文章编号:1001-506X(2010)01-0179-04

基于本体的 Web 服务连接研究

吴 奎[1],周献中[2],萧毅鸿[2],施爱博[2],田卫萍[3]
(1. 南京理工大学自动化学院,江苏 南京 210094;
2. 南京大学工程管理学院,江苏 南京 210093;
3. 北方自动控制技术研究所,山西 太原 030006)

摘 要:针对 Web 服务组装过程中服务输入-输出接口连接程度量化和多分量连接对应关系判定问题,提出了一种新的连接度量化方法。该方法基于领域本体中的概念条件出现概率,将服务连接程度定义为输入-输出接口概念的替换可能性。在此基础上,利用二部图稳定匹配算法给出了服务接口各分量之间连接对应关系判定算法。最后讨论了不同连接样式下复合服务连接度计算,并分析算法的时间性能。

关键词:服务组合;服务连接度;概念出现概率;稳定匹配;复合服务质量
中图分类号:TP 311 **文献标志码**:A

Research on Web service connection based on ontology

WU Kui[1], ZHOU Xian-zhong[2], XIAO Yi-hong[2], SHI Ai-bo[2], TIAN Wei-ping[3]
(1. *School of Automation, Nanjing Univ. of Science and Technology, Nanjing 210094, China*;
2. *School of Management & Engineering, Nanjing Univ., Nanjing 210093, China*;
3. *North Inst. of Automatic Control Technology, Taiyuan 030006, China*)

Abstract: Aiming at connectivity computation and interfaces correspondence analysis in Web service connection, the concept subjective support and confidence definition is presented based on the concept's conditional appearance probability in domain ontology. And then the connective degree between two service interfaces is quantified. Furthermore, an adapted stable matching algorithm is introduced to solve the interfaces correspondence judgment. Finally the service connection is discussed in varied cases.

Keywords: service composition; service connectivity; concept appearance probability; stable matching; composite service connectivity

0 引 言

面向服务架构以服务为核心业务逻辑,提出了一种可即需交换信息的软件架构体系,以便更快更高效地构建企业信息系统。随着应用的深入,面临的业务领域问题越来越复杂,单个服务越来越难以满足应用需求,服务组合成了一种新的开发范例,如何更快地实现服务组合也越来越被研究人员关注。服务组合包括服务的描述、实现、查找、匹配、组装、评估等过程,由于涉及问题较多,目前尚未见成熟应用。在服务的组合过程中如何分析服务之间的连接关系,以便更好地给编排人员提供参考,更好地为服务自动组合提供判定依据,是实现服务组合的基础问题。本文针对该问题,提出了一种基于领域本体中概念条件出现概率的服务连接度计算指标,并结合二部图稳定匹配实现了多分量间连接对应关系判定算法。

1 相关工作

Web 服务连接本质上是将一个服务的输出以另一个服务所理解的形式输入到对方,因而借助领域本体描述服务接口语义,通过计算服务接口语义之间的关联程度,就能够分析出服务间的连接关系,从而为服务组合提供参考和判定依据,因此服务接口之间关联程度量化指标的定义和计算便成为服务组合的关键问题。

针对服务接口连接程度量化问题,文献[1]提出了一种基于接口语义父子关系的逻辑推理方法,将服务之间的连接程度分为 Exact、Plug-in、Subsume、Fail 四个级别,文献[2]进一步提出了 Intersection 级别,文献[3]在此基础上采用因果关系网络实现了服务组合。但是,几个离散的级别不能很好地量化服务之间连接程度。文献[4-5]以服务接口之间的概念语义相似度为连接度指标,然而概念 A 和概念 B 相似

收稿日期:2008-07-23; **修回日期**:2009-06-16。
基金项目:总装"十一五"背景项目(10408)资助课题
作者简介:吴奎(1979-),男,博士研究生,主要研究方向为 Web 服务自动组合。E-mail:witkin@163.com

