

带准备时间和截止期约束的云服务 workflow 调度算法

沈虹^{1,2}, 李小平¹

(1. 东南大学 计算机科学与工程学院, 江苏 南京 210096; 2. 南京审计学院 金审学院, 江苏 南京 210029)

摘要: 带准备时间和截止期约束的云服务 workflow 费用优化是一个新的云计算资源优化分配问题。分析该 NP-hard 问题特征, 建立相应的整数规划数学模型。构建有效的变量取值概率模型和更新机制, 提出高质量初始群体的启发式生成方法; 提出混合的分布估计算法 (HEDA), 引入个体向全局最优解学习的策略, 提高算法的全局搜索和局部优化能力。模拟实验结果表明此提出的方法在合理的 CPU 时间内可有效减少 workflow 费用。

关键词: 云服务; workflow 调度; 分布估计算法

中图分类号: TP393

文献标识码: A

Algorithm for the cloud service workflow scheduling with setup time and deadline constraints

SHEN Hong^{1,2}, LI Xiao-ping¹

(1. School of Computer Science & Engineering, Southeast University, Nanjing 210096, China;

2. Jinshen College, Nanjing Audit University, Nanjing 210029, China)

Abstract: Cost minimization with deadline constraints and setup time is a new resource allocation optimization problem in cloud computing. An integer programming model is constructed for this NP-hard problem. A novel probability model and updating mechanism is suggested for sampling the solution space. A heuristic method is investigated for generating good initial population. A hybrid estimation of distribution algorithm (HEDA) is proposed for the considered problem. The global best solution is adopted to balance the diversification and intensification of HEDA. Experimental results show that the proposed method is effective to optimize the workflow running cost within reasonable CPU time.

Key words: cloud service; workflow scheduling; estimation of distribution algorithm

1 引言

云计算是分布式计算、并行计算和网格计算的商业化发展。通过虚拟化技术将分布在不同地理位置的计算资源和存储资源封装成服务, 向互联网应用提供基础设施即服务(IaaS)、平台即服务(PaaS)和软件即服务(SaaS)的三大服务模式^[1]。近年来, 基于云计算环境实现跨平台、跨部门、跨地区多任务间的协同处理已成为热点科学问题。

云计算环境下的许多大型应用常被分解成若干个彼此联系的活动(任务), 活动及活动间的复

杂约束关系通过 workflow 来表示。为实现 workflow 任务间的协同管理, 需为各任务选择合适的资源, 以满足不同的 workflow 目标。图 1 给出云计算平台 workflow 调度框架结构。云平台的服务提供商向用户提供多种性能的资源 (CPU、存储、数据库、I/O 等), 不同类型的资源具有不同的 QoS 特性。调度管理中心根据用户提出的优化目标 (最小化完工时间、最小化租赁费用等), 为 workflow 中各活动选择需要的服务 (虚拟机、磁盘存储、数据库管理等)。

云服务工作流调度问题的本质是实现 workflow

收稿日期: 2014-05-21; 修回日期: 2014-10-12

基金项目: 国家自然科学基金资助项目(61272377); 教育部高等学校博士学科点专项科研基金资助项目 (20120092110027)

Foundation Items: The National Natural Science Foundation of China (61272377); The Research Fund for the Doctoral Program of Higher Education of China (20120092110027)

活动与资源的映射，使 workflow 调度目标最优化。不同类型的资源具有不同租赁成本。通常情况下，计算和存储性能高的资源具有更高的租赁费用，计算和存储性能低的资源租赁费用则便宜些，即执行时间短的服务一般具有更高的租赁费用^[2]。该问题是离散的时间费用优化问题，大部分情况下是 NP-hard 问题^[3]。

离散的时间费用优化问题 (DTCTP, discrete time-cost tradeoff problem) 在项目管理领域已得到广泛研究。已有研究成果多采用分支定界、动态规划等精确方法求解，文献[4]详细介绍了该问题的研究现状。近 10 年，网络计算环境下的 workflow 调度优化方法主要包括启发式方法和元启发式方法两类。Yu 等^[5]基于不同截止期/成本约束提出求解独立任务调度的启发式算法 DTL。苑迎春等^[6,7]提出截止期约束的逆向分层调度算法 DBL，该算法根据各活动的逆向深度分层，将 workflow 截止期转化为活动的时间区间，通过增大活动的时间区间以优化费用。苑迎春等^[8]利用 DAG 图中串行活动特征给出串归约概念，基于分层算法对串归约组的时间窗口重定义，并提出动态规划的求解策略实现组内费用最优化。Abrishami 等^[9]提出基于部分关键路径的启

发式方法，以截止期为 workflow 完工时间，采用后向迭代方式为各活动查找部分关键路径，并为部分关键路径上各活动设置子截止期；在子截止期的约束下为各活动选择最便宜的服务。张晓东等^[10]提出离散的粒子群算法，引入粒子随机极值扰动、位置回退、交叉繁殖等机制解决服务 workflow 调度优化。Chen 等^[11]用蚁群算法求解具有不同 QoS 约束的 workflow 调度优化。Yu 等^[12]用遗传算法解决截止期约束的 workflow 费用优化和成本约束的 workflow 完工时间最小化 2 个问题。

云计算环境下 workflow 调度与网格环境 workflow 调度的本质区别在于：云计算环境的资源可认为是无限的，用户可在任何时刻采用按需付费的方式租用资源。现有文献对云计算环境下的 workflow 调度问题主要考虑截止期约束的 workflow 费用最小化问题，多采用启发式的方法。Cai 等^[13]提出基于关键路径的迭代启发式方法，利用动态规划法优化关键路径上活动的服务选择方案，最小化 workflow 费用。Abrishami 等^[14]用部分关键路径的启发式方法将 workflow 截止期转化为各活动的子截止期，在满足活动子截止期的情况下，为活动选择合适的服务以最小化 workflow 租赁费用。Van den Bossche 等^[15]考虑公

