



# 计算机科学

COMPUTER SCIENCE

## 基于边缘智能的频谱地图构建与分发方法

刘兴光, 周力, 刘琰, 张晓瀛, 谭翔, 魏急波

引用本文

刘兴光, 周力, 刘琰, 张晓瀛, 谭翔, 魏急波. [基于边缘智能的频谱地图构建与分发方法](#)[J]. 计算机科学, 2022, 49(9): 236-241.

LIU Xing-guang, ZHOU Li, LIU Yan, ZHANG Xiao-ying, TAN Xiang, WEI Ji-bo. [Construction and Distribution Method of REM Based on Edge Intelligence](#)[J]. Computer Science, 2022, 49(9): 236-241.

---

## 相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

### [基于值分解的多智能体深度强化学习综述](#)

Overview of Multi-agent Deep Reinforcement Learning Based on Value Factorization

计算机科学, 2022, 49(9): 172-182. <https://doi.org/10.11896/jsjcx.210800112>

### [智能博弈对抗方法:博弈论与强化学习综合视角对比分析](#)

Methods in Adversarial Intelligent Game:A Holistic Comparative Analysis from Perspective of Game Theory and Reinforcement Learning

计算机科学, 2022, 49(8): 191-204. <https://doi.org/10.11896/jsjcx.220200174>

### [基于多智能体强化学习的端到端合作的自适应奖励方法](#)

Adaptive Reward Method for End-to-End Cooperation Based on Multi-agent Reinforcement Learning

计算机科学, 2022, 49(8): 247-256. <https://doi.org/10.11896/jsjcx.210700100>

### [基于深度强化学习的边云协同资源分配算法](#)

Edge-Cloud Collaborative Resource Allocation Algorithm Based on Deep Reinforcement Learning

计算机科学, 2022, 49(7): 248-253. <https://doi.org/10.11896/jsjcx.210400219>

### [基于深度确定性策略梯度的服务器可靠性任务卸载策略](#)

Server-reliability Task Offloading Strategy Based on Deep Deterministic Policy Gradient

计算机科学, 2022, 49(7): 271-279. <https://doi.org/10.11896/jsjcx.210600040>

# 基于边缘智能的频谱地图构建与分发方法

刘兴光 周力 刘琰 张晓瀛 谭翔 魏急波

国防科技大学电子科学学院 长沙 410073

(liuxingguang20@nudt.edu.cn)

**摘要** 频谱地图可协助认知用户准确感知和利用频谱空洞,实现网络节点间的干扰协调,提升无线网络的频谱效率和鲁棒性。然而,当认知用户在利用和共享频谱地图时,面临着计算复杂度高和分发时延开销大的问题,限制了认知用户对空间频谱态势的实时感知能力。为了解决该问题,提出了一种边缘智能网络中基于强化学习的频谱地图构建与分发方法。首先,在频谱地图构建上,采用了一种克里金插值和超分辨率相结合的低复杂度构建技术;其次,通过引入边缘计算,将频谱地图构建与分发过程中的计算迁移策略选择问题建模为一个混合整数非线性规划问题;最后,将人工智能和边缘计算相结合,采用了一种集中式训练、分布式执行的强化学习框架,对不同网络场景下的频谱地图构建和分发策略进行学习。实验结果表明,所提方法具备良好的适应性,可有效降低频谱地图构建与分发的能耗和时延,支持认知用户在移动边缘网络场景下对频谱地图的近实时级应用。

**关键词:** 频谱地图;边缘智能;计算迁移;强化学习

**中图法分类号** TN915

## Construction and Distribution Method of REM Based on Edge Intelligence

LIU Xing-guang, ZHOU Li, LIU Yan, ZHANG Xiao-ying, TAN Xiang and WEI Ji-bo

College of Electronic Science and Technology, National University of Defense Technology, Changsha 410073, China

**Abstract** Radio environment map (REM) can assist cognitive users to accurately perceive and utilize spectrum holes, achieve interference coordination between network nodes, and improve the spectrum efficiency and robustness of wireless networks. However, when cognitive users utilize and share REM, there are problems of high computational complexity and high distribution delay overhead, which limit cognitive users' ability to perceive spatial spectrum situation in real time. To solve this problem, this paper proposes a reinforcement learning-based REM construction and distribution method in mobile edge intelligence networks. First, we employ a low-complexity construction technique that combines kriging interpolation and super-resolution for REM construction. Second, we model the computational offload strategy selection problem during REM construction and distribution as a mixed-integer nonlinear programming problem by using edge computing. Finally, we combine artificial intelligence technology and edge computing technology, and propose a centralized training, distributed execution reinforcement learning framework to learn REM construction and distribution strategies in different network scenarios. Simulation results show that the proposed method has good adaptability, and it can effectively reduce the energy consumption and delay of REM construction and distribution, and support the near real-time application of REM by cognitive users in mobile edge network scenarios.

