Journal of Computer Applications
计算机应用,2018,38(10):2753 -2758
ISSN 1001-9081
CODEN JYIIDU
2018-10-10
http://www.joca.cn

文章编号:1001-9081(2018)10-2753-06
DOI:10.11772/j.issn.1001-9081.2018041187

基于不确定服务质量感知的云服务组合方法

王思臣[1],涂 辉[2],张以文[1*]
(1.计算智能与信号处理教育部重点实验室(安徽大学),合肥 230031;
2.瓦斯治理国家工程研究中心(淮南矿业集团),安徽 淮南 232001)
(* 通信作者电子邮箱 zhangyiwen@ahu.edu.cn)

摘 要: 针对不确定服务质量(QoS)感知的云服务组合优化问题的求解,提出一种不定长时间序列(ULST)模型和锦标赛策略的改进遗传算法(T-GA)。首先,基于用户对服务不同时间段的访问规律,将服务质量的长期变化构建为不定长时间序列模型,该模型能够准确地描述一段时间内用户对服务的真实 QoS 访问记录。其次,提出一种基于不确定 QoS 模型的改进遗传算法,该算法采用锦标赛选择策略代替基本遗传算法中的轮盘赌选择策略。最后,在真实数据上进行了大量实验,所提的不定长时间序列模型能够有效地解决不确定 QoS 感知云服务组合问题,而锦标赛策略的改进遗传算法在寻优结果和稳定性方面均优于基于精英选择策略的遗传算法(E-GA)算法,且运行速度提高近 1 倍,是可行、高效且稳定的算法。

关键词: 不确定 QoS; 长期; 服务组合; 优化算法
中图分类号: TP18 **文献标志码:** A

Cloud service composition method based on uncertain QoS-aware ness

WANG Sichen[1], TU Hui[2], ZHANG Yiwen[1*]
(1. *Key Laboratory of Intelligent Computing and Signal Processing of Ministry of Education (Anhui University), Hefei Anhui* 230031, *China*;
2. *National Engineering Research Center for Coal Mine Gas Controlling (Huainan Mining (Group) Company Limited), Huainan Anhui* 232001, *China*)

Abstract: To solve the problem of uncertain Quality of Service (QoS)-aware cloud service composition optimization, an Uncertain-Long Time Series (ULST) model and Tournament strategy based Genetic Algorithm (T-GA) was proposed. Firstly, based on different access rules of users to services in different periods, the long-term change of QoS was modeled as an uncertain-long time series, which can accurately describe the users' actual QoS access record to service over a period of time. Secondly, an improved genetic algorithm based on uncertain QoS model was proposed, which used tournament strategy instead of basic roulette wheel selection strategy. Finally, a lot of experiments were carried out on real data. The uncertain-long time series model can effectively solve the problem of uncertain QoS-aware cloud service composition; the proposed T-GA is superior to the Genetic Algorithm based on Elite selection strategy (E-GA) in optimization results and stability, and the execution speed is improved by almost one time, which is a feasible, high efficient and stable algorithm.

Key words: uncertain QoS; long-term; service composition; optimization algorithm

0 引言

随着云计算的发展,越来越多的资源以服务的形式发布和使用,如何选择高性能的服务来构建增值的应用已成为研究热点。服务通常由功能属性和非功能属性构成,非功能属性也称为服务质量(Quality of Service, QoS),描述了服务在完成其功能的过程中表现出的执行特性,由响应时间、吞吐量、延迟、成功率、花费、信誉度等一系列属性构成。然而,在许多情况下,单个服务是无法满足用户复杂需求的,需要根据用户需求进行服务组合。由于功能属性相似而非功能属性不同的服务数量迅速增加,服务选择变得越来越复杂,使得服务组合问题逐渐演变为 NP(Non-derministic Polynomial)完全问题。为了高效地找到最优服务组合,许多方法被相继提出。常用的方法有整数线性规划(Integral Linear Programming, ILP)、混合线性规划(Mixed Linear Programming, MLP)和人工智能(Artificial Intelligence, AI)算法[1-7]。然而,这些方法大多只考虑确定性 QoS,而忽略现实中 QoS 的不确定性。文献[8]考虑到不确定的服务质量,将具有不确定 QoS 的服务组合问题建模为一个区间数多目标优化问题(Interval Number Multi-Objective Optimization Problem, IMOP),并提出了一种基于分解的非确定性多目标进化算法。以上这些方法只考虑组合服务在执行阶段的 QoS 性能,即组合服务的瞬时 QoS 性能,而忽略了云环境的不确定性。服务组合的成功在很大程度上取决于不同组件服务保持长期服务质量(QoS)满足用户需求的能力,与传统的服务组合相比,云服务组合通常是从长期角度考虑的。文献[9]定义了一种长期 QoS 感知云服务组合的方

