

doi: 10.12052/gdutxb.220152

移动边缘计算系统的双服务器协同与 计算通信资源联合优化

李宇龙, 梁静轩, 王 丰

(广东工业大学 信息工程学院, 广东 广州 510006)

摘要: 为充分利用移动边缘计算(Mobile Edge Computing, MEC)系统的计算资源, 本文设计了双MEC服务器协同与计算通信资源联合优化方案。通过建模双服务器协同多用户任务计算优化问题, 以系统计算时延和用户能耗的加权和最小化为准则, 优化设计多用户计算卸载发射功率和计算任务分割。提出一种较低计算复杂度的联合设计方案, 将原问题解耦为计算卸载优化和计算任务分割设计的2个子问题, 分别通过内点法和单纯形法实现快速数值求解。仿真结果表明, 本文所提算法的系统性能优于已有启发式基准算法方案, 且在较少的算法运行时间下, 联合优化算法方案能取得与最优拉格朗日乘子法基准方案相同的系统性能。

关键词: 移动边缘计算; 计算卸载; 双服务器协同; 资源优化

中图分类号: TN929.5

文献标志码: A

文章编号: 1007-7162(2024)01-0093-08

Optimized Design and Resource Allocation for Dual-server Mobile Edge Computing Systems

Li Yu-long, Liang Jing-xuan, Wang Feng

(School of Information Engineering, Guangdong University of Technology, Guangzhou 510006, China)

Abstract: In order to make full use of the computing resources of Mobile Edge Computing (MEC) system, this paper designs an optimization scheme of collaboration between two MEC servers and joint computing and communication resources. The scheme proposes the optimization problem of dual-server collaborative multi-user task calculation, where the weighted sum between system computing delay and user energy consumption is minimized. The multi-user computing unloaded transmitting power and task segmentation are optimized in the proposed scheme. A joint design scheme with low computational complexity is proposed. The original problem is decoupled into two sub-problems of computational offload optimization and computational task segmentation design, both of which can be solved by interior point method and simplex method respectively. The simulation results show that the system performance of the proposed scheme is better than the existing heuristic benchmark algorithm scheme. And the joint optimization algorithm scheme can get the similar system performance as compared with the basic scheme of the optimal Lagrange multiplier method with less computation time.

Key words: mobile edge computing; computation offloading; dual-server collaboration; resource optimization

随着互联时代的快速发展, 众多计算密集型和时延敏感型应用兴起, 如在线游戏、虚拟/增强现实、智能驾驶等。这使计算能力有限、电池电量不足和存储容量有限的移动终端设备面临着巨大的挑战。云

计算可以将到达终端的任务卸载到云端服务器进行处理, 其强大处理能力能有效地降低计算时延, 并显著降低设备耗电量的速度。近几年, 随着无线终端设备数量的剧增, 海量数据被上传到云端服务器, 造

收稿日期: 2022-10-09

基金项目: 国家自然科学基金资助项目(61901124); 广东省自然科学基金资助项目(2021A1515012305); 广州市科技计划项目(202102020856)

作者简介: 李宇龙(1999-), 男, 硕士研究生, 主要研究方向为移动边缘计算资源管理

通信作者: 王丰(1987-), 男, 副教授, 主要研究方向为通信信号处理、移动边缘计算资源管理等, E-mail: fengwang13@gdut.edu.cn

成网络堵塞,服务效率降低。移动边缘计算提供了一种解决上述问题的解决方案,通过在靠近无线终端设备的网络边缘端部署服务器来处理设备卸载的计算任务,分担云端服务器的任务处理压力,进一步降低任务处理时延和设备的电量消耗^[1-4]。

移动边缘计算中的计算卸载策略主要分为二进制卸载和部分卸载^[5-8],二进制卸载将计算任务完整地进行卸载,而部分卸载将计算任务分割为多个部分后卸载其中部分任务。在二进制卸载策略下,文献[9]针对MEC系统网络中移动终端设备因计算能力不足,在处理低时延、高可靠应用时产生的高时延问题,提出一种基于博弈论的求解方法,用于联合优化计算卸载与资源分配;文献[10]考虑了最小化系统的执行延迟和用户能耗的加权和,通过联合优化任务卸载决策、卸载调度和功率分配,设计了一种两级交替优化框架来求解非凸混合整数优化问题;文献[11]考虑了最小化用户卸载决策下的时延,通过联合优化计算卸载和资源分配问题;文献[12]从用户的角度考虑了最小化用户能耗和传输、计算成本,通过联合优化计算卸载和资源分配问题,提出了一种基于博弈论的求解方法;文献[13]考虑了两层MEC系统,以最小化系统服务时延和用户能耗为目标,联合优化了服务缓存和计算卸载策略;文献[14]考虑了一种能量感知卸载方案,通过最小化能耗和时延之间的权衡,在能量有限和延迟敏感的条件下联合优化通信和计算资源分配。上述文献中,任务的决策会取决于终端设备的运行状态,系统的性能对设备的要求较高。

部分卸载策略下,文献[15]考虑了在随机应用请求、不可预测的无线设备状态、时变的信道状态以及计算资源等因素的基础上最小化系统的时延与能耗;联合优化多用户的部分卸载决策、传输调度和计算资源分配;并提出了一种在线资源协调与分配方案。文献[16]考虑了最小化系统总延迟,联合优化计算卸载和无线传输调度问题;文献[17]考虑了MEC服务器部署在宏基站和小型基站下的多用户场景,通过联合优化用户与基站之间的计算卸载、用户与小型基站的连接、用户的CPU (Central Processing Unit) 频率和发射功率,达到MEC系统的总能耗最小化的目的。文献[18]通过联合优化用户的任务分配比例和卸载发射功率来最小化设备之间的任务延迟。文献[19]将非正交多址技术引入到MEC系统中,通过联合优化任务卸载和资源分配以最大化系统处理能力。文

