

支持服务关联的全局最优服务选择算法^①

聂 凯^② 汪厚祥 何 皎

(海军工程大学电子工程学院 武汉 430033)

摘 要 针对面向服务的军事综合电子信息系统候选服务的服务质量往往依赖于其他候选服务的问题,提出了一种支持服务关联的全局最优服务选择方法。该方法将基于 QoS 的军事信息服务选择问题建模为带 QoS 约束的多目标组合优化问题,采用基于独立成分分析(ICA)的多目标分布估计算法同时优化多个目标函数,最终产生一组满足约束条件的 Pareto 最优解服务组合集。对比实验结果表明,基于 ICA 的多目标分布估计算法通过概率图模型可以描述服务间的相互关系,可以更好地解决存在服务关联的组合服务选择问题,具有良好的有效性和可行性。

关键词 军事信息服务组合, 服务选择, 服务关联, 多目标优化, 分布估计算法(EDA), 独立成分分析(ICA)

0 引 言

服务计算和面向服务的体系结构(service-oriented architecture, SOA)在软件开发和集成上具有很多优点,这使其成为未来军事综合电子信息系统开发和集成的发展方向,军事信息服务的概念也因此应运而生^[1]。系统设计者需要在现有的多个满足系统功能需求的候选服务中,选择特定服务质量的服务,通过这些服务的组合,构建有服务质量保障且最大程度满足军事指挥员需求的增值服务,这就是基于 QoS 的组合服务选择问题。现有的组合服务选择模型认为各候选服务是相互独立的,其服务质量独立于其他服务。而面向服务的军事应用中,候选服务的服务质量往往依赖于其他候选服务,如信息处理服务作为火控解算服务的输入,其服务质量的高低直接决定着火控解算服务的精度,这类服务之间的关系称为服务关联^[2-4],现有组合服务选择方法很难应对这种场景。面向服务的军事综合电子信息系统如何快速灵活地选择出满足用户需求的最优服务组合,已成为亟待解决的问题。针对这种情况,本文提出了一种支持服务关联的全局最优服务选择算法。

1 相关研究

国内外学者已从不同的角度将基于 QoS 的组合服务选择问题转换为某种数学问题,如整数或线性规划问题^[5]、多选择背包问题^[6]、多属性决策问题^[7]、带 QoS 约束的单目标组合优化问题^[8]、带 QoS 约束的多目标组合优化问题^[9]等,在此基础上构造模型并设计相应的选择算法。传统的基于 QoS 的服务选择算法主要采用穷举算法和贪婪算法,近来较流行的是采用启发式算法设计的选择算法,如遗传算法(genetic algorithm, GA)^[10]、蚁群优化(ant colony optimization, ACO)^[11]及一些融合算法等。

整数规划模型没有考虑服务的关联关系^[5]。多维 0-1 背包模型认为每个物件的重量是固定不变的,实际上服务的质量会因其它服务类所选择的候选服务的变化而变化,因而该模型不适用于存在服务关联的应用场景^[6]。带 QoS 约束的单目标组合优化模型将各个 QoS 约束参数线性加权转化为一个单目标函数,不能真正做到同时优化多个 QoS 参数目标且对权重向量非常敏感^[8]。带 QoS 约束的多目标组合优化问题对每个 QoS 参数分别设置相应的目标函数,它的解并非唯一,用户可依其偏好选

① 国防预研基金(4010601010201)和中国博士后科学基金(20090461460)资助项目。

② 男,1983 年生,博士生,工程师;研究方向:服务计算与云计算,多目标优化算法;联系人, E-mail: 1999104133@163.com (收稿日期:2012-06-20)

择最满意的解。因此,军事综合电子信息系统服务选择问题可以建模为存在关联关系的带 QoS 约束的多目标组合优化问题。

关于服务间的关联关系及组合服务选择算法,文献[2]提出了支持服务关联的服务选择方法,包括基于整数规划求解最优解的方法和基于启发求解次优解的方法;文献[3]结合社会网络和语义 Web 技术,将可用 Web 服务组织成具有丰富语义信息、业务上下文和交互关系的服务生态系统,将语义推理和关系演算引入到服务的发现、组合等过程中。文献[4]基于服务间的关联关系,提出了一种基于路径分支关联规则和服务执行顺序序列模式的服务选取方法,但该方法不支持动态服务选择。总的来看文献[2-4]对服务关联关系的描述不统一且复杂,影响了它们的广泛应用,需要寻找新的服务选择方法。

