

超密集网络中基于移动边缘计算的任务卸载和资源优化

张海波^① 李虎^{*①} 陈善学^① 贺晓帆^②

^①(重庆邮电大学通信与信息工程学院 重庆 400065)

^②(美国德克萨斯州拉玛尔大学电子工程系 美国 77710)

摘要: 移动边缘计算(MEC)通过在无线网络边缘为用户提供计算能力, 来提高用户的体验质量。然而, MEC的计算卸载仍面临着许多问题。该文针对超密集组网(UDN)的MEC场景下的计算卸载, 考虑系统总能耗, 提出卸载决策和资源分配的联合优化问题。首先采用坐标下降法制定了卸载决定的优化方案。同时, 在满足用户时延约束下采用基于改进的匈牙利算法和贪婪算法来进行子信道分配。然后, 将能耗最小化问题转化为功率最小化问题, 并将其转化为一个凸优化问题得到用户最优的发送功率。仿真结果表明, 所提出的卸载方案可以在满足用户不同时延的要求下最小化系统能耗, 有效地提升了系统性能。

关键词: 超密集组网; 移动边缘计算; 计算卸载; 资源分配

中图分类号: TN929.5

文献标识码: A

文章编号: 1009-5896(2019)05-1194-08

DOI: 10.11999/JEIT180592

Computing Offloading and Resource Optimization in Ultra-dense Networks with Mobile Edge Computation

ZHANG Haibo^① LI Hu^① CHEN Shanxue^① HE Xiaofan^②

^①(School of Communication and Information Engineering, Chongqing University of Posts and Telecommunications, Chongqing 400065, China)

^②(Department of Electronic Engineering, Lamar University, TX 77710, USA)

Abstract: Mobile Edge Computing (MEC) improves the quality of users experience by providing users with computing capabilities at the edge of the wireless network. However, computing offloading in MEC still faces some problems. In this paper, a joint optimization problem of offloading decision and resource allocation is proposed for the computation offloading problem in Ultra-Dense Networks (UDN) with MEC. To solve this problem, firstly, the coordinate descent method is used to formulate the optimization scheme for the offloading decision. Meanwhile, the improved Hungarian algorithm and greedy algorithm are used to allocate the channels to meet the user's delay requirements. Finally, the problem of minimizing energy consumption is converted into a problem of minimizing power. Then it is converted into a convex optimization problem to get the user's optimal transmission power. Simulation results show that the proposed scheme can minimize the energy consumption of the system while satisfying the users' different delay requirements, and improve effectively the performance of the system.

Key words: Ultra-Dense Networks (UDN); Mobile Edge Computing (MEC); Computing offloading; Resource allocation

收稿日期: 2018-06-13; 改回日期: 2019-01-21; 网络出版: 2019-02-14

*通信作者: 李虎 976502889@qq.com

基金项目: 国家自然科学基金(61771084, 61601071), 长江学者和创新团队发展计划基金(IRT16R72), 重庆市基础研究与前沿探索项目(cstc2018jcyjAX0463)

Foundation Items: The National Natural Science Foundation of China (61771084, 61601071), The Foundation for Changjiang Scholars and Innovative Research Team in University (IRT16R72), The Basic Research and Frontier Exploration Projects in Chongqing (cstc2018jcyjAX0463)

1 引言

随着移动互联网和普适计算的不断发展, 移动用户越来越多地运行资源密集型应用, 如: 交互式游戏、增强现实等等^[1,2]。然而, 移动设备在有限的电量和计算资源的情况下并不能完全满足这些应用对于能耗和时延的需求。近年来, 移动边缘计算(Mobile Edge Computing, MEC)的提出为解决这些问题提供了新思路。MEC系统允许设备将计算任务卸载到网络边缘节点, 如基站、无线接入点

等,既满足了终端设备计算能力的扩展需求,同时弥补了云计算时延较长的缺点^[3]。

此外,未来5G网络需要满足连续广域覆盖、热点高容量、低时延低功耗等服务需求,而超密集网络通过更加密集化的网路部署,能够满足局部热点区域的增强覆盖,增加系统容量,同时密集的网络部署能够提升频谱利用率、减少端到端的延时^[4,5]。因此,在未来的5G网络中,为了使用户能够广泛地享受到超密集网络和MEC所带来的性能上的提升,超密集组网联合MEC的网络架构将会成为未来无线网络的典型范例。

针对MEC的任务卸载问题,许多学者做了相关研究。文献^[6]在考虑前程和回程链路容量约束以及用户的最大时延约束条件下,通过最小化网络总能耗提出了一种有效卸载方案。文献^[7]在权衡能耗和时延下,提出了一种能量感知的计算卸载方案,并将智能设备电池的剩余能量引入能量消耗和延迟的加权因子的定义中,有效地降低了系统的总消耗。考虑到任务卸载的等待时间和可靠性之间的折中,文献^[8]研究了将用户设备的任务分割成子任务并依次卸载到附近边缘节点。但是以上文献并没有对有限的无线和计算资源进行合理的分配。文献^[9]在多用户的MEC系统下,以最小化用户和MEC服务器的平均能量消耗为目标,提出了一种在线的任务卸载算法。文献^[10]和文献^[11]考虑系统的总能耗最小化,研究了卸载决定、无线资源和计算资源分配的联合优化问题。但是上述文献并没有考虑时延对系统性能的影响。

