

P2P 环境下基于蚁群优化的构件组装方法

A Component Composition Method Based on Ant Colony Optimization in P2P Networks

马 华¹ 张红宇²

(1.湖南涉外经济学院 计算机学部 湖南 长沙 410205; 2.中南大学 商学院 湖南 长沙 410083)

摘 要: 如何实现复杂的 Internet 环境下构件组装的服务质量最优, 是网构软件开发面临的一个重要问题。提出了一种 P2P 环境下基于蚁群优化的构件组装方法, 该方法基于一个由领域特征簇构成的双层 P2P 网络模型(CC-P2P), 从而将构件组装的全局优化问题转化为 CC-P2P 中的路径搜索问题, 同时, 定义了综合考虑构件和链路服务质量的多维指标, 结合用户需求偏好将 QoS 值换算为路径长度, 并将其引入启发函数, 设计了一种适用于全局优化构件组装的改进蚁群算法。仿真实验证明了方法的有效性和可行性。

关键词: 网构软件 构件组装 蚁群算法 P2P

1 引言

基于 Internet 中丰富的软件服务实体, 出现了未来开放软件协同环境中的一种新型软件形态, 即网构软件(Internetware)^[1], 其构建依赖于对开放环境中各种分布、异构和自治的软件服务实体间的有效协同^[2]。构件组装是实现网构软件有效协同的基本手段之一, 而如何有效保证构件组装的最优服务质量是网构软件研究中的一个重要问题。然而, Internet 的开放、动态和难控的特性, 大大增加了构件组装优化的难度。现有研究主要集中在构件组装中单一构件的局部优化^[2,3], 在 Internet 环境下如何实现构件组装的全局优化则更加复杂。

P2P(Peer-to-Peer)网络作为 Internet 上的一种层叠网(Overlay Network)^[4], 能够实现分布式计算环境下构件资源的有效组织和管理, 从而为构件组装过程的全局优化提供有力支持。基于网构软件的领域特征分析, 本文提出了一种 P2P 计算环境下构件组装的全局优化方法, 该方法综合构件和链路服务质量的多维 QoS 指标, 以及用户的 QoS 需求

偏好等对构件组装进行建模, 并使用改进的蚁群优化算法求解。最后, 通过实验仿真和分析, 证明了该方法的有效性和可行性。

2 支持构件组装的P2P网络模型

2.1 总体结构

领域特征体现了特定系统具有的能力或特点, 是一种功能性的需求或对系统质量属性的要求^[5]。基于网构软件的领域特征分析, 本文提出了一种支持构件组装的 P2P 网络模型 CC-P2P(Component Composition P2P network), 如图 1 所示。

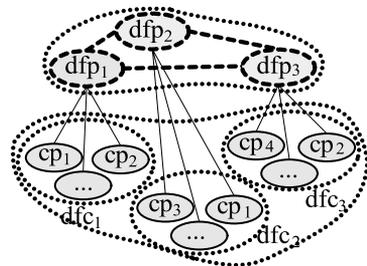


图 1 CC-P2P 的网络结构

基金项目:湖南省教育科学“十一五”规划课题(XJK08CXJ001);湖南省教育厅资助科研项目(07C425)

收稿时间:2009-03-04

CC-P2P 将 Internet 环境中的构件资源划分成不同的领域特征簇(domain feature cluster, dfc)。由领域特征对等体(domain feature peer, dfp)和普通对等体(common peer, cp)构成 CC-P2P 的上下两层结构。

2.2 CC-P2P 对构件组装的支持

CC-P2P 中, dfp 由可靠性高和性能好的节点组成, 它们构成一个结构化的领域特征网络。每个 dfp 负责一类领域特征, 为下层提供符合特定领域特征的构件的注册、发布和检索等服务。下层是由 cp 组成的 P2P 网络。cp 根据自己向外提供服务的构件的领域特征, 通过邻近的上层对等体, 注册到相应的 dfp 上, 发布的信息包括构件的功能特性以及部分 QoS 属性。CC-P2P 中, 上、下层之间在逻辑上通过领域特征簇(dfc)聚集在一起。如果某些 cp 提供了多种领域特征的相关构件, 该 cp 可注册到多个 dfc 中。如图 1 中, cp1 同时注册了 dfc1 和 dfc2, 而 cp2 同时注册了 dfc1 和 dfc3。

