文章编号:1671-4229(2024)01-0084-13

基于多载波非正交多址接入的协作移动边缘计算系统优化研究

罗至威,黄栩蔚,叶 炜,黄高飞*
(广州大学 电子与通信工程学院,广东 广州 510006)

摘 要:文章研究了基于多载波非正交多址接入(Multicarrier Non-Orthogonal Multiple Access, MC-NOMA)的协作移动边缘计算系统,该系统由一个用户端、一个协作节点与一个配备边缘计算服务器的接入点(Access Point, AP)组成,旨在给定任务执行时延约束下,最小化用户端与协作节点的通信与计算能耗。为此,首先提出一种新颖的MEC协议,使用户端和协作节点充分利用MC-NOMA技术以高频谱效率的方式进行任务卸载;然后在该协议下,建立系统优化问题,联合优化任务卸载时间与计算时间分配、子载波分配、各个子载波的发射功率分配、任务划分和MEC服务器CPU频率。该问题是一个非凸的混合整数规划问题,求解颇具挑战性。通过利用代数变换、大M表达法、分层优化、连续凸近似和黄金分割法,对该问题进行求解,从而提出一种高效的系统优化算法。仿真结果表明,与已有的仅采用多载波技术研究方案相比,采用MC-NOMA技术可节省50%~60%的能耗。

关键词:移动边缘计算;协作通信;非正交多址接入;多载波传输;凸优化
中图分类号:TN92 **文献标志码**:A

Optimization algorithm for cooperative mobile edge computing systems with multicarrier non-orthogonal multiple access

*LUO Zhi-wei, HUANG Xu-wei, YE Wei, HUANG Gao-fei**
(School of Electronics and Communication Engineering, Guangzhou University, Guangzhou 510006, China)

Abstract: This paper studies a multicarrier non-orthogonal multiple access (MC-NOMA)-based mobile edge computing (MEC) system which consists of a user terminal, a cooperative node, and an access point equipped with an MEC server. The goal is to minimize the energy consumed for communication and computation at the user terminal and the cooperative node. To this end, a novel MEC protocol is first proposed, in which the user terminal and the cooperative node are enabled to offload their computation tasks with high spectral efficiency by employing the MC-NOMA technique. Then, under the proposed protocol, an optimization problem is formulated to jointly optimize the allocation of task offloading time and computation time, subcarrier allocation, power allocation over each subcarrier, task partition and the CPU frequency of the MEC server. The formulated problem is a highly non-convex and mixed integer programming problem, which is quite challenging to solve. By using algebraic

收稿日期:2022-11-22;**修回日期**:2023-02-10
基金项目:国家自然科学基金资助项目(61902098, 61872084);大学生创新训练资助项目(202111078006)
作者简介:罗至威(1997—),男,硕士研究生. E-mail:476021739@qq.com
*通信作者. E-mail:huanggaofei@gzhu.edu.cn
引文格式:罗至威,黄栩蔚,叶炜,等. 基于多载波非正交多址接入的协作移动边缘计算系统优化研究[J]. 广州大学学报(自然科学版),2024,23(1):84-96.

transformation, big-M formulation, layered optimization, successive convex approximation and golden section techniques, the solution of the problem is achieved, and an efficient optimization algorithm is developed. Simulation results show that by employing the MC-NOMA technique, the proposed scheme can save 50 to 60 percent energy consumption for the MEC system as compared with the scheme in the literature where only the multicarrier technique is employed.

Key words: mobile edge computing; cooperative communication; non-orthogonal multiple access; multicarrier transmission; convex optimization

移动边缘计算(Mobile Edge Computing,MEC)把计算密集型任务下沉到网络边缘节点,与传统的移动云计算相比更适合于 5G 网络中具有超低时延需求的各种新应用,因此,如何为这些应用程序部署 MEC 服务器已成为 5G 网络中的关键问题。然而,要获得高性能的 MEC 系统,必须对系统中有限的计算资源和通信资源进行高效的分配。如何结合各种无线通信技术,为 MEC 系统设计高效的资源分配方案,成为目前无线通信领域的一个热点课题[1-17]。

协作中继可以以较低的成本扩展无线通信范围,因此,如何在协作中继网络上部署 MEC 一直是备受关注的研究领域[3-7]。文献[3-4]分别研究了基于单载波和多载波协作中继的 MEC 系统,其中,协作中继既可以协作信息转发,也可以协助任务计算。其研究结果表明,通过联合优化通信和计算资源,可有效地降低用户端和协作中继用于任务卸载和计算的总能耗。文献[5]研究了一个单载波两跳协作中继网络,通过在协作中继处部署 MEC 服务器进行数据压缩和解压,以提高两个移动对等设备端到端信息传输的吞吐量。文献[6]研究了以无人机作为全双工协作中继和提供边缘计算服务的 MEC 系统,其中,无人机以单载波传输方式协助地面物联网节点进行任务卸载和计算。文献[7]研究了一种基于无线多协作中继的 MEC 系统,其中,终端节点和协作中继节点由从环境能源获取的能量供电,其目标是通过优化协作中继选择以及计算与通信资源联合分配,以降低任务执行时延和失败概率。

多载波技术(Orthogonal Frequency Division Multiplexing, OFDM)能有效地解决无线信道的频率选择性衰落问题,是宽带无线通信中的一项关键技术。因此,已经有大量研究工作结合多载波技术设计高性能的 MEC 系统[8-12]。文献[8]研究了一种基于 OFDMA(orthogonal frequency division multiple access)网络的 MEC 系统,旨在给定的计算时延约束条件下,最小化多个用户终端的能耗。假设存在恶意窃听者,文献[9]进一步研究了采用 OFDMA 技术进行安全任务卸载的 MEC 系统,通过考虑计算时延约束和保密卸载率约束,优化了计算与通信资源分配,以最小化用户端的能耗。文献[10]和文献[11]则研究了基于 OFDMA 的无线供能 MEC 系统的计算与通信资源联合分配策略,以最小化系统能耗或最大化计算任务量。文献[12]研究了一种基于 OFDMA 进行任务卸载的 D2D(device-to-device)MEC 系统,其中,MEC 服务器或近端设备均可协助任务计算,其目标是通过优化任务卸载的通信资源分配,以最小化任务计算时延。

非正交多址接入(Non-Orthogonal Multiple Access,NOMA)是一种可实现高频谱效率的无线通信新技术,如何在基于 NOMA 无线通信网络部署 MEC 系统也引起了学者们的极大兴趣[13-17]。为在给定时延约束下最小化多个物联网终端的总能耗,文献[13]针对基于多载波非正交多址接入无线通信网络下行链路的 MEC 系统,通过允许两个以上的用户终端复用相同的子载波频带,提出了最小化基站发送信息能耗的系统设计方案。文献[14]研究了由一个远端用户、一个近端用户和一个 AP 组成的 MEC 系统,其中,远端用户通过 NOMA 方式把任务数据同时卸载到近端用户和 AP(access point),同时近端用户也在本地进行任务计算和卸载任务至 AP,并针对该系统,分别提出了最小化两个用户的总能耗和最大化卸载任务量的优化方案。基于文献[14],文献[15]进一步研究了两个用户由无线供能的 MEC 系统,其目标是通过联合优化任务分配、卸载时间和计算频率最小化系统能耗。文献[16]研究了基于 NOMA 的

蜂窝辅助 MEC 系统,其中,多个边缘用户通过复用传统蜂窝用户的授权频率信道,以 NOMA 方式将其任务数据卸载至边缘服务器。文献[17]研究了基于 NOMA 的安全 MEC 系统设计,其中,一个边缘用户为避免在任务卸载过程中受到恶意窃听者的窃听攻击,在任务卸载时,与一组传统无线设备组成一个 NOMA 传输组,把任务数据卸载到边缘计算服务器,其目标是通过优化边缘用户的任务卸载数据量、卸载时间和边缘服务器计算资源分配,最小化任务执行时延。

由上述分析可知,目前已有研究结合了协作中继、多载波和 NOMA 3 种技术中的一种或两种对 MEC 系统进行设计。显然,如果能结合上述 3 种技术对 MEC 系统进行设计,其系统性能有望进一步提高,然而目前尚未有相关报道。由此,本文结合 3 种技术,对基于 MC-NOMA 协作 MEC 系统设计进行研究,系统模型见图 1。

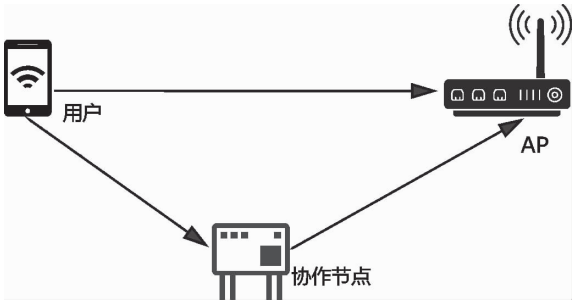

图 1 系统模型

Fig. 1 System mode

如图 1 所示,该系统由一个用户、一个协作节点和一个配备 MEC 服务器的 AP 组成,用户和协作节点均有需执行的计算任务,由于计算能力有限,均需 MEC 服务器协助完成计算任务。由于用户与 AP 之间存在障碍物且距离较远,直达链路质量较差,因此,距离 AP 较近的协作节点除了执行本地的计算任务和卸载其自身的计算任务数据至 AP 外,还作为协作中继辅助用户卸载任务数据至 AP。同时,在任务卸载时,采用 MC-NOMA 技术以充分利用频域和功率域的自由度提高任务数据传输的频谱效率,从而降低用户端和协作节点的能耗。与文献[4]仅使用多载波技术和文献[14]仅使用 NOMA 技术的协作 MEC 系统相比,本文的工作更具挑战性,其原因体现在以下两个方面。首先,基于 MC-NOMA 传输时,可以在一个子载波上通过功率域复用方式传输多个数据流。因此,包括文献[4]和文献[14]的已有研究所使用的任务卸载协议,将无法适用。由此,如何为 MC-NOMA 协作 MEC 系统设计合适的任务卸载协议,并在新协议下建立相应的系统优化问题,是本文首先要解决的新问题。其次,由于多载波传输涉及子载波分配和 NOMA 传输存在串行干扰,而且建立的系统优化问题的优化变量(包括计算任务量分配、计算时间分配、子载波分配和功率分配等)互相耦合,因此,该问题是一个非凸的混合整数规划(Mixed Integer Programming,MIP)问题。如何对该问题进行求解,也是本文需要解决的一个难题。

为应对上述 MC-NOMA 协作 MEC 系统设计中存在的挑战,本文提出一种新颖的协作 MEC 协议,使得系统可灵活地利用频域和功率域的自由度卸载用户端和协作节点的任务数据,通过结合代数变换、大 M 表达(big-M formulation)法、分层优化、连续凸优化和黄金分割法,对建立的高度非凸 MIP 问题进行求解,提出一种低复杂度的系统优化算法,从而最小化用户端和协作节点的总能耗。