**Keywords** Radio environment map, Edge intelligence, Computation migration, Reinforcement learning

## 1 引言

随着5G、物联网和智慧城市的快速推进,海量用频设备被投入使用,电磁频谱环境呈现快速复杂化的趋势,用频矛盾日益尖锐。电磁频谱作为网络信息传输的关键载体和重要资源,对人们的生活日趋重要<sup>[1]</sup>。频谱地图是认知复杂电磁环境的重要工具,它结合地理信息从时间、频率、空间和功率等多维对电磁环境进行描述<sup>[2]</sup>,可支持对电磁频谱资源分布

状况的有效获取,挖掘电磁频谱的多维度可复用价值,并以直观、灵活的方式辅助无线资源的管理,在认知无线电<sup>[3-5]</sup>、干扰管理<sup>[6]</sup>和预期网络中的主动资源分配<sup>[7-9]</sup>等领域已有较为广泛的应用。

传统的频谱地图应用主要是利用固定监测站、车载接收机、手持式频谱仪、无人机等带有频谱感知功能的设备采集频谱数据,并将数据上传到云服务器平台进行频谱地图构建和分析,然后将结果分发给认知用户(Cognitive User, CU)进行

到稿日期:2022-04-17 返修日期:2022-05-13

基金项目:国家自然科学基金(62171449,62001483,U19B2024)

This work was supported by the National Natural Science Foundation of China(62171449,62001483,U19B2024).

通信作者:周力(zhouli2035@nudt.edu.cn)

应用。但是频谱地图的构建是一种复杂型计算任务,同时频谱地图从云服务器分发给用户需要消耗较高的带宽和大量时间,在瞬息万变的电磁环境中,较大的时延阻碍了用户对频谱数据的实时级应用。

近年来,随着边缘智能的发展,融合边缘计算和人工智能两种技术有效地促进了边端协同<sup>[10]</sup>。移动边缘计算(Mobile Edge Computing, MEC)将最初部署在中心云上的各种功能部署在靠近用户的边缘服务器上,可以极大地减少传输时延<sup>[11]</sup>。同时以强化学习为代表的人工智能技术可以对边缘网络中的计算迁移和资源分配进行合理优化<sup>[12]</sup>,以提高边缘计算效率。文献[13]提出了一个空地一体化联邦学习框架,通过空地协同可以有效提高边缘智能网络中无人机等空中节点按需部署。文献[14]提出了两种基于多智能体的强化学习算法以协调优化边缘智能网络中的计算卸载和资源分配。边缘智能技术通过将计算、存储等资源向用户端推进的方式,赋予了网络边缘实体更强大的信息处理与内容投递能力,极大地改变着移动应用的功能与网络资源的利用模式,为构建与分发高带宽消耗、高能耗和时延敏感的频谱地图提供了新思路。

从频谱地图构建的角度来看,降低频谱地图构建的复杂度可以降低频谱地图的计算能耗和时延。文献[15]提出的基于卡尔曼滤波的方法、文献[16]提出的基于回归克里金以及增量聚类的方法,均降低了频谱地图构建的复杂度;文献[17]提出了一种基于超分辨率算法的频谱地图构建方法,该方法在提高频谱地图精度的同时大幅缩短了构建时间。

从频谱地图分发的角度来看,在边缘智能网络中优化计算迁移和资源分配可以降低频谱地图在传输过程中的能耗和时延。文献[18]使用深度强化学习算法对边缘网络中视频流传输过程中的能量消耗和用户体验进行同步优化;文献[19]提出了一种基于边缘网络的可感知用户体验的视频流对等迁移方法,该方法将视频流的传输、计算和迁移问题转化为用户体验最大化问题,并使用反模糊粒子群优化算法来进行优化;文献[20]使用分布式边缘计算优化视频流传输过程中的带宽消耗。以上边缘智能网络中视频流传输的方法虽然取得了不错的效果,但是并不适用于计算复杂度高、带宽需求高、时延敏感性强的频谱地图。频谱地图构建与分发过程中资源有限及时延敏感的问题亟待解决。

本文将频谱地图的构建和分发相结合,在运用文献[17]中低复杂度的频谱地图构建方法基础上,对边缘网络中的计算迁移和资源分配进行优化,可以同时降低频谱地图构建过程和分发过程中的能耗和时延。本文的主要贡献如下:

(1)采用克里金插值技术和超分辨率技术相结合的方式,根据带宽资源和计算资源限制,将频谱地图构建任务分为边缘构建、部分迁移和本地构建3种分发方式。

(2)将频谱地图构建与分发过程中的计算与迁移过程进行建模,定义为一个混合整数非线性规划(Mixed Integer Nonlinear Programming, MINLP)问题,优化联合迁移和资源分配策略,最小化计算、通信总能耗和时间的加权和。

(3)针对以上建模问题,提出了一种基于集中式训练、分布式执行强化学习框架的频谱地图构建与分发算法,提出了一种基于深度Q网络(Deep Q Network, DQN)的优化算法来对该问题进行优化。

## 2 系统模型和问题表述

### 2.1 网络模型

本文考虑了一个带有边缘服务器的基站和 $n$ 个CU的MEC网络场景,如图1所示。CU集合用 $N = \{1, 2, \dots, n\}$ 表示。CU具有射频接收器,可以获取频谱数据,然后将频谱数据和位置信息发送给基站。为了从整体上降低频谱地图构建与分发成本,频谱地图构建任务被分解迁移到边缘服务器和CU中,从而将分发地图的任务转化为分发模型和少量数据的任务,实现用算力换取带宽的效果。边缘服务器可以根据无线链路带宽和CU计算资源确定是否需要将频谱地图构建任务下行迁移到CU上。分发方式可用迁移决策变量 $m_i \in \{0, 1, 2\}$ 表示,分为以下3种。