收稿日期:2018-03-28;修回日期:2018-06-12;录用日期:2018-06-14。 基金项目:国家自然科学基金资助项目(61872002);国家科技支撑计划项目(2015BAK24B01);安徽省自然科学基金资助项目(1808085MF197);安徽省科技重大专项(16030901062)。
作者简介:王思臣(1991—),男,河南商丘人,硕士研究生,主要研究方向:服务计算、进化计算; 涂辉(1984—),男,江西宜春人,硕士研究生,主要研究方向:大数据、云计算; 张以文(1976—),男,安徽马鞍山人,副教授,博士,CCF 会员,主要研究方向:服务计算、大数据、云计算。

法，引入了三种元启发式算法，即遗传算法（Genetic Algorithm，GA）、模拟退火算法和禁忌搜索算法，以选择具有最佳平均长期 QoS 的服务；然而，他们认为每个服务的时间序列的长度是固定的，通过随机采样对服务进行选择。

基于以上分析，本文认为云服务组合是一个长期优化过程，在长期优化过程中，由于云环境的动态性，使得服务的 QoS 可能发生变化，需要在用户长期使用过程中最大化满足用户需求。本文中，首先考虑现实世界中用户对服务的访问规律，即一段时间内单个服务的被访问频率，提出一种基于不确定 QoS 感知的服务组合时间序列模型，该模型使用不定长时间序列描述一段时间内 QoS 值的变化，能够准确地描述用户对服务的真实访问记录；其次，提出一种基于不确定 QoS 模型的改进遗传算法，该算法通过采用锦标赛选择策略改进选择操作，使得种群保持了多样性，提高了群体收敛速度。

1 相关工作

近年来，基于 QoS 感知的服务选择和组合在面向服务的应用中得到了广泛的关注。一般来说，QoS 感知服务组合问题可以分为两类：一是确定 QoS 感知的服务组合，二是不确定 QoS 感知的服务组合。对于确定 QoS 感知服务组合，最流行的方法有整数线性规划（ILP）、混合线性规划（MLP）和人工智能（AI）算法。然而，这些服务组合方法主要针对确定的 QoS 值，而忽略了现实中 QoS 的不确定性。文献［10］考虑到服务质量的不确定性，提出了一种不确定 QoS 感知的服务选择模型，并利用非参数检验和对服务质量的稳定性进行了比较，提出了一种基于简单的目标函数值均值或方差比较的决策模型，并给出了实例分析和原型系统。

云服务组合通常是长期的。文献［11］中考虑服务组合的长期性，从基于用户的角度考虑云服务组合，使用离散贝叶斯网络表示终端用户的经济模型，将云服务组合问题建模为一个影响图问题，并提出了一种基于影响图的云服务组合方法。文献［12］中从长期的角度考虑了服务组合，提出了用时间序列来表示经济模型，然后将云服务组合建模为多维时间序列数据库中的相似搜索问题，并使用 QoS 属性和 QoS 关系两种结构来进一步提高相似性搜索方法的效率。文献［13］中提出了一种基于长期经济模型驱动的云服务选择和组合方法，利用 QoS 历史数据来预测服务提供商的长期 QoS 行为，利用 QoS 属性之间的相关性来降低预测值和真实值之间的误差，将预测的时间序列组表示为经济模型，然后将云服务组合建模为相似搜索问题。文献［14］中使用深度循环长期短期记忆（Long Short Term Memory，LSTM）来预测未来的 QoS 值变化，将预测的 QoS 值推荐给长期服务组合中的组件和替代品。文献［15］考虑基于时间序列的感知 QoS 的云计算服务组合，将服务的 QoS 偏好随时间不断变化的过程纳入云服务组合的研究范围，将云服务组合建模成时间序列的相似度对比问题。