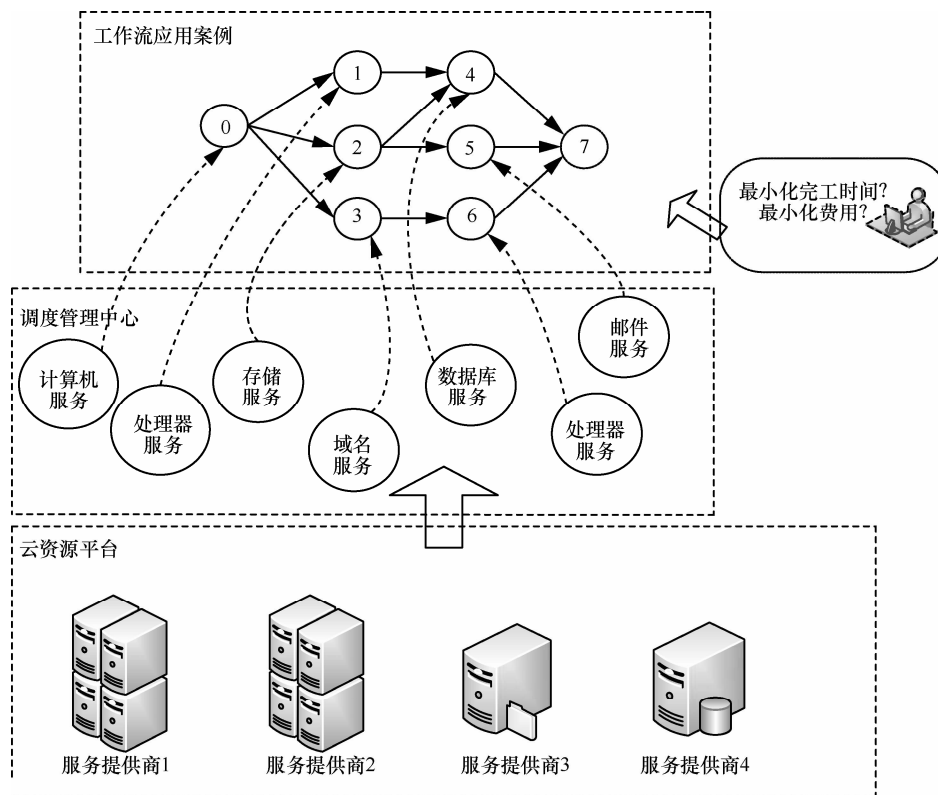


图 1 云计算平台 workflow 调度框架结构

有云和私有云混合环境下的截止期约束 workflow 费用最小化问题。启发式方法具有执行时间短、优化速度快等优点，但通常情况下得到问题的近似最优解。

与已有研究不同，本文考虑云计算环境下截止期约束具有准备时间的工作流费用最小化问题，即实际云计算 workflow 活动在执行时具有不同的准备时间（如初始化虚拟机、数据传输等），该问题是对已有研究模型^[13-15]的扩展，是 NP-hard 问题。分布估计算法^[16]（EDA, estimation of distribution algorithm）是一个基于概率统计的群智能优化算法，已被成功地运用于项目调度优化^[17]，该算法在云服务 workflow 调度领域尚无相关研究。本文结合问题特征构造启发式方法生成高质量的初始解，提出混合的分布估计算法（HEDA）进行求解，构建有效的问题变量概率向量，将粒子群算法中个体向全局最优解学习的思想引入 HEDA 中，提出新的变量取值统计方法以更新变量概率向量。实验结果表明本文提出的混合分布估计算法能更好地平衡全局搜索和局部优化过程。

2 带截止期和准备时间的云服务 workflow 优化模型

云计算环境下的应用通常由多个活动相互协作完成，活动及活动间的约束关系可模型化为有向无环图 DAG 描述的工作流。有向无循环图 DAG 记作 $G = \{V, E\}$ ，其中节点集 $V = \{0, 1, 2, \dots, n, n+1\}$ 表示图中所有活动集合； $E = \{(i, j) | i < j\}$ 是有向边集合，表示活动间的偏序约束关系，即对 $\forall (i, j) \in E$ ，活动 j 在活动 i 执行完后才可开始执行。其中 0 号节点

为开始节点标识用户向云平台提交 workflow 应用请求。 $n+1$ 号活动为虚节点标志 workflow 的结束，设 R_i 表示活动 i 的直接先驱活动集，则 $|R_i|$ 为直接先驱活动个数；设 Q_i 表示其直接后继活动集，则 $|Q_i|$ 为直接后继活动个数。 l_i 表示活动 i 的开始执行时间， f_i 表示活动 i 的完成时间。

workflow 的每个活动 i ，有若干可选服务（如虚拟机资源）可完成该任务。可选服务集称为活动 i 的服务池 $S_i = \{S_i^1, S_i^2, \dots, S_i^{m_i}\}$ 。每个服务 S_i^k 完成活动 i 所需时间包括准备时间 r_i^k 和执行时间 t_i^k 两部分，记 $d_i^k = r_i^k + t_i^k$ ，所需费用为 c_i^k 。各活动服务池中的可选服务按照时间 d_i^k 递增、费用 c_i^k 递减的顺序排序。图 2 是一个具有 12 个活动的工作流实例，0 号节点上方为初始化 workflow 的时间和费用，其余节点上方为该活动的可选服务构成的服务池 $\{(d_i^k, c_i^k)\}$ 。

截止期约束具有准备时间的云服务 workflow 费用优化问题可描述为：给定 workflow 的截止期 δ_{n+1} ，为各活动选择合适的服务，使 workflow 费用最小化。定义布尔变量 $x_{ik} \in \{0, 1\}$ ， $\forall i \in V, \forall k \in \{1, 2, \dots, m_i\}$ ，如果活动 i 选择服务池中的第 k 个服务来执行，则 $x_{ik} = 1$ ；否则 $x_{ik} = 0$ 。用整数规划方法对本问题形式化描述如下

$$\text{Minimize } \sum_{i \in V} \sum_{1 \leq k \leq m_i} x_{ik} c_i^k \quad (1)$$

$$\text{s.t. } \sum_{k=1}^{m_i} x_{ik} = 1, \forall i \in V \quad (2)$$

$$f_j - f_i - \sum_{k=1}^{m_i} x_{ik} \cdot (r_i^k + t_i^k) \geq 0, \forall (i, j) \in A \quad (3)$$

