

基于多目标遗传算法和理想点法的 Top-k 服务组合研究^①

刘 丽^{②*} ** *** 方金云* 梁 对****

(* 中国科学院计算技术研究所计算机应用研究中心 北京 100190)

(** 中国科学院大学 北京 100049)

(*** 中国电子设备工程系统工程公司研究所 北京 100141)

(**** 中国电子设备工程系统工程公司 北京 100840)

摘 要 针对组合服务 QoS 优化中单目标优化建模需要精确定义 QoS 权重和多目标优化建模返回的 Pareto 最优解集数目过多的问题,借鉴多目标遗传算法(MOGA)和多属性决策领域的理想点法(TOPSIS),提出了一个能够快速返回 Top-k 最优组合服务、适用于较大规模服务组合问题求解的采用了 TOPSIS 的多目标遗传算法——T_MOGA。该算法在 MOGA 的方案评估阶段引入理想点法对组合服务进行评估,并通过选择、交叉、变异等操作对种群迭代以获取 QoS 更好的组合服务方案。T_MOGA 有效简化了 MOGA 的非支配排序过程,减少了算法运行时间,并只返回 Pareto 最优解集 Top-k 最优组合服务,方便用户选择。通过实验,从算法的运行时间及得到集合的质量两个方面验证了 T_MOGA 的有效性。

关键词 服务组合优化,Top-k,理想点法(TOPSIS),多目标遗传算法(MOGA),服务质量(QoS)

0 引 言

Web 服务技术的成熟发展使得服务的数量不断增多,服务组合越来越多地需要考虑服务的非功能属性,即服务质量(QoS),基于 QoS 的服务组合成为研究热点。如何从服务组合各抽象任务的候选服务集合中分别选出一个具体的服务,使得选中的这组服务在满足用户对组合服务所有约束的同时,组合服务的整体 QoS 最高,则成为追求的目标^[1]。基于 QoS 的服务组合是一个组合优化问题,属于 NP 难题^[2]。其求解算法主要分为两类^[3]:一类是基于整数规划^[4]、图论^[5-7]、凸包集^[8]等非进化算法求解;一类是基于遗传算法^[9]、蚁群算法^[10]、粒子群优化^[11]等进化算法求解。非进化算法能够找到确定的最优解,但时间复杂度高,仅适合小规模的服务组合。进化算法能够以较快的速度找到近似最优解,在求解速度和求解质量上得到了平衡,其中遗传算法由于对初始可行解要求不高,涉及参数少,是最广

泛使用的算法。另外,QoS 涉及响应时间、信誉等多个维度,基于 QoS 的服务组合还是一个多属性决策问题。多数研究^[2,5-7,9-13]为每个属性分配一个表示属性相对重要性的权重,将多个 QoS 需求转化为单目标优化问题。这种方法对权重值非常敏感,并且仅能产生唯一的最优解。近年来,越来越多的研究^[14-20]通过找到满足约束条件的 Pareto 最优解集来实现对多个目标的同时优化,用户根据自己的偏好进行选择。这种方法无需用户事先给出权重值,并且能够得到多个不可比较的组合方案。但是这种方法得到的组合方案的数目(k)是不可控的,当 k 过大时,用户同样难以选择。针对上述情况,本文运用多目标优化模型对服务组合问题进行建模,提出了一种使用了理想点法(technique for order preference by similarity to ideal solution,TOPSIS)的多目标遗传算法(multi-objective genetic algorithm,MOGA),简称为 T_MOGA。该算法借鉴了 MOGA 的思想,并将 TOPSIS 用于组合方案的适应度评估中,能够快速返回有限的 k 个最优组合服务,在不对组合服务方案

① 863 计划(2011AA120302)资助项目。

② 女,1982 年,博士生;研究方向:Web 服务,遗传算法,模糊多属性决策,空间信息服务组合,空间信息服务质量;联系人,E-mail:liuli@ict.ac.cn

(收稿日期:2013-05-06)