2 问题描述

2.1 相关概念

下面介绍云计算服务组合问题中的一些相关概念，并给出云服务、服务质量、服务组合模型以及组合服务（combined service）的形式化描述。

定义1 云服务。云服务是基于互联网的相关服务的增加、使用和交互模式，通常涉及通过互联网来提供动态易扩展且经常是虚拟化的资源。可用一个二元组表示，$S=\{F,NF\}$，其中 F 和 NF 分别代表服务的功能和非功能属性。

非功能属性用 QoS 表示，所以服务又可以表述为$\{F,Q\}$，其中 Q 代表服务的 QoS 值。

定义2 服务质量。服务质量指服务的非功能属性，可用一个二元组$\{RT,TH\}$表示，其中 RT 和 TH 分别代表响应时间（response time）和吞吐量（throughput）。

1）响应时间 RT。响应时间是从用户提交服务请求到服务执行完成并返回结果的时间间隔，包括服务执行时间 T_{exe}，网络传输时间 T_{trans} 和其他时间花费 T_{oth}：

$$RT = T_{exe} + T_{trans} + T_{oth}$$

2）吞吐量 TH。吞吐量是指服务单位时间内可处理的请求数量。

QoS 属性可以分为两类：正相关和负相关属性。正相关属性意味着属性值越高，质量越好，例如：可靠性、可用性。负相关属性意味着与正相关属性相反。并且由于 QoS 中每个元素的取值有较大差异，数值不是一个数量级，因此在进行服务选择和服务组合之前，需要对 QoS 各属性值进行归一化处理，通常每一个 QoS 属性转换成 0 ~ 1 的值。因为正相关属性值很容易转换成负相关属性值，为了简单，本文只考虑负相关属性值。本文按以下公式进行归一化处理。

1）效益型：

$$q_{ij} = \begin{cases} \dfrac{q_j^{\max} - q_{ij}}{q_j^{\max} - q_j^{\min}}, & q_j^{\max} \neq q_j^{\min} \\ 1, & q_j^{\max} = q_j^{\min} \end{cases} \tag{1}$$

2）成本型：

$$q_{ij} = \begin{cases} \dfrac{q_{ij} - q_j^{\min}}{q_j^{\max} - q_j^{\min}}, & q_j^{\max} \neq q_j^{\min} \\ 1, & q_j^{\max} = q_j^{\min} \end{cases} \tag{2}$$

其中：$q_j^{\max}$ 和 $q_j^{\min}$ 分别代表当前第 j 个 QoS 属性的最大值和最小值，随着云服务的加入或退出，它们的值将发生动态变化。对于正相关的 QoS 参数属性指标采用式（1）效益型计算公式进行转换；相反，对于负相关的 QoS 参数属性指标采用式（2）成本型计算公式进行转换。响应时间 RT 属于负相关属性，故采用式（1）成本型计算公式转换；吞吐量 TH 则采用式（2）效益型计算公式进行转换。

定义3 组合服务（Combined Service，CS）。组合服务是将多个单一服务组合起来构建复杂、功能强大的服务。一个抽象的组合服务可用一个抽象的组合请求表示，$CS=\{S_1,S_2,$

$\cdots, S_n\}$，其中 CS 表示请求服务类。一个具体的组合服务可以定义为抽象组合服务的实例化。每个抽象服务类包含功能相同、QoS 值不同的服务，通过服务选择将 CS 中每个抽象服务类实例化，可以得到一个具体的组合服务。组合服务 QoS 计算方法如表 1。本文采用顺序结构模型讨论 QoS 感知云服务组合问题。

2.2　QoS 模型描述

本文考虑的 QoS 属性为响应时间和吞吐量，分别用 q_1、q_2 表示。不同于传统服务组合问题，组合服务的整体 QoS 性能需要从长期角度衡量。本文采用不定长时间序列描述服务的长期 QoS 值数据变化，时间序列是指将相同统计指标的数值按其发生的先后顺序排列而成的数列。时间序列分析的主要目的是根据已有的历史数据对未来进行预测。现实中，由于云环境的动态变化，服务提供商所提供服务质量可能会发生变化，在不同的时间段内，用户对同一服务具有不同的 QoS 体验。由于一段时间内每个服务被调用次数不同，为更真实地反映 QoS 的变化，将服务不确定 QoS 值表示为 $\boldsymbol{TSG}_{m\times 2}$:

$$\boldsymbol{TSG}_{m\times 2} = \begin{bmatrix} \boldsymbol{Q}_1 \\ \boldsymbol{Q}_2 \\ \vdots \\ \boldsymbol{Q}_m \end{bmatrix} = \begin{bmatrix} q_{1,1} & q_{1,2} \\ q_{2,1} & q_{2,2} \\ \vdots & \vdots \\ q_{m,1} & q_{m,2} \end{bmatrix} \tag{3}$$