针对超密集组网中的任务卸载,当过多的用户同时选择将任务卸载计算时,计算的瓶颈以及小区间的干扰将会严重制约MEC的可扩展性。针对这些问题,文献^[12]研究了多用户多任务场景下的计算卸载问题,考虑无线和计算资源的价格成本,进行有效的资源分配来最小化用户的总消耗。文献^[13]和文献^[14]研究了小小区网络架构下的计算卸载问题,提出了一种基于博弈的卸载算法。但是以上文献并没有考虑需求的差异性来进行合理的资源分配。文献^[15]考虑小区间干扰对系统性能的影响,将卸载决定、物理资源块分配和MEC计算资源分配作为一个联合优化问题,提出了一种有效的卸载方案,该方案中MEC首先根据负载评估做出初始卸载决定,然后根据卸载决定和用户的负载采用图着色算法进行资源分配,虽然有效地控制了干扰并且得到了较优的性能,但是没有考虑每个用户的业务需求来进行合理的资源分配,也没有考虑功率分配对系统性能的影响。

不同于上述文献,本文研究了超密集组网的MEC场景下的计算卸载问题,考虑不同的时延需求以及功率分配对系统性能的影响,分析讨论了卸载和资源分配的联合优化问题,主要贡献如下:

(1) 考虑用户的时延和功率约束,制定了卸载决定和资源分配的联合优化问题,最优化用户的总能耗。

(2) 针对上述问题,首先采用坐标下降法进行卸载决定的更新操作,考虑到密集小区下干扰对系统性能的影响,在每一次卸载决定更新下采用改进的匈牙利算法和贪婪算法进行用户级的子信道分配,然后将能耗问题转化为时延约束下的功率最小化问题,并通过凸优化方法得到用户最优的功率分配,最后通过有限次迭代得到最优的卸载决定和功率分配结果。

2 系统模型

搭建一个宏基站和 N 个小基站(Small Base Station, SBS)组成的网络。如图1所示,在网络的边缘部署有MEC服务器,MEC服务器可以同时执行多个计算任务。SBS都以同频的方式部署,并与宏基站以有线的方式连接,每个SBS的频带被分成 K 个正交的子信道,定义 $\mathcal{K} = \{1, 2, \dots, K\}$ 代表子信道的集合,不同小区的用户可以复用相同的子信道,所以不同小区的用户之间会产生干扰。为了方便分析,假设每个SBS只有1个用户的情况,定义 $\mathcal{N} = \{1, 2, \dots, N\}$ 代表所有用户的集合。在这个网络中每个用户 n 都有一个计算任务需要完成,定义 $a_n \in \{0, 1\}$ 为用户的卸载决策,0代表用户选择本地执行,1代表用户选择卸载到MEC执行。因此,用 $\mathbf{A} = \{a_1, a_2, \dots, a_N\}$ 表示所有用户的卸载决定。

2.1 通信模型

当用户选择卸载任务时考虑上行传输时相邻SBS用户的干扰,当用户 n 分配子信道 k 进行数据传输时,用户 n 在子信道 k 上的信干噪比为

$$\text{SINR}_n^k = \frac{p_n^k h_{n,n}^k}{\omega_0 + \sum_{m=1, m \neq n}^N p_m^k h_{m,n}^k} \quad (1)$$

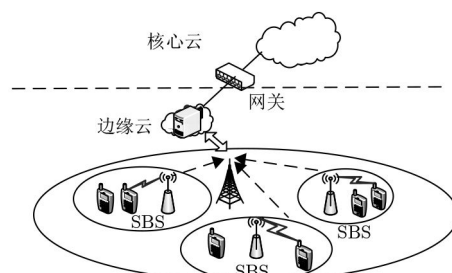


图1 系统模型

计算得到用户 n 在子信道 k 上的传输速率为

$$r_n^k = c_n^k \cdot B \log_2(1 + \text{SINR}_n^k) \quad (2)$$

则用户 n 卸载任务时上行传输的总速率为

$$R_n = \sum_{k \in K} r_n^k \quad (3)$$

其中, B 表示子信道带宽, c_n^k 表示若子信道 k 被分配给用户 n 则 $c_n^k=1$, 反之 $c_n^k=0$ 。 p_n^k 和 p_m^k 分别表示用户 n 和 m 在子信道 k 上的发送功率, $h_{n,n}^k$ 和 $h_{m,n}^k$ 分别表示 SBS n 的用户 n 和 SBS m 的用户 m 到 SBS n 的信道增益, ω_0 表示背景噪声功率。

2.2 计算模型

每个用户 n 有一个计算任务, 表示为 $t_n = (w_n, d_n, T_n^{\max})$, w_n 表示计算该任务所需的 CPU 周期, d_n 代表输入数据的大小, $T_n^{\max}$ 表示用户能容忍的最大时延。

