

基于服务质量的 Web 服务优化选择算法及仿真

王俊丽, 柳先辉, 卫 刚

(同济大学 企业数字化技术教育部工程研究中心, 上海 200092)

摘要: 提出一种基于服务非功能属性的 Web 服务优化选择的多目标遗传算法. 考虑到在服务选择时组合方案中的控制结构、具体服务之间的关联性和多个服务质量指标之间的折衷这三方面的约束, 对服务选择建立分层模型给出优化选择的形式化定义. 在此基础上, 采用多目标遗传算法求解优化服务选择. 首先设计了适合的染色体编码方式, 以表示可行的服务选择方案, 改进基本的遗传算子, 采用两种遗传策略来避免算法陷入局部最优, 以保证种群的收敛性. 仿真结果表明, 该方法能在有限进化代数内产生一组满足用户服务质量需求的非劣服务组合方案解, 遗传策略在提高算法收敛性和维护种群多样性方面也发挥了作用.

关键词: Web 服务选择; 服务质量; 多目标遗传算法

中图分类号: TP 393

文献标识码: A

Quality of Service-based Web Services Optimal Selection Algorithm and Its Simulation

WANG Junli, LIU Xianhui, WEI Gang

(Engineering Research Center for Enterprise Digital Technology of the Ministry of Education, Tongji University, Shanghai 200092, China)

Abstract: The paper presents a quality of service-based(QoS) Web services optimal selection method by multi-objective genetic algorithm by taking three aspects of constrains in services selection process into consideration, such as control structure within a composition plan, relationship between concrete services, and tradeoff among multiple QoS indexes. Moreover, a layered services selection model is constructed for formalizing optimal services selection problem, and a multi-objective genetic algorithm is first adopted. In particular, a chromosome coding method is first designed to represent a feasible services selection solution, and then genetic operators and strategies are developed to maintain diversity of population and also avoid getting trapped in local optima.

Finally, experimental results show that this algorithm can generate a set of nondominated Pareto optimal solutions which satisfy to users' QoS requirements within a finite number of evolving generation. The adopted genetic strategies have good effect on maintaining diversity of population and increasing convergence of algorithm.

Key words: Web services selection; quality of service; multi-objective genetic algorithm

Web 服务是一种新兴的分布式技术, 解决如何让计算机系统使用 Web 应用所提供的服务. 然而, 单独的 Web 服务很可能无法满足用户的请求, 需要通过合成现有的 Web 服务来创造出新的服务功能. 因而 Web 服务组合^[1]受到了学术界和工业界的广泛关注. 但是 Web 服务的自动组合需要对服务进行一定的语义描述. WSDL(Web service description language, Web 服务描述语言)主要集中于数据交换和服务发布的语法标准, 缺乏对服务描述的语义理解. 为此, 研究者们提出了 OWL-S(ontology web language for service, Web 服务本体语言)语言规范^[1]. 当前, 基于 OWL-S 的服务组合主要是研究服务的功能属性匹配, 如输入、输出、前提和效果的匹配, 以产生若干可行的服务组合方案^[1-2], 组合方案由一些抽象服务和控制结构组成. 在此基础上, 需要考虑的是如何根据已有的组合方案选择最佳的、满足用户需求的具体服务, 此时就需要借助 Web 服务的非功能属性如服务分类和 QoS(quality of service, 服务质量)指标等^[3-4], 对方案中的抽象服务选择绑定相应的具体服务^[5-10].

在选择服务时一般需要考虑以下三方面的约束: ①组合方案中的控制结构, 如顺序、并行和选择

收稿日期: 2010-01-05

基金项目: 上海市科委项目(09DZ1122701, 09DZ1120403)

第一作者: 王俊丽(1978—), 女, 讲师, 工学博士, 主要研究方向为计算机应用. E-mail: junliwang@tongji.edu.cn

通讯作者: 柳先辉(1979—), 男, 讲师, 工学博士, 主要研究方向为计算机应用. E-mail: tjercedt@126.com

等,会影响具体服务的选择,例如并行关系对于服务响应时间指标的影响;②具体服务之间的关联性,如联盟或互斥关系,也会对服务选择的结果产生影响;③多个 QoS 指标之间的冲突性.多个目标之间往往是相互制约的,很难同时最优,因此,需要多目标的协调和折衷,使各子目标函数尽可能达到最优.