服务被调用 m 次后，其不确定 QoS 值可以形式化一个时间序列 $\boldsymbol{TSG}_{m\times 2}$，$\boldsymbol{Q}_i(i = 1,2,\cdots,m)$ 代表某个具体服务在第 i 次被调用时 QoS 的瞬时记录，$q_{ij}(i = 1,2,\cdots,m;\ j = 1,2)$ 代表第 j 个 QoS 指标在第 i 次调用时的记录值。

对于每一个候选应用服务 $CAS_i^j \in CA_i(1 \leqslant i \leqslant n)$，服务的 QoS 属性用时间序列组表示，记为 $\boldsymbol{TSG}_{m\times 2} = (CAS_i^j)$。

组合服务(CS) 的 QoS 属性用链表表示，记为 $List(CS)$。服务组合中的每个节点表示对应抽象服务下所选择的具体服务，即链表中每个节点代表一个具体的候选服务。

表 1　组合服务 QoS 计算方法

Tab. 1　QoS computation formula of four basic composition patterns

模型	QoS 属性	
	响应时间 RT	吞吐量 TP
顺序模型	$\sum_{i=1}^{n} rt_i$	$\min_{i=1}^{n}(th_i)$
选择模型	$\sum_{i=1}^{n}(rt_i \times p_i)$	$\min_{i=1}^{n}(th_i)$
并行模型	$\max_{i=1}^{n}(rt_i)$	$\min_{i=1}^{n}(th_i)$
循环模型	$rt \times \theta$	th

2.3　目标函数

为了评估服务组合解的质量，首先计算每一个参数的均值，记为 $\overline{q} = (\overline{q_1}, \overline{q_2})$:

$$\overline{q_j} = \frac{1}{L_k}\sum_{i=1}^{L_k} \boldsymbol{TSG}(cs)\ [i, j];\ j = 1,2;\ k = 1,2,\cdots,n \tag{4}$$

根据式(1) ~ (2) 对 QoS 值进行归一化处理，表示为 $\overline{\boldsymbol{Q}} = (\overline{\boldsymbol{Q}_1}, \overline{\boldsymbol{Q}_2})$，基于上述描述，服务组合解 cs 定义为:

$$E(cs) = \sum_{i=1}^{2} \omega_i \times \overline{\boldsymbol{Q}_i} \tag{5}$$

其中: ω_i 是参数 q_i 的权重，$\sum_{i=1}^{2}\omega_i = 1$。根据用户约束条件重要性不同，为每一个 QoS 属性设定一个权重，本文中响应时间和吞吐量的权重均取 0.5。

3　算法设计

3.1　基本遗传算法

由于遗传算法(GA) 实现步骤简单，并且具有良好的全局收敛能力、自适应能力和并行能力，目前已经被广泛成功地应用于多个领域。基本遗传算法流程如图 1 所示。

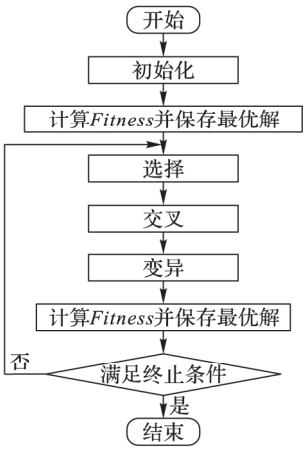


图 1　基本遗传算法流程

Fig. 1　Flow chart of basic genetic algorithm

采用遗传算法求解服务组合问题的优点是规则简单、编码简单，并且具有较强的全局寻优能力。其缺点是: 稳定性差，遗传算法属于随机性算法，需要进行多次运算，结果的可靠性差; 易早熟收敛，即对新空间探索能力有限，易收敛到局部最优解。因此，在求解现实问题时，都需要对遗传算法进行适当改进，提高算法效率。