只是从术语学的角度表明这两个概念之间的语义相关性，服务调用过程中并不能将概念 A 的个体当作概念 B 的个体在服务连接数据流中传递。文献[6]等提出根据接口参数名的聚类分析衡量服务之间的关联程度，然而一旦采用不同的参数命名规则或存在拼写错误等，则结果的正确性将大大降低。同时，上述几种方法均未解决服务接口多分量之间连接对应关系匹配问题。

2 服务连接度量化及稳定匹配

2.1 基本假设与定义

不失一般性，此处对研究内容做如下假设：

(1) 服务语义假设：每个服务的输入输出接口由领域本体的概念描述，当传来的数据匹配于输入接口时，服务是可运行的。当服务运行后，其输出数据匹配于输出接口。

(2) 服务接口分量概念独立性假设：服务的输入输出接口通常具有多个分量，每个接口分量的概念由设计者根据业务需求确定，不能根据某个分量获知其他分量的概念。

(3) 服务无状态假设：所涉及的服务都是状态无关的，服务的输出只与当前输入有关。

在此基础上给出如下定义：

定义 1 本体(ontology)可以被表示为 $O=(C,R,I,A)$。其中，C 表示概念集合；R 表示关系集合；I 表示个体集合；A 表示相关公理集合。

定义 2 概念(concept)是指具有相似特性的实体集合的名称。如果概念 A 具有概念 B 的特性，称概念 A 是概念 B 的子概念，记为 $A\sqsubseteq B$，它们之间的关系$\sqsubseteq$称为子类关系。如果概念 C 是概念 A、B 的子概念，则称概念 C 为概念 A、B 的公共子概念，记为 $C=A\sqcap B$。

定义 3 Web 服务由输入输出接口构成，表示为 $S(\boldsymbol{I},\boldsymbol{O})$。其中，$S$ 为服务标识符；$\boldsymbol{I}$ 表示服务的输入概念向量，$\boldsymbol{O}$ 表示输出概念向量。

2.2 服务连接度量化

如果服务 P 的输出 $\boldsymbol{O}_P$ 能够传递到 Q 的输入 $\boldsymbol{I}_Q$ 中，则概念 $\boldsymbol{O}_P$ 的某些个体属于 $\boldsymbol{I}_Q$。如果服务 Q 的输入 $\boldsymbol{I}_Q$ 可来自于 P 的输出 $\boldsymbol{O}_P$，则概念 $\boldsymbol{I}_Q$ 的某些个体属于 $\boldsymbol{O}_P$。因此有如下定义：

定义 4 在领域本体中，概念 X 对概念 Y 的替换支持度表示当个体 a 属于概念 X 时属于概念 Y 的概率，反映了概念 X 中的个体能够被解释为概念 Y 个体的可能性，即

$$S(X,Y)=P(a\in Y\mid a\in X)=\frac{P(a\in X\cap a\in Y)}{P(a\in X)} \tag{1}$$

概念 X 对概念 Y 的替换置信度为个体 a 属于概念 Y 时又属于概念 X 的概率，反映了概念 Y 的个体来自于概念 X 个体的可能性，即

$$C(X,Y)=P(a\in X\mid a\in Y)=\frac{P(a\in X\cap a\in Y)}{P(a\in Y)} \tag{2}$$

式中，$P(a\in X)\triangleq P(X)$ 表示概念 X 的出现概率，可近似地看作领域统计样本中概念 X 个体的出现次数与所有概念个体出现次数的比值。

通常 Web 服务为多个输出到多个输入的联合连接，因此有概念向量的替换支持度和置信度，定义如下：

定义 5 设 $\boldsymbol{X}$、$\boldsymbol{Y}$ 为两个相同维数的概念向量，$\boldsymbol{X}$ 对 $\boldsymbol{Y}$ 的替换支持度定义为

$$S(\boldsymbol{X},\boldsymbol{Y})=P(\boldsymbol{a}\in\boldsymbol{Y}\mid\boldsymbol{a}\in\boldsymbol{X}) \tag{3}$$