目前,已经提出一些 QoS 服务选择的优化方法^[5-8].文献[5]提出一种采用整数规划的穷举算法,根据多个 QoS 指标,动态地选择最适合用户需求的服务选择方案.该方法考虑了组合方案中控制结构对服务选择的约束,但是没有涉及具体服务之间的关联性.另外,已经证明基于 QoS 的 Web 服务选择是 NP(non-deterministic polynomial)难问题^[11].一般情况下,不存在多项式时间的服务选择穷举计算方法,目前更多地是采用非线性方法.其中,遗传算法^[12-13]具有线性时间复杂度和全局收敛的特点,应用比较广泛.文献[7]提出了一种基于遗传算法的服务选择优化方案,将多个 QoS 指标采用加权和方法聚合成一个单目标函数,但是,这种传统的方法有其固有的缺点,对于非劣解集的某些区域不可能求出^[13],文中也没有考虑控制结构对服务选择结果的影响.文献[8]针对服务的单个 QoS 指标(服务费用),讨论了具体服务之间存在关联关系约束时,如何运用遗传算法来求解服务选择问题,但是没有涉及控制结构和多个 QoS 之间的冲突性对服务选择的影响.综上所述,现有的基于遗传算法的 Web 服务选择方法在一定程度上解决了服务选择问题,但缺少整体考虑上述三方面的约束,不能很好地解决 Web 服务优化选择问题,以满足用户对多个 QoS 指标的需要.

针对上述问题,笔者提出并设计了一个多目标遗传算法来求解服务选择优化问题.首先,提出了面向服务选择问题的多层服务模型架构,刻画了服务分类属性、抽象服务组合方案和具体服务集合之间的层次关系,界定抽象服务之间的控制结构约束和具体服务之间的关联性约束.然后,设计一个多目标遗传算法来求解服务优化选择问题.具体为:①在染色体编码过程中,将抽象服务之间的控制结构关系和具体服务之间的关联性隐含,在可行染色体的编码机制中,避免无效的选择方案进入优化种群,节省种群进化时间;②设计适用于服务优化选择问题的遗传算子和遗传策略,避免算法陷入局部最优,而且保证了种群的收敛性;③运用 Pareto 最优关系来优化处理多个相互冲突的 QoS 指标,生成一组满足用

户多 QoS 需求的服务选择方案.最后,用仿真实验模拟本算法,通过实验结果分析算法的性能和优化效果.结果表明,该方法能在有限进化代数内产生一组满足用户 QoS 需求的非劣服务选择方案解.

1 Web 服务优化选择描述

在 OWL-S 规范中^[14],服务轮廓提供了描述服务特征的信息.其中,除了定义服务的功能属性,例如信息转换(输入和输出)和服务执行所产生的状态改变(前提和效果),还允许描述服务的非功能属性,主要包括服务分类和 QoS.OWL-S 提供了表达 QoS 指标 I_{QoS} 的机制,但没有明确定义具体的 I_{QoS} .因而很多研究者针对服务选择的需要,定义了具体的 I_{QoS} 及其描述^[15],以允许用户关于服务的非功能请求.这里主要考虑以下三种 I_{QoS} :服务费用、响应时间和可靠性.

1.1 服务选择模型

为了清晰地描述服务选择,构建分层服务模型,包括服务分类层、抽象服务层和具体服务层.

(1)服务分类层 按标准的分类系统——联合国标准产品与服务分类代码(UNSPSC)中的分类标准建立领域本体,将服务所属领域内的概念分类.故当服务提供者在注册服务时需指明 UNSPSC 分类.

(2)抽象服务层 提供一个服务组合方案——由控制结构连接一组抽象服务得到,实现一定功能以满足特定用户需求.控制结构包括抽象服务之间的顺序、选择和并行关系.每个抽象服务都隶属于某个服务分类. $S_{aw} = \{S_1, S_2, \dots, S_n\}$ 表示抽象服务层中的所有抽象服务集合.

