

基于改进 DAG 的 Web 服务组合优化^①

段立, 侯兴哲, 陈俐冰, 慕福林, 王懿, 孙洪亮

(重庆市电力公司, 重庆 400015)

摘要: 针对当前 Web 服务组合过程的复杂性, 提出了一种基于改进 DAG 的 Web 服务组合方法. 该方法通过构造有向无环图(DAG)表示服务之间可能的匹配, 并对有向无环图进行优化, 删除无效的组合路径及冗余服务, 采取逆向搜索并结合服务的 QoS, 减少处理的服务数量, 对筛选得到的组合服务通过比较 QoS 值选取最优的服务. 为了验证以上方法的有效性, 以智能用电互动化支撑平台的企业用能分析业务的生成设计实验, 并针对不同规模的服务集进行了实验分析. 结果证明本文提出的基于改进 DAG 的 Web 服务生成方法具有良好的可行性, 能够生成满足用户需求且具有较高服务质量的优化服务, 适用于智能用电互动化支撑平台的 Web 服务组合生成.

关键词: 智能用电; 服务组合; QoS; 有向无环图; 用能分析

Optimizing Web Service Composition Based on Improved DAG

DUAN Li, HOU Xing-Zhe, CHEN Li-Bing, MU Fu-Lin, WANG Yi, SUN Hong-Liang

(Chongqing Electric Power Company, Chongqing 400015, China)

Abstract: For the complexity of web service composition in current, this paper proposes a web service composition method based on improved DAG (Directed Acyclic Graph). This method presents the possible service matching through constructing DAG. Then it optimizes DAG by deleting the invalid composited-pathways and redundant service and adopting backward-search strategy with QoS to reduce the amount of web service. At last it calculates the QoS of choosing service to acquire the best service. In order to verify the feasibility of the above strategy, this paper makes an experiment on enterprise energy-consuming analysis service composition on interactive plat for smart power, and takes experimental analysis for services of different size. The results show that the improved strategy is more effective to generate optimizing service with high satisfaction and quality, and strategy applies to service composition on interactive plat for smart power.

Key words: smart power; service composition; QoS; DAG; energy-consuming analysis

当前 Web 服务的执行环境复杂多变, 且数量快速增长, 单个 Web 服务所提供的功能相对简单无法适应某些特定环境, 而用户需求却变得复杂, 因此需要将简单服务组合成适用的可靠组合服务, 实现多个 Web 服务的交互协作, 从而满足用户的需求^[1].

针对 Web 服务之间的集成问题, 国内外展开了大量的研究工作, 并取得了许多成果. 传统的 Web 服务组合方法无法很好地兼顾服务质量和组合效率问题, 容易造成 Web 组合服务生成过程中服务质量 QoS 的降低以及过多的能耗. 本文提出了一种基于改进的有向无环图(Directed Acyclic Graph, DAG)的 Web 服务

生成方法, 通过构造有向无环图表示服务之间的匹配, 并采用逆向搜索删除冗余服务对有向无环图进行优化, 对筛选得到的组合服务通过比较 QoS 值选取最优服务. 该方法应用在智能用电领域, 针对其分散业务整合的需求, 将用电信息采集系统、95598 系统、SG186 营销业务系统等外围系统提供的服务接入到智能用电互动化支撑平台, 对功能相对单一的简单服务进行快速高效地组合, 满足智能用电互动化支撑平台的复杂业务需求. 以企业用能分析服务的生成为例, 本文的 Web 服务生成方法能够在保证组合服务 QoS 的前提下, 快速高效地生成满足企业用能分析的优质组合服务.

