



移动边缘计算中资源分配和定价方法综述

邹璐珊^{1,2}, 黄晓雯^{1,3}, 杨敬民^{1,4}, 郑艺峰^{1,2}, 张光林³, 张文杰^{1,2}

- (1. 闽南师范大学计算机学院, 福建 漳州 363000;
2. 闽南师范大学数据科学与智能应用福建省高校重点实验室, 福建 漳州 363000;
3. 东华大学信息科学与技术学院, 上海 200051;
4. 台北科技大学电子工程系, 台湾 台北 106344)

摘要: 移动边缘计算 (mobile edge computing, MEC) 通过将计算任务卸载至边缘服务器, 降低网络负荷, 减少传输时延, 提升用户体验。因此, MEC 受到了广泛关注, 并成为 5G 的关键技术。资源分配作为 MEC 的主要问题, 在提升能量效率、缩短任务时延和节约成本方面具有非常重大的研究意义。首先, 介绍了 MEC 的基本概念、参考架构和技术优势; 然后, 从技术层面和经济层面归纳总结了 MEC 中最新的资源分配和定价策略; 最后, 讨论了 MEC 资源分配和定价策略中可能存在的问题与挑战, 并提出了一些可行的解决方案, 为后续研究发展提供参考。

关键词: 移动边缘计算; 资源分配; 定价方法; 经济模型; 博弈论

中图分类号: TP393

文献标志码: A

doi: 10.11959/j.issn.1000-0801.2022022

Review on resources allocation and pricing methods in mobile edge computing

ZOU Lushan^{1,2}, HUANG Xiaowen^{1,3}, YANG Jingmin^{1,4}, ZHENG Yifeng^{1,2},
ZHANG Guanglin³, ZHANG Wenjie^{1,2}

1. School of Computer Sciences, Minnan Normal University, Zhangzhou 363000, China
2. Key Laboratory of Data Science and Intelligence Application, Fujian Province University, Minnan Normal University, Zhangzhou 363000, China
3. School of Information Science and Technology, Donghua University, Shanghai 200051, China
4. Department of Electronic Engineering, Taipei University of Technology, Taipei 106344, China

Abstract: Mobile edge computing (MEC) has ability of reducing network load and transmission delay, and improving user service experience by offloading computing tasks to edge servers. Therefore, MEC has attracted extensive attention and become a key technology of 5G. As one of the main issues of MEC, resource allocation has great

收稿日期: 2021-09-07; 修回日期: 2022-01-26

通信作者: 张文杰, zhan0300@ntu.edu.sg

基金项目: 国家自然科学基金资助项目 (No.62141602); 福建省自然科学基金资助项目 (No.2021J011002, No.2021J011004, No.2020J01813)

Foundation Items: The National Natural Science Foundation of China (No.62141602), The Natural Science Foundation of Fujian Province of China (No.2021J011002, No.2021J011004, No.2020J01813)



significance in improving energy efficiency, shortening task delay and reducing cost. Firstly, the basic concept, reference architecture and technical advantages of MEC were introduced. Then, the up-to-date achievements of resources allocation and pricing strategies were summarized from both technical and economic aspects. Finally, some possible problems and challenges associated with resource allocation and pricing strategies in MEC were discussed, and the enabling methods to address these challenges were also proposed to provide reference for subsequent research and development.

Key words: mobile edge computing, resource allocation, pricing method, economic model, game theory

0 引言

近年来,伴随着移动互联网(mobile internet, MI)以及物联网(Internet of things, IoT)的不断发展,自动驾驶、智慧交通等海量应用不断涌现,移动互联网用户的数量急剧增加,进而对网络质量、传输速率、请求反馈时延等网络服务性能提出更高的要求。中国移动通信有限公司研究院推进组发布的数据显示,到2030年全球移动通信有限公司数据流量预计将增长约2万倍^[1]。爆炸式的流量增长给网络带来了前所未有的压力,导致网络堵塞而引发的时延增长,极大地影响了用户体验质量(quality of experience, QoE)。

为了应对移动互联网和物联网的飞速发展,5G网络需要同时满足多样化业务在时延、能耗、可靠性等方面的要求。另一方面,由于移动设备体积小,计算能力、电池容量和存储空间有限,无法处理高能耗、高复杂度的计算任务,需要将应用程序部分或全部任务卸载到云计算数据中心的服务器端完成^[2]。需要注意的是,将任务迁移至地理位置较远的云服务器上,不仅会产生大量的数据流,增加网络的负担,而且会显著加大通信的时延,给一些时延敏感型业务(如智能驾驶、增强现实)带来一定的影响^[3-4]。因此,尽管云计算有强大的计算能力,但是距离用户较远,导致其传输能耗增加,传输时延尤甚。为了有效缓解未来网络面临的高负荷、高带宽、高移动性、低时延等方面的挑战,提出了移动边缘计算(mobile edge computing, MEC)的概念。

与云计算不同,MEC允许用户将计算任务卸载到部署在网络边缘的服务器上,使业务本地化,大大减少了远程数据的传输量,降低了网络传输时延,在提升用户服务质量的同时,降低网络的运营成本(网络运营商的运营成本包含计算与传输过程所消耗的电、设备成本、放置成本和后期维护成本,此处降低的成本通常指消耗的电,即能量)^[5-6]。目前,MEC已经成为5G网络的关键技术之一,其能够有效解决未来网络的时延、拥塞和容量等问题,受到学术界和产业界的广泛关注^[7]。虽然将计算任务卸载至边缘服务器上,可以在一定程度上缓解终端用户对网络性能服务的需求,但是边缘服务器的资源是有限的,这使得网络弹性不足,需要引入资源分配和任务调度机制对边缘网络进行统筹管理,从而延长边缘网络的寿命。

为了充分利用边缘网络的资源,中外学者对于MEC的资源分配问题展开了深入的研究,主要涉及计算任务的卸载问题,在提升能量效率、缩短任务时延和节约成本方面具有非常重大的研究意义。MEC中的资源分配问题主要有:通信资源分配,根据无线系统环境,高效地分配通信资源,优化传输效率,降低用户间的干扰;计算资源分配,根据用户的任务需求,有效地分配计算资源,最小化任务时延和能耗,提高卸载效率;存储资源分配,根据用户的存储需求,分配对应的存储资源,避免用户存储空间不足引发的设备卡顿,提升用户QoE。因此,资源分配是解决干扰、降低时延、提高服务性能的核心,如何根据MEC的网络环境、终端的卸载请求以及资源的情况,合

理地进行资源分配是解决干扰问题、减少任务时延的有效途径之一。当前对 MEC 资源分配问题的研究已取得很多相关的成果。具体而言,主要从两个方面进行分析:技术与理论层面和经济层面。前者以最小化任务时延和能耗为目标,或根据不同应用的需求在两个优化目标之间权衡,进行资源的分配;后者则利用经济学的理论机制对资源进行定价和分配,刺激边缘服务器共享资源,以实现利益最大化。本文从技术层面和经济层面对 MEC 的资源分配和定价方法进行详细的归纳概括,讨论和分析了现有方法存在的不足之处,并对现存几个尚待解决的挑战进行展望。

1 MEC 的相关知识

1.1 MEC 的概念和框架

MEC 可利用无线接入网络就近为用户提供 IT 所需要的服务和云计算功能,从而创造出一个具备高性能、低时延、大带宽的服务环境,加速网络中各项内容、服务、应用的快速下载,让用户享有高质量网络服务体验。根据欧洲电信标准学会(ETSI)的定义,MEC 为移动网边缘提供 IT 服务环境和云计算能力,以此减少网络操作和服务时延,从而提高用户体验。根据边缘计算产业联盟的定义,边缘计算是靠近终端的网络边缘侧,也可以说,移动边缘计算是运行在移动边缘网络的云服务器,可以处理传统网络基础架构所不能处理的任务,如机器对机器/人(machine to machine/man, M2M)网关、控制功能、智能视频加速等。MEC 开放平台融合了网络、计算、存储、应用核心能力,可就近提供边缘智能服务,以满足行业数字化在敏捷连接、实时业务、数据优化、应用智能、安全与隐私保护等方面的关键需求。MEC 架构如图 1 所示,边缘计算服务器比云服务器更接近终端用户。因此,尽管边缘计算服务器的计算能力低于云服务器,但仍然为终端用户提供更好的 QoE 和更低的时延^[8]。一般而言,边缘计算的结构

可分为前端、近端和远端 3 个部分^[9-10],以下对这 3 部分详细地描述。

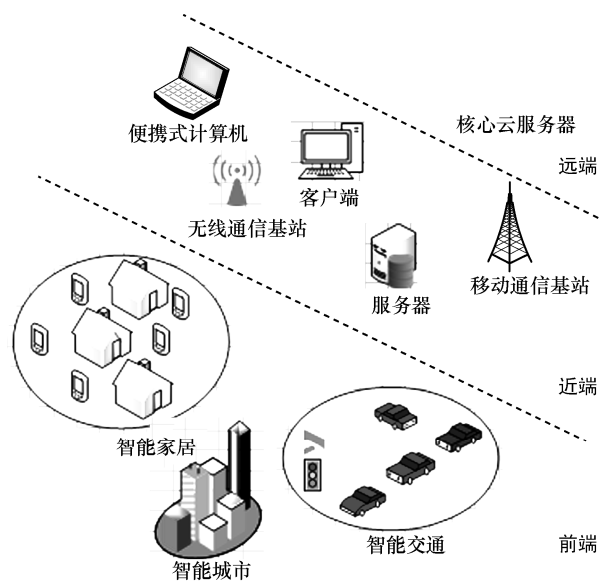


图 1 MEC 架构

前端:终端设备(如传感器、执行器等)被部署在边缘计算结构的前端。前端环境可为用户提供更多的交互能力和更好的响应能力。边缘计算凭借大量邻近终端设备提供的计算能力,可为一些应用提供实时服务。然而,由于终端用户设备的容量有限,大多数要求在前端环境下均无法得到满足。因此,在这些情况下,终端设备必须将资源需求转发给服务器。

近端:部署在近端环境中的网关支持网络中的大多数流量流。边缘服务器可以响应许多资源需求,例如实时数据处理、数据缓存和计算卸载。在边缘计算中,大部分的数据计算和存储都会迁移到这个近端环境中。这样,用户就可以在数据计算和存储方面获得更好的性能。