(3)具体服务层 是由提供商在 UDDI(universal description discovery and integration,统一描述、发现和集成)注册中心中注册的可以执行的服务.其中的每个抽象服务 S_i 对应于一组功能相似的具体服务 $\{s_{i1}, s_{i2}, \dots, s_{ik(i)}\}$,其中 $k(i)$ 表示第 i 个抽象服务对应的具体服务数目.二者的对应关系记为 $f(s_{ij}) = S_i$.对应于抽象服务集合 S_{aw} 的具体服务集合为 $S_{cw} = \{s_{11}, \dots, s_{1k(1)}; s_{21}, \dots, s_{2k(2)}; s_{n1}, \dots, s_{nk(n)}\}$,简记为 $S_{cw} = \{s_1, s_2, \dots, s_m\}$, $m = \sum_{i=1}^n k(i)$.另外,具体的服务之间存在着两种约束关系:联盟关联性和互斥关联性.联盟关系是同一个提供商的服务可能是绑定的,要求它们执行时必须同时发生;互斥关系是某些服务可能是由相互竞争的

提供商提供的,因而不允许在同一个组合服务中执行.

图1给出了旅行安排服务选择的多层服务模型框架:①在服务分类层,通过建立旅行服务领域本体,对该领域的相关概念分类.②抽象服务层对应于一个旅行服务组合方案,包括旅游信息、住宿预订、天气状况查询、票务预订等服务.这个方案存在着三种控制结构:选择结构、并行结构和顺序结构.例如,票务预订中有三个可选择的服务——汽车票、火车

票和机票预定,而天气查询和旅行信息服务是并行结构,住宿服务和天气查询服务是顺序结构.③具体服务层是UDDI中的一组可执行的服务,存在着两类约束关系:联盟关系(例如 s_{22} 和 s_{42})和互斥关系(例如 s_{33} 和 s_{43}).层之间的关系用虚线表示,如从服务分类层到抽象服务层之间的虚线,表示每个抽象服务都隶属某个服务分类;而从抽象服务层到具体服务层之间的虚线,表示每个抽象服务都对应于一组功能相似的具体服务.

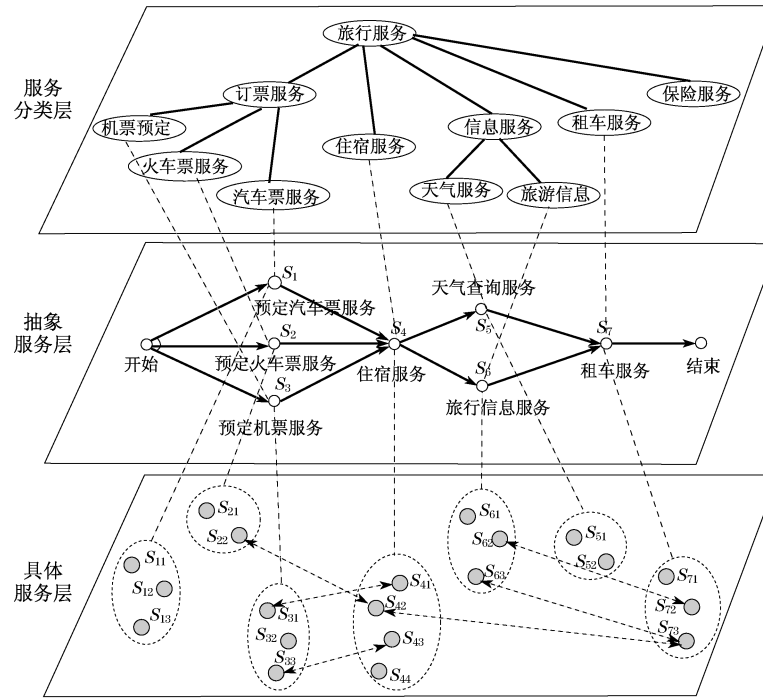


图1 分层服务模型框架

Fig.1 Framework model for layered service

1.2 服务优化选择问题描述

首先,给出服务组合方案的归纳定义:

定义1 设 $S_{aw} = \{S_1, S_2, \dots, S_n\}$ 是一个抽象服务集合,每个抽象服务 S_i 隶属于某个服务分类,则

① $S_i \in S_{aw}$ 是一个服务组合方案,称为基本服务组合方案;

② $S_{eq}(S_1, S_2, \dots), P(S_1, S_2, \dots)$ 和 $C(S_1, S_2, \dots)$ 也是服务组合方案,称为简单服务组合方案.

③ 如果 S'_1 和 S'_2 是服务组合方案,那么, $S_{eq}(S'_1, S'_2, \dots), P(S'_1, S'_2, \dots)$ 也是服务组合方案,称为复合服务组合方案.

④ 对于每个服务组合方案,必须满足 $|S_i| \leq 1$, $|S_i|$ 是 S_i 在组合方案中出现的次数.