① 收稿时间:2014-05-26;收到修改稿时间:2014-07-07

1 Web服务及服务组合的相关研究

1.1 Web 服务及生成过程

Web 服务体系结构包括服务提供者、服务请求者和服务注册中心^[2]。服务的生成过程一般分为服务的发现和服务的组合优化过程。在服务发现过程中, 服务提供者发布服务相关信息至 UDDI^[3](Universal Description Discover Integrity, 通用描述、发现和集成)注册中心, UDDI 受理请求并通过相应的格式转换对服务的功能以及 QoS 信息进行发布与存储, 以被服务查找。服务请求者发送查询请求, UDDI 受理请求并通过相应的格式转换进行服务之间的匹配, 并对请求服务与候选服务进行匹配, 返回一系列具有相同功能的候选服务, 并从中筛选出符合用户要求的服务。

1.2 服务组合的研究

Web 服务组合, 能够通过 Web 服务间的交互协作实现服务的增值, 解决更为复杂的问题^[4,5]。传统的 Web 服务只能在语法层面描述, 无法支持 Web 服务之间的动态组合和交互协作^[6], 而语义 Web 服务^[7]由于增加语义信息则可以支持, 其增强了人与机器、机器与机器之间的交互能力^[8]。

近年来, 关于语义 Web 服务组合优化的研究已经成为 Web 服务研究领域的一个重要课题, 并产生了许多创新成果。目前的语义 Web 服务组合方法有许多不同的分类方法, 一般主要有基于工作流、基于人工智能以及基于其他方法。

基于工作流的 Web 服务组合方法是基于业务流程, 将 Web 服务技术与工作流技术相结合, 实现流程的自动化处理, 获得有效的、可执行的流程式组合服务。文献[9]提出了一种业务级的基于服务模板的大粒度服务组合方法 TROLL, 通过服务模板规约组合目标, 并结合业务过程中的多约束下的参数进行求解, 满足动态变化的业务需求, 且简单易用。文献[10]提出了一种工作流环境下能耗感知的多路径服务组合方法, 在满足 QoS 约束的情况下充分考虑了服务的能耗优化。

基于人工智能(Artificial Intelligence, AI)的语义 Web 服务组合将 Web 服务组合转换为 AI 规划问题, 通过 IOPE(input, output, precondition 和 effect)描述 Web 服务的功能和行为, 并通过形式化描述和推理来组合服务。文献[11]对 Web 服务进行基于规则的建模, 并采用消除语义冲突和后向演绎, 此方法能够使用户通过最小的执行代价来产生最优组合方案。文献[12]以

最优服务组合的 QoS 为约束条件来求最优解, 但是此方法的应用范围较小。

基于形式化演算的语义 Web 服务组合主要是借助数学方法和形式化工具, 对 Web 服务进行推理和自动组合。文献[13]从 Web 服务执行的角度, 利用 Petri 的理论和方法, 将各子 Web 服务的 Petri 网合并, 自动生成 Web 组合服务。文献[14]提出了一种扩展时间 Petri 网, 分析计算了 Web 服务组合的时间成本, 检测了变迁冲突, 并能有效解决服务组合中的时间成本问题。

目前的 Web 服务组合方法由于在实现过程中处理的服务数量较多, 造成占用过多的处理时间, 如何快速高效地生成满足需要的组合服务, 显得尤为重要。因此本文针对 Web 服务组合, 基于有向无环图进行了研究, 提出了基于改进 DAG 的 Web 服务组合方法。

2 基于改进DAG的Web服务组合方法

基于改进 DAG 的 Web 服务生成方法是通过构造有向无环图, 并逐次筛选与用户需求相匹配的服务对有向无环图进行扩展, 初步得到满足用户需求的几组 Web 服务。采用逆向搜索对有向图进行优化, 删除冗余服务, 减少处理的服务数量, 并通过比较组合服务的 QoS, 得到满足用户需求的最优服务。

2.1 有向无环图构造

有向无环图(Directed Cyclic Graph, DAG)是指一个无环的有向图, 被广泛应用于描述业务逻辑。鉴于 DAG 的简单直观, 用其对 Web 服务工作流进行建模。

有向无环图的点代表服务, 边代表流程。每条起点到终点的路线都是一条可能的完整执行流程。构造有向无环图表示服务之间可能的匹配。

2.1.1 有向无环图初始化

有向无环图中有两种类型的节点, 服务节点和选择节点, 服务节点包括具体的服务, 选择节点用于决定各层服务之间的连接关系。服务节点的服务输入、输出以及服务质量 QoS 都可以在有向图中表示, 如服务 S 可表示为 $S=\{I_s, O_s, Q_s\}$, 其中 I 为输入, O 为输出, Q 为服务质量。为用户的服务请求设置起始节点, 提供输入的请求 I_R ; 为处理结束设置终止节点, 向外提供输出的处理结果 O_R 。