献[20]在考虑卸载延迟和保密约束的前提下,通过联合优化通信和计算资源分配实现总保密卸载数据最大化的部分卸载比。文献[21]在多用户多服务器MEC系统下,利用改进的深度强化学习算法,优化服务器的选择和卸载任务比例以最小化计算时延与能耗。然而在部分卸载模式下,大部分工作仅考虑了时延或能耗的单目标优化,未考虑时延和能耗的均衡优化。

不同于以上研究,本文考虑在双MEC服务器与多用户协同场景下,以最小化系统计算时延和用户能耗加权和为目标,在有限的时间内完成用户的任务并将任务计算结果发送到远处数据驱动端进行驱动。由于用户和单MEC服务器进行卸载时会有其他MEC服务器空闲,造成资源浪费,因此考虑双服务器协同任务处理方式,最终完成数据驱动端的驱动。本文通过联合优化计算任务分割和卸载发射功率以最小化系统时延与用户能耗加权和,提出了一种较低计算复杂度的算法方案实现最优任务分割和发射功率。

本文的主要贡献总结如下:(1) 针对双MEC服务器协同多用户场景,基于用户能耗与系统计算时延加权和最小化准则,建模本地计算和双MEC服务器协同计算的联合优化问题;在计算时延和发射功率约束条件下,联合优化用户计算任务分割和计算卸载发射功率。(2) 针对该联合设计问题,提出一种较低计算复杂度的联合优化算法方案,将原问题解耦成计算卸载和计算任务分割的2个子问题,分别采用内点法和单纯形法实现快速求解。(3) 大量数值仿真实验表明,本文所提联合优化算法的系统性能优于已有启发式基准方案,并在较少的算法运行时间下,本算法方案能取得与最优拉格朗日乘子法基准方案相同的系统性能。

1 系统模型与问题构造

考虑一个包含数据驱动端的双MEC服务器协同多用户系统。如图1所示,该模型包含 K 个无线设备、2个部署了MEC服务器的边缘基站和1个数据驱动端。假设每个用户通过频分多址(Frequency Division Multiple Access, FDMA) 技术接入基站,用 $i \in \mathcal{K}$ 表示图1中第 i 个无线设备,其中 $\mathcal{K} \triangleq \{1, \dots, K\}$ 表示 K 个无线设备组成的集合。每个无线设备通过本地和卸载处理计算任务^[15-21]。本文考虑通过无线设备和双MEC服务器协同计算来完成到达无线设备的计算任务,并将任务计算结果传输到数据驱动端。假设基站1和基站2之间、基站2和数据驱动端之间通过光纤进行通信^[17]。

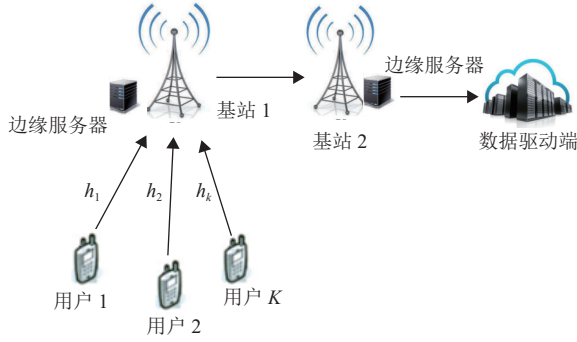


图1 基于双服务器协同的多用户系统模型

Fig.1 Multi-user system model based on cooperation of two servers

1.1 任务模型

本文采用三元物理量 $(\alpha_i, \beta_i, \gamma_i)$ 来表示第 i 个无线设备的计算任务,其中 α_i 为无线设备 i 的计算任务的输入数据量(单位:bit), β_i 为完成计算任务所需的CPU周期数以及 γ_i 为完成计算任务后输出的结果数据量(单位:bit)。假设这些计算任务是可以被任意分割的^[22],且对应的输入数据量、所需CPU周期数和输出数据量以相同的比例进行分割。令 $x_{1,i}$ 、 $x_{2,i}$ 和 $x_{3,i}$ 分别为无线设备 i 、基站1和基站2需要完成的计算任务相应的比例,并且有式(1)的限制条件。

$$\begin{cases} x_{1,i} + x_{2,i} + x_{3,i} = 1 \\ x_{j,i} \in [0, 1], j \in \{1, 2, 3\} \end{cases} \quad (1)$$

式中: $i = 1, \dots, K$ 。

1.2 时延模型

1.2.1 无线设备计算、传输的时延

第 i 个无线设备处理任务时,产生的时延主要由任务处理和传输输入数据、输出数据3部分组成。第 i 个无线设备进行本地计算时所需的时间为

$$t_{i,x_1,\beta_i} = \frac{x_{1,i}\beta_i}{f_i} \quad (2)$$

式中: $i = 1, \dots, K$; $x_{1,i}\beta_i$ 为第 i 个无线设备需要通过本地计算完成的计算量; β_i 为完成相应计算任务所需的CPU周期数; f_i 为其进行本地计算时的CPU频率。

接下来建模无线设备传输输出数据和部分输入数据所需的时间。记 r_i 为第 i 个无线设备和基站1无线通信时的传输速率。本文采用基于FDMA的多用户计算卸载通讯模式,记 B_i 为系统分配给第 i 个无线设备的上下行链路的频谱带宽。根据香农公式,第 i 个无线设备的上下行链路传输速率 r_i 为

$$r_i = B_i \log_2 \left(1 + \frac{p_i h_i}{\sigma^2} \right) \quad (3)$$

式中: $i = 1, \dots, K$; p_i 为该无线设备通信时的发射功率; $h_i > 0$ 为该无线设备与基站1的上下行链路信道增益; σ^2 为基站接收机的高斯白噪声功率。因此,第 i 个无线设备的输入输出数据的传输时间表示为^[22]

$$t_{i,tran} = \frac{(x_{2,i} + x_{3,i})\alpha_i + x_{1,i}\gamma_i}{r_i} \quad (4)$$