进行非支配排序的前提下,高效返回 Pareto 最优支配集合中 TOPSIS 意义上的 Top-k 最优组合服务,而且适用于较大规模服务组合。

1 组合服务的 QoS

组合服务 (composition service, CS) 是由单个服务根据一定的流程组合得到的,其 QoS 是单个服务 QoS 的聚合,与 QoS 属性以及服务组合的流程有关。典型的服务组合流程包括顺序、并行、条件和循环 4 种。T_MOGA 算法与具体的 QoS 属性无关,可以扩展到任意多个 QoS 属性,本文使用价格、响应时间、可用性、可靠性和声誉 5 种常用的 QoS 属性进行研究^[2]。价格是服务执行所需要的费用,响应时间是从发出服务请求至收到服务响应所需要的时间,可用性是服务可以被使用的概率,可靠性是服务成功执行的概率,信誉是服务的可信性。本文使用一个五元组来表示服务的 QoS,记为 $q(s) = \{q^1(s), q^2(s), q^3(s), q^4(s), q^5(s)\}$ 。

组合服务的 QoS 用表 1 中的公式计算。

表 1 组合服务 QoS 聚合函数

	顺序	并行	选择	循环
价格	$\sum_{i=1}^n q^1(i)$	$\sum_{i=1}^n q^1(i)$	$\sum_{i=1}^n (q^1(i) \cdot p_i)$	$q^1(1) \cdot k$
响应时间	$\sum_{i=1}^n q^2(i)$	$\max_i q^2(i)$	$\sum_{i=1}^n (q^2(i) \cdot p_i)$	$q^2(1) \cdot k$
可用性	$\prod_{i=1}^n q^3(i)$	$\prod_{i=1}^n q^3(i)$	$\prod_{i=1}^n (q^3(i) \cdot p_i)$	$q^3(1) \cdot k$
可靠性	$\prod_{i=1}^n q^4(i)$	$\prod_{i=1}^n q^4(i)$	$\prod_{i=1}^n (q^4(i) \cdot p_i)$	$q^4(1) \cdot k$
信誉	$\sum_{i=1}^n q^5(i)$	$\sum_{i=1}^n q^5(i)$	$\sum_{i=1}^n (q^5(i) \cdot p_i)$	q^5
	n	n		

2 组合服务的评估方法

2.1 Pareto 评估

多目标服务组合优化问题是对多个属性同时进行优化,而各个属性可能会存在“冲突”,比如价格和响应时间,响应时间小的服务往往价格较高,因此往往无法同时达到最优。Pareto 评估是一种对多目标个体进行评估的成熟方法。

定义 1 Pareto 支配关系^[21]。对于组合服务 CS 的任意两个可能的方案 cs_i 和 cs_j ,如果 cs_i 在所有的 QoS 属性上均优于或者等于 cs_j ,并且至少在一个

维度上完全优于 cs_j ,则 cs_i Pareto 支配 cs_j ,记为 $cs_i > cs_j$ 。 $cs_i > cs_j \leftrightarrow (\forall k \in [1, r], Q^k(cs_i) \leq Q^k(cs_j)) \wedge (\exists t \in [1, r], Q^t(cs_i) < Q^t(cs_j))$ 。

定义 2 组合服务的 Pareto 最优解集。在组合服务 CS 的所有可能组合方案中,不被任何其他方案 Pareto 支配的组合方案所构成的集合,为组合服务的 Pareto 最优解集。 $ParetoSet(CS) = \{cs_i \in CS \mid \neg \exists cs_j \in CS, cs_i > cs_j\}$ 。

Pareto 评估能够同时考虑各个方案的 QoS 向量的所有维度,所确定的 Pareto 最优解集中的服务,相互之间是不可比较的,即集合中的服务不可能在所有的 QoS 指标上都比另一个服务好。

2.2 理想点法评估

理想点法 (TOPSIS) 是求解多属性决策问题的一种非常有效的方法,其中心思想是先选定一个正理想解和一个负理想解,然后根据方案与正、负理想解的距离对各个方案排序。最佳方案一定更靠近正理想解并且更远离负理想解,这也符合人们决策时的逻辑。TOPSIS 能够在考虑多个属性的情况下对方案进行排序,已广泛应用于供应链管理、电子商务等多个领域^[22]。使用 TOPSIS 对组合服务方案进行评估的步骤如下^[23]：

(1) 建立各个组合服务的 QoS 决策向量。

QoS 指标可分为利益型和代价型两类。前者属性值越大,对应的服务质量越好,比如可用性、可靠性和声誉;后者属性值越小,对应的服务质量越好,比如价格和响应时间。为了消除不同物理量纲对评估结果的影响,也为了使所有的 QoS 指标具有一致的趋势,我们使用文献[24]中的公式(5)对代价型指标,使用公式(6)对利益型指标分别进行归一化(优化目标为最小化)。各个组合方案标准化后的决策向量记为 $V(cs)$ 。同样,用户的需求也可以表示成一个 QoS 向量,并在标准化后参加优选过程。