网构软件开发过程中, 在完成了领域特征模型分析后, 即可通过所在 cp 向 CC-P2P 网络中的各个相关簇中的 dfp 发出检索请求, 从而获取符合构件组装功能需求的构件信息。

3 构件组装问题描述

CC-P2P 网络中, 由各个 dfc 提供符合特定领域特征需求的构件信息。这样, 一个由 cp 发起装配 n 个构件的构件组装问题如图 2 所示。

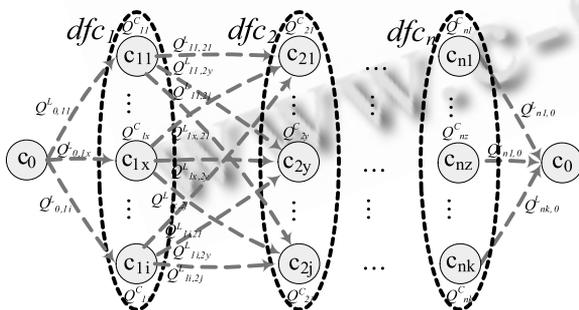


图 2 基于领域特征簇的构件组装

图 2 中, $c_{11} \sim c_{nk}$ 代表满足特定领域特征功能的构件, 它们分别由领域特征簇 $dfc_1 \sim dfc_n$ 中的相应普通对等体提供。 c_0 被称为初始装配点, 即发起构件装配过程的普通对等体。如果 c_0 选择组装了 dfc_1 中的

c_{11} (由普通对等体 cp_{11} 提供), 则 c_{11} 成为新的装配点, 继续进行下一步装配。 $QL_{i,j}$ 表示 c_i 、 c_j 间链路(i, j)的 QoS 值, 它由链路的端到端时延等因素决定。 QC_i 表示构件 c_i 的 QoS 值, 它由构件的服务费用等因素决定。如果某个普通对等体同时注册了多个领域特征簇, 即可以提供多个相应的构件, 则可能存在 $QL_{i,j}=0$ 。

位于同一个 dfc 中的构件能够提供无差别的业务功能, 但服务质量却可能存在较大差异。从而, 对多个 dfc 中构件装配路线的不同选择, 组装后的网构软件将获得不同的 QoS 值。这样, 构件组装的全局优化问题转化为一个 CC-P2P 网络中的路径搜索问题, 其目标是保证由 n 个构件组装而成的网构软件的 QoS 值最优。该问题的解空间为 $\prod_{i=1}^n |dfc_i|$ 。由于网构软件环境中基础软件资源非常丰富, 当待组装构件的数量有几十个, 各 dfc 中构件的数量有几十或上百个时, 该问题的解空间将非常庞大。因而, 在有限时间内搜索到可行的优化解, 对于 CC-P2P 中的构件组装问题非常重要。

4 基于蚁群优化的构件组装方法

蚁群算法作为一种群智能计算技术, 被广泛应用于各类复杂问题的求解^[6], 它能有效利用关联性信息提高解搜索的效率, 在动态优化方面具有突出的优势。基于此, 本文提出了一种改进蚁群算法求解构件组装的全局优化问题。

4.1 路径长度的换算

如图 2 所示, 当蚂蚁位于装配点 c_i 时, 接下去是否选择下一个 dfc 中的 c_j 作为新的装配点, 将受到链路(i, j)的 QoS 值 $QL_{i,j}$ 和构件 c_j 的 QoS 值 QC_j 的影响。本文通过对 $QL_{i,j}$ 和 QC_j 进行换算实现路径长度的度量, 并将其引入蚁群算法中的启发函数, 以支持构件组装的决策过程。

4.1.1 QoS 指标的定义

构件 QoS 定义为 $Q^L = \{Q^{ct}, Q^{cc}, Q^{ca}\}$, 其中, Q^{ct} 表示执行时间, Q^{cc} 表示服务代价, Q^{ca} 表示可用性。链路 QoS 定义为 $Q^C = \{Q^{lb}, Q^{ld}, Q^{la}\}$, 其中, Q^{lb} 表示带宽, Q^{ld} 表示端到端时延, Q^{la} 表示可用性。以上指标中, 除 Q^{ct} 和 Q^{cc} 由构件提供者指定外, 其它均由普通对等体通过网络监测获取。