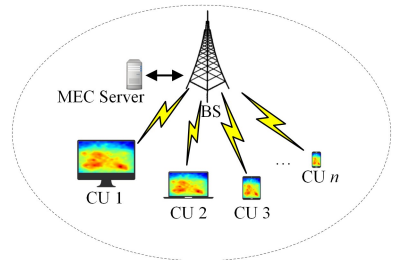


图1 MEC网络模型

Fig. 1 MEC network model

(1)边缘构建(mode 0,  $m_i = 0$ ):在该模式下,频谱地图构建任务仅发生在边缘服务器侧。首先,由边缘服务器完成高分辨率频谱地图的构建;然后,根据CU的精度需求和带宽限制,边缘服务器将频谱地图进行降分辨率及压缩处理后分发给CU。该模式适用于分发链路带宽充裕或CU处理能力不足的情况。

(2)部分迁移(mode 1,  $m_i = 1$ ):在该模式下,CU需要执行部分频谱地图构建任务。首先,由边缘服务器完成低分辨率频谱地图的构建和超分辨率模型的训练;然后,边缘服务器将压缩好的低分辨率频谱地图和超分辨率模型分发给CU,CU仅需完成超分辨率转化任务。该模式适用于分发链路带宽充裕且CU具备一定处理能力的情况。

(3)本地构建(mode 2,  $m_i = 2$ ):在该模式下,CU利用边缘服务器分发的所有CU的原始频谱感知数据以及训练好的超分辨率模型,完成低分辨率频谱地图的构建和超分辨率模型转化任务。该模式适用于分发链路带宽受限或CU具备一定处理能力的情况。

本文采用 $\tau_i = \{s_i, c_i, T_i^{\max}\}$ 模型来描述CU $i$ ( $i \in N$ )的频谱地图构建计算任务。其中, $s_i$ 表示计算任务的输入数据量大小, $c_i$ 代表任务执行所需的CPU周期数, $T_i^{\max}$ 是最大容限等待时间。我们认为计算任务是可以分解的,频谱地图构建任务可以分解为低分辨率频谱地图构建和超分辨率转化两个

子任务。在 mode 0 和 mode 1 中, 还需考虑频谱地图的压缩任务。MEC 网络中的计算任务如表 1 所列。

表 1 MEC 网络中的计算任务

Table 1 Computation tasks in MEC networks

计算任务	输入数据量	CPU 周期
原始数据—>低分辨率地图	$s_{i,1}$	$c_{i,1}$
低分辨率地图—>高分辨率地图	$s_{i,2}$	$c_{i,2}$
频谱地图—>压缩后的频谱地图	$s_{i,3}$	$c_{i,3}$

## 2.2 构建与分发模型

### 2.2.1 边缘服务器侧

假设边缘服务器分配给 CU  $i$  的 CPU 频率为  $F_i$ 。本文计算能耗的表达式为:

$$e_i^{\text{server}} = k F_i^\alpha c_i^{\text{server}} \quad (1)$$

其中,  $k$  是 CPU 的有效开关电容, 由 CPU 硬件体系结构所确定, 典型值为  $k = 10^{-27}$ ,  $c_i^{\text{server}}$  为在服务器端执行计算任务的 CPU 周期数, 根据分发模式确定。在 mode 0 下, 在边缘服务器进行将原始数据构建为低分辨率频谱地图、将低分辨率频谱地图构建为高分辨率频谱地图以及频谱地图的压缩任务, CPU 周期数为  $c_i^{\text{server}} = c_{i,1} + c_{i,2} + c_{i,3}$ 。在 mode 1 下, 在边缘服务器进行将频谱数据构建为低分辨率频谱地图和频谱地图的压缩任务, CPU 周期数为  $c_i^{\text{server}} = c_{i,1} + c_{i,3}$ 。在 mode 2 下, 在边缘服务器不进行计算, 因此  $c_i^{\text{server}} = 0$ 。在服务器端的计算时间为:

$$t_i^{\text{server}} = c_i^{\text{server}} / F_i \quad (2)$$

### 2.2.2 CU 侧

CU  $i$  的 CPU 频率为  $f_i$ , 计算能耗为:

$$e_i^{\text{local}} = k f_i^\alpha c_i^{\text{local}} \quad (3)$$

其中,  $c_i^{\text{local}}$  为在 CU  $i$  执行的计算任务的 CPU 周期数, 由分发模式确定。在 mode 0 下, CU 没有参与建图计算任务, 其 CPU 周期数为  $c_i^{\text{local}} = 0$ 。在 mode 1 下, CU 进行了超分辨率模型转化的计算任务, CPU 周期数为  $c_i^{\text{local}} = c_{i,2}$ 。在 mode 2 下, CU 完成了低分辨率频谱地图构建和超分辨率模型转化两个任务, 因此  $c_i^{\text{local}} = c_{i,1} + c_{i,2}$ 。CU 的计算时间为:

$$t_i^{\text{local}} = c_i^{\text{local}} / f_i \quad (4)$$

### 2.2.3 通信模型

基站到 CU 的数据传输速率可以表示为:

$$r_i = W_i \log_2 \left( 1 + \frac{p h_i}{\sigma^2} \right) \quad (5)$$

其中,  $W_i$  表示给 CU  $i$  分配的信道带宽;  $p$  表示基站的发射功率, 恒定不变;  $h_i$  表示 CU  $i$  与基站之间的信道增益;  $\sigma^2$  是背景噪声功率。数据传输延时可以表示为:

$$t_i^{\text{trans}} = d_i / r_i \quad (6)$$

其中,  $d_i$  为基站到 CU  $i$  之间的下行传输数据, 根据分发模式确定。在 mode 0 下, 基站分发的是压缩后的地图, 传输数据量为  $d_i = d_i^{\text{mode } 0}$ ; 在 mode 1 下, 基站分发的是低分辨率频谱地图, 传输数据量为  $d_i = d_i^{\text{mode } 1}$ ; 在 mode 2 下, 基站分发的是原始频谱感知数据, 传输数据量为  $d_i = d_i^{\text{mode } 2}$ , 如表 2 所列。基站传输数据发射能耗可以表示为:

$$e_i^{\text{trans}} = p t_i^{\text{trans}} \quad (7)$$

表 2 不同模式下的计算与分发数据

Table 2 Computation and distribution of data in different modes

迁移模式	边缘服务器计算量	CU 计算量	传输数据量
mode 0	$c_i^{\text{server}} = c_{i,1} + c_{i,2} + c_{i,3}$	$c_i^{\text{local}} = 0$	$d_i^{\text{mode } 0}$
mode 1	$c_i^{\text{server}} = c_{i,1} + c_{i,3}$	$c_i^{\text{local}} = c_{i,2}$	$d_i^{\text{mode } 1}$
mode 2	$c_i^{\text{server}} = 0$	$c_i^{\text{local}} = c_{i,1} + c_{i,2}$	$d_i^{\text{mode } 2}$

## 2.3 问题描述

CU  $i$  的频谱地图构建与分发的计算和通信的总能量开销可表示为:

$$E_i = \omega_{\text{server}}^e e_i^{\text{server}} + \omega_{\text{local}}^e e_i^{\text{local}} + \omega_{\text{trans}}^e e_i^{\text{trans}} \quad (8)$$

其中  $\omega_{\text{server}}^e$ ,  $\omega_{\text{local}}^e$ ,  $\omega_{\text{trans}}^e$  分别为边缘服务器计算、CU 计算和通信能量消耗的权重。计算与通信的总时间表示为:

$$T_i = \omega_{\text{server}}^t t_i^{\text{server}} + \omega_{\text{local}}^t t_i^{\text{local}} + \omega_{\text{trans}}^t t_i^{\text{trans}} \quad (9)$$

其中  $\omega_{\text{server}}^t$ ,  $\omega_{\text{local}}^t$ ,  $\omega_{\text{trans}}^t$  分别为边缘服务器计算、CU 计算和通信时间的权重。

本文以最小化边缘网络的能耗和时延为目标, 对频谱地图构建与分发过程中计算卸载和资源卸载问题进行优化。基于以上计算模型和通信模型, 频谱地图分发中的能量和时间的优化问题可建模为:

$$\begin{aligned} & \min_{m, W, F} \sum_{i=1}^M [(1-\omega)E_i + \omega T_i] \\ & \text{s. t. } C1: m_i \in \{0, 1, 2\} \\ & \quad C2: k f_i^2 c_i^{\text{local}} \leq e_i^{\text{local}} \\ & \quad C3: \sum_i F_i \leq F_{\max} \\ & \quad C4: \sum_i W_i \leq W_{\max} \\ & \quad C5: 0 \leq T_i^{\text{sum}} \leq T_{\max} \end{aligned} \quad (10)$$

其中,  $\omega$  为总时间的权重约束条件, C1 表示 CU 可选择的迁移方式, C2 是确保任务执行所消耗的能量不能超过 CU 设备的剩余能量, C3 表示边缘服务器分配给所有 CU 的计算资源不能超过边缘服务器的最大计算资源, C4 表示分配给所有 CU 的带宽之和不能超过边缘网络的最大可用带宽, C5 表示地图构建与分发的时间不超过任务最大容忍时间,  $T_i^{\text{sum}}$  表示给 CU  $i$  构建与分发频谱地图的总时间:

$$T_i^{\text{sum}} = t_i^{\text{server}} + t_i^{\text{local}} + t_i^{\text{trans}} \quad (11)$$