(2) 得到加权规范化向量。

由于我们的目标是在用户给出权重前进行服务的预筛选,防止用户因权重设置不准确而损失真正需要的组合服务,因此我们不考虑各个指标的权重,即认为加权后的决策矩阵仍为 $V(cs)$ 。

(3) 确定正理想解 S^+ 和负理想解 S^- 。

对于利益型指标,正理想解使指标值最大,负理想点使指标值最小;对于代价型指标,正理想点使指标值最小,负理想点使指标值最大。组合服务由单个服务经过串行、并行、选择和循环等结构组成,并行、选择和循环流程可根据表 1 分别计算该流程的

QoS,由此约简为单个的服务,与其它服务按照顺序流程进行聚合^[25]。因此,组合服务在 5 个维度的聚合函数具有单调性,对应某个维度上的最大(最小)值,可以由每个任务候选服务的最大(最小)值进行聚合得到。由此可得到所有服务组合方案的服务质量构成的矩阵 $Q(cs)$ 正、负理想点: $S^+ = \{Q_{\min}^1, Q_{\min}^2, Q_{\max}^3, Q_{\max}^4, Q_{\max}^5\}$; $S^- = \{Q_{\max}^1, Q_{\max}^2, Q_{\min}^3, Q_{\min}^4, Q_{\min}^5\}$ 。

$V(cs)$ 是 $Q(cs)$ 标准化后的决策向量(以最小化为目标),其取值范围为 $[0, 1]$, 因此 $V(cs)$ 的正、负理想点为: $S^+ = \{0, 0, 0, 0, 0\}$; $S^- = \{1, 1, 1, 1, 1\}$ 。

(4) 计算方案 cs_i 到正、负理想解之间的距离。

$$D_i^+ = \sqrt{\sum_{j=1}^5 (v_j^+)^2} \quad D_i^- = \sqrt{\sum_{j=1}^5 (1 - v_j^+)^2}$$

(5) 计算组合服务的 TOPSIS 评估值。

$$C_i^* = D_i^- / (D_i^+ + D_i^-) \quad (1)$$

组合服务的优劣即是将服务按其 C_i^* 值的大小进行排序。

2.3 TOPSIS 评估与 Pareto 评估

采用 TOPSIS 评估对各组合服务进行排序的过程中,实际上就已经完成了组合服务的 Pareto 评估。

定理 1 若一个组合服务 Pareto 支配另一个组合服务,则前者的 TOPSIS 评估值大于后者。即若 $cs_i > cs_j$, 则 $C^*(cs_i) > C^*(cs_j)$ 。

证明:下面以二维空间(价格,响应时间)为例进行证明,高维空间的证明类似。

假设组合服务 $cs_1(x_1, y_1)$ Pareto 支配组合服务 $cs_2(x_2, y_2)$, 它们的关系如图 1 所示。设组合服务 cs_1 到正理想点的距离为 S_1^+ , 到负理想点距离为 S_1^- , 组合服务 cs_2 到正理想点的距离为 S_2^+ , 到负理

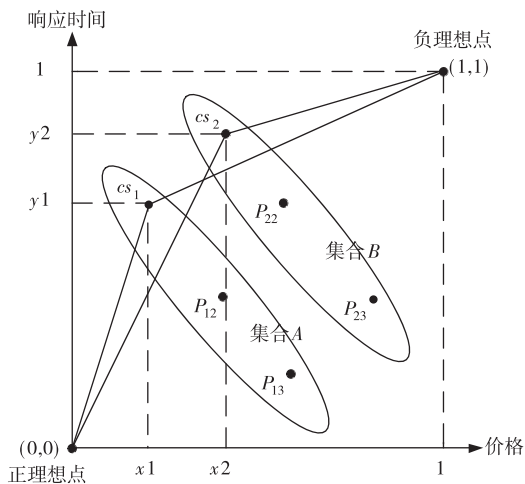


图 1 组合服务的 Pareto 支配集合

想点距离为 S_2^- , 则 $cs_1 > cs_2 \Rightarrow x_1 \leq x_2, y_1 \leq y_2$ (不全取等号) $\Rightarrow S_1^+ < S_2^+, S_2^- < S_1^- \Rightarrow \frac{S_1^+}{S_1^-} < \frac{S_2^+}{S_2^-} \Rightarrow \frac{S_1^+ + S_1^-}{S_1^-} < \frac{S_2^+ + S_2^-}{S_2^-} \Rightarrow \frac{S_1^-}{S_1^+ + S_1^-} > \frac{S_2^-}{S_2^+ + S_2^-} \Rightarrow C^*(cs_1) > C^*(cs_2)$ 。

