

基于分布感知的 Web 服务组合方法研究

蔡红¹, 孙宏彬²

CAI Hong¹, SUN Hong-bin²

1. 上海行健职业学院, 上海 200072

2. 东华大学 信息科学与技术学院, 上海 201620

1. Shanghai Xingjian College, Shanghai 200072, China

2. College of Information Sciences and Technology, Donghua University, Shanghai 201620, China

E-mail: hbsun@mail.dhu.edu.cn

CAI Hong, SUN Hong-bin. Study on composition of Web service based on distributed awareness. *Computer Engineering and Applications*, 2008, 44(5): 143-145.

Abstract: A distributed awareness method of Web service composition is proposed. The mobile agent represents Web service and constitutes an autonomic unit. The Web service composition is achieved by the evolution of the mobile agent based on the semantic negotiation. Genetic composition algorithm has the limitation of dealing with complex route and the definition of penalty function. Thus, with relation matrix that can express the relation of Web services and QoS matrix, the optimization of the Web service composition is translated into a constrained optimization problem. A new nonparametric penalty function for evolutionary optimization is introduced to solve the problem. The simulation results show the enhancement of the performance and fitness.

Key words: mobile agent; constrained optimization; Web service; composition

摘要: 提出了一种分布感知的 Web 服务组合方法。移动 Agent 代理 Web 服务, 构成一个自治的感知单元, 移动 Agent 通过语义的协商确认匹配关系, Web 服务组合可以通过移动 Agent 的演化完成。常见的遗传组合算法存在着复杂路径表示和罚函数难于确定的局限性, 采用感知基础上的服务关系矩阵和服务相对质量的矩阵, 最优服务组合的建立转化为带约束的最优化问题, 并给出不带参数的罚函数动态演化算法解决方案, 仿真表明这种方法可以提高适应度和性能。

关键词: 移动 agent; 约束优化; Web 服务; 组合

文章编号: 1002-8331(2008)05-0143-03 **文献标识码:** A **中图分类号:** TP391

1 引言

Web 服务采用 WSDL、UDDI 和 SOAP 等基于 XML 的标准和协议, 具有互操作性、跨平台性和松耦合的特点, 语义, 特别是本体的提出, 为 Web 服务的自动发现、合成提供了一种新的思想。在语义 Web 服务的方式下, 服务的发现、执行和合成不需要由人来完成, 而是可以由 Agent 来自动完成用户提出的复杂任务。服务数量的增加使集中管理变得非常困难, 同时大量的服务更新、移动服务、普适计算使服务一直处在变化中, 所以 Web 服务合成和动态管理是非常复杂的。典型系统如: eFlow 系统^[1], DynFlow^[2]和 SELF-SERV^[3]。SHOP2 服务合成方法^[4], 在互联网中, 这种系统会很快形成瓶颈。Agent 被成功用于服务发现、组合, 并应用于下一代的 Web 服务动态管理框架^[5], Hai^[6]总结出其具有的主要特征: 服务访问控制的分布性、自动的合成策略、服务变化的自适应、自扩展及对服务合成的自调整等。

Web 服务组合主要涉及两个问题: 一个是服务之间的关系

确定, 另一个是基于 QoS 服务选择。目前, Web 服务组合大多采用基于 QoS 的一维染色体编码和遗传算法(GA)来解决^[7,8], 这种方法可以描述和实现服务简单流程等情况。但对于复杂的用户需求, 特别是带有多条件选择、循环等状况的 Web 服务组合问题难于处理。同时一维染色体采用的 GA 算法, 一般是通过带参数罚函数法解决服务质量的约束问题, 但这种方法难于准确地确定参数, 往往要通过大量的试验, 只在解决特定问题上有优势。

本文提出了一种基于移动 Agent 协商的 Web 服务组合方法。移动 Agent 具有消息匹配接口代理 Web 服务, 移动 Agent 通过语义的感知准确地描述各个服务之间的关系, 服务组合可以描述为一种带约束型的多属性优化问题。针对一维染色体编码存在的问题, 采用了非参数惩罚函数演化算法^[9-11]处理约束条件, 在演化搜索中, 采用可行解和不可行解的适当比例生成群体逼近最优解, 解决了常规带参数罚函数带来的问题, 仿真表明这种方法具有良好的性能和适应性。