3.2　算法改进策略

文献[9] 中将精英选择策略应用到遗传算法，以代替标准遗传算法中的轮盘赌选择策略，并指出精英选择策略具有能够防止最优个体的丢失、提高群体收敛速度等优点。但是，相对于轮盘赌选择策略，精英选择策略每次都需要从种群中选择最好的部分个体，该操作将会使得种群多样性丢失，极易导致早熟收敛，从而降低寻优精度。另外，采用精英选择策略，每次选择都需要对种群所有个体进行排序，时间开销过大，造成运行时间很不理想。为此，本文提出一种改进遗传算法 T-GA(Tournament strategy GA)，以提高遗传算法的普适性和寻优性能。

T-GA 中采用锦标赛选择策略代替基本遗传算法中轮盘赌选择策略，每次进行选择时，适应度较好的个体被选择的概率较大。同时，由于它只是使用适应值的相对值作为选择的

标准，而与适应度的数值大小不成直接比例，从而能避免超级个体的影响，在一定程度上既使得种群保持了多样性，防止早熟收敛，又提高了群体收敛速度。

锦标赛选择策略每次从种群中选取出一定数量的个体，然后选择其中最好的一个进入子代种群，重复该操作，直到新的种群规模达到原来的种群规模。具体的操作步骤如下：

1) 确定每次选择的个体数量 R。

2) 从种群中随机选择 R 个个体(每个个体入选的概率相同)，根据每个个体的适应度值，选择其中适应度值最好的个体进入子代种群。

3) 重复步骤 2) N 次，得到的个体构成新一代种群。

3.3 算法描述

实验中，组合服务是一个五元组，即每个 CS 记录包含五个子服务过程。每个子服务作为染色体上的一个基因位，每个 CS 包含五个基因位。每个基因位置属于不同的候选服务集。根据适应度函数计算适应度值，通过迭代，本文算法能够找到最优组合方案。本文算法中种群采用实数编码，具体算法描述如下。

算法 1 T-GA。

```
Initialize s members randomly as population P
Evaluate each member in P
P′ = NULL
while stop criterion not met do
   Select (1 − P_c) * s members from P by tournament and insert into
   P′
   Randomly select p_c* s members from P with no duplicate
   Crossover(P_c)
   Insert the offspring into P′
   Randomly select u* s members from P′
   for each selected member i do
     Mutation(u)
   end for
   P = P′
   Evaluate each member in P
end while
return bestmember
```

其中：s 为初始种群数量，P 为父代种群，P' 为子代种群。在选择操作中采用轮盘赌选择策略从种群中选择 $(1-P_c)*s$ 个父代个体直接遗传到子代，然后，将父代剩余个体进行交叉操作 Crossover(P_c)，采用单点交叉，其中交叉位置随机产生，将交叉得到的个体保存到子代种群 P' 中。然后，从 P' 中随机选择 $u*s$ 数量的个体进行变异操作 Mutation(u)，同样，变异位置通过随机产生。通过选择、交叉和变异操作，本文算法不仅能够使得种群保持了多样性，防止早熟收敛，而且提高了种群收敛速度。

4 实验分析

为验证所提出时间序列模型的可行性，以及本文算法 T-GA的性能，以公共数据集 WSDream[3] 为测试数据，将本文算法与基于精英选择策略的遗传算法(Genetic Algorithm based on Elite selection strategy，E-GA)[3] 进行对比。

4.1 实验数据与环境

为了使实验更具代表性，实验数据采用 Zhang 等的公用数据集 WSDream[3]，包含 142 个用户在 64 个时间间隔对 4 532个 Web 服务的响应时间和吞吐量的真实记录值，现实世界中，大部分用户都只调用过少数的 Web 服务，所以每个服务的真实调用记录并不都包含 64 个时刻记录。实验环境为：64 位 Windows 7 OS，Intel Core i7-4790 CPU @3.60 GHz，4 GB RAM，Visio Studio 2010。

由于数据集较为庞大，需要对数据进行处理，从中选取一部分数据作为实验的真实数据。数据抽取操作步骤如下：

1) 从数据源中抽取 Service User ID 为 0 的数据。

2) 对数据进行降噪处理。

3) 将数据集按 Web Service ID 提取出对应的 QoS 值时间序列，得到用户 0 对 1 610 个 Web Service ID 的调用 QoS 值时间序列。

4) 按顺序选择前 1 500 个 Web Service ID 作为实验数据，将得到的数据集分为 5 组，每组 300 个 Web Service ID 对应的 QoS 值时间序列，每组中的服务提供相同的功能。