推论 1 如果集合 A Pareto 支配集合 B , 则集合 A 中任意一个组合服务的 TOPSIS 评估值高于集合 B 中组合服务的 TOPSIS 评估值。即若 $ParetoSet(A) > ParetoSet(B) \Rightarrow \forall (i \in A, j \in B), C^*(A_i) > C^*(B_j)$ 。

证明:反证法。假设集合 B 中存在一个方案 j , 集合 A 中存在某个方案 i , 方案 j 的 TOPSIS 评估值 B_j 大于方案 i 的评估值 A_i 。因为集合 A Pareto 支配集合 B , 所以 A_i 支配 B_j , 根据定理 1, 评估值 A_i 大于评估值 B_j , 与假设矛盾, 得证。

3 T_MOGA 服务组合算法

遗传算法是一种广泛使用的随机搜索算法,它利用对搜索空间小得多的种群规模,通过若干代的选择进化操作使种群快速收敛于目标空间,被广泛应用于各种大规模优化问题。现有的多目标遗传算法,如 NSGA2^[26]、SPEA 等,都基于 Pareto 优化来实现种群的进化。本文借鉴 NSGA2 算法的思想,设计并实现了 T_MOGA 算法。

3.1 组合方案的编码与目标向量计算

组合服务的任务数和每个任务候选服务的个数都是整数,根据这一特点,我们使用整数编码。其中每个基因位对应组合方案中的一个任务,基因值为候选服务的序号,染色体的长度等于任务的个数。按 T_MOGA, 每个染色体对应于一个可能的组合方案,对组合方案的评价就是对其 QoS 目标向量进行评估。目标向量的计算包括下面 3 个步骤:(1) 根据染色体得到构成组合方案的每个候选服务的 QoS 向量;(2) 根据表 1 计算组合方案的聚合 QoS 向量;(3) 按照 2.2 中的方法标准化 QoS 向量。

3.2 组合方案评价

遗传算法的进化选择过程发生在由染色体构成的种群之上,选择的依据是每个染色体的适应性。文献[27]使用 NSGA2 算法来求解服务组合问题,由于 NSGA2 算法是针对连续数值优化问题设计的,它的目标是找到处于或者尽量接近于 Pareto 前沿的一个具有代表性的子集,并使得结果集中的方案均

均匀分散在整个 Pareto 边界上,因此找到的组合方案是各个 QoS 指标值比较分散的 Pareto 支配集合,而不是我们所需要的 Top-k 最优方案。但是可以借鉴 NSGA2 算法对个体的评估方法来设计我们的算法。我们对每个个体设定了两个评估值:级别 *rank* 和距离 *distance*。每个个体的级别 *rank* 用来区分当前种群是否能够满足用户的 QoS 约束,如果满足,则其 *rank* = 1,否则 *rank* = 2。每个个体的距离 *distance* 是采用式(1)计算的 TOPSIS 评估值。

3.3 进化操作

遗传算法通过选择、交叉、变异等操作对种群迭代,由此得到适应度函数值高的解。

在选择操作中,采用二进制竞赛选择机制,即在种群中随机挑选两个个体,优者进入下一代。个体 P_i 优于个体 P_j 的条件为

$$P_i \text{ 优于 } P_j \leftrightarrow (P_i.rank < P_j.rank) \vee ((P_i.rank = P_j.rank) \wedge (P_i.distance > P_j.distance)) \quad (2)$$

前一个选择条件促使选择的结果能够快速收敛于满足用户 QoS 约束的解。后一个选择条件则促使选择结果能够快速收敛于 Pareto 最优解中整体性能更优的组合方案。虽然 T_MOGA 算法并没有对组合方案进行非支配排序,但是根据 2.3 中的推论 1, Pareto 最优支配集中的组合方案比其它组合方案具有更大的 TOPSIS 评估值,会优先被选中。对于同属于 Pareto 前沿集合中的组合方案,选择具有更大 *distance* 的个体应意味着选择了整体 QoS 性能更优的组合方案。同时,采用单点交叉和位变异操作。