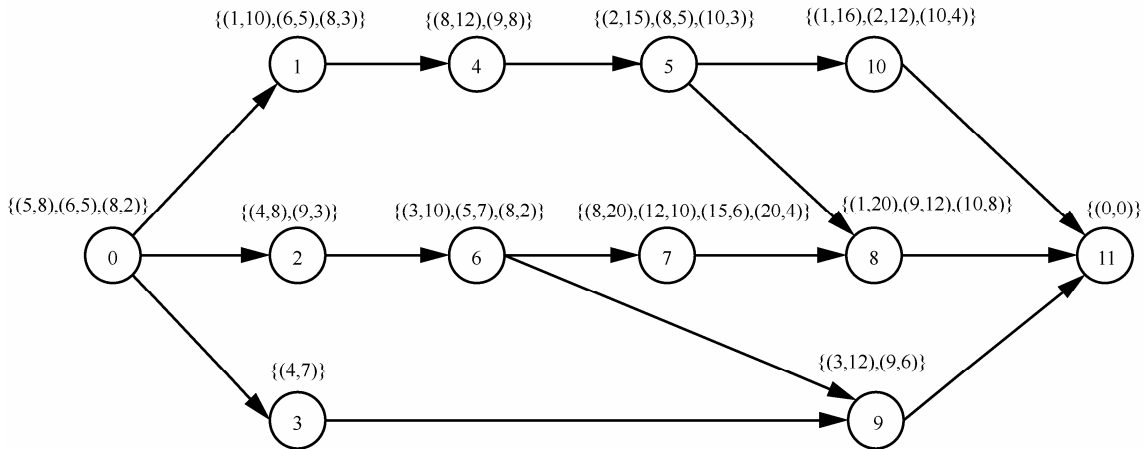


图 2 具有 12 个活动的工作流 DAG 图

$$f_{n+1} \leq \delta_{n+1} \quad (4)$$

$$x_{ik} \in \{0,1\}, \forall i \in V, \forall k \in \{1,2,\dots,m_i\} \quad (5)$$

$$t_i^k \in I^+, \forall i \in V, \forall k \in \{1,2,\dots,m_i\} \quad (6)$$

$$r_i^k \in I^+, \forall i \in V, \forall k \in \{1,2,\dots,m_i\} \quad (7)$$

式(1)表示本问题的目标是最小化工作流的执行费用；式(2)约定每个活动只能选择一个合适的服务来执行；活动间的偏序约束关系在式(3)中给出；式(4)是工作流截止期约束。

3 提出的混合分布估计优化方法

分布估计算法通过概率统计的方法来预测群体在解空间中的搜索方向，根据精英解计算各变量在解空间中取值的概率分布，并用该概率分布值指导新群体的生成。与其他智能算法（如遗传算法、粒子群算法）相比，它避免了这些算法在搜索过程中的随机性，可生成更高质量的子代群体。然而，分布估计算法初始群体中个体以变量在解空间中中等概率取值的方式获得，群体多样性好但质量较差，尤其在解决多约束条件的优化问题时初始群体的不可行解数量多。为了提高初始群体中可行解的数量和质量，本文构建基于问题特征的启发式算法生成初始群体 Ω ，结合粒子群算法子代群体同时向过去最优解和全局最优解学习的思想提出新颖的精英选择方法，在更新变量概率分布向量时不仅统计过去最优解中变量的取值情况，还统计全局最优解中变量的取值情况，用混合分布估计算法求解带准备时间和截止期约束的云服务工作流调度问题。

3.1 解的表示与概率分布向量

混合分布估计算法群中每个个体代表问题的一个解，即工作流各活动所选择的服务。假设工作流中的任何一个活动 i ，可完成该活动的服务集合所构成的服务池为 $S_i = \{S_i^1, S_i^2, \dots, S_i^{m_i}\}$ ，群 Ω 中每个个体 I_j 代表问题的一个解， $I_j = (I_{j0}, I_{j1}, I_{j2}, \dots, I_{jn})$ ，其向量长度为 $n+1$ ，其中 I_{j0} 表示用于初始化工作流的服务编号， I_{j1} 表示 1 号活动选取的服务编号， I_{j2} 表示 2 号活动选取的服务编号，以此类推。显然， I_j 中的每一维 I_{ji} 的取值满足 $1 \leq I_{ji} \leq m_i$ 。基于 I_j 的服务方案，采用前向调度方法 (FSM) 计算各活动的完工时间：节点 0 的开始

时间为 0，如果一个活动 i 所有的先驱活动已经完成，则 $l_i = \max_{a \in R_i} \{f_a\}$ ， $f_i = l_i + d_i^{I_{ji}}$ 。虚节点 $n+1$ 号的结束时间 f_{n+1} 为工作流的完工时间。如果该完工时间满足截止期的约束，则为可行解，否则为不可行解。工作流的执行费用依据式(1)中目标函数的计算得到。

混合分布估计算法中概率分布向量表示所求解问题各变量取值在解空间中的概率分布，根据本问题特点，需构造 $n+1$ 个向量，第 i 个向量 $P_i^t = (P_{i0}^t, P_{i1}^t, P_{i2}^t, \dots, P_{im_i}^t)$ 表示算法第 t 代优化结束后活动 i 选取服务池中各服务的概率分布。显然，对于任意概率分布向量 P_i^t ，满足 $\sum_{k=1}^{m_i} P_{ik}^t = 1$ 。初始化时， $P_i^0 = (1/m_i, 1/m_i, \dots, 1/m_i)$ ， $\forall i \in \{0,1,2,\dots,n\}$ ，各活动服务选择满足均匀分布。

3.2 基于特征的初始解生成方法

对于给定具有 n 个活动的工作流，一共有 $\prod_{i=1}^n m_i$ 种调度方案，其中满足截止期约束的解为可行解，其余为不可行解。为提高初始群中的质量，构建基于问题特征的启发式方法生成满足截止期约束的相对较优解，其主要思想是：首先为所有活动选择执行时间最短的服务，则该服务分配方案一定满足截止期的约束，而且工作流租赁费用最高；然后通过局部调整部分可延长执行时间而不影响截止期的活动服务选择方案，以降低工作流的费用。假设活动 i 选择服务池中的第 k 个服务，定义其最早开始时间 $EST_i = \max_{v \in R_i} \{EFT_v\}$ ，最早完成时间

$$EFT_i = EST_i + d_i^k, \text{ 最晚结束时间 } LFT_i = \min_{v \in Q_i} \{LST_v\},$$

最晚开始时间 $LST_i = LFT_i - d_i^k$ 。由于各活动服务池中的可选服务按照执行时间递增、费用递减的顺序排序，则第 k 个服务和第 $k+1$ 个服务的执行时间满足 $d_i^{k+1} > d_i^k$ ，费用满足 $c_i^k > c_i^{k+1}$ ， $1 \leq k \leq m_i - 1$ 。