初始化输入参数集合 I_R 且 $I_R=\{I_1, I_2, \dots, I_i\}$, 作为服务节点的源节点。

2.1.2 有向图扩展

根据服务的输入输出参数, 对有向无环图进行扩展. 后续节点的服务是在前续节点服务的基础上, 通过匹配确定的, 后续节点服务的输入条件需要包含前续节点服务的输出条件.

设某节点 i 的服务 S_i 输出参数概念集合为 $O_i = S_{out}^i$, 其后续节点为 $i+1$, 其服务 S_{i+1} 的输入概念集合为 $I_{i+1} = S_{in}^{i+1}$, 如果 $S_{in}^{i+1} \subseteq S_{out}^i$, 则 S_i 和 S_{i+1} 可以相互连接匹配.

两个概念集合 O_i 和 I_{i+1} 之间的语义距离 $Dist(O_i, I_{i+1})$ 为连接它们的最短路径上的 n 条边的权值的总和, 即:

$$Dist(O_i, I_{i+1}) = \sum_{i=1}^n weight_i \quad (1)$$

其语义相似度为:

$$sim(O_i, I_{i+1}) = \frac{1}{e^{Dist(O_i, I_{i+1})}} \quad (2)$$

服务 S_i 和 S_{i+1} 的输入输出参数集合的匹配是将 O_i 集合内的每个输出概念与 I_{i+1} 集合内的每个输入概念逐一匹配, 计算所有求得的概念匹配度的平均值, 就是输入输出概念集合的匹配度.

设置匹配度阈值, 将符合输入输出匹配度的服务作为后续节点加入有向无环图, 对有向无环图进行扩展.

2.1.3 有向无环图生成

从起始节点依次扩展到终止节点, 这样就可以得到数条由起点出发的路径, 每条能到达终止节点的路径上的服务组合就是符合要求的组合服务. 但是并不是所有由起始节点出发的路径都能到达终止节点, 即存在冗余服务, 并且 QoS 没有得到最优, 需要对有向无环图进一步进行优化.

2.2 有向无环图优化

针对有向无环图的冗余服务过多、处理的服务数量过多等缺点, 应用以下规则对有向无环图进行优化, 在保证服务组合质量的同时提高组合速度.

规则 1. 在有向无环图中, 有些服务路径无法到达终止节点, 即此组合服务不包含最终的输出参数, 则舍弃该路径, 逆向搜索时不再沿此路径搜索. 若有服务节点只存在于此无效路径, 而在其他服务组合路径中没有参与, 则从有向无环图中去除该服务节点.

规则 2. 逆向搜索从终止节点开始, 添加与之相连

的服务, 并计算每条路径上服务的 QoS 值.

1) 如果只有一条路径, 则此路径上的服务为最优服务.

2) 如果大于两条路径, 则比较每条路径上服务组合的 QoS 值, 选取 QoS 值大的为最优组合服务. 若不同路径上的两个节点具有相同的前续节点, 则比较两个节点的 QoS 值, 选取 QoS 值较大的节点继续进行搜索, QoS 值较小的节点则舍弃, 其所在的服务组合路径不再考虑.

以每个服务的 QoS 计算组合 QoS 的方法包括相加、 Σ 相乘、求最大值 MAX 和求最小值 MIN. 典型的 QoS 参数有响应时间、可靠性、执行费用及信誉度等, 其中响应时间和执行费用越小越好, 取最小值; 可靠性和信誉度越大越好, 取最大值.