为了求解式(10)中的上述问题, 需要找到迁移决策向量  $m = \{m_1, m_2, \dots, m_n\}$ 、带宽资源分配向量  $W = \{W_1, W_2, \dots, W_n\}$ 、计算资源分配向量  $F = \{F_1, F_2, \dots, F_n\}$  在给定的时延约束条件下的最优结果, 使系统的总能量消耗和时延最小化。值得注意的是, 迁移决策变量是整数变量, 而资源分配变量  $W_i$  和  $F_i$  是动态变化的。系统需要收集大量的网络状态信息, 并根据网络的当前状态对计算迁移和资源分配进行全局决策。因此, 目标函数是一个混合整数非线性规划问题, 这无疑是 NP 难的问题。该问题的可行集不是凸的, 复杂度总是随着 CU 数量的增加呈指数增长。由于传统的基于模型的方法可能无法适应动态特性并做出智能决策, 因此本文提出了基于强化学习的方法来解决上述问题。

### 3 基于 DQN 的联合迁移和资源分配算法

#### 3.1 强化学习框架

强化学习(Reinforcement Learning, RL)作为人工智能领域的一部分,是一种通过在智能体和环境的交互过程中不断试错和学习得到最大收益策略的方法。因此本文提出了一个集中式训练、分布式执行的强化学习框架。在边缘服务器进行集中式训练,将训练好的策略分发给 CU, CU 分布式执行的策略并在执行过程中与环境交互。同时,该强化学习框架包括在线执行模块和离线训练模块,前者利用学习到的迁移策略进行频谱地图的分步构建,而后者则根据 CU 的迁移计算结果对前者进行动态更新,如图 2 所示。

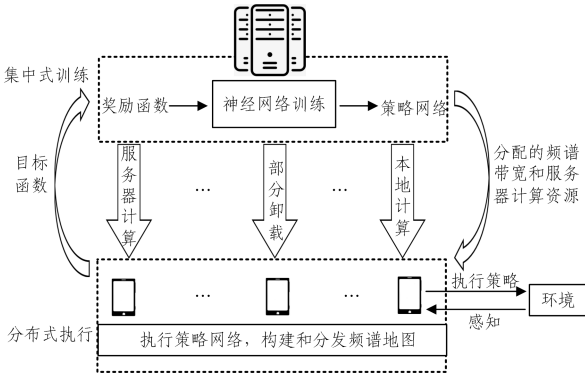


图 2 强化学习框架

Fig. 2 Framework of reinforcement learning

#### 3.2 强化学习的 3 要素

根据本文提出的强化学习框架,设计了一种基于 DQN 的任务迁移和资源分配算法,该算法的状态空间、动作空间和奖励空间如下。

##### 3.2.1 状态空间

$$S = \{s_t | s_t = (e_t^{\text{local}}, F_t^{\text{remain}}, W_t^{\text{remain}}), t \in M\} \quad (12)$$

其中,  $s_t$  表示在时间步  $t$  时的网络状态。在本文的系统中,在  $t$  时的可用资源状态由  $F_t^{\text{remain}} = F_{\text{max}} - \sum_i F_i$  和  $W_t^{\text{remain}} = W_{\text{max}} - \sum_i W_i$  表示,其中前者是边缘服务器的可用计算资源,而后者是当前网络可用的频谱带宽资源。观察它们的目的是保证满足计算容量和通信信道容量的约束。此外,本文还需要观察每个时刻 CU 的剩余能量  $e_t^{\text{remain}} = \{E_i^{\text{local}} - k(f_i)^2 c_i^{\text{local}}, i \in N\}$ , 以免 CU 的能量不足以完成下一时段分配的计算任务。

##### 3.2.2 动作空间

$$A = \{a_t | a_t = (m_t, F_t, W_t), t \in M\} \quad (13)$$

在本文提出的 MEC 系统中,边缘服务器负责决定每个时间段内各个 CU 的迁移策略。此外,还需要确定每个时刻分配给 CU 的计算和通信资源。因此,动作空间由 3 部分组成:迁移决策向量  $m_t = \{m_1^t, m_2^t, \dots, m_n^t\}$ , 计算资源分配向量  $F_t = \{F_1^t, F_2^t, \dots, F_n^t\}$ , 频谱资源分配向量  $W_t = \{W_1^t, W_2^t, \dots, W_n^t\}$ 。为了简化动作空间,分配给 CU  $i$  的计算资源是根据  $c_i^{\text{server}}$  在边缘服务器中总计算量的占比分配的,而分配给 CU  $i$  的带宽资源是根据  $d_i$  在网络中总传输数据量的占比分配的。

##### 3.2.3 奖励函数

一般来说,即时网络奖励函数与目标函数有关。假设在  $t-1$  时刻系统的总消耗为  $cost_{t-1}$ , 在  $t$  时刻系统的总消耗为  $cost_t$ , 那么在  $t$  时刻执行策略  $a_t$  得到的奖励为:

$$r_t = cost_{t-1} - cost_t, t \in M \quad (14)$$

在已考虑的优化问题中,本文的目标是使整个 MEC 系统的总消耗最小化,这是强化学习的反目标,试图实现最大的累计折扣奖励(Cumulative Discounted Reward, CDR),即:

$$C = \sum_{t=1}^M \gamma^{t-1} \cdot r_t = \sum_{t=1}^M \gamma^{t-1} \cdot (cost_{t-1} - cost_t) \quad (15)$$

当折扣因子  $\gamma=1$  时,  $C = cost_0 - cost_M$ , 其中  $cost_0 = 0$  表示频谱地图构建与分发任务未开始时系统总消耗为 0。因此  $C = -cost_M$ , 即实现最大的累计折扣奖励和实现系统的总消耗最小化的目标是一致的。