定义 1 若活动 i 选择第 k 个服务 (即 $x_{ik} = 1$)，且最晚结束时间和最早开始时间的差值满足 $LFT_i - EST_i \geq d_i^{k+1}$ ， $1 \leq k \leq m_i - 1$ ，称该活动为可延长执行活动。

定义 1 表明可延长执行活动 i 可以选择费用更便宜的第 $k+1$ 个服务而不影响工作流的完工时间。

定义 2 可延长活动集 Ψ 是所有可延长执行活

动组成的集合, $\Psi = \{i | 1 \leq i \leq n, x_{ik} = 1, LFT_i - EST_i \geq d_i^{k+1}, 1 \leq k \leq m_i - 1\}$ 。

定义 3 可延长执行活动的性价比指标 τ_i 为 $\tau_i = ((c_i^k - c_i^{k+1}) / (d_i^{k+1} - d_i^k)) / (|Q_i| + |R_i|)$ 。

从 Ψ 集中选择可重新分配服务方案的活动时, 充分考虑活动可选服务的性价比、工作流图的结构特点, 以最大程度降低工作流费用。定义 3 中 $(c_i^k - c_i^{k+1}) / (d_i^{k+1} - d_i^k)$ 体现了活动 i 选择第 k 个服务和第 $k+1$ 个服务的性价比。由于 $c_i^k > c_i^{k+1}$, 且 $d_i^{k+1} > d_i^k$, 满足 $(c_i^k - c_i^{k+1}) / (d_i^{k+1} - d_i^k) > 0$ 。如果 $(c_i^k - c_i^{k+1}) / (d_i^{k+1} - d_i^k)$ 值越大, 则活动 i 选择第 $k+1$ 个服务的性价比越高; $|Q_i| + |R_i|$ 为该活动直接后继和直接先驱的数量之和, 如果 $|Q_i| + |R_i|$ 值越小, 则为活动 i 选择第 $k+1$ 个服务对其他活动的开始时间和结束时间影响越小。

基于问题特征的启发式初始解生成算法 HISGM 生成具有 POP 个个体的初始群体, 其中个体 I_1 为启发式方法生成的满足截止期约束的较优服务分配方案。为增强群体多样性, 在该较优解 I_1 的基础上, 随机变换 I_1 中若干活动的服务模式生成其余 $POP-1$ 个初始个体。

算法 1 基于问题特征的启发式初始解生成算法(HISGM)

```

1)  $\Omega \leftarrow \phi, k \leftarrow 1, I_1 = (k, k, \dots, k), flag \leftarrow true$ 
2) Do while ( $flag = true$ )
3)  $\Psi \leftarrow \phi$  //置可延长执行活动集  $\Psi$  为空
4)  $EST_0 \leftarrow 0, EFT_0 \leftarrow EST_0, LFT_{n+1} \leftarrow \delta_{n+1}, LST_{n+1} \leftarrow LFT_{n+1}$ 
5) For  $i=0$  to  $n$ 
6)  $EST_i \leftarrow \max_{j \in R_i} \{EFT_j\}, EFT_i \leftarrow EST_i + d_i^k$ 
//计算各活动最早开始时间和最早完工时间
7) End for
8) For  $i=n$  to  $0$ 
9)  $LFT_i \leftarrow \min_{j \in Q_i} \{LST_j\}, LST_i \leftarrow LFT_i - d_i^k$ 
//计算各活动最晚开始时间和最晚完工时间
10) End for
11) For  $i=0$  to  $n$ 
12) If ( $k \leq m_i - 1$ ) and ( $LFT_i - EST_i \geq d_i^{k+1}$ )
//满足定义 2 的为可延长执行活动

```

```

13)  $\Psi \leftarrow \Psi \cup \{i\}$ 
14) End if
15) End for
16) If  $\Psi \neq \phi$ 
17) For each  $j \in \Psi$ 
18)  $\tau_j \leftarrow ((c_j^k - c_j^{k+1}) / (d_j^{k+1} - d_j^k)) / (|Q_j| + |R_j|)$ 
//计算  $\Psi$  中各活动的性价比指标  $\tau_i$ 
19) End for
20)  $a \leftarrow \arg \max_{j \in \Psi} \{\tau_j\}, I_{1a} \leftarrow k+1$ 
//为  $\tau_i$  值最大的活动更改服务方案为  $k+1$ 
21) Else
22)  $flag \leftarrow false$ 
23) End if
24) End do
25)  $\Omega \leftarrow \Omega \cup \{I_1\}$ 
26) For  $j=2$  to  $POP$ 
27) 随机变换  $I_1$  中若干活动服务方案生成  $I_j,$ 
 $\Omega \leftarrow \Omega \cup \{I_j\}$ 
28) End for
29) Return  $\Omega$ 

```

3.3 解的度量与择优

在评价个体适应度时, 不仅要考虑目标函数值, 还需考虑该服务选择方案下的工作流完工时间是否满足截止期的要求。根据当前个体 $I_j = (I_{j0}, I_{j1}, I_{j2}, \dots, I_{jn})$ 中各活动选择的服务, 采用 FSM 方法对工作流进行调度, 得到工作流的完工时间为 $CMAX(I_j)$, 利用式(1)计算目标函数值为 $COST(I_j)$, 定义个体 I_j 的适应度值 $f(I_j)$ 如下

$$f(I_j) = \begin{cases} COST(I_j), & CMAX(I_j) \leq \delta_{n+1} \\ COST(I_j) + (CMAX(I_j) - \delta_{n+1}) \cdot \theta, & CMAX(I_j) > \delta_{n+1} \end{cases} \quad (8)$$

