

多用户移动边缘计算迁移的能量管理研究

王璐瑶, 张文倩, 张光林

(东华大学信息科学与技术学院, 上海 201620)

摘要: 在移动边缘计算系统中, 通过将计算任务从移动设备迁移到移动边缘计算服务器, 可以大幅度提高计算质量。考虑将可再生能源纳入多用户移动边缘计算系统中, 并在模型中加入电池作为能量收集装置以实现能量收集和存储。通过提出的基于强化学习的资源管理算法制定了移动边缘计算系统中的任务分配策略, 实现了移动设备成本最优化(包括时延成本和能耗成本)。仿真结果表明, 与其他算法相比, 该算法显著减少了移动设备的成本。

关键词: 能量收集; 可再生能源; 移动边缘计算; 成本优化; 强化学习

中图分类号: TN929.5

文献标识码: A

doi: 10.11959/j.issn.2096-3750.2019.00091

Research on energy management of multi-user mobile edge computing offloading

WANG Luyao, ZHANG Wenqian, ZHANG Guanglin

College of Science Technology and Information, DongHua University, Shanghai 201620, China

Abstract: In mobile edge computing system, the quality of computing experience can be improved greatly by offloading computing tasks from mobile devices to mobile edge computing servers. Consider incorporating renewable energy into a multi-user mobile edge system. Moreover, a battery as an energy harvesting device was added to the model to harvest energy and storage. The task allocation strategy in mobile edge computing system was formulated through the resource management algorithm based on reinforcement learning, which achieved the cost minimization of mobile devices (including delay cost and computing cost). The simulation results show that the proposed algorithm significantly minimizes the cost of mobile devices compared with other algorithms.

Key words: energy harvesting, renewable energy, mobile edge computing, cost optimization, reinforcement learning

1 引言

1.1 背景介绍

近年来, 移动设备(如智能手机、笔记本电脑等)成为人们学习、娱乐和社交的重要工具。由于移动设备的资源限制(处理能力、电池使用寿命和存储容量等), 单纯的移动设备已不能给予用户较好的使用体验。

移动设备通过云计算可以快速访问集中式可配置计算资源池, 并将所有分布式数据和计算密集型应用程序转移到位于远程大型数据中心的云。由

于数据必须在移动设备和云数据中心之间进行传输和接收, 导致网络负载过高以及网络带宽需求不断增长。这不仅对当今已经拥塞的骨干网造成沉重负担, 还会导致传输时延较大, 从而降低服务质量^[1-3]。

由此, 移动边缘计算(MEC, mobile edge computing)^[1-3]应运而生, 它可以直接在网络边缘处理某些工作负载, 不需要将其迁移到云端再进行处理。

在 MEC 系统中, 如基站、接入点和路由器等网络边缘设备被赋予类似云的计算和存储能力, 作

为云的替代物为用户服务。同时,由于其被放置在靠近移动设备终端和数据源位置,因此可以显著减少传输时延。

能量收集装置可以从外界收集可再生能源,如太阳辐射、无线电信号和风能等,逐渐成为绿色通信系统的重要组成部分。考虑电网产生的碳足迹以及不断上涨的电价等因素,通过风能或太阳辐射收集的可再生能源被视为急待利用的新兴绿色能源。

尽管在边缘处理计算任务的行为减少了移动设备端的能耗,但是将任务从移动设备迁移至边缘设备会产生一定的时延成本。通过制定合理的移动设备迁移策略(即将多少计算任务分配给边缘设备处理),以优化移动设备的时延成本和能耗成本。

1.2 相关研究

MEC 近年来受到越来越多的关注,其概念在 2014 年被提出。MEC 作为一个新平台,为靠近移动用户的无线接入网络提供 IT 服务环境和云计算功能^[4]。雾计算^[5]与 MEC 相关,两者的应用领域重叠,术语也经常互换。与移动云计算相比,MEC 的优势包括时延低、节约能源以及安全性高等^[3]。

MEC 服务器是小型数据中心,与传统的云量级数据中心相比,能耗大幅度降低^[6]。随着 MEC 服务器的广泛部署,系统的能源消耗成为人们关注的焦点^[7],因此,实现绿色 MEC 的创新技术是必要的。可再生能源(如太阳能和风能)最近成为各种 IT 系统中有前途的电力来源^[8-9],与燃煤电厂产生的传统电网能源相比,可再生能源显著减少了碳排放量。对于具有能量收集功能的移动设备,文献[10]提出基于 Lyapunov 优化算法^[11]的动态计算浮动策略。

此外,MEC 系统的任务迁移策略近年来引起业界的广泛关注^[10,12-18]。对于具有严格截止时间要求的应用,采用动态电压和频率调整(DVFS, dynamic voltage and frequency scaling)技术将本地的执行能耗降至最低,并且使用文献[12]的数据传输调度优化了闪存计算的能耗。文献[13]提出了用于飞秒计算系统的通信和计算资源的联合分配,其中,每个计算任务应在截止日期前完成。在文献[10]中,为了满足具有能量收集功能的 MEC 系统严格执行时延的要求,开发了动态计算的浮动政策。文献[14]提出了分散计算的冲击算法,以最大限度减少多用户 MEC 系统的计算开销。为了减少长期平均能耗,文献[15]提出了随机控制算法,该算法可以确定软

件组件对 MEC 系统的影响。在文献[16]中,针对单用户 MEC 系统开发了用时延最优的随机任务调度算法。文献[17]提出了用在线任务调度算法来研究具有多核移动设备 MEC 系统的能耗与时延的折中,该研究被扩展到具有不同类型的移动应用场景^[18]中。