(1) 当用户选择本地执行时, 用 f_n^l 代表用户设备的计算能力, 则本地计算的时延为

$$T_n^l = \frac{w_n}{f_n^l} \quad (4)$$

本地计算能量消耗为

$$E_n^l = \kappa (f_n^l)^2 w_n \quad (5)$$

其中, κ 值的大小取决于移动设备的芯片结构^[7], 取 $\kappa=10^{(-26)}$ 。考虑到本地能量消耗随着用户计算能力 f_n^l 的变化而变化, 通过动态电压缩放技术动态调节用户的计算能力可以有效减少用户的计算能耗^[11,16]。因此, 在时延约束下, 本地计算时最优的计算能力 f_n^l 可以表示为

$$f_n^l = \min \left\{ \frac{w_n}{T_n^{\max}}, F_n^l \right\} \quad (6)$$

其中, F_n^l 表示用户 n 的最大计算能力。

(2) 当用户选择将任务卸载计算时, 通过无线网络传输任务时会产生相应的传输时延和能耗, 根据通信模型, 用户 n 卸载任务时上行传输时延为

$$T_n^t = \frac{d_n}{R_n} \quad (7)$$

则传输计算任务时用户的能量消耗为

$$E_n^c = P_n^t \cdot \frac{d_n}{R_n} + P_n^i \cdot \frac{w_n}{f^c} \quad (8)$$

其中, $P_n^t = \sum_{k \in K} c_n^k p_n^k$ 表示用户的总发送功率; P_n^i 表示用户空闲状态下的电路功率消耗。

计算任务卸载到 MEC 后, MEC 会为该任务分配一定的计算资源, 考虑 MEC 为每个用户分配的

计算资源是固定的^[7]。用 f^c 表示 MEC 分配的计算资源, MEC 执行该任务的时延为

$$T_n^c = \frac{w_n}{f^c} \quad (9)$$

本文主要考虑用户端的能耗和时延, 所以省略了 MEC 服务器的计算能耗。返回时的数据量较小, 因此省略了接收返回数据时用户的能耗和时延^[7,13]。

3 问题形成

在任务卸载的过程中每个用户将会对本地计算的花费进行评估, 然后上报给 MEC。同时, MEC 也会评估每个用户卸载时的花费。然后, MEC 通过比较本地和卸载的花费, 做出相应的卸载决定, 卸载决定表示为

$$\left. \begin{aligned} a_n &= 0, E_n^l \leq E_n^c \\ a_n &= 1, E_n^l > E_n^c \end{aligned} \right\} \quad (10)$$

本文用 N_c 表示卸载的用户数目, 用 $\mathcal{N}_c$ 表示卸载的用户集合, 则本地计算的用户的数目为 $N - N_c$ 。

考虑到用户的时延和能耗需求, 本文将通过优化卸载决定矩阵 $\mathbf{A}$ 和子信道分配矩阵 $\mathbf{C}$ 以及功率分配矩阵 $\mathbf{P}$ 来最小化用户的总能耗, 本文给出了优化目标函数如式(11)

$$\left. \begin{aligned} \min Z(\mathbf{A}, \mathbf{C}, \mathbf{P}) &= \sum_{n=1}^N a_n E_n^c + (1 - a_n) E_n^l \\ \text{s.t. C1: } (1 - a_n) T_n^l + a_n (T_n^t + T_n^c) &\leq T_n^{\max}, \forall n \in \mathcal{N} \\ \text{C2: } \sum_{k \in K} c_n^k p_n^k &\leq P_{\max}, \forall n \in \mathcal{N} \\ \text{C3: } p_n^k &\geq 0, \forall n \in \mathcal{N} \\ \text{C4: } c_n^k &\in \{0, 1\}, \forall n \in \mathcal{N} \\ \text{C5: } a_n &\in \{0, 1\}, \forall n \in \mathcal{N} \end{aligned} \right\} \quad (11)$$

其中, C1 表示用户所能容忍的最大时延要求, C2 表示用户的发送功率约束, C3 表示每个子信道上的发送功率是非负的, C4 表示信道的分配状态, C5 表示卸载决定是一个二进制变量。

由于整数约束的存在, 使式(11)成为一个混合整数的非线性规划问题, 是一个 NP-hard 问题。因此, 为了降低求解的难度, 将上述问题转化为卸载决定、信道分配和功率分配 3 个子问题, 并采用一个迭代求解的方法得到问题的最优解。

3.1 基于坐标下降法的卸载决定优化

用 $\mathbf{A} = [a_1, a_2, \dots, a_N]$ 表示所有用户的卸载决

定, 给定初始卸载决定 $\mathbf{A}^0$ 为全1矩阵, $\mathbf{A}^{l-1}$ 表示第 $l-1$ ($l=1, 2, \dots, L$) 次迭代时的卸载决定, L 表示最终的迭代次数, 相应地用 $V(\mathbf{A}^{l-1})$ 表示在给定卸载决定为 $\mathbf{A}^{l-1}$ 时目标函数式(11)的最优值, 定义 Q_n^l 为第 l 次迭代时改变当前的卸载决定后所获得的收益, 则

$$Q_n^l = V(\mathbf{A}^{l-1}) - V(\mathbf{A}^{l-1}(n)) \quad (12)$$

其中, $\mathbf{A}^{l-1}(n)$ 表示用户 n 改变当前决定后的卸载决定, 更新规则如式(13)

$$\mathbf{A}^{l-1}(n) = [a_1^{l-1}, a_2^{l-1}, \dots, a_n^{l-1} \oplus 1, \dots, a_N^{l-1}] \quad (13)$$

