

基于全局 QoS 和免疫粒子群算法的 WEB 服务组合研究

肖强华¹, 宁 丹², 廖 颖³

(1. 南华大学 数理学院, 湖南 衡阳 421001; 2. 南华大学 计算机科学与技术学院, 湖南 衡阳 421001;
3. 重庆邮电大学 软件学院, 重庆 400065)

摘 要:针对现有 web 服务组合在可靠性和效率方面存在缺陷,提出了一种基于全局 QoS 和免疫粒子群算法的 WEB 服务组合优化的策略,给出了构建 Web 服务及选择最佳服务的方法,有效的克服了传统方法利用穷举计算及遗传算法解决此类问题的局限性,实验结果表明,该方法的服务组合时间短,且服务组合效率高,对解决 WEB 服务组合这样的 NP 难题具有非常好的效果.

关键词:Web 服务组合;服务选取;粒子群算法

中图分类号:TP393 **文献标识码:**B

Research on Web Services Composition Based on GlobalQoS and Particle Swarm Optimization

XIAO Qiang-hua¹, NING Dan², LIAO Ying³

(1. School of Mathematics and Physics, University of South China, Hengyang, Hunan 421001, China;
2. School of Computer Science and Technology, University of South China, Hengyang, Hunan 421001, China;
3. Software College, Chongqing University of Posts and Telecommunications, Chongqing 400065, China)

Abstract: Based on the defects in the reliability and efficiency in existing web service combination, a strategy of Optimization WEB services based on global QoS and immune Particle Swarm algorithm is presented in this paper. Further, we give a best method of building and selecting Web services, which will overcome the limitations of brute-force computation and genetic algorithms in the traditional method effectively. Our experimental results show that the time of combination of service is short, and the efficient combination of service is higher, which is good to solve the NP problem such as WEB services portfolio.

key words: web services combination; services selection; particle swarm optimization

0 引 言

Web 服务是一种自包含、自描述、模块化的程序,通过采用 WSDL UDDI 和 SOAP 等标准和协议,具有高度的互操作性、跨平台性和松耦合的特点.然而单一的 Web 服务功能毕竟有限,难以满足实际应用中的需求.因此有必要将单个的 Web 服务组合起来,其目的在于通过单一服务的彼此协同合作来满足实际需求.这就促使 Web 服务组合研究的兴起.为了充分利用已有的 Web 服务,近年来国内外学术界和行业界围绕着服务组合开展了大量的研究工作,并发布了一系列用来描述 Web 服务组合的语言,如:微软公司的 XLANG, IBM 公司的 WSFL, BEA 公司的 WSCI, IBM、微软和 BEA 公司联合推出的 BPEIAWS, 乔治亚大学提出的 DAML2S 等.这些研究在 Web 组合服务的功能实施方面取得了很大的成就,然而组合服务的质量 QoS(Quality of Service)却往往被忽略,如何提高组合服务的质量是一个非常具有挑战性的问题,它对于 Web 服务组合的成败起着决定性的作用.由于穷举计算法的局限性,它需要在计算出所有可能解的情况下才能得出最优解.因此,采用

穷举计算的组合优化方法存在扩展性差、计算量相对较大的弊端^[1-3].

WEB 服务具有很多优点,如大粒度、松耦合以及后期绑定机制.随着 WEB 服务数量的增长,不同的 WEB 服务提供商可提供具有相同功能且可替换的不同 WEB 服务,人们可根据这些 WEB 服务的服务质量来选择适合自己的 WEB 服务.基于 QoS 的 WEB 服务组合问题是 NP 难题,仿生优化算法可以很好地解决这类问题.用遗传算法来解决 WEB 服务组合问题并且取得了较好的结果,但是与遗传算法相比,粒子群算法具有很多优点,例如,收敛速度快、容易实现等.与此同时 WEB 服务组合也要求处理速度必须快,特别是实时交互系统而言,因此粒子群算法在 WEB 服务组合问题上具有很大的优势^[4-6].

1 WEB 服务组合的 QoS 计算

服务质量是组合 WEB 服务的评价标准,组合 WEB 服务质量可由组合服务的基本服务的质量来描述,本文中服务质量的计算模型如表 1 所示,分别给出了三种类型:顺序类型、分支类型和循环类型.