#### 3.3 基于 DQN 的联合迁移和资源分配算法

传统的 Q-Learning 算法使用 Q 表来存储  $t$  时刻的状态、动作下所获得的 Q 值,通过遍历 Q 表,可以得到在  $s_t$  状态下奖励最大的动作。然而随着状态和动作空间的增加,存储 Q 值所需的内存空间将呈指数型增长,产生维度灾难。为了缓解这一问题, DQN 采用神经网络来近似估计不同状态动作组合下的 Q 值,将 Q 表的更新问题变成一个函数拟合问题。因此,我们设计了一种基于 DQN 的计算迁移和资源分配算法,如算法 1 所示。

##### 算法 1 基于 DQN 的联合迁移和资源分配算法

输入: 评估网络 Q, 参数  $\theta$ , 目标网络  $\bar{Q}$ , 参数  $\bar{\theta}$ , 经验池  $\mathcal{B}$   
输出:  $(m, F, W) / *$  最优的联合迁移和资源分配策略  $*$  /

1. 初始化: 初始化经验池  $\mathcal{B}$
2. For  $k=1$  to EPISODE do
3. 初始化状态  $s_t = (e_t^{\text{local}}, F_t^{\text{remain}}, W_t^{\text{remain}})$ ;
4. For step  $t=1$  to  $M$  do
5. 采用  $\epsilon$ -贪婪策略来选择动作  $a_t$ ;
6. 执行动作  $a_t$ , 观察下一状态  $s_{t+1}$  和奖励  $r_t$ , 转移样本  $\phi_t = (s_t, a_t, r_t, s_{t+1})$  存入经验池  $\mathcal{B}$ ;
7.  $s_t = s_{t+1}$
8. 从  $\mathcal{B}$  中随机抽取小批量样本集  $\Pi$ ;
9. For  $(s_j, a_j, r_j, s_{j+1})$  in  $\Pi$  do
10. If 转移终止 then 设置  $y_j = r_j$ ;
11. Else 设置  $y_j = r_j + \gamma \max_a \bar{Q}(s_{j+1}, a'; \bar{\theta})$ ;
12. 在  $(y_j - Q(s_j, a'; \theta))^2$  上执行梯度下降来更新评估网络参数  $\theta$ ;
13. 每隔  $\bar{M}$  步, 更新目标网络参数:  $\bar{\theta} = \theta$ ;
14. End for
15. End for

边缘服务器将服务器的可用计算资源、网络可用的频谱带宽资源和 CU 的剩余能量组合成系统状态输入到 DQN 网络中, 采用  $\epsilon$ -贪婪策略来选择动作(步骤 5), 即:

$$\alpha_t = \begin{cases} \text{在 } A \text{ 中随机选择,} & \text{Rand}(0,1) < \epsilon \\ \arg \max_{a \in A} Q(s_t, a_t), & \text{其他} \end{cases} \quad (16)$$

其中,  $\text{Rand}(0,1)$  表示  $(0,1)$  之间的随机数,  $\epsilon \in (0,1)$  是该策略

的参数,可以使本文的算法快速收敛,同时不失随机性。根据  $\alpha_t$  进行计算迁移和资源分配,得到下一状态  $s_{t+1}$  和奖励  $a_t$ ,将  $(s_t, a_t, r_t, s_{t+1})$  存入经验池  $\mathcal{B}$  (步骤 6)。然后从  $\mathcal{B}$  中随机抽取小批量样本集 (步骤 8),更新网络参数 (步骤 9—步骤 13)。

## 4 仿真结果

### 4.1 参数设置

本文实验中,考虑一个带有边缘服务器的基站和 5 个 CU 的 MEC 网络。假设通信带宽  $W=10$  MHz, CU 随机分布在距离基站  $[50, 200]$  米的区域内。边缘服务器的计算能力为  $F=5$  GHz,每个 CU 的 CPU 频率在  $[0.5, 1.5]$  GHz 范围随机分布。频谱地图构建与分发的最大容忍时延  $T_{\max}=1000$  ms。基站的发射功率为  $P_n=500$  mW。

### 4.2 结果分析

为了评估所提方案的性能,将其与其他 3 种方案进行了比较:1)边缘构建,所有 CU 的迁移决策变量  $m_i=0$ ,选择在边缘服务器进行频谱地图的构建,并分发压缩的频谱地图;2)本地构建,所有 CU 的迁移决策变量  $m_i=2$ ,选择分发全局的频谱信息并在用户设备上完成频谱地图的构建;3)随机迁移,所有 CU 在服务器构建、CU 构建、部分迁移 3 种构建方式中随机选择一种进行计算迁移,即  $m_i \in \{0, 1, 2\}$ ;4)基于 Q-learning 方法,使用基于 Q-learning 的强化学习算法,得到计算迁移和资源分配的最佳策略<sup>[21]</sup>。