其中, $\oplus$ 表示模二加方法。

坐标下降法每一次沿变量 a_n 的方向连续优化, 从而找到目标函数的局部最小值, 所以通过有限次迭代, 算法可以达到收敛, 从而得到最优的卸载决定。在第 l 次迭代后获得卸载决定 $\mathbf{A}^l$, 通过式(12)计算, 若收益 $Q_{n_l^*}^l > 0$, 则 $\mathbf{A}^l = \mathbf{A}^{l-1}(n_l^*)$, 其中, $n_l^* = \operatorname{argmax}_{n=1, \dots, N} Q_n^l$, 表示在第 l 次迭代中获得收益最大的用户。

3.2 卸载用户的信道分配

在子信道分配阶段, 假设用户都以最大发送功率 $P_{\max}$ 进行传输, 在每个子信道上用户都以平均功率进行数据传输, 即 $p_n^k = P_{\max} / \sum_{k \in \mathcal{K}} c_n^k$ 。采用部分频率复用的方式来进行子信道分配, 在分配的过程中, 为了更好地评估干扰, 在每次的子信道分配时, MEC将会根据基站报告的监测信息构建用户之间的干扰图 $G(V, E)$, 其中 V 表示用户, E 表示干扰关系, 边权重表示用户之间的干扰大小, 边权重的计算如式(14)

$$e_{n,m} = \begin{cases} 0, & n = m \quad n, m \in \mathcal{N}_c \\ p_n h_{n,m}, & n \neq m \quad n, m \in \mathcal{N}_c \end{cases} \quad (14)$$

其中, $h_{n,m}$ 表示SBS m 的用户到SBS n 的信道增益。因此, $e_{n,m}$ 表示SBS n 的用户受到SBS m 用户的干扰大小。

同时, 在子信道分配前, MEC将会构建子信道分配矩阵 $\mathbf{C}_{N_c \times K}$ 和干扰矩阵 $\boldsymbol{\varphi}_{N_c \times K}$ 。在子信道分配前 $\mathbf{C}_{N_c \times K}$ 和 $\boldsymbol{\varphi}_{N_c \times K}$ 均初始化为全零矩阵。其中, 子信道分配矩阵 $\mathbf{C}_{N_c \times K}$ 的每个元素 c_n^k 为0或1, 干扰矩阵中的每一个元素 φ_n^k 表示的是用户 n 在子信道 k 上受到来自其他卸载用户的总干扰。因此, 由式(14)可以将 φ_n^k 表示为

$$\varphi_n^k = \sum c_n^k e_{n,m} \quad (15)$$

在子信道分配的过程中, 本文每次将信道质量

最好的子信道分配给用户来最大化用户的上行传输速率。同时, 在满足用户时延需求下, 为每个用户分配尽可能少的子信道来避免由于用户过多的频率复用而产生严重的干扰。因此, 根据约束条件C1和C3, 每个用户的子信道分配问题可以规划为如式(16)

$$\left. \begin{aligned} & \min \sum_{k=1}^K c_n^k \\ & \text{s.t. C1: } c_n^k \in \{0, 1\}, \forall n \in \mathcal{N}_c \\ & \text{C2: } \sum_{k=1}^K c_n^k \cdot B \log_2(1 + \text{SINR}_n^k) \geq R_n^{\min}, \\ & \quad \forall n \in \mathcal{N}_c \\ & \text{C3: } R_n^{\min} = \frac{d_n f^c}{T_n^{\max} f^c - w_n}, \forall n \in \mathcal{N} \\ & \text{C4: } 0 < \sum_{k=1}^K c_n^k \leq K, \forall n \in \mathcal{N}_c \end{aligned} \right\} \quad (16)$$

对于上述子信道分配问题, 可以等效为 N_c 个用户与 K 个子信道的匹配问题, 首先采用改进的匈牙利算法进行一次信道匹配, 然后采用贪婪算法在满足最低速率需求下继续为用户分配足够的子信道, 算法步骤如下

(1) 根据式(2)构建第1次迭代所需的效益矩阵 $\mathbf{R}_{N_c \times K} = B \log_2(1 + \text{SINR}_n^k)$ 。

(2) 若用户数大于子信道数, 即 $N_c > K$, 则添加 $N_c - K$ 个虚拟子信道, 将效益矩阵变为 $N_c \times N_c$ 方阵, 若用户数小于子信道数目, 即 $N_c < K$, 则添加 $K - N_c$ 个用户, 将效益矩阵变为 $K \times K$ 的方阵。

(3) 采用匈牙利算法进行最大权重匹配得到1次信道分配。

(4) 根据分配的子信道结果由式(14)和式(15)更新子信道分配矩阵 $\mathbf{C}_{N_c \times K}$ 和干扰矩阵 $\boldsymbol{\varphi}_{N_c \times K}$ 。

(5) 查看每个用户是否满足最低速率需求, 若满足则算法终止。若不满足, 更新需要继续分配子信道的用户为 $\mathcal{N}'_c$ 。

(6) 检查信道分配矩阵 $\mathbf{C}_{N_c \times K}$, 对 $\mathcal{N}'_c$ 中的每一个用户采用贪婪算法从剩余子信道中选择产生干扰最小的子信道分配给该用户。