式中: α_i 和 γ_i 分别为到达无线设备 i 的计算任务的输入数据量和计算结果的数据量。结合式(2)和式(4),第 i 个无线设备进行本地计算和数据传输所产生的时延,记为

$$T_{WD_i} = t_{i,x_1,\beta_i} + t_{i,tran} \quad (5)$$

1.2.2 基站产生的计算、传输的时延

基站1产生的时延包括计算部分任务、传输部分输入数据和传输输出数据3部分。基站1计算部分任务产生的时延为

$$t_{mec1,\beta_i} = x_{2,i}\beta_i / f_{mec1} \quad (6)$$

式中: $i = 1, \dots, K$; f_{mec1} 为基站1使用其CPU计算部分任务时使用的CPU频率; β_i 为完成相应计算任务所需的CPU周期数。

接下来分别建模基站1和基站2传输输入数据和输出数据所需的时间。本文假设基站端和数据驱动端采用光纤通信,因此,基站与基站间、基站与数据驱动端进行数据传输的传输速率都为 R 。基站1需要传输的数据内容包括来自多个无线设备和基站1的任务结果数据以及基站2进行任务处理需要的输入数据,该传输过程所需的时间为

$$t_{mec1,\alpha_i,\gamma_i} = \frac{(x_{1,i} + x_{2,i})\gamma_i}{R} + \frac{x_{3,i}\alpha_i}{R} \quad (7)$$

式中: $i = 1, \dots, K$; α_i 和 γ_i 分别为到达无线设备 i 的计算任务的输入数据量和计算结果的数据量。结合式(6)和(7)处理第 i 个无线设备所需的时间 $T_{mec1,i}$,其表达式为

$$T_{mec1,i} = t_{mec1,\beta_i} + t_{mec1,\alpha_i,\gamma_i} \quad (8)$$

基站2产生的时延同样包括任务计算和数据传输两部分,区别在于任务分配比例。基站2完成对应比例的任务计算量所需的时间以及传输输出数据到数据驱动端所需的时间分别为

$$\begin{cases} t_{mec2,\beta_i} = \frac{x_{3,i}\beta_i}{f_{mec2}} \\ t_{mec2,\gamma_i} = \frac{\gamma_i}{r} \end{cases} \quad (9)$$

式中: $i = 1, \dots, K$; β_i 和 γ_i 分别为完成相应计算任务所需的CPU周期数和计算结果的数据量; f_{mec2} 为基站

2计算任务分配的CPU频率。记 $T_{\text{mec2},i}$ 为基站2完成任务计算和完成数据传输所需的时间和,其表达式为

$$T_{\text{mec2},i} = t_{\text{mec2},\beta_i} + t_{\text{mec2},\gamma_i} \quad (10)$$

式中: $i = 1, \dots, K$ 。

结合式(5)、(8)和(10)可得系统处理第 i 个无线设备的所需的总时间。记 T_i 为系统完成到达第 i 个无线设备的计算任务并将计算结果传输到数据驱动端所需的总时间,考虑下一部分任务计算需要上一部分任务计算的结果进行驱动,其表达式为^[23]

$$T_i = T_{\text{WD},i} + T_{\text{mec1},i} + T_{\text{mec2},i} \quad (11)$$

式中: $i = 1, \dots, K$ 。系统在给定的时限内完成所有的计算任务并将计算结果传输到数据驱动端。建模计算任务完成所需时延约束为

$$T_i \leq T_{\text{DL}} \quad (12)$$

式中: $i = 1, \dots, K$; T_{DL} 是给定的计算任务完成时限。

1.3 能耗模型

本文考虑无线设备在任务计算以及数据传输过程的能量消耗。无线设备 i 由于任务计算产生的能耗表示为

$$E_i^{\text{loc}} = \kappa x_{1,i} \beta_i f_i^2 \quad (13)$$

式中: $i = 1, \dots, K$; κ 为与处理器的结构相关的能耗常数^[24]。第 i 个无线设备传输输入数据和输出数据的所需时间由式(4)给出,由此可得数据传输能耗为

$$E_i^{\text{off}} = p_i t_{i,\text{tran}} \quad (14)$$

结合式(13)和(14),计算第 i 个无线设备进行任务计算和数据传输的总能耗,具体为

$$E_i = E_i^{\text{loc}} + E_i^{\text{off}} \quad (15)$$

1.4 优化问题构造

分别记 w_1, w_2 为用户执行任务的能耗和系统时延的偏好因子,并且 $w_1 + w_2 = 1$,可以根据任务完成时间的需求和剩余电池寿命来进行设置。本文以最小化系统时延和无线设备能耗的加权和为目标提出以下优化问题。

$$\begin{aligned} \min_{x_{1,i}, x_{2,i}, x_{3,i}, p_i} & \sum_{i=1}^K (w_1 E_i + w_2 T_i), \\ \text{s.t.} & \begin{cases} \text{C1: } \sum_{j=1}^3 x_{ji} = 1, \forall i \in \mathcal{K} \\ \text{C2: } p_i \leq P^{\text{max}}, \forall i \in \mathcal{K} \\ \text{C3: } T_i \leq T_{\text{DL}}, \forall i \in \mathcal{K} \\ \text{C4: } 0 \leq x_{ji} \leq 1, j \in \{1, 2, 3\}, \forall i \in \mathcal{K} \end{cases} \end{aligned} \quad (16)$$