由于抽象服务之间存在顺序、选择和并行结构,

因而每个服务选择方案必须满足两个约束条件,一个选择控制结构模块中的所有抽象服务,只能允许有一个抽象服务绑定唯一的具体服务;其余的每个抽象服务绑定唯一的具体服务.形式化描述为:

定义2 设 $S_{aw} = \{S_1, S_2, \dots, S_n\}$ 是抽象服务集合, P_c 是定义在 S_{aw} 上的一个服务组合方案, $S_{cw} = \{s_1, s_2, \dots, s_m\}$ 是对应 S_{aw} 的具体服务集合,设 $S_o = (x_1, x_2, \dots, x_m)$, 其中 $x_i = 0$ 或 $x_i = 1$, $x_i = 0$ 表明未选择具体服务 s_i , $x_i = 1$ 表明 s_i 被选择, m 是具体服务个数.如果——

① 对于 P_c 的任意一个选择结构模块 $B_c(S_{i1}, S_{i2}, \dots, S_{ik})$, 对于任意满足 $f(s_u) = S_{il}$ 的 s_u , $1 \leq l \leq k$, 满足 $\sum_n x_u = 1$;

② 对于其他的 S_i , 任意满足 $f(s_v) = S_i$ 的 s_v , 满足 $\sum_v x_v = 1$.

则称 S_o 是对应于 P_c 的一个服务选择方案.

定义 3 三元组的 QoS 指标 $I_{QoS}(s) = (C(s), T(s), R(s))$ 称为具体服务 s 的 I_{QoS} , $C(s)$ 表示 s 的费用, $T(s)$ 表示 s 的响应时间; $R(s)$ 表示 s 的可靠性.

在定义 3 的基础上, 可以定义服务选择方案 S_o 的 QoS 指标为 $I_{QoS}(S) = (C(S), T(S), R(S))$. 因为每个具体的服务 s_i 都有其相应的 I_{QoS} . 则根据表 1 中的公式, 可以通过具体服务 I_{QoS} 计算出总体 $I_{QoS}(S)^{[5]}$. 表 1 中, s_i 表示相应控制结构模块中的某个具体服务, w 是模块中的服务数目, p_i 表示 s_i 是否被选择, 取值为 0 或 1.

另外, 为了刻画具体的服务之间的约束关系, 定义如下具体服务关联矩阵:

定义 4 设 $S_{cw} = \{s_1, s_2, \dots, s_m\}$ 是具体服务集合, 则称 $A = \{r_{ij}\}$ 是 S_{cw} 的关联矩阵. 其中, r_{ij} 表示 s_i 和 s_j 之间的约束关系, 定义为

$$r_{ij} = \begin{cases} 1, & s_i \text{ 和 } s_j \text{ 存在联盟关系} \\ -1, & s_i \text{ 和 } s_j \text{ 存在互斥关系} \\ 0, & \text{其他} \end{cases}$$

定义 5 设 $S_o = (x_1, x_2, \dots, x_m)$ 是关于 P_c 的一个服务选择方案, A 是 S_{cw} 的关联矩阵, $I_{QoS}(S) = (C(S), T(S), R(S))$ 为 S_o 的 I_{QoS} , 服务优化选择问题的形式化表示为: $\min C(S), \min T(S), \max R(S)$; 且必须满足如下条件: 当 $r_{ij} = 1$ 时, $x_i = x_j$; 当 $r_{ij} = -1$ 时, $x_i \neq x_j$ 或 $x_i = x_j = 0$.

$$\begin{aligned} \text{s.t.} \quad & x_i = x_j, \quad \text{当 } r_{ij} = 1; \\ & x_i \neq x_j \text{ 或 } x_i = x_j = 0, \quad \text{当 } r_{ij} = -1 \end{aligned}$$

表 1 服务选择方案 S_o 的总体 QoS 计算

Tab.1 Integrated computing of QoS for a selection solution

QoS 属性	顺序结构	并行结构	选择结构
服务费用	$C(S) = \sum_{i=1}^w C(s_i)$	$C(S) = \sum_{i=1}^w C(s_i)$	$C(S) = \sum_{i=1}^w p_i C(s_i)$
响应时间	$T(S) = \sum_{i=1}^w T(s_i)$	$T(S) = \max_{i=1, \dots, w} (T(s_i))$	$T(S) = \sum_{i=1}^w p_i T(s_i)$
可靠性	$R(S) = \prod_{i=1}^w R(s_i)$	$R(S) = \prod_{i=1}^w R(s_i)$	$R(S) = \prod_{i=1}^w p_i R(s_i)$

2 基于多目标遗传算法的 Web 服务优化选择

定义 5 中的多个 QoS 优化目标之间通常是相互制约的, 最优解往往有无穷个, 如何在最优解集中求

出一组 Pareto 最优解供决策者选择, 十分重要.