3.4 算法主流程

首先生成初始化种群 *pop*,对种群进行目标向量计算和 TOPSIS 评估,得到种群个体的级别及每个级别内各个体的 TOPSIS 评估值。然后对种群进行选择、交叉和变异操作。整个算法的伪代码如下:

Algorithm T_MOGA(*M*,*N*,*R*,*P_c*,*P_m*,*popsiz*,*maxGen*,*k*)

输入:

M—任务的个数;*N*—候选服务的个数;

R—QoS 指标的个数;*P_c*—交叉概率;*P_m*—变异概率

popsiz—种群大小;*maxGen*—最大迭代代数

k—返回组合服务方案的个数

输出: TOPSIS 意义上最优的 *k* 个组合方案

```

1  opt = InitOpt( M, N, R, Pc, Pm, popsiz, maxGen );
2  pop = InitPop( opt );
3  pop = verifyrepeat( opt, pop );
4  pop = evaluate( opt, pop );
5  [ opt, pop ] = rankdis( opt, pop );
6  ngen = 0;
7  while( ngen < opt. maxGen )

```

```

8  ngen = ngen + 1;
9  newpop = selectOp( opt, pop );
10 newpop = crossoverOp( opt, newpop );
11 newpop = mutationOp( opt, newpop );
12 combinepop = [ pop, newpop ];
13 combinepop = verifyrepeat( opt, combinepop );
14 combinepop = evaluate( opt, combinepop );
15 [ opt, combinepop ] = rankdis( opt, combinepop );
16 pop = extractPop( opt, combinepop );
17 end
18 pop_k = output( pop, k );

```

在进化过程中, T_MOGA 采取了精英保留机制,即将父代种群 *pop* 和新产生的子代种群 *newpop* 合并,从合并后的种群 *combinepop* 中选择出适应度更高的个体作为下一代种群,这样能够保留父代的优势个体,提高进化效率和质量。在从 *combinepop* 中选择新的个体时,优先选择 *rank* 值小的个体,以使种群尽快收敛到满足用户 QoS 需求的解,在 *rank* 值相同的情况下,优先选择 TOPSIS 值更大的个体,直到所拷贝的个体与种群大小相等。*pop* 填满后开始下一次迭代过程。在遗传操作过程中可能产生冗余个体,即 *pop* 和 *newpop* 中包含相同的个体,当冗余个体恰为精英个体时,精英机制会使得这些冗余个体被反复保留,同时遗传操作会进一步降低精英种群的多样性,导致精英种群中的冗余个体不断增多,使结果集质量严重下降。为避免冗余, T_MOGA 对合并后的种群进行去重操作,使 *combinepop* 中没有染色体完全相同的个体,由 *verifyrepeat*(*opt*, *combinepop*) 完成。在极限情况下, *combinepop* 中冗余个体过多使得所有不同个体的数量小于种群的规模,导致没有足够的个体拷贝入 *pop*,从而影响下一次迭代过程。通过在产生初始种群时消除冗余来避免上述极限情况的发生。最后,根据用户的要求,输出前 *k* 个组合方案作为 Top-k 方案。

4 实验验证

我们将 T_MOGA 与 NSGA2 服务组合算法^[27]以及 FuMuCom 算法^[28]进行了比较,从算法的效率和结果集质量两个方面进行验证。三个算法均在 Matlab R2010b 环境下实现,运行环境为 Windows 7 Professional Service Pack 1 操作系统,硬件环境为 Intel i5 - 2410M CPU, 4GB RAM 内存。对于 NSGA2 算法^[27]和 T_MOGA 算法,种群大小 *popsiz* = 50, 最大遗传代数 *maxGen* = 50, 交叉概率 *P_c* = 0.9, 变异概

率 $P_m = 0.01$ 。默认情况下,服务组合的任务个数为 5,每个任务的候选服务个数为 10,约束条件的个数为 2,QoS 属性的个数为 5,返回 Top-10 服务组合方案。所有试验均独立运行 20 次并统计结果。

4.1 问题规模与运行时间

分别考查任务个数、候选服务个数及 QoS 指标个数三个因素对算法运行时间的影响。任务的个数变化范围是 3~50,每个任务候选服务数目变化范围是 3~50,QoS 维度变化的范围是 3~15,实验结果如图 2 至图 4 所示(采用默认服务规模时,FuMuCom 算法运行时间过长,没有在图 4 中显示)。