分布估计算法(estimation of distribution algorithm, EDA),它是一种新的进化模式,没有交叉、变异等传统遗传操作,而是采用概率图模型的学习和采样。EDA 给解决优化问题提供了新工具,它是进化算法与统计学习相结合的产物,它通过概率模型可以描述变量之间的相互关系,从而对解决非线性、变量耦合的优化问题更加有效^[12,13]。它在解决优化问题时表现出了优良性能,因此被越来越多地用于多目标优化。多目标 EDA 就是将进化多目标优化算法处理多目标优化的策略与 EDA 算法相结合,根据 EDA 算法的特点设计求解 Pareto 解集的更有效的算子^[12]。mBOA^[14]将贝叶斯优化算法(BOA)与 NSGA-II 中的选择策略相结合解决带有欺骗的多目标优化问题,仿真结果优于 NSGA-II 本身。多目标 IDEA^[15]是基于混合高斯分布和高斯核函数,它是 IDEA 的扩展,它主要针对多峰值问题,但没有充分考虑变量之间的关系。文献[16]提出了基于规则模型的多目标分布估计算法(regularity model-based multi-objective estimation of distribution algorithm, RM-MEDA),它在多样性方面优于 NSGA-II-PCX 和 GDE3,但在解决复杂的多峰强耦合、非线性多目标优化问题上仍不能达到满意结果。独立成分分析(independent component analysis, ICA)^[17,18]是在主成分分析(principal component analysis, PCA)的基础上发展而来的,不同的是 PCA 基于信号二阶统计特性,用于去除信号各分量之间的相关性,而 ICA 是基于信号高阶统计特性,经 ICA 分解出的各信号分量之间是相互独立的。正是因为这一特点,使 ICA 受到了广泛的关注。RM-MEDA 采用局部主成

分分析算法对模型进行迭代训练,但收敛时间较长;文献[19]采用概率主成分分析(probabilistic principal component analysis, PPCA)对概率图模型进行建模,但只能针对高斯分布。文献[20]将 ICA 与单变量边缘分布算法(UMDA)相结合,取得了好于 UMDA 的效果,但对多峰强耦合问题性能一般;文献[21]将动态贝叶斯混合 ICA 模型与 EDA 算法相结合,仿真结果表明它优于以前的基于单独 ICA 或混合高斯模型的 EDA 算法。

本文针对以上服务关联关系的 QoS 描述模型不统一且复杂的问题,采用 ICA 分离出独立的各分量来避免复杂的模型从而简化工作量。因此首先将军事综合电子信息系统服务选择问题建模为存在关联关系的带 QoS 约束的多目标优化问题,然后采用基于 ICA 的多目标分布估计算法同时优化多个目标函数,产生一组满足约束条件的 Pareto 最优解服务组合集,最后进行实验验证算法的有效性和可行性。

2 问题描述

军事信息服务的描述模型为 $W = \{F, Q\}$, 其中 F 为功能属性集合, Q 为非功能性集合即质量属性集合。服务请求者根据 F 来判断服务是否满足其功能需求,并依据 QoS 从一个服务类的多个候选服务中选择一个合适的服务。下面给出四个定义:

候选服务 W_{ij} : $W_{ij}(i = 1, 2, \dots, m; j = 1, 2, \dots, n)$ 是构成组合服务的基本单元,如图 1 中的 W_{1j}, W_{2j}, W_{3j} 等。利用候选服务,按照特定的业务逻辑可以构建复杂的组合服务。1 个候选服务包含 1 个 QoS 向量 $Q_{ij} = (Q_{ij}^1, Q_{ij}^2, \dots, Q_{ij}^l)$, l 为 QoS 属性个数。

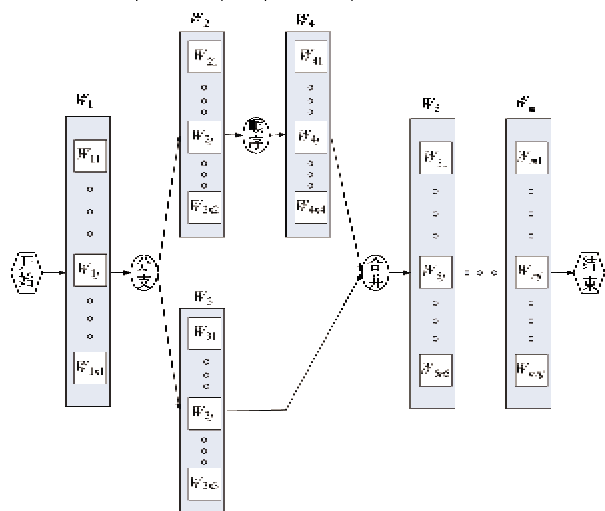


图1 候选服务与服务类示意图

服务类:具有相同功能和不同 QoS 的候选服务构成一个服务类 W_i , 如图 1 中的 W_1 、 W_2 、 W_3 等。

组合服务:组合服务(CS)由 $W_1, W_2, \dots, W_m, m$ 个服务类按照某种特定的业务逻辑组成,每个服务类 $W_i(i = 1, 2, \dots, m)$ 中有 j 个候选服务,且 W_i 中只有一个 $W_{ij}(j = 1, 2, \dots, n)$ 被选中,组合服务的 $Q_{cs} = (Q_{cs}^1, Q_{cs}^2, \dots, Q_{cs}^n)$ 满足用户的要求。