(7) 重复步骤(4)–(6), 直到所有的用户都满足最低速率需求或 $\mathcal{N}'_c = \emptyset$, 则算法终止。

3.3 卸载用户的功率分配

完成信道分配后, 原始的优化问题转化为求解最优的发送功率

$$\left. \begin{aligned} \mathbf{P1}: \min Z(\mathbf{P}) &= \sum_{n=1}^{N_c} \frac{d_n \sum_{k \in \mathcal{K}} c_n^k p_n^k}{\sum_{k \in \mathcal{K}} c_n^k B \log_2(1 + \text{SINR}_n^k)} \\ &\quad + P_n^i \cdot \frac{w_n}{f^c} \\ \text{s.t. C1: } T_n^t + T_n^c &\leq T_n^{\max}, \forall n \in \mathcal{N}_c \\ \text{C2: } \sum_{k \in \mathcal{K}} p_n^k &\leq P_{\max}, \forall n \in \mathcal{N}_c \\ \text{C3: } p_n^k &\geq 0, \forall n \in \mathcal{N}_c \end{aligned} \right\} \quad (17)$$

上述问题 $\mathbf{P1}$ 是非凸的优化问题, 考虑到约束条件C1是最大时延约束, 因此, 能耗的最小化问题可以转化为时延约束下的最小功率消耗问题, 因此我们将 $\mathbf{P1}$ 转化为

$$\left. \begin{aligned} \mathbf{P2}: \min_{p_n^k} \sum_{n=1}^{N_c} \sum_{k \in \mathcal{K}} p_n^k \\ \text{s.t. C1: } R_n &\geq \frac{d_n f^c}{T_n^{\max} f^c - w_n}, \forall n \in \mathcal{N}_c \\ \text{C2: } \sum_{k \in \mathcal{K}} p_n^k &\leq P_{\max}, \forall n \in \mathcal{N}_c \\ \text{C3: } p_n^k &> 0, \forall n \in \mathcal{N}_c \end{aligned} \right\} \quad (18)$$

由于优化问题 $\mathbf{P2}$ 仍然是一个非凸的优化问题, 通过变量替换, 令 $p_n^k = e^{S_n^k}$, 则

$$r_n^k = B \log_2 \left(1 + \frac{h_{n,n}^k e^{S_n^k}}{\omega_0 + \sum_{m=1, m \neq n}^{N_c} h_{m,n}^k e^{S_m^k}} \right) \quad (19)$$

将优化问题 $\mathbf{P2}$ 转化为

$$\left. \begin{aligned} \mathbf{P3}: \min_{S_n^k, r_n^k, R_n} \sum_{n=1}^{N_c} \sum_{k \in \mathcal{K}} e^{S_n^k} \\ \text{s.t. C1: } R_n &\geq \frac{d_n f^c}{T_n^{\max} f^c - w_n}, \forall n \in \mathcal{N}_c \\ \text{C2: } \sum_{k \in \mathcal{K}} e^{S_n^k} &\leq P_{\max}, \forall n \in \mathcal{N}_c \\ \text{C3: } r_n^k &\leq B \log_2 \left(1 + \frac{h_{n,n}^k e^{S_n^k}}{\omega_0 + \sum_{m=1, m \neq n}^{N_c} h_{m,n}^k e^{S_m^k}} \right), \\ &\quad \forall n \in \mathcal{N}_c \\ \text{C4: } R_n &= \sum_{k \in \mathcal{K}} r_n^k, \forall n \in \mathcal{N}_c \end{aligned} \right\} \quad (20)$$

将式(19)的等式关系变成了不等式约束, 目的是在于将这个非凸的问题转化为凸问题, 这个改变并不影响问题的最优解, 因为用户 n 在子信道 k 上的数据速率在最优性上不能小于 $B \log_2(1 + h_{n,n}^k e^{S_n^k} / \omega_0 + \sum_{m=1, m \neq n}^N h_{m,n}^k e^{S_m^k})$ 。

定理 问题 $\mathbf{P3}$ 在高信干噪比下是一个凸优化问题。

证明 由于目标函数是指数求和的形式, 因此是一个凸函数, 同时可以看出约束条件C1, C2, C4都是凸的, 对于约束C3, 由于吞吐量函数的非凸性, 所以约束C3是非凸的。在高信干噪比条件下, 对于吞吐量函数的非凸性, 一种通用的处理办法是有效近似, 使得 $\log_2(1+x) \approx \log_2(x)$ ^[17]。因此, 约束条件C3可以转化为

$$r_n^k + \log_2 \left(\omega h_{n,n}^{k-1} e^{-S_n^k} + \sum_{n=1, n \neq m}^N h_{n,n}^{k-1} h_{n,m}^k e^{S_n^k - S_m^k} \right) \leq 0 \quad (21)$$

其中, $\log_2 \left(\omega h_{n,n}^{k-1} e^{-S_n^k} + \sum_{n=1, n \neq m}^N h_{n,n}^{k-1} h_{n,m}^k e^{S_n^k - S_m^k} \right)$ 是对数的指数和形式, 因此是凸的。综上得到, 在高信干噪比条件下, $\mathbf{P3}$ 是一个凸优化问题。证毕

