

# 面向组合服务收益优化的动态服务选择方法

王显志 徐晓飞 王忠杰

(哈尔滨工业大学计算机科学与技术学院 哈尔滨 150001)

**摘 要** 服务组合是面向服务计算环境中实现价值增值的有效途径, 现有研究多关注于顾客价值的提高, 而对提供服务组合方案的服务提供者或代理商的价值缺乏充分考虑. 一方面, 在服务级别协议下, 超过约定级别的质量不会给提供者带来额外收益, 也不能有效提升顾客价值, 因此是没有必要的; 另一方面, 服务本身及其交付环境的不确定性导致服务质量也具有一定的不确定性, 使得实际服务可能无法达到约定的服务级别, 甚至完全失效. 对此基于服务级别协议对服务收益和策略进行分析, 给出面向收益优化的服务选择模型, 并将阶段估计与即时反馈相结合, 对需求进行贪婪调度, 最终基于模拟退火算法实现组合服务的优化选择. 实验结果表明, 与传统方法相比, 该方法不仅能够获取更多的服务收益, 而且在不同需求分布下均具有更高的效率.

**关键词** 服务组合; 收益优化; 服务级别协议; 贪婪调度; 模拟退火

**中图法分类号** TP311 **DOI 号:** 10.3724/SP.J.1016.2010.02104

## A Profit Optimization Oriented Service Selection Method for Dynamic Service Composition

WANG Xian-Zhi XU Xiao-Fei WANG Zhong-Jie

(School of Computer Science and Technology, Harbin Institute of Technology, Harbin 150001)

**Abstract** Service composition is an effective means of building value-added service in service-oriented computing environment. Current research focuses on the fulfillment of customer value, while neglects the value procurable by service broker, which is the compositor of individual services as well as the provider of composite services. On the one hand, over-optimized service quality will not bring additional profit to the service provider as well as no remarkable improvement to customer satisfaction, thus is unnecessary for the value of both sides of service participants in SLA environment; on the other hand, due to the uncertainty of both services and the environment for delivering services, real quality of service-oriented applications exhibits as uncertain, too. So real services may not meet the quality requirement of negotiated service level, or even fail. Profit and service strategies are studied for SLA, and a novel service selection model is proposed for profit optimization. Based on periodical estimation of service cost and instant feedbacks, service requirements are greedily scheduled and optimized service selection is realized for dynamic service composition based on simulated annealing algorithm. Experimental results show that this approach does not only promote the profit of composite services, but also have superior efficiency in procuring optimized results under different circumstances of requirements distribution, compared with traditional approaches.

**Keywords** service composition; profit optimization; service level agreement; greedy scheduling; simulated annealing

收稿日期: 2010-06-08; 最终修改稿收到日期: 2010-08-28. 本课题得到国家自然科学基金(61033005, 60803091, 70971029)、国家“八六三”高技术研究发展计划项目基金(2009AA04Z151)资助. 王显志, 男, 1984 年生, 博士研究生, 主要研究兴趣为双边服务资源整合和服务组合. E-mail: wangxianzhi@hit.edu.cn. 徐晓飞, 男, 1962 年生, 博士, 教授, 博士生导师, 主要研究领域包括服务科学与工程、企业计算、服务计算等. 王忠杰, 男, 1978 年生, 博士, 副教授, 主要研究兴趣为服务工程、面向服务的体系结构、服务组合和软件工程.

## 1 引言

服务组合是面向服务计算环境中将现有服务进行组合以实现服务价值增值的有效途径. 服务提供者将自身业务功能包装为服务并对外发布; 服务代理<sup>[1]</sup> (service broker) 根据具体业务需求从多个提供者处选取适当的服务并将组合后的整体提供给顾客. 其中从服务业务功能单元到具体服务的绑定通常是在组合服务执行时动态进行的, 称为动态服务选择.

在正常运行的服务系统中, 客户和服务提供商的利益必须是健康的, 即双方都能在服务中实现必要的价值, 才能维持服务体系的正常运行. 已有文献主要从服务质量<sup>[2-5]</sup>、信任<sup>[6-7]</sup>等方面研究如何为顾客提供更好的组合服务. 然而组合服务的“价值增值”是双边的, 它不仅体现为顾客所获得的增值服务, 还体现为服务代理在提供组合服务中的收益增值, 因此有必要探讨相应的策略和方法来帮助服务代理获取更多的收益.

在现实服务中, 由于服务执行、交付环境的不确定性, 实际服务的质量往往达不到预期要求, 甚至服务失败, 使得服务代理的收益受到影响. 服务级别协议 (SLA) 作为服务双方关于服务内容、质量等方面的约定, 是保障服务质量和顾客满意度的有效途径<sup>[8]</sup>. 没有 SLA 时, 用户支付相同费用所获得的服务质量却可能大不相同; 由于缺乏质量承诺和规范的罚金机制, 顾客与服务提供者之间难以建立稳定的服务关系. 而在 SLA 情况下, 组合服务变成了有保证的服务, 即便服务质量达不到约定级别时, 也有规范的赔偿.

虽然 SLA 为服务提供者提供了一种竞争依据, 但更多时候, 用户对于高质量并无特别追求, 服务只要达到一定的标准即可实现顾客价值<sup>[9]</sup>. 图 1 所示的边际效用递减规律表明, 超出标准的质量对于顾客满意度的提高并无显著影响; 而对于服务代理

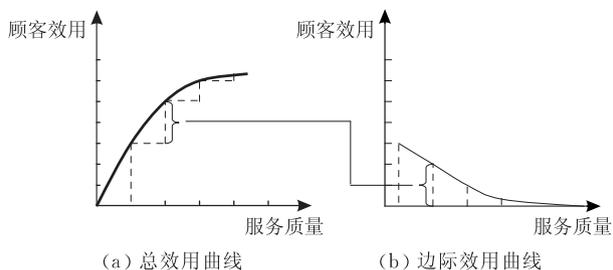


图 1 总效用曲线和边际效用曲线

而言, 却消耗大量的资源, 缩小了赢得更多利润的空间.

以海运物流服务中的服务组合为例: 货代为满足货主的需求, 组合了订舱位、预订集装箱、取重箱、返空箱、集港、报关、装船等一系列服务, 其服务价值主要体现在提供整箱出口服务中获取的金钱. 因此在整箱出口服务中, 只考虑如何为货主提供更加优质的服务是不符合现实的. 因此在服务组合中, 一方面在服务级别协议的支持下保障服务顾客一方的价值, 另一方面优化组合服务提供者的收益的策略是保证服务能够健康持续进行的有效途径.

本文针对服务级别协议下服务质量的实际特点, 研究服务组合中的收益优化方法. 相比于其它工作, 本文研究具备如下特点: (1) 相比于单纯追求顾客满意度, 本文从服务代理的角度出发, 探讨其作为服务组合者和组合服务提供者的服务策略与收益方法; (2) 基于服务级别协议, 提出确定及不确定质量条件下的服务选择模型; (3) 面向组合服务的收益优化, 提出高效的动态服务选择方法.

## 2 动态服务选择的定义

文献<sup>[1]</sup>给出了 broker 的体系结构, 这也是当前服务组合所采用的通用架构. 该架构中, 服务代理执行服务的选择与组合作为组合服务的提供者向顾客提供组合后的服务, 以此赚取收益. 基于 SLA 的服务组合中, 顾客请求根据服务级别的不同形成多个队列, 如图 2 所示. 在组合服务执行时, 通过服务流程的每个具体功能单元 (即服务任务) 与具体服务进行绑定, 最终形成满足顾客需求的可执行的组合服务实例.

