

S-ABC——面向服务领域的人工蜂群算法范型

徐晓飞 刘志中 王忠杰 闵寻优 刘睿霖 王海芳

(哈尔滨工业大学计算机科学与技术学院 哈尔滨 150006)

摘 要 服务计算优化问题(如服务选择、服务组合、服务资源调度等)随着云计算、物联网、大数据的快速发展而变得日益复杂.另一方面,各服务行业在其长期演化中逐渐形成了特有的领域特性(如服务先验性、关联性、相似性等).这些特性对服务优化问题求解有重要影响,如果对其考虑不充分,将导致服务优化问题求解的效率与效果不理想.因此,如何构建面向服务领域的服务优化算法范型及高效求解算法与优化策略成为亟待解决的关键问题.文中分析了服务领域特性对服务优化问题求解的影响规律,据此改进了人工蜂群算法的优化策略,提出了面向服务领域的人工蜂群算法范型(Service domain-oriented Artificial Bee Colony algorithm paradigm, S-ABC),阐述了 S-ABC 范型的优化机理,并通过验证实验证实了该算法范型的优化效果.该研究工作为利用服务领域特性指导服务优化问题求解算法的设计提供了新的研究思路和方法,深化了群体智能算法在服务领域的应用,扩展了群体智能算法的优化理论.

关键词 服务领域特性;人工蜂群算法;算法范型;服务选择;服务组合;服务资源调度

中图法分类号 TP393 **DOI号** 10.11897/SP.J.1016.2015.02301

S-ABC—Service Domain-Oriented Artificial Bee Colony Algorithm Paradigm

XU Xiao-Fei LIU Zhi-Zhong WANG Zhong-Jie MIN Xun-You LIU Rui-Lin WANG Hai-Fang

(School of Computer Science and Technology, Harbin Institute of Technology, Harbin 150006)

Abstract In the service computing field, the typical service optimization problems (such as service selection, service composition and service resource scheduling) become more and more complicated with the rapid development of cloud computing, internet of things and big data. Meanwhile, in many service sectors, the service domain features (such as service priori, correlation and similarity) have been gradually formed with the long-term evolution of service business. These domain features have strong influences on solutions of service optimization problems. If the service domain features are not considered adequately, the service optimization problems can not be solved effectively and efficiently. Therefore, how to design service domain-oriented optimization algorithm paradigm, efficient algorithms and optimization strategies become the critical challenges. This paper analyzes the influences on service optimization problems by service domain features. Then, based on the improved strategies of artificial bee colony algorithm, a service domain-oriented artificial bee colony algorithm paradigm (S-ABC) is presented. The optimization principle of S-ABC paradigm is described in detail. The better optimization results are verified by means of confirmatory experiment. This research work shows a new and better method for solving service optimization problems

收稿日期:2014-12-23;最终修改稿收到日期:2015-05-27.本课题得到国家自然科学基金(61033005,61472106,61272187,61300124)资助.
徐晓飞,男,1962年生,博士,教授,博士生导师,主要研究领域为服务计算与软件服务工程、企业计算、数据挖掘与商务智能等. E-mail: xiaofei@hit.edu.cn. 刘志中,男,1981年生,博士,讲师,主要研究方向为服务计算、群体智能算法. 王忠杰,男,1978年生,博士,教授,博士生导师,主要研究领域为服务计算、服务工程、软件架构及其演化. 闵寻优,女,1988年生,博士研究生,主要研究方向为服务计算. 刘睿霖,男,1989年生,博士研究生,主要研究方向为服务计算. 王海芳,女,硕士研究生,主要研究方向为服务计算.

with the support of service domain features, and to extend the theory of swarm intelligence optimization in service computing field.

Keywords service domain features; artificial bee colony algorithm; algorithm paradigm; service selection; service composition; service resource scheduling

1 引 言

随着服务计算、云计算、物联网、大数据等高新技术的快速发展与成熟,各个服务行业与领域涌现出大规模的智能服务.海量的服务交织成跨网、跨域、跨世界的服务生态系统,为广大用户提供了丰富的服务资源.越来越多的用户更加依赖于网络服务完成各种业务.而服务的分布性、聚集性与跨域性、服务资源的动态性以及用户需求的多变性等复杂因素使得服务应用日益繁琐,也使得服务系统的优化设计与高效运行更加复杂.如何在复杂动态的云计算及网络服务环境下,从跨网、跨域的海量服务中,高效地完成服务选择、服务组合、服务资源调度,构建出满足大规模个性化用户需求的最优服务方案,从而实现服务增值、用户满意的最终目标,是服务系统优化设计与高效运行面临的关键科学问题之一.

在各个服务行业领域中,随着服务应用与服务生态系统的发展演化,逐步形成了服务行业特有的领域特性(如先验性、关联性、相似性等).服务领域特性是面向服务领域的服务生态系统中存在的客观规律,对服务优化问题的求解有着重要影响.因此,在求解服务优化问题时,应深入挖掘服务领域特性对服务优化问题求解的影响规律,以此指导服务优化问题的求解,从而进一步提高服务优化问题求解的效率与效果.

当前,国内外学者针对服务选择、服务组合、服务资源调度等服务领域中重要的优化问题展开了大量的研究,提出了基于整数线性规划^[1]、混合线性规划^[2]、启发式算法^[3-4]、遗传算法^[5-6]、蚁群算法^[7-8]、粒子群算法^[9-10]、人工蜂群算法^[11-13]等多种求解方法.已有的研究工作取得了丰富的研究成果,促进了服务计算技术的快速发展与广泛应用.

然而,已有的服务优化问题研究工作通常将服务优化问题抽象成数学问题进行求解,只关注问题求解算法的设计与改进,而没有考虑服务领域特性(先验性、关联性、相似性等)对服务优化问题求解过程的影响,所设计的求解算法往往缺乏对服务领域

特性的感知,导致服务优化问题的理论研究与服务实际应用脱节,使得服务优化问题求解的效率和效果不够理想.因此,如何应用服务领域特性对服务优化问题求解的影响规律指导服务优化问题求解算法的设计,提出面向服务领域的优化算法范型及优化策略,为服务优化问题的求解提供算法模型及优化技术指导,进而提高服务优化问题求解的效率与效果,成为服务计算领域亟需解决的关键问题之一.

在服务优化问题求解方法中,群体智能算法凭借其自适应、自组织、协同性、强鲁棒性和良好的分布性并行性等优点,在求解服务优化问题时表现较为出色^[14],已经成为求解服务领域典型优化问题的主要方法.人工蜂群算法(Artificial Bee Colony algorithm, ABC)是由 Karaboga^[15]于 2005 年提出的新型群智能算法,并且在求解优化问题时表现较为出色^[15-17].

目前,已有学者将 ABC 算法用于服务优化问题的求解^[11-13].ABC 算法的优化机理与服务计算中服务优化问题的求解存在很强的映射关系.在 ABC 算法中,蜂群搜索最优蜜源的过程对应于寻找服务优化问题最优解的过程,食物源对应于服务优化问题的可行解,适应度函数对应于服务优化问题的目标函数,雇佣蜂阶段与观察蜂阶段对应于服务优化问题解空间的邻域搜索,侦查蜂阶段对应于服务优化问题解空间的全局搜索,算法的终止条件对应于服务优化问题求解结束条件.

鉴于服务优化问题求解中存在的问题以及人工蜂群算法的优越性,本文选择人工蜂群算法作为算法模型基础,在深入分析人工蜂群算法优化机理与挖掘服务领域特性对服务优化问题求解影响规律的基础上,设计了面向服务领域的人工蜂群算法范型(Service domain-oriented Artificial Bee Colony algorithm paradigm, S-ABC),为服务优化问题求解提供了算法模型与优化策略支持.最后,以服务组合问题为例,验证了 S-ABC 范型的有效性与优越性.该研究工作深化了人工蜂群算法在服务领域的应用,扩展了群体智能算法的优化理论.

本文第 2 节介绍服务领域优化问题的相关研

究工作;第3节阐述服务领域特性;第4节给出S-ABC范型设计分析;第5节给出服务空间的建立与表示;第6节首先详细阐述S-ABC范型中的关键优化策略,之后给出S-ABC范型的详细描述;第7节以服务组合问题为例,验证S-ABC范型的有效性。最后,对本文工作进行总结并对后期研究工作进行展望。

2 相关工作

近年来,国内外学者针对服务计算领域中的服务选择、服务组合、服务资源调度等优化问题开展了大量的研究,并取得了丰富的研究成果。然而,在服务行业长期的发展演化中,逐渐形成了特有的领域特性(先验性、关联性与相似性),这些领域特性对于提升服务优化问题的求解效果具有重要的“启发与指导”作用。当前,已经存在一些应用服务领域特性指导服务优化问题求解的研究工作。

文献[18]提出了一种基于组合服务执行信息的服务选取方法,该方法基于组合服务历史执行信息、路径分支关联规则和服务执行顺序序列模式,挖掘出知识模式来协助服务选择。张明卫等人^[19]提出了一种基于划分的组合服务选取方法,该方法首先基于日志记载的质量信息选取那些性能优良的组合服务执行实例,在此之上发现被频繁一起使用的具体服务集合,据此产生对组合服务的划分,形成组合服务点和对应的点模式,把点模式集作为点的备选服务集,以点为单位进行组合服务选取。文献[20]通过挖掘Web服务的历史使用信息,获取用户对Web服务的兴趣与QoS偏好,之后提出了一种基于用户QoS偏好与用户对服务兴趣点的服务推荐方法。文献[21]提出了一种基于用户使用服务的历史信息的服务推荐方法,历史信息包括功能需求信息与QoS信息。

上述研究工作^[18-21]通过挖掘组合服务或服务的历史使用信息来指导服务组合与服务推荐。本文提出的先验性与上述研究工作思想类似,它们之间的相同点是通过挖掘服务的历史使用信息获取有用的经验知识指导服务优化问题的求解;不同点在于:已有研究工作仅利用历史使用信息解决具体的服务组合或服务选择问题,没有分析挖掘先验性对服务优化问题求解的影响规律以及利用这些知识指导服务优化问题求解算法模型及优化策略的设计;此外,本文应用先验性对候选服务空间进行了划分,缩小了

算法的搜索空间;在先验性服务空间内构建了服务之间的偏序关系,并以此为指导,改进了人工蜂群算法中的搜索策略。

文献[22]在进行服务组合时,考虑了服务之间组合关联关系、商业实体关联关系以及统计关联关系,并给出了支持这3种关联关系的QoS度量模型。文献[23]提出了一种支持服务关联关系的QoS描述模型,刻画了可选服务的服务质量对其他可选服务的依赖关系,在此基础上,提出了支持服务关联的组合服务选择方法。文献[24]为了提高制造云服务的服务质量,重点对服务之间的QoS统计关联进行了研究,提出了支持服务统计关联的服务组合优化方法。文献[25]给出了服务质量关联与选择关联的形式化描述,提出了一种关联关系驱动、QoS感知的服务选择方法。