为便于计算和分析, 将组装特定构件时的网络链

路和该构件合称为“链路-构件对”，其 QoS 定义为 $QoS^{LC} = \{Q^s, Q^c, Q^b, Q^a\}$ ，其中， Q^s 代表链路-构件对的响应时间，它是构件服务时间和链路时延的总和，即 $Q^s = Q^{ct} + Q^{ld}$ ； Q^c 代表链路-构件对的成本， $Q^c = Q^{cc}$ ； Q^b 代表链路-构件对的带宽， $Q^b = Q^{lb}$ ； Q^a 代表链路-构件对的可用性，它是构件可用性和链路可用性的累乘，即 $Q^a = Q^{ca} \cdot Q^{la}$ 。

设 CC-P2P 中链路总数为 nc ，通过对多维 QoS 指标进行归一化操作，得到所有链路-构件对的 QoS 矩阵 Q^{LC} ：

$$Q^{LC} = \begin{pmatrix} Q(lc_1) \\ \dots \\ Q(lc_{nc}) \end{pmatrix} = \begin{pmatrix} Q_{11}, & Q_{12}, & Q_{13}, & Q_{14} \\ \dots & \dots & \dots & \dots \\ Q_{nc1}, & Q_{nc2}, & Q_{nc3}, & Q_{nc4} \end{pmatrix} \quad (1)$$

4.1.2 基于 QoS 偏好的换算方法

构件组装时，用户对满足特定领域特征要求的构件的 QoS 需求偏好，可用四维行矩阵表示： $W^r_i = \{W_1, W_2, W_3, W_4\}$ 。 W_i 是对第 i 个待组装构件及调用链路 QoS 指标的权重偏好， $0 < W_i < 1$ ，并且， $W_1 + W_2 + W_3 + W_4 = 1$ 。

根据 QoS 偏好 W^r_i ，由式(1)得到 QLC 的加权值：

$$Q^{LC'} = \begin{pmatrix} Q(lc_1)' \\ \dots \\ Q(lc_{nc})' \end{pmatrix} \cdot W^r_i = \begin{pmatrix} Q(lc_1)' \\ \dots \\ Q(lc_{nc})' \end{pmatrix} \quad (2)$$

令 $Q^{max} = \max\{Q(lc_i)'\}$ ， $Q^{min} = \min\{Q(lc_i)'\}$ ， $0 < i < nc$ ，令 $L_{i,j}$ 为构件 c_i 和 c_j 间的“路径”长度，则 $L_{i,j}$ 定义如下：

$$L_{i,j} = \begin{cases} (Q(lc_i)' - Q^{min}) / (Q^{max} - Q^{min}), & \text{if } Q^{max} - Q^{min} \neq 0 \\ 1, & \text{if } Q^{max} - Q^{min} = 0 \end{cases} \quad (3)$$

4.2 改进的蚁群优化算法

4.2.1 初始化条件

令待组装构件的总数为 n ，蚁群中蚂蚁的总数为 m ， $\tau_{ij}(t)$ 表示时刻 t 链路 (i, j) 上的信息量。初始时刻，所有蚂蚁均聚集在初始装配点 c_0 (即普通对等体 cp_0) 上，各条链路上信息量相等，并设 $\tau_{ij}(0)$ 为一个常量。

4.2.2 路径选择方法

蚂蚁 k 从 c_0 出发，按照指定的构件组装顺序从相应领域特征簇中获取所需构件。将蚂蚁第 i 步行走时到达的领域特征簇记为 dfc_i 。在搜索过程中，蚂蚁根据当前装配点到达 dfc_i 中各构件的链路上的信息量，

及启发信息决定构件装配路线。时刻 t ，蚂蚁 $k(k=1, 2, \dots, m)$ 选择下一装配点 c_j 的规则是：

$$j = \begin{cases} \arg \max_{s \in dfc_t} \{ [\tau_{is}(t)]^\alpha \cdot [\eta_{is}(t)]^\beta \}, & \text{if } q \leq q_0 \\ j', & \text{otherwise} \end{cases} \quad (4)$$