上述部分文献关注了单用户 MEC 系统,部分文献采用离线算法。考虑了多用户的 MEC 系统,并将可再生能源纳入多用户 MEC 系统中,通过强化学习在线算法来研究多用户 MEC 迁移的能量管理。

1.3 本文研究内容

本文考虑具有多个移动设备的通用 MEC 系统,其中,计算任务以随机方式到达移动设备。主要研究怎样合理进行任务分配以达到移动设备端成本最低的目标,并实现将可再生能源纳入多用户 MEC 系统中。在此基础上,提出了基于强化学习的资源管理算法,实现了移动设备成本最优化(包括时延成本和能耗成本)。本文的主要贡献如下。

- 1) 将待解决问题转化为马尔可夫决策过程(MDP, Markov decision processes)问题,通过强化学习方法推导 Q-learning 策略^[19],在时变和未知环境中实现了实时的移动设备成本最优化。
- 2) 本文设计的奖赏函数不仅反映了实时的移动设备成本最优化,还包含历史信息。与直接的即时奖赏函数相比,本文设计的奖赏函数可以显著提高系统学习效率。
- 3) 通过仿真,将强化学习算法与短视优化算法和静态迁移算法进行比较,证明本文所提的强化学习算法消耗的成本更低。

2 系统模型

MEC 系统模型如图 1 所示。在 MEC 系统中, N 个移动设备由 MEC 服务器共同处理随机到达的任务,且由电池或可再生能源统一供电。移动设备可以通过无线信道访问边缘系统,将任务迁移至边缘设备,由边缘系统执行部分计算任务^[20]。通过使用在线学习算法将部分计算任务迁移至 MEC 服务器,使移动用户可以享受到更高质量的计算体验^[20]。

2.1 计算任务模型

考虑离散时间模型,该模型将工作周期分成具有相等长度 τ 的 T 个时隙。假设有 N 个移动设备,定义为 $N \triangleq \{1, \dots, N\}$,时隙定义为 $T \triangleq \{0, 1, \dots\}$ 。从第 t 个时隙开始,计算到达第 i 个移动设备的任务

$A_i(t)$ 。一般情况下,假设在不同时隙内到达的任务 $A_i(t)$ 是在 $[A_{i,\min}, A_{i,\max}]$ 内的独立同分布, $A_{i,\min}$ 和 $A_{i,\max}$ 分别为任务最小可能到达率和最大可能到达率。

在每个时隙内,系统通过使用在线学习算法决定在本地移动设备处理的工作 $v_i(t) \leq A_i(t)$, 剩余工作 $\lambda_i(t) = A_i(t) - v_i(t)$ 将被迁移至边缘设备进行处理。 N 个移动设备到达的总任务为 $A(t)$, 即

$$\sum_{i \in N} A_i(t) = A(t) \quad (1)$$

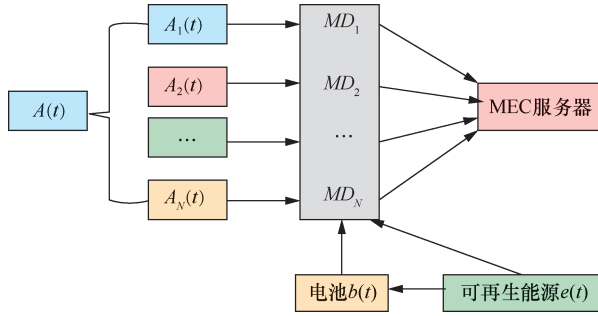


图 1 MEC 系统模型

2.2 时延模型

2.2.1 计算时延成本

对于本地处理的工作负载来说,由于移动设备的计算能力有限,时延成本 $d_{i,\text{delay}}(t)$ 主要处理时延成本 $d_{i,\text{lo}}(t)$ 。为了量化服务的时延性能,使用 $d_{i,\text{lo}}(v_i(t))$ 表示移动设备在时隙 t 内处理任务 $v_i(t)$ 所产生的处理时延成本。使用移动设备任务处理的平均响应时间乘以任务到达率来表示时延成本,具体可表示为

$$d_{i,\text{lo}}(v_i(t)) = \frac{v_i(t)}{\rho - v_i(t)} \quad (2)$$

其中, ρ 为移动设备处理任务的速率,假设每个移动设备的处理速率相同。

2.2.2 传输时延成本

对于迁移到 MEC 端被处理的工作负载,时延成本 $d_{i,\text{delay}}(t)$ 主要是在传输过程中产生的传输时延成本 $d_{i,\text{off}}(t)$,该时延是由迁移至边缘设备端进行处理的任务 $A_i(t) - v_i(t)$ 和上行无线信道条件决定的,具体可表示为

$$d_{i,\text{off}}(A_i(t), v_i(t)) = \beta \cdot (A_i(t) - v_i(t)) \quad (3)$$

其中, β 为无线信道单位传输时延,取决于上行无线信道条件。由式(3)可以看出,当任务全部在本地

处理时,即 $v_i(t) = A_i(t)$ 时,移动设备端产生的传输时延成本 $d_{i,\text{off}}(t)$ 为 0。

2.2.3 总时延成本

总时延成本 $d_{i,\text{delay}}(t)$ 是计算时延成本 $d_{i,\text{lo}}(t)$ 和传输时延成本 $d_{i,\text{off}}(t)$ 的总和,具体可表示为

$$d_{i,\text{delay}}(A_i(t), v_i(t)) = d_{i,\text{lo}}(v_i(t)) + d_{i,\text{off}}(A_i(t), v_i(t)) \quad (4)$$

