme in cellular/WLAN interworking [J]. IEEE Transactions on Wireless Communications, 2007, 6(5): 1932-1952.
- [15] 吴利平, 王双双, 马彬. 改善用户体验的垂直切换算法[J]. 电子与信息学报, 2022, 44(8): 2824-2832.
WU Liping, WANG Shuangshuang, MA Bin. Vertical handoff algorithm for improving user experience [J]. Journal of Electronics & Information Technology, 2022, 44(8): 2824-2832. (in Chinese)
- [16] 马彬, 钟世林, 谢显中, 等. 考虑负载均衡和用户体验的垂直切换算法[J]. 电子与信息学报, 2022, 44(12): 4218-4228.
MA Bin, ZHONG Shilin, XIE Xianzhong, et al. Vertical handoff algorithm considering load balance and user experience[J]. Journal of Electronics & Information Technology, 2022, 44(12): 4218-4228. (in Chinese)
- [17] DAS S K, JAYARAM R, KAKANI N K, et al. A call admission and control scheme for quality-of-service (QoS) provisioning in next generation wireless networks[J]. Wireless Networks, 2000, 6: 17-30.
- [18] 周建明. 无线异构网络的垂直切换[J]. 通信技术, 2011, 44(5): 97-98.
ZHOU Jianming. Vertical handoff of wireless heterogeneous network [J]. Communications Technology, 2011, 44(5): 97-98. (in Chinese)
- [19] SGORA A, VERGADOS D D, CHATZIMISIOS P. An access network selection algorithm for heterogeneous wireless environments[C]//Proceedings of 2010 IEEE Symposium on Computers and Communications. [S.l. :s.n.], 2010: 890-892.
- [20] LEE W J, KIM E K, KIM J H, et al. Movement-aware vertical handoff of WLAN and mobile WiMAX for seamless ubiquitous access[J]. IEEE Transactions on Consumer Electronics, 2007, 53(4): 1268-1275.
- [21] CHEN Y H, CHANG C J, HUANG C Y. Fuzzy Q-learning admission control for WCDMA/WLAN heterogeneous networks with multimedia traffic[J]. IEEE Transactions on Mobile Computing, 2009, 8(11): 1469-1479.

作者简介

丁雨

女, 1998年生, 博士研究生, 研究方向为无人机通信、智能通信

E-mail: 1624879917@qq.com



李晨凯

男, 1996年生, 硕士研究生, 研究方向为异构网络接入、智能通信

E-mail: 1040526094@qq.com



卢为党

男, 1984年生, 博士, 教授, 博士研究生导师, 研究方向为无人机通信、移动边缘计算

E-mail: luweid@zjut.edu.cn



张清清

女, 1990年生, 博士, 副教授, 研究方向为无线通信网络智能化、物联网

E-mail: qingqingzhang@zjut.edu.cn



任元红

男, 1984年生, 高级工程师, 研究方向为指控系统设计

E-mail: 281829382@qq.com



高原

男, 1986年生, 博士, 副研究员, 研究方向为无人机通信

E-mail: yuangao08@tsinghua.edu.cn



责任编辑 董莉