4.2 实验结果与分析

下面评估所提出不定长时间序列(Uncertain-Long Time Series，ULST)模型的可行性和所改进的算法的有效性、效率和稳定性。实验参数设置如表 2 所示。

表 2 实验参数设置

Tab. 2 Experimental parameter settings

参数	含义	数值
Sizepop	种群规模	100
T	进化代数	500
P_c	交叉概率	0.8
u	变异概率	0.15

4.2.1 ULST 模型的可行性

为了验证不定长时间序列模型的可行性，采用 4.1 节数据处理所得到的数据，从每个服务对应的时间序列中随机选取一个 QoS 值作为该服务的瞬时访问记录，将得到的数据分为 5 组，每组 300 个 Web Service ID 对应的 QoS 值瞬时访问记录，每组中的服务提供相同的功能。

图 2 中：Instant1 和 Instant2 分别代表两次随机操作的结果，即瞬时 QoS 性能。从图 2 可以看出，本文方法较瞬时访问

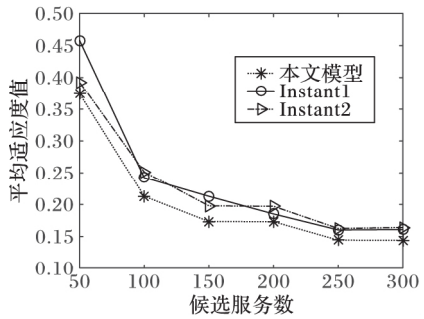


图 2 ULST 模型可行性

Fig. 2 The feasibility of ULST model

结果有所提高。

4.2.2　算法有效性分析

为了进行算法的有效性对比,在如下实验参数设置下进行 100 次重复实验,取其均值为最终衡量指标,实验结果如表 3所示。其中: 种群规模为 100,迭代次数为 500,变异概率为 0.15,交叉概率为 0.8,QoS 的维度为 2,任务数量为 5,每个任务的候选服务数以 50 为步长从 50 递增到 300。

从表 3 可以看出,随着候选服务数量的增加,本文算法获得的最优解优于 E-GA[3],表明本文算法的寻优精度高。候选服务数在 50 和 100 时,由于候选服务数量少,两种算法都能够收敛到最优解; 但随着候选服务数量继续增大,本文算法在收敛精度上具有明显的优势。

表 3　不同候选服务数据集大小的适应度值对比

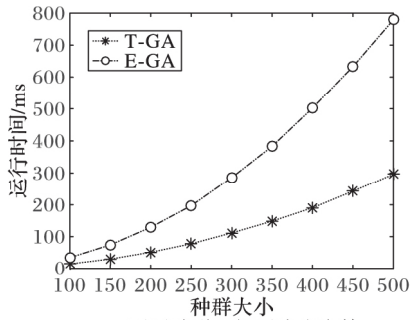
Tab. 3　Fitness value comparison with different candidate service dataset sizes

候选服务数据集大小	E-GA	T-GA
50	0.3740	0.3740
100	0.2129	0.2129
150	0.1748	0.1733
200	0.1736	0.1731
250	0.1464	0.1443
300	0.1462	0.1441

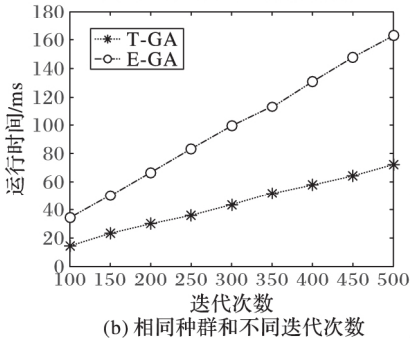
4.2.3　算法的时间效率分析

为进一步分析 T-GA 和 E-GA 的时间开销,本文分别进行了两组对比实验:

1) 具有不同的种群大小和相同的迭代次数(100) 进行优化 100 次的平均运行时间(如图 3(a) 所示) ;



(a) 不同种群和相同迭代次数



(b) 相同种群和不同迭代次数

图 3　算法在不同实际条件下的平均运行时间对比

Fig. 3　Average running time comparison of algorithms under different actual conditions