基金项目: 国家自然科学基金(the National Natural Science Foundation of China under Grant No.60534020); 国家教育部新世纪人才支持计划(the New Century Excellent Talent Foundation from MOE of China under Grant No.NCET-04-415); 教育部科技创新工程重大项目培育资金项目(No.706024); 上海市国际科技合作基金项目(No.061307041)。

作者简介: 蔡红(1969-), 男, 讲师, 主要研究领域为计算机网络, 数据库; 孙宏彬(1969-), 男, 博士, 副教授, 主要研究领域为智能服务, 网络智能。

收稿日期: 2007-05-31 **修回日期:** 2007-07-30

2 基于语义的服务感知设计

移动 Agent 是一种自治的单元(代理 Web 服务),它由属性、行为和功能组成。属性描述关于移动 Agent 的信息,包括全局标识、服务状态描述集、过程状态和组标识。功能实现移动 Agent 代理的服务运算,包括有语义匹配、条件决策、关系数据记录、服务质量评估等。行为实现与服务相关的处理操作,包括有信息发布、数据接口操作、QoS 计算、匹配计算。移动 Agent 是网络应用中的服务单元。它可以接受服务请求消息,通过消息匹配的方法感知并完成演化构建。

Web 服务组合是多个服务在不同结构下的组合,存在多个具有不同 QoS 属性的服务组合方案。为了移动 Agent 演化进行,这里提出移动 Agent 的感知设计,Web 服务的关系描述通过移动 Agent 的感知协商自治完成。移动 Agent 关系是通过代理的 Web 服务的匹配能力计算得到的。设两个移动 Agent 代理的服务为:WS1 和 WS2, 移动 Agent 的服务匹配定义为代理的 Web 服务匹配消息的能力,表示为 $MS_{bio-entity}^{res}$, 设 WS 为移动 Agent 代理的 Web 服务,WM 为另一个移动 Agent 服务的输出消息,其匹配能力 $MS_{bio-entity}^{res}(WM,MS)$ 的计算是由 WM 和 WS 的接口概念匹配度决定。进一步表示为: $WM \Rightarrow WM(C_1, C_2, C_3, \dots, C_i)$ 和 $WS \Leftrightarrow WS(C'_1, C'_2, \dots, C'_j)$, 其中 C_i 和 C'_j 分别代表组成消息接口服务描述的相关概念。这里首先给出概念相似度^[12] $SM(C_1, C_2)$ 的计算公式为:

$$SM(C_1, C_2) = \begin{cases} 1 & C_1=C_2 \\ SM_{con}(C_1, C_2) & C_1 < C_2, C_1 > C_2 \\ 0 & C_1 \neq C_2 \end{cases} \quad (1)$$

$$SM_{con}(C_1, C_2) = \frac{\beta \times (L_1 + L_2)}{[Dis(C_1, C_2) + \beta] \times \text{Max}(|L_1 - L_2|, 1)}$$

其中: $Dis(C_1, C_2)$ 是在由 WordNet 和 HowNet 构建的概念层上的最小概念距离,这里假设概念 C_1 在 L_1 层,在 L_2 层。 β 是一个常数,在两个词的相似度为 0.5 时,可以用来代表概念间的距离。

- $C_1=C_2$ 表示两个概念相等;
- $C_1 > C_2$ 表示概念 C_1 包含概念 C_2 ;
- $C_1 < C_2$ 表示概念 C_2 包含概念 C_1 ;
- $C_1 \neq C_2$ 表示概念 C_2 和概念 C_1 无直接联系。

服务需求的匹配强度 $MS_{bio-entity}^{res}$ 表示为

$$MS_{bio-entity}^{res}(WM, WS) = \frac{1}{n} \left[\sum_{j=1}^n \text{Max}_{i=1}^m (SM(C_i, C'_j)) \right] \quad (2)$$

$$MS_{bio-entity}^{res}(WM, WS) \in [0, 1]$$

系统设定一个阈值 λ , 当 $MS_{bio-entity}^{res}(WM, WS) \geq \lambda$ 可确认移动 Agent 间有直接前后匹配的关系。