图 3 给出了随着 CU 总数的增加的 MEC 系统的总成本,其中边缘服务器的计算能力为  $F=5$  GHz。整体来看,当 CU 数量不断增加时,4 种方案的总成本增加,两种强化学习方案的总消耗相对较小,这是因为两种方案都对服务器计算资源和网络带宽资源进行了分配。然而由于基于 Q-learning 的方案在迭代时需要遍历 Q 表,随着 CU 数量的增加,动作空间呈指数增加,所提基于 DQN 的方案引入神经网络得到了不同状态下的动作 Q 值,可以加快收敛,因此总消耗更小。

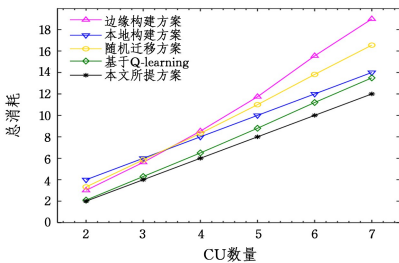


图 3 CU 数量对总消耗的影响

Fig. 3 Effect of number of CUs on total consumption

图 4 给出了 MEC 系统相对于边缘服务器计算能力增加的总成本,其中 CU 的数量为 5。从图 3 可以看出,本文提出的基于 DQN 的算法可以达到最佳效果,本地构建的曲线没有随着边缘服务器计算资源的增加而变化,这是因为 CU 没有使用边缘服务器的计算资源。其他曲线随着边缘服务器计算能力的增加而减小,因为每个 CU 分配到更多的服务器计算资源,计算时间会相应变短。此外,当  $F>8$  GHz 时,边缘服务器构建、基于 Q-learning 的方案和本文方案的总成本下降缓慢。结果表明,当边缘服务器上的计算资源远远多于

CU 的计算资源时,边缘网络的总消耗主要受网络带宽资源等其他因素的限制。

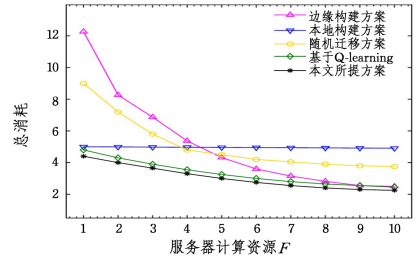


图 4 服务器计算能力对总消耗的影响

Fig. 4 Effect of server computing power on total consumption

图 5 给出了 MEC 系统相对于网络带宽增加的总成本,其中 CU 的数量为 5,服务器的计算能力为 8 GHz。图中边缘构建的曲线没有随着网络总带宽的增加而变化,这是因为边缘构建没有使用带宽资源进行迁移。其他曲线随着网络总带宽的增加而减小,因为每个 CU 可以分配到更多的带宽资源,通信传输时间会变短。从图 5 可以看出,两种强化学习方案优于其他方案,与基于 Q-learning 的方案相比,本文提出的 DQN 方法可以训练得到更好的策略。

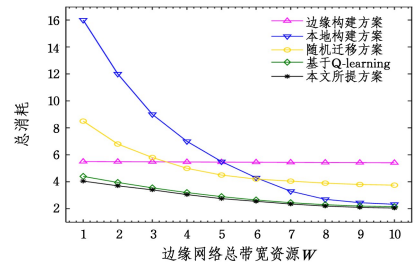


图 5 网络带宽对系统总消耗的影响

Fig. 5 Effect of network bandwidth on total consumption

图 6 给出了在 CU 的数量为 5、边缘服务器的计算能力为 8 GHz、网络带宽为 10 MHz 的场景下向每个 CU 分发频谱地图的平均时延。所提方案的时延为 358 ms,比边缘构建方案减少了 62.63%,比本地构建方案减少了 40.33%,比随机迁移方案减少了 48.85%,比基于 Q-learning 方案减少了 17.65%,可以满足近实时级应用的要求。

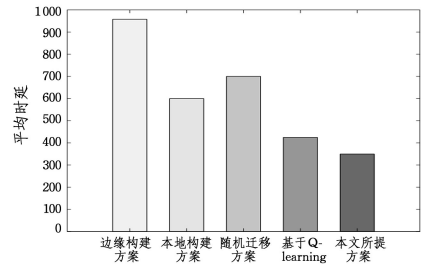


图 6 频谱地图分发的平均时延

Fig. 6 Average delay of REM distribution

综上所述,在解决频谱地图构建与分发的问题中,与其他方案相比,本文提出的基于 DQN 的联合迁移和资源分配算法可以同时最小化计算和传输的能耗和时延。由于所提算法对边缘网络中的资源进行了合理分配,因此随着计算资源和带宽资源等环境的变化,所提算法体现出了良好的适应性。

此外,优化后的策略支持 CU 对频谱地图的近实时级应用。

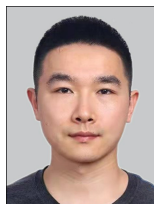
**结束语** 频谱地图可为认知用户获取空间频谱态势提供重要数据基础,然而,其构建复杂度和分发时延限制了其实时应用。鉴于此,本文将频谱地图构建与分发任务建模为一个混合整数非线性规划问题。为了解决该问题,本文提出了一种强化学习框架,采用基于 DQN 的联合迁移和资源分配算法对问题进行求解。通过与其他方案进行实验对比,验证了所提方法具备良好的适应性,可有效降低频谱地图构建与分发的能耗和时延,大大降低认知用户获取频谱地图的成本。