对于上述凸优化问题 $\mathbf{P3}$, 最优的功率分配结果可以使用内点法进行求解^[18]。

表1给出了本文求解最优的卸载决定、子信道分配和功率分配的具体迭代求解步骤。

4 仿真结果及分析

假设一个集中式的MEC网络, MEC服务器位

表1 任务卸载和资源分配算法

输入: 用户数 N , $t_n = (w_n, d_n, T_n^{\max})$, f^c , 初始卸载决定 $\mathbf{A}^0$ 。
初始化: $l \leftarrow 0$,
Repeat
$l \leftarrow l + 1$
for $n = 1:N$
根据式(13)得到 $\mathbf{A}^{l-1}(n)$;
采用改进的匈牙利算法和贪婪算法得到子信道分配矩阵 $\mathbf{C}_{N_c \times K}$;
根据凸优化问题 $\mathbf{P3}$ 采用内点法求解得到每个子信道上的最优的发送功率 p_n^k ;
根据式(12)计算 Q_n^l ;
end
$q_l^* \leftarrow \max_{n=1, \dots, N} Q_n^l$ 和 $n_l^* \leftarrow \operatorname{argmax}_{n=1, \dots, N} Q_n^l$;
更新 $\mathbf{A}^l \leftarrow \mathbf{A}^{l-1}(n_l^*)$;
Until $q_l^* \leq 0$;
输出: 卸载决定矩阵 $\mathbf{A}^*$, 信道分配矩阵 $\mathbf{C}_{N_c \times K}^*$, 功率分配矩阵 $\mathbf{P}^*$ 。

于宏基站附近, $N(N=5, 10, \dots, 30)$ 个SBS随机地分布在这个网络中, 每个SBS的覆盖范围为30 m, 本文使用的路径损耗因子是参考标准的SBS模型^[19], 其它仿真参数如表2所示。本文分析了所提算法的性能与用户数目、时延约束和输入数据等之间的关系, 并与其它算法进行相关性能的对比。

表2 仿真参数

参数	取值
子信道带宽 B	0.2 MHz
子信道个数	20
用户最大发送功率 $P_{\max}$	23 dBm
空闲时电路功率消耗 P^i	10 mW
背景噪声功率 ω_0	-100 dBm
用户的计算能力 f_n^l	0.1~1 GHz/周期
计算任务的大小 d_n	400~1200 kB
需要的CPU周期 w_n	0.2~1 GHz
用户容忍最大时延 $T_n^{\max}$	1~4 s
MEC的计算能力 f^c	4 GHz/周期

考虑到仿真参数取值范围对本文算法性能的影响, 规定了3种不同类型的时延约束取值范围: 类型1=1~3 s, 类型2=1~4 s, 类型3=1~5 s。然后在不同时延取值范围下分析了本文算法的系统性能。从图2中可以看到在时延约束范围比较小时, 更多的用户选择将任务卸载到MEC计算, 随着时延约束范围取值的增大, 选择卸载的用户越来越少。这是因为时延约束范围越小, 时延敏感的用户越多, 而时延较为敏感时, 本地计算CPU消耗较大, 能耗较高, 卸载往往比本地计算性能更优, 因而此时选择卸载的比例越高。而随着时延约束范围增大, 时延敏感用户的比例在下降, 时延约束越大本地CPU消耗越小, 相应的本地计算能耗越小, 用户更倾向于在本地计算。

图3是在不同时延约束范围下本文算法的系统

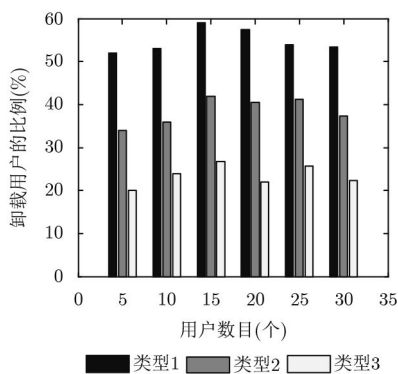


图2 不同时延约束范围下卸载用户的比例

总能耗的对比图, 可以看出随着时延约束范围的增大系统的总能耗是在减小。结合图2分析, 这是因为随着时延约束范围增大, 选择本地计算的用户越来越多, 而时延约束越大则本地计算节能越多。另外, 在时延约束范围较小时, 时延敏感的用户较多, 对于这些用户, 由于卸载时分得的无线资源较多, 基于频率复用而造成的同频干扰将更严重, 由此将会造成系统能耗的增加。因此, 随着时延约束范围的增大系统的总能耗越来越小。

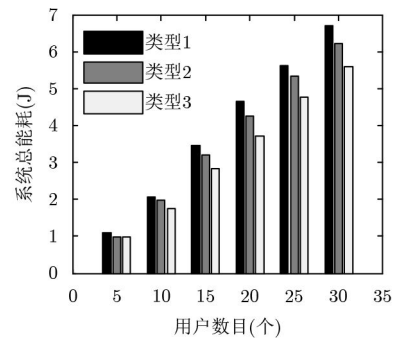


图3 不同时延约束范围下的系统总能耗

图4中在 $N=20$ 下分析了时延约束对本文算法性能的影响。将本文算法与本地计算、全部卸载和文献[15]的资源分配算法进行性能对比。可以看到随着时延约束的增大, 本文算法、本地计算和文献[15]算法的系统总能耗都是呈下降趋势, 这是因为随着时延约束的增大, 本地计算将会有更优的性能, 更多的用户将选择本地计算, 所以能耗呈下降趋势。另外从图4中可以看到本文算法相比于文献[15]算法能得到更低的能耗, 当时延约束较小时, 更能体现本文算法的优越性, 这是因为时延约束越小, 基于用户的最低速率需求, 用户将会分得较多的资源, 因此频率复用的程度就会越大, 小区之间的同频干扰将会更严重。而本文算法相比于文献[15]算法在信道分配的基础上进一步优化了用户的功率分配, 减轻干扰的同时降低了能耗, 因此得到