1 系统模型与问题建模

1.1 系统模型

图 1 所示的 MEC 系统中,协作节点采用解码转发(Decode-and-Forward, DF)方式协助用户端卸载任务数据至 AP。协作节点是由网络中的中继设备兼任,同时该中继设备也有计算任务需要卸载。用户端和协作节点均由电池供电,AP 和 MEC 服务器则由电网供电。此外,由于部分卸载在通信和计算资源分配上比二元卸载更灵活[18],因此,假设用户端和协作节点采用部分卸载方式进行任务卸载。设用户端和协作节点需执行的任务数据量分别为 L_U 和 L_H 比特,而且以部分卸载进行任务卸载时,用户端和协作节点的本地计算任务数量分别为 l_U 和 l_H 比特,AP 处协助用户端和协作节点执行的任务数据量为 l_A 比特,则以下等式成立:

$$l_U + l_H + l_A = L_U + L_H。 \tag{1}$$

假设系统各链路信道为频率选择性信道,为解决任务卸载中无线信道频率选择性问题,提高

信息传输频谱效率,用户端和协作节点基于 MC-NOMA(即通过多载波子信道以 NOMA 的传输方式)卸载任务数据。设子载波数量为 N,而且子载波集合表示为$\mathbb{N}\triangleq\{1,\cdots,N\}$。此外,假设协作节点和 AP 均配备串行干扰消除(Successive Interference Cancellation, SIC)电路,因此,它们可使用 NOMA 技术对每个子载波上的接收信息进行解码。

假设任务时延约束为 T s,在此称之为一个时间帧。假设系统各链路信道在一个时间帧内保持不变,各个时间帧则可独立变化。h_n^{UH}、h_n^{UA} 和 h_n^{HA} 分别表示为用户端至协作节点(User-to-Helper, UH)、用户端至 AP(User-to-AP, UA)和协作节点至 AP(Helper-to-AP, HA)的第 n 个子载波信道系数。假设每个子载波信道服从非视距(Non Light of Sight, NLoS)瑞利平坦衰落。因此,在系统各个信道链路,第 n 个子载波信道系数可以表示为 $h_n^{\mathrm{ab}}=\sqrt{\rho_0 d_{\mathrm{ab}}^{-\alpha_{\mathrm{ab}}}}\tilde{h}_n^{\mathrm{ab}}$,其中,$\mathrm{ab}\in\{\mathrm{UH},\mathrm{UA},\mathrm{HA}\}$,$\rho_0$ 是参考距离 $d_0=1$ m 时的路径损耗,d_{ab}是节点 a 和节点 b 之间的距离,α_{ab}是从节点 a 到节点 b 的信道路径损耗指数,$\tilde{h}_n^{\mathrm{ab}}$ 为均值为 0、方差为 1 的复高斯随机变量,其代表相关信道的小尺度衰落。

1.2　协议设计

为充分利用 MC-NOMA 技术实现高频谱效率任务卸载,本文提出的新协作 MEC 协议时间帧结构如图 2 所示。

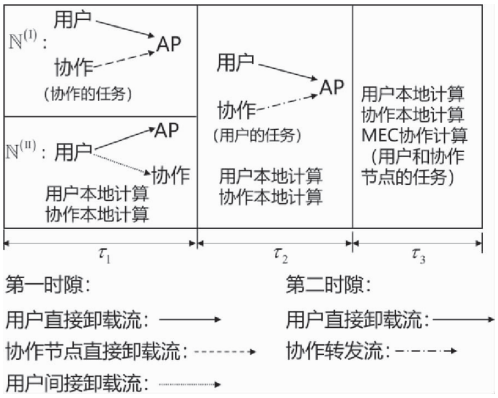

图 2　新协议的时间帧结构

Fig. 2　Time frame structure of the new protocol

其中,一个时间帧分为 3 个时隙,其长度分别表示为 τ_1、τ_2和 τ_3。因此,以下不等式成立:

$$\tau_1+\tau_2+\tau_3\leqslant T。\tag{2}$$

需要指出的是,由于计算结果的数据量比计算任务数据量通常要少很多,因此,在图 2 的时间帧结构中,忽略了 MEC 服务器把计算结果回传至用户端和协作节点的时间。该假设已在许多已有研究被采用,如文献[7-9]。

根据图 2 所示的新协议,在第一个时隙,用户和协作节点开始本地计算。在该时隙中,存在 3 个方向的任务卸载数据流,即用户直接卸载至 AP 的数据流(记为用户直接卸载流),以及协作节点直接卸载至 AP 的数据流(记为协作节点直接卸载流)、用户发送给协作节点且后续由协作节点转发至 AP 的数据流(记为用户间接卸载流)。为了限制多个数据流在一个子载波上造成的干扰,并减少处理时延和硬件复杂性[19],假设基于 NOMA 方式发送数据时,上述 3 个数据流中只有两个在同一个子载波上传输。具体而言,把用户直接卸载流和协作节点直接卸载流复用到同一个子载波上传输,其对应的子载波集合记为$\mathbb{N}^{(\mathrm{I})}$;同时,把用户直接卸载流和用户间接卸载流复用到同一个子载波上传输,其对应的子载波集合记为$\mathbb{N}^{(\mathrm{II})}$。因此,有$\mathbb{N}=\mathbb{N}^{(\mathrm{I})}\cup\mathbb{N}^{(\mathrm{II})}$且$\mathbb{N}^{(\mathrm{I})}\cap\mathbb{N}^{(\mathrm{II})}=\varnothing$,其中,$\varnothing$表示空集。设 $\omega_n^{(j)}\in\{0,1\}$ 是$\mathbb{N}^{(j)}$的子载波分配指示子,其中,$j\in\{\mathrm{I},\mathrm{II}\}$,即如果 $\omega_n^{(j)}=1$,则将第 n 个子载波分配给$\mathbb{N}^{(j)}$,反之亦然。由于每个子载波只能分配给一个子载波集合,因此有

$$\omega_n^{(\mathrm{I})}+\omega_n^{(\mathrm{II})}\leqslant 1,\ \forall n\in\mathbb{N}。\tag{3}$$

在第二个时隙,在每个子载波上,存在两个方向的任务卸载数据流,即用户直接卸载流和协作转发流。与第一个时隙相同,用户直接卸载流为用户直接卸载至 AP 的数据流;而协作转发流则是协作节点在第一个时隙的子载波上解码用户间接卸载流后,转发至 AP 的任务数据。

最后,在第三个时隙,MEC 服务器基于 AP 在第一个时隙和第二个时隙解码的任务数据执行用户端和协作节点上传的计算任务。

1.3　信号模型

根据图 2 的第一个时隙所示,令 $p_{\mathrm{UA},n}^{(1,\mathrm{I})}$和 $p_{\mathrm{HA},n}^{(1,\mathrm{I})}$为$\mathbb{N}^{(\mathrm{I})}$的子载波 n 上用户直接卸载流和协作节点直接卸载流的发射功率,令 $p_{\mathrm{UA},n}^{(1,\mathrm{II})}$和 $p_{\mathrm{UH},n}^{(1,\mathrm{II})}$为用户端在$\mathbb{N}^{(\mathrm{II})}$的子载波 n 上用户直接卸载流和用户间接卸载流的发射功率。由此,在第一个时隙的$\mathbb{N}^{(\mathrm{I})}$子载波 n 上,AP 接收到的信号为

$$y_{\mathrm{A},n}^{(1,\mathrm{I})}=\sqrt{p_{\mathrm{UA},n}^{(1,\mathrm{I})}}h_n^{\mathrm{UA}}x_{\mathrm{U},n}^{(1,\mathrm{I})}+\sqrt{p_{\mathrm{HA},n}^{(1,\mathrm{I})}}h_n^{\mathrm{HA}}x_{\mathrm{H},n}^{(1,\mathrm{I})}+z_{\mathrm{A},n},\tag{4}$$

其中,$x_{\mathrm{U},n}^{(1,\mathrm{I})}$表示用于传输用户直接卸载流的信号,$x_{\mathrm{H},n}^{(1,\mathrm{I})}$表示用于传输协作直接卸载流的信号,$z_{\mathrm{A},n}$为均值为0、方差为$\sigma_{\mathrm{A}}^2$的复加性高斯白噪声。而在$\mathbb{N}^{(\mathrm{II})}$的子载波$n$上,AP接收到的信号为

$$y_{\mathrm{A},n}^{(1,\mathrm{II})}=\sqrt{p_{\mathrm{UA},n}^{(1,\mathrm{II})}}h_n^{\mathrm{UA}}x_{\mathrm{U},n}^{(1,\mathrm{II})}+\sqrt{p_{\mathrm{HA},n}^{(1,\mathrm{II})}}h_n^{\mathrm{HA}}x_{\mathrm{H},n}^{(1,\mathrm{II})}+z_{\mathrm{A},n},\tag{5}$$

协作节点接收到的信号为

$$y_{\mathrm{H},n}=\sqrt{p_{\mathrm{UH},n}^{(1,\mathrm{II})}}h_n^{\mathrm{UH}}x_{\mathrm{H},n}^{(1,\mathrm{II})}+\sqrt{p_{\mathrm{UA},n}^{(1,\mathrm{II})}}h_n^{\mathrm{UH}}x_{\mathrm{A},n}^{(1,\mathrm{II})}+z_{\mathrm{H},n},\tag{6}$$

其中,$x_{\mathrm{U},n}^{(1,\mathrm{II})}$表示用于传输用户直接卸载流的信号,$x_{\mathrm{H},n}^{(1,\mathrm{II})}$表示用于传输用户间接卸载流的信号,$z_{\mathrm{H},n}$是均值为0、方差为$\sigma_{\mathrm{H}}^2$的复加性高斯白噪声。

根据图2的第二个时隙所示,令$p_{\mathrm{UA},n}^{(2)}$和$p_{\mathrm{HA},n}^{(2)}$表示子载波$n\in\mathbb{N}$上用户直接卸载流和协作转发流的发射功率,则AP在子载波n接收到的信号为

$$y_{\mathrm{A},n}^{(2)}=\sqrt{p_{\mathrm{UA},n}^{(2)}}h_n^{\mathrm{UA}}x_{\mathrm{U},n}^{(2)}+\sqrt{p_{\mathrm{HA},n}^{(2)}}h_n^{\mathrm{HA}}x_{\mathrm{H},n}^{(2)}+z_{\mathrm{A},n},\tag{7}$$

其中,$x_{\mathrm{U},n}^{(2)}$表示用于传输用户直接卸载流的信号,$x_{\mathrm{H},n}^{(2)}$表示用于传输协作转发流的信号。

1.4 任务卸载与能耗模型

在第一个时隙,当AP接收来自$\mathbb{N}^{(\mathrm{I})}$子载波上的用户直接卸载流和协作节点直接卸载流时,对这两个数据流的解码顺序可以是任意的[20-21]。不失一般性,假设AP先解码协作节点的信号,然后解码用户的信号。记每个子载波的频率带宽为B,则在$\mathbb{N}^{(\mathrm{I})}$的子载波上,用户传输到AP的任务数据量可表示为

$$C_{\mathrm{UA}}^{(1,\mathrm{I})}=B\tau_1\sum_{n=1}^{N}\omega_n^{(\mathrm{I})}\log\left(1+\frac{p_{\mathrm{UA},n}^{(1,\mathrm{I})}\left|h_n^{\mathrm{UA}}\right|^2}{\sigma^2}\right),\tag{8}$$