表 1 服务组合 QoS 的计算模型
Table 1 Calculation modeling of QoS for service composition

QoS 属性	顺序执行活动	选择执行活动	循环执行活动
服务费用(C)	$C(s) = \sum_{i=1}^m C(i)$	$C(s) = \sum_{i=1}^m p_i c(i)$	$C(s) = k \times C(s)$
响应时间(T)	$T(s) = \sum_{i=1}^m T(i)$	$T(s) = \sum_{i=1}^m p_i T(i)$	$T(s) = k \times T(s)$
可用性(A)	$A(s) = \prod_{i=1}^m A(i)$	$A(s) = \sum_{i=1}^m p_i A(i)$	$A(s) = A(s)^k$
可靠性(R)	$R(s) = \prod_{i=1}^m R(i)$	$R(s) = \sum_{i=1}^m p_i R(i)$	$R(s) = R(s)^k$

如表 1 所示^[7-10],假设分支类型共有 m 个分支,对应的执行概率分别是 $P_1, P_2, \cdots, P_m$, 且满足 $\sum_{i=1}^m P_i = 1$. 如图 1 所示,假设组合 WEB 服务定义如下:

$$W = (w_1, w_2, \cdots, w_m). \tag{1}$$

同时假设每个基本服务 W_i 有一系列的候选基本服务,例如 $W_i (1 \leq i \leq m)$ 有 $n_i (1 \leq i \leq m)$ 个

候选 WEB 服务,这些候选 WEB 服务功能相同并且可相互替换. WEB 服务组合的目的就是在用户和服务提供商的条件约束下寻找一条最佳的 WEB 服务组合方式,使得组合服务的服务质量最好.图 1 中每个候选 WEB 服务的服务质量定义为 $Q_{i,j} (1 \leq i \leq m, 1 \leq j \leq n_i)$. $Q_{i,j}$ 是综合多种质量属性的综合评价结果.

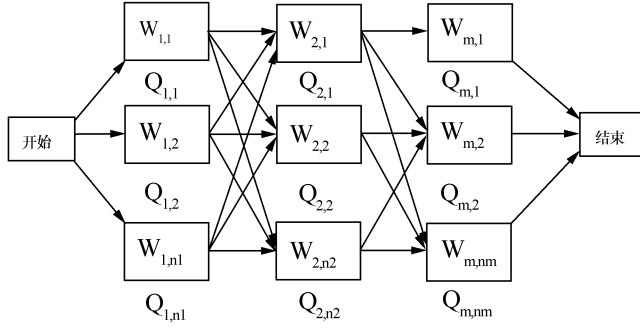


图 1 基于 QoS 的 WEB 服务组合问题

Fig.1 Web services composition problem based on QoS

2 免疫粒子群优化算法

免疫粒子群优化算法是将免疫的重要特性及免疫机理应用在粒子群算法中,在保留原粒子群算法优良特性的前提下,力图有选择、有目的地利用相关免疫理论求解优化问题,以此抑制其优化过程出现的粒子群后阶段退化现象.基于免疫算子的粒子群算法基本流程如下:

1) 确定参数数值. 学习因子 c_1, c_2 和粒子的惯性权重 w , 粒子(抗体)群体个数 N , 进化代数 K .

2) 初始化. 针对待求问题的特征,从记忆库中搜寻该类问题的记忆粒子(抗体)或随机产生 N 个粒子(抗体) x_i , 及其“飞行”速度 v_i , ($i=1, 2, \dots, N$), 形成初始粒子(抗体)群体 X_0 .

3) 产生免疫记忆粒子(抗体). 计算当前粒子(抗体)群体 X_k ($k=1, 2, \dots, K$) 中粒子(抗体)的适应度及 P_i ($i=1, 2, \dots, N$), P_g 并把 P_g 作为免疫记忆粒子(抗体)存入记忆库中,判断是否满足结束条件,若满足结束条件,则停止运行并输出结果,否则继续.

4) 新粒子的生成. 新粒子的产生主要基于以下两种方式:(1) 粒子群优化算法公式(1), 公式(2)产生 N 个新粒子;(2) 随机产生 N 个新粒子.

5) 基于浓度的粒子选择. 用计算步骤 4) 中生成的 $N+M$ 个新粒子的选择概率,依概率大小选择 N 个粒子形成粒子群 X'_k ($k=1, 2, \dots, K$).

6) 粒子群更新. 将记忆库中的免疫记忆粒子替换粒子群 X'_k 中适应度较差的若干粒子,开成新一代粒子群体 X_{k+1} ($k=1, 2, \dots, K$), 然后,转步骤 3).