其中， q 是一个随机变量， q_0 是一个常量， $0 \leq q, q_0 \leq 1$ 。当 $q > q_0$ 时，根据式(5)计算 dfc_i 中各个构件的状态转移概率，并按照特定策略确定 j' ：

$$P^k_{ij}(t) = \begin{cases} \frac{[\tau_{ij}(t)]^\alpha \cdot [\eta_{ij}(t)]^\beta}{\sum_{s \in dfc_t} [\tau_{is}(t)]^\alpha \cdot [\eta_{is}(t)]^\beta}, & \text{if } j \in c^C_t \\ 0, & \text{otherwise} \end{cases} \quad (5)$$

其中， dfc_t 是蚂蚁 k 在时刻 t (即蚂蚁进行第 t 步走) 将组装的领域特征簇。 α 为信息启发式因子，表示历史轨迹的相对重要性，其值越大，蚂蚁越可能选择其它蚂蚁走过的链路。 β 为期望启发式因子，表示能见度的相对重要性，其值越大， $P^k_{ij}(t)$ 越接近于贪心规则； $\eta_{ij}(t)$ 为启发函数，表示蚂蚁经过链路 (i, j) 的期望程度，定义为： $\eta_{ij}(t) = 1/L_{i,j}$ 。其中， $L_{i,j}$ 为构件 c_i 和 c_j 间的“路径长度”，通过式(3)计算得到。

4.2.3 信息素更新方法

在 $t+n$ 时刻，所有蚂蚁返回到初始装配点 c_0 后，需要对链路 (i, j) 上的信息素进行如下调整：

$$\tau_{ij}(t+n) = (1-\rho) \cdot \tau_{ij}(t) + \Delta\tau_{ij}(t) \quad (6)$$

式(6)中， ρ 为信息素挥发系数， $0 \leq \rho < 1$ 。 $\Delta\tau_{ij}(t)$ 表示本次行走中链路 (i, j) 上的信息素增量，初始时刻 $\Delta\tau_{ij}(0) = 0$ ， $\Delta\tau_{ij}(t) = \sum_{k=1}^m \Delta\tau^k_{ij}(t)$ 。表示第 k 只蚂蚁在本次行走中留在链路 (i, j) 上的信息素增量，第 k 只蚂蚁经过链路 (i, j) 时， $\Delta\tau^k_{ij}(t) = Q/L_k$ ， Q 表示信息素强度。 L_k 表示蚂蚁 k 在本次行走中经过的链路总长。蚂蚁未经过链路 (i, j) 时 $\Delta\tau^k_{ij}(t)$ 取值为 0。

4.3 算法实现

求解构件组装全局优化的改进蚁群算法实现步骤如下：

(1) 初始化参数 $n, m, q_0, \alpha, \beta, \rho$ ，以及最大迭代次数 n^{max} 。

(2) 派遣蚂蚁进行迭代求解。

For $u=1$ to n^{max}

 初始化各链路上的信息量。

 由初始装配点 c_0 所在的普通对等体 cp_0 创建 m 只蚂蚁

For k=1 to m

蚂蚁 k 开始在 CC-P2P 网络内的搜索。

For t=1 to n

获取 dfct 中各构件及相关链路的 QoS 值。

按式(3)计算各链路的路径长度。

按式(4)从 dfct 选择装配点 cj(普通对等体

cpj)。

蚂蚁走到新的装配点 cj,记录 j,累计路径长度。

度。

End For

记录蚂蚁 k 本次行走的路径总长。

End For

计算本代最好解,如路径长度小于当前最优解,则替换当前最优解。

根据解的停滞或进化情况,动态调整信息素挥发因子。

按式(6)更新链路上的信息素。

End For

(3)输出构件组装问题的最优解。

5 仿真实验

对第 4 节提出的实现构件组装的方法进行了仿真实验,实验环境如下:CPU 为奔腾 1.6G、内存为 2G、OS 为 Window XP。开发语言是 C#。实验采用伪随机方式生成构件及链路的多维 QoS 属性。改进蚁群算法的参数设置如表 1 所示。

表 1 参数设置

参数名	参数值	说明
m	10	蚂蚁总数
α	0.1	信息启发式因子
β	2	期望启发式因子
ρ	0.1	信息素挥发系数
Q	1	信息素强度
q_0	0.9	路径选择的决策常量