式中: $x_{1,i}, x_{2,i}, x_{3,i}$ 为对所有任务的分配比例;约束C1保证系统通过协同完成所有任务;约束C2为任务卸载

发射功率的范围;约束C3为全部任务完成并传输到数据驱动段所需的时间必须不大于时限 T_{DL} 。在偏好因子确定的情况下,通过设置合理的时延约束 T_{DL} 来实现MEC的性能偏好。

2 问题(16)优化分析与求解

由于问题(16)目标函数中任务分配比例和发射功率是耦合的,因此该问题是一个非凸优化问题。为此,本文设计一种较低计算复杂度的算法方案,将原问题解耦成任务分割和计算卸载2个子问题,首先利用内点法求解计算卸载子问题,再利用单纯形法求解最优任务分割子问题。

2.1 传输功率优化

将原问题解耦成计算卸载子问题,其中问题(16)可以变换为子问题(17)。

$$\begin{aligned} \min_{\{p_i\}_{i=1}^K} & \sum_{i=1}^K \left[\left(\frac{p_i}{r_i} + \frac{1}{r_i} \right) \right], \\ \text{s.t.} & \text{C1: } p_i \leq P^{\text{max}}, \forall i \in \mathcal{K} \end{aligned} \quad (17)$$

考虑到目标函数关于传输功率的复杂性,引入辅助变量 $o_i = \frac{1}{r_i}, \forall i \in \mathcal{K}$,则问题(17)可以变换为问题(18)。

$$\begin{aligned} \min_{\{o_i\}_{i=1}^K} & \sum_{i=1}^K \left[\left(\frac{(2^{o_i B_i} - 1) \sigma^2}{h_i} o_i + o_i \right) \right], \\ \text{s.t.} & \text{C1: } o_i \geq M_i, \forall i \in \mathcal{K} \end{aligned} \quad (18)$$

式中: $M_i = \frac{1}{B_i \log_2(1 + P^{\text{max}} h_i / \sigma^2)}$ 。问题(18)是一个凸问题,可以使用内点法进行求解^[25]。

证明 假设 $\chi(o) \triangleq o(2^{(\frac{1}{B_i o})} - 1)$,关于 $\chi(o)$ 的二阶导数为 $\frac{d^2 \chi(o)}{do^2} \triangleq 2 \left(\frac{1}{B_i o} \frac{\ln^2 2}{B_i^2 o^2} \right) \geq 0, \forall o > 0$,因此问题(18)的目标函数为凸,其约束条件为线性函数,所以(18)是一个凸问题。利用问题(18)构造惩罚函数为

$$\varphi(o_i, v_i) = f(o_i) - v_i \ln(o_i - M_i) \quad (19)$$

式中: $f(o_i)$ 为式(18)中的目标函数, v_i 为迭代过程中的惩罚因子。算法1对求解式(18)所采用的内点法进行了总结。

算法1 求解问题(18)的内点法。

- (1) 输入:输入惩罚因子 $v_i > 0$;允许误差 $\varepsilon > 0$;选取初始点 $o_i^0, k = 1$;递减系数 $c = 0.1$ 。
- (2) 迭代:令式(19)一阶导数为零,计算函数 $\varphi(o_i, v_i)$ 的极值点 $(o_i^*, v_i^{(k)})$,更新 $v_i^{k+1} = c v_i^k, o_i^k = o_i^* v_i^k, k = k + 1$ 。
- (3) 直到: $o_i^k - o_i^{k-1} \leq \varepsilon o_i^* = o_i^k$ 。
- (4) 输出:输出问题(18)的最优解 (o_i^*) 。

2.2 任务分配比例优化

在取得最优发射功率 p_i^* 的情况下,问题(16)可

以转化为求解任务分割子问题。将优化问题(16)的C1约束转化为 $x_{3,i} = 1 - x_{1,i} - x_{2,i}$,并与目标函数和其他约束条件结合,形成下述优化问题(20)。

$$\begin{aligned} \min_{\{x_{1,i}, x_{2,i}\}_{i=1}^K} & \sum_{i=1}^K (A_{1,i}x_{1,i} + C_{1,i}x_{2,i}) + D_{1,i}, \\ \text{s.t.} & \begin{cases} \text{C1: } x_{1,i} + x_{2,i} \leq 1, \forall i \in \mathcal{K} \\ \text{C2: } F_i x_{1,i} + G_i x_{2,i} \leq \\ \quad T_{\text{DL}} - \frac{\gamma_i}{R} - \left(\frac{\alpha_i}{R} + \frac{\beta_i}{f_{\text{mec2}}} + \frac{\alpha_i}{r_i} \right), \\ \quad \forall i \in \mathcal{K} \\ \text{C3: } x_{1,i}, x_{2,i} \in [0, 1], \forall i \in \mathcal{K} \end{cases} \end{aligned} \quad (20)$$

式中:

$$\begin{cases} F_i = \left(\frac{\beta_i}{f_i} + \frac{\gamma_i}{r_i} + \frac{\gamma_i}{R} \right) - \frac{\alpha_i}{R} - \frac{\beta_i}{f_{\text{mec2}}} - \frac{\alpha_i}{r_i} \\ G_i = \left(\frac{\beta_i}{f_{\text{mec1}}} + \frac{\gamma_i}{R} \right) - \left(\frac{\alpha_i}{R} + \frac{\beta_i}{f_{\text{mec2}}} \right) \\ A_{1,i} = \left(\kappa \beta_i f_i^2 + \frac{p_i \gamma_i}{r_i} + \frac{\beta_i}{f_i} + \frac{\gamma_i}{r_i} + \frac{\gamma_i}{R} \right) - \\ \quad \left(\frac{\alpha_i p_i}{r_i} + \frac{\alpha_i}{R} + \frac{\beta_i}{f_{\text{mec2}}} + \frac{\alpha_i}{r_i} \right) \\ C_{1,i} = \left(\frac{\alpha_i p_i}{r_i} + \frac{\beta_i}{f_{\text{mec1}}} + \frac{\gamma_i}{R} \right) - \\ \quad \left(\frac{\alpha_i p_i}{r_i} + \frac{\alpha_i}{R} + \frac{\beta_i}{f_{\text{mec2}}} \right) \\ D_{1,i} = \left(\frac{\alpha_i p_i}{r_i} + \frac{\alpha_i}{R} + \frac{\beta_i}{f_{\text{mec2}}} + \frac{\alpha_i}{r_i} \right) + \left(\frac{\gamma_i}{R} \right) \end{cases}$$