相应地，也有概念向量替换置信度的定义 $C(\boldsymbol{X},\boldsymbol{Y})=P(\boldsymbol{a}\in\boldsymbol{X}\mid\boldsymbol{a}\in\boldsymbol{Y})$。

定理 1 概念向量 $\boldsymbol{X}$ 对 $\boldsymbol{Y}$ 的替换支持度为各对应分量的替换支持度之积，即

$$S(\boldsymbol{X},\boldsymbol{Y})=\prod_i S(X_i,Y_i)。 \tag{4}$$

证明 根据服务输入输出分量概念独立假设，服务 S 的任意一组输出 $\boldsymbol{a}$，其中分量 a_i 与 a_j 属于哪个概念是独立的，即有 $I(a_1\in Y_1,\{a_2\in X_2,\cdots,a_n\in X_n\}\mid a_1\in X_1)$ 和 $I(a_1\in Y_1,a_2\in Y_2\mid a_1\in X_1)$。

根据条件独立弱联合定理[7]，则有 $I(a_1\in Y_1,a_2\in Y_2\mid\{a_1\in X_1,\cdots,a_n\in X_n\})$，即所有的分量 $a_i\in Y_i$ 相互独立于条件$\{a_1\in X_1,\cdots,a_n\in X_n\}$。

同样根据独立条件，有

$$P(a_i\in Y_i\mid\boldsymbol{a}\in\boldsymbol{X})=P(a_i\in Y_i\mid a_1\in X_1,\cdots,a_n\in X_n)=P(a_i\in Y_i\mid a_i\in X_i)$$

则

$$P(\boldsymbol{a}\in\boldsymbol{Y}\mid\boldsymbol{a}\in\boldsymbol{X})=P(a_1\in Y_1,\cdots,a_n\in Y_n\mid a_1\in X_1,\cdots,a_n\in X_n)=\prod_i P(a_i\in Y_i\mid a_1\in X_1,\cdots,a_n\in X_n)=\prod_i P(a_i\in Y_i\mid a_i\in X_i)$$

因此

$$S(\boldsymbol{X},\boldsymbol{Y})=P(\boldsymbol{a}\in\boldsymbol{Y}\mid\boldsymbol{a}\in\boldsymbol{X})=\prod_i P(a_i\in Y_i\mid a_i\in X_i)=\prod_i S(X_i,Y_i)$$

同理

$$C(\boldsymbol{X},\boldsymbol{Y})=\prod_i C(X_i,Y_i)$$ 证毕

定义 6 服务连接支持度表示服务 P 的输出 $\boldsymbol{O}_P$ 与 Q 的输入 $\boldsymbol{I}_Q$ 之间的概念向量替换支持度，记为 $S(P,Q)=S(\boldsymbol{O}_P,\boldsymbol{I}_Q)$；同样，服务连接置信度表示服务 P 的输出 $\boldsymbol{O}_P$ 与 Q 的输入 $\boldsymbol{I}_Q$ 之间的概念向量替换置信度，记为 $C(P,Q)=C(\boldsymbol{O}_P,\boldsymbol{I}_Q)$。

2.3 服务连接的稳定匹配

Web 服务的输入输出接口通常具有多个分量，两个可连接的服务接口之间分量个数和排列次序并非一一对应的，如图 1 所示。图中服务 P 的输出接口有 m 个分量，服务 Q 的输入接口有 n 个分量，其中 O_{pi} 可能与 $\boldsymbol{I}_q$ 中多个分量的支持度不为 0，根据分量概念独立假设，实际连接中只可能是一一连接的，故确定概念向量之间的连接对应关系后才能计算连接度大小。