研究工作^[22-25]通过挖掘服务之间的关联关系来提高服务组合的效果与效率。本文中所提到的“关联性”与上述研究工作中的“关联性”的定义与用途相同,即在求解服务优化问题时要充分考虑服务之间的关联关系,提高问题的求解效果。

文献[26]将服务组合问题转换成查找、选择与服务组合规范最接近的组合服务问题,该研究工作应用相似度衡量给定的服务组合与具体描述之间的接近程度,并提出了一系列用于Web服务组合模式的相似度度量公式。文献[27]提出了服务流程相似性定义,并利用服务流程的相似性协助解决服务组合中的服务替换问题,取得了较好的效果。文献[28]将服务流程中可利用的逻辑归纳定义为可在用户之间共享的模式,以方便新组合服务的构建。文献[29]对服务流程中重复出现的过程片段进行提取,期望在服务选择中重用。

文献[26-29]利用服务之间功能与性能的相似性,提高了服务推荐与组合的效果,与其不同的是,本文利用服务的相似性对候选服务空间进行了划分,缩小了算法的搜索范围;在相似服务空间内构建了服务之间的偏序关系,有助于提高算法的搜索效率。

通过分析可以得出,当前已有的研究工作仅采用单一的服务领域特性来改善服务优化问题的求解效果^[30-32],而综合应用3个服务领域特性设计求解服务优化问题算法模型的研究工作还很鲜见,鉴于服务领域特性对服务优化问题求解的重要性,以及当前研究存在的空白,本文在挖掘服务领域特性对服务优化问题求解过程影响规律的基础上,改进了

ABC 算法中的关键优化策略,提出了面向服务领域的人工蜂群算法范型,以期为服务优化问题的求解提供高效的算法模型与优化策略。

3 服务领域的重要特性

在服务领域中,服务生态系统在其长期的发展演化过程中,形成了特有的服务领域特性,主要包括:先验性、关联性与相似性。

(1) 先验性(Priori)

先验性是指在长期的服务应用中产生的经验知识,即针对某一典型的服务应用需求类 D ,通常存在一些使用频率较高且用户比较满意的服务(或服务方案),这一特性称为服务领域的先验性。

服务领域中某服务(或服务方案) s 与服务应用需求类 D 的先验性关系 $Pris\langle s, D \rangle$ 的定义如式(1)所示:

$$Pris\langle s, D \rangle \Leftrightarrow (PF(s) \geq P \wedge US(s) \geq E) \quad (1)$$

其中, $PF(s)$ 为服务的先验性度量函数, P 为先验性阈值; $US(s)$ 为用户满意度函数, E 为用户满意度阈值。

(2) 关联性(Correlation)

关联性是指服务与服务(服务与资源)之间在业务协作中的关联关系。设 s_i, s_j 为服务领域内的两个服务, s_i 与 s_j 之间的关联关系 $Cor\langle s_i, s_j \rangle$ 的定义如式(2)所示。

$$Cor\langle s_i, s_j \rangle \Leftrightarrow p(s_i, s_j) \geq P \quad (2)$$

其中, $p(s_i, s_j)$ 表示服务 s_i 与 s_j 的协作概率; P 为服务 s_i 与 s_j 的协作概率阈值。服务之间的关联关系通常包括业务关联、统计关联、服务质量关联等。

(3) 相似性(Similarity)

服务领域的相关服务中存在大量功能相同、服务质量相似的服务,该特性被称为服务领域的相似性。设 s_i 与 s_j 是服务领域内的两个服务,它们之间的相似关系 $Sim\langle s_i, s_j \rangle$ 可以通过式(3)来确定。

$$Sim\langle s_i, s_j \rangle \Leftrightarrow (Func(s_i) = Func(s_j)) \wedge Dif_{QoS}(s_i, s_j) \leq Q \quad (3)$$

其中, $Func(s_i)$ 表示服务 s_i 的功能集; $Dif_{QoS}(s_i, s_j)$ 表示服务 s_i 与 s_j 之间服务质量的差异性函数, Q 表示服务质量差异性阈值。

4 S-ABC 范型的设计分析

S-ABC 范型的设计思想是在遵循 ABC 算法优

化机理与计算流程的基础上,充分利用服务领域特性对服务优化问题求解过程的影响规律,改进 ABC 算法中的关键优化策略,从而形成面向服务领域的人工蜂群算法范型。下面首先分析人工蜂群算法的优化机理,之后给出服务领域重要特性对服务优化问题求解过程的影响。

4.1 ABC 算法优化机理分析

在 ABC 算法中,人工蜂群分为雇佣蜂、观察蜂和侦查蜂,优化问题的解及相应的函数值抽象为蜜源的位置和花蜜的质量。蜂群寻找最优蜜源的过程可描述为:雇佣蜂发现蜜源并记忆,在各蜜源附近搜索新蜜源,根据前后蜜源的花蜜质量选择较优蜜源并标记;雇佣蜂释放与标记蜜源质量成正比的信息,用来招募观察蜂,观察蜂在某种机制下选取合适的标记蜜源并在其附近搜索新蜜源,与标记蜜源进行比较,选取较优异的蜜源作为本次循环的最终标记蜜源,反复循环寻找最佳蜜源。但如果在采蜜过程中,蜜源经若干次搜索后质量没有改变,相应的雇佣蜂变成侦查蜂,随机搜索新蜜源取代原来的蜜源。

在 ABC 算法中,雇佣蜂、观察蜂通过邻域搜索对蜜源进行开采,侦查蜂通过发现新的蜜源以避免蜜源种类过少。人工蜂群算法通过角色的分工、转换、协作与交流,很好地结合了开采和探索两种搜索方式,实现了全局搜索和局部搜索之间的平衡。已有研究表明,ABC 算法概念简单,控制参数少,易于实现,收敛性较好,且比粒子群算法、蚁群算法、遗传算法、差分算法等优化技术在解决某些复杂问题上更有效^[33-39]。因此,将人工蜂群算法作为设计面向服务领域优化算法范型基础具有一定的优势。

4.2 服务领域特性对服务优化问题求解的影响

服务领域特性是服务应用中存在的客观规律,对服务优化问题的求解具有重要的指导作用。下面分别分析领域特性对服务优化问题求解的影响。

(1) 先验性对服务优化问题求解的影响

首先,先验性对于服务优化问题初始解的生成具有重要的指导作用,基于领域先验知识可以生成一些质量较高的初始解,从而提高算法的收敛速度;其次,根据先验性可以划分候选服务空间,缩小算法的搜索范围,从而提高算法的求解效率;此外,根据先验性可以设计启发式的邻域搜索策略,避免随机盲目搜索。

(2) 关联性对服务优化问题求解的影响

服务领域的关联性对于初始食物源生成、适应

度函数以及邻域搜索具有一定影响。在食物源的生成、适应度函数设计以及蜂群邻域搜索中考虑服务间的关联关系,可以使问题的求解更加贴近服务应用的实际,改善问题的求解效果。

(3) 相似性对问题求解的影响

利用服务的相似性可以找出先验性服务的相似服务,建立具有较大潜在应用价值的候选服务集,有助于缩小算法的搜索空间;基于相似性,可以设计启发式的邻域搜索策略,进而提高问题的求解效率。

利用服务领域的特性对服务优化问题的解空间进行预处理,能够为设计启发式的邻域搜索与全局搜索策略提供知识准备,下面给出面向服务领域的服务空间的建立与表示。

5 面向领域的服务空间建立与表示

在服务应用中,通常可以根据服务使用的历史数据对用户请求与候选服务空间进行预处理。在服

务系统中,通常存在与用户服务请求相对应的服务流程,而根据用户服务请求偏好可以将功能相同的用户请求分成若干类;针对每一类请求 D ,根据服务使用的历史数据,可以挖掘出该类需求使用频率较高的服务方案(先验性服务方案, *Priori Service Schemes, PriSS*)以及服务流程中每个任务经常使用的服务(先验性服务)。

由于网络上存在海量的可用服务,服务流程中每个任务节点具有一个庞大的候选服务集,根据先验性与相似性,可以将每个任务的候选服务集划分成先验性服务集 $PriS$ (*Priori Service Set*)、相似性服务集 $SimS$ (*Similar Service Set*)与一般服务集 $GenS$ (*General Service Set*),具体操作过程如图 1 所示。其中, $D_i (1 \leq i \leq k)$ 表示不同的用户请求类; *Service Process* 表示与每一类用户请求相对应的服务流程; *Priori Service Scheme* 表示先验性服务方案集; $PriS$ 、 $SimS$ 与 $GenS$ 分别表示任务节点的先验性服务集、相似性服务集与一般性服务集。

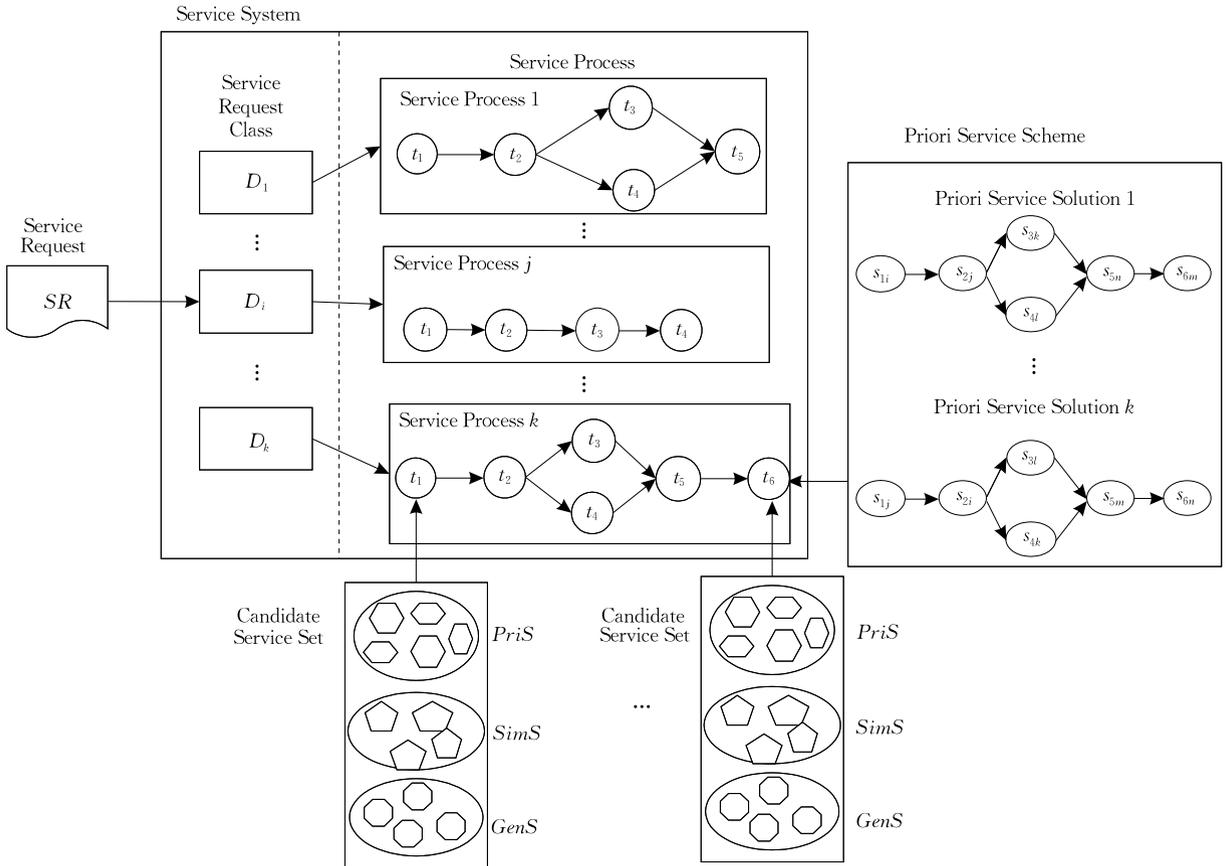


图 1 服务需求处理以及服务空间划分示意图

本文假设每一个服务流程的先验性服务方案集、先验性服务集、相似性服务及与一般性服务集已经存在,在进行服务优化问题求解时,可以直接导入

使用。服务空间的划分为设计具有领域知识指导的优化策略提供知识准备。下面给出面向领域的服务空间的建立与表示。

5.1 服务空间的建立与表示

设某一典型的服务应用需求类 D , 其对应的服务流程中某一任务 T_i 的候选服务集为 S_i , 则任务 T_i 的先验性服务集、相似性服务集与一般性服务集的定义如下.