远端:由于云服务器部署在离终端设备较远的地方,传输时延在网络中就变得十分明显。尽管如此,远端环境中的云服务器仍可提供更多的计算能力和数据存储。例如,云服务器可以提供海量并行数据处理、大数据挖掘、大数据管理、机器学习等。



1.2 MEC 3 层逻辑实体

MEC 服务平台 3 层逻辑实体如图 2 所示, 包括 MEC 基础设施管理系统、MEC 应用平台管理系统和应用管理系统^[11]。

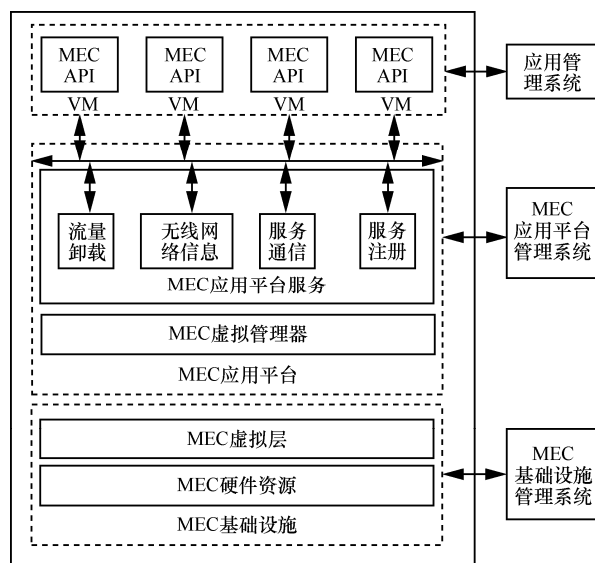


图 2 MEC 服务平台 3 层逻辑实体

MEC 基础设施管理系统包括 MEC 硬件资源和 MEC 虚拟层。其中, 硬件资源主要支持底层的运算、存储及控制功能, 而虚拟层则主要完成运算处理、信息缓存、虚拟交换及象形的管理功能。

MEC 应用平台管理系统由 MEC 虚拟管理器和 MEC 应用平台服务组成。其中, MEC 虚拟管理器为应用创造高效灵活、多用户的服务环境, 即基础设施即服务 (infrastructure as a service, IaaS) 功能。MEC 应用平台服务则承载了对外接口服务功能, 为上层运行于 MEC 服务器的应用提供了若干中间服务, 例如, 涵盖通信服务 (communication service, CS) 和服务注册 (service registry, SR) 的基础设施服务 (infrastructure service, IS)、无线网络信息服务 (radio network information services, RNIS)、流量卸载功能 (traffic offload function, TOF) 等。本层具有相应的底层数据包解析、无线信息交互、内容路由选择、上层应用注册管理等基本功能。

应用管理系统, 即 MEC 应用层, 是基于虚拟机 (virtual machine, VM) 应用构架, 将 MEC 功能组件集成到虚拟应用中, 且通过标准的组装接口开放, 进而完成无线网络的能力开放与调度。虚拟设备的应用程序运行于 IaaS 上的虚拟机映像, 各虚拟机与 MEC 应用平台间保持相同的应用程序接口 (application programming interface, API)。

1.3 MEC 的优点

MEC 的研究, 对于未来网络的发展有着重要的实际意义, 其优点有以下 3 个。

(1) 低时延

边缘计算实时性高、周期性短的数据分析能更好地支持对时延要求高的计算任务的处理与执行。MEC 通过将计算任务“下沉”至更靠近用户的无线接入网侧, 在本地的 MEC 服务器上执行。由于与用户的距离更近, 用户请求可以及时处理, 从而大大降低通信时延^[12]。此外, MEC 的低时延特性在智能交通系统、人脸识别、视频编辑、智慧城市等对时延具有极高需求的新兴应用中优势尤其显著。以智能交通系统为例, 在传统情况下, 高额传感成本极大地限制了智能交通系统, 使其不能全面、高效地收集交通系统中的种种信息, 无法动态和精准地反映交通系统的状态; 而且各系统采集信息不能互通, 不同设备商的系统间或设备间接口不开放, 信息难以协调, 独立系统的判决结果不具备综合性和全局性。在引入 MEC 之后, 可通过 MEC 低时延的特性以及“智能分布式”控制网络平衡技术, 对路口和区域的车流情况、道路交通流的饱和情况、车辆通行速度、总时延和车辆排队长度等, 进行区域管理和交通流绿波控制, 从而实现智能交通系统^[13]。

(2) 缓解流量压力

MEC 在进行云端传输数据时, 先通过边缘服务器进行一部分简单的数据预处理, 减少数据带宽的消耗; 然后, 将提炼出来的精简数据传输到

云端,减轻云端的压力,从而缓解流量压力。以视频编辑为例,在视频传输中,对于某些点击率较高的视频,如天气预报等,通常以直播的方式发布。这种高并发的发布方式使得在相同时间段将会有大量用户接入并且请求同一资源,一般的带宽和链路状态难以满足这种需求。借助部署在网络边缘的 MEC 服务器,能够将这类视频直播实时缓存在更靠近用户的地方,在本地为用户进行请求反馈,进而减小发生链路拥塞和网络故障的可能性。

(3) 信息安全

网络边缘数据往往涉及一些个人隐私信息,如家用摄像机缓存的视频等,传统的云服务模式需要将这些数据信息上传至云端,这使用户隐私承受更大的泄露风险。MEC 模式可以在靠近用户的边缘节点直接对数据源进行管理、归档、分析,以过滤掉隐私数据,实现对一些敏感数据的保护与隔离。以人脸识别为例,不同于传统的视频监控,“MEC+视频监控”模式能够对视频图像信息进行预处理,弥补了云服务响应不及时和能量消耗高等问题,并且在用户隐私方面满足了用户隐私信息安全保护等方面的诉求以及安防行业实时性的需求。边缘计算的数据可使用加密算法后,再打包传回云端,同时,MEC 可以改进部分公网易破解、个人隐私容易被窃取等缺陷,极大地提升整个行业的性能和信息安全。

2 MEC 中的资源分配问题

目前,MEC 正处于高速发展阶段,许多问题仍然需要进一步深入研究。资源分配作为 MEC 的一大关键问题,引起了广大研究者的关注,主要体现在如何有效地分配 MEC 服务器上的空闲资源,以提高网络的整体性能。一方面,利用资源分配可以通过降低计算与传输消耗的能量,进而达到降低运营商成本的目的;另一方面,用户端的资源(计算资源、通信资源、存储资源等)也

是有限的,并且用户对网速(网络的数据传输速率)也有相应的要求。对不同场景进行合理的资源分配,可以在确保用户 QoE(例如,降低时延,提高网络数据传输速率避免视频播放卡顿;降低计算、传输和存储的能量消耗,延长终端设备的待机时间)的前提下最小化运营商的成本。MEC 中的资源分配问题主要有:通信资源分配,根据无线系统环境,高效地分配通信资源,优化传输效率,降低用户间的干扰;计算资源分配,根据用户的任务需求,有效地分配计算资源,最小化任务时延和能耗,提高卸载效率;存储资源分配,根据用户的存储需求,分配对应的存储资源,避免用户存储空间不足引发的设备卡顿,提升用户 QoE。

为了最大限度地利用有限的资源,最小化任务时延/能量消耗,MEC 系统需要根据网络环境、终端的卸载请求以及资源的情况,合理地进行资源分配,MEC 资源分配如图 3 所示。用户终端将计算任务提交至 MEC 系统中的本地调度程序,调度程序检查是否有计算节点能够提供足够的资源。如果存在可处理应用程序的计算节点,则会在该节点上分配资源,最后在该节点上处理应用程序,并将处理结果发回给用户;若不存在该计算节点,则调度应用程序给远端云服务器。

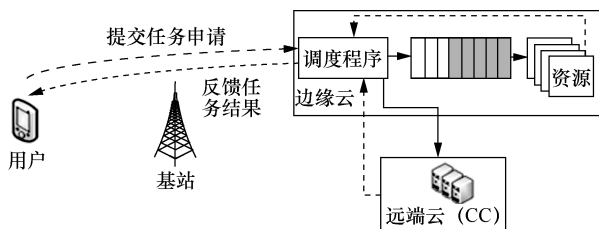


图 3 MEC 资源分配

计算资源与通信资源的综合管理是 MEC 系统设计中必不可少的部分。针对不同 MEC 系统的设计目标,需要采用不同的策略解决资源管理问题。当前对 MEC 资源分配方法的研究大致从两个方面进行:技术层面和经济层面。其中,技术层



面主要涉及将计算任务卸载到哪个边缘节点，致力于平衡通信资源和计算资源，以提升能量效率、缩短任务时延和节约成本；而经济层面不仅需要考虑如何分配资源，更需要考虑市场竞争性。用户不仅需要评估资源出价，更需要考虑向哪个服务器购买资源，以使自身效益最大化。因此，经济层面的资源分配问题计算复杂度大，综合管理更加复杂。对此，研究者一般采用动态分布式算法、拍卖、博弈等经济分析方法求解。