而协作节点传输到AP的任务数据量可表示为

$$C_{\mathrm{HA}}^{(1,\mathrm{I})}=B\tau_1\sum_{n=1}^{N}\omega_n^{(\mathrm{I})}\log\left(1+\frac{p_{\mathrm{HA},n}^{(1,\mathrm{I})}\left|h_n^{\mathrm{HA}}\right|^2}{p_{\mathrm{UA},n}^{(1,\mathrm{I})}\left|h_n^{\mathrm{UA}}\right|^2+\sigma^2}\right)。\tag{9}$$

在$\mathbb{N}^{(\mathrm{II})}$的子载波上,用户端卸载至AP的任务数据量可表示为

$$C_{\mathrm{UA}}^{(1,\mathrm{II})}=B\tau_1\sum_{n=1}^{N}\omega_n^{(\mathrm{II})}\log\left(1+\frac{p_{\mathrm{UA},n}^{(1,\mathrm{II})}\left|h_n^{\mathrm{UA}}\right|^2}{p_{\mathrm{UH},n}^{(1,\mathrm{II})}\left|h_n^{\mathrm{UA}}\right|^2+\sigma^2}\right),\tag{10}$$

而用户端传输至协作节点的任务数据量可表示为

$$C_{\mathrm{UH}}^{(1,\mathrm{II})}=B\tau_1\sum_{n=1}^{N}\omega_n^{(\mathrm{II})}\log\left(1+\frac{p_{\mathrm{UH},n}^{(1,\mathrm{II})}\left|h_n^{\mathrm{UH}}\right|^2}{\sigma^2}\right)。\tag{11}$$

第一个时隙用户端用于任务卸载的能耗可通过式(12)得到:

$$E_{\mathrm{O}}^{(1)}=\tau_1\sum_{n=1}^{N}\left[\omega_n^{(\mathrm{I})}\left(p_{\mathrm{UA},n}^{(1,\mathrm{I})}+p_{\mathrm{HA},n}^{(1,\mathrm{I})}\right)+\omega_n^{(\mathrm{II})}\left(p_{\mathrm{UH},n}^{(1,\mathrm{II})}+p_{\mathrm{UA},n}^{(1,\mathrm{II})}\right)\right]。\tag{12}$$

在第二个时隙,用户端和协作节点在每个子载波上均传输用户直接卸载流和协作转发流至AP。其中,用户端发送至AP的用户直接卸载流任务数据量可表示为

$$C_{\mathrm{UA}}^{(2)}=B\tau_2\sum_{n=1}^{N}\log\left(1+\frac{p_{\mathrm{UA},n}^{(2)}\left|h_n^{\mathrm{UA}}\right|^2}{\sigma^2}\right),\tag{13}$$

而协作节点转发至AP的协作转发流任务数据量可表示为

$$C_{\mathrm{HA}}^{(2)}=B\tau_2\sum_{n=1}^{N}\log\left(1+\frac{p_{\mathrm{HA},n}^{(2)}\left|h_n^{\mathrm{HA}}\right|^2}{p_{\mathrm{UA},n}^{(2)}\left|h_n^{\mathrm{UA}}\right|^2+\sigma^2}\right)。\tag{14}$$

此外,在该时隙中,用户端和协作节点用于任务卸载的能耗可通过式(15)得到:

$$E_{\mathrm{O}}^{(2)}=\tau_2\sum_{n=1}^{N}\left(p_{\mathrm{UA},n}^{(2)}+p_{\mathrm{HA},n}^{(2)}\right)。\tag{15}$$

1.5 任务计算能耗模型

令ξ_{U}、ξ_{H}和ξ_{A}分别表示在用户端、协作节点和AP处计算1比特任务数据所需的计算资源量(即CPU周期数)。因为用户端和协作节点的计算任务数据量分别为l_{U}比特和l_{H}比特,所以根据文献[4]中的引理1,在计算能耗最小化时,可得用户和协作节点完成计算任务所需的CPU频率为$f_{\mathrm{U}}=\xi_{\mathrm{U}}l_{\mathrm{U}}/T$和$f_{\mathrm{H}}=\xi_{\mathrm{H}}l_{\mathrm{H}}/T$。因此,在用户和协作节点处任务计算所消耗的能量可分别计算如下:

$$E_{\mathrm{U,c}}=\kappa_{\mathrm{U}}\xi_{\mathrm{U}}l_{\mathrm{U}}f_{\mathrm{U}}^2=\frac{\kappa_{\mathrm{U}}\xi_{\mathrm{U}}^3l_{\mathrm{U}}^3}{T^2},\tag{16}$$

$$E_{\mathrm{H,c}}=\kappa_{\mathrm{H}}\xi_{\mathrm{H}}l_{\mathrm{H}}f_{\mathrm{H}}^2=\frac{\kappa_{\mathrm{H}}\xi_{\mathrm{H}}^3l_{\mathrm{H}}^3}{T^2},\tag{17}$$

其中,κ_{U}和κ_{H}分别是由CPU芯片结构决定的用户和协作节点处的有效电容系数[22]。

此外,假设$f_{\mathrm{A},k}$为MEC服务器第k个周期的CPU频率,那么MEC服务器的计算时间为$\tau_3=\sum_{k=1}^{\xi_{\mathrm{A}}l_{\mathrm{A}}}\frac{1}{f_{\mathrm{A},k}}$。因此,式(2)中的时间分配约束可改写为

$$\tau_1+\tau_2+\sum_{k=1}^{\xi_{\mathrm{A}}l_{\mathrm{A}}}\frac{1}{f_{\mathrm{A},k}}\leqslant T。\tag{18}$$

1.6 问题建模

针对图1所示的MC-NOMA协作MEC系统,

本文的目标是通过联合优化一个时间帧的 3 个时隙时间分配、子载波分配、功率分配、任务划分和 MEC 服务器的 CPU 频率,最小化用户端和协作节点处的任务卸载和计算总能耗。根据上述能耗模型,用户端和协作节点的总能耗可表示为

$$\begin{aligned}E_{\text{total}} &= E_{\text{o}}^{(1)} + E_{\text{o}}^{(2)} + E_{\text{U,c}} + E_{\text{H,c}} = \\ &\tau_1 \sum_{n=1}^{N} \left[\omega_n^{(\text{I})} \left(p_{\text{UA},n}^{(1,\text{I})} + p_{\text{HA},n}^{(1,\text{I})} \right) + \right. \\ &\left. \omega_n^{(\text{II})} \left(p_{\text{UH},n}^{(1,\text{II})} + p_{\text{UA},n}^{(1,\text{II})} \right) \right] + \\ &\tau_2 \sum_{n=1}^{N} \left(p_{\text{UA},n}^{(2)} + p_{\text{HA},n}^{(2)} \right) + \frac{\kappa_{\text{U}} \xi_{\text{U}}^3 l_{\text{U}}^3}{T^2} + \frac{\kappa_{\text{H}} \xi_{\text{H}}^3 l_{\text{H}}^3}{T^2}。\end{aligned} \tag{19}$$

令 $\boldsymbol{\Omega} = \{ \omega_n^{(\text{I})}, \omega_n^{(\text{II})}, \forall n \}$, $\boldsymbol{P} = \{ p_{\text{UA},n}^{(1,\text{I})}, p_{\text{HA},n}^{(1,\text{I})}, p_{\text{UA},n}^{(1,\text{II})}, p_{\text{UH},n}^{(1,\text{II})}, p_{\text{UA},n}^{(2)}, p_{\text{HA},n}^{(2)}, \forall n \}$, $\boldsymbol{l} = \{ l_{\text{U}}, l_{\text{H}}, l_{\text{A}} \}$, $\tau = \{ \tau_1, \tau_2 \}$ 和 $\boldsymbol{f}_{\text{A}} = \{ f_{\text{A},k}, \forall k \}$,则本文所研究的系统优化问题可表示为

$$(\text{P0}): \min_{\Omega, \text{P}, \text{l}, \tau, f_{\text{A}}} E_{\text{total}}, \tag{20-1}$$

s. t.

$$\omega_n^{(\text{I})} + \omega_n^{(\text{II})} \leqslant 1, \forall n, \tag{20-2}$$

$$\omega_n^{(\text{I})}, \omega_n^{(\text{II})} \in \{0,1\}, \forall n, \tag{20-3}$$

$$\tau_1 + \tau_2 + \sum_{k=1}^{\xi_{\text{A}} l_{\text{A}}} \frac{1}{f_{\text{A},k}} \leqslant T, \tag{20-4}$$

$$\tau_1, \tau_2 \geqslant 0, \tag{20-5}$$

$$p_{\text{UA},n}^{(1,\text{I})}, p_{\text{HA},n}^{(1,\text{I})}, p_{\text{UA},n}^{(1,\text{II})}, p_{\text{UH},n}^{(1,\text{II})}, p_{\text{UA},n}^{(2)}, p_{\text{HA},n}^{(2)} \geqslant 0, \forall n, \tag{20-6}$$

$$\sum_{n=1}^{N} \left[\omega_n^{(\text{I})} \left(p_{\text{UA},n}^{(1,\text{I})} + p_{\text{UA},n}^{(1,\text{II})} \right) + \omega_n^{(\text{II})} \left(p_{\text{UH},n}^{(1,\text{II})} \right) \right] \leqslant p_{\text{U,max}}, \tag{20-7}$$

$$\sum_{n=1}^{N} \left[\omega_n^{(\text{II})} \left(p_{\text{HA},n}^{(1,\text{II})} \right) \right] \leqslant p_{\text{H,max}}, \tag{20-8}$$

$$\sum_{n=1}^{N} p_{\text{UA},n}^{(2)} \leqslant p_{\text{U,max}}, \sum_{n=1}^{N} p_{\text{HA},n}^{(2)} \leqslant p_{\text{H,max}}, \tag{20-9}$$

$$l_{\text{A}} = L_{\text{U}} - l_{\text{U}} + L_{\text{H}} - l_{\text{H}}, \tag{20-10}$$

$$l_{\text{U}}, l_{\text{H}}, l_{\text{A}} \geqslant 0, \tag{20-11}$$

$$L_{\text{H}} - l_{\text{H}} \leqslant C_{\text{HA}}^{(1,\text{I})}, \tag{20-12}$$

$$L_{\text{U}} - l_{\text{U}} \leqslant \min\left(C_{\text{UH}}^{(1,\text{II})}, C_{\text{HA}}^{(2)} \right) + C_{\text{UA}}^{(1,\text{I})} + C_{\text{UA}}^{(1,\text{II})} + C_{\text{UA}}^{(2)}, \tag{20-13}$$

$$f_{\text{A},k} \in [0, f_{\text{A,max}}], \forall k, \tag{20-14}$$

其中,$p_{\text{U,max}}$ 和 $p_{\text{H,max}}$ 分别表示用户端和协作节点的最大发射功率。在问题(P0)中,(20 - 3)是第一个时隙子载波分配的整数约束,(20 - 5)是时间分配的非负约束,(20 - 6)是用户端和协作节点的发射功率非负约束,(20 - 7)和(20 - 8)分别是在第一个时隙用户端和协作节点的最大发射功率限制,(20 - 9)是在第二个时隙用户端和协作节点的最大发射功率限制,(20 - 10)表示 MEC 服务器计算任务量为用户卸载至 MEC 服务器的任务量(即 $L_{\text{U}} - l_{\text{u}}$)与协作节点卸载至 MEC 服务器的任务量(即 $L_{\text{H}} - l_{\text{h}}$)之和,(20 - 11)是对用户端、协作节点和 AP 任务分配的非负约束,(20 - 12)确保协作节点分配给 MEC 服务器的计算任务量不能大于第一个时隙协作节点发送至 AP 的最大可达数据量,(20 - 13)确保用户端分配给 MEC 服务器的计算任务量不能大于第一个时隙和第二个时隙用户端发送至 AP 的最大可达数据量之和,(20 - 14)是对 AP 处 MEC 服务器 CPU 频率的限制。