定义 1. 先验性服务集 $PriS$ (Priori Service Set).

$$PriS(T_i, D) =$$

$$\{s_{ij} | s_{ij} \in S_i \wedge PF(s_{ij}) \geq P \wedge (US(s_{ij}) \geq E)\} \quad (4)$$

其中, s_{ij} 为任务 T_i 候选服务集 S_i 中的一个服务, S_i 中满足式(4)中约束条件的服务所构成的集合称为关于服务应用需求类 D 、任务 T_i 的先验性服务集.

定义 2. 相似性服务集 $SimS$ (Similar Service Set).

$$SimS(T_i, D) = \{s_{ip} | \forall s_{iq} \in PriS(T_i, D), s_{ip} \in S_i \wedge s_{iq} \notin PriS(T_i, D) \wedge Dif_{QoS}(s_{ip}, s_{iq}) \leq Q\} \quad (5)$$

相似性服务集是指在任务 T_i 的候选服务集 S_i 中, 与先验性服务集中的服务具有相似服务质量的服务构成的集合. $Dif_{QoS}(s_i, s_j)$ 表示服务 s_i 与 s_j 之间服务质量的差异性函数, Q 表示服务质量差异性阈值.

定义 3. 一般性服务集 $GenS$ (General Service Set).

$$GenS(T_i, D) = S_i - PriS(T_i, D) - SimS(T_i, D) \quad (6)$$

任务 T_i 的一般性服务集为其候选服务集 S_i 中除去先验性服务与相似性服务后剩余服务所构成的服务集合.

5.2 服务之间的偏序关系

在先验性服务集与相似性服务集内, 服务之间存在一些潜在的偏序关系, 通过建立服务之间的偏序关系, 为设计启发式的邻域搜索策略提供领域知识指导, 从而提高算法的效率.

(1) 先验性服务集内的偏序关系

在先验性服务集内, 以服务的先验性与用户满意度的拟合值作为依据, 建立服务之间的偏序关系. 其中, 拟合值的计算公式如式(7)所示.

$$O(s_{ij}) = w'_1 PF(s_{ij}) + w'_2 US(s_{ij}) \quad (7)$$

其中, s_{ij} 为先验性服务集内的服务, w'_1 与 w'_2 分别为服务的先验性与用户满意度的权重.

(2) 相似性服务集内的偏序关系

为了建立相似性服务集内的偏序关系, 首先在先验性服务集内找出具有最优先验性值的服务, 称之为最佳先验性服务; 之后, 计算相似服务集内的服务与最佳先验性服务之间的服务质量相似度, 依据相似度建立服务之间的偏序关系.

6 S-ABC 范型

基于服务领域特性对服务优化问题求解的影响规律, 本文对人工蜂群算法中的关键优化策略进行了改进, 形成了面向服务领域的人工蜂群算法范型 (S-ABC), 主要包括面向服务领域的 ABC 算法框架、服务空间搜索策略、领域特性感知的初始食物源生成策略、面向服务领域的适应度函数、领域特性驱动的雇佣蜂阶段、观察蜂阶段与侦查蜂阶段、面向服务领域的算法判优准则. S-ABC 范型架构如图 2 所示.

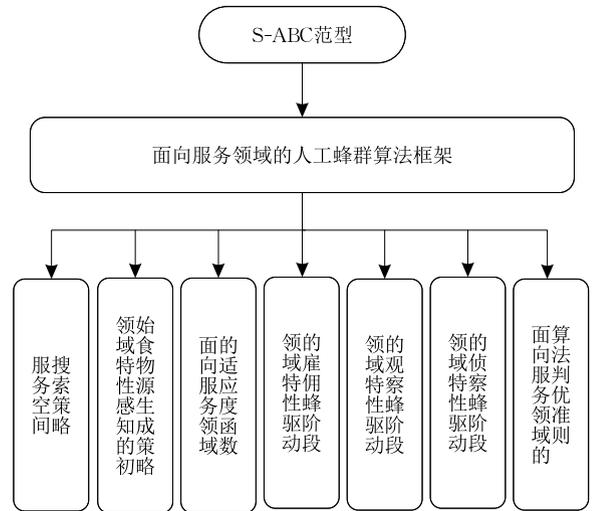


图 2 S-ABC 范型架构

6.1 服务空间搜索策略

在服务领域, 服务优化问题先验性服务集的丰富程度决定了服务空间的搜索策略. 当服务优化问题的先验性服务集比较丰富时, 优先在先验性服务集内搜索则能以较大的概率找出满足用户需求的解. 在这种情况下, 算法按照一定的规则依次在先验性服务集、相似性服务集与一般性服务集内进行搜索, 能够提高问题的求解效率, 本文把这种搜索策略称为服务空间优先级搜索策略, 记为 \mathbb{P} .

当服务优化问题的先验性服务集不丰富时, 在先验性服务集内找到满足用户需求解的概率较小, 而同时在 3 个服务空间内开展搜索找到满足需求解的概率较大. 在这种情况下, 算法在 3 个空间内同时展开搜索, 本文把这种搜索策略称为服务空间均衡搜索策略, 记为 \mathbb{E} .

在求解服务优化问题时, 应首先根据服务优化问题先验性服务集的丰富程度确定服务空间的搜索策略, 之后, 选择与服务空间搜索策略对应的优化操

作. 服务空间搜索策略的确定规则如 Rule(1) 所示.

$$\begin{cases} \text{If}(\|PriS\| \geq CP) \text{ Then Search Strategy} \leftarrow \mathbb{P} \\ \text{If}(\|PriS\| < CP) \text{ Then Search Strategy} \leftarrow \mathbb{E} \end{cases} \quad \text{Rule(1)}$$

其中, $\|PriS\| = \frac{|PriS|}{|S|}$ 表示服务优化问题先验性服务集的丰富度, $|PriS|$ 表示先验性服务集的基数; $|S|$ 表示服务优化问题候选服务集的基数; CP 为先验性服务集的丰富度阈值.

6.2 领域特性感知的初始食物源生成策略

S-ABC 范型中食物源的编码方式如式(8)所示:

$$X_t = \langle s_{1_t}, \dots, s_{i_t}, \dots, s_{n_t} \rangle \quad (8)$$

其中, X_t 表示一个由 n 维向量构成的食物源, 其中, s_{i_t} 为食物源 X_t 中第 i 维变量的取值, 亦为第 i 维变量对应的候选服务集中的第 j 个候选服务; 这种食物源编码方式能够适应服务优化问题离散性, 可以用来表示多种优化问题的解.

优化问题初始解的质量对于群体智能算法的收敛速度及求解质量有着重要影响. 通常地, 在没有任何先验知识的情况下, 采用随机的方法生成初始解; 当存在可用的领域知识时, 则采用启发式方法生成高质量的初始解, 从而提高算法的求解效率. 本文针对 S-ABC 范型中两种服务空间搜索策略, 分别设计了相应的初始食物源生成策略, 即基于先验性的初始食物源生成策略与初始食物源均衡生成策略.

(1) 基于先验性的初始食物源生成策略

设初始食物源的规模为 SN , 基于先验性生成的初始食物源由两部分组成, 一部分是基于先验性服务方案集生成的初始食物源, 设该部分食物源占总食物源的比例为 α . 另一部分为基于先验性服务集生成的初始食物源, 设该部分食物源占总食物源的比例为 β , 并且, $\alpha + \beta = 1$. 基于先验性生成的初始食物源的集合如式(9)所示:

$$FS = FS_{PriSS} \cup FS_{PriS} \quad (9)$$

其中, FS_{PriSS} 表示基于先验性服务方案集生成的初始食物源, FS_{PriS} 表示基于先验性服务集生成的初始食物源.

基于先验服务方案集生成初始食物源是将满足一定条件的先验性服务方案直接作为初始食物源, FS_{PriSS} 的定义如式(10)所示:

$$\begin{aligned} FS_{PriSS} = \{ & X_i = ps_i \mid ps_i \in PS \wedge (PF(ps_i) \geq P) \wedge \\ & (US(ps_i) \geq E) \}, \\ \forall X_i, X_j \in & FS_{PS}, Het(X_i, X_j) \geq H \quad (10) \end{aligned}$$

其中, $PriSS = \{ ps_1, ps_2, \dots, ps_m \}$ 表示先验性服务方

案集, $Het(X_i, X_j)$ 为两个食物源之间的差异性函数, H 为两个食物源之间的差异性阈值. 基于先验性服务方案的初始食物源生成方法($FS(PriSS)$)的具体操作如下:

1. 计算 $PriSS$ 中先验服务方案的先验性值与用户满意度;
2. 以先验性值为依据, 对先验性服务方案进行降序排序;
3. 从上至下选出满足式(10)中条件的先验性服务方案作为初始食物源;
4. 重复步 3 直至获取 $\alpha \times SN$ 个初始食物源.