3 MEC 技术层面的资源分配方法

按照 MEC 系统规模，本节将 MEC 应用场景分为单用户系统、多用户系统和异构服务器系统，并对这 3 种场景在资源分配上的研究现状进行归纳总结。

3.1 单用户场景下的资源分配方法

单用户 MEC 系统一般指一个边缘服务器为单用户提供服务的场景，单用户 MEC 系统如图 4 所示。

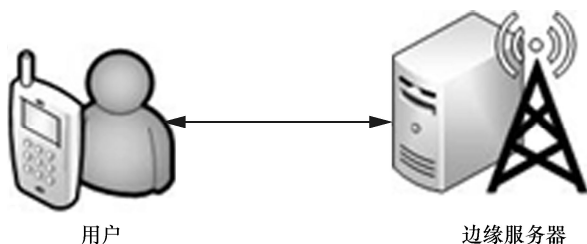


图 4 单用户 MEC 系统

在单用户 MEC 系统中，用户不但占有边缘服务器的全部计算资源，且无须考虑用户之间通信资源的竞争，是一种比较简单的场景。目前，国内外已有一些研究人员对单用户 MEC 系统的资源分配问题进行研究。文献[14]在不考虑网络状态、频繁刷新内容进而造成高能量消耗的背景下，提出了基于马尔可夫决策过程的能耗优化模型。文献[15]考虑在一个具有能量收集装置的绿色 MEC 场景中，提出了基于李雅普诺夫优化的动态计算卸载 (Lyapunov optimization based dynamic

computing offloading, LODCO) 算法。相比基准贪婪策略，该算法能明显降低执行成本（执行时延与任务失败时所消耗的能量），且在一定的时延约束下减少计算失败概率。文献[16]采用马尔可夫决策方法处理单用户 MEC 系统的计算任务调度问题。文献[17]同时考虑计算时延与通信时延，并对这两种时延进行建模，提出了一个网络时延估计决策分析，以找出网络时延估计误差对作业完成时间的影响。为研究 MEC 服务器自身的安全性问题，文献[18]利用空间泊松点过程刻画多个潜在窃听者场景，在给定安全传输条件约束下追求最小化文件传输的时延与能耗，得到了最优码本速率设计和计算任务分配方案。文献[19]研究了绿色 MEC 在移动客户端 QoE 和回传流量的联合优化方面的启用。作者设计了一种具有自调整参数化机制的基于贪婪的算法 (greedy-based algorithm with self-adjusting parameterization mechanism, GASPM) 解决公式化的问题。结果表明，启用绿色 MEC 的动态自适应视频流系统不仅有助于降低电网能耗，还能明显减少回传流量并提高客户端的平均视频比特率。单用户 MEC 系统结构是规模最小、最简单的系统模型，在网络环境中，多个用户并存才是常态。因此，将简单的单用户 MEC 系统扩展到多用户场景中研究资源分配问题是十分有必要的。

3.2 多用户场景下的资源分配方法

多用户的 MEC 系统一般指一个边缘服务器为多个终端用户提供服务的系统模型，多用户 MEC 系统如图 5 所示。

在多用户的 MEC 场景中，系统需要依据有限的资源协调用户之间的竞争，制定资源分配和计算卸载策略，满足大量用户的任务请求。文献[20]以提高边缘服务器的能量使用效率 (power usage efficiency, PUE)、实现绿色计算为目标，提出了基于设备间通信 (device-to-device communication, D2DC) 技术和能量收集 (energy harvesting, EH)

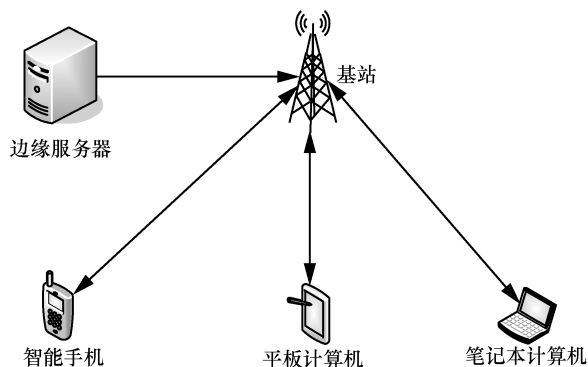


图5 多用户 MEC 系统

技术的绿色任务卸载框架，通过李雅普诺夫技术优化在线任务卸载和资源分配算法。文献[21]提出了云计算与边缘计算结合模型结构（architecture model combines cloud computing and edge computing, CENAM）的新型 3 层网络架构模型，降低边缘层与云计算中心的通信时延、边缘节点间的通信时延和边缘层的计算时延。文献[22]旨在最大化长期效用性能，同时考虑降低计算能耗以及时延，提出了一种运用深度确定性策略梯度（depth deterministic policy gradient, DDPG）和基于候选网络优化边缘计算优化卸载（edge computing optimization offloading, ECOO）算法的新型学习算法，用于解决随机任务卸载问题。文献[23]采用遗传算法和基于分割时间槽的资源分配算法搜索最优决策，以减少云端通信流量。文献[24]提供了一种新颖的 MEC 代理扩展和实施缓存基础设施，以实现高效可靠的内容分发。代理将下载会话切换到替代内容分发网络（content delivery network, CDN）以确保 QoE 速率，从而实现基于实时连接统计的 CDN 动态选择。文献[25]考虑 Wi-Fi 和蜂窝网络覆盖的场景，将线性规划与交替技术相结合，有效地解决任务分流决策的非凸问题，该方案能够极大地提高系统性能。

上述研究成果忽略异构 MEC 系统的资源分配问题，特别是在 5G 场景中，相同区域内存在大量的边缘设备，以及若干数量的边缘服务器。这些边缘服务器的运行方式、功能以及结构往往不

尽相同，如果没有合理的资源分配机制将极大地影响用户的 QoS。因此，研究异构 MEC 系统的资源分配问题已成为研究热点之一。

3.3 异构服务器场景下的资源分配方法

异构服务器 MEC 系统如图 6 所示，包括多个边缘服务器，这为设计高效的资源分配方案带来诸多挑战，包括服务器的选择与协作、计算迁移等。

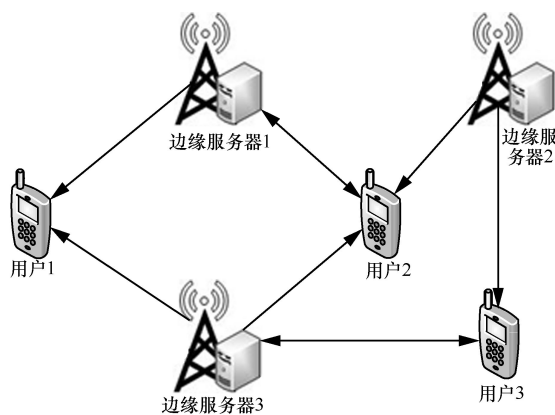


图6 异构服务器 MEC 系统

伴随着车联网行业的迅猛发展，车联网应用场景对任务时延有着很高的要求。文献[26]提出了一种基于车联网多跳传输的移动边缘计算卸载策略，通过预测车辆行驶轨迹，保证在一定时延条件下完成任务卸载。文献[27]运用最佳停止理论（optimal stopping theory, OST）原则解决数据质量感知卸载顺序的决策问题，通过对比众多 OST 随机模型，该方法能够极大地减少车联网分析卸载任务优先级调度的需求。文献[28]提出了改进 min-min 算法 TPMM（task priority-based min-min）。该算法能在一定时延约束的条件下，根据用户任务的优先级、服务器调度次数制定资源匹配方案。文献[29]考虑时间敏感任务的 QoE 要求，研究了无人机辅助的多任务 MEC 网络，其中，无人机为物联网设备提供协作卸载服务。结果表明，协同计算和缓存卸载可以有效降低物联网设备的总能耗，同时满足不同类型任务的 QoE 需求。在文献[30]中，作者开发了两种机器学习模型，并纳入 MEC 应用程序，用于流行视频预测和无线电信



道质量预测。实验证明作者构建的模型在 MEC 体系中能够对提高用户的 QoE 产生效果。除此之外,深度强化学习在优化计算任务卸载策略方面成效显著。文献[31]提出了基于深度强化学习的计算访问点 (computing access point, CAP) 辅助计算的策略,通过将一些计算任务卸载到 CAP 降低系统的时延和能耗。文献[32]则研究了具有软时间窗口的 MEC 系统的卸载策略,并设计了遗传学和生物地理学集成算法处理多个同时执行的任务的卸载问题。

3.4 技术层面资源分配方法总结

MEC 中技术层面资源分配方法归类总结见表 1。在 MEC 中计算卸载分为 3 种情况:本地执行 (计算任务完全在本地进行)、完全卸载 (计算

任务全部由边缘服务器负责)和部分卸载 (计算任务一部分在本地执行一部分上传至边缘服务器执行)。本地卸载的能量消耗主要来源于 CPU 的功耗,而上传至边缘服务器进行计算卸载的能量消耗则由传输能耗与边缘服务器计算能耗组成。相似地,本地卸载的时延通常指 CPU 完成计算卸载所消耗的时间,上传到边缘服务器进行计算卸载的时延则包括传输时延与计算时延。除此之外,以 QoE 为优化目标的研究成果通常以量化的方式表达用户对当前网络的体验和感受,从而反映当前业务情况与终端用户期望之间的差距。上述列举的研究成果主要从网络环境、技术方面研究 MEC 的资源分配问题,旨在最小化能量消耗和任务执行时延,或找到这两个度量之间的权衡,但

表 1 MEC 中技术层面资源分配方法归类总结

MEC 场景	优化目标	参考文献	卸载方式	计算卸载方案
单用户	能耗	文献[14]	完全卸载	基于马尔可夫决策过程的能耗优化模型
		文献[18]	完全卸载	最优码本速率设计和计算任务分配方案
	时延	文献[16]	部分卸载	马尔可夫决策
		文献[17]	部分卸载	不等式分析法
	成本	文献[15]	部分卸载	基于李雅普诺夫优化的动态计算卸载算法 (LODCO)
多用户	QoE	文献[19]	完全卸载	基于自调整参数化技术的贪婪算法
	能耗	文献[20]	部分卸载	基于李雅普诺夫优化技术在线卸载算法设计了任务卸载分配策略
		文献[21]	部分卸载	云计算和边缘计算结合 (CENAM) 模型
	时延	文献[23]	部分卸载	遗传算法和基于分割时间槽的资源分配算法
		文献[22]	部分卸载	运用 DDPG 和 ECOO 算法的新型学习算法
		文献[25]	部分卸载	线性规划与交替优化技术结合的迭代算法
	QoE	文献[24]	部分卸载	多 CDN 视频分发和本地存储的网络代理方案
多服务器	时延	文献[26]	完全卸载	基于车联网多跳传输的移动边缘计算卸载策略
		文献[27]	完全卸载	最佳停止理论方法
		文献[28]	完全卸载	改进 min-min 算法 (TPMM)
	时延和能耗	文献[31]	部分卸载	基于深度强化学习的 CAP 辅助计算
	QoE	文献[29]	部分卸载	基于块坐标下降算法的迭代算法
		文献[30]	完全卸载	机器学习模型
	能耗	文献[32]	完全卸载	遗传学与生物地理学集成的算法

