



面向5G网络云原生应用资源调度的博弈优化策略

赵文君,周金和,王 晶

(北京信息科技大学 信息与通信工程学院,北京 100101)

摘 要:随着5G网络和云原生技术的发展,面向服务的5G云原生核心网应运而生,传统应用正朝着云原生化方向发展。目前云原生服务提供商和云原生应用商数量众多且关系复杂,使得应用在云原生化过程中的资源调度面临新挑战。提出一种5G网络云原生应用资源调度优化策略,将云原生应用商和云原生服务提供商构建为多主多从的Stackelberg博弈模型,对传统收益进行具体描述并联合能耗构建利润函数和策略空间,证明给定一组微服务资源定价的情况下存在云原生应用商的纳什均衡点。在此基础上,引入柯西分布对策略进行优化,提高其收敛性能,通过分布式迭代方法得到云原生服务提供商的最佳微服务定价和云原生应用商的最佳微服务租用比例。仿真结果表明,相比ACA算法、QOS PA算法以及GOS策略,该策略能够有效提高网络收益和用户体验质量,同时降低应用开发能耗。

关键词:5G网络;云原生应用;微服务;Stackelberg博弈;网络能耗

开放科学(资源服务)标志码(OSID):



中文引用格式:赵文君,周金和,王晶.面向5G网络云原生应用资源调度的博弈优化策略[J].计算机工程,2021,47(4):40-47.

英文引用格式:ZHAO Wenjun,ZHOU Jinhe,WANG Jing.Game optimization strategy of cloud native application resource scheduling for 5G network[J].Computer Engineering,2021,47(4):40-47.

Game Optimization Strategy of Cloud Native Application Resource Scheduling for 5G Network

ZHAO Wenjun,ZHOU Jinhe,WANG Jing

(School of Information and Communication Engineering, Beijing Information Science and Technology University, Beijing 100101, China)

[Abstract] The development of 5G network and cloud native technologies has led to the emergence of service-oriented 5G cloud native core networks, and traditional applications are also developing in the cloud native direction. However, the large number of cloud native service and application providers and their complex relationships impose a new challenge to the resource scheduling of applications in cloud native development. To address the problem, this paper proposes a game optimization strategy of cloud native application resource scheduling for 5G network. The cloud native service and application providers are built as a multi-leader and multi-follower Stackelberg game model to specify the traditional revenue, which is combined with the energy consumption to build a profit function and strategy space. On this basis, it is proven that there is a Nash equilibrium point for cloud native application providers when a set of microservice resource pricing is given. Then the Cauchy distribution is introduced to optimize the strategy and improve its convergence performance. Through a distributed iterative method, the best microservice price and the optimal microservice lease ratio of cloud native application providers are obtained. The simulation results show that compared with the ACA algorithm, QOS PA algorithm and GOS strategy, the proposed strategy can effectively improve network revenue and user experience quality, and reduce energy consumption for application development.

[Key words] 5G network; cloud native application; microservices; Stackelberg game; network energy consumption

DOI: 10.19678/j.issn.1000-3428.0059050

基金项目:国家自然科学基金“5G超密集接入网智能动态资源分配及其优化方法研究”(61872044)。

作者简介:赵文君(1996—),男,硕士研究生,主研方向为5G网络缓存技术、云计算;周金和,教授;王 晶,硕士研究生。

收稿日期:2020-07-24 **修回日期:**2020-09-25 **E-mail:**2018020189@mail.bistu.edu.cn

0 概述

随着5G网络和云原生技术的发展,2018年中国数字经济的规模已达4.73万亿元,其中70%以上的企业已经使用容器技术或正在测试相关的应用环境,2020年将有约50%的老旧应用以云原生化的方式被改进,云原生应用将成为新常态^[1]。云原生被证明是建设和连续运行世界上最大云的有效加速技术。在当前复杂的云计算环境中,传统应用需要快速响应用户需求,从而提高了应用的复杂性,同时,5G网络产生的新需求也对应用提出了更高的挑战。微服务体系结构的出现成为解决该问题最有效的方案之一,其将应用划分为多个小服务,其中,单个服务专注于特定的功能,并且独立运行在隔离的环境中,然后使用轻量级通信机制进行集成从而形成高内聚低耦合的服务结构^[2]。云原生技术使应用具有更强的可靠性、更短的交付时间、更简化的运维操作以及更高的运营效率。

5G云原生技术的发展使得云原生应用引起研究人员的广泛关注。在大规模应用系统中,微服务数量较多,微服务之间存在复杂的交互关系,导致云原生应用在运行时的资源调度面临新的挑战。微服务之间存在对网络资源的竞争关系,5G云原生服务提供商缺乏充足的资源来满足所有应用程序的需求^[3],因此,将可用的服务资源合理地分配给不同的应用程序具有重要意义。文献[4]研究云分布虚拟资源上的最佳虚拟功能放置问题,使用分支定界和模拟退火2种启发式算法来实现最佳布局,其降低了执行复杂性并缩短了服务迁移的延迟。文献[5]考虑系统调整的延迟性,提出一种基于控制理论前馈和反馈的虚拟资源动态分配方法。文献[6]针对5G云无线电接入网络提出一种动态资源管理策略,设计2种贪婪启发式算法,从而有效降低网络中的服务器总数并提高终端用户的体验质量。文献[7]从计算资源成本的角度出发,以最小化所有背包成本之和为目标研究动态背包问题,实现了服务部署成本的最优化。