基于先验性服务集的初始食物源生成方法($FS(PriS)$)的具体操作为: 随机地从食物源每维变量对应的先验性服务集中选出一个候选服务, 构成满足约束条件的初始食物源. 基于先验性服务集生成的初始食物源集合如式(11)所示.

$$FS_{PriS} = \{ X_t = \langle s_{1_t}, \dots, s_{i_t}, \dots, s_{n_t} \rangle \mid s_{i_t} \in PriS(T_i, D) \} \quad (11)$$

(2) 初始食物源均衡生成策略

初始食物源均衡生成策略是指分别在先验性服务方案集、先验性服务集、相似性服务集与一般性服务集内生成一定数量的初始食物源, 将这些食物源放在一起作为初始食物源. 其中, 每部分食物源占总食物源的比例分别为 $\alpha, \beta, \lambda, \delta$, 且 $\alpha + \beta + \lambda + \delta = 1$. 基于均衡策略生成的初始食物源集合如式(12)所示.

$$FS = FS_{PriSS} \cup FS_{PriS} \cup FS_{SimS} \cup FS_{GenS} \quad (12)$$

其中, FS_{SimS} 、 FS_{GenS} 分别为基于相似性服务集与一般服务集生成初始食物源集合. FS_{SimS} 与 FS_{GenS} 如式(13)、(14)所示.

$$FS_{SimS} = \{ X_t = \langle s_{1_t}, \dots, s_{i_t}, \dots, s_{n_t} \rangle \mid s_{i_t} \in SimS(T_i, D) \} \quad (13)$$

$$FS_{GenS} = \{ X_t = \langle s_{1_t}, \dots, s_{i_t}, \dots, s_{n_t} \rangle \mid s_{i_t} \in GenS(T_i, D) \} \quad (14)$$

食物源 FS_{SimS} 的生成方法($FS(SimS)$)为采用随机的方法在相似性服务集上生成初始食物源. 食物源 FS_{GenS} 的生成方法($FS(GenS)$)为采用随机方法在一般性服务集上生成初始食物源. 对于每一种初始食物源生成策略, 可以根据先验性服务集的丰富程度和问题求解经验确定参数 $\alpha, \beta, \lambda, \delta$ 的取值.

6.3 面向服务领域的适应度函数

在 S-ABC 范型中, 食物源的适应度值是新旧食物源取舍的标准, 也是判断算法所搜索到的解能否满足用户需求的重要依据. 同时, 由于服务领域存

在一些领域知识约束, 适应度值也要能反应食物源满足服务领域约束的程度. 下面给出面向服务领域的适应度函数的定义.

定义 4. 面向服务领域的适应度函数

$$Fit(X_t) = \omega_1 \times fit_{US}(X_t) + \omega_2 \times fit_C(X_t) + \omega_3 \times fit_{DC}(X_t) \quad (15)$$

适应度函数 $Fit(X_t)$ 由用户满意度评价子函数 $fit_{US}(X_t)$ 、服务关联性评价子函数 $fit_C(X_t)$ 与领域约束满足程度评价子函数 $fit_{DC}(X_t)$ 组成; $\omega_j (1 \leq j \leq 3)$ 分别表示用户满意度、关联性满足程度与领域知识约束满足度在评价食物源中的重要程度, 并且

$$\sum_{j=1}^3 \omega_j = 1.$$

该适应度函数具有一定的可扩展性与可定制性, 在求解服务优化问题时, 可以增加其他服务领域特性的评价子函数(比如服务价值等), 也可以根据实际应用的需要选择或修改其中的部分评价子函数, 使得食物源的适应度值能真实地反应服务满足应用的程度. 在构造适应度函数时, 首先需要为每个子评价函数进行归一化. 下面分别给出 3 个评价子函数的定义.

(1) 用户满意度评价子函数

$$fit_{US}(X_t) = \sum_{j=1}^k \omega_j'' \times \frac{|c_j - q_j|}{q_{j_{\max}}} \quad (16)$$

$fit_{US}(X_t)$ 用于计算食物源满足用户 QoS 需求的程度. 其中, X_t 表示第 t 个食物源, q_j 表示食物源第 j 个 QoS 属性的值, c_j 表示用户对第 j 个 QoS 属性提出的约束, $q_{j_{\max}}$ 表示第 j 个 QoS 属性的最大值, k 表示 QoS 属性的数目, ω_j'' 表示用户对第 j 个 QoS 属性的偏好, 并且 $\sum_{i=1}^k \omega_j'' = 1$.

(2) 关联关系满足度评价子函数

$$fit_C(X_t) = \frac{\sum_{j=1}^m \omega_j}{m} \quad (17)$$

$fit_C(X_t)$ 用来计算食物源 X_t 满足服务之间关联性的评价子函数. m 为鼓励性关联关系(如统计关联关系与服务质量关联关系)的数目, $\omega_j (1 \leq j \leq m)$ 表示第 j 个关联关系满足度的取值 $\omega_j = p$, p 表示服务之间关联概率.

(3) 领域约束满足度评价子函数

$$fit_{DC}(X_t) = \sum_{j=1}^k \varphi_j \quad (18)$$

$fit_{DC}(X_t)$ 用来计算食物源 X_t 满足领域约束的程度. k 为服务领域约束的数目, $\varphi_j (1 \leq j \leq k)$ 表示食

物源 X_t 关于第 j 个领域约束的取值, 当其被满足时 $\varphi_j = 1$, 或者 $\varphi_j = 0$.

6.4 领域特性驱动的雇佣蜂阶段

在 ABC 算法中, 雇佣蜂通过邻域搜索生成一新食物源, 然后基于贪婪选择保留较好的食物源. 本文针对 3 个服务空间分别设计了相应的邻域搜索策略. 这里的邻域搜索是指在食物源某一维变量所对应的候选服务集上的搜索.

(1) 先验性服务集内的邻域搜索策略(NS(PriS))

由于先验性服务集内的服务之间存在一定的偏序关系, 基于这一领域知识, 设计了一种具有方向指导的自适应邻域搜索方法, 具体操作为: 在进行邻域搜索时, 首先确定最佳搜索方向, 然后根据一定的搜索步长进行邻域搜索并生成新的食物源. 下面给出搜索方向和搜索步长的确定方法.

① 搜索方向的确定

设当前食物源 X_t 在第 i 维的取值为 s_{ij} , 以 s_{ij} 为参照点, 以步长 η 在参照点的一侧选择候选服务(设为 $s_{ij+\eta}$) 替换 s_{ij} 并生成新食物源 X'_t , 通过 Rule(2) 确定在第 i 维变量对应的先验性服务集内的邻域搜索方向.

$$\begin{cases} \text{If}(Fit(X_t) < Fit(X'_t)) \text{ Then } sd = s_{ij} \rightarrow s_{ij+\eta} \\ \text{If}(Fit(X'_t) < Fit(X_t)) \text{ Then } sd = s_{ij+\eta} \rightarrow s_{ij} \end{cases} \text{Rule}(2)$$

② 搜索步长的确定

为了提高邻域搜索的效率, 邻域搜索步长应根据搜索结果进行动态调整. 搜索步长调整规则为 Rule(3):

If(sd remains unchanged) Then $\eta = \eta + \delta$ Rule(3)
其中, sd 表示搜索方向, δ 为一整数并且 $0 < \delta < \eta$, η 与 δ 的取值可根据先验性服务集中服务的数量及问题求解经验确定.

在先验性服务空间内的邻域搜索方法描述如下:

1. 确定邻域搜索方向 sd ;
2. 根据搜索方向与搜索步长生成新食物源

$$X'_t = \langle s_{1t}, \dots, s_{ij \pm \eta}, \dots, s_{nt} \rangle;$$

3. 根据 Rule(2) 与 Rule(3) 确定下次邻域搜索的方向与步长, 返回步 2.

(2) 相似性服务集内的邻域搜索策略(NS(SimS))

相似性服务集内的邻域搜索与先验性服务集内的邻域搜索方法相同, 具体操作为: 首先确定搜索方向和搜索步长, 然后在确定的搜索方向上以一定的搜索步长进行邻域搜索.

(3) 一般性服务集内的邻域搜索策略(NS(GenS))

鉴于一般性服务集内没有可利用的领域知识,

因此采取随机的邻域搜索方法,具体操作如下:设当前食物源为 $X_i = \langle s_{i_1}, \dots, s_{i_j}, \dots, s_{i_n} \rangle$, 随机地从第 i 维变量对应的服务 s_{i_j} 所属的一般性服务空间 $Gens(T_i, D)$ 中选择一候选服务 s_{i_p} 来替换 s_{i_j} , 生成一新食物源 $X'_i = \langle s_{i_1}, \dots, s_{i_p}, \dots, s_{i_n} \rangle$.

在雇佣蜂阶段,蜜蜂在执行完邻域搜索后,比较原食物源 X_i 与新食物源 X'_i 并保留较好的食物源. 这里所设计的邻域搜索策略适用于 S-ABC 范型中两种服务空间搜索策略.

6.5 面向服务领域的观察蜂阶段

在 S-ABC 范型中的观察蜂阶段采用 ABC 算法中基于轮盘赌的选择操作,赋予每个观察蜂一个食物源. 轮盘赌选择操作中食物源选择概率的计算公式如式(19)所示.

$$p_k = \text{Fit}(X_k) / \sum_{i=1}^{SN} \text{Fit}(X_i) \quad (19)$$

观察蜂在获取一个食物源之后,根据 6.4 节中设计的邻域搜索策略进行邻域搜索生成新食物源,并基于贪婪选择操作保留较好的食物源.

6.6 面向服务领域的侦查蜂阶段

针对 S-ABC 范型中两种不同的服务空间搜索策略,这里设计了两种侦查蜂阶段的新食物源生成方法.

(1) 面向服务空间优先级搜索策略的侦查蜂阶段

- ① 当蜂群处于先验性服务集时,则基于先验性服务集生成新的食物源;
- ② 当蜂群处于相似性服务集时,则基于相似性服务集生成新的食物源;
- ③ 当蜂群处于一般性服务集时,则基于一般性服务集生成新的食物源.

(2) 面向服务空间均衡搜索策略的侦查蜂阶段

基于先验性服务集、相似性服务集与一般性服务集生成新的食物源.

6.7 面向服务领域的算法判优准则

本文根据服务领域中服务应用的特点,归纳出以下 4 个面向服务领域的算法判优准则,分别为

(1) 时间约束下的最优解 (*MaxTime*)

算法搜索在达到用户提出的时间约束时输出所找到的最优解. 该判优准则如 Rule(4)所示:

If $((\text{Time} - \text{RunningTime}) \rightarrow 0)$ then Output X_{GB}
Rule(4)

(2) 用户约束下的满意解 (*UserStasifaction*)

当算法搜索到满足用户提出的满意度阈值 (*ST*) 的解时,停止搜索并输出该解. 该判优准则如 Rule(5)

所示:

If $(\text{Fit}(X_{GB}) \geq \text{ST})$ then Output X_{GB} Rule(5)

(3) 最大重复次数下的最优解 (*MaxRepeation*)

在算法的搜索过程中,算法找到的当前最优解达到最大重复次数时输出该解. 该判优准则如 Rule(6)所示.

If $(|\text{Fit}(X_{GB}^{n_1}) - \text{Fit}(X_{GB}^{n_2})| < \Delta_{fit} \wedge (n_2 - n_1) = \text{max}_{\text{repeat}})$
then Output X_{GB} Rule(6)

(4) 最大迭代次数下的最优解 (*MCN*)

当算法的迭代次数达到预先设定的最大迭代次数时,算法停止搜索并输出所找到的最优解. 该判优准则如 Rule(7)所示.