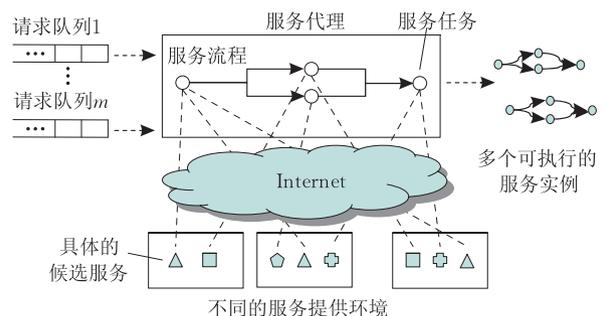


图 2 网络环境下基于 SLA 的服务组合

对于包含有  $n$  个任务  $Tasks = \{task_1, task_2, \dots, task_n\}$  的服务流程, 每个  $task_i (i=1, 2, \dots, n)$  对应一个候选服务集合  $CS_i = \{ws_{i,1}, ws_{i,2}, \dots, ws_{i,n_i}\}$ , 每个具体服务定义  $ws_{i,j} = (Qset_{i,j}, price_{i,j}) (i=1, 2, \dots,$

$n; j=1, 2, \dots, n_i$ ), 其中  $Qset = \{q_1, q_2, \dots, q_k\}$  是质量指标集,  $price$  为服务价格. 任一组合服务需求用  $r$  表示, 服务选择方案  $Plan = \{\langle task_1, ws_1 \rangle, \langle task_2, ws_2 \rangle, \dots, \langle task_n, ws_n \rangle\}$ , 其中  $ws_i \in CS_i, Plan \subseteq \bigcup_{i=1}^n (\langle task_i \rangle \times CS_i)$ . 组合服务的不同质量指标  $q_l (Plan) = g_l(q_l(ws_1), q_l(ws_2), \dots, q_l(ws_n)) (l=1, 2, \dots, k)$ , 其中  $q_l(ws_i)$  为  $ws_i$  的  $q_l$  指标值,  $g_l$  为计算组合服务  $q_l$  指标值的与服务流程有关的函数.

由于成员服务自身及其所处的运行环境、网络环境等不确定因素,  $q_j(ws_i)$  通常是一个不确定值, 导致组合服务的质量也具有一定的不确定性, 因此在实际提供的服务很可能无法达到预期的质量, 甚至失效. 然而在更多的情况下, 某些服务质量的不如意并不会影响用户对于服务的使用, 即提供的质量相对劣一些的服务对顾客依然有意义, 则极端情况下: 为赚取更多收益, broker 总是优先满足那些更能获利的需求, 而故意降低对其它需求的服务质量.

### 3 基于 SLA 的服务级别分析

现有的服务质量交付机制缺乏 QoS、SLA 及定价三者之间的联系<sup>[10]</sup>, 而现实中服务的价格、罚金以及实际质量等因素都会影响服务提供者的收益. 以海运物流服务中的服务组合为例: 若货主的订舱委托是由 Web 服务完成的. 该服务存在两个服务级别: 级别 A 规定的价格和响应时间为 (100 元, 0.5s), 级别 B 为 (50 元, 2s). 此时有如下可能:

(1) 若服务的实际响应时间小于 0.5s; 顾客支付订舱费用 100 元;

(2) 由于服务策略或负载原因导致顾客服务的实际响应时间为 1.5s; 顾客获得级别 A 规定金额的罚金, 并支付网上订舱服务费用 50 元;

(3) 服务响应时间很长, 实际质量达不到任何服务级别; 服务提供者向顾客支付级别 A 规定的罚金.

因此充分考虑服务级别协议下服务收益的特点, 建立适合于收益优化的策略、模型与方法成为保障组合服务收益的有效途径. 本文将 SLA 具体化为一系列的服务级别, 每个级别规定服务质量、价格等方面的参数<sup>[11]</sup>, 并给出其收益规则 (包括定价与罚金规则), 以支持基于 SLA 的服务组合.

#### 3.1 多服务级别及其组织

在实际应用中, broker 通常定义一系列的服务

级别<sup>[12]</sup>, 形如:  $SLASet = \{SLA_1, SLA_2, \dots, SLA_m\}$ . 每个服务级别  $SLA_i = (Qset_i, price_i, penalty_i) (i=1, 2, \dots, m)$ , 其中各项含义如下.

(1)  $Qset_i = (q_1^i, q_2^i, \dots, q_k^i)$  为服务级别相应的质量标准的集合, 其中  $q_j^i (j=1, 2, \dots, k)$  是单个质量指标值. 不同质量指标具有不同意义, 如可靠性应不低于  $Qset$  中规定的指标值; 而响应时间则应小于等于规定值.

(2)  $price$  为服务价格;  $penalty$  表示服务失效或者未达到约定质量时的罚金.

当前服务组合研究<sup>[13-14]</sup>多侧重于表达不同服务级别对于服务质量与价格的权衡, 而一合理的服务级别还应该能够表达对于不同质量指标的侧重, 因此不同服务级别间的质量优劣关系可以表示为图 3 所示的有向无环图, 称为多服务级别的质量关系图, 记为  $G$ .

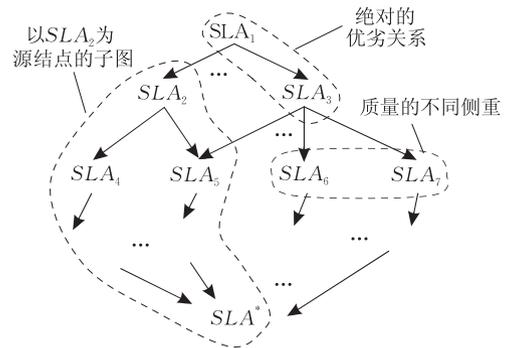


图 3 多服务级别的质量关系图

图中箭头表示源服务级别的质量标准 ( $Qset$ ) 优于箭头指向的服务级别. 若用“ $>$ ”表示不同级别质量间的“优于”关系, 则图 3 中  $SLA_1 > SLA_3$ ; 同层结点 (如  $SLA_6$  与  $SLA_7$ ) 的质量标准不存在绝对的优劣关系, 即必然一个服务级别在某些质量指标上优于后者, 而后者又在另一些指标上优于前者. 可见, 该有向无环图在其结点集合上对应一个严格的偏序关系, 形成偏序集  $(\{SLA_1, SLA_2, \dots, SLA_m\}, >)$ .

尽管服务资源不足时 broker 可以策略性地拒绝部分顾客需求, 但由于服务本身、环境等不确定因素, 实际服务可能只能达到较低的级别, 甚至失效. 以下分别用服务级别变迁图  $STG$  和失效服务级别  $SLA^*$  表示服务级别变迁和失效两种情况:

(1) 设组合服务需求  $r$  的约定服务级别为  $SLA_x$ , 某具体服务选择方案  $Plan$  实际能够达到  $SLA_y$  的质量标准, 超出的质量对于服务收益并无影响, 因此将  $Plan$  实际达到的服务级别定义为

$$SLA(Plan) = \begin{cases} SLA_x, & SLA_y > SLA_x \\ SLA_y, & SLA_x \geq SLA_y \end{cases}$$

设  $G^2$  为图  $G$  的所有子图的集合, 则  $G_x(V_x, E_x)$  ( $\in G^2$ ) 表示  $G$  中以  $SLA_x$  为源的子图(图 3 中所示为以  $SLA_2$  为源子图)。从  $SLA_x$  到  $G_x$  中任一结点都存在一条有向路径, 即  $\forall SLA' \in V_x, \exists path(\in G^2): SLA_x \rightarrow \dots \rightarrow SLA'$ , 也即  $SLA_x \geq SLA'$ , 因此  $G_x$  描述了约定服务级别为  $SLA_x$  时  $SLA(Plan)$  的取值范围, 称为  $SLA(Plan)$  的服务级别变迁图, 记为  $STG$ 。

(2) 若  $\exists SLA^*$ , 使得  $\forall SLA_i (i=1, 2, \dots, m)$ , 满足  $SLA_i > SLA^*$ , 则称  $SLA^*$  为失效服务级别。它是一个虚拟的服务级别, 仅用于表征组合服务未成功执行的情况(质量标准为 0), 没有任何用户会请求该级别, 因此服务级别集合可以进一步写为  $SLASet = \{SLA_1, SLA_2, \dots, SLA_m\} \cup \{SLA^*\}$ 。

为保证定义的服务级别集合能够完整表达实际服务质量的所有可能, 要求对于任一服务选择方案  $Plan$ , 有且仅有唯一的一个服务级别与之对应。体现在  $G$  图中, 即要求图中的两层子结构满足:

(1) 分支结构。达到父级别的方案都必然能够满足子级别的质量标准:  $U_0 = U_1 \cap U_2 \cap \dots \cap U_t (t > 0)$ ; 任意一个能够同时达到两个子级别质量标准的方案必然属于父级别:  $B_i = U_i - U_0, B_1 \cap B_2 \cap \dots \cap B_t = \emptyset$ 。