虽然上述研究均针对资源利用率和网络延迟2个方面进行了优化,但是都未充分考虑网络收益问题。在资源调度的过程中,能耗是5G网络中亟待解决的问题之一。文献[8]充分考虑网络中的能耗问题,通过提供按需内容交付服务解决了部署成本与服务可用性之间的权衡问题,然后进一步提出多项式时间启发式方法,以对计算资源进行合理分配。

目前,博弈论已经广泛应用于资源分配与优化任务,如互联网定价和5G网络切片资源分配等。文献[9]针对云原生应用运行时计算资源有限的问题,采用拥塞博弈理论对运行时的微服务资源进行调度管理,其有效提高了云原生应用的整体性能。文

献[10]提出支持卸载和迁移的智能网关方法,该方法采用非合作博弈重新调度云环境中的计算任务,降低了能耗和网络延迟,但是其未考虑用户的需求,如用户对应用程序的偏好性等。文献[11]通过对多个用户构建非合作博弈模型,考虑传输时间和网络成本的影响从而确定用户的利润,其提高了用户的体验质量,但未充分考虑资源有限和云提供商的利润问题。由于云环境下应用程序之间的资源竞争行为与经济学中的自由竞争市场类似,因此基于博弈论的方法能够构建资源管理中的竞争关系^[12]。考虑微服务之间的协作与竞争关系,可以将博弈论引入微服务中,通过对有限的网络资源进行竞争以获取更多的网络效益。

云原生应用资源调度的目的包括优化应用时间、保证服务质量、优化费用以及实现高效节能等。本文提出一种面向5G网络云原生应用的资源调度博弈优化策略。将云原生应用程序划分为多个微服务,微服务资源由云原生服务提供商所有,云原生应用为云原生应用商所有。云原生应用商通过租用云原生服务提供商的微服务来构建应用,从而提高收益,而云原生服务提供商通过收取对应的费用作为收益。为了使云原生应用商与云原生服务提供商的收益最大化并最大限度地降低能耗,本文将云原生应用资源调度问题建模为多主多从的Stackelberg博弈模型,构建收益函数对传统收益进行具体化描述,并引入用户偏好性指标以提高两者的传统收益。为解决能耗问题,本文对能耗和传统收益进行联合建模,并证明纳什均衡解的存在。在此基础上,引入柯西分布提高策略的收敛性,使用分布式迭代方法确定云原生应用商的最佳租用比例和云原生服务提供商的最佳定价,最后通过Stackelberg博弈实现效用最大化。

1 系统模型

根据5G网络架构^[13],本文将云原生服务提供商和数据网络中的云原生应用商构建为图1所示的系统模型。

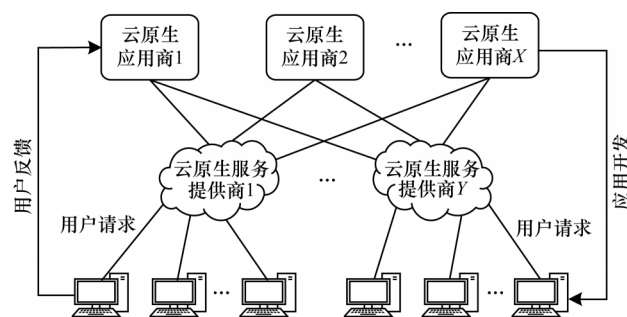


图1 系统模型

Fig.1 System model

由图1可知,本文系统模型主要由 X 个云原生应用商 $N_{\text{CNA}_i} (i=1,2,\dots,X)$ 和 Y 个云原生服务提供商 $N_{\text{CSP}_j} (j=1,2,\dots,Y)$ 构成。云原生应用商拥有有限的微服务资源,可以将微服务资源租给云原生应用商,以便对云原生应用进行开发。其中,微服务资源服从密度为 φ 的异构齐次泊松点分布。应用开发的资源调度过程为:

1)云原生服务提供商给定微服务资源的定价,并告知云原生应用商。

2)云原生应用商根据应用的性能和微服务的定价等多种因素确定自身租用微服务资源的比例,并将自身应用开发为云原生应用。

3)用户依据自身偏好和满意度对应用进行访问。

1.1 云原生应用流行度

在现实中,用户对不同的应用具有不同的偏好,因此,应用之间的流行度存在一定差别,应用流行度越高,用户的访问频率越高,相应的收益越大。因此,引入应用流行度这一指标可以更合理地确定微服务的租用比例,从而提高云原生应用的收益。假设存在 N 个应用,用集合表示为 $A=\{A_1, A_2, \dots, A_N\}$,第 n 个应用的流行度为 $D_n, n=\{1,2,\dots,N\}$ 。不同的应用有不同的流行度,假设云原生应用商的应用集依据应用流行度由高到低降序排列,则用户向应用请求服务的概率为 D_n ,其服从Zipf分布^[14]。