If $(\text{IterationsTime} = \text{MCN})$ then Output X_{GB}
Rule(7)

结合 S-ABC 范型框架及关键优化策略的详细设计,给出 S-ABC 范型的详细描述如算法 1 所示.

算法 1. S-ABC Paradigm.

1. 确定服务空间搜索策略

If $(\| \text{PriS} \| > \text{CP})$ // *CP* 为丰富度阈值;

选择服务空间优先级搜索策略 \mathbb{P} ;

ElseIf

选择服务空间均衡搜索策略 \mathbb{E} ;

EndIf

2. 领域特性感知的初始食物源生成

If (服务空间搜索策略 = \mathbb{P})

分别在先验性服务方案集与先验性服务集内生成 $\alpha \times SN$ 与 $\beta \times SN$ 个初始食物源;

EndIf // α 与 β 为食物源生成比例且 $\alpha + \beta = 1$;

If (服务空间搜索策略 = \mathbb{E})

分别在先验性服务方案集、先验性服务集、相似性服务集与一般性服务集生成 $\alpha \times SN$ 、 $\beta \times SN$ 、 $\gamma \times SN$ 与 $\mu \times SN$ 个食物源;

EndIf // α, β, γ 与 μ 为食物源生成比例, $\alpha + \beta + \gamma + \mu = 1$;

Repeat

3. 领域特性驱动的雇佣蜂阶段(局部优化)

For $(i = 1$ to $SN)$ // 对于每一个食物源

If (当前搜索空间 = PriS)

依据最佳搜索方向与搜索步长生成新的食物源;

评价新的食物源并保留较好的食物源;

If (当前食物源质量没有改变)

$\text{counter}_i = \text{counter}_i + 1$

EndIf

EndIf

If (当前搜索空间 = SimS)

依据搜索方向与搜索步长生成新的食物源;

评价新的食物源并保留较好的食物源;

If (当前食物源质量没有改变)

```

    counteri = counteri + 1
  EndIf
EndIf
If(当前搜索空间 == GenS)
  采用随机的方式生成新的食物源;
  评价食物源并保留较好的食物源;
  If(当前食物源质量没有改变)
    counteri = counteri + 1
  EndIf
EndIf
EndFor
4. 领域特性驱动的观察蜂阶段(局部优化)
For(i=1 to SN) //对于每一个食物源
  通过轮盘赌的形式选择一个食物源赋与观察蜂;
EndFor
Carryout 3.
5. 领域特性驱动的侦察蜂搜索阶段(全局优化)
For(i=1 to SN) //对于每一个食物源
  If(counteri > limit && 当前搜索的空间 == PriS)
    依据方法 FS(PriS)生成新的食物源;
  ElseIf(counteri > limit && 当前搜索的空间 == SimS)
    依据方法 FS(SimS)生成新的食物源;
  ElseIf(counteri > limit && 当前搜索的空间 == GenS)
    依据方法 FS(GenS)生成新的食物源;
  EndIf
EndFor
6. 更新全局最优解(全局优化)
Until
7. 满足判优准则
  If(满足判优准则)
    Return 全局最优解
  EndIf

```

S-ABC 范型的创新之处有以下 3 点:

(1) 基于服务领域特性对服务优化问题的解空间进行了划分,在一定程度上缩小了蜂群搜索的范围,使算法能够在较短的时间内找到满足用户需求的较好解。

(2) 基于服务领域特性,改进了初始食物源生成方法,基于先验性的食物源生成策略能够产生质量较好的初始解,有助于提高问题求解的效率。

(3) 利用服务空间中服务之间的关系,设计了自适应领域搜索方法,避免了盲目随机搜索,提高了算法的搜索效率。

7 S-ABC 范型验证实验

S-ABC 范型为求解服务优化问题提供了算法

模型与优化技术支持,基于 S-ABC 范型与服务优化问题的特征,可以生成求解服务优化问题的算法. 本文以服务组合问题为例,阐述 S-ABC 范型的应用并验证 S-ABC 范型的有效性及其优越性。

7.1 实验 1: S-ABC 范型有效性验证

该实验以常用的服务组合案例来验证 S-ABC 范型的有效性. 设某顺序型服务组合问题由 6 个任务节点组成,每个任务节点具有 200 个候选服务,该服务组合流程如图 3 所示. 组合流程中每个任务候选服务的 QoS 属性值取自于文献[40]提供的真实 Web 服务 QoS 数据,这里考虑了 Web 服务的反应时间(Response Time)、可靠性(Reliability)、吞吐量(Throughput)以及成功率(Successability)4 个 QoS 属性,这里对候选服务的 QoS 属性值进行了标准化处理,并将成本型 QoS 转化成效益型 QoS. 设用户对这 4 个 QoS 属性的偏好分别为 0.5, 0.3, 0.1 与 0.1.

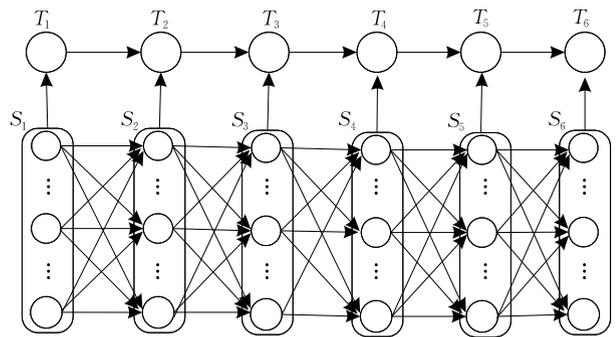


图 3 服务组合流程图

用户提出的全局 QoS 约束条件分别为

$$T \leq 4s, Rel \geq 0.05, Throu \geq 3, Suc \geq 0.05.$$

该服务组合问题可以描述为:如何从多个候选服务集中找出一组具体的服务,使得由这组具体服务所构成的组合服务既能满足用户提出的全局 QoS 约束,又具有最优的全局服务质量?该服务组合问题的数学模型如下:

$$\text{Objective Function: } \begin{cases} \text{Max } Fit(CS_i) \\ Fit(CS_i) = \sum_{k=1}^4 \omega_k \times q'_k \end{cases} \quad (20)$$

$$\text{Constraints: } \begin{cases} q'_1 \leq C_1 \\ q'_2 \geq C_2 \\ q'_3 \geq C_3 \\ q'_4 \geq C_4 \end{cases}$$

其中, CS_i 表示某一组合服务, q'_k ($1 \leq k \leq 4$) 表示组合服务 CS_i 4 个 QoS 属性的聚合值^[22], C_k ($1 \leq k \leq 4$) 表示用户提出的 QoS 约束。

7.1.1 基于 S-ABC 范型的生成服务组合算法

设该服务组合问题具有较为丰富的先验性服务集,则基于 S-ABC 范型生成的服务组合问题求解算法(S-ABC_{SC})的具体配置如表 1 所示。

表 1 基于 S-ABC 范型生成具体服务组合算法

算法要素	优化策略
1. 服务空间搜索策略	服务空间优先级搜索策略;
2. 食物源编码格式	$X_i = \langle s_{i_1}, \dots, s_{i_j}, \dots, s_{i_n} \rangle$
3. 初始食物源生成方法	基于先验性的初始食物源生成策略;
4. 适应度函数	考虑用户满意度的适应度函数;
5. 雇佣蜂阶段	搜索步长 $\eta=3, \delta=0$;
6. 观察蜂阶段	基于轮盘赌的食物源选择法;
7. 侦查蜂阶段	面向服务空间优先级策略的侦查蜂阶段;
8. 算法判优准则	Rule(4), Rule(6), Rule(7)

为了获取服务组合问题的先验性服务方案集与先验性服务集,该实验首先通过迭代遗传算法 100 次获取 200 个服务组合问题的可行解.然后,从可行解中随机地选择 20 个解作为服务组合问题的先验性服务方案集.最后,按照先验性服务集、相似性服务集以及一般性服务集的定义,找出服务组合流程中每个任务中的 3 个服务集.其中,先验性服务集的规模设为 20,相似性服务集的规模设为 40,一般性服务集的规模设为 140.将生成的先验性服务方案集、先验性服务集、相似性服务集以及一般性服务集存储到文本里,在求解服务组合问题时,S-ABC_{SC}算法可直接读取这部分数据。

在该实验中,S-ABC_{SC}算法的初始食物源规模设为 $SN=100$,设 $\alpha=0.2, \beta=0.8$.服务空间搜索转移规则为:若 S-ABC_{SC}算法对 3 种服务空间的开采度大于 70%时,进行搜索空间转移;设置食物源放弃阈值 $Limit=3$.服务组合问题求解算法 S-ABC_{SC}采用 C++编程语言实现,实验平台是 PC 机,具体配置为 OS: Microsoft Windows XP 2002; CPU: Intel(R),G550@2.60GHz; Memory: 3GB.

7.1.2 实验结果

该实验通过改变服务组合问题的先验性服务集的可信度,验证先验性对于提高算法求解效率的影响.服务组合先验性的可信度是通过改变用于生成先验性服务方案集与先验性服务集的遗传算法的迭代次数实现的,遗传算法迭代次数越多,表示服务组合问题先验性的可信度越强。

在该实验中,分别通过迭代遗传算法不同次数获得具有不同可信度的先验性服务集,之后,基于这些不同的先验性服务集,分别采用 S-ABC_{SC}算法求解服务组合问题.该实验中 S-ABC_{SC}算法的迭代次

数为 10,实验结果如图 4 所示.其中,横坐标表示服务组合问题先验性的可信度,纵坐标表示在不同的先验性可信度下,S-ABC_{SC}算法搜索到的最优解的适应度值。

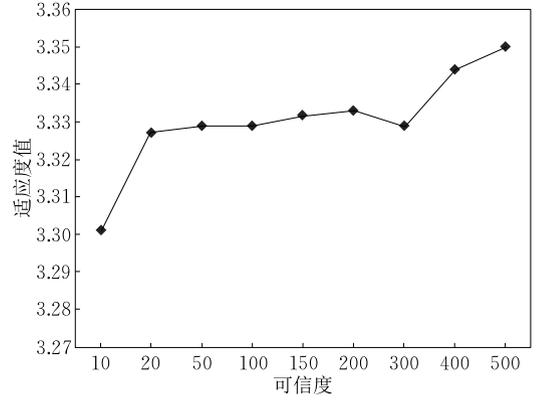


图 4 先验性有效性验证

从图 4 可以看出,随着问题先验性可信度的增强,S-ABC_{SC}算法的搜索能力逐渐增强,由此可以得出,服务组合问题先验性的可信度越高,S-ABC_{SC}算法的搜索能力越强.因此,服务领域的先验性是影响服务优化问题求解效率的一个重要因素.所以,在求解服务优化问题时,考虑先验性有助于提高问题的求解效率。

为了验证服务组合算法 S-ABC_{SC}的性能,该实验同时采用 S-ABC_{SC}算法、离散型的 ABC 算法(D-ABC)^[41]以及遗传算法(GA)同时对该服务组合问题进行求解。