2.1 Pareto 最优解

定义 6^[16] 对于 $\min f(X) = [f_1(X), f_2(X), \dots, f_n(X)]$, 设其定义域为 $\Omega, X^* \in \Omega$, 若不存在 $X \in \Omega$, 使得 $f_i(X) \leq f_i(X^*) (i = 1, 2, \dots, n)$, 则称 X^* 为该问题的 Pareto 最优解 (非劣解).

对于多目标优化问题, Pareto 最优解一般不是一个, 而是一组 Pareto 最优解的集合, 彼此无优劣之分, 为 Pareto 最优解集. 理论上, Pareto 最优解集中的任意一个解都可能成为“最优解”, 这取决于决策者的偏好. 而尽可能求出逼近 Pareto 最优解集的一个子集是获得令决策者满意的折衷解的基础.

现采用多目标遗传算法求解多目标优化问题的 Pareto 最优解集合, 逐一介绍遗传算法涉及的染色体编码、目标函数、遗传算子和遗传策略的解决方案.

2.2 染色体编码方式

根据定义 2 给出的服务选择方案 $S_o = (x_1, x_2, \dots, x_m)$, 每个染色体 $C = c_1 c_2 \dots c_m$ 表示一个可能满足用户请求的解决方案. 对于任意的 $1 \leq i \leq m$, $c_i = x_i$. 染色体中的每一个基因对应于一个具体的服务, 它有两种取值: 取 1 表示此服务被选中, 0 表示没有选中. 染色体编码采用长度为 m 的二进制串.

但是, 由于服务选择过程中存在着多个约束条件, 并非每个染色体都是可行的服务选择方案, 因而需要判定染色体的可行性. 主要判定方法是: 如果染色体 $C = c_1 c_2 \dots c_m$ 同时满足定义 2 和定义 5 中的约束条件, 那么 C 为可行染色体, 即个体.

2.3 目标函数

在服务选择时, 每个服务选择方案的性能都是通过其 I_{QoS} 来度量. 所以, 本算法中的优化目标函数即是定义 5 中给出的服务优化选择问题的三个 QoS 优化函数.

2.4 遗传算子

遗传算法主要包括选择、交叉和变异三个遗传算子. 这里采用的选择算子是排序选择方法^[17]. 首先, 根据“Pareto 最优个体”的概念来对种群中所有个体排序: 设当代种群中所有非劣个体的等级为 1, 从种群中暂时移去; 余下的种群中将所有非劣个体的等级设为 2, 也暂时移去. 如此循环直至所有个体都被赋值. 最后, 将个体等级值赋为适应度值. 进化过程中依据这个排列次序来选择运算, 使排在前面的 Pareto 最优个体有更多机会遗传到下一代种群.

传统的交叉算子是根据交叉概率随机选择交叉

点,然后交换交叉点前的串,形成两个新个体.但是,由于对染色体有一定的限制,例如对应于某个抽象服务的所有具体服务中最多只能有一个被选中,因而交叉点只能发生在抽象服务之间.此时,把对应相同抽象服务的具体服务看作一个不可分的子串,交叉点只能在子串之间,而不能选择子串内的点.

以图2为例,两个个体 C_1 和 C_2 (为了便于区分, C_1 用黑色粗体表示),确定交叉点后,生成2个新的个体 C'_1 和 C'_2 .虽然, C'_1 是由 C_2 的前段子串和 C_1 的后段子串组成, C'_2 由 C_1 的前段子串和 C_2 的后段子串组成.

同样,此处的变异算子也不同于传统的变异算子,变异时随机选择某一个子串变异,变异采用传统的变异方法,如图3所示.图中,个体 C_1 中的黑色粗体部分表示选中的变异之串.

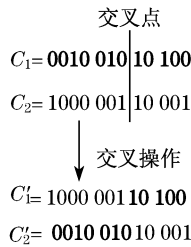


图2 交叉算子

Fig.2 Crossover operator

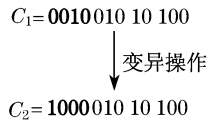


图3 变异算子

Fig.3 Mutation operator

2.5 遗传策略

为保证算法收敛性,并克服算法陷入局部最优,采用如下两种策略:最优保存策略和小生境技术.