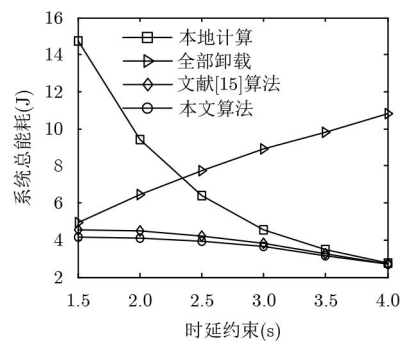


图4 系统的能耗与时延约束

了更优的系统性能。而对于全部卸载的情况,可以看到随着时延约束增大系统总能耗是上升的,这是因为时延约束越大,用户卸载时延越大,相应的能耗就越高。

考虑到输入数据大小对本文算法的性能影响,图5中描述了在 $N = 20$ 时随着输入数据增大系统总能耗的变化情况。将本文算法与其它算法进行对比,可以看出随着输入数据的增大本文算法相比其它算法有更低的总能耗。此外也可以看到随着输入数据的不断增大,用户趋向于全部卸载,这是因为随着输入数据增大,用户所需的CPU周期也不断增大,本地计算能耗不断增加,由于本文的资源分配方法可以保证时延并且最小化能耗,这时卸载将比本地计算性能更好,所以越多的用户将会选择将任务卸载计算。

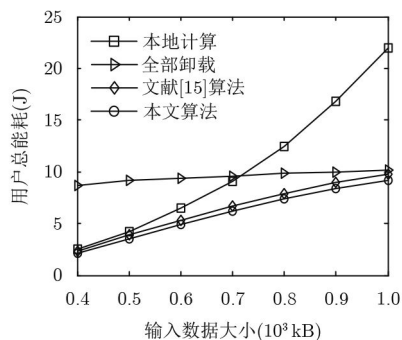


图5 系统的能耗与输入数据大小

图6描述的是随着用户数量增加系统总能耗的变化情况。与其它算法对比可以看出本文算法有更低的总能耗。其中相比于文献[15]的算法,本文算法考虑整体的任务卸载优化方案,在满足时延约束下为用户进行了有效的信道分配,并且在考虑系统总能耗下最小化用户的发送功率。所以本文算法相比于文献[15]的算法能得到更优的卸载决定和资源分配方案,因此在系统性能上有明显的提升。

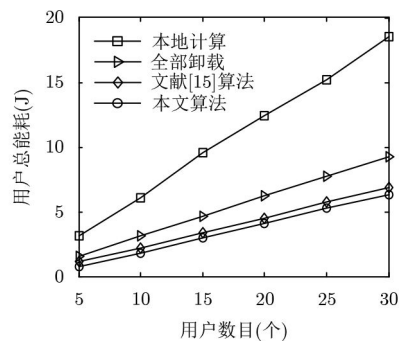


图6 系统的能耗与用户数目

5 结论

本文讨论了超密集组网下基于移动边缘计算的任务卸载和资源分配问题。考虑用户时延的需求下,提出了以最小化用户的总能耗为目标的优化问题,然后采用了基于坐标下降法的任务卸载方案,并在每次迭代的过程中对用户进行合理的子信道和功率分配。最终,迭代求解得到了最优的卸载决定和资源分配结果,有效地降低了系统的能耗,提升了系统的整体性能。