将 i 节点的服务 S_i 的 QoS 用 (T_i, R_i, M_i, C_i) 描述, 其中, T 指响应时间, R 指可靠性, M 指费用, C 指信誉度. 该节点所在的服务组合路径有 N 个服务, 则该定义目标函数如下:

$$Q(T) = \sum((1-T_1), \dots, (1-T_i), \dots, (1-T_N))$$

$$Q(R) = \prod(R_1, \dots, R_i, \dots, R_N)$$

$$Q(M) = \sum((1-M_1), \dots, (1-M_i), \dots, (1-M_N)) \quad (3)$$

$$Q(C) = \prod(C_1, \dots, C_i, \dots, C_N)$$

$$GlobalQoS(S) = W_1 * Q(T) + W_2 * Q(R) + W_3 * Q(M) + W_4 * Q(C)$$

其中, $W_1 + W_2 + W_3 + W_4 = 1$, W_i 具体的数值根据服务的具体特点来选取, 范围均是 $[0, 1]$.

基于优化策略的 Web 服务生成方法, 将服务请求的输入参数作为初始状态构造有向无环图, 并对有向图进行优化, 删除无效的组合路径及冗余服务, 采取逆向搜索并结合服务的 QoS 参数, 减少处理的服务数量, 从而快速高效地生成具有较高服务质量的优化服务.

3 实例分析

为证明基于改进 DAG 的 Web 服务组合方法的有效性, 本文以重庆市智能用电互动化支撑平台的企业用能分析服务生成为例来说明. 智能用电互动化支撑平台生成 Web 服务的结构如图 1 所示:

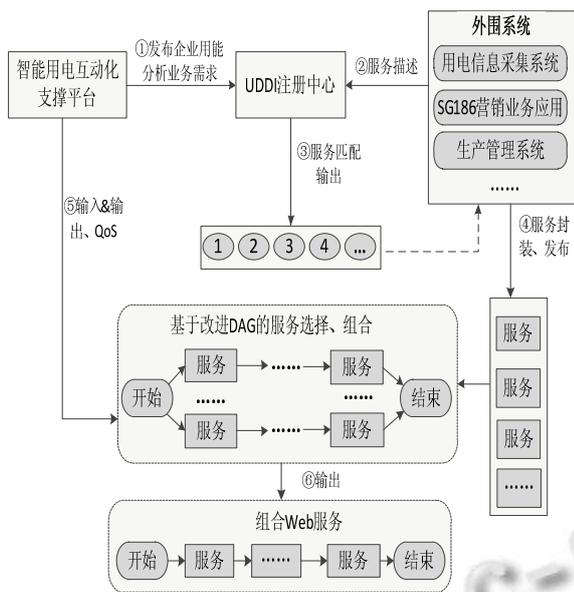


图 1 智能用电互动化支撑平台 Web 服务生成结构

平台发布企业用能分析业务需求(包括用户负荷分析、用电设备信息采集、企业用电量分析等服务的 QoS 描述和可提供的输入及期望的输出), 发送到 UDDI 注册中心。

外围系统(用电信息采集系统、SG186 营销业务应用系统、生产管理系统等)在互动化支撑平台发布需求后, 参照发布的业务功能需求在 UDDI 注册中心上查找其可发布的业务功能。平台将业务功能需求转换为领域本体, 并利用领域本体与外围系统服务功能描述进行服务匹配。外围系统将功能需求匹配成功的服务进行封装、发布。智能用电互动化平台对每个服务提出输入输出、QoS 需求, 并基于优化策略进行服务选

择组合, 输出符合用户需求的最优组合 Web 服务。

3.1 有向无环图构建

企业用能分析服务的组合生成首先进行有向无环图的构建。用 $\{I_R\}$ 表示服务请求的输入参数集合, $\{O_R\}$ 表示服务请求的输出参数集合。设置服务匹配度的阈值为 $\alpha=0.9(\alpha$ 的取值可自定义, 用于调整服务匹配的精确度, 取值范围为 $[0,1]$, 值越大匹配度越高), 构造有向无环图:

1)初始化参数 $\{I_R\}$ 作为有向无环图起始节点, 其中用电信息采集系统提供的用户负荷分析服务 S_1 、负控系统提供的用户负荷分析服务 S_2 、生产管理系统提供的用户负荷分析服务 S_3 同 O_0 的输入输出匹配度分别为: 0.925、0.938、0.912, 均大于阈值 α , 因此将服务 S_1 、 S_2 、 S_3 加入有向无环图。

2)计算服务之间的输入输出匹配度, 其中用电信息采集系统提供的企业用电设备信息采集服务 S_4 与有向无环图中服务 S_1 的匹配度为 0.915, 生产管理系统提供的企业用电设备信息采集服务 S_5 与有向图中服务 S_2 、 S_3 的匹配度为 0.937、0.926, 均大于阈值 α , 因此将服务 S_4 、 S_5 分别作为 S_1 、 S_2 、 S_3 的后续节点加入有向无环图。

3)依次扩展, SG186 营销系统、用电信息采集系统、费控系统分别提供的用电量服务 S_6 、 S_7 、 S_8 作为 S_4 的后续节点, S_7 与 S_8 作为 S_5 的后续节点, 将 S_6 、 S_7 、 S_8 加入有向图。

4)其中 S_7 与 S_8 包含输出参数集合 $\{O_R\}$, 到达终止节点, 停止有向图的扩展。

初步的有向无环图构建完成, 如图 2 所示:

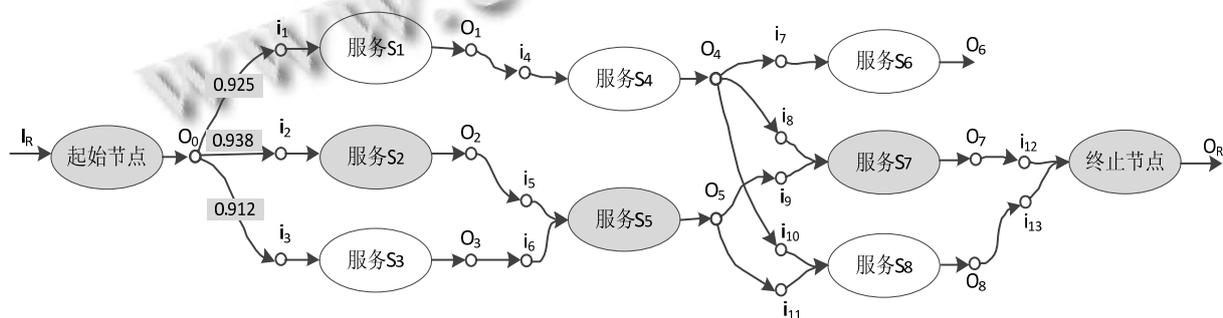


图 2 企业用能决策分析服务的有向无环图

图中所涉及到的各服务的具体描述如下:

S_1 : 用电信息采集系统根据采集到的测量点负荷

曲线进行分析, 提供用户负荷分析服务。

S_2 : 负控系统利用采集的负控户总加组负荷信息

进行分析, 提供负控户负荷分析服务.

S₃: 生产管理系统对专线用户的负荷信息进行分析, 提供专线路的负荷分析服务.

S₄: 用电信息采集系统提供对电力用户的用电设备的信息采集服务;

S₅: 生产管理系统提供关口考核点的实时负荷信息.

S₆: SG186 营销系统提供用户月度结算电量服务;

S₇: 用电信息采集系统能够提供用户采集计量点的电量服务;

S₈: 费控系统能够提供费控户的购电量、用电量及剩余电量服务.

3.2 有向无环图优化

优化策略 1. 有向无环图中 S₆ 节点无法到达终止节点, 也就是不包含最终的输出参数, 所以舍弃 S₆ 所在的路径, 逆向搜索时候不予考虑. 其中 S₆ 节点只存在于此条路径, 不参与其他路径, 所以将 S₆ 节点从有向无环图中去除; 该路径中的节点 S₁、S₄ 参与其他路径的服务组合, 保留 S₁、S₄.