其中, θ 为惩罚因子。实验中 θ 取较大的值, 以保证可行解个体的适应度值总是优于不可行解个体的适应度值。

混合分布估计算法中子代群体的生成依赖于问题各变量取值的概率分布向量, 而概率分布向量的更新与精英集密切相关, 故精英集的质量对算法性能有很大影响。精英集由当前群体的部分个体构成, 该解集可看作是群体的过去最优解。结合粒子群算法中个体同时向过去最优解和全局最优解学

习的思想,可在统计分析时同时考虑全局最优解中变量的取值情况。本文提出的精英选择方法(ES, elite selection)可形式化表示为: $[\beta\varepsilon + (1-\beta)\varepsilon] + 1$, $0 \leq \beta \leq 1$ 。从当前群体选出 $\beta\varepsilon + (1-\beta)\varepsilon$ 共 ε 个个体,与算法的全局最优解 G_{Best} (G_{Best} 表示算法到本次迭代结束找到的最好解)共同组成精英集。当 $\beta=1$ 时,从群体中选择适应度值最小的 ε 个个体与 G_{Best} 组成精英集;当 $\beta=0$ 时,从群体中随机选出 ε 个个体与 G_{Best} 组成精英集;当 $0 < \beta < 1$ 时,从群体中选择适应度值最小的 $\beta\varepsilon$ 个个体、随机选择 $(1-\beta)\varepsilon$ 个个体与 G_{Best} 组成精英集。

3.4 概率分布向量的更新

分布估计算法是一种基于概率模型的进化算法,概率模型的选取对算法的性能影响很大。常用的概率模型可分为变量无关、变量相关两类,UMDA 算法(univariate marginal distribution algorithm)是解决变量无关问题的常用方法^[14]。本文提出的方法中个体编码表示 workflow 各活动的服务选择方案,各维变量间可看作是与变量无关的,故可采用 UMDA 方法来更新各活动服务选择的概率分布向量 $P_i^t = (P_{i1}^t, P_{i2}^t, \dots, P_{im_i}^t)$ 。

基于精英集中的解,对活动 i 服务选择的概率分布向量 P_i^t 中的每一维 P_{ik}^t ,提出如下更新策略

$$P_{ik}^{t+1} = (1-\alpha)P_{ik}^t + \frac{\alpha}{\varepsilon+1} \sum_{j=1}^{\varepsilon+1} y_{ji}^k \quad (9)$$

其中, y_{ji}^k 是一个二元变量。如果精英集中的第 j 个个体为活动 i 选择第 k 个服务,则 $y_{ji}^k = 1$; 否则 $y_{ji}^k = 0$ 。 α 为学习因子, $0 < \alpha < 1$ 。

式(9)表明 P_{ik}^{t+1} 的值与群迭代优化过程中以往的经验值 P_{ik}^t 和精英集中为活动 i 选择第 k 个服务的数量有关。 α 值越大, P_{ik}^{t+1} 越依赖于精英集中的解;相反, α 值越小, P_{ik}^{t+1} 越依赖于已有的经验值 P_{ik}^t 。更新后的 P_{ik}^{t+1} 值越大,则在生成新个体时为活动 i 选择第 k 个服务的概率越大。

根据所有活动更新后的服务选择概率分布向量,采用轮盘赌的策略为各活动选择服务,生成新个体。算法2详细描述了针对本问题的概率分布向量的更新与子代群体的生成方法。

算法2 概率分布向量的更新与新群体的生成(PUPG)

1) 采用 ES 策略从当前群 Ω 中选择精英解

2) For $i=0$ to n

3) For $k=1$ to m_i

$$4) \quad P_{ik}^{t+1} \leftarrow (1-\alpha) \cdot P_{ik}^t + \frac{\alpha}{\varepsilon+1} \cdot \sum_{j=1}^{\varepsilon+1} y_{ji}^k$$

//更新各活动服务选择的概率分布向量

5) End for

6) End for

7) $\Omega \leftarrow \phi$ //置新群体为空

8) For $j=1$ to POP

9) For $i=0$ to n

10) $\mu \leftarrow \text{Rand}(0,1)$

//产生[0, 1]区间的随机数 μ

11) For $k=1$ to m_i

$$12) \quad \mu \leftarrow \mu - P_{ik}^{t+1}$$

13) If $\mu \leq 0$

$$14) \quad I_{ji} \leftarrow k$$

//为活动 i 选择第 k 个服务

15) End if

16) End for

17) End for

18) $\Omega \leftarrow \Omega \cup I_j$

//将新个体 I_j 添加到新群体 Ω 中

19) Endfor

20) Return Ω

3.5 混合分布估计算法

综上所述,本文提出混合分布估计算法求解云服务工作流调度问题,算法3详细描述了迭代优化过程。首先初始化各变量取值的概率分布向量,用 HISGM 算法生成初始群体 Ω ,重复执行以下步骤直到终止条件满足:根据个体的服务选择方案,采用 FSM 方法计算工作流的完工时间 $C_{\text{MAX}}(I_j)$,依据目标函数式(1)计算工作流的租赁费用 $C_{\text{Ost}}(I_j)$,按式(8)计算个体的适应度值 $f(I_j)$;根据群中所有个体的适应度值更新全局最优解 G_{Best} ;用 ES 策略选择精英集;根据精英集的解更新各变量取值的概率分布向量并生成子代群体。

算法3 混合分布估计算法(HEDA)

1) $t \leftarrow 0$, $P_i^t = (1/m_i, 1/m_i, \dots, 1/m_i)$ ($i=0,1,2,\dots,n$)

- 2) 调用启发式算法 HISGM 生成初始群体 Ω , $G_{\text{Best}} \leftarrow I_1$
- 3) Do while ($t < iter$)
- 4) For $j=1$ to POP
- 5) 用 FSM 方法计算个体 I_j 服务选择方案下的 workflow 完工时间 $C_{\text{MAX}}(I_j)$
- 6) 基于式 (1) 计算个体 I_j 服务选择方案下的 workflow 费用 $C_{\text{OST}}(I_j)$
- 7) 基于式 (8) 计算个体 I_j 的适应度值 $f(I_j)$
- 8) If ($f(I_j) < f(G_{\text{Best}})$) and ($C_{\text{MAX}}(I_j) \leq \delta_{n+1}$)
- 9) $G_{\text{Best}} \leftarrow I_j$
- 10) End if
- 11) End for
- 12) 调用 PUPG 算法更新概率分布向量并生成新群体 Ω
- 13) $t \leftarrow t+1$
- 14) End do
- 15) Return G_{Best}