执行路径:对由 m 个服务类 $\{W_1, W_2, \dots, W_m\}$ 构成的组合服务,执行路径 P 是一个由 m 个二元组组成的集合 $\{ \langle W_1, W_{1n_1} \rangle, \langle W_2, W_{2n_2} \rangle, \dots, \langle W_m, W_{mn_j} \rangle \}$, 其中 $W_{mn_j}(j = 1, 2, \dots, n)$ 是服务类 W_i 的一个候选服务, n_j 是 j 的扩展,是每个服务类中被选中的候选服务的下标。

3 组合服务流程基本模型及其 QoS 计算方法

本文将组合服务建模为基于服务的工作流,将工作流的 4 种基本模型作为组合服务的基本模型,大部分组合服务流程都可以由这 4 种基本模型组合而成,组合服务的 QoS 参数也可以通过基本模型的 QoS 计算方法来得到。

目前,服务的通用 QoS 主要有执行时间 T (time)、执行代价 C (cost)、可靠性 R (reliability)、可用性 A (availability) 和信誉等级 Rep (reputation) 等,在军事综合电子信息系统领域,服务的一些特殊质量指标是服务选择的重要依据,因此我们选择 5 种 QoS 参数来描述和度量军事信息服务,即执行时间 T 、执行代价 C 、可靠性 R 、精度 A (accuracy) 和信誉等级 Rep ,其中精度以各种误差来度量,如目标探测误差、火控解算误差、射击概率误差等。设 CS 为多个服务形成的组合服务, W_i 为组合服务的服务类, W_i 和 CS 的服务质量模型分别为 $Q_{W_i} = (T_i, C_i, R_i, A_i, Rep_i)$, $Q_{CS} = (T_{CS}, C_{CS}, R_{CS}, A_{CS}, Rep_{CS})$, 组合服务流程基本模型如图 2,其 QoS 参数计算方法如下:

$$\text{顺序: } T_{cs} = \sum_{i=1}^m T_i, C_{cs} = \sum_{i=1}^m C_i, A_{cs} = \sum_{i=1}^m A_i,$$

$$Rep_{cs} = \frac{\sum_{i=1}^m Rep_i}{n}。$$

$$\text{并行: } T_{cs} = \text{Max}(T_1, T_2, \dots, T_m), C_{cs} = \sum_{i=1}^m C_i,$$

$$R_{cs} = \text{Min}(R_1, R_2, \dots, R_m), A_{cs} = \text{Max}(A_1, A_2, \dots,$$

$$A_m), Rep_{cs} = \frac{\sum_{i=1}^m Rep_i}{n}。$$

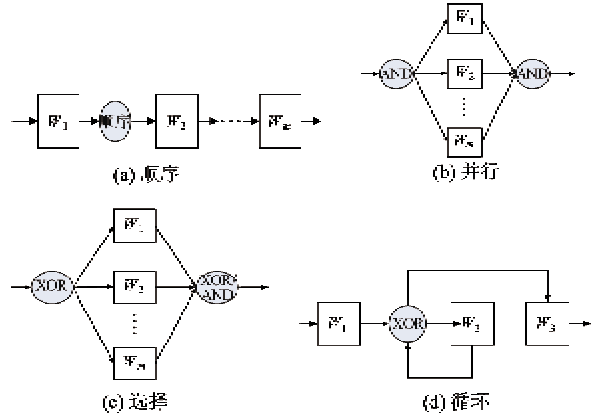


图 2 组合服务流程基本模型

选择:设第 i 个分支被选中的概率是 α_i , $\sum_{i=1}^m \alpha_i$

$$= 1, \text{ 则 } T_{cs} = \sum_{i=1}^m T_i \alpha_i, C_{cs} = \sum_{i=1}^m C_i \alpha_i, R_{cs} = \sum_{i=1}^m R_i \alpha_i,$$

$$A_{cs} = \sum_{i=1}^m A_i \alpha_i, Rep_{cs} = \sum_{i=1}^m Rep_i \alpha_i。$$

循环:设循环次数为 k , 则 $T_{cs} = k \cdot \sum_{i=1}^m T_i, C_{cs} =$

$$k \cdot \sum_{i=1}^m C_i, R_{cs} = R_i^k, A_{cs} = k \cdot \sum_{i=1}^m A_i, Rep_{cs} = \frac{\sum_{i=1}^m Rep_i}{n}。$$