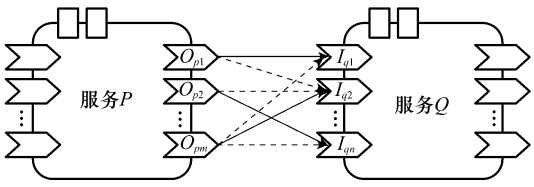

图 1 服务连接示意图

从输出的角度看,输出分量 O_{Pi} 倾向于与使得 s_{ij} 最大的 I_{Qj} 相连;从输入角度看,输入分量 I_{Qj} 倾向于与使得 c_{ji} 最大的 O_{Pi} 相连。因而,服务连接的接口对应关系判定可以转化为二部图稳定匹配问题[8],即以服务接口中的概念为节点,构造二部图 $G(\boldsymbol{O}_P,\boldsymbol{I}_Q,E)$,并以概念之间的支持度和置信度大小为优先选择顺序,寻找 G 中的一个匹配,使得连接最稳定。面向服务连接的二部图稳定匹配算法对传统稳定匹配算法做了如下改进:

(1) 传统的稳定匹配问题必须是对等匹配,而服务连接的双方通常是数量不等的,也不一定能让某一方饱和,因此算法中只要某次迭代结果相对上次不变,则算法终止。

(2) 如果某个概念替换支持度 s_{ij} 小于某个阈值 τ,则令 $s_{ij}=0$,置信度也做同样预处理。

(3) 若存在某对 $s_{ij}=0$ 或 $c_{ji}=0$,则 O_{Pi} 不能与 I_{Qj} 相连。

面向支持度的最优稳定匹配算法过程如下:

输入　服务 P 和服务 Q 相连时的概念替换支持度和置信度矩阵。

输出　服务 P 的输出接口到服务 Q 输入接口之间的连接对应关系。

步骤 1　每个 O_{Pi} 向与它的支持度最大的 I_{Ok} 发出连接请求,每个 I_{Ok} 在向它发出请求的 O_{Pi} 中选择置信度最大的分量连接。

步骤 2　所有剩余的自由的 O_{Pk} 向各自次优的 I_{Ol} 发出请求,每个 I_{Ol} 在它上次迭代的 O_{Pi} 和本次迭代的 O_{Pk} 中选择置信度最大的分量相连。

步骤 3　重复步骤 2 直至两次迭代的结果不变,此时获得了一个面向支持度的最优稳定匹配。可以证明,本算法最大迭代次数为 mn[8]。

2.4　应用示例

如在一体化指挥系统中侦察服务 R 的输出为时间、位置、目标类型和重要度,火炮弹药量计算服务 K 的输入为炮目距离、装甲目标、毁伤指标,则服务 R 和 K 输入输出接口之间的概念替换支持度和置信度矩阵如表 1 所示。

表 1　服务 R 和 K 接口概念替换支持度与置信度矩阵

	时间	位置	目标类型	重要度
距离	0, 0	0, 0	0, 0	0, 0
装甲目标	0, 0	0, 0	0.7, 0.9	0, 0
毁伤指标	0, 0	0, 0	0, 0	0.6, 0.5

本例中各个概念之间的支持度和置信度以 WordNet[9] 为数据源,统计各个单词的出现次数和概念层次关系,经计算获得。实际应用中可由领域样本统计或由领域专家指定概念出现概率。两个概念之间支持度和置信度均为 0 表示在领域本体中二者不具备父子关系。

若采用语义相似度分析服务 R 和 K 之间的连接关系,则得到表 2 所示接口语义度矩阵,表 2 中相似度数值大小根据 WordNet::Similarity[9-10] 计算得出。

表 2　服务 R 和 K 接口概念语义相似度矩阵

	时间	位置	目标类型	重要度
炮目距离	0.769 2	0.769 2	0.533 3	0.800 0
装甲目标	0.571 4	0.500 0	0.666 7	0.533 3
毁伤指标	0.800 0	0.666 7	0.533 3	0.947 4