在求解(P0)之前,有以下引理,以便对(P0)进行简化。

引理 1　对于问题(P0),AP 处 MEC 服务器的最优 CPU 频率为其最大值,即 $f_{\text{A},k}^* = f_{\text{A,max}}, \forall k \in \{1, \cdots, \xi_{\text{A}} l_{\text{A}}\}$。当问题(P0)取得最优解时,时间分配约束(21)是起作用的,即

$$\tau_1 + \tau_2 + \frac{\xi_{\text{A}} l_{\text{A}}}{f_{\text{A,max}}} = T。 \tag{21}$$

证明　引理 1 可以通过反证法证明。假设当问题(P0)取得最优解时,式(18)的等号不成立,此时 τ_2 的最优解为 τ_2^*,则存在 $\Delta\tau > 0$,使得 $\tau_2' = \tau_2^* + \Delta\tau$ 时,问题(P0)的所有约束仍然成立。

在问题(P0)中,令 $\tilde{p}_{\text{UA},n}^{(2)} = \tau_2 p_{\text{UA},n}^{(2)}$,则基于 $C_{\text{UA}}^{(2)}$,见式(13),有

$$\begin{aligned}&\tau_2 \sum_{n=1}^{N} \log\left(1 + \frac{p_{\text{UA},n}^{(2)} \left| h_n^{\text{UA}} \right|^2}{\sigma^2} \right) = \\ &\tau_2 \sum_{n=1}^{N} \log\left(1 + \frac{\tilde{p}_{\text{UA},n}^{(2)} \left| h_n^{\text{UA}} \right|^2}{\tau_2 \sigma^2} \right) \triangleq l',\end{aligned}$$

从而有

$$\tilde{p}_{\text{UA},n}^{(2)} = \frac{\left(e^{\frac{l'}{\tau_2}} - 1 \right) \sigma^2}{\left| h_n^{\text{UA}} \right|^2} \tau_2 \triangleq J(\tau_2),$$

由此得 $J(\tau_2)$ 的一阶导数为

$$J'(\tau_2) = \frac{1}{\left| h_n^{\text{UA}} \right|^2} \left(\left(e^{\frac{l'}{\tau_2}} - 1 \right) \sigma^2 - \frac{l' \sigma^2 e^{\frac{l'}{\tau_2}}}{\tau_2} \right)。$$

令 $x = \frac{l'}{\tau_2}$,则 $J'(\tau_2)$ 可表示为

$$W(x)=\frac{\sigma^2}{|h_n^{UA}|^2}(e^x-xe^x-1),$$

而 $W(x)$ 的一阶导数为

$$W'(x)=-\frac{\sigma^2 xe^x}{|h_n^{\mathrm{UA}}|^2}。$$

当 $\tau_2>0$、$p_{\mathrm{UA},n}^{(2)}>0$ 时,有 $l'>0$ 且 $x=\frac{l'}{\tau_2}>0$,从而 $W'(x)<0$。因此,$W(x)$,为 x 的单调递减函数。

又因为 $W(0)=0$,因此,当 $x>0$ 时,有 $W(x)<0$,即 $\tau_2>0$、$p_{\mathrm{UA},n}^{(2)}>0$ 时,$J'(\tau_2)<0$。因此,$J(\tau_2)$ 为 τ_2 的单调递减函数。从而,当 τ_2^* 增加为 $\tau_2'=\tau_2^*+\Delta\tau$ 时,$\tilde{p}_{\mathrm{UA},n}^{(2)}$ 减少,导致目标函数会减小。这与 τ_2^* 为问题(P0)的最优解矛盾。引理1得证。

根据引理1,问题(P0)简化为

$$(\mathrm{P1}):\min_{\Omega,\boldsymbol{P},l,\tau} E_{\mathrm{total}},$$

$$\text{s.t. } (20-2),(20-3),(20-5)-(20-13),(21), \tag{22}$$

其中,$\boldsymbol{f}_{\mathrm{A}}$ 不再是优化变量。因此,以下只需求解问题(P1)。

2 求解问题(P1)

在问题(P1)中,由于目标函数是非凹的,而且(20-3)是整数约束,(20-7)、(20-8)和(20-13)均是非凸约束,因此是一个难以处理的非凸MIP问题。为解决此问题,以下将首先采用代数变换和大M表达法对问题进行转换,然后通过分层优化技术对问题进行求解。在分层优化中,内层问题以连续凸优化方法求解,外层优化问题以黄金分割法求解。

2.1 问题转换

为了解决子载波分配涉及的整数规划和式(8)~(11)中发射功率与子载波分配指示子耦合导致问题(P1)非凸的问题,把式(8)~(11)分别改写为

$$\tilde{C}_{\mathrm{UA}}^{(1,\mathrm{I})}=B\tau_1\sum_{n=1}^{N}\log\left(1+\frac{\omega_n^{(\mathrm{I})}p_{\mathrm{UA},n}^{(1,\mathrm{I})}|h_n^{\mathrm{UA}}|^2}{\sigma^2}\right), \tag{23}$$

$$\tilde{C}_{\mathrm{HA}}^{(1,\mathrm{I})}=B\tau_1\sum_{n=1}^{N}\log\left(1+\frac{\omega_n^{(\mathrm{I})}p_{\mathrm{HA},n}^{(1,\mathrm{I})}|h_n^{\mathrm{HA}}|^2}{\omega_n^{(\mathrm{I})}p_{\mathrm{UA},n}^{(1,\mathrm{I})}|h_n^{\mathrm{UA}}|^2+\sigma^2}\right), \tag{24}$$

$$\tilde{C}_{\mathrm{UA}}^{(1,\mathrm{II})}=B\tau_1\sum_{n=1}^{N}\log\left(1+\frac{\omega_n^{(\mathrm{II})}p_{\mathrm{UA},n}^{(1,\mathrm{II})}|h_n^{\mathrm{UA}}|^2}{\omega_n^{(\mathrm{II})}p_{\mathrm{UH},n}^{(1,\mathrm{II})}|h_n^{\mathrm{UA}}|^2+\sigma^2}\right), \tag{25}$$

$$\tilde{C}_{\mathrm{UH}}^{(1,\mathrm{II})}=B\tau_1\sum_{n=1}^{N}\log\left(1+\frac{\omega_n^{(\mathrm{II})}p_{\mathrm{UH},n}^{(1,\mathrm{II})}|h_n^{\mathrm{UH}}|^2}{\sigma^2}\right)。 \tag{26}$$

进一步地,令

$$e_{\mathrm{ab},n}^{(1,\mathrm{I})}=\omega_n^{(\mathrm{I})}p_{\mathrm{ab},n}^{(1,I)},\ \forall \mathrm{ab}\in\{\mathrm{UA},\mathrm{HA}\}, \tag{27}$$

以及

$$e_{\mathrm{ab},n}^{(1,\mathrm{II})}=\omega_n^{(\mathrm{II})}p_{\mathrm{ab},n}^{(1,\mathrm{II})},\ \forall \mathrm{ab}\in\{\mathrm{UA},\mathrm{UH}\}, \tag{28}$$

则问题(P1)的目标函数,即 E_{total} 可改写为

$$\tilde{E}_{\mathrm{total}}=\tau_1\sum_{n=1}^{N}\left[e_{\mathrm{UA},n}^{(1,\mathrm{I})}+e_{\mathrm{HA},n}^{(1,\mathrm{I})}+e_{\mathrm{UH},n}^{(1,\mathrm{II})}+e_{\mathrm{UA},n}^{(1,\mathrm{II})}\right]+\tau_2\sum_{n=1}^{N}\left(p_{\mathrm{UA},n}^{(2)}+p_{\mathrm{HA},n}^{(2)}\right)+\frac{\kappa_{\mathrm{U}}\xi_{\mathrm{U}}^3 l_{\mathrm{U}}^3}{T^2}+\frac{\kappa_{\mathrm{H}}\xi_{\mathrm{H}}^3 l_{\mathrm{H}}^3}{T^2}。 \tag{29}$$

同时,约束(20-7)和(20-8)可以分别改写为

$$\sum_{n=1}^{N}\left[e_{\mathrm{UA},n}^{(1,\mathrm{I})}+e_{\mathrm{UA},n}^{(1,\mathrm{II})}+e_{\mathrm{UH},n}^{(1,\mathrm{II})}\right]\leqslant p_{\mathrm{U,max}}, \tag{30}$$

$$\sum_{n=1}^{N}\left(e_{\mathrm{HA},n}^{(1,\mathrm{II})}\right)\leqslant p_{\mathrm{H,max}}, \tag{31}$$

而约束(20-12)和(20-13)可分别改写为

$$L_{\mathrm{H}}-l_{\mathrm{H}}\leqslant\bar{C}_{\mathrm{HA}}^{(1,\mathrm{I})}=B\tau_1\sum_{n=1}^{N}\log\left(1+\frac{e_{\mathrm{HA},n}^{(1,\mathrm{I})}|h_n^{\mathrm{HA}}|^2}{e_{\mathrm{UA},n}^{(1,\mathrm{I})}|h_n^{\mathrm{UA}}|^2+\sigma^2}\right), \tag{32}$$

$$L_{\mathrm{U}}-l_{\mathrm{U}}\leqslant\min\{\tilde{C}_{\mathrm{UH}}^{(1,\mathrm{II})},C_{\mathrm{HA}}^{(2)}\}+\tilde{C}_{\mathrm{UA}}^{(1,\mathrm{I})}+\tilde{C}_{\mathrm{UA}}^{(1,\mathrm{II})}+C_{\mathrm{UA}}^{(2)}=\min\{\tau_1\Psi_{\mathrm{UH}}^{(1,\mathrm{II})},\tau_2\Psi_{\mathrm{HA}}^{(2)}\}+\tau_1\Psi_{\mathrm{UA}}^{(1,\mathrm{I})}+\tau_1\Psi_{\mathrm{UA}}^{(1,\mathrm{II})}+\tau_2\Psi_{\mathrm{UA}}^{(2)}, \tag{33}$$