(1)最优保存策略 每次保留当前代中等级数为1的个体.此策略保证当前种群中适应度最高的个体不参与交叉和变异运算,用来替代下一代中适应度最低的个体.这样,从算法开始到结束生成的所有个体适应度最大的个体能够保存到最后一代,有效保证算法的收敛,但可能导致局部最优个体不被淘汰而快速扩散.

(2)小生境技术^[17] 在标准遗传算法中引入小生境概念,限制相同个体或类似个体的数量,确保相似个体较少的个体有更多机会遗传下一代,可以有效地维持种群多样性,防止算法早熟收敛.具体的实现步骤如下.

首先,为了反映个体编码串之间的相似程度,定义了如下共享函数:

$$F_s(C_i, C_j) = m - \sum_{1 \leq k \leq m} |c_{ik} - c_{jk}| \quad (1)$$

式中: C_i 和 C_j 是种群中的任意两个个体; c_{ik} 是个体

C_i 是第 k 位; c_{jk} 是个体 C_j 是第 k 位; m 是染色体的长度.然后,为了度量每个个体在种群中共享程度,定义个体的共享度

$$M_s(C_i) = \sum_{1 \leq j \leq N, j \neq i} F_s(C_i, C_j) \quad (2)$$

并依据下式调整种群中各个个体 C_i 的适应度:

$$F(C_i) = F(C_i)/M_s(C_i) \quad (3)$$

进而依据调整后的新适应度来选择运算,就能够限制种群中相似个体的数量,从而维护了群体的多样性,创造出小生境的进化环境.

在完成了算法的几个重要设计基础上,给出本服务优化选择的多目标遗传算法如下:

开始

遗传参数初始化(最大迭代次数 n ,交叉概率 p_c ,变异概率 p_m ,选择概率 p_s)随机产生 N 个个体(可行的染色体),形成初始种群.对迭代次数 $1 \sim n$,操作如下:

- 目标函数计算 对种群中每个个体,分别计算 I_{Qos} 子目标函数.
- 排序 利用Pareto最优解的定义,对种群中的个体排序,并将排序等级作为原始适应度.
- 小生境策略 计算种群中每个个体共享度,并根据式(3)调整其适应度.
- 选择 根据选择概率 p_s ,从种群中随机选择 $N/2$ 个个体,得到临时种群.
- 遗传操作 对临时种群中的个体,采用交叉、变异操作,产生新 $N/2$ 个个体,与临时种群中的个体一起进入下一代种群.
- 最优保存策略 保留当前代中等级数为1的个体,直接进入下一代种群.

结束.

3 仿真分析

本节用仿真实验模拟本算法,通过实验结果分析算法的性能和优化效果.下面给出设计的模拟实验方案:模拟实验环境和实验结果.

3.1 仿真环境

硬件环境:CPU为Intel Pentium IV 2.4 GHz,内存为1 GB,操作系统为Windows XP,算法实现工具为VC.

实验数据:鉴于目前没有相关的标准平台和标准测试数据集,采用随机生成的模拟数据作为测试用例.各服务之间的流程关系如图2所示.共有7个

抽象服务,模拟实验中共生成了 3 组测试数据集,其中每个抽象服务对应的具体服务数目分别为 5, 10 和 15, 即 35, 70 和 105. 每个具体服务的 I_{QoS} 的范围分别是: 服务费用 0~300 元; 响应时间 0~10 s; 可靠性 0~100%. 服务关联矩阵记录了任两个具体服务之间的关系, 是通过随机生成 0, 1, -1 作为矩阵中的元素. 关联度定义为: 矩阵中元素值为 1 和 -1 的个数除以矩阵元素个数.

遗传算法参数: $N=100$; $n=500$; $p_c=0.85$; $p_m=0.01$; $p_s=0.65$.

3.2 仿真结果

实验一, 选用的遗传算法有两种策略: 最优保存策略和小生境策略. 首先通过实验比较了在三组不同具体服务个数情况下, 采用不同策略时算法的优化所需代数. 实验中有四种的算法. 算法 1, 没有采用以上两种策略的; 算法 2, 采用了最优保存策略; 算法 3, 小生境策略; 算法 4, 同时采用最优保存策略和小生境策略, 即本文提出的方法.