概念语义相似度计算有多种算法,此处选取了其中 WUP 算法的计算结果,由于概念相似度具有主观性,其数值大小非实质性的,能提供排序准则以便优选即可。

对照表 1 和表 2 内容不难看出,尽管表 2 中“时间”与“指标”两个概念之间具有一定的语义相似度,但在服务调用过程中并不能将二者所对应的数据信息在服务接口中传递。而概念支持度和置信度反映了两概念所对应个体数据能正确转义的概率,因而与语义相似度方法比,其数学意义更明确,也能提供更好的组合参考依据。

3　复合服务连接质量分析

服务连接支持度和置信度刻画了两两服务之间的依赖程度,复合服务是多个服务的组合,本文将业务需求所提供的输入数据抽象为初始服务 $\perp$,将期望输出抽象为目标服务 $\top$。在复合服务中,初始服务经过服务组合后对目标服务的支持程度称为服务支持度,反映了目标服务的可运行概率;目标服务经过服务组合后对初始服务的需求程度称为服务置信度,反映了初始服务的价值程度。不失一般性,此处只研究服务支持度计算。

3.1　基本连接样式

如果服务 Q 的所有输入分量均来自于服务 P,则称服务 P 与 Q 饱和相连。本文先研究基本服务连接方式,包括顺连、分连和汇连,如图 2 所示。当已知前驱服务的可运行概率时,如何根据服务连接支持度计算后继服务的可运行概率如下所述。

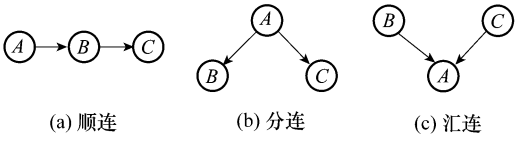

图 2　基本服务连接方式

(1) 顺连

顺连结构如图 2(a)所示,服务 B 只接受 A 的输出,服务 C 只接受 B 的输出,且 AB 之间、BC 之间均为饱和相连。

定理 2　在服务组合层次结构中,服务 B 只与 A 相连且为饱和连接,连接支持度为 $S(A,B)$。若服务 A 的可运行概率为 $P(A)$,则服务 B 的可运行概率为

$$P(B)=S(A,B)P(A) \tag{5}$$

证明　根据假设 4,有 $P(A)=P(\boldsymbol{a}\in\boldsymbol{I}_A)=P(\boldsymbol{b}\in\boldsymbol{O}_A)$,其中 $\boldsymbol{a}$ 和 $\boldsymbol{b}$ 为个体向量,又由于服务 B 仅与 A 相连,则 B 运行时 A 必然将数据传递给 B,即

$$P(B)=P(\boldsymbol{a}\in\boldsymbol{I}_B\cap\boldsymbol{a}\in\boldsymbol{O}_A)$$

因此

$$S(A,B)=P(\boldsymbol{a}\in\boldsymbol{I}_B\mid\boldsymbol{a}\in\boldsymbol{O}_A)=\frac{P(\boldsymbol{a}\in\boldsymbol{I}_B\cap\boldsymbol{a}\in\boldsymbol{O}_A)}{P(\boldsymbol{a}\in\boldsymbol{O}_A)}=\frac{P(B)}{P(A)}$$

即 $P(B)=S(A,B)P(A)$。

推论　若服务 $Q_1,Q_2,\cdots,Q_n$ 依次顺连,各自连接支持度为 $S(Q_1,Q_2)=S_1,\cdots,S(Q_{n-1},Q_n)=S_{n-1}$,则有

$$P(Q_i)=S(Q_{i-1},Q_i)\cdot P(Q_{i-1})=P(Q_1)\cdot\prod_{j=1}^{j=i-1}S_j \tag{6}$$