(2) 汇聚结构。达到任一父级别的方案必然能够满足子级别的质量标准:  $U_1 \cup U_2 \cup \dots \cup U_t \subset U_0 (t > 0)$ ; 达到任一你级别的方案必然不能达到其它的父级别:  $B_0 = U_0 - (U_1 \cup U_2 \cup \dots \cup U_t), B_1 \cap B_2 \cap \dots \cap B_t = \emptyset$ 。

其中,  $B_i = \{Plan \mid SLA(Plan) = SLA_i\}$  表示实际级别为  $SLA_i$  的服务选择方案的集合,  $U_i = \{Plan \mid SLA(Plan) \geq SLA_i\}$  表示实际质量达到级别  $SLA_i$  的质量标准的服务选择方案集合。显然对于任意两个服务级别  $SLA_x$  与  $SLA_y$ , 若  $SLA_x > SLA_y$ , 则  $U_x \subset U_y$ 。

### 3.2 基于 SLA 的服务定价与罚金

由于服务选择的不同, 具有相同服务级别的组合服务可能具有不同的服务成本。因此服务级别的定价直接关系着组合服务的收益, 不同的定价规则将导致不同的服务策略和收益优化方法。在航运物流服务中, 服务级别的罚金完全取决于获取该级别服务所需支付的价格, 即采用固定价格的罚金策略。本文在此基础上对该策略进行了扩展, 提出 3 种可能的服务策略, 使得本文方法能够应用到更加广泛的场景。本文考虑基于 SLA 的 3 种定价规则:

#### (1) 固定价格

每种服务级别具有固定的价格, 即对于任何需求, 不管获取的实际质量如何, 只要达到约定级别, 就支付该级别的费用。显然, 服务级别越高, 价格也越高:  $\forall SLA_x, SLA_y (x \neq y)$ , 若  $SLA_x > SLA_y$ , 则  $price_x > price_y$ 。

#### (2) 固定收益率

所有服务级别的收益率相同, 即服务价格只与具体服务组合成本有关, 与服务级别无关。设  $Plan = \{\langle task_1, ws_1 \rangle, \langle task_2, ws_2 \rangle, \dots, \langle task_n, ws_n \rangle\}$  为任一服务选择方案  $Plan$ , 其收益定义为服务价格与成本之差:  $profit(Plan) = price(Plan) - cost(Plan)$ 。收益率  $\alpha = profit(Plan) / cost(Plan)$ , 则  $price(Plan) = (1 + \alpha) \cdot cost(Plan)$ , 且  $cost(Plan) = g(price(ws_1), price(ws_2), \dots, price(ws_n))$ ,  $g$  为由成员服务价格计算组合服务总价格的与服务流程有关的函数。

#### (3) 可变收益率

在商品市场中, 商品价格越高, 其利润通常也越高。借鉴到服务场景, 质量越高的服务级别, 其收益率也越高。设  $SLA_x, \alpha(Plan_x)$  分别为服务选择方案  $Plan_x$  的约定服务级别和收益率,  $SLA_y (y \neq x), \alpha(Plan_y)$  分别为  $Plan_y$  的约定服务级别和收益率, 若  $SLA_x > SLA_y$ , 则  $\alpha(Plan_x) > \alpha(Plan_y)$ 。

从服务提供者的角度出发, 大多数情况下不会无偿提供服务, 因此服务价格与组合服务的实际质量、服务级别的变迁及定价策略有关; 从顾客角度, 则要求有保障的服务价值实现, 因而需要相应的规则来确定未达到服务级别时的罚金, 以促使服务提供者确保服务质量, 努力满足顾客需求。无论罚金如何设置, 都可归结为 3 种可能:

#### (1) 罚金大于收益

组合服务未达到服务级别规定的质量时, 提供者所需支付的罚金不小于获取的收益, 即  $\forall SLA_x, SLA_y (x \neq y)$ , 若  $SLA_x > SLA_y$ , 则  $penalty_x \geq profit(SLA_y)$ 。

#### (2) 罚金小于收益

组合服务未达到服务级别规定的质量时, 提供者所需支付的罚金小于获取的收益, 即  $\forall SLA_x, SLA_y (x \neq y)$ , 若  $SLA_x > SLA_y$ , 则  $penalty_x < profit(SLA_y)$ 。

#### (3) 混合情况

针对每次具体的服务选择, 根据服务成本的不同, 组合服务未达到约定质量标准时的罚金可能大于、等于或者小于获取的服务收益。

## 4 基于 SLA 的组合服务收益优化

现实中任何一个持续运行的服务系统所执行的服务组合都是多需求的服务组合. 在海运订舱服务中, 顾客的需求是连续/断续到来的, 而在较短时间内, 可能有多个服务级别的多个需求同时需要得到满足. 在此情况下, 服务选择的目标是使 broker 在较长时期内获取尽可能多的服务收益. 在有限的服务资源下, 组合服务的收益优化问题转化为基于 SLA 的以收益最大化为目标的组合服务动态选择.

### 4.1 基于 SLA 的服务选择模型

设某时期可被划分为一系列的不相关的时间区间:  $T = \{T_1, T_2, \dots\}$ , 相应于每个区间的需求同样形成序列:  $R = \{R_1, R_2, \dots\}$ . 对于每个时间区间  $T_i$ , 针对其中每个需求进行优化服务选择而获取的服务收益

$\sum_{r \in R_i} profit'(r)$  显然大于未优化的收益  $\sum_{r \in R_i} profit(r)$ , 即其差值为

$$\sum_{r \in R_i} \Delta profit(r) = \sum_{r \in R_i} profit'(r) - \sum_{r \in R_i} profit(r) \geq 0, \quad i=1, 2, \dots$$

对于两个相邻时间段  $T_i$  和  $T_{i+1}$ , 令  $\Delta T_i = T_i \cap T_{i+1}$ , 则  $\Delta T_i$  内所有需求的集合  $\Delta R_i = R_i \cap R_{i+1}$ , 于是有

$$\begin{aligned} \sum_{r \in R} \Delta profit(r) &= \sum_{r \in R} (profit'(r) - profit(r)) \\ &= \sum_{r \in R} profit'(r) - \sum_{r \in R} profit(r) \\ &= \sum_{R_i \in R} \left( \left( \sum_{r \in T_i} profit'(r) - \sum_{r \in \Delta R_i} profit'(r) \right) - \left( \sum_{r \in R_i} profit(r) - \sum_{r \in \Delta R_i} profit(r) \right) \right) \\ &= \sum_{R_i \in R} \left( \left( \sum_{r \in R_i} profit'(r) - \sum_{r \in R_i} profit(r) \right) - \left( \sum_{r \in \Delta R_i} profit'(r) - \sum_{r \in \Delta R_i} profit(r) \right) \right) \\ &= \sum_{R_i \in R} \left( \sum_{r \in R_i} \Delta profit(r) - \sum_{r \in \Delta R_i} \Delta profit(r) \right). \end{aligned}$$

显然  $\Delta R_i \subseteq R_i$ , 其差值  $R_i^- = \{r | r \in R_i \setminus \Delta R_i\}$ , 则  $\sum_{r \in R} \Delta profit(r) = \sum_{R_i^- \in R} \sum_{r \in R_i^-} \Delta profit(r) \geq 0$ . 因此对于在每个时间区间的收益优化能够保证在很长以及无限长期间内持续优化组合服务的收益. 以下将在一个有限时间段内讨论.

设一定时间段内的组合服务需求形成序列  $R = \{r_1, r_2, \dots, r_c\}$ , 相应的约定服务级别同样形成序列:  $SLASeq = \{SLA_1, SLA_2, \dots, SLA_c\}$ , 这些需求的

服务选择方案序列:  $PlanSeq = \{Plan_1, Plan_2, \dots, Plan_c\}$ , 每个方案实际达到的服务级别:  $SLASeq' = \{SLA'_1, SLA'_2, \dots, SLA'_c\}$ . 设  $PlanSet$  为  $PlanSeq$  的服务方案集, 它由 3 种方案构成:  $PlanSet = PlanSet_s \cup PlanSet_d \cup PlanSet_f$ . 其中  $PlanSet_s = \{Plan_i | Plan_i \in PlanSet \wedge SLA_i = SLA'_i\}$ ,  $PlanSet_d = \{Plan_i | Plan_i \in PlanSet \wedge SLA_i > SLA'_i\}$  分别表示达到和未达到约定级别的服务选择方案集合,  $PlanSet_f = \{Plan_i | Plan_i \in PlanSet \wedge Plan_i = Plan^\emptyset\} = Tasks \times \emptyset$  是仅包含一个空方案  $Plan^\emptyset$  的集合, 用于表征 broker 拒绝服务需求的情况. 由 3.1 节中  $SLA(Plan)$  的定义可知, 收益计算中任一服务选择方案的实际服务级别必然低于或者等于约定的服务级别, 即  $SLA'_i \leq SLA_i$ , 且  $SLA'_i \in V_i$ , 其中  $G_i(V_i, E_i)$  是图  $G$  中以  $SLA_i$  为源的子图.