随着移动终端的普及,一些流行的应用开始为人们所熟知,如抖音、快手等。本文引入调节因子 θ ,以调整应用的流程度,从而提高云原生应用的流行度与云原生应用商的收益。 D_n 的计算公式如下:

$$D_n = \frac{\theta/n^\alpha}{\sum_{i=1}^N (1/i^\alpha)}, \forall n \quad (1)$$

其中, α 为云原生应用的流行度指数,流行度随着 α 的增大而提高,即云原生应用1最受用户欢迎,云原生应用 N 最不受用户欢迎。

1.2 用户偏好性

用户偏好性直接影响云原生网络的收益,一般而言,如果用户偏好性低,云原生应用的收益就会降低,反之亦然。因此,本文引入用户偏好性指标来衡量应用的服务质量,通过用户反馈来调整云原生应用商的收益。用户偏好性表示用户使用云原生应用的体验水平,偏好性直接受到应用时延、能耗大小等多方面因素的影响。文献[15]建立用户偏好性和网络服务质量之间的多维函数关系,以此为参考,本文对用户偏好性进行定义。假设每个应用有 M 个评价指标,每个评价指标都有 R 个离散的取值,则每个应用对应一个应用偏好矩阵 G :

$$G = \begin{bmatrix} g_{11} & g_{12} & \dots & g_{1R} \\ g_{21} & g_{22} & \dots & g_{2R} \\ \vdots & \vdots & & \vdots \\ g_{M1} & g_{M2} & \dots & g_{MR} \end{bmatrix}$$

其中, g_{mr} 表示用户对应用的第 m 个特性的偏好程度, $g_{mr} \in [0,1]$,0代表用户反感,1代表用户喜欢, $m \in \{1,2,\dots,M\}, r \in \{1,2,\dots,R\}$ 。 $W = [W_1 \ W_2 \ \dots \ W_M]$ 为 M 个评价指标的权重值矩阵,用户对第 n 个应用的偏好性如式(2)所示:

$$B_n = \frac{\lambda}{\lambda + \varepsilon} \frac{1}{R} \sum_{r=1}^R W_r G_r \quad (2)$$

其中, ε 为用户偏好性的影响因子, λ 为云原生应用商给出的微服务的租用比例。评价指标的取值数量 R 可以自行设定,依据大数定律, R 趋于无穷大时, B_n 收敛到常数值,具体如下:

$$\lim_{R \rightarrow \infty} B_n = \lim_{R \rightarrow \infty} \frac{\lambda}{\lambda + \varepsilon} \frac{1}{R} \sum_{r=1}^R W_r G_r = \frac{\lambda}{\lambda + \varepsilon} \quad (3)$$

1.3 云原生应用商能耗

云原生应用商提供云原生应用的过程主要包括:1)云原生应用的开发,其中主要为微服务的开发;2)维持云原生应用的运行。因此,云原生应用商能耗被定义为微服务开发能耗 E_d 和应用运行能耗 E_r 之和,单位为瓦特。应用运行能耗 E_r 包括微服务运行能耗 E_{mr} 和传统应用运行能耗 E_{tr} 。微服务开发能耗是开发时间的函数,表示为 $t_{\text{dev}} e_{\text{dev}}$,其中, e_{dev} 为单位时间内微服务开发所需能耗。综上,云原生应用商的总能耗计算公式如式(4)所示:

$$E_x = \sum_{y=1}^Y \lambda_{xy} E_d + E_r + E_{\text{idle}} = \sum_{y=1}^Y \left[\lambda_{xy} t_{\text{dev}} e_{\text{dev}} + \lambda_{xy} t_{\text{run}} e_{mr} + (1 - \lambda_{xy}) t_{\text{run}} e_{tr} + (T_{\text{total}} - t_{\text{dev}} - t_{\text{run}}) e_{\text{idle}} \right] \quad (4)$$

其中, e_{mr}, e_{tr} 分别为单位时间内微服务运行和传统应用运行的能耗, $t_{\text{dev}}, t_{\text{run}}$ 分别为微服务开发和应用运行的时间, e_{idle} 为空闲时应用所消耗的能耗, T_{total} 为应用开发和运行的总时间。

2 博弈创建

面向5G网络的云原生应用资源调度策略采用Stackelberg博弈进行优化。Stackelberg博弈模型至少包括博弈方、博弈的策略空间和博弈的效用函数^[16]3个元素。Stackelberg博弈是一种符合主从关系的博弈,博弈方一般由主导者和追随者构成,博弈双方都存在各自的效用函数和策略空间^[17]。本文的博弈过程可描述为:

- 1)主导者根据效用函数做出决策。
- 2)追随者根据主导者的效用函数和决策给出自

己的决策。

3)主导者根据追随者的决策和自己的效用函数调整自己的决策。

4)博弈双方通过调整各自的决策从而达到纳什均衡,最终使双方的效用达到最优。

2.1 博弈构成

Stackelberg博弈模型的3个元素具体如下:

1)博弈方:主导者是 X 个云原生应用商,追随者是 Y 个云原生服务提供商。

2)博弈的策略空间:主导者的策略空间是云原生应用商给出的微服务的租用比例 λ_{xy} ,追随者的策略空间是云原生服务提供商制定的微服务价格 $P = \{P_1, P_2, \dots, P_Y\}$ 。