4 实验结果与分析

为评估本文所提出的混合分布估计算法的性能,用 Java 对算法编程,在大量具有不同参数值的 DAG workflow 应用实例上进行模拟测试。模拟实验环境为 Windows 7, AMD E2-3000M 1.8GHz, RAM 为 2 GB 的 IBM PC 机。采用文献[18]的方法随机生成活动数 $|V| \in \{30,60,90,120,150,180,210,240,270,300\}$ 的 workflow 案例, workflow 拓扑结构的复杂度用 OS 值 (order strength) 来衡量,本文实验中的 OS 取值为 0.1。workflow 中任一活动 i 的服务提供者数量 m_i 的取值在区间 $[2,10]$ 内服从均匀分布;每个服务 S_i^k 执行活动 i 所需的准备时间 r_i^k 在区间 $[1,10]$ 中随机生成,执行时间 t_i^k 按照文献[19]的方法在时间区间 $[3,120]$ 内随机生成,则第 k 个服务总共所需时间 $d_i^k = r_i^k + t_i^k$,执行费用 c_i^k 根据 d_i^k 采用凸函数^[19]在费用区间 $[5,100]$ 中随机生成。实验中截止期 δ_{n+1} 以 workflow 下界完工时间为准,每个实例按 5% 递增取 10 个不同截止期,实验结果为所有实例的平均值。

4.1 算法参数设置

混合分布估计算法中的参数群体大小 POP 设

置为 $2 \times |V|$, 算法迭代次数 $iter = 30$ 。式(9)中变量取值概率分布向量更新的学习因子 α 是算法的一个重要参数。为了较为公平且快速地测试参数取值,实验中选择具有 120 个活动的中等规模 workflow 案例,学习因子 $\alpha \in \{0.1,0.2,0.3,0.4,0.5,0.6,0.7,0.8,0.9\}$,从当前群体中选择的精英数量 ϵ 取经验值 10,不同学习因子设置下算法随机运行 300 次。利用 SPSS 统计分析软件对学习因子取值的实验结果进行统计分析,结果如图 3 所示。

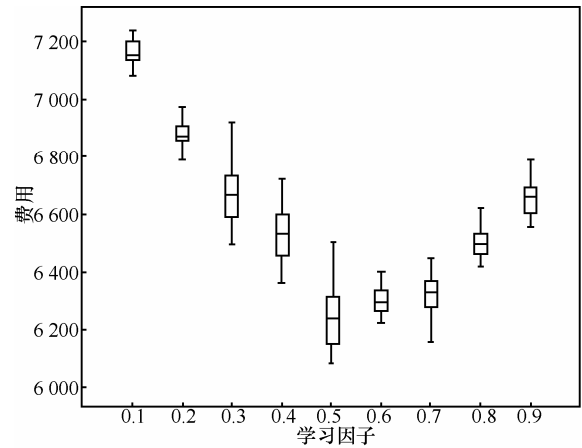
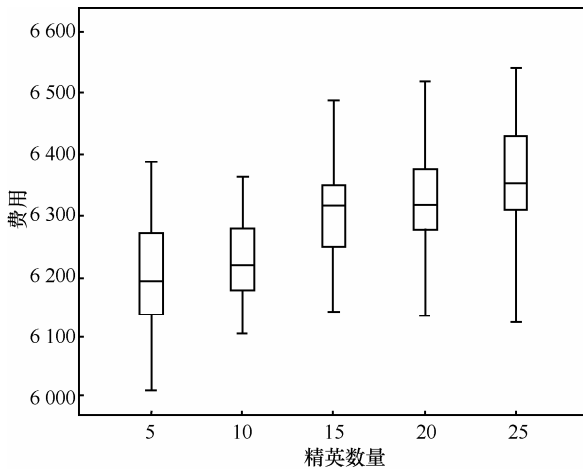


图3 学习因子 α 对算法性能的影响

式(9)的学习因子 $\alpha \in (0,1)$, α 值越小,活动 i 选择服务 S_i^k 的概率值 P_{ik}^{t+1} 越依赖于已有的经验值 P_{ik}^t ;相反, α 值越大, P_{ik}^{t+1} 越依赖于精英集中为活动 i 选择服务 S_i^k 的个体数量。图 3 的统计结果表明,当 $\alpha \leq 0.5$ 时,随着 α 的增大, workflow 的平均费用降低,算法性能有显著提高;而当 $\alpha > 0.5$ 时,随着 α 的增大, workflow 的平均费用增大。总体来看,当 α 取值为 0.5 时,获得的工作流费用均值最小。所以,在以下的实验中 α 取值为 0.5。

在保持学习因子 α 取最优值 0.5 的情况下,采用 SPSS 统计分析软件对精英数量 $\epsilon \in \{5,10,15,20,25\}$ 不同取值下算法性能进行统计分析,结果如图 4 所示。从均值来看, ϵ 不同取值下对算法平均性能差异不大。当 ϵ 取值为 10 时,实验数据越集中,算法稳定性越好性能更优。结果表明增大精英数量,并不能提高算法的性能。事实上,精英数量越大,算法更倾向于向最优解学习,使群体的局部搜索能力增强,而全局搜索能力变弱。在后续实验中,从当前群体中选择的最优解数量 ϵ 均设置为 10。

图4 精英数量 ϵ 对算法性能的影响

4.2 精英选择方法的比较

本文提出的精英选择方法中精英集包含 $\beta\epsilon$ 个从当前群体选出的最优个体, $(1-\beta)\epsilon$ 个随机个体和全局最优解 G_{Best} 。显然, 如果 β 取值越大, 精英集中最优个体的数量越多。本节对于不同的 β 取值进行实验比较, 从当前群体中选择的精英数量 $\epsilon=10$, workflow 规模 $|V| \in \{30, 60, 90, 120, 150, 180, 210, 240, 270, 300\}$, 实验结果如图5所示。