4 基于 ICA 的多目标分布估计算法

在面向服务的军事综合电子信息系统中,执行时间、可靠性和精度对完成作战任务十分重要,因此将执行时间、可靠性和精度作为三个目标准则,希望组合流程执行时间极短、可靠性和精度极高;费用和信誉等级作为两个约束条件, Rep_0 和 C_0 分别表示所要求的组合服务路径的最低信誉等级和最大费用值。对于 QoS 属性来说,有些是越大越好,有些越小越好,在此我们统一求 $Minf(x)$ 。其中精度通常以各种误差来度量,误差越小则精度越高,因此目标函数精度极高就变成了误差极小。在可靠性计算公式的前面加一负号将可靠性极高转化为负的可靠性极低,则带约束条件的多目标服务组合优化问题描述如下:

$$MinF(P) = (T(P), A(P), -R(P)) \quad (1)$$

式中执行路径 P 受两个条件约束,即

$$\begin{cases} C(P) \leq C_0 \\ Rep(P) \geq Rep_0 \end{cases} \quad (2)$$

$T(P), R(P), A(P), Rep(P)$ 和 $C(P)$ 对应组合服

务执行路径 P 的 QoS 参数计算公式,可以通过组合服务基本模型 QoS 计算方法来得到。优化解集中各条路径之间的比较采用了相对优概念,采用文献[9]的个体优劣比较准则来判断解空间中不同个体的相对优关系。

4.1 ICA 模型

ICA 是非高斯数据的一种线性变换,对观测信号进行某种线性分解后得到统计独立的成分,因此 ICA 假设独立成分是非高斯分布的。例如 $\mathbf{x} = \mathbf{A}_c \mathbf{s}_c + \mathbf{e}_c$ 中, $\mathbf{x}$ 为 S 维观测信号向量, $\mathbf{A}_c$ 为未知的 $S \times L_c$ ($L_c \leq S$) 信号的混合矩阵, $\mathbf{s}_c$ 为独立的 L_c 维未知源信号向量, $\mathbf{e}_c$ 为均值为 0 的 S 维加性噪声, λ_c 为其特征值, c 代表第 c 个 ICA 模型。

通常使用信号的非高斯性作为分离信号之间独立性的度量,常用度量方法有非高斯的最大化、微熵、负熵、互信息的最小化和最大似然估计等。这里我们采用负熵作为度量非高斯性的方法,即 $J(\mathbf{x}) = H(\mathbf{x}_{\text{Gauss}}) - H(\mathbf{x})$, $\mathbf{x}_{\text{Gauss}}$ 为高斯随机向量, $\mathbf{x}_{\text{Gauss}}$ 与 $\mathbf{x}$ 的协方差相同。负熵计算较困难,常采用负熵的近似。将每一代种群中所有个体的目标函数作为一个序列,当做 ICA 的输入,输出为分离出的独立分量。

4.2 算法框架

步骤 1:初始化:令 $t = 0$, 产生一个初始种群 $\text{Pop}(0)$, 设定算法的最大迭代次数,并计算在 $\text{Pop}(0)$ 中每个个体的目标函数值。

步骤 2:停止条件:如果停止条件(达到最大迭代次数)满足,则停止运算并返回 $\text{Pop}(t)$ 中的非支配解和它们对应的目标函数值。

步骤 3:建模:采用 ICA 对 $\text{Pop}(t)$ 中的解进行建模,获得 $\text{Pop}(t)$ 的概率模型。

步骤 4:重采样:利用步骤 3 建立的概率模型产生新解集 Q , 评价 Q 中每个解的目标函数值。

步骤 5:选择:从 $Q \cup \text{Pop}(t)$ 中选择 N 个解产生 $\text{Pop}(t+1)$ 。

步骤 6:令 $t = t + 1$, 跳转到步骤 1。

4.3 基于 ICA 的建模

我们基于最大熵原理来近似负熵,其中:

$$J(\mathbf{x}) \approx k_i [E\{G_i(\mathbf{x})\} - E\{G_i(\mathbf{v})\}]^2 \quad (3)$$

这里 k_i 是常数且大于零, $\mathbf{v}$ 是均值为零、方差为 1 的高斯变量。此处假设 $\mathbf{x}$ 是均值为零、方差为 1 的变量, G_i 是某一非二次函数。为了提高精度,我们仅用非二次函数 G_i , 即

$$J(\mathbf{x}) \propto [E\{G(\mathbf{x})\} - E\{G(\mathbf{v})\}]^2 \quad (4)$$

为了提高近似度,我们合理地选择 G_i :

$$G_1(\mathbf{x}) = \frac{1}{a_1} \log \cosh a_1 \mathbf{x} \quad (5)$$

$$G_2(\mathbf{x}) = -\exp(-\mathbf{x}^2/2) \quad (6)$$

其中常数 a_1 在 $1 \leq a_1 \leq 2$ 范围内取比较合适,通常取 1。接着采用 fastICA^[18] 算法,其算法框架如下:

步骤 1:对信号中心化使其均值为 0;步骤 2:对信号进行白化得到 $\mathbf{z}$;步骤 3:选择要估计的独立成分的个数 L_c ;步骤 4:初始化所有的 $\mathbf{w}_i, i = 1, 2, \dots, L_c$, 其中每个 $\mathbf{w}_i$ 都具有单位范数,用下面的步骤 6 的方法对矩阵 $\mathbf{W}$ 进行正交化;步骤 5:对于每个 $i = 1, 2, \dots, L_c$, 更新 $\mathbf{w}_i$:

$$\mathbf{w}_i \leftarrow E\{z(g(\mathbf{w}_i^T \mathbf{z}))\} - E\{g'(\mathbf{w}_i^T \mathbf{z})\} \mathbf{w}_i \quad (7)$$

函数 g 的定义如式

$$g_1(x) = \tanh(a_1 x) \quad (8)$$

$$g_2(x) = x \exp(-x^2/2) \quad (9)$$

$$g_3(x) = x^3 \quad (10)$$

所示。常数 a_1 取值同公式(5);步骤 6:对矩阵 $\mathbf{W} = (\mathbf{w}_1, \dots, \mathbf{w}_{L_c})^T$ 进行正交化: $\mathbf{W} \leftarrow (\mathbf{W}\mathbf{W}^T)^{-1/2} \mathbf{W}$;步骤 7:标准化 $\mathbf{w}_i$, 即 $\mathbf{w}_i \leftarrow \mathbf{w}_i / \|\mathbf{w}_i\|$;步骤 8:如果尚未收敛,则返回公式(7),继续更新 $\mathbf{w}_i$ 。

4.4 重采样

利用 ICA 建立的概率分布模型,产生新解集,评价 Q 中每个解的目标函数值。首先随机产生一个整数 $r, r \in \{1, 2, \dots, L_c\}$, 接着随机产生一个信号向量 $\mathbf{s}', \mathbf{s}'$ 基于 ICA 建立的概率分布模型,随机产生一个噪声向量 $\mathbf{e}', \mathbf{e}' \in N(0, \sigma^{L_c} \mathbf{I})$, 令 $\mathbf{x}' = \mathbf{s}' + \mathbf{e}'$ 。按此方法重复 N 次,就可产生 N 个新解。

4.5 选择

选择操作是基于拥挤距离排序和 NSGA-II 的非支配排序^[16]。

5 实验设计及结果

5.1 实验设计

以面向服务的海战场综合电子信息系统为例,验证算法的可行性和有效性。如图 3 所示,海战场综合电子信息系统含有 8 个服务类,分别为 $\mathbf{W}_1$ 水面

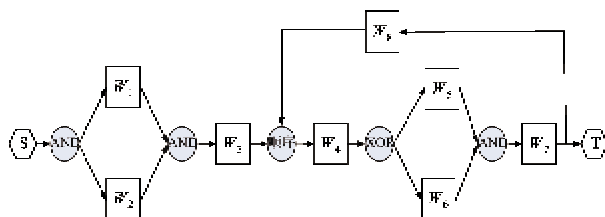


图3 面向服务的海战场综合电子信息系统组合流程

目标探测服务、 W_2 水下目标探测服务、 W_3 信息处理服务、 W_4 火控解算服务、 W_5 导弹发控服务、 W_6 鱼雷发控服务、 W_7 效能评估服务和 W_8 重复攻击服务。 S 和 T 分别为开始和结束服务。

我们设计了面向服务海战场综合电子信息系统的半实物仿真系统,实验环境为 100M 局域网加无线电通信网,选择算法采用 C++ 进行编程,运行在 CPU 为 Intel(R) Core(TM2) 2.50GHz 的计算机上。关联关系的生成过程如下:(1)随机选择一定比例的候选服务作为支持服务关联的候选服务,为其每个 QoS 属性生成两个 QoS 值,其他候选服务只生成一个 QoS 值;(2)确定与本服务类有前后关联关系的两个服务类,比如对于信息处理服务,选择火控解算和两类目标探测服务;(3)在选择的服务类中随机选择 4~12 个候选服务;(4)对于第一步生成的支持服务关联的候选服务,将其中一个作为默认的 QoS 值,另外一个作为依赖于第(3)步选择的候选服务的 QoS 值。参数取值范围假定为 $0 < T \leq 20s$, $0 < C \leq 1000$ 万元, $0 < Rep \leq 5$, $0 < R \leq 1$, $0 < A \leq 50m$, 信誉等级最小为 2,可靠性最小为 0.2。