(2) 分连

分连结构如图 2(b)所示,服务 B 和 C 分别只接受 A 的输出,且 AB、AC 之间为饱和连接。当已知服务 A 的可运行

概率时，服务 B 和 C 的可运行概率可由定理 2 分别计算，即

$$P(B) = S(A,B)P(A)$$
$$P(C) = S(A,C)P(A) \quad (7)$$

(3) 汇连

汇连结构如图 2(c)所示，服务 A 必须同时接受服务 B 和 C 的输出才可运行。当形成汇连组合时，服务 B 和 C 各自独立不重合地将数据传入 A 中。

定理 3 在服务组合层次结构中，服务 A 的输入由服务 B 和 C 的输出给出，且服务 B 和 C 到 A 的汇连是饱和连接，连接支持度分别为 $S(B,A)$ 和 $S(C,A)$，已知服务 B 和 C 的可运行概率分别为 $P(B)$ 和 $P(C)$，则服务 A 的可运行概率为 $P(A)=S(B,A)P(B)\cdot S(C,A)P(C)$。

证明 由于服务 A 的输入来自于服务 B 和 C 的输出，则 $P(A)=P(\boldsymbol{a}\in \boldsymbol{I}_A, \boldsymbol{b}\in \boldsymbol{O}_B, \boldsymbol{c}\in \boldsymbol{O}_C)$，其中 $\boldsymbol{b}$ 和 $\boldsymbol{c}$ 为服务 B 和 C 的输出，$\boldsymbol{a}$ 为 $\boldsymbol{b}$ 和 $\boldsymbol{c}$ 的联合，考虑到概念独立性假设，有

$$P(\boldsymbol{a}\in \boldsymbol{I}_A \mid \boldsymbol{b}\in \boldsymbol{O}_B \cap \boldsymbol{c}\in \boldsymbol{O}_C) =$$
$$\frac{P(\boldsymbol{a}\in \boldsymbol{I}_A \cap \boldsymbol{b}\in \boldsymbol{O}_B \cap \boldsymbol{c}\in \boldsymbol{O}_C)}{P(\boldsymbol{b}\in \boldsymbol{O}_B \cap \boldsymbol{c}\in \boldsymbol{O}_C)} = \frac{P(A)}{P(B)P(C)} =$$
$$\frac{P(\boldsymbol{b}\in \boldsymbol{I}_A \cap \boldsymbol{c}\in \boldsymbol{I}_A \cap \boldsymbol{b}\in \boldsymbol{O}_B \cap \boldsymbol{c}\in \boldsymbol{O}_C)}{P(\boldsymbol{b}\in \boldsymbol{O}_B \cap \boldsymbol{c}\in \boldsymbol{O}_C)} = S(B,A)S(C,A)$$

因此 $P(A)=P(A|B)P(B)P(A|C)P(C)=S(B,A)\cdot P(B)S(C,A)P(C)$。 证毕

3.2 混合连接

复合服务的组合层次结构由上述三种基本连接样式混合而成，复合服务有且仅有唯一的初始服务和目标服务，初始服务的可运行概率必然为 1。根据基本连接样式计算公式，从初始服务开始按照宽度优先搜索算法计算后续服务的可运行概率，直至得出目标服务的可运行概率 $P(\top)$，该值便是复合服务支持度。同样，从目标服务开始，依次计算前驱服务的可运行概率，直至计算出初始服务的可运行概率 $P(\perp)$，该值便是复合服务置信度的大小。

4 算法实验与结果分析

本文提出的方法解决了服务连接的接口匹配问题和连接量化问题，鉴于目前没有相关的标准平台和标准测试数据，此处参考文献[4]采用随机生成的模拟 Web 服务数据作为测试用例，在不同数量的概念集中，分别生成若干组服务，分析两两服务之间的连接程度，对算法的时间性能进行评估。实验结果如图 3 所示。