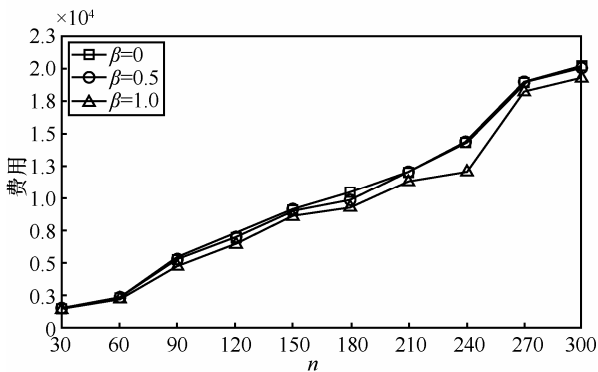


图5 不同精英选择方法的比较

从实验结果来看, 最优解在精英解中的比例对算法性能影响较大, β 取值越大, 即最优个体在精英集中数量越多, 算法性能越好。在所有的 workflow 规模上 $\beta=1.0$ 时算法性能最好。 $\beta=1.0$ 时 workflow 的平均费用比 $\beta=0$ 时降低了 8.69%, 而 $\beta=0.5$ 时 workflow 的平均费用比 $\beta=0$ 时降低了 3.02%。这样的现象可解释为: β 取值大则群体总是向最优解学习, 学习速度快, 易趋于局部最优解。 β 取值小则学习过程的随机性大, 算法收敛慢。

4.3 HEDA 算法与其他算法的比较

将本文提出的 HEDA 算法与 DTL^[5]算法、

DBL^[6]算法、改进后 PSO^[10]算法等在大量不同 DAG 结构的工作流应用实例上进行模拟测试。根据 4.1 节对学习因子 α 的实验分析结果, 以下实验中 HEDA 算法的学习因子 α 取值 0.5。为保证算法比较的公平性, HEDA 算法与 PSO 算法中群体大小 POP 均设置为 $2 \times |V|$, 算法迭代次数 $iter=30$ 。另外, 为分析本文提出的精英选择方法对 HEDA 算法的影响, 实验中比较 2 种不同统计学习方法: 从群体中选择适应度值最小的 $\epsilon+1$ 个个体组成精英集, 算法标记为 HEDA-1; 从群体中选择适应度值最小的 ϵ 个体, 与算法的全局最优解 C_{Best} 共同组成精英集, 算法标记为 HEDA-2。每个 workflow 案例设置 10 个不同的截止期, 不同截止期约束下算法随机运行 10 次, 取平均值作为算法得到的最终结果。

图6中给出了5个算法在不同规模的工作流上获得的平均费用。所有案例的运行结果显示, 随着 workflow 规模的扩大, workflow 的运行成本在增加。用平均偏移量 $ARPD = (C_A - C_{\text{Best}}) / C_{\text{Best}} \times 100\%$ 来衡量各算法与最好解的偏差, 其中 C_A 为算法 A 所获得平均费用, C_{Best} 为 5 个算法找到的最好解。在规模为 30 和 60 的工作流案例上, HEDA-2 与 HEDA-1 的平均便宜量为 2.34%, 在其余规模的工作流案例上, HEDA-2 总是能比其他算法找到更好的最优解。在所有的案例上, DTL、DBL、PSO 相对于最优解的平均偏移量分别为 51.76%、43.89%、17.62%, 即 HEDA-2 算法比 DTL 算法平均降低了 51.76% 的工作流费用, 比 DBL 算法降低了 43.89% 的费用, 比 PSO 算法平均降低了 17.62% 的工作流运行成本。同时, HEDA-1 算法相对于最优解 HEDA-2 的平均偏移量为 7.48%。结果表明统计学习过程中引入对全局最优解的学习可以提高算法的性能。

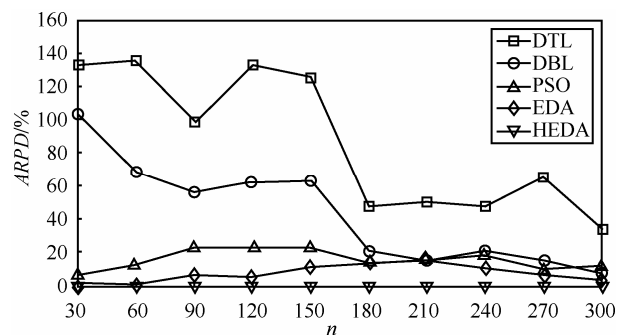


图6 不同算法性能比较

图 7 给出了具有 300 个节点的工作流案例在不同截止期约束下 5 个算法的实验结果。截止期 δ_{n+1} 与求解问题解空间大小与成正比。工作流任务越紧急, 也即截止期 δ_{n+1} 越小, 问题可行解的数量越少。工作流截止期 δ_{n+1} 越大, 则可行解空间越大。从实验结果来看, 随着工作流截止期 δ_{n+1} 逐渐变大, DTL、DBL、PSO、HEDA-1 和 HEDA-2 算法获得的工作流费用都在降低。

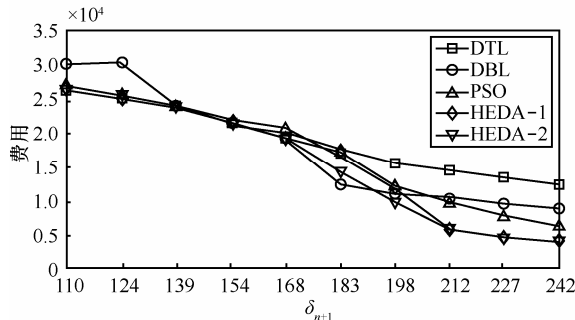


图 7 不同截止期对算法性能的影响

4.4 算法执行时间

对于群智能算法而言, 一般难以评价其时间复杂度, 常采用事后测量其计算时间来评价其时间开销。从表 1 中比较了 PSO、HEDA-1 和 HEDA-2 3 个算法的运行时间。随着问题规模的扩大, 算法所需时间在增加, 这是因为 PSO 算法和 EDA 算法都是基于群体进化的智能算法, 群体大小的设置与问题规模成正比。在工作流规模小于 210 的情况下, 3 种算法的执行时间均小于 5 s。所有案例运行结果显示, EDA 算法的执行时间略少于 HEDA 算法, PSO 算法的执行时间最短。