3 种算法均采用 C++编程语言实现,其运行环境也相同.其中,对于 D-ABC 算法,设置初始食物源规模 $SN=100$,采用随机的方法在候选服务集上生成初始食物源,设置 $Limit=3$.遗传算法的初始群体规模设为 100,交叉概率与变异概率分别为设为 0.8 与 0.5.下面分别验证在不同算法判优准则下 3 种算法性能的优劣。

(1) 时间约束下的最优解(Rule(4))

该判优准则是在达到用户给出的最大求解时间时,给出算法搜索到的当前最优解.在执行该实验时关闭了实验环境中的其他运行程序,独立运行每一个求解算法.为了克服实验平台软、硬件环境的不确定性带来的时间误差,提高实验结果的准确度,在每个时间约束下,分别独立运行算法 20 次,取 20 次运行结果的平均值为该时间约束下最优解的适应度值.实验结果如图 5 所示.其中,纵坐标表示算法搜索到的解的适应度值,横坐标表示用户提出的不同时间约束,单位为毫秒(ms)。

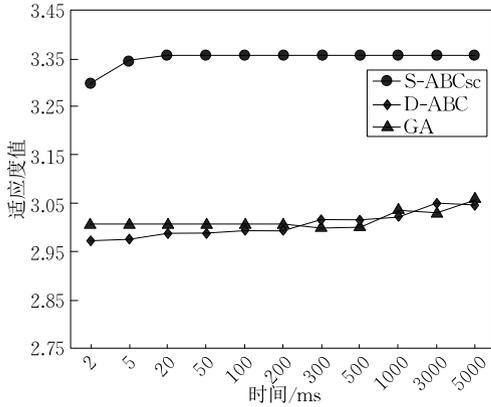


图 5 时间约束下的最优解

由图 5 可以看出,在相同时间约束下,S-ABC_{sc} 算法搜索到解优于 GA 算法与 D-ABC 算法搜索到的解. 当时间约束值较小时,GA 算法优于 D-ABC 算法;当时间约束值较大时,D-ABC 算法优于 GA 算法. 该从实验结果可以看出,S-ABC_{sc} 算法能在较短的时间内找到较好的解.

(2) 最大重复次数下的最优解(Rule(6))

该判优准则是当算法搜索的解达到给定的最大重复次数时,输出该解. 实验结果如图 6 所示,其中,纵坐标表示算法搜索到的解的适应度值,横坐标表示给定的最大重复次数.

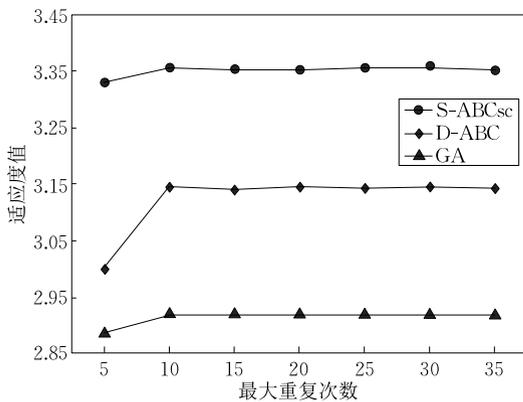


图 6 最大重复次数下的最优解

从图 6 可以看出,在相同的最大重复次数下,S-ABC_{sc} 算法搜索到解优于 GA 算法与 D-ABC 算法搜索到的解;D-ABC 算法搜索到的解优于 GA 算法搜索到的解.

(3) 最大迭代次数下的最优解(Rule(7))

该判优准则是当算法达到给定的最大迭代次数时,输出所搜索的最优解. 该实验结果如图 7 所示,其中,纵坐标表示算法搜索到的解的适应度值,横坐标表示算法的迭代次数.

从图 7 可以看出,在相同的迭代次数下,S-ABC_{sc} 算法搜索到解优于 GA 算法与 D-ABC 算法搜索到

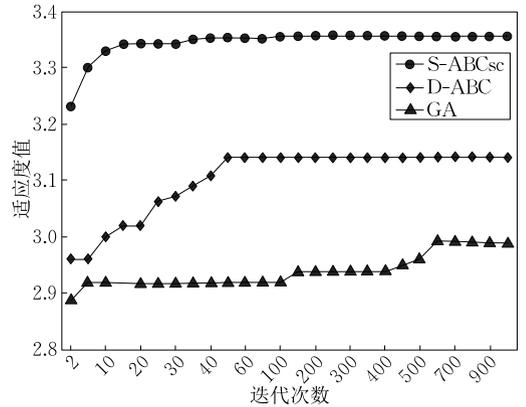


图 7 最大迭代次数下的最优解

的解,而 D-ABC 算法搜索到的解优于 GA 算法搜索到的解. 为了充分测试 3 个算法的搜索能力,分别迭代 S-ABC_{sc} 算法、D-ABC 算法与 GA 算法 20 000 次,3 种算法搜索到的解的适应度值分别为 3.356、3.06 与 3.04. 从该实验结果可以得出,S-ABC_{sc} 算法的搜索能力优于 D-ABC 算法与 GA 算法. 通过对比分析上述实验结果可以得出,在 3 种判优准则下,S-ABC_{sc} 算法的性能均优于 D-ABC 算法与 GA 算法. 与其他两种算法相比,S-ABC_{sc} 算法能在较短的时间内找到较好的解. 由于可以得出,S-ABC 范型及其优化策略是有效的.

7.2 实验 2: S-ABC 范型影响因素分析验证

S-ABC 范型的主要影响因素是先验性服务集与相似性服务集的主要特征,本文定义先验性服务集与相似性服务集的主要特征包括:先验性服务集的大小(SP)、先验性服务集的置信度(CP)、先验性服务集的使用频率(FP)、相似性服务集的大小(SS)、相似性服务集的相似度(SDS). 这些特征的计算公式如下:

$$(1) SP = \frac{1}{D} \times \sum_{i=1}^D N_{P_i}, N_{P_i} \text{ 表示第 } i \text{ 个任务节点先验性服务集的数量, } D \text{ 表示服务组合流程中任务节点的数量.}$$

(2) $CP = \frac{1}{D} \times \sum_{i=1}^D CP_i, CP_i \text{ 表示第 } i \text{ 个任务节点先验性服务集的置信度, } CP_i = \frac{1}{N_{P_i}} \times \sum_{j=1}^{N_{P_i}} \frac{US_j}{PU_j}, US_j \text{ 表示第 } j \text{ 个先验性服务满足用户需求的程度, } PU_j \text{ 表示第 } j \text{ 个先验性服务被使用的次数.}$

$$(3) FP = \frac{1}{D} \times \sum_{i=1}^D FP_i, FP_i = \left(\sum_{j=1}^{N_{P_i}} PU_j \right) / hdSize, hdSize \text{ 表示历史数据的规模.}$$

(4) $SS = \frac{1}{D} \times \sum_{i=1}^D N_{S_i}, N_{S_i} \text{ 表示第 } i \text{ 个任务节点相似服务的规模, } N_{S_i} = \left| U \left(\sum_{j=1}^{N_{P_i}} Sim(sp_j) \right) \right|, sp_j \text{ 表}$

示第 j 个先验性服务, $Sim()$ 是计算候选服务与 s_{P_j} 的相似度的函数, $U()$ 表示计算第 i 个任务节点相似服务集的综合相似度的函数.

$$(5) SDS = Avg \left(\sum_{j=1}^{N_{P_i}} \sum_{l=1}^N CED(s_{P_j}, s_l) \right), CED()$$

表示计算服务之间欧氏距离的函数, $Arg()$ 计算满足下述表达式的数据的平均值, $CED(s_{P_j}, s_l) \geq t_s, t_s$ 为相似度阈值.

该实验采用图 8 所示服务组合问题作为实验对象, 该服务组合流程共包括 11 个任务节点, 由顺序结构、选择结构以及并行结构组成; 设每个任务节点具有 5000 个候选服务, 考虑了 5 个 QoS 属性, 分别为反应时间、费用、可靠性、可用性以及吞吐量, 每个服务的 QoS 属性值在以下范围内随机生成: $[1, 10]$, $[20, 100]$, $[0.6, 0.99]$, $[0.8, 0.99]$, $[1, 20]$; 设用户

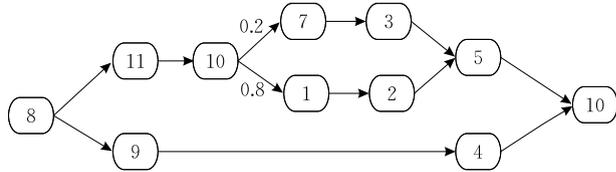


图 8 服务组合流程图

提出的全局 QoS 约束为 $\langle 50, 400, 0.2, 0.85, 3 \rangle$; 用户对每个 QoS 属性的偏好为 0.2.

该服务组合问题的目标函数如式(21)所示:

$$F(CS) = \sum_{k=1}^5 \omega_k \times F_k \times ((RQ_k - CQoS_k(CS, SerPro)) / RQ_k) \quad (21)$$

其中, $CQoS_k$ 表示第 k 个 QoS 属性值的聚合值; RQ_k 表示用户提出的关于第 k 个 QoS 属性的约束; 当第 k 个 QoS 属性为效益型时 $F_k = 1$, 或者, $F_k = -1$; $\sum_{k=1}^5 \omega_k = 1$.

基于 S-ABC 范型生成具体服务组合算法的配置与表 1 相同.

在该实验中, 采用 ABC 算法求解图 8 所示的服务组合问题, 通过随机设置不同的最大迭代次数来获取不同的服务使用历史数据集. 在该实验中, 共产生 200 个服务使用历史数据集, 随机地从中抽取 20 个数据集作为实验对象, 其编号记为 $HD1, HD2, \dots, HD20$. 之后, 计算出每个数据集的特征值, 如图 9 所示, 其中, 横坐标表示数据集的序号, 纵坐标表示不通的特征值.