2.3 能耗模型

每个时隙内的能量消耗是功率与时间长度的乘积,而模型中每个时隙长度相等,因此将能量等价于功率。

移动设备在一个时隙内的总能耗由两部分组成:首先是移动设备的基本操作和传输功率需求,其次是移动设备的计算功率需求。其中,移动设备的基本操作和传输功率需求独立于制定的迁移策略,只与到达的总任务量有关,表示为 $c_{i,\text{op}}(A_i(t)) = c_{i,\text{sta}} + c_{i,\text{dyn}}(A_i(t))$ 。其中, $c_{i,\text{sta}}$ 是静态功耗,为常数; $c_{i,\text{dyn}}(A_i(t))$ 是动态功耗,取决于到达的总任务量。假设动态功耗 $c_{i,\text{dyn}}(A_i(t))$ 与到达的总任务量 $A_i(t)$ 成正比,移动设备计算功率需求则取决于移动设备处理的工作负载。使用 $c_{i,\text{com}}(v_i(t))$ 表示计算功率需求,并假设计算功率需求 $c_{i,\text{com}}(v_i(t))$ 与在本地(移动设备)处理的工作负载 $v_i(t)$ 成正比。因此,在时隙 t 内移动设备的总能耗是基本操作和传输功率需求 $c_{i,\text{op}}(A_i(t))$ 与计算功率需求 $c_{i,\text{com}}(v_i(t))$ 的总和,可表示为

$$c_i(A_i(t), v_i(t)) = c_{i,\text{op}}(A_i(t)) + c_{i,\text{com}}(v_i(t)) \quad (5)$$

2.4 电池模型

能量收集装置被用于平衡电量供应需求。在“太阳能+风能”系统中,光伏模块和风力涡轮机可以将其输出能源相结合作为可再生能源,为移动设备供电并为能量收集装置充电^[21]。由可再生能源供电的无线通信系统通常利用蓄电池实现能量收集和存储,提供能量用于用户数据传输和处理。当可再生能源不能满足移动设备的用电需求时,由电池供电,以确保移动设备的正常运行。用 $b(t) \in [0, B]$ 表示时隙 t 开始时的电池电量。其中, B 是电池容量,单位以功率表示。考虑系统性能与安全,要求电池电量一旦低于特定的电量阈值 0,则电池必须

与负载断开、停止供电。

由于可再生能源的不确定性, 需要保证移动设备的计算功耗满足一定条件, 即 $c_{\text{com}}(v(t)) \leq \max\{b(t) - c_{\text{op}}(A(t)), 0\}$ 。如果当前电池电量不能满足移动设备的基本操作和传输功率需求时, 则激活备用电源, 边缘设备承担所有计算任务。

1) 当 $c_{\text{op}}(A(t)) \geq b(t)$ 时, 系统将所有工作负载迁移到边缘系统, 使用备用电源来满足移动设备的基本操作和传输功率需求。因此, 产生激活备用电源的成本, 定义为 $c_{\text{ba}}(t) = \varphi \cdot c_{\text{op}}(A(t))$, 其中, $\varphi > 0$, 表示由于使用备用电源而导致的高成本。此时, 电池将存储可再生能源所产生的电量, 直到电池电量达到最大容量 B , 因此, 下一个时隙的电池电量为 $b(t+1) = \min\{b(t) + e(t), B\}$ 。

2) 当 $c_{\text{op}}(A(t)) < b(t)$ 时, 移动设备可以在本地处理部分工作 $v_i(t) \leq A_i(t)$, 但计算功率需求必须满足 $c_{\text{com}}(v(t)) \leq b(t) - c_{\text{op}}(A(t))$ 。对于不同的可再生能源 $e(t)$ 和计算功率需求 $c_{\text{com}}(v(t))$, 电池相应地被充电或放电。

如果 $e(t) \geq c(v(t), A(t))$, 则完成任务处理后, 剩余的可再生能源将持续给电池充电, 直到电池电量达到最大容量 B , 因此, 下一个时隙的电池电量为

$$b(t+1) = \min\{b(t) + e(t) - c(v(t), A(t)), B\} \quad (6)$$

如果 $e(t) < c(v(t), A(t))$, 那么可再生能源将不能满足用电需求, 此时电池会放电以完成任务处理。因此, 下一个时隙的电池电量为

$$b(t+1) = b(t) + e(t) - c(v(t), A(t)) \quad (7)$$

由于电池寿命的影响, 电池可放电次数有限, 使用时隙 t 内的放电功率 $c_{\text{de}}(t)$ 表示电池损耗成本, 具体可表示为

$$c_{\text{de}}(t) = \omega \cdot \max\{c(v(t), A(t)) - e(t), 0\} \quad (8)$$

其中, $\omega > 0$ 是标准化的单位电池损耗成本。

3 问题描述

在本文模型中, 将移动设备的总成本 (即移动设备的时延成本和能耗成本) 作为衡量所提算法性能好坏的指标。

3.1 成本最小化问题

由于移动设备的电池由可再生能源提供电能, 因

此电池电量的消耗不属于成本, 移动设备的总成本是时延成本、电池损耗成本和备用电源成本的总和。

当电池电量刚好满足甚至不能满足移动设备的基本操作和传输功率需求时, 即 $c_{\text{op}}(A(t)) \geq b(t)$, 系统将所有工作负载迁移到边缘系统。此时, 备用电源将满足移动设备的基本操作和传输功率需求。因此, 移动设备的总成本为时延成本与备用电源成本之和。