对于序列  $R$  的任一需求  $r_i$  的服务选择方案  $Plan_i = \{\langle task_1, ws_1^i \rangle, \langle task_2, ws_2^i \rangle, \dots, \langle task_n, ws_n^i \rangle\}$ , 有如下 3 种可能:

(1)  $Plan_i \in PlanSet_s$ , 此时 broker 的服务收益:  $profit_s(Plan_i) = price_i - g(price(ws_1^i), price(ws_2^i), \dots, price(ws_n^i))$ , 其中  $price_i$  为约定服务级别  $SLA_i$  的价格,  $g$  为由成员服务价格计算组合服务价格的与服务流程有关的函数.

(2)  $Plan_i \in PlanSet_d$ , broker 在服务中的收益:  $profit_d(Plan_i, SLA'_i) = price'_i - g(price(ws_1^i), price(ws_2^i), \dots, price(ws_n^i)) - penalty_i$ ,

其中  $price'_i$  为  $Plan_i$  实际达到的服务级别  $SLA'_i$  的价格,  $penalty_i$  是约定服务级别  $SLA_i$  中规定的罚金.

(3)  $Plan_i \in PlanSet_f$ , 此时 broker 在服务中的收益为

$$profit_f(Plan_i) = profit(Plan^\emptyset) = 0.$$

综上, 得到具体服务选择方案  $Plan_i$  的收益:  $profit(Plan_i) =$

$$\begin{cases} profit_s(Plan_i), & SLA_i = SLA'_i \\ profit_d(Plan_i, SLA'_i), & SLA_i > SLA'_i \\ 0, & r_i \text{ 被拒绝} \end{cases}$$

考虑服务质量的不确定性. 虽然 broker 可以决定是否接受服务需求, 然而一旦接受, 组合服务质量的不确定性将导致 broker 对具体服务选择方案的收益估计也具有一定的不确定性.

设  $ws_l^j (i=1, 2, \dots, c, j=1, 2, \dots, n)$  为  $Plan_i$  在第  $j$  个服务任务上的具体服务, 其第  $l (l=1, 2, \dots, k)$

个质量指标值为随机变量  $q_{i,j}^l$ , 则组合服务的第  $l$  个指标  $q_l(Plan_i) = g_l(q_{i,1}^l, q_{i,2}^l, \dots, q_{i,n}^l)$ , 进而  $Plan_i$  的实际质量  $\{q_1(Plan_i), q_2(Plan_i), \dots, q_k(Plan_i)\}$ . 其中  $g_l$  为计算组合服务  $q_l$  指标值的与服务流程有关的函数.

设  $SLA_i$  的质量参数集合  $QSet_i = (q_i^1, q_i^2, \dots, q_i^k)$ ,  $\varphi_{i,l}$  为  $q_{i,1}^l, q_{i,2}^l, \dots, q_{i,n}^l$  的联合概率密度函数, 且所有质量指标均为正向参数<sup>[3]</sup> (负向参数可以转为正向参数), 则 broker 在方案  $Plan_i$  的期望收益为  $E(\text{profit}(Plan_i)) =$

$$\begin{aligned} & p(SLA_i = SLA'_i) \cdot \text{profit}_s(Plan_i) + \\ & p(SLA_i > SLA'_i) \cdot \text{profit}_d(Plan_i, SLA'_i) \\ = & p(SLA_i = SLA'_i) \cdot \text{profit}_s(Plan_i) + \\ & \sum_{SLA_w \in V_i} p(SLA_w = SLA'_i) \cdot \text{profit}_d(Plan_i, SLA_w), \end{aligned}$$

其中,  $SLA_i$  是约定的服务级别,  $SLA'_i$  是  $Plan_i$  实际达到的服务级别,  $V_i$  是图  $G$  中以  $SLA_i$  为源的子图  $G_i(V_i, E_i)$  的结点集合,  $w = 1, 2, \dots, |V_i|$ . 参照图 4 (b),  $Plan_i$  的实际服务级别  $SLA'_i$  等于任一既定级别  $SLA_x$  的概率遵从容斥原理, 即

$$\begin{aligned} & p(SLA'_i = SLA_x) = \\ & p(SLA'_i \geq SLA_x) - \sum_{w=1,2,\dots,t} p(SLA'_i \geq SLA_w) + \\ & \sum_{e \neq y; e, y=1,2,\dots,t} p(SLA'_i \geq SLA_w) \cdot p(SLA'_i \geq SLA_e) - \dots + \\ & (-1)^{t+1} \cdot \prod_{w=1}^t p(SLA'_i \geq SLA_w) \quad (\text{共 } 2^t - 1 \text{ 项}), \end{aligned}$$

其中  $Plan_i$  达到任一级别  $SLA_x$  的概率为

$$\begin{aligned} & p(SLA'_i \geq SLA_x) = \prod_{l=1}^k p(q_l(Plan_i) \geq q_x^l) \\ & = \prod_{l=1}^k \int \int \dots \int_{q_l(Plan_i) \geq q_x^l} \varphi_{i,l} \cdot d(q_{i,1}^l) d(q_{i,2}^l) \dots d(q_{i,n}^l). \end{aligned}$$



图 4 多服务级别质量关系图的两种子结构

因此在服务质量具有不确定性的情况下, broker 服务选择方案  $Plan_i$  的期望收益为

$$\text{profit}^E(Plan_i) = \begin{cases} E(\text{profit}(Plan_i)), & r_i \text{ 被接受} \\ 0, & r_i \text{ 被拒绝} \end{cases}$$

设组合服务流程的每个服务任务  $task_j \in Tasks$  对应的候选服务集合为  $CS_j = \{ws_{j,1}, ws_{j,2}, \dots, ws_{j,n_j}\}$ , 其中每个服务的可用并发数分别为  $\{N_{j,1}, N_{j,2}, \dots, N_{j,n_j}\}$ ; 时间段  $T_x$  内的所有需求  $R =$

$\{r_1, r_2, \dots, r_c\}$ , 相应的服务选择方案为  $PlanSeq = \{Plan_1, Plan_2, \dots, Plan_c\}$ , 这些方案在  $T_x$  内对于  $CS_j$  中每个服务的调用次数总和分别为  $\{N'_{j,1}, N'_{j,2}, \dots, N'_{j,n_j}\}$  ( $task_{i,j}$  为  $Plan_i$  的第  $j$  个任务,  $CN_i = \{\langle task_{i,j}, \omega_s \rangle \mid \langle task_{i,j}, \omega_s \rangle \in Plan_i \wedge \omega_s = \omega_{s_j,v}\}$ ,  $N'_{j,v} = \sum_{Plan_i \in PlanSeq} |CN_i|$ ), 则在时间段  $T_x$  内追求 broker 的最大收益, 即在有限服务能力下的优化组合服务的选择为

$$\begin{aligned} & \max \sum_{i=1}^c \text{Revenue}(Plan_i) \\ \text{s. t. } & N'_{j,v} \leq N_{j,v} \quad (j=1, 2, \dots, n \wedge v=1, 2, \dots, n_j), \end{aligned}$$

其中  $\text{Revenue}(Plan_i)$  为  $Plan_i$  的估计收益. 质量确定时,  $\text{Revenue}(Plan_i) = \text{profit}(Plan_i)$ ; 否则  $\text{Revenue}(Plan_i) = \text{profit}^E(Plan_i)$ . 该优化问题的

搜索空间为  $c \cdot \prod_{j=1}^n |CS_j|$ , 至少属于 NP 难.