是都忽略了经济层面的分析。除了优化系统的能耗与时延以外,经济模型能够最大限度地利用边缘网络的资源使得系统利润最大化。

4 MEC 在经济层面的资源分配与定价方法

在一个由多个资源提供者、大量多样的应用和用户需求组成的 MEC 系统中,很多的资源所有者可能不会有动力与 MEC 供应商合作,因此,需要设计适当的激励机制进行资源的分配和定价,以刺激资源所有者最大限度地共享其通信和计算能力,实现最高利润。此外,不同的边缘节点有不同的管理模式、业务策略和服务定价。因此,从经济层面上对 MEC 资源分配和定价问题进行研究是很有必要的。虽然在云计算领域已提出很多基于市场的资源分配机制,由于 MEC 服务器资源有限性和分布式的特点,为云计算系统设计的资源分配和定价机制并不能直接适用于 MEC 系统。最近,国内外学者开始为 MEC 系统设计基于市场的资源分配和定价策略,主要基于博弈论/拍卖(auction-based)和基于市场定价等经济模型对资源交易过程进行建模,经济与定价模型如图 7 所示。

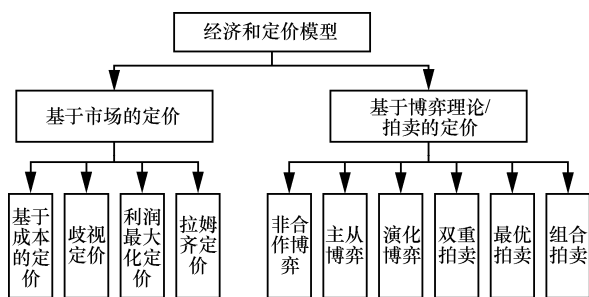


图7 经济与定价模型

4.1 基于市场定价的资源分配方案

定价常用于研究商品的价格制定和变更策略。定价策略框架如图 8 所示,目前用于 MEC 系统的定价策略主要包括基于成本的定价(cost-based pricing)、歧视定价(differential pricing)、利润最大化定价(profit maximization pricing)以及拉姆齐定价(ramsey pricing)。

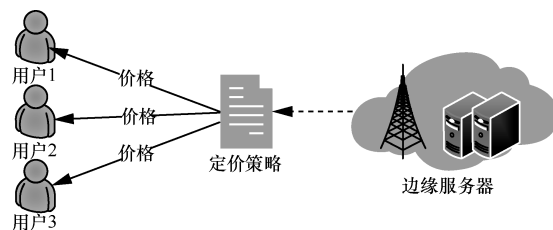


图8 定价策略框架

(1) 基于成本的定价

基于成本的定价是最常见、最简单的定价策略,通过计算服务的总成本,再加上成本的百分比作为“期望利润”确定服务的价格^[33]。采用这种定价策略能够保证服务提供者(即边缘服务器)至少不亏损,并且基于成本的定价具有易于确定价格的优点,适用于比较简单的市场结构。为了在保留用户体验的同时最小化运营商成本,在车联网环境中,文献[34]结合部分卸载决策和边缘服务器的定价对系统进行成本优化。文献[35]对资源进行基于成本的统一定价,并通过仿真实验表明在统一定价下的最大价格就是资源供应商获得最大利润的最优价格。文献[36-38]亦得出最优价格总是接近最大价格结论。因此,这种定价策略容易使资源供应商为获得最大利润而不断提高资源的价格。基于成本的定价策略易于实现,但是没有考虑其他供应商的策略、买家的价格感知能力和支付意愿等市场外部因素,只适用于简单的市场机制。

(2) 歧视定价

歧视定价又称为差别定价。在一个资源市场中,依据终端用户不同的需求和购买意愿,资源供应商可以向他们收取不同的价格^[33]。歧视定价本质上是将用户剩余(即用户愿意支付的总金额与用户实际支付的总金额间的差额)转移到资源供应商,这种定价模型能够使资源供应商利润最大化。文献[35]结合区块链技术构建基于云挖掘机制的算力交易模型,设计一种低梯度迭代算法用于完成边缘服务器与“矿工”之间算力交易的利润最大化。同样在区块链场景中,文献[36]采用两



阶段的 **stackelberg** 博弈实现不同矿工的利润最大化和各自效用最大化。文献[37]提出基于定价策略的分布式方法用以管理用户的计算卸载任务, 建立 **stackelberg** 博弈模型分析边缘服务器与用户间的交互。服务器根据有限的计算能力设置定价以最大化收益, 用户根据卸载决策以最小化成本。文献[38]以资源总需求、竞争环境、对象属性和对象需求等价格导向因素为基石, 制定一致性定价与弹性定价这两种定价方案, 运用 **stackelberg** 博弈模型解决了资源需求以及资源价格只能静态定价的问题, 并通过改进的强化学习状态动作奖励状态动作 (state action reward state action, SARASA) 算法获得了最优解。歧视定价模型虽然能保证资源供应商的利润, 但是让不同类型的终端用户支付不同的价格是不公平的。

(3) 利润最大化定价

利润最大化定价是一个确认资源供应商的资源量与其相应的价格能够产生最高利润的过程^[30]。利润最大化定价如图 9 所示, 假设存在一个资源供应商需要确定其资源量, 资源供应商的利润可表示为:

$$\pi = R(P, Q) - C(Q) \quad (1)$$

其中, Q 表示资源量, P 表示终端用户的出价, R 表示总收益函数, C 表示总成本函数。若以 Q^* 指代最优资源量, 则在利润最大化时 Q^* 可以表示为:

$$Q^* = \arg \max_Q \pi \quad (2)$$

通常资源量与资源价格的关系曲线可表示为:

$$P = a - bQ \quad (3)$$

其中, a 、 b 为适当的参数。基于市场需求曲线, 当 $Q=Q^*$ 时, 可以通过最优资源量找到最优价格:

$$P^* = a - bQ^* \quad (4)$$

文献[39]提出一个基于市场的框架有效地将容量受限边缘节点的资源分配给多个竞争服务。通过引入分布式算法得出了市场需求曲线, 对地理上分布的边缘节点进行合理定价, 提出一个市场均衡和计算均衡的解决方案, 并进一步证明了均衡分配是帕累托最优的, 且满足期望的公平性, 包括共享激励、比例性和无嫉妒性。然而, 在实际市场中, 由于用户需求是随机的, 确定需求曲线是比较困难的。另外, 利润最大化定价模型在确定资源数量和资源价格时, 往往会忽略市场竞争问题。

(4) 拉姆齐定价

拉姆齐定价指针对不同市场对商品的不同需求采用不同的定价策略^[33]。与歧视定价的不同之处在于, 拉姆齐定价是在保证一定的资源供应商利润的前提下最大化社会福利。文献[40]提出一种可以通过任务卸载降低成本的效用函数, 经过实验证实该效用函数可以达到均衡, 且该定价策略能够

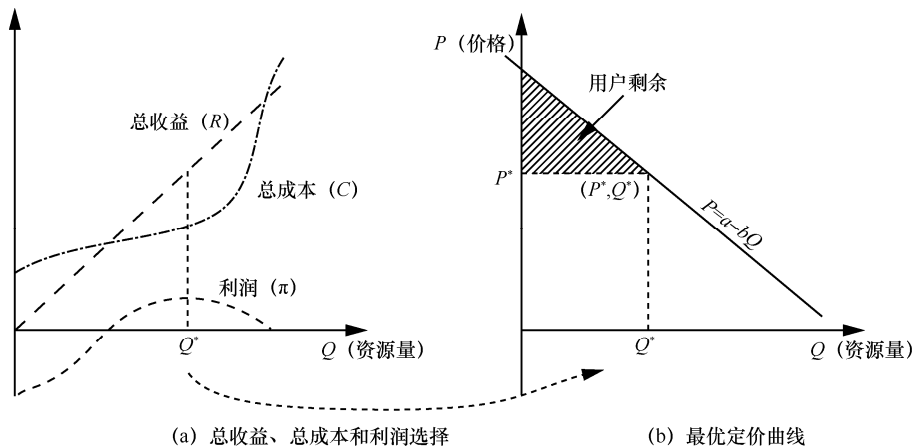


图9 利润最大化定价

在满足个人均衡的同时达到社会均衡。文献[41]提出一种基于学习的定价机制,在达到均衡时,通过接入点(access point, AP)诱导自利用户选择正确的优先级类别以达到社会均衡。拉齐姆定价模型局限性在于,通常考虑的市场是一个无弹性需求的市場。但是,在现实的市場中,用户的需求会根据资源的价格以及其他因素的变化而发生改变,资源供应商无法利用无限抬高价格获取相应的利润。

4.2 基于博弈理论的资源分配方案

博弈论是研究关于激励结构间的相互作用(即包含相互依存情况中的理性行为)的学科,是现代数学的一个分支。在边缘计算环境下,参与者可以是边缘服务器或者终端用户。本节将介绍博弈论模型在 MEC 资源分配中的应用。

(1) 非合作博弈(non-cooperative game)

非合作博弈指在策略环境中所有参与者都是单独进行决策,与环境中的其他参与者无关,非合作博弈框架如图 10 所示。由于每个参与者都具有利己性,所有的参与者都会为了最大化自身利益相互竞争,当模型具有纳什均衡特性时,则表示存在最优策略使得系统利益最大化^[33]。在 MEC 系统中,边缘服务器作为卖家争相出售资源给用户,而用户作为主要参与者会为了最小化时延和能耗而争夺有限的资源^[42]。类似于背包问题,在给定的资源(包括计算资源、通信资源和存储资源等)中,每份资源都有自己的价值,并且资源供应商(即 MEC 服务器)对每份资源也有相应的定价。然而终端用户的“购买力”是有限的,所以终端用户需要在自身有限的“购买力”内购买能够使自身利益最大化的资源量。例如,在一个“资源市场”中,边缘服务器作为资源售卖者向终端用户出售资源,但是卖家通常是自私的,他们之间不会达成协议,但同时都希望自身利益最大化。假设在博弈中所有卖家同时宣布自己的定价策略,这样的“资源市场”就可以建模成非合作博弈。假

设有 N 个参与者,参与者 i 的定价策略集合用 P_i 表示, $P = P_1 \times \cdots \times P_N$ 是策略集合的笛卡儿积。 p_i 表示参与者 i 的定价策略,且 $p_i \in P_i$,则 N 个参与者的定价策略向量可以表示为 $p = (p_1, \cdots, p_N)$,其对应的收益向量为 $\pi = (\pi_1(p), \cdots, \pi_N(p))$,其中, $\pi_i(p)$ 是参与者 i 在已知自己和其他参与者的策略条件下得到的收益。所有参与者都倾向于选择自己的最佳策略 p_i^* ,从而获得最大的利益。当任何一个参与者在其他参与者策略未发生改变的情况下都不能通过调整自身策略取得更高的收益,即:

$$\forall i, p_i \in P_i : \pi_i(p_i^*, \bar{P}_i^*) \geq \pi_i(p_i, \bar{P}_i^*) \quad (5)$$