当电池电量可以完全满足移动设备的基本操作和传输功率需求时, 即 $c_{\text{op}}(A(t)) < b(t)$, 移动设备可以在本地处理部分工作 $v_i(t) < A_i(t)$, 但是计算功率需求必须满足 $c_{\text{com}}(v(t)) \leq b(t) - c_{\text{op}}(A(t))$ 。由于可再生能源 $e(t)$ 和计算功率需求 $c_{\text{com}}(v(t))$ 不同, 电池相应地被充电或放电。由于电池寿命的影响, 电池可放电次数有限, 因此电池在放电过程中产生了电池损耗成本, 此时移动设备的总成本为时延成本与电池损耗成本之和。因此本文的成本最小化问题可描述为

$$\min C(t) = d_{\text{delay}}(A(t), v(t)) + c_{\text{ba}}(A(t)) + c_{\text{de}}(t) \quad (9)$$

服从关系为

$$A_{i,\min} \leq A_i(t) \leq A_{i,\max} \quad (10)$$

$$0 \leq v_i(t) \leq A_i(t) \quad (11)$$

$$0 \leq b(t) \leq B \quad (12)$$

3.2 问题分析

一般来说, 本文所提的成本最小化问题是随机优化问题, 为此, 需要在每个时隙内为移动设备确定迁移到边缘设备处理的任务量。这个问题很难解决, 因为需要制定的策略与时间相关, 当前的迁移策略会影响之后的系统性能。而下一时刻的状态, 如电池电量、可再生能源和到达的任务量均未知, 为解决该问题增加了难度。

4 强化学习算法设计

4.1 强化学习算法

为了解决成本最优化问题, 利用在线强化学习算法推导最优策略 $\pi(s)$ 。

在强化学习中, 智能体作为学习系统, 在获取外部环境当前的状态信息 s 后, 对环境采取试探动作 a , 从而进入下一状态 s' , 环境通过奖赏形式对智能体的动作进行反馈与评价。如果智能体的某个

动作 a 导致环境给予正奖赏，那么智能体以后进行此动作的趋势将会增强；如果智能体的某个动作 a 导致环境给予负奖赏，那么智能体以后进行此动作的趋势将会减弱。智能体通过不断与环境进行交互试错的方式，达到优化系统性能的目的。强化学习模型如图 2 所示。

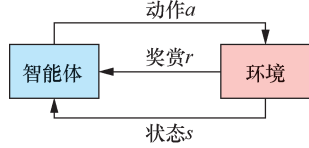


图 2 强化学习模型

任务迁移问题具有以下 4 个性质：1) 在运行时间内，不同的到达任务会导致不同的电池电量和任务分配（本地处理或迁移至 MEC 系统），未来的电池电量取决于当前迁移策略；2) 多用户 MEC 系统旨在最小化移动设备的总成本；3) 多用户 MEC 系统不确定未来到达的任务量，但可以确定过去到达的任务量及过去的迁移策略；4) 实际每一个时隙内任务的到达量不固定。因此，任务迁移问题可以通过强化学习算法来解决。

由于本文所提问题是顺序型决策问题，因此可以使用 MDP 问题解决，由 4 个元素组成：状态集合 S 、动作集合 A 、状态转移概率 $P_s(s(t+1)|s(t))$ 和奖赏函数 r_t ，其中， $\forall s \in S, \forall a \in A$ 。

4.2 状态集合

状态集合表征了移动设备当前时刻的状态，由集合 $s(t) \triangleq (A(t), b(t), e(t))$ 表示，这些状态的值在每个时隙开始时已知。在 3 个状态元素中， $A(t)$ 和 $e(t)$ 是外部状态，与迁移策略的制定无关，而电池状态 $b(t)$ 则随着可再生能源提供的电量和所制定的迁移策略改变。为了便于处理随机控制问题，以上状态元素的值被假定为有限的，因此可以将 $A(t)$ 、 $e(t)$ 演化为有限状态的马尔可夫链。令 $P_A(A(t+1)|A(t))$ 和 $P_e(e(t+1)|e(t))$ 分别表示 $A(t)$ 和 $e(t)$ 的转换概率，此假设与文献[23]和文献[24]中的假设类似。

4.3 动作集合

动作集合是移动设备在当前状态下，根据任务迁移策略所采取的动作。尽管移动设备在每个时隙 t 内采取的实际行动是 $v(t)$ （与迁移策略有关），但由于在本地处理的任务 $v(t)$ 与该时隙的计算功率需

求相互对应，因此定义动作稽核为每个时隙 t 内的计算功率需求，由 $a(t) \in A$ 表示，其中， A 是有限值的集合 $[0, 1, \dots, B]$ 。根据计算功率需求可以确定最佳迁移策略，为了在最坏情况下保持基本操作，需要满足条件为 $a(t) \leq \max\{b(t) - c_{op}(A(t)), 0\}$ ，此条件决定了在每个时隙内可制定的迁移策略。

4.4 状态转移概率

给定当前状态 $s(t)$ 、计算功率需求 $a(t)$ 和可再生能源 $e(t)$ ，下一个时隙的电池状态为

$$b(t+1) = \begin{cases} \lceil b(t) + e(t) \rceil_0^B, & c_{op}(A(t)) \geq b(t) \\ \lceil b(t) - c_{op}(A(t)) - a(t) + e(t) \rceil_0^B, & c_{op}(A(t)) < b(t) \end{cases} \quad (13)$$

其中， $\lceil \cdot \rceil_0^B$ 表示 $\max\{\min\{\cdot, B\}, 0\}$ ，系统在下一个时隙的状态转变为 $s(t+1)$ 。在给定 $a(t)$ 的情况下，从状态 $s(t)$ 转变为状态 $s(t+1)$ 的概率可以表示为

$$P_s(s(t+1)|s(t), a(t)) = P_A(A(t+1)|A(t)) P_e(e(t+1)|e(t)) \times \mathbf{1}\{\xi(t)\} \quad (14)$$