5.2 确定目标函数和约束函数

计算如图 3 所示的组合服务流程确定其 QoS 参数计算公式,作为算法的目标函数和约束函数。其中 W_1 和 W_2 构成一个并联模型,然后和 W_3 串联, W_3 和 W_6 是选择关系,根据作战效能评估结果判定作战任务没有完成时,需要再次发起攻击,所以 W_8 与 W_4 、 W_5 、 W_6 和 W_7 是循环关系,循环次数 k 是正整数,一般不会很大。算法目标函数和约束函数如下:

$$T_p = \text{Max}(T_1, T_2) + T_3 + k[T_4 + \alpha T_5 + (1 - \alpha)T_6 + T_7 + T_8]$$

$$R_p = \text{Min}(R_1, R_2)R_3R_4^k[\alpha R_5 + (1 - \alpha)R_6]^kR_7^kR_8^k$$

$$A_p = \text{Max}(A_1, A_2) + A_3 + k[A_4 + \alpha A_5 + (1 - \alpha)A_6 + A_7 + A_8]$$

$$C_p = \sum_{i=1}^3 C_i + k[C_4 + \alpha C_5 + (1 - \alpha)C_6 + C_7 + C_8]$$

$$Rep_p = (\sum_{i=1}^2 Rep_i + 2 \sum_{j=3}^4 Rep_j + 2[\alpha Rep_5 + (1 - \alpha)Rep_6] + 2 \sum_{m=7}^8 Rep_m)/12$$

从执行时间 T_p 和精度 A_p 的计算式可以看出,二者是一样的,因此原来的三个目标函数就变成了双目标函数。

5.3 实验分析

5.3.1 有效性实验

通过算法运行的 CPU 执行时间可以验证基于 ICA 的多目标分布估计算法的有效性。实验分别考虑了服务类中候选服务规模为 4,8 和 12,进化代数取 100,200,300 和 400, $k = 2, \alpha = 0.4$ 的情况下,利用基于 ICA 的多目标分布估计算法求解满足约束的 Pareto 优化服务执行路径的 CPU 执行时间。对每一种情况,算法分别运行 30 次取平均值。由图 4 可以看出:随着候选服务数目的增加,在不同的进化代数下,CPU 执行时间并没有大量增加。本文研究的面向服务海战场综合电子信息系统应用流程包含 8 个服务类,在候选服务规模为 12 的情况下,算法迭代 400 次得到优化解的执行时间在 18s 左右,这一速度可以满足大部分服务组合求解需要和作战要求。

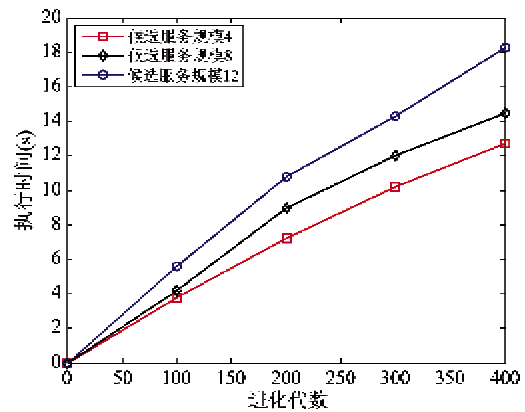


图4 ICAMOEDASS 平均执行时间

5.3.2 可行性实验

实验目的是验证基于 ICA 的多目标分布估计算法找到 QoS 全局最优服务执行路径的可行性。我们与文献[2]的支持服务关联关系及求解次优解的启发式算法做对比。图 5 为在候选服务规模分别为 4,8,12, $k = 2, \alpha = 0.4$ 的情况下,基于 ICA 的多

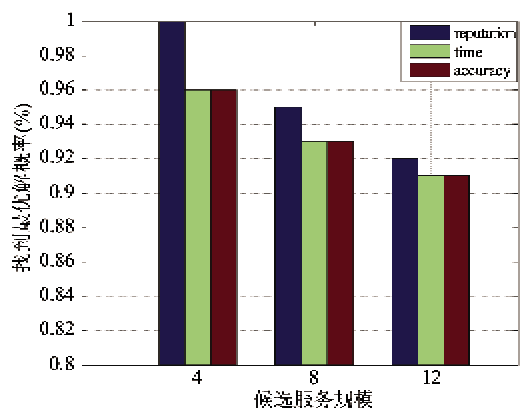


图5 找到 Pareto 最优解概率

目标分布估计算法迭代 400 代所获得的各优化指标的全局最优结果的百分率。百分率定义为在相同服务类和候选服务规模的条件下算法运行 30 次所得非劣解集中存在某优化指标的启发式算法最优值的统计概率。从图 5 可以看出:在 400 代条件下找到各优化指标最优值的概率都在 91% 以上,因此基于 ICA 的多目标分布估计算法解决 QoS 全局优化问题是可行的。

候选服务规模为 12, $k = 2, \alpha = 0.4$, 算法迭代 400 代产生的一组 Pareto 最优解如图 6 所示。从单个目标准则来看,这些解可能不是最优的,但同时从服务执行路径的执行时间、精度和可靠性来看,这些解是非劣的且分布也是均匀的。指挥员可以根据实际经验和作战任务需要选择最合适的 Pareto 解,因此 ICA 多目标分布估计算法是有有效的。