这种状态称为纳什均衡,相应地, $p^* = (p_1^*, \cdots, p_N^*) \in P$ 被称为纳什均衡解,其中, $\bar{P}_i^* = (p_1^*, \cdots, p_{i-1}^*, \cdots, p_{i+1}^*, p_N^*)$ 表示除去参与者 i 其他参与者策略向量。

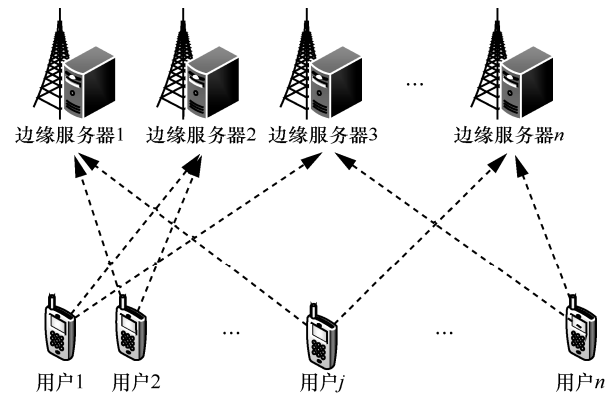


图 10 非合作博弈框架

文献[43]在确保用户计算任务时延需求的条件下,将博弈模型转换为势博弈模型,构造了基于有限改进性质的分布式博弈方法,用于寻找纳什均衡解,从而降低移动设备端的能耗。文献[44]通过对基站群的通信资源、计算资源以及任务队列进行建模,将任务卸载过程的能源成本进行量化。利用潜在博弈模型解决分布式任务卸载的问题,结果表明,所提模型可以收敛到纳什均衡,并且同其他算法相比,该算法可以在时延受限下有效降低系统能耗。文献[45]提出基于博弈的云资



源和计算资源联合分配方案 (joint resource allocation scheme based on evolutionary game, JRA-EG), 通过对这两种资源的分配实现最小化移动终端任务卸载的代价。作者提出基于复制动态的集中式算法和基于 Q-learning 的分布式算法, 实验证实两种算法均能收敛至均衡解。文献[46]提出云边协作任务迁移策略, 借此充分利用 MEC 服务器及云服务器的资源。作者还提出了一种运用博弈论的两阶段任务迁移算法降低系统开销。文献[47]则重点考虑移动终端设备的时延问题, 设计了一个非合作博弈模型并证实了纳什均衡的存在。该算法运用拉格朗日乘子法以及迭代法, 在资源受限的情况下最小化计算任务的时延。

上述文献所考虑的非合作博弈模型建立在参与者同时公布自身定价策略的前提下, 然而这在实际市场中并不总是成立。在主从博弈 (stackelberg game) 中参与者可按照事先确定的顺序公布自己的策略, 因此可以考虑根据主从博弈模型的资源分配方案。

(2) 主从博弈

主从博弈是非合作博弈的一种特殊情况, 在博弈中领导者具有领导优势能够占据先机, 跟随者需要在领导者之后做出博弈^[33]。假设“资源市场”中存在两个资源卖家 S1 和 S2, P_1 和 P_2 分别表示它们的定价策略集合。S1 从集合 P_1 中选择 p_1 以最大化其收益函数 $\pi_2(p_1, p_2)$, 其中, S2 从集合 P_2 中选择 p_2 。不失一般性, 假设 S2 是领导者, S1 是跟随者, 即 S2 优先选择自身策略, 当跟随者用他的最佳策略响应领导者时, 对于领导者来说主从博弈策略是最优解^[48]。设 $D_1 = \{(p_1, p_2) \in P_1 \times P_2 : p_1 = Fp_2\}$ 表示 S2 选择策略 $p_2 \in P_2$ 时 S1 的理性反应集合, S1 被称为理性参与者。当 S1 为领导者时, 同理可得 S2 的理性反应集合。对于领导者来说, 主从博弈策略能够保证其收益至少能够和纳什均衡同样好, 因为在选择主从博弈策略时领导者会对自己施加一个利己的方案。所以,

当 MEC 系统场景符合某一方具有优先权可抢占先机的情况时, 可运用 stackelberg 博弈模型解决资源分配问题。主从博弈过程可以分为两个阶段, stackelberg 博弈模型如图 11 所示。第一个阶段是边缘服务器作为领导者制定资源的价格; 第二个阶段, 根据边缘服务器制定的价格用户作为跟随者提出自己的服务需求。领导者与跟随者在博弈过程中可不断调整策略以实现利益最大化。

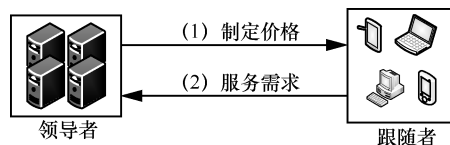


图 11 stackelberg 博弈模型

文献[49]考虑一个边缘服务器和多个终端用户的 MEC 场景, 设计一个以终端用户作为领导者、边缘服务器充当跟随者的 stackelberg 博弈模型。在车联网场景中, 文献[50]将优化目标放在时延上, 提出基于 stackelberg 博弈模型的资源分配框架, 并利用分布式迭代算法对边缘节点的定价问题和边缘服务器的资源策略进行分析。文献[51]则针对多边缘服务器与多用户的异构 MEC 系统, 构建了“多主多从”的 stackelberg 博弈模型, 证明了 stackelberg 平衡的存在, 并通过设计最佳需求设计算法 (optimal demand computation algorithm, ODCA) 寻找终端用户的最佳资源诉求策略和均衡价格, 但是没有考虑动态场景。

单纯运用 stackelberg 博弈模型往往忽略参与者并非完全理性这一事实, 演化博弈模型则将这一问题考虑在内。

(3) 演化博弈 (evolutionary game)

演化博弈不同于传统的博弈理论, 认为参与者并非完全理性的, 在博弈中通过不断试错, 最后达到均衡。演化博弈多用于多用户计算卸载场景, 在动态环境中, 每个设备都无法得知其他参与者的策略以及定价, 参与者通过不断地调整改善自身的策略最终实现群体利益最大化, 演化博

弈模型如图 12 所示。与遗传算法的理念类似,在遗传算法中,通过层层迭代一组候选解会向更好的解进化。当迭代代数到了某个标准或者种群达到符合的适应度水平(即上文提到的纳什均衡)时,停止迭代。也有一些学者会通过贪婪算法求解这类问题(即通过在每一个阶段都求解局部最优解的启发式方法实现寻找全局最优解),但是贪心策略通常仅能产生近似全局最优解的局部最优解,而不会产生全局最优解。例如,在“资源市场”中存在 N 个“资源”卖家和 n 个买家,初始时,每个买家随机选择一个卖家购买所需要的资源。然后,每个买家计算其收益 μ 并将此信息共享给其他的买家,收到所有效用信息后,每个买家能够计算其的平均收益 $\bar{\mu}$,若自身收益低于平均收益则改变策略,即选择其他的卖家购买资源。重复上一步直到没有买家需要改变其策略为止,此时算法达到平衡。在参与者具有各自的计算卸载任务,并且不了解其他参与者的卸载策略时,运用演化博弈理论进行资源分配是十分合适的。

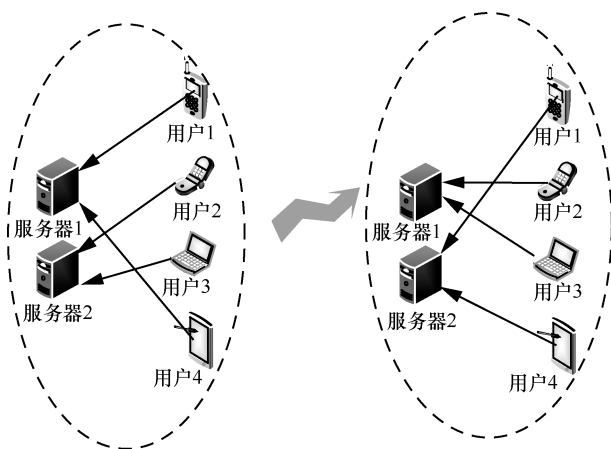


图 12 演化博弈模型

文献[52]考虑在多用户多服务器场景下计算和通信资源的分配问题,利用演化博弈对移动用户的动态行为进行建模,建立确定性和随机性模型,并通过设计进化算法得到进化均衡解。文献[53]将多用户卸载问题看作一个演化博弈模型,经过分析策略选择过程,达到演化均衡。之后结合实

际应用场景,设计基于强化学习的进化博弈卸载策略(evolutionary game theory based on Q-learning, EGT-QL),在能耗方面达到稳定收敛。文献[54]考虑多个边缘云与中央云协作的异构 MEC 系统,根据去中心化的思想,提出了基于演化博弈的分散迭代算法。通过模拟终端用户的动态行为,在确保终端用户 QoE 的前提下,实现运营成本最小化。同样在 3 层异构 MEC 系统下,文献[55]综合考虑 MEC 系统的能耗、时延以及经济成本,提出人口进化的集中算法,该算法能够快速实现均衡。为研究边缘服务器的地理位置对 MEC 系统性能的影响,文献[56]提出了基于概率演化博弈框架的动态卸载策略。文献[57]提出一种创新的均值场进化方法,该方法考虑了信道干扰、平均时延、边缘服务器间的平衡和公平性,但是没有考虑资源十分有限的情况。

4.3 基于拍卖理论的资源分配方案

拍卖经济理论最初归属于经济理科学与运筹学,是最久远的价格发现机制之一。拍卖理论也是研究 MEC 资源分配和定价的有效方法。

(1) 双重拍卖(double auction)

双重拍卖通常是买家和卖家双方同时向拍卖人提出竞价和要价,最后由拍卖人确定成交价格的一种拍卖程序^[33]。例如在一个“资源市场”中,边缘服务器为“资源”卖家,终端设备为买家,拍卖人将买家出价按照降序排列,卖家要价按照升序排列。拍卖人从中寻找出价高于定价的最大指数 k ,即 $p_k^a \leq p_k^b$ 。最终成交价可表示为 $p^* = (p_k^a + p_k^b) / 2$ 。此时买家得到资源,卖家得到付款。不断重复上述过程,直至所有的买家与卖家都能获得相应的资源与付款为止。实际上双重拍卖的过程类似于由边缘服务器的要价与用户的出价组成的离散化的供需模型。假设用户的要价向量为 $\mathbf{a} = \{a_1, a_2, \dots, a_m\}$,服务器的要价向量为 $\mathbf{b} = \{b_1, b_2, \dots, b_n\}$,不失一般性有:



$$\begin{aligned} a_1 &\leq a_2 \leq \dots \leq a_m \\ b_1 &\geq b_2 \geq \dots \geq b_n \end{aligned} \quad (6)$$