3)博弈的效用函数:主导者的效用函数是云原生应用商的效用 U_x ,追随者的效用函数是云原生服务提供商的效用 U_y 。云原生应用商的效用函数 U_x 由云原生应用商所得收益 U_x^{prof} (如式(5)所示)与用户租用微服务的租金 U_x^{rent} 和运行能耗 E_x 之差构成,如式(6)所示。

$$U_x^{\text{prof}} = \sum_{y=1}^Y \sum_{n=1}^N TSD_n B_n \quad (5)$$

$$U_x = U_x^{\text{prof}} - U_x^{\text{rent}} - E_x \quad (6)$$

$$U_x^{\text{rent}} = \sum_{y=1}^Y \lambda_{xy} \phi_y P_y \quad (7)$$

其中, T 为单位时间内应用的被访问次数, S 为每次服务的单价。

云原生服务提供商的效用函数 U_y 由租用微服务收取的费用 $P_y \phi_y \lambda_{xy}$ 和微服务管理成本 $C_y \phi_y \lambda_{xy}$ 之差构成,如式(8)所示:

$$U_y = \sum_{x=1}^X (P_y - C_y) \phi_y \lambda_{xy} \quad (8)$$

2.2 问题创建

在微服务分配的过程中,云原生应用商通过租用云原生服务提供商的微服务来开发自身应用,以降低应用的时延和能耗,从而为用户提供优质的服务并提高自己的收益。云原生服务提供商可以通过租用微服务来获取一定的收益,最终通过博弈使云原生应用商与云原生服务提供商的收益最大化。

云原生服务提供商的优化问题可以视为使其效用函数最大化的问题,表达式为:

$$\text{Max } U_y(P, \lambda) \quad (9)$$

云原生应用商的优化问题也可以创建为其效用函数最大化问题,即:

$$\text{Max } U_x(P, \lambda) \quad (10)$$

效用函数最大化意味着云原生应用商收益最大

化且能耗达到最优。通过对式(7)进行推导,可以得到云原生应用商的效用函数,即:

$$\begin{aligned} U_x &= U_x^{\text{prof}} - U_x^{\text{rent}} - E_x = \\ &= \sum_{y=1}^Y \sum_{n=1}^N TSD_n B_n - \sum_{y=1}^Y \lambda_{xy} \phi_y P_y - \sum_{y=1}^Y \lambda_{xy} E_d + E_r = \\ &= \sum_{y=1}^Y \sum_{n=1}^N \frac{\lambda TSD_n \sum_{r=1}^R WG}{R(\lambda + \epsilon)} - \sum_{y=1}^Y \lambda_{xy} \phi_y P_y - \\ &= \sum_{y=1}^Y [E_{tr} + \lambda_{xy} (E_d + E_{mr} - E_{tr})] \end{aligned} \quad (11)$$

通过博弈构建过程可知,云原生服务提供商的博弈目的是最优化自己的效用函数值。云原生应用商依据收益和能耗租用一定比例 λ 的微服务, λ_{xy} 的大小取决于云原生服务提供商的微服务定价 P_y ,如果微服务的定价过高,云原生应用商将不会租用微服务;反之,如果微服务定价过低,云原生服务提供商难以获得利润。因此,云原生服务提供商只有进行合理地定价,云原生应用商才能选择合适的租用比例,使双方效用函数最优化,从而提高用户体验质量并使能耗达到最优。

2.3 纳什均衡

Stackelberg博弈由多个云原生应用商和多个云原生服务提供商构成,由于云原生服务提供商资源有限,因此多个云原生应用商之间存在竞争关系。同时,每个云原生应用商的租用策略会影响云原生服务提供商的微服务定价,从而影响其他云原生服务提供商的租用策略,因此,云原生应用商之间存在非合作博弈关系。

纳什均衡是非合作博弈的最优解,是博弈参与者之间策略空间的稳定状态,即不存在一个博弈参与者能够通过改变对应的策略空间来取得更多的收益^[18]。

对于非合作博弈而言,如果博弈满足如下条件:

1)博弈方的集合有限。

2)博弈参与者的策略空间集合为欧氏空间中的封闭、有界以及非空的凸集。

3)博弈的效用函数在博弈参与者的策略空间中满足连续的性质且为凹函数。

则可以证明该博弈过程存在纳什均衡解,每一个博弈方的效用函数都可以达到最优化,且任何博弈的参与者不能通过改变自己的策略来获得更高的收益^[19]。

纳什均衡的证明过程如下:

1)设主导者(云原生应用商)为 X 个,追随者(云原生服务提供商)为 Y 个,则对应的集合有限。

2)博弈参与者的策略空间显然是欧氏空间中的有界非空闭集,且效用函数在策略空间上连续。

3)对效用函数 U_x 求一阶和二阶偏导数分别可得:

$$\frac{\partial U_x}{\partial \lambda} = \frac{\varepsilon \sum_{n=1}^N TSD_n \sum_{r=1}^R WG}{R(\lambda + \varepsilon)^2} - (\varphi_y P_y + E_d + E_{mr} - E_{tr}) \quad (12)$$

$$\frac{\partial^2 U_x}{\partial \lambda^2} = \frac{-2\varepsilon \sum_{n=1}^N TSD_n \sum_{r=1}^R WG}{R(\lambda + \varepsilon)^3} < 0 \quad (13)$$