本文利用单纯形算法求解变换后的优化问题(20)。

首先,设置优化问题的一个初始可行解;其次,在初始可行解条件下,如果存在非初始变量 $s_i \geq 0$,则找出最大非初始变量作为换入变量;然后,从初始变量中找到换出变量,进行初等行列变换,将新的初始变量转换为单位矩阵。最后,循环迭代此过程直到非初始变量为非正。因此,对于问题(20)添加非初始变量 s_i ,即可得到问题(21),表示为

$$\begin{aligned} \min_{\{x_{1,i}, x_{2,i}, s_i\}_{i=1}^K} & \sum_{i=1}^K (w_1 E_i + w_2 T_i), \\ \text{s.t.} & \begin{cases} \text{C1: } x_{1,i} + x_{2,i} + s_{1,i} = 1, \forall i \in \mathcal{K} \\ \text{C2: } A_{2,i}x_{1,i} + C_{2,i}x_{2,i} + s_{2,i} = D_{2,i}, \forall i \in \mathcal{K} \\ \text{C3: } x_{1,i}, x_{2,i} \in [0, 1], s_i \geq 0, \forall i \in \mathcal{K} \end{cases} \end{aligned} \quad (21)$$

式中:

$$\begin{cases} A_{2,i} = \left(\frac{\beta_i}{f_i} + \frac{\gamma_i}{r_i} + \frac{\gamma_i}{R} \right) - \left(\frac{\alpha_i}{R} + \frac{\beta_i}{f_{\text{mec2}}} + \frac{\alpha_i}{r_i} \right) \\ C_{2,i} = \left(\frac{\beta_i}{f_{\text{mec1}}} + \frac{\gamma_i}{R} \right) - \left(\frac{\alpha_i}{R} + \frac{\beta_i}{f_{\text{mec2}}} \right) \\ D_{2,i} = T_{\text{DL}} - \frac{\gamma_i}{R} - \left(\frac{\alpha_i}{R} + \frac{\beta_i}{f_{\text{mec2}}} + \frac{\alpha_i}{r_i} \right) \end{cases}$$

优化问题(21)是一个标准的线性规划问题,算法2对求解问题(21)所使用的单纯形法的求解步骤进行了总结。

算法2 求解问题(21)的单纯形法。

(1) 输入:输入初始可行解 $\mu^{(0)} = (x_{1,i}^{(0)}, x_{2,i}^{(0)}, s_{j,i}^{(0)})$ 和对应的索引集; $m = 0$ 。

(2) 迭代:变换初始变量并保持可行性;若 $\mu^{(m)}$ 的每个邻可行解大于 $\mu^{(m)}$,则 $\mu^{(m+1)}$ 为无边界。否则 $\mu^{(m+1)} = \mu^{(m)}$ 的每个邻可行解小于 $\mu^{(m)}$ 。

(3) 直到: $\mu^{(m)}$ 为局部最优。则有

$$\mu^* = \mu^{(m)} = (x_{1,i}^{(*)}, x_{2,i}^{(*)}, s_{j,i}^{(*)})$$

(4) 输出:输出问题(21)的最优解 $(x_{1,i}^{(*)}, x_{2,i}^{(*)})$ 。

3 仿真结果分析与对比

仿真结果验证了本模型的有效性和优越性。考虑在双MEC服务器协同多用户场景下,将系统任务计算结果发送到远处数据驱动端进行驱动,其中每个基站覆盖范围为100 m, d_i 为无线设备至基站的距离,其信道功率增益设置为 $h_i = \Gamma_i d_i^{-3.5} \bar{h}_i$ [26],其中 $\bar{h}_i \sim \mathcal{CM}(0, 1)$ 为小尺度衰落效应, $\Gamma_i = -32$ dBm为用户 i 参考距离为1 m时的路径损耗。未作额外说明时,该模型的系统参数设置如下:系统分配给第 i 个无线设备的上下行链路的频谱带宽 $B_i = 20$ MHz[9],基站接收机的高斯白噪声功率 $\sigma^2 = 10^{-9}$ W,任意无线设备 i 采用 $f_i = 1.5$ GHz的CPU频率来处理任务;第1个和第2个基站服务器的计算频率为 $f_{\text{mec1}} = 2$ GHz、 $f_{\text{mec2}} = 4$ GHz;无线设备的最大通信发射功率 $p_{\text{max}} = 1$ W[17]; $\kappa = 10^{-27}$ 为无线设备进行本地计算时与处理器的结构相关的能耗常数[25];最大时延约束为 $T_{\text{DL}} = 1.3$ s;偏好因子 $\omega_1 = 0.5, \omega_2 = 0.5$;基站与基站间、基站与数据驱动端进行数据传输的传输速率 $R = 1$ Gbps。