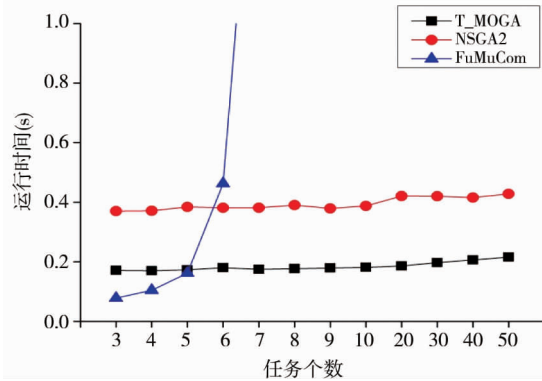


图2 任务数对运行时间的影响

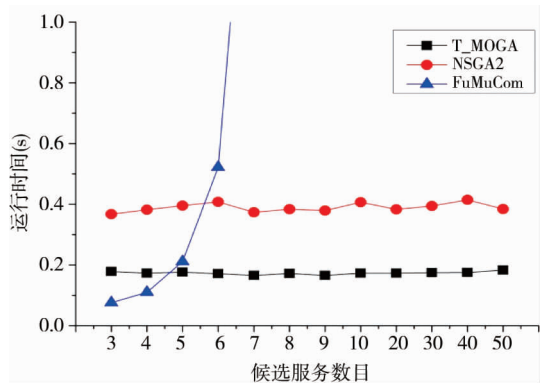


图3 候选服务数对运行时间的影响

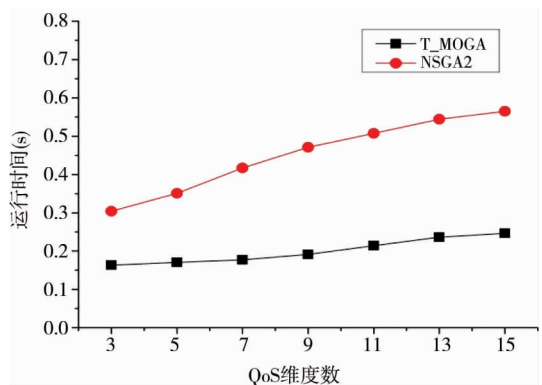


图4 QoS 维数对运行时间的影响

从实验结果可以看出,FuMuCom 的运行时间过长,仅适用于服务规模较小的情况;T_MOGA 和 NSGA2 算法的运行时间较短,可适用于服务规模较大的情况。T_MOGA 和 NSGA2 算法的运行时间与种群的大小和迭代的次数紧密相关。因为它们每次只从可能的空间中随机选择一个参加组合,因此运行时间不受服务组合规模的影响,当服务任务数和候选服务数增大时,运行时间基本没有发生改变。同时,T_MOGA 使用 TOPSIS 方法对个体进行评估,而没有进行 NSGA2 中的非支配排序,因此运行时间明显小于 NSGA2。此外,随着 QoS 维数的增加,需要计算的适应性函数的个数不断增多,并且约束的个数也不断增多,需要更多时间比较,因此 T_MOGA 和 NSGA2 算法的运行时间都有轻微增长,但 T_MOGA 的时间始终小于 NSGA2。随着任务数或候选服务数的增长,可能的组合服务的个数将增加,保持种群大小和迭代次数不变,将减少可能搜索到的解的可能性,会在一定程度上影响结果的准确性。

4.2 T_MOGA 算法的质量

为验证算法得到的结果的质量,我们对默认的空间进行搜索,在此设定下,组合服务的个数将达到 10^5 个,能够满足用户的服务组合需要。采用 FuMuCom 算法能够得到精确的 Top-k 结果,作为 T_MOGA 算法和 NSGA2 算法的参照结果。将 T_MOGA 算法或 NSGA2 算法返回的 Top-k 结果集的 TOPSIS 评估值与 FuMuCom 算法返回的 Top-k 个结果集的 TOPSIS 评估值的比值作为算法的贴近度,实验结果如图 5 和图 6 所示。