其中， $\mathbf{1}\{\cdot\}$ 为指标函数， $\xi(t)$ 表示由式(13)定义的事件。

4.5 奖赏函数

奖赏函数是每个时隙环境根据智能体的动作返回给智能体的信号，定义了智能体在该状态下执行该动作的好坏，智能体可以根据奖赏函数调整策略。

本文的目标是实现移动设备成本最小化，与 MDP 问题中要求奖赏最大化相矛盾。为了解决此问题，使用成本的相反数作为奖赏函数，将移动设备成本最小化问题转化为最大化问题，具体表示为

$$r_t^1 = -C(t) \quad (15)$$

式(15)为迁移策略的即时奖赏（奖赏 1），奖赏 1 是直接的，但并不有效。原因在于：负奖赏会使移动设备保守地进行学习过程，因此无法探索迁移策略。为了设计更有效的奖赏函数，在奖赏中引入平均成本。平均成本作为简单的指标来确定当前成本是低于历史成本还是高于历史成本，新奖赏（奖赏 2）被定义为

$$r_t^2 = \bar{C}(t) - C(t) \quad (16)$$

其中, 平均价格通过下式得出

$$\bar{C}(t) = (1 - \eta)\bar{C}(t-1) + \eta C(t) \quad (17)$$

其中, η 是平滑参数, $\bar{C}(t)$ 不是一个简单衡量过去所有成本的平均值。在式(16)中使用移动平均数, 不仅可以利用过去的成本信息, 还可以适应当前的成本变化。

从奖赏 2 中可知, 当移动设备以低于平均成本的成本, 即 $C(t) < \bar{C}(t)$ 来制定迁移策略时, 得到的奖赏 $\bar{C}(t) - C(t) > 0$, 那么移动设备以后产生此动作的趋势就会增强; 如果当前成本高于平均成本, 移动设备得到的奖赏 $\bar{C}(t) - C(t) < 0$, 那么移动设备以后产生此动作的趋势就会减弱。通过此方式, 可以鼓励移动设备以更低的成本制定迁移策略。奖赏 2 在探索更多迁移策略方面优于奖赏 1, 同时减轻了成本的非平稳性, 因为相比于以前的成本, 更看重现在的成本。

4.6 Q-learning 算法

在状态、动作、状态转移概率和奖赏函数都定义后, 使用 Q-learning 算法^[22]获得任务迁移策略。在该策略中, 移动设备获得一个状态动作值矩阵 Q , 其中, 每一项 $Q(s, a)_t$ 被定义为每对状态 s 和动作 a 。当移动设备在当前状态下进行任务迁移时, 状态动作值矩阵 Q 通过以下方式更新

$$Q(s, a)_t = \sum_{s(t+1) \in S} P_s(s(t+1)|s(t), a(t)) \left\{ (1 - \alpha)Q(s, a)_{t-1} + \alpha \left[r_t + \gamma \max_{a(t+1)} Q(s, a)_{t+1} \right] \right\} \quad (18)$$

其中, 参数 $\alpha \in [0, 1]$ 是加权过去值和新奖赏的学习率, $\gamma \in [0, 1]$ 是决定未来奖赏重要性的折合因子。在采取行动 $a(t)$ 后, 状态从 $s(t)$ 转移到 $s(t+1)$, 并且移动设备更新状态动作值矩阵 Q 。随着时间的推移, 移动设备可以学习所有状态下采取任意动作的值。当 $Q(s, a)$ 收敛于最优状态动作值时, 可以得到最优任务迁移策略。Q-learning 算法可以得到的迁移策略为

$$a^* = \pi(s) = \arg \max Q(s, a) \quad (19)$$

针对有限 MDP 问题, 这是对于最优迁移策略的保证^[22]。对于任何状态 s , 移动设备总是选

择使得状态动作值矩阵 $Q(s, a)$ 最大化的最佳动作 a^* 。

算法 1 Q-learning 算法

- 1) 在每个时隙 $t \in T$ 内, 设置参数 α 和 γ
- 2) 初始化 Q 矩阵为 0
- 3) 重复
- 4) 对每一次操作
 - 观察状态 $s(t)$
 - 根据 Q 矩阵, 使用 ε -贪婪算法决定最佳动作 $a(t)$
 - 进行这个可能动作, 到达下一个状态
 - 根据奖赏函数 2 求解奖赏 r_t
 - 根据式(18)更新状态动作值矩阵 Q
 - 设置下一个状态作为当前状态
- 5) 当 $t=T$ 时, 结束循环

算法 1 中提出了迁移策略的逐步 Q-learning 算法。为了避免学习算法陷入次优解, 采用 ε -贪婪算法^[23]。该算法不仅能根据式(18)得到最优动作, 还探索了可能更好的其他动作。使用 ε -贪婪算法, 以 $\varepsilon \in [0, 1]$ 的概率随机选择动作, 以 $1 - \varepsilon$ 的概率根据式(18)选择最优动作。设置 $\varepsilon = 0.1$, 即移动设备有 90% 的概率根据状态动作值矩阵 Q 选择当前状态下的最优动作, 有 10% 的概率随机选择动作。

5 仿真结果与分析

5.1 参数设置

在仿真中, 假设每个时隙为 15 min, 任务到达量 $A_i(t)$ 是服从均匀分布的随机变量。电池的最大容量默认为 1.2 kWh, 默认电池初始电量为 0, 移动设备静态功耗设置为 60 W。其他重要参数设置为: 标准化的电池单位损耗成本为 $\omega = 0.1$, 备用电源成本系数为 $\varphi = 1$, 加权过去值和新奖赏的学习率 $\alpha = 0.5$, 决定未来奖赏重要性的折合因子 $\gamma = 0.8$ 。将本文所提基于强化学习的资源管理算法与以下两种算法进行比较。