为了比较算法性能,本文考虑了5种已有的启发式基准方案和1种最优拉格朗日法基准方案,即:(1)启发式基准方案1:完全本地计算,所有任务都在无线设备端进行本地计算。(2)启发式基准方案2:单服务器卸载,所有任务仅由无线设备和第一个基站分配。(3)启发式基准方案3:混合比例卸载,所有任务由本地和服务器按比例 $x_{1,i} : x_{2,i} : x_{3,i} = \frac{1}{2} : \frac{1}{4} : \frac{1}{4}$ 完成任务分配和计算。(4)启发式基准方案4:固定无线设备发射功率,即令 $p_i = 0.4$ W。(5)启发式基准方案5:粒子群算法[27],其中学习因子 $C_1 = 2, C_2 = 2$,惯性权重 $w = 0.4$ 。(6)最优拉格朗日乘子法基准方案6:拉格朗日乘子法[22],采用内点法进行迭代求解。

图2显示了本文所提方案与最优拉格朗日乘子法基准方案6的运行时间随不同无线设备数目变化的曲线。随着无线设备数目 K 的增加,本文所提方案和最优拉格朗日法基准方案6的运行时间均增加。本文所提方案的运行时间明显少于最优拉格朗日乘子

法基准方案6。随着 K 增大,本文所提方案和最优拉格朗日法基准方案6的运行时间差距呈现增大的趋势。例如,较最优拉格朗日法基准方案6,当 $K=17$ 时,本文所提算法能节省1.5 s的运行时间,当 $K=45$ 时,本文所提算法能节省2.6 s的运行时间。

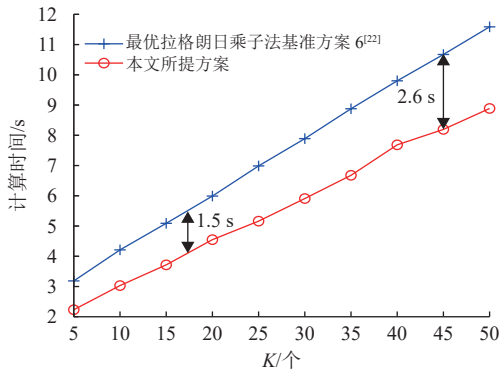


图2 不同无线设备数目 K 下的系统运行时间

Fig.2 System runtime under different number of wireless device K

图3显示了能耗与系统时延加权随不同输入数据量 α_i 变化的性能曲线,其中 $K=10$,输出数据量 $\gamma_i=10^6$ bit,由图可知,6种方案的时延与能耗的加权和性能均随着输入数据量的增大而增加,其中本文提出的优化方案优于其他基准方案1、2、3和4。考虑到时延约束的影响,当输入数据较小时,可以全卸载进行计算,基准方案5可以很快收敛到最优系统性能,随着输入数据的增大,本文所提方案的时延与能耗加权和会逐渐接近基准方案3,数据会分配到本地进行处理;基准方案5中的粒子群可能会陷入局部最优,但可以通过较低的复杂度接近本文所提方案的系统性能;本文提出方案能取得和基准方案6相似的系统性能,但在大规模用户的情况下,该算法的计算复杂度远大于本文提出方案。

图4显示了能耗与系统时延加权随不同输出数据量 γ_i 变化的性能曲线,其中 $K=10$,输入数据量 $\alpha_i=10^7$ bit。由图可知,随着输出数据量的增加,基准方案具有缓慢的增长趋势。本文提出的方案在时延约束下优于其他已有的启发式基准方案。基准方案1的时延与能耗加权和不如本文所提方案和其他基准方案,这说明将计算任务卸载到边缘服务器进行计算能有效地减少时延与能耗的加权和。且本文所提方案在大规模用户条件下能以较低的复杂度达到与基准方案6相同的系统性能。

图5显示了能耗与系统时延加权随不同时限 T_{DL} 变化的性能曲线,其中无线设备数量 $K=10$,任务输入数据量 $\alpha_i=10^7$ bit,任务的输出数据量 $\gamma_i=10^6$ bit。

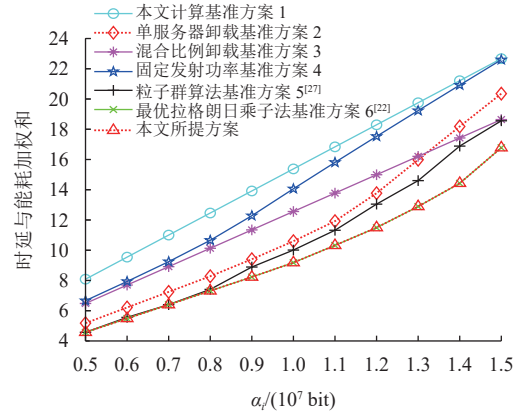


图3 不同输入数据量 α_i 下的时延与能耗加权和性能

Fig.3 Time delay and energy consumption weighted sum versus different input data size α_i

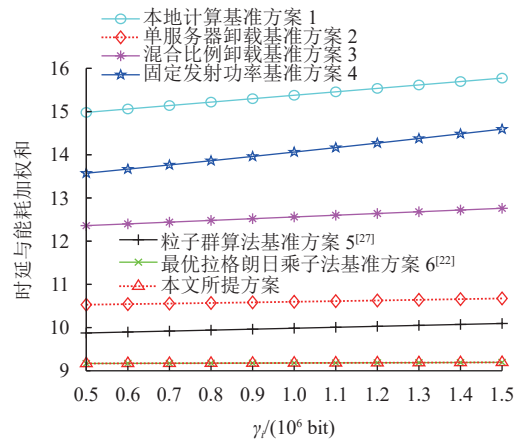


图4 不同输出数据量 γ_i 下的时延与能耗加权和性能

Fig.4 Time delay and energy consumption weighted sum versus output data size γ_i

由图5可知,随着所需时延约束的增加,本文所提方案和基准方案2、4、5和6都在减小,系统将数据更多地卸载到边缘端进行计算,使得总体性能提升,当 T_{DL} 取值较大时,虽然系统性能有所提升,但会使得任务完成总时延过大。当时延约束大于1.5 s时,本文提出方案和基准方案5、6均达到最优,但会造成过多的时延浪费,当时延约束小于1.5 s时,本文所提优化方案的性能优于其他基准方案1、2、3、4、5,并以较低的复杂度达到基准方案6所示的最优解。