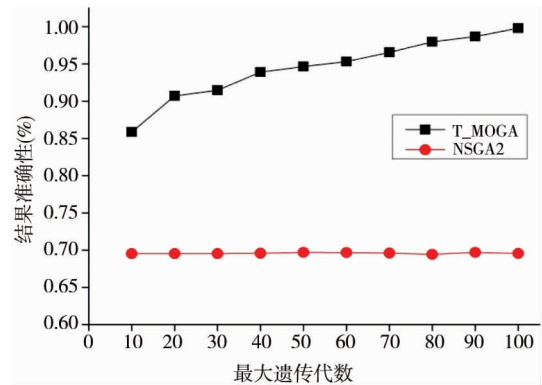


图5 T_MOGA 与 NSGA2 结果的准确性比较

从试验结果可以看出,T_MOGA 算法得到的准确性远远高于 NSGA2 算法,在默认的代数下能够达到 90% 以上。NSGA2 算法更强调得到的结果在各

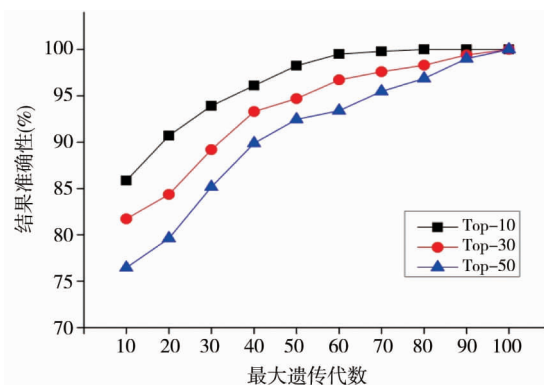


图6 Top-K 集合准确性与 K 和迭代次数的关系

个 QoS 维度上的分散性,而没有考虑结果在 TOPSIS 评估下的整体最优效果,所得到的结果整体贴适度并不高。T_MOGA 算法得到的结果的贴适度与种群的大小和迭代次数密切相关,随着两者的增加,准确性不断增加。此外,即使对于较大的 k,算法也能够返回质量较好的结果集,当迭代次数达到 50 时,对于较大的 k 值($k=50$),所得到的结果的贴适度也能够达到 90% 以上。

5 结论

目前,基于 QoS 的服务组合优化算法或者将服务组合优化问题建模为单目标优化问题,需要用户为每个 QoS 指标精确分配权重,或者虽然使用了多目标优化模型,但是返回给用户过多的服务,用户难以选择。为了解决此难题,本文提出了 T_MOGA 算法,一种基于多目标遗传算法和 TOPSIS 评估方法的服务组合优化算法。试验证明,T_MOGA 算法的运行时间少,得到的集合质量好,能够适应较大规模的服务组合需要。下一步的研究是将模糊控制机制引入到 T_MOEA 算法中,根据群体适应度来动态调整交叉概率 P_c 和变异概率 P_m ,防止由于 P_c 和 P_m 参数值设置不当引起的算法早熟和局部优化等问题,并进一步提高算法的性能。

参考文献

- [1] 李金忠,夏洁武,唐卫东等. 基于 QoS 的 Web 服务选择算法综述. 计算机应用研究, 2010, 27(10): 3622-3627
- [2] Zeng L, Benatallah B, Ngu A, et al. QoS-aware middleware for web services composition. *IEEE Transactions on Software Engineering*, 2004, 30(5): 311-327
- [3] Pejman E, Rastegari Y, Majlesi P, et al. Web service composition methods: a survey. In: Proceedings of the Interna-

tional MultiConference of Engineers and Computer Scientists, Hong Kong, China, 2012. 335-350