Web 服务组合的建立不仅需要语义的匹配,而且在可选服务中需确认最优的 QoS 组合,服务质量数值差别较大,为了服务质量的运算,对服务质量进行相应的标准化。Web 服务的质量定义为一个向量 $Q_{bio-entity} = (Q_1(Latency), Q_2(Cost), Q_3(Availability), Q_4(Trust))$, 这里取的主要参数为:服务时间(Latency), 可靠性(Availability), 费用(Cost), 服务信任(Trust), 通过合并候选服务的质量向量构成了一个质量矩阵 $Q_{bio-entity} = (Q_j; 1 \leq i \leq$

$n, 1 \leq j \leq 4)$, 其中行向量 Q_j 代表服务。通过 SAW^[13]方法,移动 Agent 感知协商可以对服务质量进行相对化处理,把决定服务质量的各项属性值转化为相对数值。由于 Latency, Cost 属于否定性属性,所以使用公式(3)量化,而 Availability, Trust 属于肯定性属性,因此通过公式(4)量化。

$$P_{ij} = \begin{cases} \frac{Q_j^{\max} - Q_{ij}}{Q_j^{\max} - Q_j^{\min}} & \text{if } Q_j^{\max} - Q_j^{\min} \neq 0 \\ 1 & \text{if } Q_j^{\max} - Q_j^{\min} = 0 \end{cases} \quad (3)$$

$$P_{ij} = \begin{cases} \frac{Q_{ij} - Q_j^{\min}}{Q_j^{\max} - Q_j^{\min}} & \text{if } Q_j^{\max} - Q_j^{\min} \neq 0 \\ 1 & \text{if } Q_j^{\max} - Q_j^{\min} = 0 \end{cases} \quad (4)$$

$Q_j^{\max} = \text{Max}(Q_j), 1 \leq i \leq n, Q_j^{\min} = \text{Min}(Q_j), 1 \leq i \leq n$ 。通过标准化,得到服务质量的相对数据。

3 基于服务感知的 Web Service 的组合法

为了便于移动 Agent 的演化计算描述,建立了两个矩阵(如图 1),即 Web 服务关系矩阵(B_m)和质量关系(P_{n4})矩阵。Web 服务关系矩阵(图 1)建立如下:在移动 Agent 关系矩阵中,每列表示一个移动 Agent 代表的 Web 服务, B_{ij} 代表组合的被选任务,矩阵主对角线上的位置表示服务位, B_{ii} 的取值为 0, 1, 其中 1 可以表示该服务被选入组合, 0 未选入, B_{ii} 的取值也作为演化算法中的群体选择标识。 B_j (其中 $i \neq j$) 表示任务 i 和服务 j 关系,不同取值(0, 1)描述服务间的关系是否有感知的匹配联系。服务关系矩阵编码可以描述模式结构的过程。服务质量矩阵 P_{n4} (图 1)是对移动 Agent 代理 Web 服务质量的相对度量表示,行代表每个服务,列代表质量属性。数据为相对度量值。

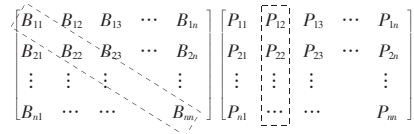


图 1 Web 服务关系矩阵(B_m)和质量关系(P_{n4})矩阵

通过 Web 服务关系矩阵的定义,可知主对角线上任务编码,代表组合服务的一个组合方案;交叉、变异操作只对矩阵主对角线上的任务位进行操作。在每次交叉、变异操作后,由于服务匹配关系,移动 Agent 会根据关系矩阵编码中的任务关系位来判断有效性,同时对服务质量进行重新计算。Web 服务组合的建立不仅满足匹配关系,而且是在满足用户服务质量要求下的服务最优组合。下面给出 Web 服务组合计算规则,通过规则可以计算出过程模型的总体服务质量的数据, $Q_{workflow}^{Global} = (Q_{Lat}^{all}, Q_{Ava}^{all}, Q_{Cos}^{all}, Q_{Tru}^{all})$, 即 $Q_{workflow}^{Global}$ 是满足用户的 QoS 的最优组合。累加统计性能指标的计算方法见表 1, 最优服务工作构建的评价函数见公式(5)。