优化策略 2. 对有向无环图进行逆向搜索, 首先计算有向无环图中剩下节点的 QoS 参数值, 如下表所示:

表 1 有向无环图中各服务的 QoS 参数

S _i	T(时间)	R(可靠性)	M(费用)	C(信誉度)
S ₁	0.15	0.78	0	0.70
S ₂	0.11	0.92	0	0.82
S ₃	0.17	0.77	0	0.75
S ₄	0.15	0.82	0	0.79
S ₅	0.12	0.95	0	0.85
S ₇	0.13	0.86	0	0.84
S ₈	0.21	0.71	0	0.74

在有向无环图中, 服务 S₇ 和 S₈ 的前续节点是一样的, 但是 QoS₇>QoS₈, 所以舍弃 S₈ 节点及其所在的服务组合路径(S₈→S₄→S₁)、(S₈→S₅→S₂)、(S₈→S₅→S₃), 选择 S₇.

与服务 S₇ 相关的路径共有三条, 计算其 QoS 值. 其中根据智能用电互动化支撑平台的企业用能分析服务的特点, 服务的费用为 0, 且对服务的响应时间和可靠性要求较高, 所以设服务的 QoS(T, R, M, C)各参数值为: W₁(T)=0.4, W₂(R)=0.4, W₃(M)=0, W₄(C)=0.2. 由公式(3)计算组合服务 S₇→S₄→S₁ 的 QoS 值为:

$$\begin{aligned}
 QoS(S_7 \rightarrow S_4 \rightarrow S_1) &= W_1 * Q(T) + W_2 * Q(R) \\
 &+ W_3 * Q(M) + W_4 * Q(C) \\
 &= W_1 * \sum((1-T_7), (1-T_4), (1-T_1)) \\
 &+ W_2 * (\prod(R_7, R_4, R_1)) + 0 \\
 &+ W_4 * (\prod(C_7, C_4, C_1))
 \end{aligned} \tag{4}$$

将表 1 的数据带入公式(4)得: QoS(S₇→S₄→S₁)=1.353; 同理得, QoS(S₇→S₅→S₂)=1.466, QoS(S₇→S₅→S₃)=1.379. 其中 QoS(S₇→S₅→S₂)=1.466 最大, 所以 S₇→S₅→S₂ 为最优路径, 此路径上的服务 S₂、S₅、S₇ 为最优组合服务. 即负控系统提供的用户负荷分析服务 S₂、生产管理系统提供的企业用电设备信息采集服务 S₅、SG186 营销系统提供的企业用电量服务组成的组合服务 S₇ 为满足企业用能分析的最优组合服务.

3.3 事件复杂度比较分析

针对企业用能分析服务的组合生成, 基于有向无环图的 Web 服务组合方法采用优化策略时, 计算其时间复杂度为: 设服务 7 和服务 8 的 QoS 比较的时间复杂度为 o(1), 每条路径上组合服务的 QoS 计算的时间复杂度为 T(n)=2o(m*n3)+o(mn). 则总的时间复杂度为 o(1)+3T(n).

未采用优化策略的基于有向无环图的 Web 服务生成方法由于未删除冗余服务, 需要计算所有路径的组合服务的 QoS 值, 其时间复杂度为 7T(n).

将两种方法处理服务的数量、处理的组合路径数、时间复杂度进行比较, 如下表所示:

表 2 采用优化策略与未采用优化策略的服务组合方法比较

	服务数量	组合路径数	时间复杂度
优化策略	5	3	o(1)+3T(n)
无优化策略	8	7	7T(n)

3.3 不同规模服务集的实验分析

为了验证本文提出的采用优化策略的基于有向图的 Web 服务组合方法的有效性, 针对不同规模的服务测试集进行实验分析.

实验 1. 在服务总数为 10、50、100、200 的 4 个实验数据集上比较两种方法的服务组合效率, 为了避免计算机的硬件系统对算法性能造成影响, 评价指标为 50 此实验的平均组合时间. 得到的实验结果如图 3 所示.