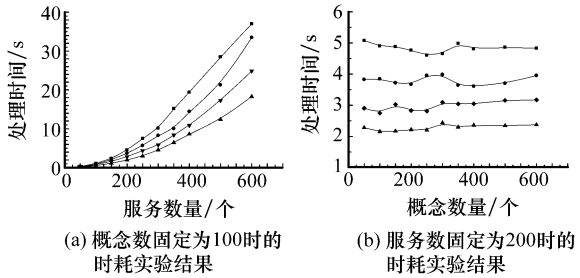

图 3 服务连接度时耗实验结果

其中，图 3(a)为概念数为 100 且服务接口数分别为 5、6、7、8 时，不同的服务数所需要的处理时间；图 3(b)为服务数为 200 且服务接口数分别为 5、6、7、8 时，不同的领域概念数所需要的处理时间。结合实验结果和程序分析发现，由于需要分析两两服务之间的连接关系，服务连接分析需要的时间与服务数、各个服务的接口数成非线性增长，而概念数影响了接口支持度和置信度的计算，程序实现中通过服务接口描述中的概念索引值直接获取相关概念及其公共子概念的出现概率，因而领域本体中的概念数对处理时间影响不大。由此可见，当服务数很多时，服务连接分析时耗很大，因此在服务组合前，可预先建立服务连接索引库，一旦新增加服务则立即更新与之相关的服务索引，这样可大大提高组合的响应时间。

5 结束语

由于服务组合涉及的问题很多，自动组合难度很大。在实际的服务组合操作过程中，当用户选择某个服务时，系统根据服务连接索引矩阵，找到可连接的前驱和后继服务，并根据连接度大小进行排序，供用户参考，提高了组合的工作效率。如何正确合理地给出服务连接关系，是实现服务组合的首要问题，为服务的人工和自动组合提供了重要的参考依据。本文根据服务输入输出接口语义，给出服务连接度的量化和接口分量对应关系判定算法，并讨论各种不同的连接方式下服务连接的度量。本文工作意义主要体现在：提出了一种新的服务接口关联程度量化指标，改进了父子关系推理方法的粗略性，且与基于语义相似度的方法相比其数学意义更明确。基于此可以深入研究服务规划组装过程，并可通过控制服务连接度阈值保证服务组合质量。另外，如何在输入输出数据流的基础上，加入服务前提和效果的控制流，从而更为精确地分析评价服务连接，将是值得深入研究的工作。

参考文献：

[1] Massimo Paolucci, Takahiro Kawamura, Terry R Payne, et al. Semantic matching of Web services capabilities[C]// *The 1st International Semantic Web Conference*, 2002:333-347.

[2] Li Lei, Horrocks Ian. A software framework for matchmaking based on semantic web technology[C]// *Proc. of the 12th International Conference on World Wide Web*, 2003:331-339.

[3] Freddy Lécué, Alain Léger. A formal model for semantic Web service composition[C]// *The 5th International Semantic Web Conference*, 2006:385-398.

[4] 李曼，王大治，杜小勇，等. 基于领域本体的 Web 服务动态组合[J]. 计算机学报，2005,28(4):644-650.

[5] 艾未华，黄敬平，周宁，等. 一种基于语义本体的 Web 服务自动组合算法[J]. 系统仿真学报，2008,20(4):935-937.

[6] 于守健，何丰，乐嘉锦. 基于接口匹配的 Web 服务自动组合[J]. 计算机科学，2007,34(3):61-68.

[7] Judea Pearl. Probabilistic reasoning in intelligent systems: networks of plausible inference[M]. *San Mateo: Morgan Kaufmann*, 1988:241-288.

[8] Fred S Roberts, Barry Tesman. 应用组合数学[M]. 冯速，译. 北京：北京机械工业出版社，2007.

[9] Princeton University Cognitive Science Laboratory. WordNet: a lexical database for the English language[EB/OL]. [2008-06-03]. *http://wordnet.princeton.edu*.

[10] Pedersen T. Wordnet::Similarity[EB/OL]. [2008-06-03]. *http://wn-similarity.sourceforge.net*.