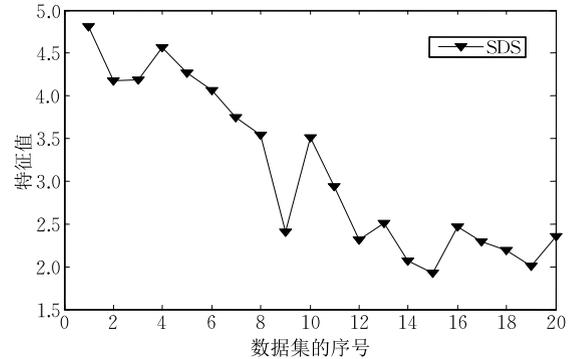
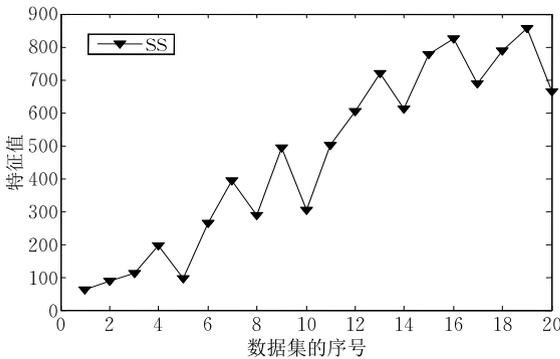
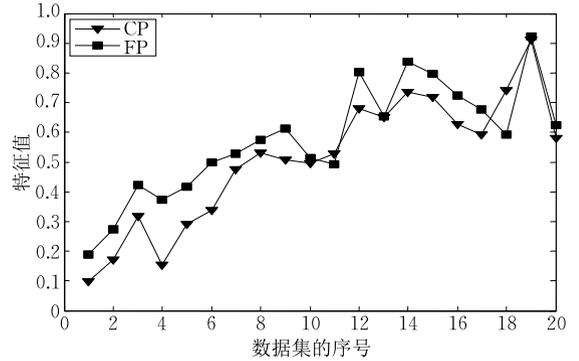
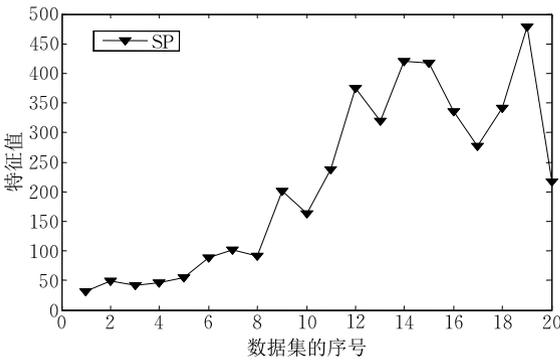


图 9 20 个数据集的特征值

从图 9 可以看出特征值 SP, CP, FP, SS, SDS 的取值范围分别为 $[0, 500], [0, 1], [0, 1], [0, 900], [0, 5]$. 为了验证 S-ABC_{sc} 算法寻优能力, 该实验采用 ABC 算法与 S-ABC_{sc} 算法同时求解图 8 所示的

服务组合问题. 采用所选出的 20 个历史数据集生成服务组合问题的先验性服务集、相似性服务集以及一般性服务, 将其作为 S-ABC_{sc} 算法求解服务组合问题的已知条件; 两个算法中参数 $SN, Limit$ 与

MCN(最大迭代次数)的取值相同,实验结果如图 10 所示,其中,横坐标表示数据集的序号,纵坐标表示

算法所搜索到的解的评价值,实验结果为运行算法 50 次求得平均值.

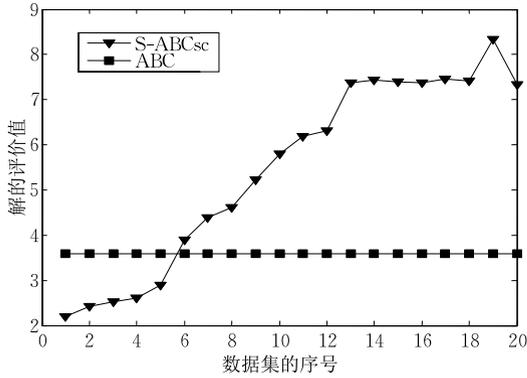
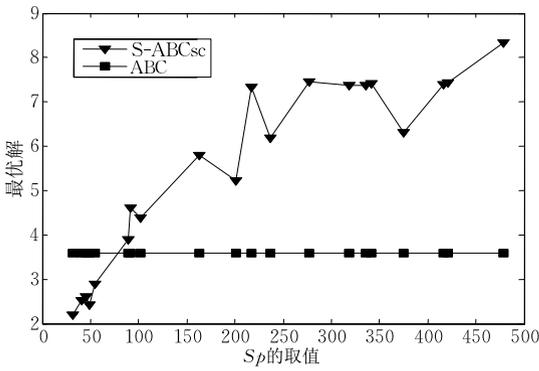
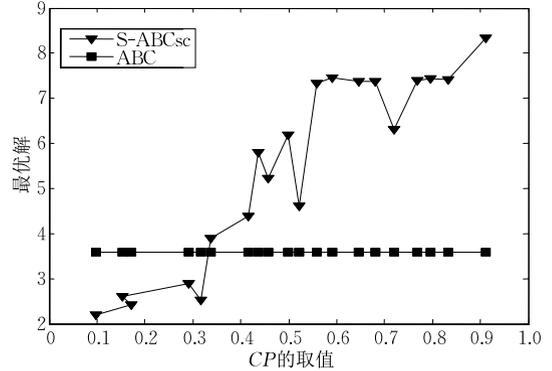


图 10 两种算法在不同数据集的最优解

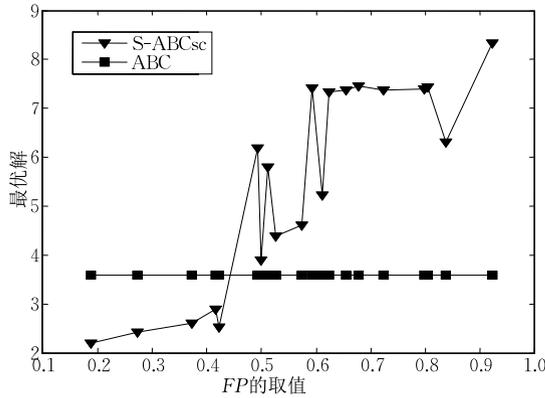
通过图 10 可以看出, S-ABC_{sc} 算法在数据集 HD1、HD2、HD3、HD4 与 HD5 上的寻优性能不如 ABC 算法, 在其余 15 个数据集上的寻优性能优于 ABC 算法. 该实验结果表明, 在不同的先验性服务集与相似性服务集上, S-ABC_{sc} 算法的寻优能力有所不同, 也即先验性服务集与相似性服务集的特征值对 S-ABC_{sc} 算法的寻优性能具有一定的影响. 下面通过实验分别验证了服务集特征 SP、CP、FP、SS 与 SDS 对 S-ABC_{sc} 算法寻优能力的影响, 实验结果如图 11 所示.



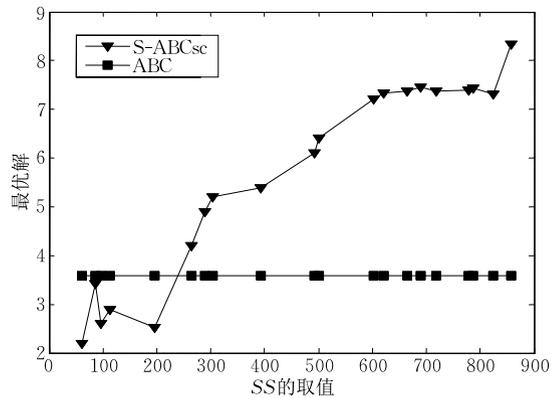
(a) 最优解与SP取值的关系



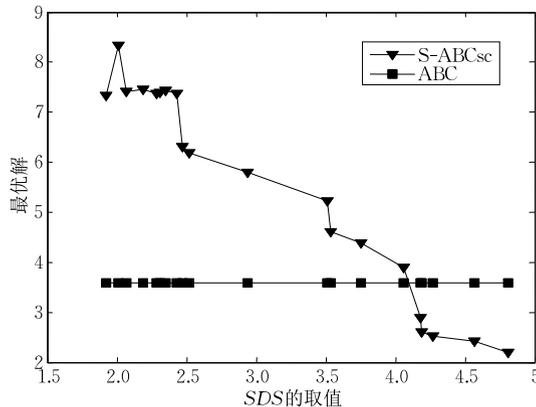
(b) 最优解与CP取值的关系



(c) 最优解与FP取值的关系



(d) 最优解与SS取值的关系



(e) 最优解与SDS取值的关系

图 11 服务集特征对 S-ABC_{sc} 算法性能的影响

从图 11 可以看出,随着 SP, CP, FP 与 SS 数值的增加, S-ABC_{SC} 算法寻优能力在逐渐增强, 这表明当先验性服务集的丰富度越高、置信度越高、服务使用频率越高以及相似性服务集越丰富, 则 S-ABC_{SC} 算法的寻优能力越强; 这与本文研究的出发点一致, 即先验性与相似性有助于提高服务优化问题求解的效果; 然而, 随着 SDS 值的增加, S-ABC_{SC} 算法的寻优能力逐渐变差, 这是因为服务之间的相似度越高, 服务之间的差异性越低, 解的多样性越低, 致使 S-ABC_{SC} 算法搜索能力变差. 通过比较分析上述实验结果可以得出: 考虑服务领域的先验性与相似性, 有助于提高服务优化问题求解的效果; 先验性服务集与相似服务集的特征值 SP, CP, FP 与 SS 越强, 也即当服务优化问题的先验性准确度可靠越丰富时, 应用 S-ABC 范型求解服务优化问题的效果越好.

8 总结与展望

随着互联网迅猛的发展和服务资源的不断丰富, 服务解决方案的优化构建变得越来越复杂. 研究者提出了各类通用优化方法来尝试求解这些问题, 但往往忽略了服务问题所在的领域环境与特性及其对优化求解的影响. 本文基于人工蜂群算法的优化机理及服务领域特性对服务优化问题求解的影响规律, 设计了领域特性驱动和优化策略, 提出了面向服务领域的人工蜂群算法范型(S-ABC), 为服务优化问题的高效求解提供了算法模型及优化技术支持.

在后期的研究中, 将进一步细化 S-ABC 范型规范及其要素, 继续深入研究针对各种具体服务优化问题的 S-ABC 范型的子范型以及用于求解典型服务优化问题的 S-ABC_x 系列算法, 进一步将 S-ABC 范型及其算法应用于具体服务应用领域(如智慧物流服务、智慧家庭服务、智慧医疗服务、智慧城市服务等)优化问题的求解, 并证实 S-ABC 范型的优越性和实用价值, 进一步丰富和改进 S-ABC 范型的理论与方法.