2) 具有相同的种群大小(100) 和不同的迭代次数进行优化 100 次的平均运行时间(如图 3(b) 所示) 。

从图 3(a) 可以看出,相同迭代次数情况下,随着种群规模的增大,算法运行时间也在增加,且算法的运行时间与种群规模大小呈线性关系; 同样,如图 3(b) ,种群大小相同,随着迭代次数的增加,算法运行时间与种群迭代次数呈线性增加。因此,在时间开销方面,本文算法具有 E-GA 无法替代的优势。

4.2.4　算法稳定性分析

为验证 T-GA 寻优结果的稳定性,采用均方差(Root Mean Square Error, RMSE) 验证不同算法的稳定性。*RMSE* 定义如下:

$$RMSE = \frac{1}{n}\sqrt{\sum_{i=1}^{n}(x_i - x)^2} \tag{6}$$

其中: x_i 表示第 i 次实验结果; x 表示 n 次实验的平均结果。

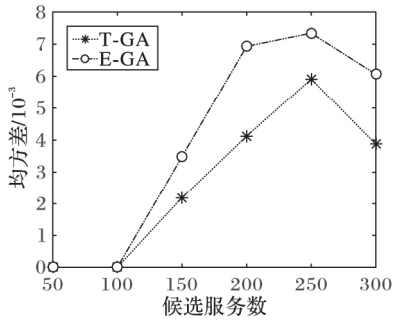


图 4　不同算法的均方差对比

Fig. 4　RMSE comparison of different algorithms

从图 4 可以看出,随着候选服务数的增加,算法的均方差值也在逐渐上升,即稳定性在降低,候选服务数越多,结果越不稳定。实验结果可看出: E-GA 寻优结果的波动性较大; 而 T-GA 寻优结果的均方差低于 E-GA 算法,表明本文算法具有更好的稳定性。

4.2.5　变异概率 u 的影响

遗传算法中,变异是模拟生物在自然环境中由于各种偶然因素引起的基因突变过程。变异概率 u 是遗传算法中的重要参数,变异运算是种群产生新个体的辅助方法,决定了遗传算法的局部搜索能力,同时保持种群多样性,避免进化停滞,出现早熟收敛,其概率分布在 0.1 ~ 1,变异体现了全局最优解的全局覆盖。在遗传算法中,为测试变异概率对算法的影响,设置候选服务规模为 300,u 的变化为 0.1 ~ 0.9,步长为 0.1。运行 100 次的平均适应度值的统计结果如图 5 所示。

从图 5 可以看出,在相同候选服务规模下,E-GA 寻优结果随 u 的不同而变化; T-GA 在不同参数设置下,即使在相应参数变化足够大时,寻优结果仍趋于不变,说明本文算法具有很好的稳定性。

以上实验结果表明: 本文提出的 ULST 模型能够有效地解决不确定 QoS 感知的云服务组合问题; 并且,结合所提出的改进遗传算法 T-GA 的有效性、效率和稳定性来看,T-GA

能够有效地解决不确定性云服务组合优化问题。

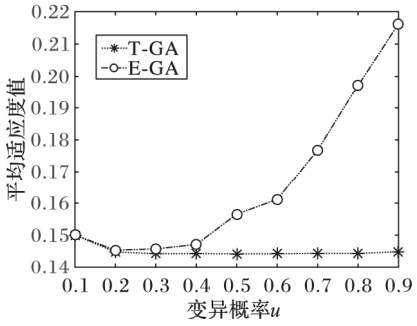


图5 变异概率对平均适应度值的影响

Fig. 5 Influence of mutation probability on average fitness value

5 结语

本文研究基于不确定 QoS 感知的云服务组合问题，首先，提出了一种基于不确定性 QoS 感知的云服务组合时间序列模型，将服务的 QoS 值随时间变化过程建模为不定长时间序列，该模型能够准确地描述用户对服务访问的真实记录；其次，采用一种改进的遗传算法找寻最优解，该算法通过引入轮盘赌选择策略改进基本遗传算法选择策略，使得算法能有效地避免早熟收敛，较好地提高算法的全局收敛能力和搜索效率。实验结果表明了本文方法的可行性和有效性。在今后的工作中，将研究跨界服务不确定 QoS 问题，寻找一种解决复杂不确定性跨界服务系统的优化均衡方法。

参考文献(References)

[1] ARDAGNA D, PERNICI B. Adaptive service composition in flexible processes [J]. IEEE Transactions on Software Engineering, 2007, 33(6): 369 – 384.