显然,存在一阶偏导数(如式(12)所示)和二阶偏导数(如式(13)所示),且二阶偏导数小于0,可得效用函数在策略空间上满足严格凹函数特性。

综上,云原生应用商之间的非合作博弈存在纳什均衡解,证毕。

3 策略优化

云原生服务提供商的微服务资源有限,因此,云原生应用商的租用比例应该符合实际情况。本文考虑租用比例的合理性,通过求最值来确定租用比例的边界条件,然后通过拉格朗日乘数法对博弈进行优化求解,最后利用迭代方法给出云原生应用商的最佳租用比例以及云原生服务提供商的最佳定价。策略优化过程具体如下:

假设微服务资源足够所有云原生应用商使用,云原生服务提供商给出一组微服务租用价格 $P = \{P_1, P_2, \dots, P_y\}$,通过求导可得云原生应用商的最佳租用比例,如下:

$$\lambda = \left[\sqrt{\frac{\varepsilon \sum_{n=1}^N TSD_n \sum_{r=1}^R WG}{R(\varphi_y P_y + E_d + E_{mr} - E_{tr})}} - \varepsilon \right]^+ \quad (14)$$

由式(14)可知,当 $\lambda=0$,即云原生应用商的租用比例为0时,云原生应用商将不会租用微服务,此时云原生服务提供商和云原生应用商都不会在云原生网络中受益,而且用户体验质量不能得到改善。因此,存在微服务定价的最大值,由 $\lambda=0$ 可以求得微服务定价的最大值为:

$$P_y^{\max} = \frac{1}{\varphi_y} \left[\frac{\sum_{n=1}^N TSD_n \sum_{r=1}^R WG}{\varepsilon R} - (E_d + E_{mr} - E_{tr}) \right] \quad (15)$$

当 $\lambda=1$,即云原生应用商的租用比例为1时,云原生应用商将租用所有微服务,由于云原生服务提供商资源有限,而且租用所有微服务会导致云原生服务提供商效用减少,因此存在微服务定价的最小值。由 $\lambda=1$ 可以求得微服务定价的最小值为:

$$P_y^{\min} = \frac{1}{\varphi_y} \left[\frac{\varepsilon \sum_{n=1}^N TSD_n \sum_{r=1}^R WG}{R(\varepsilon + 1)^2} - (E_d + E_{mr} - E_{tr}) \right] \quad (16)$$

综上,存在微服务定价的最大值和最小值,当定

价小于最小值 P_y^{\min} 时,云原生服务提供商应该上调当前的租用价格;当定价大于最大值 P_y^{\max} 时,云原生服务提供商应该下调当前的租用价格。通过多次调整来达到最佳微服务定价,最终使两者利润最大化,同时提高用户的体验质量。

在实际情况下,云原生服务提供商的微服务资源情况不可忽略,其需满足式(17),而且云原生应用商的租用比例应该满足 $0 \leq \lambda \leq 1$ 。

$$F_x = \sum_{y=1}^Y \lambda_{xy} \varphi_y \leq F_{\max} \quad (17)$$

其中, F_{\max} 是微服务资源数量的最大值。

本文利用拉格朗日乘数法对博弈进行优化求解,首先构建拉格朗日函数如下:

$$L_x = \sum_{y=1}^Y \sum_{n=1}^N TSD_n B_n - \sum_{y=1}^Y \lambda_{xy} \varphi_y P_y - \sum_{y=1}^Y [E_{tr} + \lambda_{xy} (E_d + E_{mr} - E_{tr})] + \omega \lambda_{xy} + \zeta \lambda_{xy} + v \lambda_{xy} \varphi_y \quad (18)$$

其中, ω, ζ, v 为拉格朗日函数中的拉格朗日乘数。拉格朗日函数的充分必要约束条件为:

$$\begin{cases} \frac{\partial L_x}{\partial \lambda_{xy}} = 0 \\ \omega \geq 0, \zeta \geq 0, v \geq 0 \\ \omega \lambda_{xy} = 0 \\ \zeta (\lambda_{xy} - 1) = 0 \end{cases} \quad (19)$$

经过推导可得微服务的最佳租用比例如式(20)所示:

$$\lambda_{xy}^* = \begin{cases} 1, \gamma < \frac{\varepsilon \sum_{n=1}^N TSD_n \sum_{r=1}^R WG}{R(\varphi_y P_y + E_d + E_{mr} - E_{tr}) (1 + \varepsilon)^2} - 1 \\ 0, \gamma > \frac{\sum_{n=1}^N TSD_n \sum_{r=1}^R WG}{R(\varphi_y P_y + E_d + E_{mr} - E_{tr})} - 1 \\ \left[\sqrt{\frac{\varepsilon \sum_{n=1}^N TSD_n \sum_{r=1}^R WG}{R(\varphi_y P_y + E_d + E_{mr} - E_{tr}) (1 + \gamma)}} - \varepsilon \right]^+, \text{其他} \end{cases} \quad (20)$$