以横轴表示边缘服务器所提供的资源数量，纵轴表示用户给出的定价，则双重拍卖资源要价的供需曲线如图 13 所示，当需求和供给曲线相交于一点时，达到了市场均衡，这一点被称作市场均衡点。双重拍卖具有拍卖参与者无损失、拍卖人得到收益、拍卖双方如实提供定价与要价以及群体收益最优的特点。

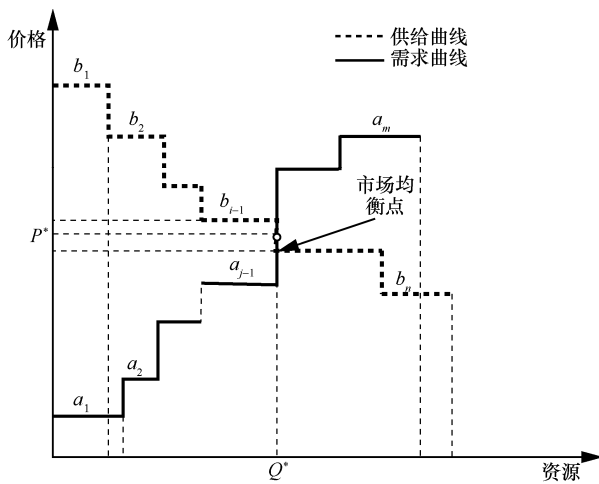


图 13 双重拍卖资源要价的供需曲线

在车联网环境中，文献[58]提出一种基于交通流量预判测的车联网边缘计算迁移方案。该方案采用 McAfee 拍卖算法解决车辆与边缘服务器之间的资源双重拍卖问题，旨在提高车辆的计算迁移率同时，改善边缘服务器的资源利用率。文献[59]针对工业物联网移动设备和边缘服务器的适配问题提出两种双重拍卖方案，分别为盈亏平衡双重拍卖 (break-even double auction, BDA) 和动态定价双重拍卖 (dynamic pricing double auction, DPDA)。这两种方案都能够满足 MEC 的经济特性，有效提高 MEC 系统资源分配效率。文献[60]提出了一种基于时延保证的资源双重拍卖算法 (latency guaranteed based resource double auction, LGRDA)，在保证计算任务时延约束的前提下，该算法能高效地将边缘服务器中的资源分配给终

端用户，从而最大化边缘服务器利润。文献[61]为解决长距离传输对时延敏感任务（如自动驾驶、工业控制等）的资源需求问题，提出双边拍卖方案，并运用经验加权吸引力 (experience weighted attraction, EWA) 算法加快不透明环境中每个参与者的学习过程。

双重拍卖要求拍卖双方如实提供定价及要价，并且在整个交易过程中，价格信息是公开的。

(2) 最优拍卖 (optimal auction)

最优拍卖参与的买家是独立且理性的，卖家无法干预买家参与竞价与否以及如何出价。在拍卖机制中的收益期望可表示为：

$$U_0(q, p) = E_v[v_0(1 - \sum_i q_i(V)) + \sum_i p_i(V)] \quad (7)$$

其中， q 表示分配函数， p 表示计价函数， v_0 表示商品的正常价格， V 表示买家对商品的估价。最优拍卖机制要解决的问题就是设计具体的 q 和 p ，实现最大化收益期望。

文献[62]兼顾终端用户的任务处理时延以及系统能耗，运用深度神经网络控制终端设备参与拍卖活动。深度神经网络能够保证拍卖结果的公平性，与随机拍卖相比该拍卖机制能够将利润提高至少 50%。文献[63]提出基于深度学习的最佳拍卖方案以解决移动区块链网络中边缘服务器资源的分配问题。考虑任务获取和传输的成本，文献[64]为最大化边缘服务器的预期收入，提出基于最优拍卖的任务缓存机制。经过仿真实验，该机制性能很好，但是受各种模型的参数影响大。

最优拍卖需要满足两个假设：每个买家都要有参与拍卖的意愿；买家会依照真实的价值估价出价。

(3) 组合拍卖 (combinatorial auction)

组合拍卖也是典型的背包问题，它不同于传统拍卖机制的地方在于，其拍卖目标为多种商品的组合^[65]，组合拍卖框架如图 14 所示。MEC 中的资源分配常常设计多种资源，包括计算资源、通信资源

和存储资源等,当要进行多种资源的分配和定价时,组合拍卖的效率比传统拍卖方式更高。

文献[66]提出由层次分析法(analytic hierarchy process, AHP)排序、任务投标、获胜者决策3个部分组合而成的多轮顺序组合拍卖机制。该机制适用于时延与容量受限,最大化服务节点效益的车载网络环境。

在组合拍卖中,胜利者(即拍卖的最后中标者)如何确定是组合拍卖的一大挑战,该问题为非确定性多项式(non-deterministic polynomial, NP)难题,在当前的研究中依旧很热门。

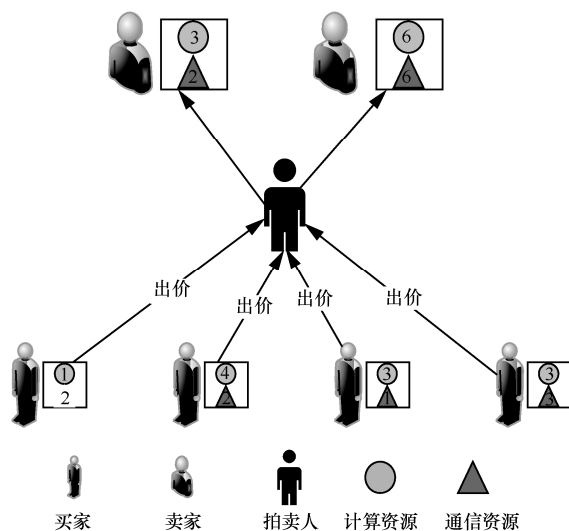


图 14 组合拍卖框架

5 问题与挑战

基于经济分析的资源分配和定价方法比较及使用场景与局限性见表 2, 尽管基于经济分析的资源分配和定价方法已经开始被研究, 国内外研究人员针对该领域方向已取得一些研究成果, 但以下几个关键问题到目前为止仍未得到有效解决。

(1) 用户移动性和任务随机性

用户高度移动是 MEC 的一大特点, 为了确保正在执行的应用程序不会中断, 任务必须能够迁移到另一个节点。当前大部分经济模型比较适用

于静态环境, 因此当用户移动时, 依赖于中心实体的分配和定价策略就不是很适用。此外, 用户不是时刻都有任务需要处理, 它们的计算任务是随机到达的。因此系统中活跃用户的数目会随着新任务到达和旧任务完成发生变化。用户移动性和任务随机性极大提升了系统资源分配和定价的难度。利用马尔可夫近似技术, 可通过设计可逆且收敛的马尔可夫链将用户的移动性问题映射为马尔可夫链的状态转移问题。考虑到任务随机性, 由于 MEC 系统的网络拓扑是随时间发生变化的, 该马尔可夫链的实现算法应该支持分布式实现。

(2) 需求不确定性和信息不对称性

由于用户的移动性和任务的随机性, 接入点无法事先知道用户的服务需求, 导致需求不确定性。此外, 在基于基础设施的网络环境中, 代理和无线接入点与用户之间的距离不同, 掌握的信息也是不一样的。接入点可能不愿意将私有信息与代理节点共享, 导致信息不对称的情况, 这给经济模型的资源预留和定价带来了很大的挑战。一般情况下, 接入点的收益会随着资源预留量的增大而增大, 因此, 为增加自身利益, 接入点会有夸大用户资源需求的倾向, 造成代理资源预留过量。借鉴合同理论, 为了激励接入点共享可信的资源需求信息, 代理可制定合同条款, 额外收取资源保留费用, 让接入点分担一些超额预留的风险成本。

(3) 用户多样性

MEC 系统中有很多种用户, 根据不同的服务需求, 有些选择注册成为合同用户, 有些是随机用户。例如, 在健康监测中, 由于终端设备位置相对固定, 注册成为合同用户是比较好的选择; 而对一些移动性比较高的设备, 由于服务需求是短暂的, 适合成为随机用户。不同的用户需要缴纳不同的费用, 享受不一样的服务, 为此需要根据用户类型制定不同的资源分配和定价策略。



表 2 基于经济分析的资源分配和定价方法比较及使用场景与局限性

经济和定价模型	定价策略	参考文献	具体方案	适用场景和局限性
基于市场的定价	基于成本的定价	文献[35]	低梯度迭代算法	适用于比较简单的市场结构，缺乏考虑供应商策略、卖家价格感知和支付意愿等市场外部因素
		文献[36]	低复杂度梯度算法	
		文献[37]	动态递推	
		文献[38]	SARASA 算法	
	歧视定价	文献[35]	低梯度迭代算法	适用于大部分市场结构，但是缺乏公平性
		文献[36]	低复杂度梯度算法	
		文献[37]	动态递推	
		文献[38]	SARASA 算法	
	利润最大化	文献[39]	引入几种分布式算法得出市场需求曲线	适用于大部分市场结构，但忽略市场竞争问题
	拉姆齐定价	文献[40]	任务卸载效用函数	适用于大部分市场结构，但未考虑弹性需求市场
		文献[41]	AP 诱导自利用用户选择正确的优先级	
基于博弈理论的定价	非合作博弈	文献[43]	基于有限改进性质的分布式博弈方法	适用于每个参与者都是自主决策与其他参与者无关的场景，但是参与者同时公布自身定价策略在实际中很难成立
		文献[44]	利用潜在博弈模型解决分布式任务卸载的问题	
		文献[45]	基于博弈的云资源和计算资源联合分配方案	
		文献[46]	两阶段任务迁移算法	
		文献[47]	拉格朗日乘子法以及迭代法	
	主从博弈	文献[49]	迭代算法	适用于某一方具有优先权可抢占先机的场景，但是要求参与者完全理性
		文献[50]	分布式迭代算法	
		文献[51]	ODCA 算法	
	演化博弈	文献[52]	迭代算法	适用于参与者具有各自的计算卸载任务并且不了解其他参与者的卸载策略
		文献[53]	基于强化学习的进化博弈卸载策略 (EGT-QL)	
		文献[54]	基于演化博弈论的分散迭代算法	
		文献[55]	人口进化集中算法	
		文献[56]	基于概率演化博弈论框架的动态卸载策略	
		文献[57]	均值场进化方法	
基于拍卖理论的定价	双重拍卖	文献[58]	McAfee 拍卖算法	适用于拍卖双方如实提供定价及要价的场景
		文献[59]	盈亏平衡双重拍卖 (BDA) 和动态定价双重拍卖 (DPDA)	
		文献[60]	基于时延保证的资源双重拍卖算法 (LGRDA)	
		文献[61]	经验加权吸引力 (EWA) 算法	
	最优拍卖	文献[62]	运用深度神经网络控制终端设备参与拍卖活动	适用于买家都有拍卖意愿并且买家会依据真实的价值估价进行出价
		文献[63]	基于深度学习的最佳拍卖方案	
		文献[64]	基于最优拍卖的任务缓存机制	
	组合拍卖	文献[66]	多轮顺序组合拍卖机制	适用于分配多种商品的场景，但是拍卖最终中标者的确定是一大难点