实验结果如图 4 所示, 随着具体服务数增加, 四种算法的优化时间都是单调增加的. 算法 1 所需的优化代数最长, 而算法 2 和 3 在问题规模较小的时性能相当. 算法 2 由于采用了最优保存策略, 相对于 1 有着较快的算法收敛速度. 但随着规模的增大, 算法 3 优势明显. 这是因为小生境策略在保持种群多

样性的同时, 具有很高的全局寻优能力和收敛速度. 从图 4 可以看出, 结合了两种策略优越性的算法 4 优化代数最短, 性能更好, 两种遗传策略在维护种群多样性和提高算法收敛性方面发挥了很好的作用.

实验二, 比较非劣解集中各优化指标的最优解随代数变化情况. 图 5a, b, c 给出了算法 4 在具体服务个数为 70 且关联度为 50% 时, 50 次随机实验中各优化指标对应等级为 1 的 Pareto 解集的平均分布情况. 图中纵轴上标注了单个目标的最优解. 这些图形反映出多目标遗传算法的优化和收敛过程. 由于使用了最优个体保存策略, 算法的渐进收敛速度得到了保证. 从图中可以看出, 随着优化代数的增加, 每个 QoS 优化目标都近似地收敛于其最优解.

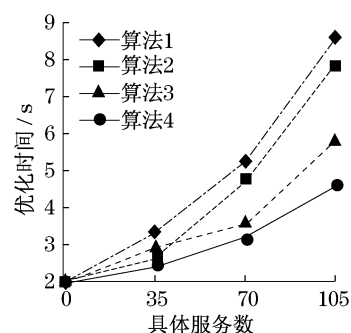


图 4 各算法优化时间比较

Fig. 4 Comparison of optimization time among different algorithms

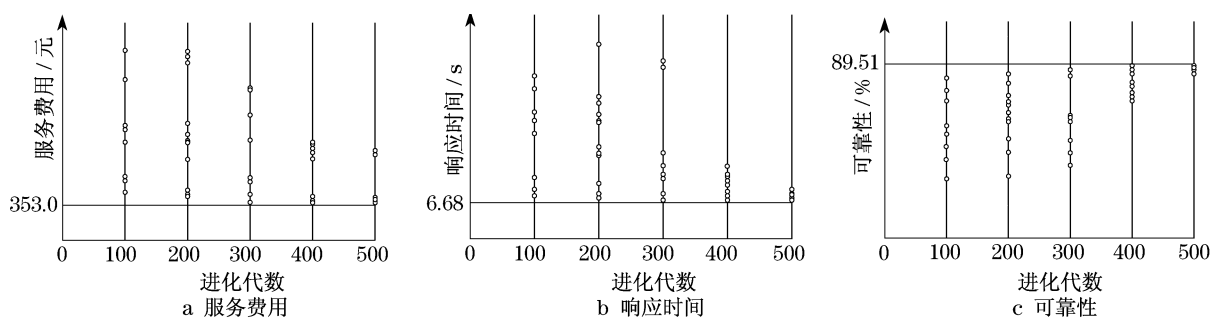


图 5 各优化指标随代数的变化结果

Fig. 5 Variation results of the optimized indexes

实验三, 算法 4 和穷举最优方法比较. 穷举法是通过列举出所有可能的解, 计算出符合单个目标最佳的目标函数作为穷举最优值. 但实际上, 枚举出的符合某个目标最优的个体往往并非是满足其他目标最优的.

算法 4 在实验结果中得到的是一组非劣解集, 这个集合中任意两个个体之间不存在优劣之分. 表 2 给出了算法在 500 代内获得的非劣解集中单个目标满足最优所占的比例. 此处的最优解是指非劣解集

内存在某优化指标的穷举最优值, 百分率是针对 50 次实验结果的统计值.

表 2 获得最优解比率

Tab. 2 Percentage of optimal solution %

服务数	关联度/%	服务费用/元	响应时间/s	可靠性/%
35	30	100(426.5)	96(9.37)	100(84.26)
70	50	97(353.0)	94(6.68)	98(89.51)
105	80	93(326.0)	91(5.72)	92(93.38)

注: 括号内数字为穷举法取得的实际最优值.

从表2中看出,本算法在500代内找到各优化指标的概率均在91%以上,因此,用来解决Web服务选择中的多个QoS优化问题是可行的.而且,由于遗传算法的并行处理特征,使其具有较好的空间搜索能力^[13].该方法可扩展至更多的服务个数的情况.