[2] HUANG A F M, LAN C W, YANG S J H. An optimal QoS-based Web service selection scheme[J]. Information Sciences, 2009, 179(19): 3309 – 3322.

[3] ZHANG Y W, CUI G M, ZHAO S, et al. IFOA4WSC: a quick and effective algorithm for QoS-aware service composition[J]. International Journal of Web and Grid Services, 2016, 12(1): 81 – 108.

[4] 倪志伟, 方清华, 李蓉蓉, 等. 改进蚁群算法在基于服务质量的 Web 服务组合优化中的应用[J]. 计算机应用, 2015, 35(8): 2238 – 2243. (NI Z W, FANG Q H, LI R R, et al. Improved ant colony optimization for QoS-based Web service composition optimization [J]. Journal of Computer Applications, 2015, 35(8): 2238 – 2243.)

[5] CREMENR M, SUCIU M, PALLEZ D, et al. Comparative analysis of multi-objective evolutionary algorithms for QoS-aware Web service composition[J]. Applied Soft Computing, 2016, 39: 124 – 139.

[6] WAGNER F, KLEIN A, KLOPPER B, et al. Multi-objective service composition with time and input-dependent QoS [C]// Proceedings of the 2012 IEEE 19th International Conference on Web Services. Piscataway, NJ: IEEE, 2012: 234 – 241.

[7] YILMAZ A, KARAGOZ P. Improved genetic algorithm based approach for QoS aware Web service composition [C]// Proceedings of the 2014 IEEE International Conference on Web Services. Piscataway, NJ: IEEE, 2014: 463 – 470.

[8] NIU S, ZOU G, GAN Y, et al. Towards uncertain QoS-aware service composition via multi-objective optimization [C]// Proceedings of the 2017 IEEE International Conference on Web Services. Piscataway, NJ: IEEE, 2017: 894 – 897.

[9] LIU S C, WEI Y, TANG K, et al. QoS-aware long-term based service composition in cloud computing [C]// Proceedings of the 2015 IEEE Congress on Evolutionary Computation. Piscataway, NJ: IEEE, 2015: 3362 – 3369.

[10] WAN C, WANG H. Uncertainty-aware QoS description and selection model for Web services [C]// SCC 2007: Proceedings of the 2007 IEEE International Conference on Services Computing. Piscataway, NJ: IEEE, 2007: 154 – 161.

[11] YE Z, BOUGUETTAYA A, ZHOU X. QoS-aware cloud service composition based on economic models [C]// ICSOC 2012: Proceedings of the 2012 International Conference on Service-Oriented Computing. Berlin: Springer, 2012: 111 – 126.

[12] YE Z, BOUGUETTAYA A, ZHOU X. QoS-aware cloud service composition using time series [C]// ICSOC 2013: Proceedings of the 2013 International Conference on Service-Oriented Computing. Berlin: Springer, 2013: 9 – 22.

[13] YE Z, MISTRY S, BOUGUETTAYA A, et al. Long-term QoS-aware cloud service composition using multivariate time series analysis[J]. IEEE Transactions on Services Computing, 2016, 9(3): 382 – 393.

[14] LABBACI H, MEDJAHED B, AKLOUF Y. A deep learning approach for long term QoS-compliant service composition [C]// ICSOC 2017: Proceedings of the 2017 International Conference on Service-Oriented Computing. Berlin: Springer, 2017: 287 – 294.

[15] 肖文娟, 段玉聪. 基于时间序列的感知 QoS 的云服务组合[J]. 计算机工程与科学, 2014, 36(11): 2061 – 2066. (XIAO W J, DUAN Y C. Time series-based QoS aware cloud service composition[J]. Computer Engineering & Science, 2014, 36(11): 2061 – 2066).

This work is partially supported by the National Natural Science Foundation of China (61872002), the National Key Technology R&D Program of China (2015BAK24B01), the Natural Science Foundation of Anhui Province (1808085MF197), the Science and Technology Major of the Ministry Project of Anhui Province (16030901062).

WANG Sichen, born in 1991. M. S. candidate. His research interests include service computing, evolutionary computing.

TU Hui, born in 1984. M. S. candidate. His research interests include big data, cloud computing.

ZHANG Yiwen, born in 1976, Ph. D., associate professor. His research interests include service computing, big data, cloud computing.