$$S_{Bio-entity} = \sum_{k=4}^4 (Q_k^{all} * W_k) \quad (5)$$

$$g_j(x) > 0, 1 \leq j \leq q \quad h_j(x) > 0, q+1 \leq j \leq m$$

其中 $W_k \in [0, 1], \sum_{k=1}^4 W_k = 1, W_k$ 代表属性权重系数。用户服务请求中可以包含多个 QoS 的约束, x 为服务质量属性变量, $g_j(x)$

为单独服务质量属性约束条件集,这些约束相互之间是“与”的关系。 $h_j(x)$ 为多服务质量属性混合约束条件,表示用户提出的全局 QoS 要求和几项属性运算满足特定的限制要求。评价函数数值越大,说明组合服务相对质量越高。

表1 Web 服务的 QoS 计算规则

Sequence	Choice	Simple Loop	Dual Loop
Q_{Ava}^{all}	$a_1 \cdot a_2$	$\text{Min}(a_i)$	$N \cdot a_0 \cdot \text{Min}(a_i)$
Q_{Lsd}^{all}	$l_1 + l_2$	$\text{Min}(l_i)$	$N \cdot (l_0 + l_0') + \text{Min}(l_i)$
Q_{Csa}^{all}	$c_1 + c_2$	$\text{Min}(c_i)$	$N \cdot (c_0) + \text{Max}(c_i)$
Q_{Tra}^{all}	$t_1 \cdot t_2$	$\text{Min}(t_i)$	$N \cdot a_0 \cdot \text{Min}(t_i)$

显然此类问题是一种带约束型的多属性优化问题,求解这类问题时,最常用算法是采用带参数罚函数的遗传算法^[7,8],罚函数处理受限优化,对该个体依其违反约束条件的程度由惩罚函数来进行惩罚,以减小它被选择的概率。但是带参数罚函数算法的性能强烈依赖于惩罚参数的选取,若参数过小,则算法找到的解远离真正的最优解;若参数过大则会引发计算的困难,以及容易产生演化算法早熟收敛。

为了解决这个问题,本文采用非参数惩罚函数算法^[10,11],通过优化目标函数和约束条件的两目标演化算法来处理约束条件。演化算法中设 $F(x)$ 为评价函数,用下面定义的 $f_j(x)$ 构造罚函数,取为个体 x 到可行区域的距离的函数。

$$F(x) = \begin{cases} f(x), & \text{如果 } x \text{ 为可行解} \\ \text{Max}\{a \cdot f_{\min}, b \cdot f_{\max}, f(x)\} - \sum_{j=1}^q f_j(x), & x \text{ 为不可行解} \end{cases} \quad (6)$$

$$f(x) = \sum_{k=1}^4 (Q_k^{all} * W_k)$$

$$f_j(x) = \begin{cases} \text{Max}\{0, g_j(x)\}, & 1 \leq j \leq q \\ |h_j(x)|, & q+1 \leq j \leq m \end{cases}$$

$\sum_{j=1}^m f_j(x)$ 表示个体违反约束的程度,也表示个体 x 到可行区域的距离。其中 a, b 为当前群体中可行解和不可行解的百分比, $a+b=1$, f_{\min} 为当前群体可行解的函数值中的最小值, f_{\max} 为当前群体可行解的函数值中的最大值,移动 Agent 的协商和选择过程中,利用前述混合编码及策略获得最优服务组合。下面给出求解演化算法流程:

(1) 根据用户需求,按照组合模式,通过移动 Agent 的语义感知匹配关系和质量关系,建立移动 Agent 关系矩阵和质量服务矩阵 移动 Agent 群体规模为 N ,置 $t=0$;