其中, γ 的表达式如式(21)所示,其为云原生服务提供商微服务资源的约束条件。

$$\gamma = \left(\frac{1}{F_{\max} - \sum_{y=1}^Y \lambda_{xy} \varphi_y} \right)^2 - 1 \quad (21)$$

对云原生服务提供商的效用函数求最大值可得微服务资源的最优定价 P_y ,将微服务的资源租用比例最佳取值 λ_{xy}^* 代入式(8)进行求导,求导结果如

式(22)所示,令导数为0可反解出 P_y ,每次更新的微服务资源定价如式(23)所示。

$$\frac{\partial U_y}{\partial P_y} = \varphi_y \lambda_{xy}^* + \varphi_y (P_y - C_y) \frac{\partial \lambda_{xy}}{\partial P_y} \quad (22)$$

$$P_y^* = C_y - \frac{\lambda_{xy}^*}{\partial \lambda_{xy} / \partial P_y} \quad (23)$$

由式(23)可知,单个微服务资源的最优定价是动态变化的,它与其他云原生服务提供商的微服务定价密切相关。因此,本文采用迭代法对微服务的定价进行求解,迭代公式如下:

$$P_y^{t+1} = C_y - \frac{\rho \lambda_{xy}^*}{\partial \lambda_{xy} / \partial P_y^t} \quad (24)$$

其中, t 为迭代次数, P_y^t 为第 t 次迭代时微服务资源的定价, C_y^t 为第 t 次迭代时微服务资源管理的成本价, ρ 为迭代步长,其为一个递减参数,取值随着迭代次数的变大而逐渐变小。在实际情况中,递减参数值过小或过大都不利于纳什均衡点逼近,因此,本文引入柯西分布^[20]对该系数进行优化,柯西分布较高的两翼概率特性使其可以产生远离原点并具有较宽分布范围的随机数,使策略可以在更宽的范围内寻找纳什均衡点,从而提高策略的收敛性能。优化后的迭代步长 ρ^* 表达式如下:

$$\rho^* = \frac{1}{2} + \frac{1}{\pi} \arctan \rho \quad (25)$$

当迭代次数为 $t+1$ 时,若云原生服务提供商和云原生应用商的效用值达到最大,则终止迭代过程;反之,进入下一个周期,直到两者效用值最大后停止迭代过程。

4 仿真分析

本文通过MATLAB R2016b对所提策略进行仿真验证,网络模型由2个云原生服务提供商和6个云原生应用商构成,云原生应用商竞争微服务资源以开发应用。具体的仿真参数设置如下:单个云原生服务提供商拥有500个微服务资源, $\lambda=0.6$, $\varepsilon=0.2$, $R=50$, $T=100$, $S=2$ 。在仿真中,将本文策略与蚁群(ACA)算法^[21]、全局最优(GOS)策略^[22]、QOS优先(QOS PA)算法^[23],分别在云原生应用商效用、用户满意度和能耗减少率3个方面进行对比分析。

图2所示为云原生服务提供商1和云原生服务提供商2之间微服务定价的关系,曲线上的点表示当前云原生服务提供商相对于另一个云原生服务提供商的最优定价策略。从图2可以看出,(0.49,0.4)为2条曲线的交点,即纳什均衡点,此处代表双方定价和效用达到最优。虽然云原生服务提供商之间的最优定价互相影响,但是总存在一组最优定价使双方效用最大化。

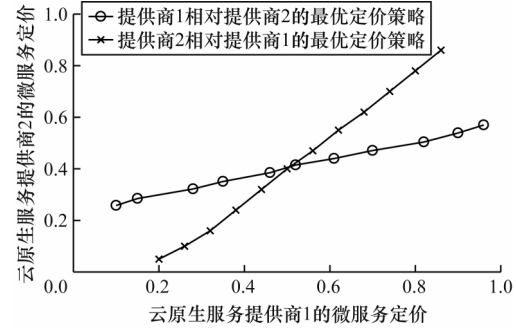


图2 2个云原生服务提供商之间的微服务定价关系

Fig.2 Microservices pricing relationship between two cloud native service providers

图3所示为不同应用流行指数下云原生应用商的收益变化情况。从图3可以看出,在相同迭代次数情况下, α 值越大,云原生应用商越受欢迎,对应的云原生应用商收益越大。在 α 一定的情况下,随着迭代次数的增大,云原生应用商收益逐渐提高,当迭代次数约为50时,云原生应用商收益值收敛到当前情况下的最大值。

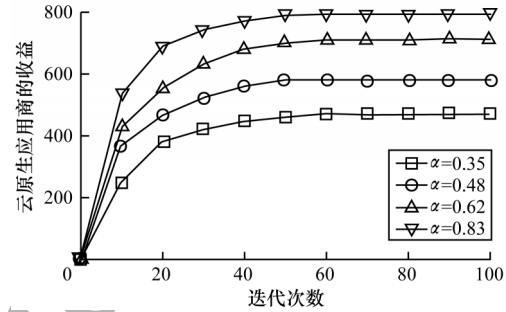


图3 云原生应用商收益与迭代次数之间的关系

Fig.3 The relationship between the benefits of cloud native application business and the number of iterations