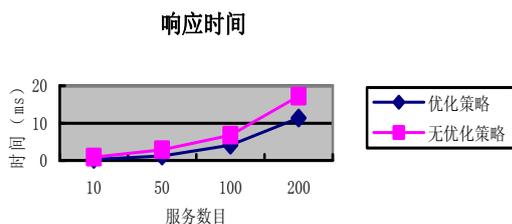


图 3 两种方法的响应时间比较

由图中两种方法的时间变化趋势可以看出,采用优化策略的方法具有更短的服务组合生成时间,其性能有所提升。

实验 2. 在服务总数为 10、50、100、200 的 4 个实验数据集上比较两种方法在 Web 服务搜索匹配时的遍历节点数目,遍历节点数目的多少直接反应了算法搜索区域的大小,由图 4 可以看出,采用优化策略的节点数目明显要少,即搜搜区域较小,此方法能够以较小的搜索区域得到最优 Web 服务。

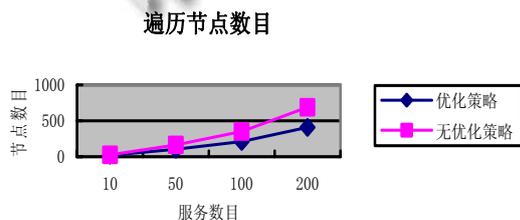


图 4 两种方法的遍历节点数目比较

由实验可以看出,本文提出的采用优化策略的基于 DAG 的 Web 组合方法能够以较少的响应时间和搜索区域得到最优服务,具有较高的性能。

3 结语

本文针对 Web 服务的组合问题提出了基于改进的有向无环图的方法,通过构造有向无环图表示服务之间可能的匹配,采取逆向搜索删除冗余服务,减少处理的服务数量,并通过比较服务的 QoS 选取最优服务。通过对智能用电互动化支撑平台中企业用能分析业务的生成进行实例分析,并针对不同规模的服务测试集进行了实验分析。结果表明,本文提出的改进方法能

够通过处理较少的服务数量,快速高效的生成具有较高 QoS 的优化服务,满足智能用电领域中分散业务整合的问题。

参考文献

- Rao J, Kungas P, Matskin M. Logic-based Web services composition: From service description to process model. Proc. IEEE International Conference on Web Services. IEEE. 2004. 446-453.
- 柴晓路,梁宇奇. Web Services 技术、架构和应用. 北京:电子工业出版社,2003:6-48.
- 马恒太,李鹏飞,颜学熊,等. Web 服务安全. 北京:电子工业出版社,2007:101-104.
- Sycara K, Paolucci M, Ankolekar A, et al. Automated discovery, interaction and composition of semantic web services. Web Semantics: Science, Services and Agents on the World Wide Web, 2003, 1(1): 27-46.
- Medjahed B, Bouguettaya A, Elmagarmid A K. Composing web services on the semantic web. The VLDB Journal, 2003, 12(4): 333-351.
- 岳昆,王晓玲,周傲英. Web 服务核心支撑技术:研究综述. 软件学报,2004,15(3):428-442.
- Ponnekanti SR, Fox A. Sword: A developer toolkit for web service composition. 11th World Wide Web Conference (Web Engineering Track). 2002. 7-11.
- 程永上,王志坚. 语义 Web 服务组合的形式化描述与验证研究. 计算机科学,2008,35(4):40-43.
- 胡海涛,李刚,韩燕波. 一种面向业务用户的大粒度服务组合方法. 计算机学报,2005,28(4):694-703.
- 朱勇,罗军舟,李伟. 一种 workflow 环境下能耗感知的多路径服务组合方法. 计算机学报,2012,35(3):627-638.
- Liu JM, Chen HF, Ning G. Web services automatic composition with minimal execution price web services. IEEE International Conference, 2005.
- 邓水光,黄龙涛,吴斌,等. 一种 QoS 最优的语义 Web 服务自动组合方法. 计算机学报,2013,36(5): 1015-1030.
- 马炳先,相东明,张正明. Web 服务组合的 Petri 网自动生成方法. 小型微型计算机系统,2013,34(2):332-337.
- 高海宁,李蜀瑜. 基于扩展时间 Petri 网的 Web 服务组合的分析与验证. 计算机应用与软件,2012,29(3):82-85.