其中,$\Psi_{\mathrm{UH}}^{(1,\mathrm{II})}=B\sum_{n=1}^{N}\log\left(1+\frac{e_{\mathrm{UH},n}^{(1,\mathrm{II})}|h_n^{\mathrm{UH}}|^2}{\sigma^2}\right)$,

$$\Psi_{\mathrm{HA}}^{(2)}=B\sum_{n=1}^{N}\log\left(1+\frac{p_{\mathrm{HA},n}^{(2)}|h_n^{\mathrm{HA}}|^2}{p_{\mathrm{UA},n}^{(2)}|h_n^{\mathrm{UA}}|^2+\sigma^2}\right),$$

$$\Psi_{\mathrm{UA}}^{(1,\mathrm{I})}=B\sum_{n=1}^{N}\log\left(1+\frac{e_{\mathrm{UA},n}^{(1,\mathrm{I})}|h_n^{\mathrm{UA}}|^2}{\sigma^2}\right),$$

$$\Psi_{\mathrm{UA}}^{(1,\mathrm{II})}=B\sum_{n=1}^{N}\log\left(1+\frac{e_{\mathrm{UA},n}^{(1,\mathrm{II})}|h_n^{\mathrm{UA}}|^2}{e_{\mathrm{UH},n}^{(1,\mathrm{II})}|h_n^{\mathrm{UA}}|^2+\sigma^2}\right),$$

$$\Psi_{\mathrm{UA}}^{(2)}=B\sum_{n=1}^{N}\log\left(1+\frac{p_{\mathrm{UA},n}^{(2)}|h_n^{\mathrm{UA}}|^2}{\sigma^2}\right)。$$

然而,式(27)~(28)的新约束导致问题(P1)仍难以求解。为解决此问题,可采用大M表达法[23]对式(27)~(28)进行等价变换。具体而言,

考虑到式(20-3)中的整数约束,约束式(27)可以等效地表达为

$$\begin{cases} e_{\mathrm{ab},n}^{(1,\mathrm{I})} \leqslant \omega_n^{(\mathrm{I})} p_{\mathrm{s,max}}, \\ e_{\mathrm{ab},n}^{(1,\mathrm{I})} \leqslant p_{\mathrm{ab},n}^{(1,\mathrm{I})}, \\ e_{\mathrm{ab},n}^{(1,\mathrm{I})} \geqslant 0, \\ e_{\mathrm{ab},n}^{(1,\mathrm{I})} \geqslant p_{\mathrm{ab},n}^{(1,\mathrm{I})} - (1-\omega_n^{(\mathrm{I})}) p_{\mathrm{s,max}}, \end{cases} \tag{34}$$

其中,当 ab = UA 时,s = U;当 ab = UH 时,s = H。同时,对 $\mathrm{ab} \in \{\mathrm{UA}, \mathrm{UH}\}$,约束式(28)可以等效地表达为

$$\begin{cases} e_{\mathrm{ab},n}^{(1,\mathrm{II})} \leqslant \omega_n^{(\mathrm{II})} p_{\mathrm{U,max}}, \\ e_{\mathrm{ab},n}^{(1,\mathrm{II})} \leqslant p_{\mathrm{ab},n}^{(1,\mathrm{II})}, \\ e_{\mathrm{ab},n}^{(1,\mathrm{II})} \geqslant 0, \\ e_{\mathrm{ab},n}^{(1,\mathrm{II})} \geqslant p_{\mathrm{ab},n}^{(1,\mathrm{II})} - (1-\omega_n^{(\mathrm{II})}) p_{\mathrm{U,max}}, \end{cases} \tag{35}$$

由此,问题(P1)可以等效地表示为

$$(\mathrm{P2}): \min_{\varepsilon, \Omega, \mathrm{P}, \mathrm{l}, \tau} \tilde{E}_{\mathrm{total}},$$

$$\text{s.t. } (20-2), (20-3), (20-5), (20-6), (20-9)-(20-11), (30)-(35), \tag{36}$$

其中,$\varepsilon = \{e_{\mathrm{UA},n}^{(1,\mathrm{I})}, e_{\mathrm{HA},n}^{(1,\mathrm{I})}, e_{\mathrm{UH},n}^{(1,\mathrm{II})}, e_{\mathrm{UA},n}^{(1,\mathrm{II})}, \forall n\}$。

为处理问题(P2)中约束(32)和(33)的非凸性,引入松弛变量 $\frac{e_{\mathrm{HA},n}^{(1,\mathrm{I})} |h_n^{\mathrm{HA}}|^2}{e_{\mathrm{UA},n}^{(1,\mathrm{I})} |h_n^{\mathrm{UA}}|^2 + \sigma^2} = s_{\mathrm{HA},n}^{(1,\mathrm{I})}$、$\frac{e_{\mathrm{UA},n}^{(1,\mathrm{II})} |h_n^{\mathrm{UA}}|^2}{e_{\mathrm{UH},n}^{(1,\mathrm{II})} |h_n^{\mathrm{UA}}|^2 + \sigma^2} = s_{\mathrm{UA},n}^{(1,\mathrm{II})}$ 和 $\frac{p_{\mathrm{HA},n}^{(2)} |h_n^{\mathrm{HA}}|^2}{p_{\mathrm{UA},n}^{(2)} |h_n^{\mathrm{UA}}|^2 + \sigma^2} = s_{\mathrm{HA},n}^{(2)}$,则约束(32)~(33)可等价地表达为式(37)~(41):

$$L_{\mathrm{H}} - l_{\mathrm{H}} \leqslant \hat{C}_{\mathrm{HA}}^{(1,\mathrm{I})} = B\tau_1 \sum_{n=1}^{N} \log(1 + s_{\mathrm{HA},n}^{(1,\mathrm{I})}), \tag{37}$$

$$L_{\mathrm{U}} - l_{\mathrm{U}} \leqslant \min\{\tau_1 \Psi_{\mathrm{UH}}^{(1,\mathrm{II})}, \tau_2 \hat{\Psi}_{\mathrm{HA}}^{(2)}\} + \tau_1 \Psi_{\mathrm{UA}}^{(1,\mathrm{I})} + \tau_1 \hat{\Psi}_{\mathrm{UA}}^{(1,\mathrm{II})} + \tau_2 \Psi_{\mathrm{UA}}^{(2)}, \tag{38}$$

$$\frac{e_{\mathrm{HA},n}^{(1,\mathrm{I})} |h_n^{\mathrm{HA}}|^2}{e_{\mathrm{UA},n}^{(1,\mathrm{I})} |h_n^{\mathrm{UA}}|^2 + \sigma^2} \geqslant s_{\mathrm{HA},n}^{(1,\mathrm{I})}, \forall n, \tag{39}$$

$$\frac{e_{\mathrm{UA},n}^{(1,\mathrm{II})} |h_n^{\mathrm{UA}}|^2}{e_{\mathrm{UH},n}^{(1,\mathrm{II})} |h_n^{\mathrm{UA}}|^2 + \sigma^2} \geqslant s_{\mathrm{UA},n}^{(1,\mathrm{II})}, \forall n, \tag{40}$$

$$\frac{p_{\mathrm{HA},n}^{(2)} |h_n^{\mathrm{HA}}|^2}{p_{\mathrm{UA},n}^{(2)} |h_n^{\mathrm{UA}}|^2 + \sigma^2} \geqslant s_{\mathrm{HA},n}^{(2)}, \forall n, \tag{41}$$

其中,$\hat{\Psi}_{\mathrm{HA}}^{(2)} = B\sum_{n=1}^{N} \log(1 + s_{\mathrm{HA},n}^{(2)})$,$\hat{\Psi}_{\mathrm{UA}}^{(1,\mathrm{II})} = B\sum_{n=1}^{N} \log(1 + s_{\mathrm{UA},n}^{(1,\mathrm{II})})$。因此,(P2)可以等效地改写为

$$(\mathrm{P3}): \min_{\mathrm{S}, \mathbb{E}, \Omega, \mathrm{P}, \mathrm{l}, \tau} \tilde{E}_{\mathrm{total}},$$

$$\text{s.t. } (20-2), (20-3), (20-5), (20-6), (20-9)-(20-11), (30), (31), (34), (35), (37)-(41), \tag{42}$$

其中,$\boldsymbol{S} = \{s_{\mathrm{HA},n}^{(1,\mathrm{I})}, s_{\mathrm{UA},n}^{(1,\mathrm{II})}, s_{\mathrm{HA},n}^{(2)}, \forall n\}$。

2.2　分层优化

在问题(P3)的目标函数和约束条件(37)~(38)中,τ 与 ε 或 $\boldsymbol{S}$ 是耦合的,因此,问题(P3)仍是一个难以直接求解的非凸问题。为此,根据式(21),令 $\tau_2 = T - \tau_1 - \frac{\xi_{\mathrm{A}} l_{\mathrm{A}}}{f_{\mathrm{A,max}}} \triangleq \varphi(\tau_1, l_{\mathrm{A}})$,则(P3)中的目标函数 $\tilde{E}_{\mathrm{total}}$(其表达式见(29)式)可重写为

$$\hat{E}_{\mathrm{total}} = \tau_1 \sum_{n=1}^{N} [e_{\mathrm{UA},n}^{(1,\mathrm{I})} + e_{\mathrm{HA},n}^{(1,\mathrm{I})} + e_{\mathrm{UH},n}^{(1,\mathrm{II})} + e_{\mathrm{UA},n}^{(1,\mathrm{II})}] + \varphi(\tau_1, l_{\mathrm{A}}) \cdot \sum_{n=1}^{N} (p_{\mathrm{UA},n}^{(2)} + p_{\mathrm{HA},n}^{(2)}) + \frac{\kappa_{\mathrm{U}} \xi_{\mathrm{U}}^3 l_{\mathrm{U}}^3}{T^2} + \frac{\kappa_{\mathrm{H}} \xi_{\mathrm{H}}^3 l_{\mathrm{H}}^3}{T^2}, \tag{43}$$

并且式(38)可改写为

$$L_{\mathrm{U}} - l_{\mathrm{U}} \leqslant \min\{\tau_1 \Psi_{\mathrm{UH}}^{(1,\mathrm{II})}, \varphi(\tau_1, l_{\mathrm{A}}) \hat{\Psi}_{\mathrm{HA}}^{(2)}\} + \tau_1 \Psi_{\mathrm{UA}}^{(1,\mathrm{I})} + \tau_1 \hat{\Psi}_{\mathrm{UA}}^{(1,\mathrm{II})} + \varphi(\tau_1, l_{\mathrm{A}}) \Psi_{\mathrm{UA}}^{(2)}。 \tag{44}$$