1) 短视优化算法: 该算法忽略了系统状态与决策之间的时间相关性, 并且利用当前所有可用的电池电量, 使得给定状态下当前时隙的成本最小化。

2) 静态迁移算法: 该算法在每个时隙内使用固定分配方法(在可能的情况下)进行迁移计算, 即在每个时隙内本地处理的计算任务量在总任务量中所占比例固定。

5.2 仿真分析

5.2.1 平均总成本比较

平均总成本比较如图3所示,比较1 000个时隙内强化学习算法与其他两个算法(短视优化算法和静态迁移算法)的平均总成本(包括时延成本和能耗成本)。

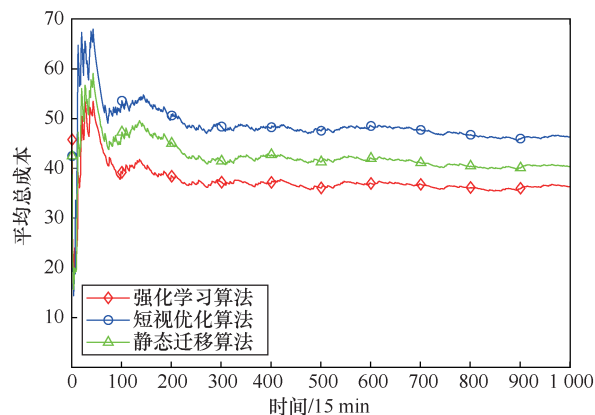


图3 平均总成本比较

由图3可以看出,本文所提的强化学习算法与其他两种算法相比,平均总成本(包括时延成本和能耗成本)显著降低,具体分析如下。

1) 短视优化算法导致累积平均成本较大,因为其忽略了系统状态与决策之间的时间相关性,并且为了使给定状态下当前时隙的成本最小化,利用当前所有可用的电池电量,导致之后时隙的电池电量不足,被迫激活备用电源,产生较大成本。

2) 由于任务分配的比例不同以及对系统参数较敏感,静态迁移算法的性能表现不同,选取迁移率为50%的分配策略进行仿真。由于系统状态先验未知,并且可能随时间改变,所以使用静态迁移算法导致性能损失显著。由于迁移率不同,静态迁移算法与短视优化算法的性能优劣无法确定。当迁移率较高时,由于减少了对备用电源的使用,静态迁移算法比短视优化算法的成本更低。

5.2.2 成本组成比较

备用电源平均成本、平均时延成本以及电池损耗平均成本分别如图4、图5和图6所示。可以看出,强化学习算法通过在电池电量较低的情况下采取保守行动(迁移大部分任务到边缘设备),避免使用备用电源,显著降低了备用电源成本和电池损耗成本。因此,在3种算法中,备用电源平均成本和电池损耗平均成本均最低。虽然强化学习算法导致迁移动作较多,因此产生的平均时延成本最高,