## 4.2 面向组合服务收益优化的服务选择方法

为获取更多的服务收益, broker 在不同的收益规则下的服务策略也有所不同, 如表 1 如示.

表 1 不同收益规则下 broker 的服务策略

方面	规则名称	broker 的服务策略
固定 收益率	固定 收益率	$\text{profit} = \alpha \cdot \sum_{j=1}^c \text{cost}(Plan_j)$ 兼顾服务尽可能多的需求, 提高达到约定服务级别的成本
		$\text{profit} = \sum_{j=1}^c \alpha(SLA_j) \cdot \text{cost}(Plan_j)$ 兼顾服务尽可能多的需求, 服务于收益率高的服务级别, 提高达到约定服务级别的成本
固定价格	固定价格	$\text{profit} = \sum_{j=1}^c (\text{price}_j - \text{cost}(Plan_j))$ 兼顾服务尽可能多的需求, 降低达到约定服务级别的成本
		罚金 $\geq$ 收益 达不到约定 SLA 时则拒绝服务
失效 罚金	失效 罚金	罚金 $<$ 收益 达不到约定 SLA 时也尽可能地服务于更多需求
		混合情况 根据具体情况决定是否接受请求

现有文献基于组合模型和图模型研究服务组合<sup>[15]</sup>, 并利用整数规则<sup>[1]</sup>、分支定界<sup>[9]</sup>等精确算法和遗传算法等启发式算法<sup>[3-4]</sup>进行组合服务的优化选择, 此外还有服务社区<sup>[16]</sup>等基于历史数据的方法. 然而在有限的服务资源下, 一定时间区间内的服务组合具有以下特点:

(1) 可能使这个需求得到满意的服务级别, 则另外一个就无法满足, 也可能两个需求都可以在较低服务级别下得到满足, 因此不同需求之间可能存

在某种优化的权衡。

(2) 为追求更多的服务收益, broker 可能需要放弃或者搁置某些需求而优先服务其它需求, 即不同需求之间可能存在某种优化的服务次序。

针对以上特点, 两种可能的优化策略是: (1) 区间内所有需求一次处理, 统一权衡, 最终达到总体收益最大化的均衡; (2) 优化需求的执行顺序, 每次贪婪获取当前可预见的最大收益, 逐步趋近全局最优收益。

本文采用第 2 种策略, 结合阶段性成本估算与执行结果的即时反馈来动态调度需求, 并基于模拟退火算法优化组合服务选择, 该方法包括“成本估算”和“服务选择”两部分。

### (1) 成本估算算法

随机获取不同级别需求的可行解, 并估计当前所处的较长时期内达到不同服务级别的组合服务成本和收益率。算法结果将作为下一步服务选择算法中需求调度的初始参考值。

算法基于模拟退火实现, 主要步骤如下。

1. 模拟一组需求  $R^- = \{r_1, r_2, \dots, r_k\}$ , 使得每种服务级别  $SLA_i (i=1, 2, \dots, k)$  都恰有一个需求  $r_i$  与之相应; 定义评价函数:

$$f(X) = \begin{cases} 1, & SLA = SLA' \\ 0, & SLA > SLA' \end{cases}$$

其中  $X$  为任一服务选择方案,  $SLA$  为约定服务级别,  $SLA'$  为  $X$  实际达到的服务级别。

2. 针对每个  $r_i \in R^-$ , 执行步 3~10。

3. 设置初始温度  $T_0 > 0$ , 当前迭代次数  $K=0$ , 最大迭代次数为  $M$ , 初始解个数  $N = |R^-| = k$ , 可行解集  $S = \emptyset$ , 退火参数  $\lambda < 1$ , 选择概率参数  $\beta > 0$ 。

4. 随机生成初始解集  $X^0 = (X_1^0, X_2^0, \dots, X_N^0)$ , 计算相应评价  $F^0 = \{f(X_1^0), f(X_2^0), \dots, f(X_N^0)\}$ ; 令  $S^0 = \{X_i^0 | X_i^0 \in F^0 \wedge f(X_i^0) > 0\}$ ,  $S = S \cup S^0$ 。

5. 产生一组随机向量  $Z^K = (Z_1^K, Z_2^K, \dots, Z_N^K)$ , 其中  $Z_l^K = (z_{l,1}^K, z_{l,2}^K, \dots, z_{l,n}^K) (l=1, 2, \dots, N)$ ,  $n$  为流程的服务任务数,  $z_{l,j}^K = \text{sign}(U_j - 5/6) \cdot \text{ceiling}(T_K \cdot (1/|U_j| - 1)) (j=1, 2, \dots, n)$ ,  $U_j$  是  $(2/3, 1)$  上的随机数,  $\text{sign}$  是符号函数,  $\text{ceiling}$  为上取整函数。

6. 计算  $y_{l,j}^K = (x_{l,j}^K + z_{l,j}^K) \% n_j (j=1, 2, \dots, n)$ , 得到  $Y_l^K = (y_{l,1}^K, y_{l,2}^K, \dots, y_{l,n}^K)$  及  $Y^K = (Y_1^K, Y_2^K, \dots, Y_N^K)$ , 计算每个  $Y_l^K (l=1, 2, \dots, N)$  的评价值  $F^K = \{f(Y_1^K), f(Y_2^K), \dots, f(Y_N^K)\}$ , 其中  $n_j = |CS_j|$ , 是服务流程的第  $j$  个任务的候选服务数目。令  $S^K = \{Y_l^K | Y_l^K \in F^K \wedge f(Y_l^K) > 0\}$ ,  $S = S \cup S^K$ 。

7. 针对每个  $Y_l^K \in S^K$ , 执行步 8, 9。

8. 如果  $f(Y_l^K) \leq 0$ , 则  $Y_l^K = X_l^K$ 。

9. 计算温度  $T_K$  下接受试探点  $Y_l^K$  的概率  $P_a(Y_l^K | X_l^K)$ ,

$T_K) = \min\{1, \exp((f(X_l^K) - f(Y_l^K))/(\beta T_K))\}$ 。如果在  $(0, 1)$  区间上的随机数  $\eta \leq P_a(Y_l^K | X_l^K, T_K)$ , 则  $X_l^{K+1} = Y_l^K$ ,  $f(X_l^{K+1}) = f(Y_l^K)$ ; 否则置  $X_l^{K+1} = X_l^K$ ,  $f(X_l^{K+1}) = f(X_l^K)$ 。

10. 如果  $K < M$ , 更新温度:  $T_{K+1} = \lambda \cdot T_K$ , 置  $K = K + 1$ , 转入步 5; 否则算法终止, 计算级别  $SLA_i$  的成本  $cost_i = \sum_{X_v \in S} cost(X_v) / |S|$  (其中  $cost(X_v)$  通过函数  $g$  计算得到, 参见第 2 节) 和  $SLA_i$  在这段时期的预期收益率  $\alpha_i = (price_i - cost_i) / cost_i$ 。

### (2) 服务选择算法

将需求按服务级别分组, 并以估算的收益率作为初始参数, 每次贪婪选择收益率最高的需求执行服务选择。每次算法执行都根据结果对收益率进行更新, 进而不断调整执行次序。同组的部分甚至全部需求可能共享同一服务选择方案, 也可能被整体调整到队列的其它位置或者一起被拒绝, 因此同一级别的需求可能被分解为更小的分组以执行相同的操作。图 5 是算法处理不同阶段需求队列变化情况的示例。

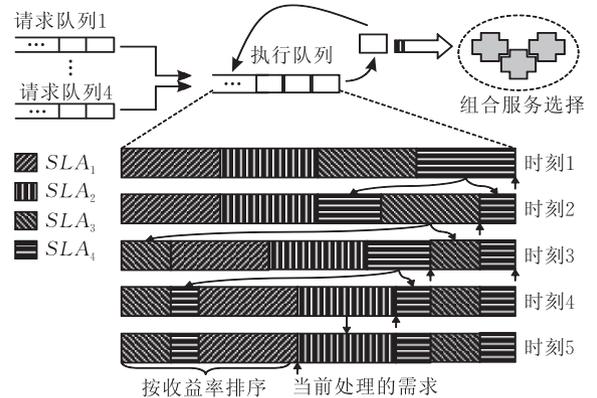


图 5 服务组合需求调度示例

具体步骤如下:

1. 按初始收益率高低对需求排序, 得到序列  $R = \{r_1, r_2, \dots, r_c\}$ , 约定服务级别  $\{SLA_1, SLA_2, \dots, SLA_c\}$ , 收益率  $\{\alpha_1, \alpha_2, \dots, \alpha_c\}$ , 有  $\alpha_1 \geq \alpha_2 \geq \dots \geq \alpha_c$ 。

2. 令  $token(r)$  表示需求  $r$  被标记为已处理。如果仍有未处理的需求, 即  $\{r | r \in R \wedge \neg token(r)\} \neq \emptyset$ , 转入下一步; 否则转入步 5。

3. 对于  $R$  中第一个未处理需求  $r_i (i > 1)$ , 若  $SLA_i \neq SLA_{i-1}$ , 转入下一步; 否则考察  $r_{i-1}$  的执行情况:

3.1. 若  $r_{i-1}$  未被处理, 即  $\neg token(r_{i-1})$ , 则将  $r_i$  直接迁移到需求序列中  $r_{i-1}$  的后面, 转入步 2;

3.2. 若  $r_{i-1}$  的服务选择方案  $Plan_{i-1}$  中所有服务的剩余并发数均大于零, 则  $Plan_i = Plan_{i-1}$ ,  $token(r_i)$ , 转入步 2; 否则转入下一步。

3.3. 若  $r_{i-1}$  实际达到服务级别  $SLA(Plan_{i-1}) = SLA^*$ , 则  $Plan_i = Plan_{i-1}$ ,  $token(r_i)$ , 转入步 2;

4. 采用与成本估算算法步 3~10 类似的过程针对  $r_i$  进行服务选择. 不同之处在于:

- 4.1. 使用新的评价函数  $f'(X) = Revenue(X)$ ;  $N=1$ ;
- 4.2. 取消可行解集  $S$  的定义, 另定义最优解  $X_{opt}$  及其评价  $f'(X_{opt})$ , 初始化:  $X_{opt} = X^0$ ,  $f'(X_{opt}) = f'(X^0)$ ; 更新方法: 每次迭代中如果  $f'(X^K) > f'(X_{opt})$  则  $X_{opt} = X^K$ ,  $f'(X_{opt}) = f'(X^K)$ ;
- 4.3. 步 8: 判断条件  $f'(Y_i^K) - f'(X_i^K) \leq 0$ ;
- 4.4. 步 10: 当迭代终止时:

如果实际方案达到约定服务级别, 即  $SLA(X_{opt}) = SLA_i$ , 则  $X_{opt}$  为  $r_i$  的近似最优服务选择方案,  $f'(X_{opt})$  为相应评价价值,  $Token(r_i)$ , 转入步 2;

否则, 重新计算服务级别  $SLA_i$  的收益率  $\alpha_i = (price_{opt} - cost_{opt} - penalty_i) / cost_{opt}$ . 若  $\alpha_i$  仍为最高收益率, 则  $X_{opt}$  为  $r_i$  的近似最优服务选择方案,  $f'(X_{opt})$  为相应评价价值,  $Token(r_i)$ , 转入步 2; 否则调整  $r_i$  在需求序列中的位置, 使所有未处理需求对应的收益率仍为非递减序列, 转入步 2.

5. 接受服务选择方案的评价价值大于零的需求, 拒绝其它需求, 算法终止.

当服务资源的数量及质量未发生重大变化时, 成本估算算法结果在较长时期内有效, 因此不计入算法复杂性; 服务选择算法的最优复杂性为  $O(m \cdot (n \cdot N \cdot M))$ , 其中,  $m$  为服务级别数,  $n$  为服务选择方案的编码长度,  $N$  为初始方案数,  $M$  为迭代次数.

## 5 实验结果及对比分析

目前服务组合研究缺乏公共的实验数据和一致的实验平台, 本文采用 Java2、Tuscany 的 XML 封装技术、axis2、ActiveBPEL 中间件实现服务的组合与选择. 所有算法均在同一平台下实现, 因此实验能够相对客观地对比不同算法的性能.

由于海运物流服务本身的复杂性限制, 本文利用已被广泛采用的旅游规划服务流程进行实验. 该服务流程如图 6 所示.

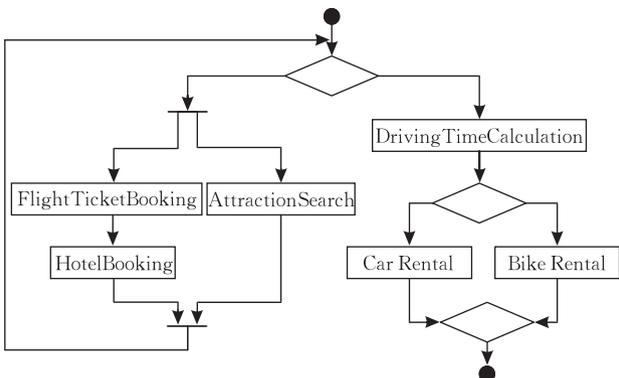


图 6 旅游规划服务流程<sup>[14]</sup>

设流程中每个任务有 8 个候选服务, 其质量指标值均为随机生成, 其中  $Reliability \in (0.9, 0.999)$ ,  $Response\_time \in (1.3, 1.5)$ , 价格  $price \in (1, 3)$ , 并发数  $concurrent\_number = 5$ ; 服务级别描述  $SLA = ((Reliability, Response\_time), price, penalty)$ , 具体级别:  $SLA_1 = ((0.75, 8), 100, 50)$ ,  $SLA_2 = ((0.7, 8.3), 80, 40)$ ,  $SLA_3 = ((0.65, 8.5), 60, 30)$ ,  $SLA_4 = ((0.6, 8.8), 40, 25)$ ,  $SLA_5 = ((0.55, 9), 20, 20)$ . 参数  $k=5, T_0=10, M=200, N=50, m=0.95, \lambda < 0.97, \beta=1$ . 由成本估算算法得到  $SLA_1 \sim SLA_5$  的实现成本分别为 10.98、11.07、10.16、10.97 和 10.17.

对半年间海运物流中整箱出口服务的统计发现, 大部分顾客订单倾向于选择中等质量的服务级别, 仅有少量顾客会选择使用服务质量较差、和服务质量很高的服务. 该结果符合我们对于一般服务级别定义的通用经验, 因此采用正态分布作为实现中需求分布的假设. 同时作为对比, 我们还引入均匀分布的服务需求.

为了便于实验比较, 本文分别基于模拟退火算法实现了传统的服务选择方法和基于 4.2 节得出的两种优化策略的服务选择方法, 并定义基于权衡的算法每两次执行的时间间隔为一个组合周期.

### (1) 传统方法

将现有服务组合方法直接应用到收益优化的场景, 针对每个需求分别进行服务选择.

### (2) 基于权衡的方法

每次选择统一处理多个需求, 通过权衡不同需求的选择方案来不断趋近于整体的最大收益.

### (3) 基于调度的方法

以成本估算为基础, 根据每次算法执行结果即时调度需求, 通过贪婪的需求调度和以组为单位的服务选择, 追求近似的最大收益. 这也是本文采用的方法.

## 5.1 需求规模对算法的影响

分 3 次实验, 每次在单个组合周期内分别发起 20、40 和 60 个服务级别呈均匀分布的随机需求, 持续 10 个周期, 如表 2 所示.