参 考 文 献

- [1] Zeng L, Benatallah B, Ngu A H H, et al. QoS-aware middleware for Web services composition. *IEEE Transactions on Software Engineering*, 2004, 30(5): 311-327
- [2] Ngoko Y, Goldman A, Milojicic D. Service selection in Web service compositions optimizing energy consumption and service response time. *Journal of Internet Services and Applications*, 2013, 4(1): 1-12
- [3] Yu T, Zhang Y, Lin K J. Efficient algorithms for Web services selection with end-to-end QoS constraints. *ACM Transactions on the Web*, 2007, 1(1): 6
- [4] Yu T, Lin K J. Service selection algorithms for Web services with end-to-end QoS constraints. *Information Systems and E-Business Management*, 2005, 3(2): 103-126
- [5] Canfora G, Di Penta M, Esposito R, et al. An approach for QoS-aware service composition based on genetic algorithms// *Proceedings of the 7th Annual Conference on Genetic and Evolutionary Computation*. New York, USA, 2005: 1069-1075
- [6] Ma Y, Zhang C. Quick convergence of genetic algorithm for QoS-driven Web service selection. *Computer Networks*, 2008, 52(5): 1093-1104
- [7] Zhang W, Chang C K, Feng T, et al. QoS-based dynamic Web service composition with ant colony optimization// *Proceedings of the IEEE 34th Annual Computer Software and Applications Conference (COMPSAC, 2010)*. Seoul, Korea, 2010: 493-502
- [8] Rostami N H, Kheirkhah E, Jalali M. An Optimized Semantic Web Service Composition Method Based on Clustering and Ant Colony Algorithm. *arXiv preprint arXiv*, 2014: 1402.2271
- [9] Zhao X, Song B, Huang P, et al. An improved discrete immune optimization algorithm based on PSO for QoS-driven Web service composition. *Applied Soft Computing*, 2012, 12(8): 2208-2216
- [10] Wen Tao, Sheng Guo-Jun, Guo Quan, Li Ying-Qiu. Web service composition based on modified particle swarm optimization. *Chinese Journal of Computers*, 2014, 36(5): 1031-1046(in Chinese)
(温涛, 盛国军, 郭权, 李迎秋. 基于改进粒子群算法的 Web 服务组合. *计算机学报*, 2014, 36(5): 1031-1046)
- [11] He J, Chen L, Wang X, et al. Web service composition optimization based on improved artificial bee colony algorithm. *Journal of Networks*, 2013, 8(9): 2143-2149
- [12] Wang X, Wang Z, Xu X. An improved artificial bee colony approach to QoS-aware service selection//*Proceedings of the IEEE 20th International Conference on Web Services (ICWS 2013)*. Santa Clara Marriott, USA, 2013: 395-402
- [13] Wang S, Wang Z, Xu X. An artificial bee colony algorithm based optimization method for service network customization// *Proceedings of the International Conference on Service Sciences (ICSS 2013)*. Shenzhen, China, 2013: 101-106
- [14] Rajeswary C. A survey on efficient evolutionary algorithms for Web service selection. *International Journal of Management, IT and Engineering*, 2012, 2(9): 177-191
- [15] Karaboga D. An idea based on honey bee swarm for numerical optimization. *Computer Engineering Department, Engineering Faculty, Erciyes University; Technical Report-tr06*, 2005
- [16] Karaboga D, Basturk B. On the performance of artificial bee colony (ABC) algorithm. *Applied Soft Computing*, 2008, 8(1): 687-697

- [17] Karaboga D, Gorkemli B, Ozturk C, et al. A comprehensive survey: Artificial bee colony(ABC) algorithm and applications. *Artificial Intelligence Review*, 2014, 42(1): 21-57
- [18] Zhang Ming-Wei, Wei Wei-Jie, Zhang Bin, Zhang Xi-Zhe, Zhu Zhi-Liang. Research on service selection approach based on composite service execution information. *Chinese Journal of Computers*, 2008, 31(8): 1398-1411(in Chinese)
(张明卫, 魏伟杰, 张斌, 张锡哲, 朱志良. 基于组合服务执行信息的服务选取方法研究. *计算机学报*, 2008, 31(8): 1398-1411)
- [19] Zhang Ming-Wei, Zhang Bin, Zhang Xi-Zhe, Zhu Zhi-Liang. A division based composite service selection approach. *Journal of Computer Research and Development*, 2012, 49(5): 1005-1017(in Chinese)
(张明卫, 张斌, 张锡哲, 朱志良. 一种基于划分的组合服务选取方法. *计算机研究与发展*, 2012, 49(5): 1005-1017)
- [20] Kang G, Tang M, Liu J, et al. Diversifying Web service recommendation results via exploring service usage history. *IEEE Transactions on Services Computing*, DOI 10.1109/TSC.2015.2415807
- [21] Kang G, Liu J, Tang M, et al. AWR: Active Web service recommendation based on usage history//Proceedings of the 2012 IEEE 19th International Conference on Web Services (ICWS). Honolulu, Hawaii, USA, 2012: 186-193
- [22] Dai Yu, Yang Lei, Zhang Bin, Gao Yan. QoS for composite Web services and optimizing. *Chinese Journal of Computers*, 2006, 29(7): 1167-1178(in Chinese)
(代钰, 杨雷, 张斌, 高岩. 支持组合服务选取的 QoS 模型及优化求解. *计算机学报*, 2006, 29(7): 1167-1178)
- [23] Ye Shi-Yang, Wei Jun, Li Lei, Huang Tao. Service correlation aware selection for composite service. *Chinese Journal of Computers*, 2008, 31(8): 1383-1397(in Chinese)
(叶世阳, 魏峻, 李磊, 黄涛. 支持服务关联的组合服务选择方法研究. *计算机学报*, 2008, 31(8): 1383-1397)
- [24] Li H, Jiang R, Ge S. Researches on manufacturing cloud service composition & optimization approach supporting for service statistic correlation//Proceedings of the 26th IEEE Control and Decision Conference (2014 CCDC). Changsha, China, 2014: 4149-4154
- [25] Wu Q, Zhu Q, Zhou M. A correlation-driven optimal service selection approach for virtual enterprise establishment. *Journal of Intelligent Manufacturing*, 2014, 25(6): 1441-1453
- [26] Bravo M. Similarity measures for Web service composition models. *International Journal on Web Service Computing*. Marzo, 2014, 5(1): 1-16
- [27] Li Xi-Tong, Fan Yu-Shun. Analyzing compatibility and similarity of Web service processes. *Chinese Journal of Computers*, 2009, 32(12): 2429-2437(in Chinese)
(李喜彤, 范玉顺. Web 服务流程相容性和相似性分析. *计算机学报*, 2009, 32(12): 2429-2437)
- [28] Zhang M, Zhang B, Na J, et al. Composite service selection based on dot pattern mining//Proceedings of the IEEE World Conference on Services-I. Los Angeles, USA, 2009: 740-746
- [29] Tang R, Zou Y. An approach for mining Web service composition patterns from execution logs//Proceedings of the 12th IEEE International Symposium on Web Systems Evolution (WSE 2010). Timisoara, Romania, 2010: 53-62
- [30] Huang Z, Huai J, Liu X, Zhu J. Business process decomposition based on service relevance mining//Proceedings of the International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT 2010). Toronto, Canada, 2010: 573-580
- [31] Li Su-Fen, Fan Yu-Shun, Li Xiu. Service selection method supporting service relation. *Computer Integrated Manufacturing Systems*, 2012, 18(3): 645-653(in Chinese)
(李素芬, 范玉顺, 李秀. 支持服务应用关联的服务选择方法. *计算机集成制造系统*, 2012, 18(3): 645-653)
- [32] Hassan O A H, Ramaswamy L, Miller J A. MACE: A dynamic caching framework for mashups//Proceedings of the IEEE International Conference on Web Services (ICWS 2009). Los Angeles, USA, 2009: 75-82
- [33] Karaboga N, Latifoglu F. Adaptive filtering noisy transcranial Doppler signal by using artificial bee colony algorithm. *Engineering Applications of Artificial Intelligence*, 2013, 26(2): 677-684
- [34] Karaboga D, Gorkemli B. A quick artificial bee colony (qABC) algorithm and its performance on optimization problems. *Applied Soft Computing*, 2014, 23: 227-238
- [35] Karaboga D, Akay B. A comparative study of artificial bee colony algorithm. *Applied Mathematics and Computation*, 2009, 214(1): 108-132
- [36] Akay B, Karaboga D. A modified artificial bee colony algorithm for real-parameter optimization. *Information Sciences*, 2012, 192: 120-142
- [37] Karaboga D, Basturk B. A powerful and efficient algorithm for numerical function optimization: Artificial bee colony (ABC) algorithm. *Journal of Global Optimization*, 2007, 39(3): 459-471
- [38] Brajevic I, Tuba M. An upgraded artificial bee colony (ABC) algorithm for constrained optimization problems. *Journal of Intelligent Manufacturing*, 2013, 24(4): 729-740
- [39] Karaboga N. A new design method based on artificial bee colony algorithm for digital IIR filters. *Journal of the Franklin Institute*, 2009, 346(4): 328-348
- [40] Al-Masri E, Mahmoud Q H. Investigating Web services on the world wide web//Proceedings of the 17th International Conference on World Wide Web. New York, USA, 2008: 795-804
- [41] Pan Q K, Fatih Tasgetiren M, Suganthan P N, et al. A discrete artificial bee colony algorithm for the lot-streaming flow shop scheduling problem. *Information Sciences*, 2011, 181(12): 2455-2468



XU Xiao-Fei, born in 1962, Ph. D. , professor, Ph.D. supervisor. His research interests include service computing and software service engineering, enterprise computing, data mining and business intelligence.

LIU Zhi-Zhong, born in 1981, Ph. D. , lecturer. His research interests include service computing and swarm intelligence algorithm.

Background

With the rapid development of Service Computing, Cloud Computing, Internet of Things, Big Data and other new information technologies, large-scale intelligent business services are available in service domains and provide abundant service resources for users. Many complicated factors of the service environment, such as distribution, aggregation and cross-domain of services, dynamics of service resource and variability of user demands, cause service applications more and more complicated. Meanwhile, in the service industries, it gradually forms the unique service domain features (such as priori, correlation and similarity) with the development and evolution of service ecosystem. These service domain features are the objective law existing in the service ecosystem and have strong influences on solving service optimization problems (service selection, service composition and service resource scheduling).

However, existing research for solving service optimization problems are often short of perceiving the service domain features and always causing inconsistent results between theoretical research and the practical service applications in service domains, lead to undesirable efficiency and effect of service optimization problems solving. Therefore, how to design service domain-oriented optimization algorithm paradigm and provide algorithm model and optimization technology for solving service optimization problems effectively becomes a key issue to be solved. For this problem, this paper proposes

WANG Zhong-Jie, born in 1978, Ph. D. , professor, Ph. D. supervisor. His research interests include service computing, service engineering, software architecture and its evolution.

MIN Xun-You, born in 1988, Ph. D. candidate. Her research interest is service computing.

LIU Rui-Lin, born in 1989, Ph. D. candidate. His research interest is service computing.

WANG Hai-Fang, born in 1989, M. S. candidate. Her research interest is service computing.

the service domain-oriented artificial bee colony algorithm paradigm (S-ABC) based on the optimization mechanism of the artificial bee colony algorithm and the influence of service domain features on solving service optimization problems. S-ABC paradigm provides algorithm model and optimization technology for solving service optimization problem effectively. This work is expected to strengthen the combination of artificial bee colony algorithms and the optimization problems of service domain, to explore a new and better method for solving service optimization problems, and to extend the theory of swarm intelligence optimization.

This work is partly supported by the National Natural Science Foundation of China (Nos. 61472106, 61272187, 61300124, 61033005). These projects aim to provide better principles for the design and management of service computing systems, our team has been working on the design and optimization of service computing systems for years. Several research papers have been published on ICWS, SCC, International Journal of Web Services Research, Service Oriented Computing and Applications, and so on. The work of this paper is the new and important research result based on our existing research work, and is the key research part of the project “Research on Service Domain-Oriented Artificial Bee Colony Algorithm Paradigm and Optimization Theory” (No. 61472106).