在粒子群体更新过程中,总是希望适应度高的粒子被保留下来,但如果此类粒子过于集中,即浓度过高,则浓度过高,则很难保证粒子的多样

性,很容易使算法陷入局部极优,而丢失那些适应度较差但却保持着较好的进化趋势的粒子. 计算浓度的公式:

$$D(x_i) = \frac{1}{\sum_{j=1}^{N+M} |f(x_i) - f(x_j)|}, \quad i = 1, 2, \dots, N+M. \quad (2)$$

选择概率为:

$$P(x_i) = \frac{\frac{1}{D(x_i)}}{\sum_{i=1}^{N+M} \frac{1}{D(x_i)}} = \frac{\frac{1}{\sum_{j=1}^{N+M} |f(x_i) - f(x_j)|}}{\sum_{i=1}^{N+M} \frac{1}{\sum_{j=1}^{N+M} |f(x_i) - f(x_j)|}}, \quad i = 1, 2, \dots, N+M, \quad (3)$$

式中 x_i 和 $f(x_i)$ 分别表示第 i 个粒子抗体及其适应度函数值.

3 基于 QoS 和免疫粒子群算法的 WEB 服务组合

3.1 适应度函数设计

适应度函数设计如下:设服务的 QoS 由 p 个 QoS 属性来刻画,即 $\{Q_1, Q_2, \dots, Q_m\}$. 根据 QoS 计算模型,可以得出基本服务评价函数 f 的计算公式:

$$f = \sum_{i=1}^p (w_i \times Q_i) \quad (4)$$

其中 w_i 为组合服务中相应 QoS 属性的权值,表示用户对该 QoS 属性的关注程度,并且 $0 \leq$

$w_i \leq 1$, $\sum_{i=1}^p w_i = 1$. Q_i 为组合服务第 i 项 QoS 属性值, m 为 QoS 属性个数. 由于不同 QoS 属性值可能会差几个数量级,而且有些 QoS 属性值越大,服务的 QoS 评价就越低,像服务价格和执行时间;有些 QoS 属性越大,服务 QoS 评价相反就越高,像可靠性和可用性. 因此对 QoS 属性值按下面公式

进行处理:

$$Q'_i = \frac{Q_i^{\max} - Q_i}{Q_i^{\max} - Q_i^{\min}},$$

(当 $Q_i^{\max} - Q_i^{\min} = 0$ 时, $Q'_i = 1$) (5)

$$Q'_i = \frac{Q_i - Q_i^{\min}}{Q_i^{\max} - Q_i^{\min}},$$

(当 $Q_i^{\max} - Q_i^{\min} = 0$ 时, $Q'_i = 1$) (6)

式(5)、式(6)中的最大值 $Q_i^{\max}$ 和最小值 $Q_i^{\min}$ 是所有组合服务当中第 i 项 QoS 属性的最大值和最小值. 在计算出组合服务的第 i 项 QoS 属性值 Q_i 后,利用公式(5)和公式(6)求出 Q_i 对应的 Q'_i . QoS 属性值和 QoS 评价成正比的利用公式(5)计算,成反比例的利用公式(6)计算. 最后组合服务的评价函数可以表示为:

$$f = \sum_{i=1}^p (w_i \times Q'_i) \tag{7}$$

其中 $0 \leq Q'_i \leq 1, 0 \leq w_i \leq 1, \sum_{i=1}^m w_i = 1$.

3.2 WEB 服务选择算法

基于 QoS 和免疫粒子群算法的 WEB 服务选择算法步骤如下:

1) 依照初始化过程,对微粒群的随机位置和速度进行初始设定;设定粒子数为 n ,迭代次数为 $n_{\max}$,加法运算概率常数为 c_x ,变异算子概率常数为 c_d ;计算每个粒子的适应值,根据当前位置由式(7)计算每个粒子的适应值 $f_i (1 \leq i \leq n)$.

2) 对每个粒子,将其适应值 f_i 与所经历过的最好位置 $PC_{i,best} (1 \leq i \leq n)$ 的适应值 $PL_{i,best} (1 \leq i \leq n)$ 进行比较,若较好, $PL_{i,best} = f_i$ (注:初始化粒子 X_i 的个体极值位置 $PC_{i,best} = X_i, PL_{i,best} = f_i$).

3) 对每个粒子,将其适应值 f_i 与全局所经历的最好位置 $PL_{i,best}$ 的适应值 $PC_{i,best}$ 进行比较,若最好, $PC_{i,best} = f_i$ (注:初始化 $GL_{best} = PL_{i,best}$, 如果 $f_i = \max\{f_k\} (1 \leq k \leq n)$; 初始化 $GC_{best} = PC_{i,best}$, 如果 $GL_{best} = PL_{i,best}$).