表 2 不同规模的服务需求在各服务级别上的分布

规模	级别				
	SLA <sub>1</sub>	SLA <sub>2</sub>	SLA <sub>3</sub>	SLA <sub>4</sub>	SLA <sub>5</sub>
20	4	4	4	4	4
40	8	8	8	8	8
60	12	12	12	12	12

图 7、图 8 分别给出了不同需求规模对于服务选择效率和服务收益的影响。

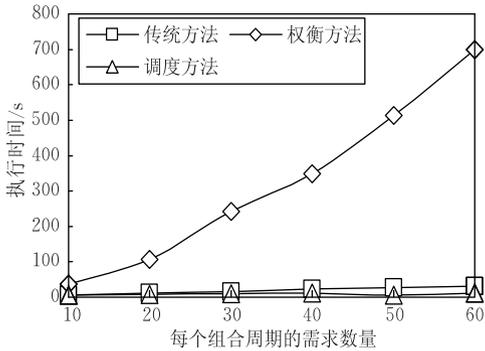


图 7 需求规模变化时不同方法的效率对比

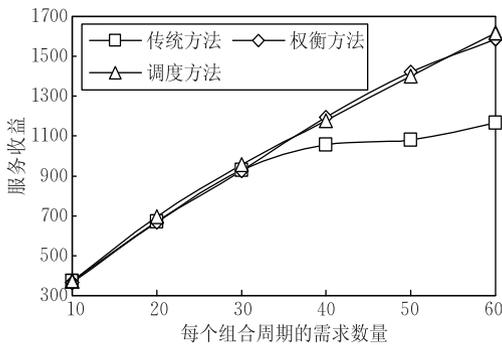


图 8 不同需求规模时不同方法得到的服务收益

图 7 所示为 3 种方法执行时间随单位时间内需求数量增加的变化情况。传统方法是针对每个需求分别执行一次服务组合；基于调度的方法实际上将需求按服务级别的不同分为一些需求分组，并以组为单位进行组合，根据每次组合结果的不同，已有分组可能被划分成为更小的组，可能被整体迁移等待执行，也可能整体从执行任务里取消，而且多个需求可能共享同一服务选择方案。需求间的均质性使得算法效率更多依赖于服务级别的多少，从而有效避免了需求规模对于算法效率的影响；基于权衡的方法将一定时间段内的需求一次进行组合，当需求数目为  $c$ ，方案编码长度为  $n$ ，初始方案数为  $N$ ，迭代次数为  $M$  时，前两种方法的时间复杂性均为  $O(c \cdot (n \cdot N \cdot M)) = O(1 \cdot ((c \cdot n) \cdot N \cdot M))$ ，而基于调度的方法在理想情况下（不需要调整需求次序）的最优时间复杂度为  $O(m \cdot (n \cdot N \cdot M))$ ，其中  $m$  为需求中包含服务级别的数目（实验中  $m=5, c=10, 20, 30, 40, 50, 60$ ），远远小于其它方法；而基于权衡的方法空间复杂度是另两种方法的  $(c \cdot n \cdot N) / (n \cdot N) = c$  倍，实际效率很低。

如图 8 所示，传统方法每次服务选择时都竭力最大化当前服务的收益，因此不能持续优化服务组

合收益，随着服务需求的增多，尤其当需求数量超过服务能力（需求数量  $\geq 40$ ）时，服务收益已基本不再增加；基于权衡的方法在每次组合时对时间段内需求统一进行服务选择，因而能够对不同需求的满足进行取舍，最终取得较好的优化效果；基于调度的方法依照收益率高低对服务需求进行贪婪的动态调度，有很大机率获得最优解，图 8 所示的实验结果表明其收益优化效果与基于权衡的方法几乎没有差别，即使当需求规模达到服务能力 1.5 倍（60/40）时，仍有进一步优化的空间。

## 5.2 需求在服务级别上的分布对算法的影响

分两次实验。每次在单个服务周期内发起 40 个服务需求，一次实验的需求呈均匀分布，另一次需求呈正态分布，各持续 10 个周期，如表 3 所示。

表 3 40 个服务需求的两种服务级别分布

分布	级别				
	SLA <sub>1</sub>	SLA <sub>2</sub>	SLA <sub>3</sub>	SLA <sub>4</sub>	SLA <sub>5</sub>
均匀分布	8	8	8	8	8
正态分布	4	10	12	10	4

图 9 和图 10 分别给出了 3 种算法在服务需求非均匀分布时的效率、效果相比于需求均匀分布时的变化情况。设  $T_e$ 、 $T_u$  分别为算法在需求均匀、非均匀分布时的执行时间，则执行时间变化率  $\phi_t = (T_u - T_e) / T_e$ ；同理，服务收益变化率  $\phi_{profit} = (Profit_u - Profit_e) / Profit_e$ 。

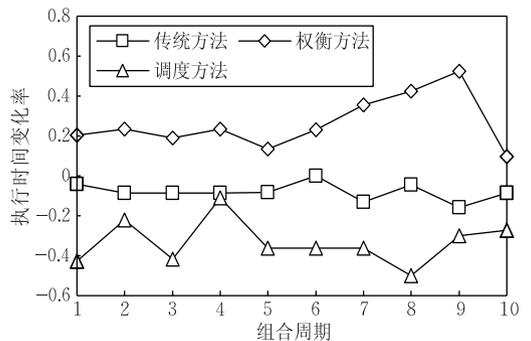


图 9 需求的非均匀分布对不同方法效率的影响

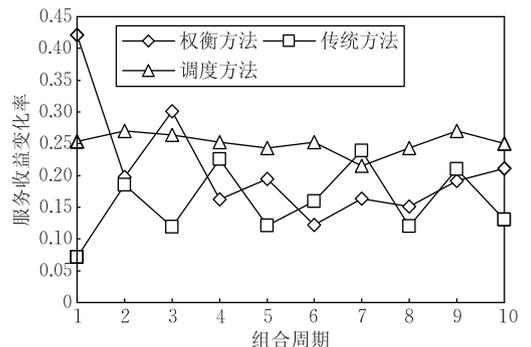


图 10 需求的非均匀分布对不同方法所得收益的影响

需求在服务级别上的非均匀分布使得需求对于资源的竞争不是分散在各服务级别之间,而是集中于少量的服务级别上.由图 9 可见,需求的集中对于传统的服务组合方法几乎无影响,而基于权衡的方法力图一次找到多个需求的服务选择方案,导致每个选择方案需要更多的冲突消解,因此其执行时间增加了逾 20%;基于调度的方法效率更依赖于服务级别的数目,而需求的集中导致某些服务级别上的需求数量很少,反而起到了类似于服务级别数目变少的效果,图 9 结果显示,在非均匀需求分布情况下,基于调度的算法效率提高近 40%.

由图 10 可见,相比于需求均匀分布的情况,非均匀需求下每种方法得到的服务收益均有所增加,这说明该部分收益增加是由需求本身特点决定的,而与具体方法无关.然而传统方法和基于权衡的方法获得的收益增量较少,平均仅为均匀需求时的 15%;而基于调度的方法获得 25%左右的稳定的收益增加值.通过与图 8 结果比较可以发现,相比于其它方法,本文提出的基于调度的服务选择方法在需求呈非均匀分布时,更易于获得相对优化的服务选择结果.

## 6 相关工作比较

当前研究多依据服务质量来进行服务组合.文献[4]提出关系矩阵编码方式以支持 QoS 感知的服务组合,解决了组合服务选择中的多路径问题;文献[15]对具有 QoS 约束的服务选择的组合模型和图模型进行比较并总结了相关的启发式算法;鉴于 SLA 在端到端服务质量保障中的有效性,文献[8]提出一个能够将业务目标、上下文和约束等策略信息逐层映射的中间件框架,以支持面向服务架构中的 SLA 协商;文献[1]将服务需求按服务级别分组,并针对每个服务级别分别利用整数规划进行服务选择;文献[13]提出一种服务覆盖网络(SON)中基于 SLA 提供简单服务时的组合服务 QoS 保障框架,实现了服务路径的快速生成和恢复;此外也有基于信任的服务组合.文献[6]利用 D-S 理论进行服务工作流的信任计算,通过基于信任的 QoS 调度,有效保证了工作流的 QoS 性能和调度的成功率;文献[7]将组合服务的可信性保证问题转换为自适应控制问题,并通过马尔可夫决策过程建模,实现了基于强化学习的组合服务自适应维护.

此外,由于服务本身负载、服务提供环境等不确

定因素的影响,组合服务质量呈现出一定的动态性和不确定性.文献[3]提出对服务的实际质量进行监测并对其质量指标进行级别标定,以支持组合服务的评价;文献[5]针对 Web 服务的随机性和 Internet 环境的动态性,提出随机 QoS 指标的度量方法和自适应 QoS 管理框架,并利用马尔可夫决策过程有效提高了服务组合成功率;文献[9]提出根据顾客获得的实际质量制定服务价格,并通过优化服务质量的评价,保证服务提供商努力达到约定的服务质量.