4 结语

服务优化选择问题的分层模型能近似地收敛于各指标的最优值,同时能在有限进化代数内产生一组Pareto最优的服务选择方案.目前,本算法与其他用于Web服务优化选择的不确定性算法(粒子群算法、蚁群算法等相关方法)的性能比较和分析工作正在开展,下一步将结合一些实际应用检验本方法.

参考文献:

- [1] Brahim Medjahed, Athman Bouguettaya, Ahmed K Elmagarmid. Composing web services on the semantic web[J]. The VLDB Journal, 2003, 12(4): 333.
- [2] Li Li, Liu Dongxi, Athman Bouguettaya. Semantic weaving for context-aware Web service composition[C]//Proceedings of the International Conference on Web Information Systems Engineering. Poznan: Springer LNCS 5802, 2009: 101-114.
- [3] Hwang S Y, Lim E P, Lee C H, et al. On composing a reliable composite Web service: a study of dynamic Web service selection[C]//Proceedings of the IEEE 2007 International Conference on Web Services. Salt Lake City: IEEE Computer Society Press, 2007: 184-191.
- [4] Weise T, Bleul S, Comes D, et al. Different approaches to semantic Web service composition[C]//Proceedings of the Third International Conference on Internet and Web Applications and Services. Athens: IEEE Computer Society Press, 2008. 90-96.
- [5] Zeng Liangzhao, Benatallah Boualem, Ngu Anne, et al. QoS-aware middleware for web services composition[J]. IEEE Transactions on Software Engineering, 2004, 30(5): 311.
- [6] LI Zhen, YANG Fangchun, SU Sen. QoS-aware semantic web service composition with uncertainties[J]. Chinese Journal of Electronics, 2008, 17: 703.
- [7] ZHANG Liangjie, LI Bing. Requirements driven dynamic services composition for web services and grid solutions[J]. Journal of Grid Computing, 2004, 2(2): 121.
- [8] CAO Lei, LI Minglu, CAO Jian. Cost-driven web service selection using genetic algorithm [C]//Proceedings of the Workshop on Internet and Network Economics. Hong Kong: Springer, 2005: 906-915.
- [9] ZHENG Huiyuan, ZHAO Weiliang, YANG Jian, et al. QoS analysis for Web service composition[C]//Proceedings of the IEEE International Conference on Service Computing. Bangalore: IEEE Computer Society Press, 2009: 235-242.
- [10] 孙萍, 蒋昌俊. 利用服务聚类优化面向过程模型的语义Web服务发现[J]. 计算机学报, 2008, 31(8): 1340.
- [11] SUN Ping, JIANG Changjun. Using service clustering to facilitate process-oriented semantic Web service discovery[J]. Chinese Journal of Computers, 2008, 31(8): 1340.
- [12] Canfora Gerardo, Penta Massimiliano Di, Esposito Raffaele, et al. A lightweight approach for QoS-aware service composition [C] // Proceedings of the 2nd International Conference on Service Oriented Computing. New York: ACM, 2004: 36-47.
- [13] Holland J. H. Adaptation in natural and artificial systems[M]. Ann Arbor: University of Michigan Press, 1975.
- [14] 崔逊学, 林闯. 多目标进化算法的研究与进展[J]. 模式识别与人工智能, 2003, 16(3): 306.
- [15] CUI Xunxue, LIN Chuang. A review of the research on multi-objective evolutionary algorithms[J]. Pattern Recognition and Artificial Intelligence, 2003, 16(3): 306.
- [16] Martin D, Burstein M, Hobbs J. OWL-S: Semantic Markup for Web Service. [EB/OL]. [2004: 11-12]. <http://www.w3.org/Submission/OWL-S>.
- [17] ZHOU Chen, CHIA Liangtien, LEE Busung. DAML-QoS ontology for web services[C]//Proceedings of the International Conference on Web Services. San Diego: IEEE Computer Society, 2004: 472-479.
- [18] Van Veldhuizen D A, Lamont G B. Multiobjective evolutionary algorithms: analyzing the state of the art[J]. IEEE Transaction on Evolutionary Computation, 2000, 8(2): 125.
- [19] Goldberg D E, Richardson J. Genetic algorithms with sharing for multimodal function optimization[C]//Proceedings of the 2nd International Conference on Genetic Algorithms and Their Applications. [S.l.]: Lawrence Erlbaum, 1987: 41-49.