4) 计算粒子群多样性 $\bar{d}_i$, 若 $\bar{d}_i < c_d$, 启动变异算子,为粒子群的所有粒子随机生成一个新速度.

5) 对粒子速度和位置进行进化.

6) 若迭代次数 $k < n_{\max}$ 或者达到预先设定的阈值,则返回步骤 2).

4 仿真实验

本文将基于粒子群算法的 WEB 服务选择一同传统的遗传算法进行比较,两种算法都采用本

文提出的 QoS 模型以及适应度函数,遗传算法个体选择采用轮盘赌法,并且保证组合服务中基本服务有相同的候选基本服务. 运行条件为种群 500,最大迭代次数 $n_{\max} = 500, c_1 = 0.2, c_2 = 0.3, c_x = 0.5, c_c = 0.2$,两种算法分别运行 60 次,记粒子群算法为 P ,一维编码遗传算法为 G ,得到的运行结果如表 2 所示.

表 2 相同迭代次数下最大适应度平均值以
及时间平均值对比结果

Table 2 Comparison of the average of maximum fitness
and time in the same iterations

组合服务中 基本服务数量	最大适应度平均值	时间/ms
10	0.183:0.191	803:792
25	0.065:0.061	2554:2785
30	0.054:0.042	2908:3204

从上面的仿真数据可以看出,基于粒子群算法的 WEB 服务组合比基于遗传算法的 WEB 服务组合用时明显减少,这是因为粒子群算法是根据群体最优进行优化,而遗传算法是群体平行地向前优化. 另外可以看出随着组合服务中基本服务数量的增加,基于粒子群算法的最大适应度要明显比基于遗传算法的解好. 若两种算法采用相同的迭代次数, $n_{\max} = 500$, 而表 3 为在相同阈值的情况下,两种算法的耗时比较.

表 3 相同阈值情况下用时比较 (P:G)

Table 3 Comparison of the used time
in the same threshold

组合服务中 基本服务数量	最大适应度 平均值(阈值)	时间/ms
10	0.191	513:792
25	0.061	2476:2785
30	0.042	4793:3204

从仿真实验可以看出,免疫粒子群算法在基于 QoS 的 WEB 服务选择中具有较好的应用特性,这为解决此类问题提供了一种新的思路,该方法的服务组合时间短,且服务组合效率高,对解决 WEB 服务组合这样的 NP 难题具有非常好的效果.

5 结束语

本文提出的基于全局 QoS 和免疫粒子群算法的 WEB 服务组合策略,对于构建基于互联网复杂组合服务的应用具有一定意义.同时也应该看到,粒子群算法虽然有收敛速度快的特点,但同时也存在易停滞早熟的现象,进一步的工作是根据服务组合的具体实际,进行基于粒子群算法的优化与改进.

参考文献:

- [1] Alrifai M, Risse T. Combining global optimization with local selection for efficient QoS-aware service composition [C] // Proceedings of the 18th International Conference on World Wide Web. New York, USA. 2009:881-890.
- [2] El Hadad J, Manouvrier M, Rukoz M. TQoS: Transactional and QoS-aware selection algorithm for automatic Web service composition[J]. Services Computing, IEEE Transactions on, 2010, 3(1):73-85.
- [3] Rosenberg F, Celikovic P, Michlmayr A, et al. An end-to-end approach for qos-aware service composition[C] // Enterprise Distributed Object Computing Conference, 2009. EDOC'09. IEEE International. Auckland, 2009:151-160.
- [4] 彭晓明, 何炎祥, 朱兵舰. 蚁群算法在 Web 服务组合中的应用[J]. 计算机工程, 2009, 35(10):182-187.
- [5] 刘冰. 基于 QoS 的服务组合关键技术研究[D]. 济南: 山东大学, 2010.
- [6] 邓水光, 黄龙涛, 尹建伟, 等. Web 服务组合技术框架及其研究进展[J]. 计算机集成制造系统, 2011, 17(2):404-412.
- [7] 陈彦萍, 李增智, 郭志胜, 等. Web 服务组合中基于服务质量的服务选择算法[J]. 西安交通大学学报, 2006, 40(8):897-900.
- [8] 陈明. 分布计算应用模型[M]. 1 版. 北京: 科学出版社, 2009.
- [9] 王俊丽, 柳先辉, 卫刚. 基于服务质量的 Web 服务优化选择算法及仿真[J]. 同济大学学报(自然科学版), 2011, 39(4):599-605.
- [10] 李金忠, 夏洁武, 唐卫东, 等. 基于 QoS 的 Web 服务选择算法综述[J]. 计算机应用研究, 2010, 27(10):3622-3627.