(4) 5G 异构网

在 5G 异构网络中, 小型基站大规模高度密集, 资源常常被共享分配给不同的网络层, 干扰也来自不同层, 使得在不同网络层之间进行经济分析变得十分困难, 因此, 5G 异构网络需要比传统网络更加灵活高效的资源分配方式。此外, 将 MEC 和网络切片融合, 满足更多用户个性化需求是未来 5G 中广泛的服务模式^[67]。随着 MEC 的引入, 资源分配和定价面临新的挑战, 此时资源不仅有无线资源, 还有计算资源和存储资源。在保持所需要的 QoE 水平的同时优化系统资源利用率是一项重大挑战, VM 的引入将资源捆绑销售, 制定组合资源价格策略是实现这一目标的关键。一般情况下, 采用 VM 机制解决组合资源分配问题可以分解为多个决策过程: 给用户分配哪些资源, 每种资源分配多少, 以及这些资源的最佳价格。在网络切片运行时, 为了能够满足用户的动态需求, 提供定制化服务, 应设计一种自适应 VM 算法, 按需调整资源的分配和定价。

6 结束语

本文主要从 MEC 的基本概念、参考架构、优点、资源分配与定价方法以及存在问题和挑战等方面进行了系统性介绍, 首先, 介绍了 MEC 的基本概念和 MEC 的参考框架, 对 MEC 的优点以及资源分配问题进行归纳并阐述; 然后, 着重从技术和经济两个层面对现有的资源分配策略和定价方法进行归纳总结, 并讨论了不同方案的缺点; 最后, 对 MEC 资源分配中现存的几个尚待解决的问题做了进一步分析, 并对相应的解决方向做了简要的展望。

参考文献:

- [1] IMT-2020(5G)推进组. 5G 愿景与需求白皮书 V1.0[R]. 2014. IMT-2020. 5G Vision and demand white paper V1.0[R]. 2014.
- [2] LIN J, YU W, ZHANG N, et al. A survey on Internet of things: architecture, enabling technologies, security and privacy, and applications[J]. IEEE Internet of Things Journal, 2017, 4(5): 1125-1142.
- [3] MACH P, BECVAR Z. Mobile edge computing: a survey on architecture and computation offloading[J]. IEEE Communications Surveys & Tutorials, 2017, 19(3): 1628-1656.
- [4] 张建敏, 谢伟良, 杨峰义, 等. 移动边缘计算技术及其本地分流方案[J]. 电信科学, 2016, 32(7): 132-139. ZHANG J M, XIE W L, YANG F Y, et al. Mobile edge computing and application in traffic offloading[J]. Telecommunications Science, 2016, 32(7): 132-139.
- [5] 李子姝, 谢人超, 孙礼, 等. 移动边缘计算综述[J]. 电信科学, 2018, 34(1): 87-101. LI Z S, XIE R C, SUN L, et al. A survey of mobile edge computing[J]. Telecommunications Science, 2018, 34(1): 87-101.
- [6] 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.
- [7] 施巍松, 张星洲, 王一帆, 等. 边缘计算: 现状与展望[J]. 计算机研究与发展, 2019, 56(1): 69-89. SHI W S, ZHANG X Z, WANG Y F, et al. Edge computing: state-of-the-art and future directions[J]. Journal of Computer Research and Development, 2019, 56(1): 69-89.
- [8] ISLAM A, DEBNATH A, GHOSE M, et al. A survey on task offloading in multi-access edge computing[J]. Journal of Systems Architecture, 2021, 118: 102225.
- [9] KE H C, WANG J, DENG L Y, et al. Deep reinforcement learning-based adaptive computation offloading for MEC in heterogeneous vehicular networks[J]. IEEE Transactions on Vehicular Technology, 2020, 69(7): 7916-7929.
- [10] 夏士超, 姚枝秀, 鲜永菊, 等. 移动边缘计算中分布式异构任务卸载算法[J]. 电子与信息学报, 2020, 42(12): 2891-2898. XIA S C, YAO Z X, XIAN Y J, et al. A distributed heterogeneous task offloading methodology for mobile edge computing[J]. Journal of Electronics & Information Technology, 2020, 42(12): 2891-2898.
- [11] PATEL M, NAUGHTON B, et al. Mobile-edge computing-introductory technical white paper[R]. European Telecommunications Standards Institute, 2015.
- [12] 李邱苹, 赵军辉, 贡毅. 移动边缘计算中的计算卸载和资源管理方案[J]. 电信科学, 2019, 35(3): 36-46. LI Q P, ZHAO J H, GONG Y. Computation offloading and resource management scheme in mobile edge computing[J]. Telecommunications Science, 2019, 35(3): 36-46.
- [13] LIU L, CHEN C, PEI Q Q, et al. Vehicular edge computing and networking: a survey[J]. Mobile Networks and Applications, 2021, 26(3): 1145-1168.
- [14] 郭延超, 高岭, 王海, 等. 移动边缘计算中基于内容动态刷新的能耗优化[J]. 计算机研究与发展, 2018, 55(3): 563-571. GUO Y C, GAO L, WANG H, et al. Power optimization based on dynamic content refresh in mobile edge computing[J]. Journal of Computer Research and Development, 2018, 55(3): 563-571.



- 563-571.
- [15] MAO Y 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.
 - [16] LIU J, MAO Y Y, ZHANG J, et al. Delay-optimal computation task scheduling for mobile-edge computing systems[C]//2016 IEEE International Symposium on Information Theory (ISIT), 2016: 1451-1455.
 - [17] MELENDEZ S, MCGARRY M P. Computation offloading decisions for reducing completion time[C]//Proceedings of 2017 14th IEEE Annual Consumer Communications & Networking Conference. Piscataway: IEEE Press, 2017: 160-164.
 - [18] 任品毅, 许茜. 基于移动边缘计算的时延能耗最小化安全传输[J]. *通信学报*, 2020, 41(11): 52-63.
REN P Y, XU Q. Delay and energy minimization for MEC-based secure communication[J]. *Journal on Communications*, 2020, 41(11): 52-63.
 - [19] YANG S R, TSENG Y J, HUANG C C, et al. Multi-access edge computing enhanced video streaming: proof-of-concept implementation and prediction/QoE models[J]. *IEEE Transactions on Vehicular Technology*, 2019, 68(2): 1888-1902.
 - [20] 马惠荣, 陈旭, 周知, 等. 绿色能源驱动的移动边缘计算动态任务卸载[J]. *计算机研究与发展*, 2020, 57(9): 1823-1838.
MA H R, CHEN X, ZHOU Z, et al. Dynamic task offloading for mobile edge computing with green energy[J]. *Journal of Computer Research and Development*, 2020, 57(9): 1823-1838.
 - [21] LI G S, WANG J P, WU J H, et al. Data processing delay optimization in mobile edge computing[J]. *Wireless Communications and Mobile Computing*, 2018, 2018: 6897523.
 - [22] ZHANG W X, DU Y W. Deep reinforcement learning-based optimization of lightweight task offloading for multi-user mobile computing[J]. *Journal of Measurement Science and Instrumentation*, 2020: 1-14.
 - [23] 李振江, 张幸林. 减少核心网拥塞的边缘计算资源分配和卸载决策[J]. *计算机科学*, 2021, 48(3): 281-288.
LI Z J, ZHANG X L. Resource allocation and offloading decision of edge computing for reducing core network congestion[J]. *Computer Science*, 2021, 48(3): 281-288.
 - [24] VIOLA R, MARTIN A, ZORRILLA M, et al. MEC proxy for efficient cache and reliable multi-CDN video distribution[C]//Proceedings of 2018 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting. Piscataway: IEEE Press, 2018: 1-7.
 - [25] FAN W H, HAN J T, YAO L, et al. Latency-energy optimization for joint Wi-Fi and cellular offloading in mobile edge computing networks[J]. *Computer Networks*, 2020, 181: 107570.
 - [26] 邸剑, 薛林, 蔡震. 基于车联网多跳传输的移动边缘计算卸载[J]. *计算机应用研究*, 2021, 38(4): 1145-1148, 1157.
DI J, XUE L, CAI Z. Mobile edge computing offloading based on multi-hop transmission of connected vehicles[J]. *Application Research of Computers*, 2021, 38(4): 1145-1148, 1157.
 - [27] ALGHAMDI I, ANAGNOSTOPOULOS C, PEZAROS D P. Data quality-aware task offloading in mobile edge computing: an optimal stopping theory approach[J]. *Future Generation Computer Systems*, 2021(117): 462-479.
 - [28] 董思岐, 李海龙, 屈毓铨, 等. 面向优先级用户的移动边缘计算任务调度策略[J]. *计算机应用研究*, 2020, 37(9): 2701-2705.
DONG S Q, LI H L, QU Y B, et al. Task scheduling policy for mobile edge computing with user priority[J]. *Application Research of Computers*, 2020, 37(9): 2701-2705.
 - [29] TAO L W, ZHAO M X, HUI W Y, et al. Collaborative offloading for UAV-enabled time-sensitive MEC networks[J]. *EURASIP Journal on Wireless Communications and Networking*, 2021, 2021(1): 1-17.
 - [30] MEHRABI A, SIEKKINEN M, YLÄ-JÄÄSKI A. Energy-aware QoE and backhaul traffic optimization in green edge adaptive mobile video streaming[J]. *IEEE Transactions on Green Communications and Networking*, 2019, 3(3): 828-839.
 - [31] ZHAO R, WANG X J, XIA J J, et al. Deep reinforcement learning based mobile edge computing for intelligent Internet of Things[J]. *Physical Communication*, 2020(43): 101184.
 - [32] WEI Z C, PAN J, LYU Z W, et al. An offloading strategy with soft time windows in mobile edge computing[J]. *Computer Communications*, 2020(164): 42-49.
 - [33] LUONG N C, WANG P, NIYATO D, et al. Resource management in cloud networking using economic analysis and pricing models: a survey[J]. *IEEE Communications Surveys & Tutorials*, 2017, 19(2): 954-1001.
 - [34] LI L H, LV T J, HUANG P M, et al. Cost optimization of partial computation offloading and pricing in vehicular networks[J]. *Journal of Signal Processing Systems*, 2020, 92(12): 1421-1435.
 - [35] 吴雨芯, 蔡婷, 张大斌. 移动边缘计算中基于 Stackelberg 博弈的算力交易与定价[J]. *计算机应用*, 2020, 40(9): 2683-2690.
WU Y X, CAI T, ZHANG D B. Computing power trading and pricing in mobile edge computing based on Stackelberg game[J]. *Journal of Computer Applications*, 2020, 40(9): 2683-2690.
 - [36] XIONG Z H, FENG S H, NIYATO D, et al. Optimal pricing-based edge computing resource management in mobile blockchain[C]//Proceedings of 2018 IEEE International Conference on Communications. Piscataway: IEEE Press, 2018: 1-6.
 - [37] LIU M Y, LIU Y. Price-based distributed offloading for mobile-edge computing with computation capacity constraints[J]. *IEEE Wireless Communications Letters*, 2018, 7(3): 420-423.
 - [38] 刘荆欣, 王妍, 韩笑, 等. 基于 Stackelberg 博弈的边缘云资源定价机制研究[J]. *计算机科学与探索*, 2020: 1-12.
LIU J X, WANG Y, HAN X, et al. Research on edge cloud resource pricing mechanism based on Stackelberg game[J]. *Jour-*