图6显示了系统时延和能耗的加权和随着无线设备数量的增长变化的性能曲线。无线设备的数量从5增加到50,其中无线设备任务输入数据量 $\alpha_i=10^7$ bit,任务的输出数据量 $\gamma_i=10^6$ bit。由图6可知,随着无线设备数量的增加,优化方案和其他基准方案的时延与能耗加权和都在增加,这是因为无线设备个数的增加使得系统需要处理的计算任务增加,完成这些增加的计算任务需要更多时间与能耗。与已有的启

发式基准方案1、2、3、4和5相比,本文所提优化方案有较大的优势;且能在较少的算法运行时间下达到最优拉格朗日乘子法基准方案6的系统性能。

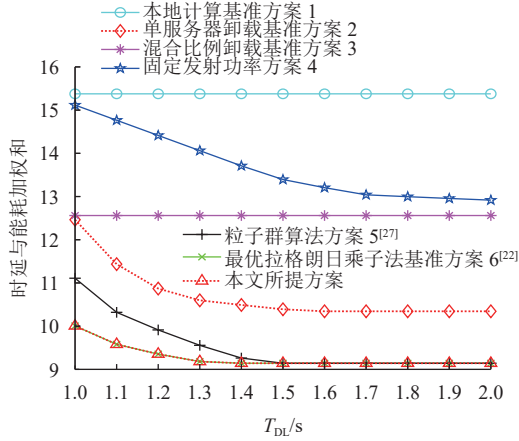


图5 不同时延长度 T_{DL} 下的时延与能耗加权和性能
Fig.5 Time delay and energy consumption weighted sum versus different delay length T_{DL}

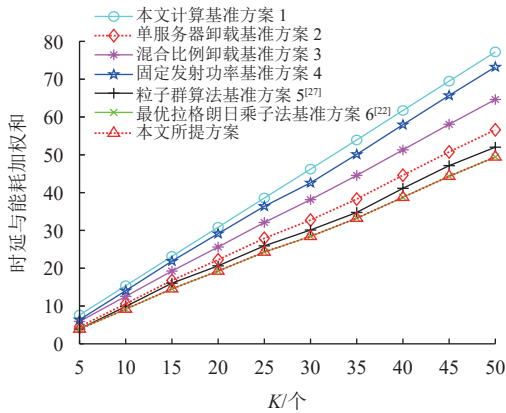


图6 不同无线设备个数 K 下的时延与能耗加权和性能
Fig.6 Time delay and energy consumption weighted sum under different number of wireless device K

图7显示了能耗与系统时延加权和随无线设备计算能力变化的性能曲线,其中无线设备数目 $K=10$,任务输入数据量 $\alpha_i=10^7$ bit,任务的输出数据量 $\gamma_i=10^6$ bit。由图7可知,随着无线设备计算能力的增长,本地能耗会增大,而计算时延会减小,基准方案1由于完全本地计算,能耗的影响会大于计算时延的影响,所以增长速率最快。基准方案2考虑到单服务器资源有限,增长速率会高于本文所提方案。基准方案4更倾向于卸载任务到基站端进行计算,无线设备计算能力的增长对其没有影响。基准方案5能够达到较低的计算复杂度,但会陷入局部最优,而本文所提方案能够以相对基准方案6较小的复杂度达到其最优解。

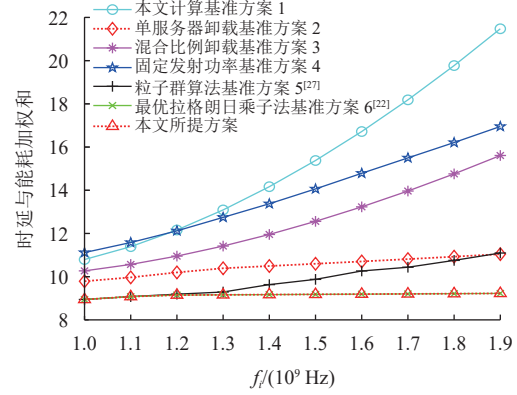


图7 不同无线设备计算能力 f_i 下的时延与能耗加权和性能
Fig.7 Time delay and energy consumption weighted sum of wireless devices versus computing capacities f_i

4 结论

本文研究了双MEC服务器协同与计算通信资源联合优化设计方案,以系统计算时延和用户能耗的加权和最小化为准则,联合优化用户计算任务分割和计算卸载发射功率。本文提出一种较低计算复杂度的算法方案,将原联合设计问题解耦为计算卸载和任务分割的2个子问题,分别采用内点法和单纯形法实现快速数值求解。仿真结果表明,本文所提算法的系统性能优于已有的启发式基准方案,并在较低的算法运行时间下,本文所提方案能取得与最优拉格朗日乘子法基准方案相同的系统性能。

参考文献:

- [1] MACH P, BECVAR Z. Mobile edge computing: a survey on architecture and computation offloading [J]. IEEE Communications Surveys & Tutorials, 2017, 19(3): 1628-1656.
- [2] MAO 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.
- [3] 赵炫宇. 资源受限的移动边缘计算系统中计算卸载问题研究[D]. 北京: 北京邮电大学, 2019.
- [4] WANG F, XU J, DING Z. Multi-antenna noma for computation offloading in multiuser mobile edge-computing systems [J]. IEEE Transactions on Communications, 2019, 67(3): 2450-2463.
- [5] HUYNH L N T, PHARN Q V, PHAM O V, et al. Efficient computation offloading in multi-tier multi-access edge computing systems: a particle swarm optimization approach [J]. Applied Sciences, 2020, 10(1): 203.
- [6] 景泽伟, 杨清海, 秦猛. 移动边缘计算中的时延和能耗均衡优化算法[J]. 北京邮电大学学报, 2020, 43(2): 110-115. JING Z W, YANG Q H, QIN M. A delay and energy