参考文献

- [1] WANG Shiqiang, ZAFER M, and LEUNG K K. Online placement of multi-component applications in edge computing environments[J]. *IEEE Access*, 2017(5): 2514–2533. doi: [10.1109/ACCESS.2017.2665971](https://doi.org/10.1109/ACCESS.2017.2665971).
- [2] MAO Yuyi, YOU Changsheng, ZHANG Jun, et al. A survey on mobile edge computing: the communication perspective[J]. *IEEE Communications Surveys & Tutorials*, 2017, 19(4): 2322–2358. doi: [10.1109/COMST.2017.2745201](https://doi.org/10.1109/COMST.2017.2745201).
- [3] PAN Jianli and MCELHANNON J. Future edge cloud and edge computing for internet of things applications[J]. *IEEE Internet of Things Journal*, 2018, 5(1): 439–449. doi: [10.1109/JIOT.2017.2767608](https://doi.org/10.1109/JIOT.2017.2767608).
- [4] YANG Bin, MAO Guoqiang, DING Ming, et al. Dense small cell networks: from noise-limited to dense interference-limited[J]. *IEEE Transactions on Vehicular Technology*, 2018, 67(5): 4262–4277. doi: [10.1109/TVT.2018.2794452](https://doi.org/10.1109/TVT.2018.2794452).
- [5] GE Xiaohu, TU Song, MAO Guoqiang, et al. 5G ultra-dense cellular networks[J]. *IEEE Wireless Communications*, 2016, 23(1): 72–79. doi: [10.1109/MWC.2016.7422408](https://doi.org/10.1109/MWC.2016.7422408).
- [6] YANG Lichao, ZHANG Heli, LI Ming, et al. Mobile edge computing empowered energy efficient task offloading in 5G[J]. *IEEE Transactions on Vehicular Technology*, 2018, 67(7): 6398–6409. doi: [10.1109/TVT.2018.2799620](https://doi.org/10.1109/TVT.2018.2799620).
- [7] ZHANG Jiao, HU Xiping, NING Zhaolong, et al. Energy-latency tradeoff for energy-aware offloading in mobile edge computing networks[J]. *IEEE Internet of Things Journal*, 2018, 5(4): 2633–2645. doi: [10.1109/JIOT.2017.2786343](https://doi.org/10.1109/JIOT.2017.2786343).
- [8] LIU Jianhui and ZHANG Qi. Offloading schemes in mobile edge computing for ultra-reliable low latency communications[J]. *IEEE Access*, 2018, 6: 12825–12837. doi: [10.1109/ACCESS.2018.2800032](https://doi.org/10.1109/ACCESS.2018.2800032).
- [9] MAO Yuyi, ZHANG Jun, SONG S H, et al. Stochastic joint radio and computational resource management for multi-user mobile-edge computing systems[J]. *IEEE Transactions on Wireless Communications*, 2017, 16(9): 5994–6009. doi: [10.1109/TWC.2017.2717986](https://doi.org/10.1109/TWC.2017.2717986).
- [10] TI N T and LE Longbao. Computation offloading leveraging

- computing resources from edge cloud and mobile peers[C]. Proceedings of 2017 IEEE International Conference on Communications, Paris, France, 2017: 1–6. doi: [10.1109/ICC.2017.7997138](https://doi.org/10.1109/ICC.2017.7997138).
- [11] ZHAO Pengtao, TIAN Hui, QIN Cheng, *et al.* Energy-saving offloading by jointly allocating radio and computational resources for mobile edge computing[J]. *IEEE Access*, 2017(5): 11255–11268. doi: [10.1109/ACCESS.2017.2710056](https://doi.org/10.1109/ACCESS.2017.2710056).
- [12] ZHANG Jing, XIA Weiwei, YAN Feng, *et al.* Joint computation offloading and resource allocation optimization in heterogeneous networks with mobile edge computing[J]. *IEEE Access*, 2018, 6: 19324–19337. doi: [10.1109/ACCESS.2018.2819690](https://doi.org/10.1109/ACCESS.2018.2819690).
- [13] GUO Jun, ZHANG Heli, YANG Lichao, *et al.* Decentralized computation offloading in mobile edge computing empowered small-cell networks[C]. Proceedings of 2017 IEEE Globecom Workshops, Singapore, Singapore, 2017: 1–6. doi: [10.1109/GLOCOMW.2017.8269049](https://doi.org/10.1109/GLOCOMW.2017.8269049).
- [14] RANADHEERA S, MAGHSUDI S, and HOSSAIN E. Computation offloading and activation of mobile edge computing servers: a minority game[J]. *IEEE Wireless Communications Letters*, 2018, 7(5): 688–691. doi: [10.1109/LWC.2018.2810292](https://doi.org/10.1109/LWC.2018.2810292).
- [15] WANG Chenmeng, YU F R, LIANG Chengchao, *et al.* Joint computation offloading and interference management in wireless cellular networks with mobile edge computing[J]. *IEEE Transactions on Vehicular Technology*, 2017, 66(8): 7432–7445. doi: [10.1109/TVT.2017.2672701](https://doi.org/10.1109/TVT.2017.2672701).
- [16] DINH T Q, TANG Jianhua, LA Q D, *et al.* Offloading in mobile edge computing: task allocation and computational frequency scaling[J]. *IEEE Transactions on Communications*, 2017, 65(8): 3571–3584. doi: [10.1109/TCOMM.2017.2699660](https://doi.org/10.1109/TCOMM.2017.2699660).
- [17] RAM S S, VEERAVALLI V V, and NEDIC A. Distributed non-autonomous power control through distributed convex optimization[C]. Proceedings of IEEE INFOCOM 2009, Rio de Janeiro, Brazil, 2009: 3001–3005. doi: [10.1109/INFOCOM.2009.5062275](https://doi.org/10.1109/INFOCOM.2009.5062275).
- [18] LIU Peng, LI Jiandong, LI Hongyan, *et al.* Convex optimisation-based joint channel and power allocation scheme for orthogonal frequency division multiple access networks[J]. *IET Communications*, 2015, 9(1): 28–32. doi: [10.1049/iet-com.2014.0409](https://doi.org/10.1049/iet-com.2014.0409).
- [19] 3GPP organizational partners. Evolved universal terrestrial radio access (E-UTRA); Further advancements for E-UTRA physical layer aspects (Release 9), document TS 36.814, 3GPP[OL]. <http://www.3gpp.org/ftp/2012>.

张海波：男，1979年生，副教授，研究方向为无线资源管理。

李 虎：男，1992年生，硕士生，研究方向为移动边缘计算、无线资源管理。

陈善学：男，1966年生，教授，研究方向为图像处理、数据压缩。

贺晓帆：男，1985年生，助理教授，研究方向为无线资源优化。