基于式(43)~(44),问题(P3)可通过分层优化方法分解为一个内嵌子问题和一个外部主问题。在问题(P3)中,固定优化变量 τ_1 和 l_{A},剩余的优化变量则可表示为 $\Upsilon = \{\boldsymbol{S}, \boldsymbol{\varepsilon}, \boldsymbol{\Omega}, \boldsymbol{P}, l_{\mathrm{H}}, l_{\mathrm{U}}\}$。而且,把约束(20-11)改写为下式:

$$l_{\mathrm{U}}, l_{\mathrm{H}} \geqslant 0, \tag{45}$$

则可得内嵌子问题为

$$(\mathrm{P4}): F^*(\tau_1, l_{\mathrm{A}}) = \min_{\Upsilon} \hat{E}_{\mathrm{total}},$$

$$\text{s.t. } (20-2), (20-3), (20-5), (20-6), (20-9), (20-10), (30), (31), (34), (35), (37), (39)-(41), (44), (45)。 \tag{46}$$

基于内嵌子问题(P4)定义的 $F^*(\tau_1, l_{\mathrm{A}})$,外部主问题是求解问题(P3)中 τ_1 和 l_{A} 的最优解,其表示为

$$(\mathrm{P5}): \min_{l_{\mathrm{A}}, \tau_1} F^*(\tau_1, l_{\mathrm{A}}),$$

$$\text{s.t. } l_{\mathrm{A}} \in [0, L_{\mathrm{U}} + L_{\mathrm{H}}], \tau_1 \in \left[0, T - \frac{\xi_{\mathrm{A}} l_{\mathrm{A}}}{f_{\mathrm{A,max}}}\right]。 \tag{47}$$

2.3　求解内嵌子问题(P4)

在问题(P4)中,由于约束(39)~(41)是非凸的,而且(20-3)是整数约束,因此,如何处理这些约束是求解问题(P4)的难点。为应对约束(39)~

(41)的非凸性,注意到有以下不等式:

$$e_{\mathrm{UA},n}^{(1,\mathrm{I})} * s_{\mathrm{HA},n}^{(1,\mathrm{I})} \overset{(a)}{\leqslant} \frac{1}{2}\Big[(s_{\mathrm{HA},n}^{(1,\mathrm{I})}+e_{\mathrm{UA},n}^{(1,\mathrm{I})})^2-(s_{\mathrm{HA},n}^{(1,\mathrm{I})(r)})^2-(e_{\mathrm{UA},n}^{(1,\mathrm{I})(r)})^2\Big]-s_{\mathrm{HA},n}^{(1,\mathrm{I})(r)}(s_{\mathrm{HA},n}^{(1,\mathrm{I})}-s_{\mathrm{HA},n}^{(1,\mathrm{I})(r)})-e_{\mathrm{UA},n}^{(1,\mathrm{I})(r)}(e_{\mathrm{UA},n}^{(1,\mathrm{I})}-e_{\mathrm{UA},n}^{(1,\mathrm{I})(r)}) \triangleq \Lambda_n^{(1,\mathrm{I})},$$

$$e_{\mathrm{UH},n}^{(1,\mathrm{II})} * s_{\mathrm{UA},n}^{(1,\mathrm{II})} \overset{(b)}{\leqslant} \frac{1}{2}\Big[(s_{\mathrm{UA},n}^{(1,\mathrm{II})}+e_{\mathrm{UH},n}^{(1,\mathrm{II})})^2-(s_{\mathrm{UA},n}^{(1,\mathrm{II})(r)})^2-(e_{\mathrm{UH},n}^{(1,\mathrm{II})(r)})^2\Big]-s_{\mathrm{UA},n}^{(1,\mathrm{II})(r)}(s_{\mathrm{UA},n}^{(1,\mathrm{II})}-s_{\mathrm{UA},n}^{(1,\mathrm{II})(r)})-e_{\mathrm{UH},n}^{(1,\mathrm{II})(r)}(e_{\mathrm{UH},n}^{(1,\mathrm{II})}-e_{\mathrm{UH},n}^{(1,\mathrm{II})(r)}) \triangleq \Lambda_n^{(1,\mathrm{II})},$$

$$p_{\mathrm{UA},n}^{(2)} * s_{\mathrm{HA},n}^{(2)} \overset{(c)}{\leqslant} \frac{1}{2}\Big[(s_{\mathrm{HA},n}^{(2)}+p_{\mathrm{UA},n}^{(2)})^2-(s_{\mathrm{HA},n}^{(2)(r)})^2-(p_{\mathrm{UA},n}^{(2)(r)})^2\Big]-s_{\mathrm{HA},n}^{(2)(r)}(s_{\mathrm{HA},n}^{(2)}-s_{\mathrm{HA},n}^{(2)(r)})-p_{\mathrm{UA},n}^{(2)(r)}(p_{\mathrm{UA},n}^{(2)}-p_{\mathrm{UA},n}^{(2)(r)}) \triangleq \Lambda_n^{(2)},$$

其中,因为凹函数在任意给定可行点的值不大于其在该点的一阶泰勒展开,所以,上述不等式的步骤(a)、(b)和(c)成立。令$\{e_{\mathrm{UA},n}^{(1,\mathrm{I})(r)}, s_{\mathrm{HA},n}^{(1,\mathrm{I})(r)}\}$、$\{e_{\mathrm{UH},n}^{(1,\mathrm{II})(r)}, s_{\mathrm{UA},n}^{(1,\mathrm{II})(r)}\}$和$\{p_{\mathrm{UA},n}^{(2)(r)}, s_{\mathrm{HA},n}^{(2)(r)}\}$是第$r$次迭代时一阶泰勒展开点,则式(39)~(41)可分别表达为

$$e_{\mathrm{HA},n}^{(1,\mathrm{I})}|h_n^{\mathrm{HA}}|^2 \geqslant \Lambda_n^{(1,\mathrm{I})}|h_n^{\mathrm{UA}}|^2 + s_{\mathrm{HA},n}^{(1,\mathrm{I})}\sigma^2,\ \forall n, \tag{48}$$

$$e_{\mathrm{UA},n}^{(1,\mathrm{II})}|h_n^{\mathrm{UA}}|^2 \geqslant \Lambda_n^{(1,\mathrm{II})}|h_n^{\mathrm{UA}}|^2 + s_{\mathrm{UA},n}^{(1,\mathrm{II})}\sigma^2,\ \forall n, \tag{49}$$

$$p_{\mathrm{HA},n}^{(2)}|h_n^{\mathrm{HA}}|^2 \geqslant \Lambda_n^{(2)}|h_n^{\mathrm{UA}}|^2 + s_{\mathrm{HA},n}^{(2)}\sigma^2,\ \forall n。 \tag{50}$$

此外,对于整数约束(20-3),可等效地表达为

$$0 \leqslant \omega_n^{(j)} \leqslant 1,\ \forall n, \forall j \in \{\mathrm{I},\mathrm{II}\}, \tag{51}$$

$$\omega_n^{(j)} - (\omega_n^{(j)})^2 \leqslant 0,\ \forall n,\ \forall j \in \{\mathrm{I},\mathrm{II}\}, \tag{52}$$

由此,以约束(48)~(50)替换式(39)~(42),以约束(51)~(52)替换式(20-3),内嵌问题(P4)可等效地改写为

$$(\mathrm{P6}):\ F^*(\tau_1, l_{\mathrm{A}}) = \min_{\boldsymbol{V}} \hat{E}_{\mathrm{total}},$$

$$\text{s.t. } (20-2),(20-3),(20-5),(20-6),(20-9)-(20-11),(30),(31),(34),(35),(37),(44),(45),(48)-(52)。 \tag{53}$$

在问题(P6)中,约束(52)不等式的左边是两个凸函数之差,因此,其是一个d.c.(difference of two convex functions)约束[24]。除约束(52)外,目标函数和其他约束都是凸的。因此,问题(P6)是一个典型的d.c.规划问题,可以通过连续凸近似(Successive Convex Approximation,SCA)方法在多项式时间计算复杂度内得到性能良好的局部最优解[25-26]。

为通过SCA求解问题(P6),将式(52)的左侧作为惩罚项引入问题(P6)中的目标函数中,即将问题(P6)重写为

$$(\mathrm{P7}):\ \min_{\boldsymbol{V}} \hat{E}_{\mathrm{total}} + \mu \sum_{j\in\{\mathrm{I},\mathrm{II}\}}\sum_{n=1}^{N}[\omega_n^{(j)} - (\omega_n^{(j)})^2]$$

$$\text{s.t. } (20-2),(20-3),(20-5),(20-6),(20-9)-(20-11),(30),(31),(34),(35),(37),(44),(45),(48)-(51), \tag{54}$$

其中,$\mu>0$是惩罚因子。根据文献[25-26]可知,存在$\mu_0>0$,当$\mu>\mu_0$时,问题(P5)等价于问题(P6)。

利用SCA方法,可求解问题(P7)。为此,在第r次迭代中,给定$\omega_n^{(j)(r)}$时,把问题(P7)目标函数的惩罚项一阶泰勒近似表示为

$$\sum_{j\in\{\mathrm{I},\mathrm{II}\}}\sum_{n=1}^{N}[\omega_n^{(j)} - (\omega_n^{(j)})^2] = \sum_{j\in\{\mathrm{I},\mathrm{II}\}}\sum_{n=1}^{N}[\omega_n^{(j)} - (\omega_n^{(j)(r)})^2 - 2\omega_n^{(j)(r)}(\omega_n^{(j)} - \omega_n^{(j)(r)})] \triangleq M(\boldsymbol{\omega})。 \tag{55}$$

然后,基于问题(P7),可得到第r次迭代所求解的凸优化问题如下:

$$(\mathrm{P8}):\ \min_{\boldsymbol{V}} \hat{E}_{\mathrm{total}} + \mu M(\boldsymbol{\omega}),$$

$$\text{s.t. } (20-2),(20-3),(20-5),(20-6),(20-9)-(20-11),(30),(31),(34),(35),(37),(44),(45),(48)-(51)。 \tag{56}$$

由于问题(P8)是凸问题,因此,可通过内点法求解。最后,可得计算问题(P4) $F^*(\tau_1, l_{\mathrm{A}})$的方法如算法1所示,其中,$\boldsymbol{e}_{(1)} = \{e_{\mathrm{UA},n}^{(1,\mathrm{I})}, e_{\mathrm{UH},n}^{(1,\mathrm{II})}, \forall n\}$,$\boldsymbol{p}_{(2)} = \{p_{\mathrm{UA},n}^{(2)}, \forall n\}$。

算法1 基于SCA计算$F^*(\tau_1, l_{\mathrm{A}})$的迭代算法

初始化$\boldsymbol{S}^{(0)}$,$\boldsymbol{e}_{(1)}^{(0)}$,$\boldsymbol{\Omega}^{(0)}$、$\boldsymbol{p}_{(2)}^{(0)}$、迭代次数$r=0$和迭代精度参数$\varepsilon$

1)循环

2)用$\boldsymbol{S}^{(r)}$,$\boldsymbol{e}_{(1)}^{(r)}$,$\boldsymbol{\Omega}^{(r)}$和$\boldsymbol{p}_{(2)}^{(r)}$求解问题(P8),得到$\boldsymbol{Y}^{*(r)}$和$F^*(\tau_1, l_{\mathrm{A}})^{(r)}$