然而当前的服务组合多从顾客角度出发,而对组合服务提供者的价值缺乏考虑.更多时候,高服务质量在给顾客带来的少量价值增值的同时,却给组合服务提供者带来惨重的代价,因此有必要从服务代理的角度探讨组合服务的收益优化.文献[9]借鉴商品市场中的竞争机制,通过成员服务的按需定价,保证了服务选择中的负载均衡;文献[17]提出了组合服务收益在成员服务间的分配方法.然而这些研究仅将服务收益作为实现其它优化目标的辅助手段,而未用其指导服务组合过程.

本文站在组合服务提供者的立场探讨服务组合问题,一方面给出服务级别协议下的收益规则及相应服务策略,提出收益优化的服务选择模型;另一方面充分考虑基于 SLA 的服务组合的特点,提出基于贪婪调度的服务选择方法,该方法将阶段性信息统计与即时反馈相结合,高效优化服务组合的收益.

## 7 结 论

双边客户价值的实现是保证服务持续、良性运营的关键.本文在传统顾客价值(体现为顾客满意度等)的基础上,更多地从组合服务的提供者(即服务代理)的角度出发,探讨如何从服务组合中获取更多的收益.服务级别协议有效保障了单次服务中顾客价值的实现,本文在对服务级别及其收益方式进行分析的基础上,提出相应的服务策略和服务选择模型.为了高效求解收益优化问题,进一步提出动态需求调度与模拟退火相结合的组合服务选择方法.理论分析和实验表明,本文算法的复杂性仅与需求中包含的服务级别数目有关,而与需求的规模无关,且与传统方法相比,能够获取更多收益.下一步将在海运物流服务平台中加以验证.

服务提供者与顾客的利益常是相悖的,各自遵从一定的收益规则;同时它们的利益又是相通的,服务的顺利进行是双方价值的共同需要.本文利用服

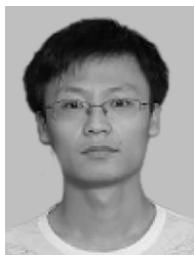
务级别协议保障顾客价值的同时,更加强调服务提供者的收益优化.而服务组合作为调节双方利益的手段,探讨一种长期上看对服务提供者有利、从每次服务组合来看对顾客更加有利的服务组合方式也许是一个值得尝试的方向.另外,本文对于服务组合的评价侧重于质量方面,并且假设服务提供者可以拒绝任何顾客需求,因此未能很好地解释拒绝服务以及服务达不到约定级别等情况对于顾客价值的影响.事实上,这些因素可能影响顾客满意度、服务提供者的信誉、以及需求数量及其在不同服务级别上的分布.如何建立相应的模型以刻画不同策略对于服务组合的影响也是值得深入探讨的问题.

### 参 考 文 献

- [1] Cardellini V, Casalicchio E, Grassi V et al. Presti. Flow-based service selection for web service composition supporting multiple QoS classes//Proceedings of the IEEE International Conference on Web Services (ICWS'07). Salt Lake City, UT, 2007; 743-750
- [2] Guan Ying, Ghose A K, Lu Zheng. Using constraint hierarchies to support QoS-guided service composition//Proceedings of the IEEE International Conference on Web Services (ICWS'06). Chicago, IL, 2006; 743-752
- [3] Jiang Zhe-Yuan, Han Jiang-Hong, Wang Zhao. An optimization model for dynamic QoS-aware Web services selection and composition. Chinese Journal of Computers, 2009, 32(5): 1014-1025(in Chinese)  
(蒋哲远, 韩江洪, 王钊. 动态的 QoS 感知 Web 服务选择和组合优化模型. 计算机学报, 2009, 32(5): 1014-1025)
- [4] Zhang Wen-Cheng, Su Sen, Chen Jun-Liang. Genetic algorithm on Web services selection supporting QoS. Chinese Journal of Computers, 2006, 29(7): 1029-1037(in Chinese)  
(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)
- [5] Fan Xiao-Qin, Jiang Chang-Jun, Wang Jun-Li, Pang Shan-Chen. Random-QoS-aware reliable Web service composition. Journal of Software, 2009, 20(3): 546-556(in Chinese)  
(范小芹, 蒋昌俊, 王俊丽, 庞善臣. 随机 QoS 感知的可靠 Web 服务组合. 软件学报, 2009, 20(3): 546-556)
- [6] Hu Chun-Hua, Wu Min, Liu Guo-Ping. QoS scheduling based on trust relationship in Web service workflow. Chinese Journal of Computers, 2009, 32(1): 42-53(in Chinese)

(胡春华, 吴敏, 刘国平. Web 服务工作流中基于信任关系的 QoS 调度. 计算机学报, 2009, 32(1): 42-53)

- [7] Guo Hui-Peng, Huai Jin-Peng, Deng Ting, Li Yang. A dependable and adaptive approach to supporting Web service composition. Chinese Journal of Computers, 2008, 31(8): 1434-1444(in Chinese)  
(郭慧鹏, 怀进鹏, 邓婷, 李扬. 一种可信的自适应服务组合机制. 计算机学报, 2008, 31(8): 1434-1444)
- [8] Zulkernine F, Martin P, Craddock C et al. A policy-based middleware for Web services SLA negotiation//Proceedings of the IEEE International Conference on Web Services (ICWS'09). Los Angeles, CA, 2009; 1043-1050
- [9] Ni Wan-Cheng, He Ling-Juan, Liu Lian-Chen et al. Commodity-market based services selection in dynamic Web service composition//Proceedings of the IEEE Asia-Pacific Services Computing Conference (APSCC'07). Tsukuba Science City, 2007; 218-223
- [10] Garg R, Saran H, Randhawa R S, Singh M. A SLA framework for QoS provisioning and dynamic capacity allocation//Proceedings of the 10th International Workshop on Quality of Service (IWQoS'02). Miami Beach, 2002; 129-237
- [11] Xiong Kai-Qi, Perros H. SLA-based service composition in enterprise computing//Proceedings of the 16th IEEE International Workshop on Quality of Service (IWQoS'08). Enschede, 2008; 30-39
- [12] Jurca Radu, Faltings Boi. Reputation-based service level agreements for Web services//Benatallah B, et al. Lecture Notes in Computer Science 3826. Springer, 2005; 396-409
- [13] Gu Xiao-Hui, Nahrstedt K, Chang R N et al. QoS-assured service composition in managed service overlay networks//Proceedings of the IEEE International Conference on Distributed Computing Systems (ICDCS'03). Providence, RI, 2003; 1-8
- [14] Zeng Liang-Zhao, Benatallah B, Ngu Anne H H et al. QoS-aware middleware for Web services composition. IEEE Transactions on Software Engineering, 2004, 30(5): 311-327
- [15] Yu Tao, Zhang Yue, Lin Kwei-Jay. Efficient algorithms for Web services selection with end-to-end QoS constraints. ACM Transactions on the Web, 2007, 1(1): 1-26
- [16] Perryea C A, Chung S. Community-based service discovery//Proceedings of the IEEE International Conference on Web Services (ICWS'06). Chicago, IL, 2006; 903-906
- [17] Chawathe S S. Strategic Web-service agreements//Proceedings of the IEEE International Conference on Web Services (ICWS'06). Chicago, IL, 2006; 119-126



**WANG Xian-Zhi**, born in 1984, Ph.D. candidate. His current research interests include service computing and service engineering.

**XU Xiao-Fei**, born in 1962, Ph. D., professor, Ph. D. supervisor. His main research interests include service science and engineering, enterprise computing and service computing.

**WANG Zhong-Jie**, born in 1978, Ph. D., associate professor. His main research interests include service engineering, service-oriented architecture, service composition and software engineering.

## Background

This work is supported by the National Natural Science Foundation of China (Nos. 61033005, 60803091 and 70971029) and the National High Technology Research and Development Program (863 Program) of China under grant No. 2009AA04Z151.

These projects are all around the joint content of service innovation and value-based service system design under the context of BIRIS. BIRIS is a bilateral customer value oriented resource integration service mode proposed by HIT-ICES based on the analysis of mash-up stories of industry. By introducing a third-party service platform/service agency, the matching between service requirements and service capabili-

ties is performed. To secure the sustainability of services, the realization of service value of both service providers and service consumers becomes necessary. Now that traditional approaches fail/neglect to cope with such circumstances, which concentrate mainly on improving service consumers' satisfaction, this work presents a profit optimization method for composite services based on Service Level Agreement (SLA) aiming at solving this problem. The work does not only make sense to BIRIS, but can also be applied to various domains such as e-commerce, manufacture and healthcare services.