(2) 使用 $n+1$ 个父体向量重组产生后代,这 $n+1$ 个向量形成 R_n 中的一个单形。将单形以一定的比例扩张,从扩张后的单形中随机地取一点即为一个后代,以这种多父体单形杂交子^[14]从第 t 代群体 $B(t)$ 中随机地选取 L 个父体;

(3) 设 P_m 为变异概率,个体 B 的第 i 个变量 B_i 变异为 $B_i' = B_i + N(0, \tau_i)$,其中 $N(0, \tau_i)$ 为独立的随机高斯数,均值为 0,标准差为 τ_i ,将 τ_i 取 B_i 的变化范围的上下界的平均值,产生个体集合 N ;

(4) 由式(2)计算 $L(t)+N$ 中个体的适应值,从中选择适应值大前 λ 个个体组成下一代群体 $L(t+1)$;

(5) 若满足条件,列出相应 Web 服务组合和确定移动 agent,移动 agent 感知协商结束,否则 $t=t+1$,转(2)。

4 仿真结果及性能分析

4.1 仿真设计

随机抽取了 20,40,60,80,100,120,140 个服务来评价方法性能,每个服务计算 4 个 QoS 属性值,具体属性值在一定的取值范围内随机生成。为了简化问题,将生成的测试数据随机分配给移动 Agent,用来测试系统的组合特性。仿真主要讨论这种方法的性能。本仿真试验采用了 Pentium 4 处理器的(2.4 GHz 和 512 MB RAM)的计算机。为了比较与一维编码 GA 算法性能,设定服务请求不变,采用如下的移动 Agent 数据分配机制,随机生成矩阵算法的初始种群数据,并且保证用户要求的个体存在并且数量相同;一维编码的算法采用标准遗传算法结构,相同的进化操作概率值, QoS 模型,个体选择采用轮盘赌法。仿真实验分别收集统计了对比数据与算法运行数据。

4.2 结果及分析

图 2 是一维编码和矩阵编码的响应时间,从图中可以看出在移动 Agent 任务数目比较小的时候,一维编码算法的响应时间比较短,随着任务数量的增加,矩阵编码方式响应时间增长速度较慢,在 40 个任务以后,时间明显低于一维编码方式,事实上,矩阵编码方式与一维编码相比扩大了搜索范围,矩阵编码的关系处理虽然占用部分处理时间,但关系矩阵的描述减少了无效组合,这种方式在多任务,复杂结构更具有优势。

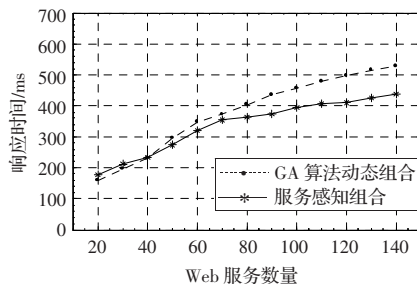


图2 服务感知组合和 GA 动态组合的响应时间比较

同时从表 2 可以看出矩阵演化算法具有更高的命中率,在 20 个任务时,两种方法成功率基本相同,在实体任务增加到 120 个的过程中,这种差距逐步增大。从上面的数据分析可知,生物实体编码矩阵的建立可以准确地描述各个服务和组合的关系,在关系建立的基础上,采用了演化算法处理约束条件,一定程度上解决了局部最优解的问题,取得较好的效果。

表2 最优解成功率比较

Web 服务数量	服务感知组合	GA 动态组合
20	100.0	100.0
40	100.0	97.2
60	100.0	96.6
80	97.4	95.2
100	95.8	92.5
120	91.5	87.8

5 结论

Web 服务是一个正在迅速发展的研究领域,本文给出 Web 服务组合问题感知演化模型,移动 Agent 代理 Web 服务,并通过关系和质量矩阵描述,使服务组合问题转化为服务匹配和动态管理问题,这种方法解决了常规 GA 方法存在的问题,仿真的特性显示这种方法的优良特性。其进一步的工作是扩充 Web 服务编码的描述功能,完善智能算法设计,提高系统计算效率,探讨其在服务中的广泛应用。