- nal of Frontiers of Computer Science and Technology, 2020: 1-12.
- [39] NGUYEN D T, LE L B, BHARGAVA V. Price-based resource allocation for edge computing: a market equilibrium approach[J]. IEEE Transactions on Cloud Computing, 2021, 9(1): 302-317.
- [40] LI L X, QUEK T Q S, REN J, et al. An incentive-aware job offloading control framework for multi-access edge computing[J]. IEEE Transactions on Mobile Computing, 2021, 20(1): 63-75.
- [41] LI L X, SIEW M, QUEK T Q S, et al. Learning-based priority pricing for job offloading in mobile edge computing[J]. ArXiv Preprint ArXiv, 2019: 1-25.
- [42] ALSKAIF T, GUERRERO ZAPATA M, BELLALTA B. Game theory for energy efficiency in Wireless Sensor Networks: latest trends[J]. Journal of Network and Computer Applications, 2015(54): 33-61.
- [43] 张艮山, 刘旭宁. 资源受限移动边缘计算任务拆分卸载调度决策[J]. 计算机应用与软件, 2019, 36(10): 268-273, 278.
ZHANG G S, LIU X N. Tasks split and offloading scheduling decision in mobile edge computing with limited resources[J]. Computer Applications and Software, 2019, 36(10): 268-273, 278.
- [44] LI Y H, JIANG C S. Distributed task offloading strategy to low load base stations in mobile edge computing environment[J]. Computer Communications, 2020, 164: 240-248.
- [45] ZHANG J, XIA W W, HUANG B N, et al. Joint resource allocation scheme based on evolutionary game for mobile edge computing[J]. Journal of Southeast University (English Edition), 2018, 34(4): 415-422.
- [46] 王艺洁, 凡佳飞, 王陈宇. 云边环境下基于博弈论的两阶段任务迁移策略[J]. 计算机应用, 2021, 41(5): 1392-1398.
WANG Y J, FAN J F, WANG C Y. Two-stage task offloading strategy based on game theory in cloud-edge environment[J]. Journal of Computer Applications, 2021, 41(5): 1392-1398.
- [47] 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 computation offloading and resource allocation in mobile edge computing[J]. Chinese High Technology Letters, 2020, 30(8): 765-773.
- [48] LEITMANN G. Multicriteria Decision Making and Differential Games[M]. Boston, MA: Springer US, 1976.
- [49] JIE Y M, TANG X Y, CHOO K K R, et al. Online task scheduling for edge computing based on repeated Stackelberg game[J]. Journal of Parallel and Distributed Computing, 2018, 122: 159-172.
- [50] LI G S, ZHANG Y, WANG M L, et al. Resource management framework based on the stackelberg game in vehicular edge computing[J]. Complexity, 2020: 8936064.
- [51] CHEN Y F, LI Z Y, YANG B, et al. A Stackelberg game approach to multiple resources allocation and pricing in mobile edge computing[J]. Future Generation Computer Systems, 2020, 108: 273-287.
- [52] HUANG X W, ZHANG W J, YANG J M, et al. Market-based dynamic resource allocation in Mobile Edge Computing systems with multi-server and multi-user[J]. Computer Communications, 2021, 165: 43-52.
- [53] CUI Y Y, ZHANG D G, ZHANG T, et al. Novel method of mobile edge computation offloading based on evolutionary game strategy for IoT devices[J]. AEU - International Journal of Electronics and Communications, 2020, 118: 153134.
- [54] DONG C W, WEN W S. Joint optimization for task offloading in edge computing: an evolutionary game approach[J]. Sensors (Basel, Switzerland), 2019, 19(3): 740.
- [55] DONG Y F, PENG Y F, GUO X P, et al. Offloading decision algorithm using evolutionary game for mobile edge computing[C]//Proceedings of 2019 IEEE 2nd International Conference on Information Communication and Signal Processing. Piscataway: IEEE Press, 2019: 210-214.
- [56] LEI Y, ZHENG W B, MA Y, et al. A novel probabilistic-performance-aware and evolutionary game-theoretic approach to task offloading in the hybrid cloud-edge environment[C]//Collaborative Computing: Networking, Applications and Worksharing, 2021: 255-270.
- [57] GAO H, LI W C, BANEZ R A, et al. Mean field evolutionary dynamics in ultra dense mobile edge computing systems[C]//Proceedings of 2019 IEEE Global Communications Conference. Piscataway: IEEE Press, 2019: 1-6.
- [58] 林艳, 闫帅, 张一晋, 等. 基于交通流量预测的车联网双边拍卖边缘计算迁移方案[J]. 通信学报, 2020, 41(12): 205-214.
LIN Y, YAN S, ZHANG Y J, et al. Flow-of-traffic prediction program based mobile edge computing for Internet of vehicles using double auction[J]. Journal on Communications, 2020, 41(12): 205-214.
- [59] SUN W, LIU J J, YUE Y L, et al. Double auction-based resource allocation for mobile edge computing in industrial Internet of Things[J]. IEEE Transactions on Industrial Informatics, 2018, 14(10): 4692-4701.
- [60] LIN J, HUANG L, ZHANG H L, et al. A novel Latency-Guaranteed based Resource Double Auction for market-oriented edge computing[J]. Computer Networks, 2021, 189: 107873.
- [61] LI Q Y, YAO H P, MAI T L, et al. Reinforcement-learning- and belief-learning-based double auction mechanism for edge computing resource allocation[J]. IEEE Internet of Things Journal, 2020, 7(7): 5976-5985.
- [62] MASHHADI F, MONROY S A S, BOZORGCHENANI A, et al. Optimal auction for delay and energy constrained task offloading in mobile edge computing[J]. Computer Networks, 2020, 183: 107527.
- [63] LUONG N C, XIONG Z H, WANG P, et al. Optimal auction for edge computing resource management in mobile blockchain



- networks: a deep learning approach[C]//Proceedings of 2018 IEEE International Conference on Communications. Piscataway: IEEE Press, 2018: 1-6.
- [64] CAO X Y, ZHANG J S, POOR H V. Mobile edge caching: an optimal auction approach[J]. ArXiv, 2017: 1-14.
- [65] PARKES D C. Iterative combinatorial auctions[M]. Cambridge: MIT press, 2005: 41-78.
- [66] 张海波, 栾秋季, 朱江, 等. 基于移动边缘计算的 V2X 任务卸载方案[J]. 电子与信息学报, 2018, 40(11): 2736-2743.
ZHANG H B, LUAN Q J, ZHU J, et al. V2X task offloading scheme based on mobile edge computing[J]. Journal of Electronics & Information Technology, 2018, 40(11): 2736-2743.
- [67] 项弘禹, 肖扬文, 张贤, 等. 5G 边缘计算和网络切片技术[J]. 电信科学, 2017, 33(6): 54-63.
XIANG H Y, XIAO Y W, ZHANG X, et al. Edge computing and network slicing technology in 5G[J]. Telecommunications Science, 2017, 33(6): 54-63.

[作者简介]



邹璐珊 (1998-), 女, 闽南师范大学计算机学院硕士生, 主要研究方向为移动边缘计算。



黄晓雯 (1995-), 女, 东华大学信息科学与技术学院博士生, 主要研究方向为移动边缘计算、资源分配问题等。



杨敬民 (1980-), 男, 台北科技大学电子工程系博士生, 闽南师范大学计算机学院副教授, 主要研究方向为软件定义网络、计算机网络体系结构和无线通信。

郑艺峰 (1980-), 博士, 闽南师范大学计算机学院讲师, 主要研究方向为移动边缘计算和深度强化学习等。

张光林 (1981-), 男, 博士, 东华大学信息科学与技术学院教授、博士生导师, 主要研究方向为 5G 车联网、智能物联网、移动边缘计算和深度强化学习等。

张文杰 (1984-), 男, 博士, 闽南师范大学计算机学院教授, 主要研究方向为物联网、移动边缘计算和深度强化学习等。