图4所示为不同策略和算法下云原生应用商的效用变化情况。从图4可以看出,本文策略的效用函数明显优于其他策略和算法,且在4种不同的策略和算法中,GOS策略的收敛性能最好,在迭代次数约为45时其收敛到最大值。

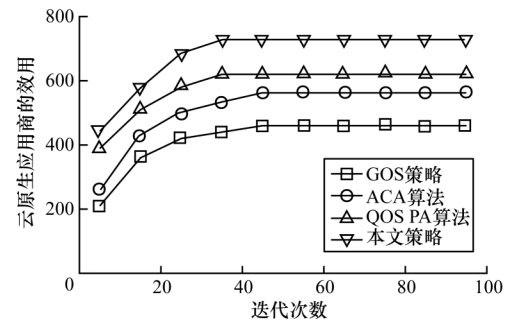


图4 云原生应用商效用和迭代次数之间的关系

Fig.4 The relationship between the utility of cloud native application business and the number of iterations

图5所示为不同策略和算法下用户平均偏好度和用户对应用的使用次数之间的关系,从图5可以看出,随着用户对应用使用次数的增加,用户的平均偏好度逐渐增加,且本文优化策略的用户平均偏好度明显优于其他3种策略和算法。

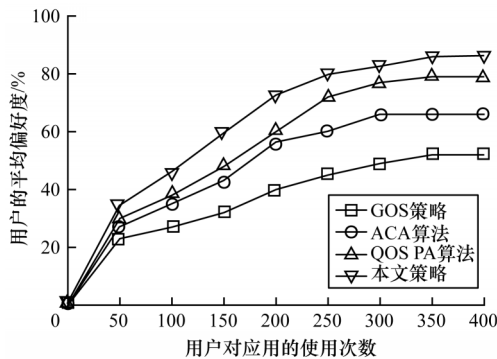


图5 应用使用次数与用户平均偏好度的关系

Fig.5 The relationship between the number of applications used and the average user preference

图6所示为不同策略和算法下应用使用次数与能耗减少率的关系,本次实验中假定不同策略和算法下用户产生的请求一致。从图6可以看出,随着应用使用次数的增加,能耗减少率逐渐提高,且本文优化策略的能耗减少率性能优于其他3种策略和算法。

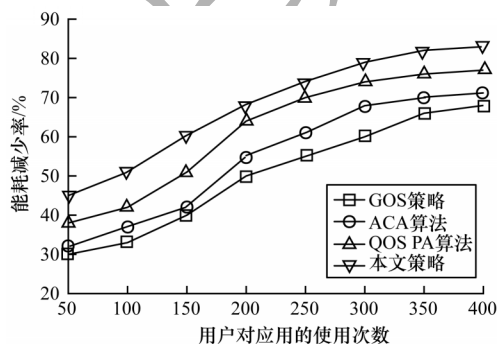


图6 应用使用次数与能耗减少率的关系

Fig.6 The relationship between the number of applications used and energy consumption reduction rate

5 结束语

本文针对5G网络中云原生应用的微服务资源调度问题进行研究,提出一种面向云原生应用资源调度的博弈优化策略,以提高云原生应用商和云原生服务提供商的收益,降低网络能耗,提高用户的体验质量。将云原生应用资源调度问题建模为多主多从的Stackelberg博弈模型,对传统收益进行具体描述,引入用户偏好性指标,以提高云原生应用商和云原生服务提供商的传统收益。针对资源调度过程中

的能耗问题,对能耗和传统收益进行联合建模,证明纳什均衡解的存在,并使用分布式迭代方法确定云原生应用商的最佳租用比例和云原生服务提供商的最佳定价,在此基础上,通过Stackelberg博弈实现效用的最大化。仿真结果表明,该策略能够有效提高网络收益并降低网络能耗。本文工作对构建未来的绿色网络有一定参考意义,下一步将在5G网络和信息中心网络中联合部署云原生应用,以提高内容分发率并降低网络能耗。