但从图3可以看出,强化学习算法的平均总成本最低。短视优化算法常导致电池电量较低,因此备用电源平均成本明显偏高。

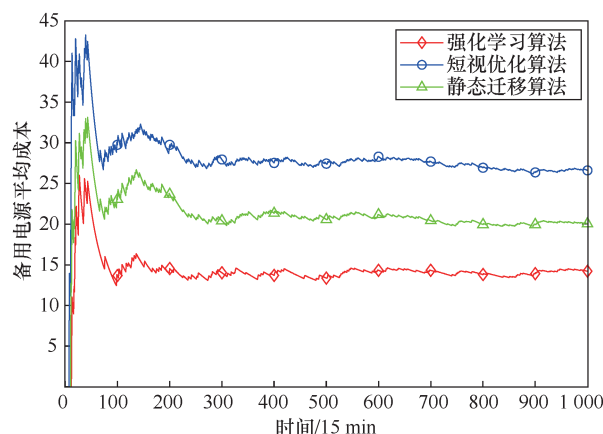


图4 备用电源平均成本

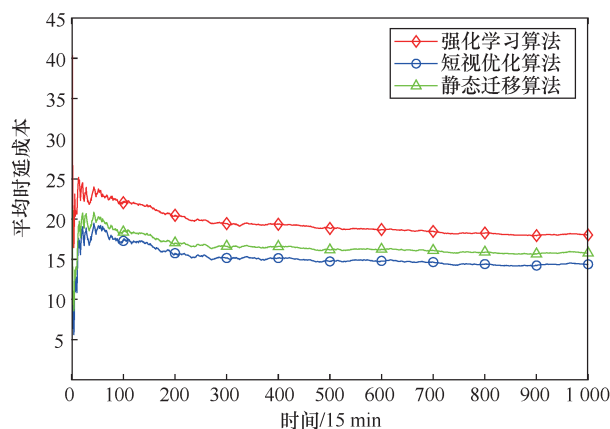


图5 平均时延成本

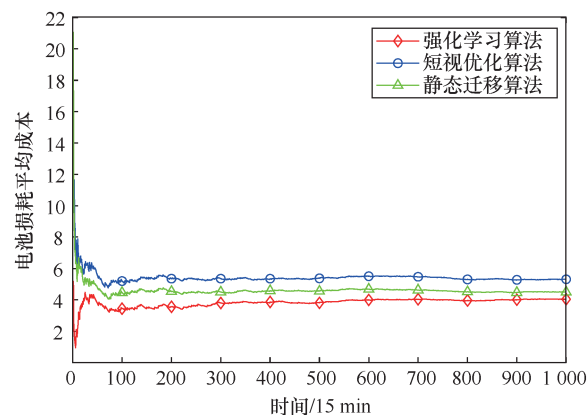


图6 电池损耗平均成本

5.2.3 计算功率需求比较

计算功率需求学习结果如图7所示,图7通过显示学习的最优策略(不同电池电量下移动设备的计算功率需求),进一步解释了强化学习算法优于

短视优化算法和静态迁移算法的原因。如果电池电量较低,强化学习算法的分配策略在使用本地计算功率时较保守,会选择迁移大部分任务到边缘设备,可以节省更多电力,避免在之后到达任务量较多时因电池电量不足而使用备用电源,从而降低系统的长期成本。另一方面,短视优化算法忽略了系统状态与决策之间的时间相关性,即使电池电量不高,为了使当前时隙给定状态下的成本最小化,会分配较多任务在移动设备进行处理。因此,即使短视优化算法在当前时隙成本方面有轻微优势,如果未来到达任务量较多,则会由于电池电量不足,导致使用备用电源,增加未来的成本,使得总成本较高。此外,图 7 表明,计算功耗需求随电池电量的增加而增加。

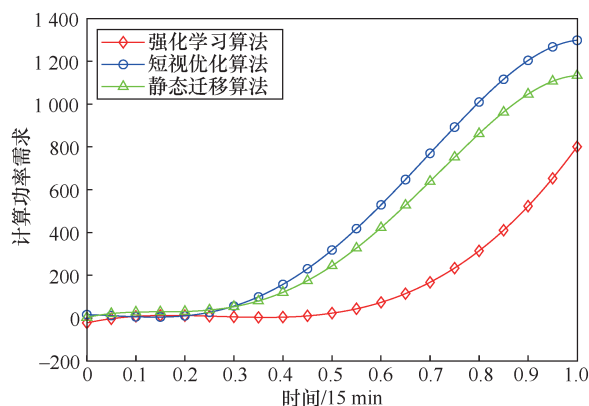


图 7 计算功率需求学习结果

5.2.4 电池状态分布比较

电池状态分布如图 8 所示,对 1 000 个时隙内电池状态的分布情况进行多项式拟合。由图 8 可以看出,在到达任务量相同时,短视优化算法、静态迁移算法相比于强化学习算法的分配策略而言,所执行的动作导致系统处于电池电量较低的状态。而强化学习算法的分配策略则在电池电量不足时,将大部分任务迁移至边缘设备,不仅可以实现更高的能量收集效率,还可以减少对备用电源的使用,降低备用电源成本。

5.3 仿真结果

通过上述 4 个方面(平均总成本、成本组成、计算功率需求和电池状态分布)的仿真,对本文所提的强化学习算法的分配策略性能进行了分析。仿真结果表明,相比于其他两种算法(短视优化算法和静态迁移算法),强化学习算法的分配策略实现了成本最优化。随着时间的推移与状态动作值矩阵

Q 的不断更新,移动设备可以学习所有状态下采取任意动作后得到的 Q 值。在已知当前时隙的状态时,移动设备通过访问状态动作值矩阵 Q 得到当前状态下应采取的最佳动作,由此实现成本最优化。

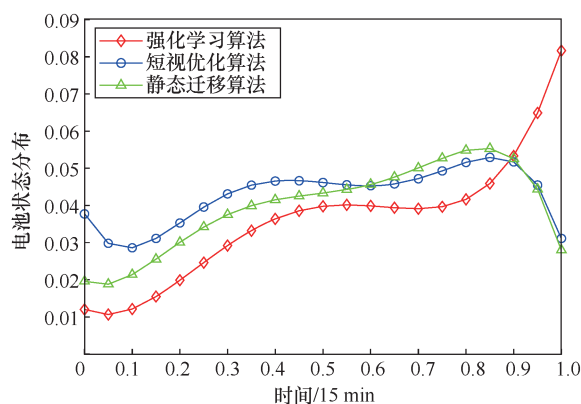


图 8 电池状态分布

相比于本文所提的强化学习算法的分配策略,短视优化算法忽略了系统状态与决策之间的时间相关性。为了使给定状态下当前时隙的成本最小,要利用当前所有可用的电池电量,导致长期成本较高。静态迁移策略则在每个时隙内使用固定分配方法(在可能的情况下)进行迁移计算,忽略了当前的电池电量与可再生能源状态,产生了较高成本。

强化学习算法的分配策略通过将大部分任务迁移至边缘设备进行处理,减少了对电池电量的使用,为未来节省了更多电力,避免在之后到达任务量较多时激活备用电源,从而降低了系统的长期成本。因此,本文所提算法可以达到移动设备成本最优化的目的,具有较好的可行性和有效性。

6 结束语

针对多用户 MEC 系统中移动设备成本的最小化问题,提出了基于强化学习的资源管理算法,制定了成本最优化情况下移动设备的任务分配策略,并将可再生能源纳入多用户移动边缘系统中,实现了移动设备成本最优化(包括时延成本和能耗成本)。

由于本文的问题是顺序型决策问题,因此可以使用 MDP 问题解决。在设计 MDP 问题四要素之一的奖赏函数时,通过引入移动平均成本概念,不仅可以利用过去的成本信息,还可以适应当前的成本变化,同时也减轻了成本的非平稳性。相比于很久之前的成本,更看重现在的成本。

本文将可再生能源纳入多用户移动边缘系统中,由可再生能源供电的无线通信系统利用蓄电池实现能量收集和存储,并提供能量用于用户数据传输和处理。因此在所提模型中添加电池作为能量收集装置,以实现能量收集和存储。在可再生能源短缺的情况下,移动设备可以由备用电源供电,在电网成本极高甚至电网尚未普及的偏远地区,此方法更实用。如果移动设备可以便捷地连接电网,并且通过从电网购买电力方式获得电能,那么本文所提算法同样适用。在此情况下,模型中购买电网电力的支付成本将取代由于激活备用电源而产生的成本。