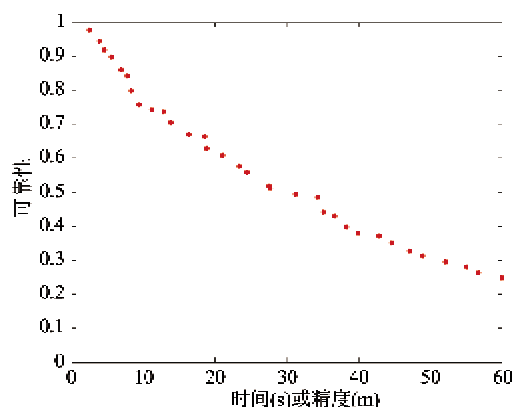


图 6 Pareto 最优解集合

5.3.3 有无关联关系对组合服务 QoS 的影响

我们这里把基于 ICA 的多目标分布估计算法与假设服务之间是相互独立的选择算法进行比较,来说明考虑服务之间关联关系对于组合服务选择问题的重要性和有效性。我们选择文献[9]的 GODSS 算法与基于 ICA 的多目标分布估计算法进行比较。图 7 为在候选服务规模分别为 4, 8, 12, $k = 2, \alpha = 0.4$ 的情况下 GODSS 与基于 ICA 的多目标分布估计算法分别迭代 400 代的执行时间比较,这里的时间为两种算法分别运行 30 次取平均。从图 7 可以看出:因为考虑了服务之间的关联关系,需要先采用 ICA 对目标函数进行处理,所以基于 ICA 的多目标分布估计算法的执行时间要大于 GODSS,但执行时间增加较为缓慢。图 8 为在服务群规模为 12、进化代数 400 的条件下, GODSS 与基于 ICA 的多目标分布估计算法所得的最优解的分布情况,可以看出,基于 ICA 的多目标分布估计算法所求得的最优解

性能和分布要优于 GODSS。因此基于 ICA 的多目标分布估计算法虽然计算复杂性高,但其得到的解质量好,适合对 QoS 性能要求严格的服务选择问题; GODSS 由于缺乏对这种关联关系的考虑,使得服务质量在关联关系出现时与其所生成的 QoS 值产生偏离,从而导致所选择的组合服务实际的 QoS 有所下降。

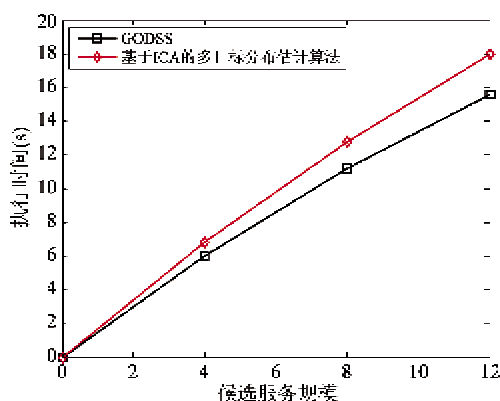


图 7 基于 ICA 的多目标分布估计算法和 GODSS 平均执行时间

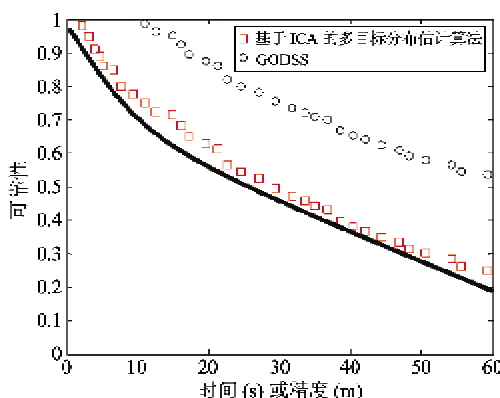


图 8 基于 ICA 的多目标分布估计算法和 GODSS Pareto 最优解集合

6 结论

针对军事综合电子信息系统服务选择问题,提出了一种支持服务关联的全局最优服务选择方法。本算法有以下优点:(1)考虑了服务之间的关联关系,减小了 QoS 值的偏离,更适用于对 QoS 性能要求严格的军事信息服务选择;(2)不是通过建立支持服务关联关系的 QoS 描述模型,而是通过 ICA 分离出独立的各分量来处理关联关系,简化了工作量,提供了解决问题的新思路;(3)提出的基于 ICA 的多目标分布估计算法用于服务选择,既考虑

了非高斯分布情形,又采用 fastICA 方法,在计算量增加的情况下,计算时间却增加缓慢,这点对实时性要求较高的作战指挥和军事综合电子信息系统的运用特别重要。实验结果表明,基于 ICA 的多目标分布估计算法具有良好的有效性和可行性。