参考文献

- [1] PELLE I, CZENTYE J, DÓKA J, et al. Towards latency sensitive cloud native applications: a performance study on AWS [C]//Proceedings of 2019 IEEE International Conference on Cloud Computing. Washington D. C., USA: IEEE Press, 2019: 272-280.
- [2] WANG Huaimin, SHI Peichang, DING Bo, et al. Online evolution of software services [J]. Chinese Journal of Computers, 2011, 34(2): 318-328. (in Chinese)
王怀民, 史佩昌, 丁博, 等. 软件服务的在线演化[J]. 计算机学报, 2011, 34(2): 318-328.
- [3] ZHOU Jinhe, ZHAO Wenjun, CHEN Shuo. Dynamic network slice scaling assisted by prediction in 5G network [J]. IEEE Access, 2020, 8: 133700-133712.
- [4] SARKAR A, MURUGAN T S. Cluster head selection for energy efficient and delay-less routing in wireless sensor network [J]. Wireless Networks, 2019, 25(1): 303-320.
- [5] BHAMARE D, ERBAD A, JAIN R, et al. Efficient virtual network function placement strategies for cloud radio access networks [J]. Computer Communications, 2018, 127: 50-60.
- [6] MISHRA D, TAMMA B R. Dynamic resource management of cloud native control units in 5G radio access networks [EB/OL]. [2020-06-15]. <http://raiith.iith.ac.in/3733/>.
- [7] YAN Mu, FENG Gang, ZHOU Jianhong, et al. Intelligent resource scheduling for 5G radio access network slicing [J]. IEEE Transactions on Vehicular Technology, 2019, 68(8): 7691-7703.
- [8] YALA L, FRANGOUDIS P A, LUCARELLI G, et al. Cost and availability aware resource allocation and virtual function placement for CDNaaS provision [J]. IEEE Transactions on Network and Service Management, 2018, 15(4): 1334-1348.
- [9] LUO Ruici, YE Wei, LIU Xueyang, et al. A runtime resource management approach of microservices based on congestion game [J]. Acta Electronica Sinica, 2019, 47(7): 1497-1505. (in Chinese)
罗睿辞, 叶蔚, 刘学洋, 等. 基于拥塞博弈的微服务运行时资源管理方法 [J]. 电子学报, 2019, 47(7): 1497-1505.
- [10] BALASUBRAMANIAN S, MEYYAPPAN T. Game theory based offload and migration-enabled smart gateway for cloud of things in fog computing [M]. Berlin, Germany: Springer, 2020.

- [11] QIN An, CAI Chengcheng, WANG Qin, et al. Game theoretical multi-user computation offloading for mobile-edge cloud computing [C]//Proceedings of 2019 IEEE Conference on Multimedia Information Processing and Retrieval. Washington D. C., USA: IEEE Press, 2019: 328-332.
- [12] LÜ Xinchun. Task offloading and resource management in mobile edge computing [D]. Beijing: Beijing University of Posts and Telecommunications, 2019. (in Chinese)
吕昕晨. 移动边缘计算任务迁移与资源管理研究 [D]. 北京: 北京邮电大学, 2019.
- [13] ZHAO Wenjun, ZHOU Jinhe. Stackelberg game cache optimization strategy for 5G networks [J]. Computer Engineering, 2019, 45(12): 52-57, 78. (in Chinese)
赵文君, 周金和. 面向5G网络的Stackelberg博弈的缓存优化策略 [J]. 计算机工程, 2019, 45(12): 52-57, 78.
- [14] QUIROZ J C, LARANJO L, TUFANARU C, et al. Empirical analysis of Zipf's law, power law, and lognormal distributions in medical discharge reports [EB/OL]. [2020-06-15]. <https://arxiv.org/vc/arxiv/papers/2003/2003.13352v1.pdf>.
- [15] MING C L, ANDREAS F. MOLISCH N S, et al. Individual preference probability modeling and parameterization for video content in wireless caching networks [J]. IEEE/ACM Transactions on Networking, 2019, 27(2): 676-690.
- [16] HAN Zhenzhen, XU Chuan, WANG Qianyun, et al. The energy saving scheme of WLAN based on Bayesian game [J]. Acta Electronica Sinica, 2019, 47(10): 2083-2088. (in Chinese)
韩珍珍, 徐川, 王倩云, 等. 基于贝叶斯博弈的WLAN节能机制研究 [J]. 电子学报, 2019, 47(10): 2083-2088.
- [17] JIE Yingmo, LI Mingchu, GUO Cheng, et al. Game-theoretic online resource allocation scheme on fog computing for mobile multimedia users [J]. China Communications, 2019, 16(3): 22-31.
- [18] KONSTANTIN A, JASPER G, BERKSAN S. A low-complexity approach to distributed cooperative caching with geographic constraints [EB/OL]. [2020-06-15]. <https://arxiv.org/pdf/1704.04465.pdf>.
- [19] GUO Fengxian, ZHANG Heli, LI Xi. Joint optimization of caching and association in energy-harvesting-powered small-cell networks [J]. IEEE Transactions on Vehicular Technology, 2018, 67(7): 6469-6480.
- [20] WANG Chuntao, XIAO Deqing, PENG Hongxing, et al. A lossy compression scheme for encrypted images exploiting cauchy distribution and weighted rate distortion optimization [J]. Journal of Visual Communication & Image Representation, 2018, 51: 122-130.
- [21] CHEN Xuan, XU Jianwei, LONG Dan. Resource scheduling algorithm of cloud computing based on ant colony optimization-shuffled frog leading algorithm [J]. Journal of Computer Applications, 2018, 38(6): 1670-1674, 1681. (in Chinese)
陈暄, 徐见伟, 龙丹. 基于蚁群优化-蛙跳算法的云计算资源调度算法 [J]. 计算机应用, 2018, 38(6): 1670-1674, 1681.
- [22] LI Jun, CHEN He, CHEN Youjia, et al. Pricing and resource allocation via game theory for a small-cell video caching system [J]. IEEE Journal on Selected Areas in Communications, 2016, 34(8): 2115-2129.
- [23] PRATEEKSHA V, YOGESH S. Characterizing application scheduling on edge, fog and cloud computing resources [EB/OL]. [2020-06-15]. <https://arxiv.org/pdf/1904.10125.pdf>.

编辑 吴云芳