因此,本文所提算法既可以用于搭建电网的地区,也可以用于电网普及率不高的偏远地区,在实际生活中应用前景较广泛。此外,通过使用可再生能源,在减少环境污染的同时,还可以满足用户需求,保持数据传输与处理稳定性。

本文所提算法仍存在很多不足,在后续研究中可以设计较完善的系统模型,结合实际情况考虑移动用户、不同处理速率、无线信道在不同时刻的数据传输效率以及多个移动设备传输任务的优先级。针对以上实际情况,本文所提算法还有待改善。

参考文献:

- [1] BECK M T, MAIER M. Mobile edge computing: challenges for future virtual network embedding algorithms[J]. The Eighth International Conference on Advanced Engineering Computing and Applications in Sciences, 2014: 65.
- [2] VAQUERO L M, RODEROMERINO L. Finding your way in the fog: towards a comprehensive definition of fog computing[J]. ACM Special Interest Group on Data Communication, 2014, 44(5): 27-32.
- [3] SHI W, CAO J, ZHANG Q, et al. Edge computing: vision and challenges[J]. IEEE Internet of Things Journal, 2016, 3(5): 637-646.
- [4] PATEL M, NAUGHTON B, CHAN C, et al. Mobile edge computing—introductory technical white paper[S]. Mobile-Edge Computer Industry Initiative, 2014.
- [5] CHIANG M, ZHANG T. Fog and IoT: an overview of research opportunities[J]. IEEE Internet of Things Journal, 2017, 3(6): 854-864.
- [6] MAO Y, YOU C, ZHANG J, et al. Mobile edge computing: survey and research outlook[J]. 2017.
- [7] ABBAS N, ZHANG Y, TAHERKORDI A, et al. Mobile edge computing: a survey[J]. IEEE Internet of Things Journal, 2017(99): 1.
- [8] SUDEVALAYAM S, KULKARNI P. Energy harvesting sensor nodes: survey and implications[J]. IEEE Communications Surveys & Tutorials, 2011, 13(3): 443-461.
- [9] RUAN T, CHEW Z J, ZHU M. Energy-aware approaches for energy harvesting powered wireless sensor nodes[J]. IEEE Sensors Journal, 2017, 17(7): 2165-2173.
- [10] MAO Y, ZHANG J, LETAIEF K B. Dynamic computation offloading for mobile-edge computing with energy harvesting devices[J]. IEEE Journal on Selected Areas in Communications, 2016, 34(12): 3590-3605.
- [11] NEELY M. Stochastic network optimization with application to communication and queueing systems[J]. Synthesis Lectures on Communication Networks, 2010, 3(1): 211.
- [12] ZHANG W, WEN Y, GUAN K, et al. Energy-optimal mobile cloud computing under stochastic wireless channel[J]. IEEE Transactions on Wireless Communications, 2013, 12(9): 4569-4581.
- [13] MUÑOZ O, PASCUAL-ISERTE A, VIDAL J. Optimization of radio and computational resources for energy efficiency in latency-constrained application offloading[J]. IEEE Transactions on Vehicular Technology, 2015, 64(10): 4738-4755.
- [14] CHEN X. Decentralized computation offloading game for mobile cloud computing[J]. IEEE Transactions on Parallel and Distributed Systems, 2015, 26(4): 974-983.
- [15] HUANG D, WANG P, NIYATO D, et al. A dynamic offloading algorithm for mobile computing[J]. IEEE Transactions on Wireless Communications, 2012, 11(6): 1991-1995.
- [16] LIU J, MAO Y, ZHANG J, et al. Delay-optimal computation task scheduling for mobile-edge computing systems[J]. IEEE International Symposium on Information Theory, 2016: 1451-1455.
- [17] JIANG Z, MAO S. Energy delay tradeoff in cloud offloading for multi-core mobile devices[J]. IEEE Access, 2015: 2306-2316.
- [18] KWAK J, KIM Y, LEE J, et al. Dream: dynamic resource and task allocation for energy minimization in mobile cloud systems[J]. IEEE Journal on Selected Areas in Communications, 2015, 33(12): 2510-2523.
- [19] WATKINS C, DAYAN P. Technical note Q-learning[J]. Machine Learning, 1992: 279-292.
- [20] SATYANARAYANAN M, BAHL P, DAVIES N. The case for VM-Based cloudlets in mobile computing[J]. IEEE Pervasive Computing, 2009, 8(4): 14-23.
- [21] LI C, HU Y, LIU L, et al. Towards sustainable in-situ server systems in the big data era[C]//ACM/IEEE, International Symposium on Computer Architecture. IEEE, 2015: 14-26.
- [22] SUTTON R S, BARTO A G. Reinforcement learning: an introduction[C]//Neural Information Processing Systems. IEEE, 1999.
- [23] GUENTER B, JAIN N, WILLIAMS C. Managing cost, performance, and reliability tradeoffs for energy-aware server provisioning[J]. Proceedings-IEEE INFOCOM, 2011, 2(3): 1332-1340.
- [24] ZHANG Y, SCHAAR M V D. Structure-aware stochastic storage management in smart grids[C]//IEEE INFOCOM 2014-IEEE Conference on Computer Communications. IEEE, 2014: 2643-2651.

[作者简介]



王璐瑶(1996—),女,山西大同人,东华大学硕士生,主要研究方向为移动边缘计算系统中任务迁移策略和能量管理。

张文倩(1994—),女,山东济宁人,东华大学博士生,主要研究方向为移动边缘计算系统中任务迁移策略和能量管理。

张光林(1981—),男,山东东营人,东华大学教授、博士生导师,主要研究方向为车联网、内容中心网络、智能电网、物联网与移动边缘计算。