选取 CC-P2P 中领域特征簇总数 n 分别为 3、6、9、12、15、18、21,在各个簇中构件数均为 10 的情况下,测试算法的执行开销和命中最优解的比率。

在算法执行开销实验中,设定改进蚁群算法迭代次数区间[200, 2000],得到 CPU 运行时间的增长曲线。由于构件-链路对的 QoS 计算代价相对稳定,为便于直观比较,本实验忽略其时间消耗。实验结果如

图 3 所示。由图可知,随着领域特征簇和算法迭代次数的不断增加,算法的执行开销并没有急剧增加。当领域特征簇总数为 21 时,算法迭代 2000 代得到优化的执行时间仍在 7s 以内。显然,这样的执行开销完全可以接受。

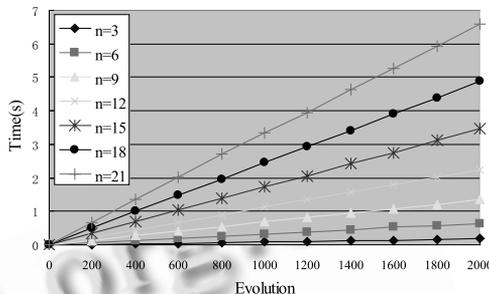


图 3 算法的执行开销

在优化解命中最优解比率的测试实验中,先通过穷举法计算出构件组装的全局最优解,将其作为改进蚁群算法得到的可行解的优化程度的评判依据。考虑到 Internet 网络环境中的多变性,定义了优化解的一个较小范围偏差系数,本实验中取值为。通过设定改进蚁群算法的迭代次数区间[100,1000],得到实验结果如图 4 所示。

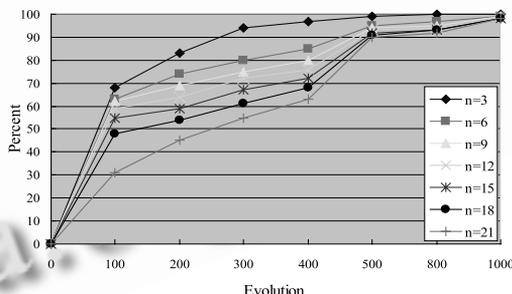


图 4 优化解命中最优解的比率

由图 5 可知,在迭代次数足够条件下,当领域特征簇不断增加、复杂度急剧增大时,优化解命中最优解的比率降低不多,这表明算法比较稳定。随着迭代次数的增加,命中率明显上升。本算法在经过 500 代迭代后搜索到的优化解命中最优解的比率达到 90%以上,经过 1000 代后,这个比率接近 100%。可见,该算法用于解决构件组装的全局优化问题是可行的。

6 结论

针对开放、动态和难控的 Internet 环境中网构 (下转第 64 页)

软件的构件组装问题, 本文提出了一种基于 P2P 环境的全局优化方法。该方法通过 CC-P2P 网络实现对构件资源按照领域特征进行分类和管理, 在构件组装过程中, 综合考虑构件和通信链路的多维 QoS 指标, 结合用户需求偏好将 QoS 值换算为路径长度, 并将其引入启发函数, 设计了一种适用于全局优化构件组装的改进蚁群算法。仿真实验表明该算法有效、可行。

参考文献

- 1 杨芙清, 吕建, 梅宏. 网构软件技术体系: 一种以体系结构为中心的途径. 中国科学(E 辑), 2008, 38(6): 818 - 828.
- 2 吴国全, 魏峻, 黄涛. 基于非确定性推理的网构软件服务质量动态评估方法. 软件学报, 2008, 19(5): 1173 - 1185.
- 3 赵丽娜, 张引, 叶修梓. 基于 P2P 网络的网构软件自适应性研究. 浙江大学学报(工学版), 2008, 42(8): 1130 - 1136.
- 4 Sanjay G, Shashishekar ST, Michael S. Service-Based P2P overlay network for collaborative problem solving. Decision Support Systems, 2007, 43(2): 547 - 568.
- 5 张伟, 梅宏. 一种面向特征的领域模型及其建模过程. 软件学报, 2003, 14(8): 1345 - 1356.
- 6 段海滨. 蚁群算法原理及应用. 北京: 科学出版社, 2005.