3)$r=r+1$

4) until 目标值的变化小于ε,即

$|F^*(\tau_1,l_A)^{(r+1)}-F^*(\tau_1,l_A)^{(r)}|\leqslant\varepsilon$

5)获得最优解 Y^* 和 $F^*(\tau_1,l_A)$

2.4　求解外部主问题(P5)

为求解外部主问题(P5),注意到式(43)中的 $\hat{E}_{\text{total}}$ 是变量 τ_1 和 l_A 的凸函数。因此,问题(P5)可通过双层黄金分割法求解。具体而言,问题(P5)可进一步分解为一个内嵌子问题和一个外部主问题。根据问题(P5),当 l_A 给定而求解 τ_1 的最优解时,可得内嵌子问题为

$$\begin{aligned}&\psi^*(l_A)=\min_{\tau_1}F^*(\tau_1,l_A),\\ \text{s.t.}\quad&\tau_1\in\left[0,T-\frac{\xi_A l_A}{f_{A,\max}}\right],\end{aligned}\tag{57}$$

而外部主问题则是求解 l_A 的最优解,其可表示为

$$\begin{aligned}&\min_{l_A}\psi^*(l_A),\\ \text{s.t.}\quad&l_A\in[0,L]。\end{aligned}\tag{58}$$

基于求解嵌入子问题(P4)的算法 1 和以双层黄金分割法求解外部主问题(P5),可得问题(P3)的求解步骤如算法 2 所示。

算法 2　问题(P3)求解算法

初始化迭代精度参数 ε

1) 令 $l_A^{\min}=0$, $l_A^{\max}=\text{L}$, $l_A^{\text{U}}=l_A^{\max}$, $l_A^{\text{L}}=l_A^{\min}$

2) while　$l_A^{\text{U}}-l_A^{\text{L}}>\varepsilon$ do

3)　　$l_A^{\text{U}}=l_A^{\min}+(l_A^{\max}-l_A^{\min})*0.618$

4)　　$l_A^{\text{L}}=l_A^{\max}+l_A^{\min}-l_A^{\text{U}}$

5)　　$l_A=l_A^{\text{U}}$

6)　　令 $\tau_1^{\text{U,min}}=0$, $\tau_1^{\text{U,max}}=T-\frac{\xi_A l_A}{f_{A,\max}}$, $\tau_1^{\text{U,U}}=\tau_1^{\text{U,max}}$, $\tau_1^{\text{U,L}}=\tau_1^{\text{U,min}}$

7)　　while $\tau_1^{\text{U,U}}-\tau_1^{\text{U,L}}>\varepsilon$ do

8)　　$\tau_1^{\text{U,U}}=\tau_1^{\text{U,min}}+(\tau_1^{\text{U,max}}-\tau_1^{\text{U,min}})*0.618$

9)　　$\tau_1^{\text{U,L}}=\tau_1^{\text{U,max}}+\tau_1^{\text{U,min}}-\tau_1^{\text{U,U}}$

10)　以 $\tau_1=\tau_1^{\text{U,U}}$ 和 $\tau_1=\tau_1^{\text{U,L}}$ 分别执行算法 1 计算得到 $F_{\text{U,U}}^*(\tau_1,l_A)$ 和 $F_{\text{U,L}}^*(\tau_1,l_A)$

11)　　if $F_{\text{U,U}}^*(\tau_1,l_A)<F_{\text{U,L}}^*(\tau_1,l_A)$ then

12)　　$\tau_1^{\text{U,min}}=\tau_1^{\text{U,L}}$

13)　　else

14)　　$\tau_1^{\text{U,max}}=\tau_1^{\text{U,U}}$

15)　　end if

16)　end while

17)　计算 $F_{\text{U}}^*(\tau_1,l_A)=\frac{F_{\text{U,U}}^*(\tau_1,l_A)+F_{\text{U,L}}^*(\tau_1,l_A)}{2}$

18)　令 $l_A=l_A^{\text{L}}$

19)　令 $\tau_1^{\text{L,min}}=0$, $\tau_1^{\text{L,max}}=T-\frac{\xi_A l_A}{f_{A,\max}}$, $\tau_1^{\text{L,U}}=\tau_1^{\text{L,max}}$, $\tau_1^{\text{L,L}}=\tau_1^{\text{L,min}}$

20)　while $\tau_1^{\text{L,U}}-\tau_1^{\text{L,L}}>\varepsilon$ do

21)　　$\tau_1^{\text{L,U}}=\tau_1^{\text{L,min}}+(\tau_1^{\text{L,max}}-\tau_1^{\text{L,min}})*0.618$

22)　　$\tau_1^{\text{L,L}}=\tau_1^{\text{L,max}}+\tau_1^{\text{L,min}}-\tau_1^{\text{L,U}}$

23)　　以 $\tau_1=\tau_1^{\text{L,U}}$ 和 $\tau_1=\tau_1^{\text{L,L}}$ 分别执行算法 1 计算得到 $F_{\text{L,U}}^*(\tau_1,l_A)$ 和 $F_{\text{L,L}}^*(\tau_1,l_A)$

24)　　If $F_{\text{L,U}}^*(\tau_1,l_A)<F_{\text{L,L}}^*(\tau_1,l_A)$ then

25)　　$\tau_1^{\text{L,min}}=\tau_1^{\text{L,L}}$

26)　　else

27)　　$\tau_1^{\text{L,max}}=\tau_1^{\text{L,U}}$

28)　　end if

29)　end while

30)　计算 $F_{\text{L}}^*(\tau_1,l_A)=\frac{F_{\text{L,U}}^*(\tau_1,l_A)+F_{\text{L,L}}^*(\tau_1,l_A)}{2}$

31)　If $F_{\text{U}}^*(\tau_1,l_A)<F_{\text{L}}^*(\tau_1,l_A)$ then

32)　　$l_A^{\min}=l_A^{\text{L}}$

33)　　else

34)　　$l_A^{\max}=l_A^{\text{U}}$

35)　　end if

36)　end while

37) 计算 $\tau_1^*=\frac{\tau_1^{\text{U,U}}+\tau_1^{\text{U,L}}+\tau_1^{\text{L,U}}+\tau_1^{\text{L,L}}}{4}$ 和 $l_A^*=\frac{l_A^{\text{U}}+l_A^{\text{L}}}{2}$,然后,用 τ_1^* 和 l_A^* 执行算法 1 得到最优解 Y^*

由于问题(P8)可通过内点法求解,且其优化变量数为 $15N+2$,因此,算法 1 的计算复杂度为 $\mathcal{O}(I(15N+2)^{3.5})$,其中,$I$ 为算法 1 收敛的所需迭代次数。由此,算法 2 的计算复杂度可表示为 $\mathcal{O}(2[\log_2(\varepsilon^{-1})]^2\text{I}(15N+2)^3.5)$。因此,算法 2 具有多项式时间的计算复杂度,是一个低复杂度的高效算法。

3　仿真结果与性能分析

本节以计算机仿真来验证提出方案的性能提

升。为此,比较了提出的方案与4种基准方案获得的系统性能,其中4种基准方案如下:

(1)全部本地计算方案:用户端和协作节点的计算任务全部都在本地计算,不进行卸载;

(2)全部卸载至AP执行方案:用户端和协作节点的计算任务全部都卸载至AP处的MEC进行计算,本地不进行计算;

(3)基于多载波传输的任务卸载方案:基于文献[4]提出的算法推导而来,其中,任务卸载不使用NOMA,仅使用多载波进行传输;

(4)无协作任务卸载方案:用户端和协作节点通过MC-NOMA方式直接卸载部分计算任务至AP,即协作节点不协助用户端转发任务数据。

在仿真中,采用图3所示的二维坐标系,其中,用户端位于原点(0 m,0 m),协作节点位于(10 m, -2 m),AP的纵坐标设置为0 m(即位于横轴上)。子载波的数量设置为 $N=16$。而且,在以下所有仿真中,为方便起见且不失一般性,令用户端和协作节点的计算任务量均为相同的,即 $L_{\mathrm{U}}=L_{\mathrm{H}}$,其值记为 L_0。给定1 m的参考距离,设平均信号功率衰减为-30 dB。与用户端与AP之间存在障碍物的场景相对应,设用户端与AP之间的链路路径损耗指数 α_{UA} 为3,而用户端与协作节点、协作节点与AP之间的链路均为自由空间,因此,设这两个链路的路径损耗指数为 $\alpha_{\mathrm{UH}}=\alpha_{\mathrm{HA}}=2$。设加性高斯白噪声的方差 $\sigma_{\mathrm{H}}^2=\sigma_{\mathrm{A}}^2=-70$ dBm,用户端和协作节点的最大发射功率 $p_{\mathrm{U,max}}=p_{\mathrm{H,max}}=1$ W,惩罚系数 $\eta=1\times10^{-3}$,计算参数 $\xi_{\mathrm{U}}=\xi_{\mathrm{H}}=\xi_{\mathrm{A}}=10^3$ cycle/bit,$f_{\mathrm{A,max}}=5$ GHz,$\kappa_{\mathrm{U}}=\kappa_{\mathrm{H}}=10^{-27}$。除非特别说明,否则上述所有配置参数在以下所有仿真中均不会改变。

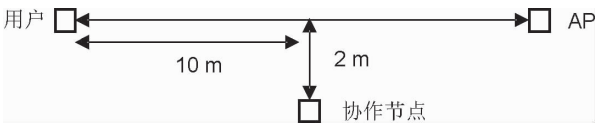

图3 仿真中各节点位置示意图

Fig. 3 The positions of the nodes in simulations

3.1 不同计算任务量下的系统性能评估

图4画出了提出的方案与4个基准方案在不同计算任务量 L_0 下的能耗性能,其中,任务时延约束设置为 $T=0.08$ s,AP的横坐标设为20 m。

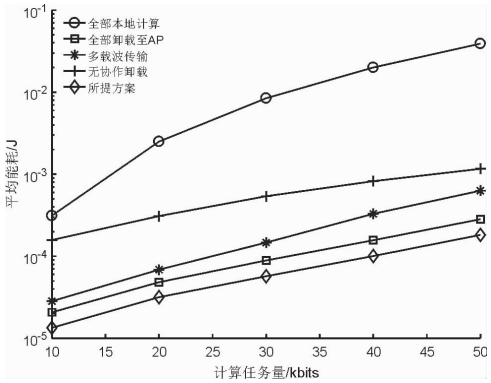

图4 不同计算任务量 L_0 下不同方案的性能比较

Fig. 4 Comparison of performance achieved by different schemes under various computation task bits L_0

从图4可以发现,在所有的方案中,提出的方案消耗能量最少。特别地,与全部卸载至AP执行方案相比,提出的方案可以节省大约35%的能耗,而且明显比全部本地计算方案所需的能耗低得多,其原因在于采用部分卸载可取得比二元卸载更好的系统性能。进一步地,与基于多载波传输的任务卸载方案相比,提出的方案可以节省大约60%的能耗,从而验证了结合多载波和NOMA进行任务卸载可显著地降低系统能耗。与无协作任务卸载方案相比,提出的方案可节省大约87%的能耗,从而说明了在基于MC-NOMA的MEC系统中采用协作任务卸载可明显提高系统性能。