表 1 算法执行时间

问题规模	算法		
	PSO	HEDA-1	HEDA-2
30	0.02 s	0.10 s	0.11 s
60	0.18 s	0.32 s	0.36 s
90	0.45 s	0.73 s	0.82 s
120	0.75 s	1.39 s	1.55 s
150	1.71 s	2.28 s	2.56 s
180	2.76 s	3.57 s	3.87 s
210	4.06 s	5.18 s	5.81 s
240	5.86 s	7.32 s	7.88 s
270	8.17 s	10.02 s	10.80 s
300	11.21 s	14.67 s	15.44 s

5 结束语

云计算环境下许多大型应用被描述成具有 QoS 约束的工作流调度, 该问题是一类复杂且难以求解的 NP-hard 问题。本文提出混合分布估计算法优化截止期约束具有准备时间的云服务 workflow 费用, 该问题是云计算环境下 workflow 调度的一个新问题。基于问题特征构建启发式方法生成初始群体, 提高初始解的质量。结合粒子群算法的思想, 提出新颖的精英集选择方法, 引入有效的概率分布向量更新方式。与已有算法在大量 workflow 实例上的实验结果表明尽管本文所提出算法比已有算法需要略多的计算时间, 但大大提高了求解质量。

在实际动态的工作流调度系统中, 由于网络和服务本身的不稳定性, 经常会在调度中出现扰动 (服务失效、费用变化), 如何用最快时间、最经济费用前提下完成对工作流优化调度的实现是下一个阶段值得研究和解决的问题。

参考文献:

- [1] 林闯, 苏文博, 孟坤等. 云计算安全: 架构、机制与模型评价[J]. 计算机学报, 2013, 36(9): 1765-1784.
LIN C, SU WB, MENG K, et al. Cloud computing security: architecture, mechanism and modeling[J]. Chinese Journal of Computers, 2013, 36(9): 1765-1784.
- [2] CHARD K, BUBENDORFER K. High performance resource allocation strategies for computational economies[J]. IEEE Transactions on Parallel and Distributed Systems, 2013, 24(1): 72-84.
- [3] DE P, DUNNE E, GHOSH J, et al. Complexity of the discrete time-cost tradeoff problem for project networks[J]. Operations Research, 1997, 45(2): 302-306.
- [4] DE P, DUNNE E, GHOSH J, et al. The discrete time-cost tradeoff problem revisited[J]. European Journal of Operational Research, 1995, 81(2): 225-238.
- [5] YU J, BUYAYA R, THAM C. Cost-based scheduling of workflow applications on utility grids[A]. Proceedings of 1st International Conference on e-Science and Grid Computing[C]. 2005. 140-147.
- [6] YUAN Y C, LI X P, WANG Q, et al. Deadline division-based heuristic for cost optimization in workflow scheduling[J]. Information Sciences, 2009, 179(15): 2562-2575.
- [7] 苑迎春, 李小平, 王茜. 基于逆向分层的网格 workflow 调度算法[J]. 计算机学报, 2008, 31(2): 282-290.
YUAN YC, LI XP, WANG Q. Bottom level based heuristic for workflow scheduling in grids[J]. Chinese Journal of Computers, 2008,

- 31(2):282-290.
- [8] 苑迎春, 李小平, 王茜. 基于串归约的网格 workflow 费用优化方法[J]. 计算机研究与发展, 2008, 45(2):246-253.
YUAN YC, LIXP, WANG Q. Cost optimization heuristic for grid workflows scheduling based on serial reduction[J]. Chinese Journal of Computer Research and Development, 2008, 45(2):246-253.
- [9] ABRISHAMI S, NAGHIBZADEH M, EPEMA D. Cost-driven scheduling of grid workflows using partial critical paths[J]. IEEE Transactions on Parallel and Distributed Systems, 2012, 23(8): 1400-1414.
- [10] 张晓东, 李小平, 王茜等. 服务工作流的混合粒子群调度算法[J]. 通信学报, 2008, 29(8): 87-93.
ZHANG XD, LI XP, WANG Q, *et al.* Hybrid particle swarm optimization algorithm for cost minimization in service-workflows with due dates[J]. Journal on Communications, 2008, 29(8): 87-93.
- [11] CHEN W N, ZHANG J. An ant colony optimization approach to a grid workflow scheduling problem with various QOS requirements[J]. IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews, 2009, 39(1): 29-43.
- [12] YU J, BUYYA R. Scheduling scientific workflow applications with deadline and budget constraints using genetic algorithms[J]. Scientific Programming, 2006, 14(3): 217-230.
- [13] CAI Z C, LI X P, GUPTA J N D. Critical path-based iterative heuristic for workflow scheduling in utility and cloud computing[A]. Proceedings of 11th International Conference on Service-Oriented Computing[C]. 2013.207-221.
- [14] ABRISHAMI S, NAGHIBZADEH M, EPEMA D. Deadline constrained workflow scheduling algorithms for IAAS clouds[J]. Future Generation Computer Systems, 2013, 29(1): 158-169.
- [15] VAN DEN BOSSCHE R, VANMECHELEN K, BROECKHOVE J. Cost optimal scheduling in hybrid IAAS clouds for deadline constrained workloads[A]. Proceedings of IEEE 3rd International Conference on Cloud Computing[C]. 2010. 228-235.
- [16] LARRAANAGA P, LOZANO J A. Estimation of Distribution Algorithms: a New Tool for Evolutionary Computation[M]. Boston: Kluwer Press, 2002.
- [17] WANG L, FANG C. An effective estimation of distribution algorithm for the multi-mode resource-constrained project scheduling problem[J]. Computers & Operations Research, 2012, 39(2): 449-460.
- [18] DEMEULEMEESTER E, VANHOUCKE M, HERROELEN W. Rangen: A random network generator for activity-on-the-node networks[J]. Journal of Scheduling, 2003, 6(1): 17-38.
- [19] AKKAN C, DREXL A, KIMMS A. Network decomposition based benchmark results for the discrete time-cost tradeoff problem[J]. European Journal of Operational Research, 2005, 165(2): 339-358.

作者简介:



沈虹 (1981-), 女, 江苏启东人, 东南大学博士生, 南京审计学院讲师, 主要研究方向为项目调度、服务计算。



李小平 (1970-), 男, 重庆人, 博士, 东南大学教授、博士生导师, 主要研究方向为机器调度、项目调度、服务计算。