参考文献

- [1] 舒振,刘静,罗雪山等. 军事信息服务的概念体系及其应用. 火力与指挥控制,2010,35(9):53-57
- [2] 叶世阳,魏峻,李磊等. 支持服务关联的组合服务选择方法研究. 计算机学报,2008,31(8):1383-1397
- [3] 陈世展,冯志勇,王辉. 服务关系及其在面向服务计算中的应用. 计算机学报,2010,33(11):2068-2083
- [4] 张明卫,魏伟杰,张斌等. 基于组合服务执行信息的服务选取方法研究. 计算机学报,2008,31(8):1398-1411
- [5] Zeng L Z, Benattallah B, Ngu A, et al. QoS-aware middleware for web services composition. *IEEE Transactions on software and engineering*, 2004, 30(5): 311-327
- [6] Yu T, Zhang Y, Lin K J. Efficient algorithms for Web services selection with end-to-end QoS constraints. *ACM Transactions on Web*, 2007, 1(1): 6-32
- [7] 杨放春,苏森,李祯. 混合 QoS 模型感知的语义 Web 服务组合策略. 中国科学 E 辑:信息科学,2008,38(10):1697-1746
- [8] Zeng L Z, Benattallah B, Dumas M, et al. Quality driven Web services composition. In: *Proceedings of the 12th International Conference on World Wide Web*. New York, USA, 2003. 411-421
- [9] 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
- [10] 张成文,苏森,陈俊亮. 基于遗传算法的 QoS 感知的服务选择. 计算机学报,2006,29(7):1029-1037
- [11] 王勇,代桂平,姜正涛等. 服务组合中信任感知的成员服务选择算法. 高技术通讯,2010,20(8):798-802
- [12] Hauschild M, Pelikan M. An introduction and survey of estimation of distribution algorithms. *Swarm and Evolutionary Computation*, 2011, 1: 111-128
- [13] 周树德,孙增圻. 分布估计算法综述. 自动化学报,2007,33(2):113-124
- [14] Pelikan M, Sastry K, Goldberg D E. Multiobjective hBOA, clustering, and scalability. In: *Proceedings of Genetic and Evolutionary Computing Conference*, New York, USA, 2005. 663-670
- [15] Bosman P A N, Thierens D. The naive MIDEA: a baseline multi-objective EA. In: *Proceedings of the 3rd International Conference on Evolutionary Multi-Criterion Optimization*. Guanajuato, Mexico, 2005. 428-442
- [16] Zhang Q F, Zhou A M, Jin Y C. RM-MEDA: a regularity model-based multiobjective estimation of distribution algorithm. *IEEE Transactions on Evolutionary Computation*, 2008, 12(1): 41-63
- [17] 杨竹青,李勇,胡德文. 独立成分分析方法综述. 自动化学报,2002,28(5):762 - 772
- [18] Hyvarinen, A. Fast and robust fixed-point algorithms for independent component analysis. *IEEE Transactions on Neural Networks*, 1999, 10(3): 626 - 634
- [19] Cho D Y, Zhang B T. Continuous estimation of distribution algorithms with probabilistic principal component analysis. In: *Proceedings of the Congress on Evolutionary Computation*, Seoul, Korea, 2001. 521-526
- [20] Zhang Q F, Allinson N M, Yin H J. Population optimization algorithm based on ICA. In: *Proceedings of the First IEEE Symposium on Combinations of Evolutionary Computation and Neural Networks*, San Antonio, USA, 2000. 33-36
- [21] Cho D Y, Zhang B T. Evolutionary continuous optimization by distribution estimation with variational Bayesian Independent Component Analyzers mixture model. In: *Lecture Notes in Computer Science*, Birmingham, UK, 2004. 212-221

A service correlation-supporting algorithm for global optimization of service selection

Nie Kai, Wang Houxiang, He Jiao

(College of Electronic Engineering, Naval University of Engineering, Wuhan 430033)

Abstract

Aiming at the problem that a service-oriented military electronic information system's QoS of a candidate service relies on other candidate services, a method for global optimization of service selection with the consideration of service correlation is proposed. The method transforms the problem of dynamic web service selection with QoS global optimal into a multiobjective services composition optimization with QoS constraints, and the multiobjective estimation of distribution algorithm based on independent component analysis (ICA) is utilized to produce a set of optimal Pareto services composition with constraint principle by means of optimizing various objective functions simultaneously. The results of the compared experiment indicate that the multiobjective estimation of distribution algorithm based on ICA can describe the correlation between candidate services through the probabilistic graphical model, and the algorithm has the better performance in solving the service selection problem with service correlation, which confirms its feasibility and efficiency.

Key words: military information service composition, service selection, service correlation, multiobjective optimization, estimation of distribution algorithm (EDA), independent component analysis (ICA)