- tradeoff optimization algorithm for task offloading in mobile edge computing network [J]. *Journal of Beijing University of Posts and Telecommunications*, 2020, 43(2): 110-115.
- [7] GONG Y. Optimal edge server and service placement in mobile edge computing [C]// 2020 IEEE 9th Joint International Information Technology and Artificial Intelligence Conference. Chongqing: IEEE, 2020: 688-691.
- [8] GUO H, LIU J. Collaborative computation offloading for multiaccess edge computing over fiberwireless networks [J]. *IEEE Transactions on Vehicular Technology*, 2018, 67(5): 4514-4526.
- [9] 龙隆, 刘子辰, 石晶林, 等. 移动边缘计算中计算卸载与资源分配的联合优化策略[J]. *高技术通讯*, 2020, 30(8): 765-773.
- LONG L, LIU Z C, SHI J L, *et al.* Joint optimization strategy of service cache and resource allocation in mobile edge network [J]. *High Technology Letters*, 2020, 30(8): 765-773.
- [10] KUANG Z, LI L, GAO J, *et al.* Partial offloading scheduling and power allocation for mobile edge computing systems [J]. *IEEE Internet of Things Journal*, 2019, 6(4): 6774-6785.
- [11] TANG L, HU H. Computation offloading and resource allocation for the internet of things in energy constrained multiaccess edge networks [J]. *IEEE Access*, 2020, 8: 47509-47521.
- [12] ZHANG J, XIA W W, ZHANG Y Y, *et al.* Joint offloading and resource allocation optimization for mobile edge computing [C]// IEEE Global Communications Conference. Singapore: IEEE, 2017: 1-6.
- [13] FENG H, GUO S, YANG L, *et al.* Collaborative data caching and computation offloading for multi-service mobile edge computing [J]. *IEEE Transactions on Vehicular Technology*, 2021, 70(9): 9408-9422.
- [14] ZHANG J, HU X P, NING Z L, *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.
- [15] PENG J, QIU H, CAI J, *et al.* D2D-assisted multi-user cooperative partial offloading transmission scheduling and computation allocation for MEC [J]. *IEEE Transactions on Wireless Communications*, 2021, 20(8): 4858-4873.
- [16] NING Z, DONG P, KONG X, *et al.* A cooperative partial computation offloading scheme for mobile edge computing enabled internet of things [J]. *IEEE Internet of Things Journal*, 2019, 6(3): 4804-4814.
- [17] BI J, YUAN H, ZHANG K, *et al.* Energy-minimized partial computation offloading for delay sensitive applications in heterogeneous edge networks [J]. *IEEE Transactions on Emerging Topics in Computing*, 2022, 10(4): 1941-1954.
- [18] FANG F, XU Y, DING Z, *et al.* Optimal resource allocation for delay minimization in NOMA-MEC networks [J]. *IEEE Transactions on Communications*, 2020, 68(12): 7867-7881.
- [19] XUE J, AN Y. Joint task offloading and resource allocation for multi-task multi-server NOMA-MEC networks [J]. *IEEE Access*, 2021, 9: 16152-16163.
- [20] XU J, ZHU P, LI J, *et al.* Secure computation offloading for multi-user multi-server MEC-enabled IoT [C]// IEEE International Conference on Communications. Montreal: IEEE, 2021: 1-6.
- [21] SHANG C, SUN Y, LUO H. A hybrid deep reinforcement learning approach for dynamic task offloading in NOMA-MEC system [C]// IEEE International Conference on Sensing, Communication, and Networking (SECON). Stockholm: IEEE, 2022: 434-442.
- [22] 代美玲, 刘周斌, 郭少勇, 等. 基于终端能耗和系统时延最小化的边缘计算卸载及资源分配机制[J]. *电子与信息学报*, 2019, 41(11): 2684-2690.
- DAI M L, LIU Z B, GUO S Y, *et al.* A computation offloading and resource allocation mechanism based on minimizing devices energy consumption and system delay [J]. *Journal of Electronics & Information Technology*, 2019, 41(11): 2684-2690.
- [23] FAN W H, HAN J T, YAO L, *et al.* Latency-energy optimization for joint wifi and cellular offloading in mobile edge computing networks [J]. *Comput Networks*, 2020, 181: 107570.
- [24] CAO X W, WANG F, XU J, *et al.* Joint computation and communication cooperation for energy-efficient mobile edge computing [J]. *IEEE Internet of Things Journal*, 2019, 6(3): 4188-4200.
- [25] LUO Z Q, MA W K, SO A M C, *et al.* Semidefinite relaxation of quadratic optimization problems [J]. *IEEE Signal Processing Magazine*, 2010, 27(3): 20-34.
- [26] 王丰, 李宇龙, 林志飞, 等. 基于计算吞吐量最大化的能量采集边缘计算系统在线资源优化配置[J]. *广东工业大学学报*, 2022, 39(4): 17-23.
- WANG F, LI Y L, LIN Z F, *et al.* Online resource allocation design for computation capacity maximization in energy harvesting mobile edge computing systems [J]. *Journal of Guangdong University of Technology*, 2022, 39(4): 17-23.
- [27] 李顺, 葛海波, 刘林欢, 等. 移动边缘计算中的协同计算卸载策略[J]. *计算机工程与应用*, 2022, 58(21): 83-90.
- LI S, GE H B, LIU L H, *et al.* Collaborative computing offloading strategy in mobile edge computing [J]. *Computer Engineering and Applications*, 2022, 58(21): 83-90.

(责任编辑: 张玮欣 英文审核: 熊荣斌)