- [4] Zeng L, Benatallah B, Dumas M, et al. Quality-driven web services composition. In: ACM International World Wide Web Conference, Budapest, Hungary, 2003. 218-239
- [5] Yu T, Lin K J. Service selection algorithms for composing complex services with multiple QoS constraints. In: International Conference on Service Oriented Computing, Amsterdam, Netherlands, 2005. 130-143
- [6] 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
- [7] 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
- [8] Mostofa A M, Sohel Rahman M, Kaykobad M, et al. Lin. Solving the multidimensional multiple-choice knapsack problem by constructing convex hulls. *Computers & Operations Research*, 2006, 33(5): 1259-1273
- [9] Canfora G, Di Penta M, Esposito R, et al. An approach for QoS-aware service composition based on genetic algorithms. In: Genetic and Evolutionary Computation Conference, Washington DC, USA, 2005. 1069-1075
- [10] 夏亚梅,程渤,陈俊亮等. 基于改进蚁群算法的服务组合优化. 计算机学报, 2012, 35(2): 270-281
- [11] 舒振,陈洪辉,罗雪山. 基于改进混合粒子群算法的服务动态选择方法. 中南大学学报(自然科学版), 2011, 42(10): 3086-3094
- [12] 刘书雷,刘云翔,张帆等. 一种服务聚合中 QoS 全局最优服务动态选择算法. 软件学报, 2007, 18(3): 646-656
- [13] 张童,刘云生,刘云生. 一种 QOS 驱动的仿真服务组合优化方法. 系统仿真学报, 2009, 21(16): 4990-4994
- [14] 胡焕耀,董渭清,符锐等. 面向 Pareto 最优遗传算法的服务组合方法. 西安交通大学学报, 2009, 43(12): 50-54
- [15] 方其庆,刘庆华,彭晓明等. QoS 全局最优的多目标 Web 服务选择算法. 计算机应用研究, 2009, 26(12): 4442-4445
- [16] Li L, Yang P, Ou L, et al. Genetic algorithm-based multi-objective optimization for QoS-aware web services composition. In: Knowledge Science, Engineering and Management Lecture Notes in Computer Science, Dalian, China, 2010. 549-554
- [17] Alrifai M., Skoutas D, Risse T. Selecting skyline services for QoS-based web service composition. In: Proceedings of the 19th International Conference on World Wide Web, Raleigh, USA, 2010. 11-20
- [18] 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. 同济大学学报(自然科学版), 2011, 39(4):599-605
- [19] Benouaret K, Benslimane D, Hadjali A. Selecting skyline web services from uncertain QoS. In: IEEE 9th International Conference on Services Computing, Hawaii, USA, 2012. 523-530
- [20] Yu Q, Bouguettaya A. Computing service skylines over sets of services. In: The 8th IEEE International Conference on Web Services, Miami, USA, 2010. 481-488
- [21] Ppen K O, Vicente-Garcia M R, Nickolay B. Fuzzy Pareto dominance and its application in evolutionary multi-objective optimization. In: Evolutionary Multi-Criterion Optimization Lecture Notes in Computer Science, Sheffield, UK, 2005. 399-412
- [22] Behzadian M, Otaghsara S, YazdanibM, et al. , A state-of-the-art survey of TOPSIS applications. *Expert Systems with Applications*, 2012, 39(17):13051-13069
- [23] Chamodrakas I, Martakos D. A utility-based fuzzy TOPSIS method for energy efficient network selection in heterogeneous wireless networks. *Applied Soft Computing*, 2012, 12(7):1929-1938
- [24] Yuhong Y, Chen M, Yubin Y. Anytime QoS optimization over the PlanGraph for web service composition. In: Proceedings of the 27th Symposium on Applied Computing, Trento, Italy, 2011. 1968-1975
- [25] 林闯. 随机 Petri 网和系统性能评价. 北京:清华大学出版社, 2005. 305-313
- [26] Deb K, Pratap A, Agarwal S. A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation*, 2002. 6(2):182-197
- [27] YaoY, Chen H. QoS-aware service composition using NSGA-II. In: Proceedings of the 2nd International Conference on Interaction Sciences: Information Technology, Culture and Human, Seoul, Korea, 2009. 358-363
- [28] 李祯, 杨放春, 苏森. 基于模糊多属性决策理论的语义 Web 服务组合算法. 软件学报, 2009, 20(3):583-596

A Top-k service composition research based on MOGA and TOPSIS

Liu Li * ** ***, Fang Jinyun *, Liang Dui ****

(* Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100190)

(** University of Chinese Academy of Sciences, Beijing 100049)

(*** Institute of China Electric Equipment and System Engineering Co. Ltd, Beijing 100141)

(**** China Electric Equipment and System Engineering Co. Ltd, Beijing 100840)

Abstract

By reference to the multi-objective genetic algorithm(MOGA) and the technique for order preference by similarity to ideal solution(TOPSIS) in multi-attribute decision making, the T_MOGA, a MOGA using the TOPSIS, was presented for optimization of composition service, without the difficulties that the modeling of single-objective optimization needs accurate QoS weight defining and the number of Pareto optimal sets is too large when modeling of multi-objective optimization. This T_MOGA can efficiently find the Top-k optimal solutions from the view of TOPSIS evaluation. It introduces the TOPSIS to evaluate the fitness of the service composition plans, and adopts the operations of binary tournament selection, one-point crossover and bit-mutation to evolve plans with better fitness. The T_MOGA simplifies the non-dominated sorting process of the MOGA, which reduces the running time of the algorithm, and it only returns Top-k optimal plans, which is convenient for users to select. The effectiveness of the algorithm in efficiency and scalability was verified by experiment.

Key words: service composition optimization, Top-k, technique for order preference by similarity to ideal solution(TOPSIS), multi-objective genetic algorithm(MOGA), quality of service(QoS)