### 参考文献

- [1] XIA H Y, ZHA S, HUANG J J, et al. Survey on the Construction Methods of Spectrum Map[J]. Chinese Journal of Radio Science, 2020, 35(4): 12.
- [2] KATAGIRI K, FUJII T. Mesh-Clustering-Based Radio Maps Construction for Autonomous Distributed Networks[C]// 2021 Twelfth International Conference on Ubiquitous and Future Networks(ICUFN). IEEE, 2021: 345-349.
- [3] BEDNAREK P, ŁOPATKA J, BICKI D. Radio environment map for the cognitive radio network simulator[J]. International Journal of Electronics and Telecommunications, 2018, 64(1): 45-49.
- [4] EZZATI N, TAHERI H. Distributed spectrum sensing in rem based cognitive radio networks[J]. Journal of Modeling in Engineering, 2019, 17(56): 223-233.
- [5] KANIEWSKI P, ROMANIK J, GOLAN E, et al. Spectrum Awareness for Cognitive Radios Supported by Radio Environment Maps: Zonal Approach[J]. Applied Sciences, 2021, 11(7): 2910.
- [6] SANTANA Y H, PLETS D, ALONSO R M, et al. Tool for Recovering after Meteorological Events Using a Real-Time REM and IoT Management Platform[J]. Wireless Communications and Mobile Computing, 2019, 2019(1): 1-13.
- [7] YIN L, LUO J, LUO H. Tasks scheduling and resource allocation in fog computing based on containers for smart manufacturing[J]. IEEE Transactions on Industrial Informatics, 2018, 14(10): 4712-4721.
- [8] KIBRIA M G, NGUYEN K, VILLARDI G P, et al. Big data analytics, machine learning, and artificial intelligence in next-generation wireless networks[J]. IEEE Access, 2018, 6: 32328-32338.
- [9] ATAWIA R, HASSANEIN H S, ALI N A, et al. Utilization of stochastic modeling for green predictive video delivery under network uncertainties[J]. IEEE Transactions on Green Communications and Networking, 2018, 2(2): 556-569.
- [10] LI M S, GAO J, ZHAO L, et al. Deep reinforcement learning for collaborative edge computing in vehicular networks[J]. IEEE Transactions on Cognitive Communications and Networking, 2020, 6(4): 1122-1135.
- [11] SHI W, CAO J, ZHANG Q, et al. Edge computing: Vision and challenges[J]. IEEE Internet of Things Journal, 2016, 3(5): 637-646.
- [12] KHORAMNEJAD F, EROL-KANTARCI M. On Joint Offloading and Resource Allocation: A Double Deep Q-Network Approach[J]. IEEE Transactions on Cognitive Communications and Networking, 2021, 7(4): 1126-1141.
- [13] QU Y B, DONG C, ZHENG J C, et al. Empowering Edge Intelligence by Air-Ground Integrated Federated Learning[J]. IEEE Network, 2021, 35(5): 34-41.
- [14] LIU Z C, ZHAO Y F, SONG J D, et al. Learn to Coordinate for Computation Offloading and Resource Allocation in Edge Computing: A Rational-based Distributed Approach [J]. IEEE Transactions on Network Science and Engineering, 2021, 1(1): 1-15.
- [15] CHOWDAPPA V P, BOTELLA C, SAMPER-ZAPATER J J, et al. Distributed radio map reconstruction for 5G automotive [J]. IEEE Intelligent Transportation Systems Magazine, 2018, 10(2): 36-49.
- [16] SERGEEV A P, TARASOV D A, BUEVICH A G, et al. High variation subarctic topsoil pollutant concentration prediction using neural network residual kriging[J]. AIP Conference Proceedings, 2017, 1836(1): 020023.
- [17] DENG Y B, ZHOU L, WANG L, et al. Radio environment map construction using super-resolution imaging for intelligent transportation systems[J]. IEEE Access, 2020, 8: 47272-47281.
- [18] DU J, YU F R, LU G, et al. MEC-assisted immersive VR video streaming over terahertz wireless networks: A deep reinforcement learning approach[J]. IEEE Internet of Things Journal, 2020, 7(10): 9517-9529.
- [19] TAHA A E M, ALI N A, CHI H R, et al. MEC resource offloading for QoE-aware HAS video streaming[C]// ICC 2021-IEEE International Conference on Communications. IEEE, 2021: 1-5.
- [20] GENDA K, ABE M, KAMAMURA S. Video Communication Optimization Using Distributed Edge Computing[C]// 2020 21st Asia-Pacific Network Operations and Management Symposium (APNOMS). IEEE, 2020: 381-384.
- [21] JIANG K, ZHOU H, LI D, et al. A Q-learning based Method for Energy-Efficient Computation Offloading in Mobile Edge Computing[C]// 2020 29th International Conference on Computer Communications and Networks (ICCCN). IEEE, 2020: 1-7.



**LIU Xing-guang**, born in 1998, post-graduate. His main research interests include radio environment map and mobile edge computing.



**ZHOU Li**, born in 1988, Ph.D, master supervisor. His main research interests include intelligent communication network, wireless resource management and edge computing.