3.2 不同时延约束下的系统性能评估

图5中画出了提出的方案与基准方案的能耗性能在不同的任务时延约束 T 下的对比结果,其中,计算任务量设置为 $L_0=20$ kbits, AP的横坐标设置为20 m。

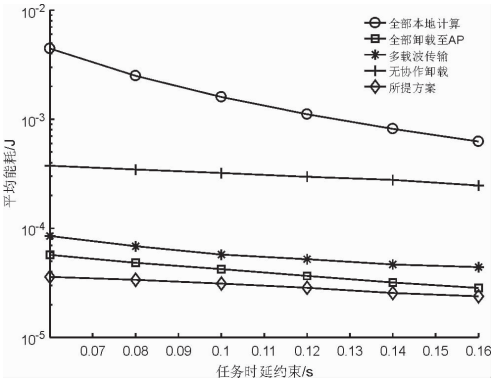

图5 不同任务时延约束下不同方案的性能比较

Fig. 5 Comparison of performance achieved by different schemes under various computation task delay constraints

在图 5 中可以发现,随着 T 的增加,能耗会随之降低。更为重要的是,图 5 的仿真结果再次验证了提出的方案优于基准方案。具体而言,与全部卸载至 AP 执行方案、基于多载波传输的任务卸载方案和无协作任务卸载方案相比,提出的方案分别节省了大约 20%、50% 和 90% 的能耗,从而验证了在多载波 MEC 系统上结合 NOMA 技术进行任务卸载的重要性,也展现了在结合 MC-NOMA 技术的 MEC 系统上采用本文提出的新协议和系统优化算法的优势。

3.3　AP 位置变化时的系统性能评估

图 6 展示了 AP 位置变化时提出的方案与基准方案获得的能耗性能比较结果,其中,计算任务量设置为 $L_0 = 20$ kbits,任务时延约束指定为 $T = 0.08$ s,AP 的横坐标从 20 m 增加到 40 m。当 AP 的横坐标增加时,AP 与用户端和协作节点的距离也会增加。

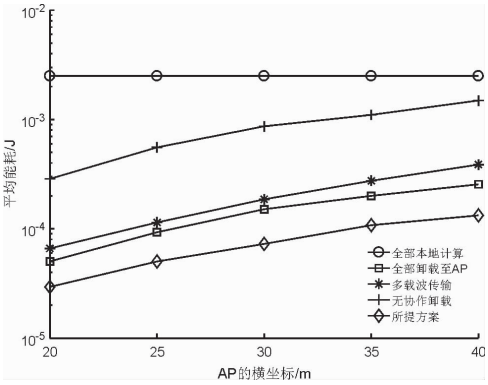

图 6　AP 位置变化时不同方案的性能比较

Fig. 6　Comparison of performance achieved by different schemes under various positions of AP

从图 6 可以看出,随着 AP 横坐标增加,系统能耗也会随之增加,其原因在于任务卸载能耗的增加。然而,与全部卸载至 AP 执行方案、基于多载波传输的任务卸载方案和无协作任务卸载方案相比,提出的方案分别节省了大约 45%、60% 和 90% 的能耗,从而进一步验证了本文所提方案的优越性。

3.4　第一时隙子载波分配结果分析

图 7 展示了不同任务时延约束下第一个时隙的子载波分配情况。

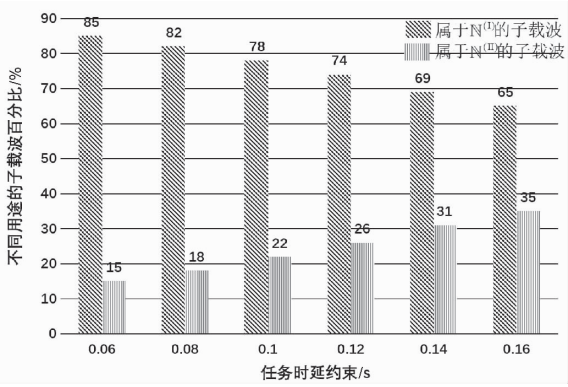

图 7　第一时隙的子载波分配结果示意图

Fig. 7　Subcarrier allocation results in the first time slot

由图 7 可知,随着任务时延约束的增加,分配至$\mathbb{N}^{(I)}$的子载波数量会减少,而分配至$\mathbb{N}^{(II)}$的子载波数量会增加。如前所述,$\mathbb{N}^{(I)}$的子载波用于传输用户直接卸载流和协作节点直接卸载流,$\mathbb{N}^{(II)}$的子载波用于传输用户直接卸载流和用户间接卸载流。因此,图 7 的结果表明,任务时延约束越宽松,就有越多的子载波用于协作转发,即用户端经协作节点转发至 AP 的任务数据就越多。

4　结束语

本文研究了基于 MC-NOMA 的协作边缘计算系统设计与优化方案,该系统由一个用户端、一个协作节点和一个配备 MEC 服务器的 AP 组成。为了提高该 MEC 系统的性能,提出了一种新的协议,该协议可充分利用多载波和 NOMA 技术的高频谱效率进行任务卸载。在新协议下,以最小化用户端和协作节点的能耗为目标,考虑用户端和协作节点的任务执行时延约束,建立了该 MEC 系统的资源分配优化问题。然后,通过利用代数变换、大 M 表达法、分层优化、SCA 和黄金分割法,对建立的优化问题进行了求解,并提出了一种高效的系统优化算法。最后,通过计算机仿真,验证了所提出的设计方案可显著地减少系统能耗。

参考文献:

[1]　曹敦, 张应宝, 邹电,等. V2X 多节点协同分布式卸载策略[J]. 通信学报, 2022, 43(2):185-195.

[2]　崔高峰, 徐媛媛, 张尚宏,等. 基于最小能耗的多无人机无线网络安全数据卸载策略[J]. 通信学报, 2021,42(5):51-62.

[3]　Cao X W, Wang F, Xu J, et al. Joint computation and communication cooperation for energy-efficient mobile edge compu-

ting[J]. IEEE Internet of Things Journal, 2019, 6(3):4188-4200.

[4] Hu D, Huang G F, Tang D, et al. Joint task offloading and computation in cooperative multicarrier relaying based mobile edge computing systems[J]. IEEE Internet of Things Journal, 2021, 8(14):11487-11502.

[5] Qin M, Chen L, Zhao N, et al. Computing and relaying: Utilizing mobile edge computing for P2P communications[J]. IEEE Transactions on Vehicular Technology, 2020, 69(2): 1582-1594.

[6] Zhang L, Ansari N. Optimizing the operation cost for UAV-aided mobile edge computing[J]. IEEE Transactions on Vehicular Technology, 2021, 70(6):6085-6093.

[7] Li M L, Zhou X B, Qiu T, et al. Multi-relay assisted computation offloading for multi-access edge computing systems with energy harvesting [J]. IEEE Transactions on Vehicular Technology, 2021, 70(10): 10941-10956.

[8] You C S, Huang K B, Chae H, et al. Energy-efficient resource allocation for mobile-edge computation offloading[J]. IEEE Transactions on Wireless Communications, 2017, 16(3):1397-1411.

[9] Xu J, Yao J P. Exploiting physical-layer security for multiuser multicarrier computation offloading[J]. IEEE Wireless Communications Letters, 2019, 8(1):9-12.

[10] Nguyen P X, Tran D H, Onireti O, et al. Backscatter-assisted data offloading in OFDMA-based wireless-powered mobile edge computing for IoT networks[J]. IEEE Internet of Things Journal, 2021, 8(11):9233-9243.

[11] Saleem U, Liu Y, Jangsher S, et al. Latency minimization for D2D-Enabled partial computation offloading in mobile edge computing[J]. IEEE Transactions on Vehicular Technology, 2020, 69(4):4472-4486.

[12] Bai T, Pan C H, Ren H, et al. Resource allocation for intelligent reflecting surface aided wireless powered mobile edge computing in OFDM systems[J]. IEEE Transactions on Wireless Communications, 2021, 20(8): 5389-5407.

[13] Zhang L, Fang F R, Huang G X, et al. Energy-efficient non-orthogonal multiple access for downlink communication in mobile edge computing systems[J]. IEEE Transactions on Mobile Computing, 2022, 21(12):4310-4322.

[14] Huang Y W, Liu Y, Chen F J. NOMA-aided mobile edge computing via user cooperation[J]. IEEE Transactions on Communications, 2020, 68(4): 2221-2235.

[15] Li B G, Si F Q, Zhao W, et al. Wireless powered mobile edge computing with NOMA and user cooperation[J]. IEEE Transactions on Vehicular Technology, 2021, 70(2): 1957-1961.

[16] Qian L P, Wu Y, Ouyang J Y, et al. Latency optimization for cellular assisted mobile edge computing via non-orthogonal multiple access[J]. IEEE Transactions on Vehicular Technology, 2020, 69(5): 5494-5507.

[17] Qian L P, Wu W C,Lu W D, et al. Secrecy based energy-efficient mobile edge computing via cooperative non-orthogonal multiple access transmission[J]. IEEE Transactions on Communications, 2021, 69(7): 4659-4677.

[18] Mao Y Y, You C S, Zhang J, et al. A survey on mobile edge computing: The communication perspective[J]. IEEE Communications Surveys & Tutorials, 2017, 19(4): 2322-2358.

[19] Saito Y, Kishiyama Y, Benjebbour A, et al. Proceeding of 2013 IEEE 77th Vehicular Technology Conference (VTC Spring), June 2-5, 2013[C]. Piscataway:IEEE, 2013.

[20] Sun Y, Ng D W K, Ding Z G, et al. Optimal joint power and subcarrier allocation for full-duplex multicarrier non-orthogonal multiple access systems[J]. IEEE Transactions on Communications, 2017, 65(3):1077-1091.

[21] Ding Z G, Fan P Z, Poor H V. Impact of user pairing on 5G nonorthogonal multiple-access downlink transmissions[J]. IEEE Transactions on Vehicular Technology, 2015, 65(8): 6010-6023.

[22] Burd T D, Brodersen R W. Processor design for portable systems[J]. Journal of VLSI Signal Processing Systems for Signal, Image and Video Technology, 1996, 13(2): 203-221.

[23] Lee J, Leyffer S. Mixed integer nonlinear programming[M]. New York:Springer, 2012.

[24] Smola A J, Vishwanathan S V N, Hofmann T. Proceeding of International Workshop on Artificial Intelligence and Statistics [C]. New York:PMLR, 2005.

[25] Rashid U, Tuan H D, Kha H H, et al. Joint optimization of source precoding and relay beamforming in wireless MIMO relay networks[J]. IEEE Transactions on Communications, 2014, 62(2):488-499.

[26] Che E, Tuan H, Nguyen H H. Joint optimization of cooperative beamforming and relay assignment in multi-user wireless relay networks[J]. IEEE Transactions on Wireless Communications, 2014, 13(59):5481-5495.

【责任编辑：卓祯雨】