

基于形式概念分析的云服务发现方法

杨晋青

(大连海事大学 信息科学技术学院, 辽宁 大连 116026)

摘要 云计算在学术界和工业界都引起了极大的兴趣。在云计算的环境中也已经开发出了越来越多的云服务。这就使得如何从大量的云服务中有效的发现合适的云服务成为了一项具有挑战性的任务。在这篇论文中, 我们提出一种基于形式概念分析的新的云服务发现方法。首先, 通过可用性参数和语义距离相似性来过滤不相关的服务, 然后使用概念格存储这一实际问题, 并使用形式概念分析来处理非功能服务质量属性, 再根据它们的服务质量级别来对这些服务进行分类, 并为用户选择合适的服务。文中使用 OWLS-TC 来验证这一服务发现方法, 结果表明这一方法不仅能够为用户提供合适的服务, 而且十分有效。

关键词 形式概念分析; 概念格; 云服务发现; 服务质量;
中图法分类号 TP369 文献标识码 A

A Cloud Service Discovery Approach Based on FCA

Yang Jin-qing

(College of Information Science and Technology, Dalian Maritime University, Dalian 116026)

Abstract Cloud computing has attracted great interest from both academic and industrial communities. More and more cloud services have been developed in the cloud computing environments. But it is a challenging task to discovery the suitable service from large number of services efficiently. A new service discovery approach based on Formal Concept Analysis (FCA) is presented in this paper. First, filter irrelevant services by availability parameter and semantic distance similarity, then use concept lattice to store the problem and handle the non-functional QoS attributes with FCA, classify services by their QoS levels and select suitable service for user. We use OWLS-TC to validate this service discovery approach and the results show that the approach not only provide suitable service for user, but also very efficient.

Key words Formal Concept Analysis; Concept Lattice; Cloud Service Discovery; Quality of service

1 引言

云计算是基于互联网的计算机技术, 是将互联网中的可动态扩展和虚拟化的资源作为一种服务来提供的技术[1]。随着云计算的快速发展, 实现同样功能的不同的云服务数量也在不断上涨。所以如何选取一个满足用户需求的服务成为了研究的重点。服务发现的主要过程分为两步: 一是根据用户功能需

求, 例如类别, 投入, 产出等来通过功能性过滤服务, 这样会有一些候选服务的功能相似; 服务质量[2]描述了服务的非功能属性, 例如成本, 带宽, 可用性等, 为了比较功能相似的服务并从其中进行挑选, 基于服务质量这一属性的服务选择方法也在很多文章中有所介绍。例如, 一些文章倾向于基于服务质量的服务发现模型构建[3], 一些文章学习如何对服务进行聚类, 还有一些文章[4-5]设计算法来选择或组合服务等。

收稿日期: 2016-11-4; 修订日期: 2016-11-15。

基金项目: 国家自然科学基金项目 (60972014)。

杨晋青 (1995—), 女, 辽宁大连人, 硕士研究生, 研究方向为智能信息处理。

形式概念分析提供了一种构建、分析和虚拟化数据的概念型框架。为了使数据易于理解，形式概念分析以格理论为基础。形式概念分析中的应用程序域根据概念格进行组织和结构化。概念是共享最大属性集合最大实体集合。概念格中的每一个概念都继承它祖先的属性，并且具有比它的后代要少的属性。

在这篇论文汇总，提出了一个基于形式概念分析的服务发现方法来寻找合适的服务。这种方法利用语义匹配来根据用户的需求过滤服务，尤其是当不相关域的服务数量较大时。使用概念格来根据服务的服务质量级别来对服务进行分类，生成的概念格有助于识别与用户指定的服务质量要求相匹配的最合适的服务。

2 基本概念

2.1 形式概念分析

形式概念分析为结构化、分析和虚拟化数据提供了一种概念型的框架，这里介绍的服务选择方法也是基于形式概念分析的。

表 1 对象集 O 和属性集 A 的形式背景

	a	b	c	d	e	f	g	h	i
Leech	x	x					x		
Fish	x	x					x	x	
Frog	x	x	x				x	x	
Dog	x		x				x	x	x
Waterweeds	x	x		x		x			
Reed	x	x	x	x		x			
Bean	x		x	x	x				
Corn	x		x	x		x			

概念就在形式概念分析中的一个形式背景中定义的，形式背景就是一个三元组 $F=(G, M, I)$ ，其中 G 是对象集， M 是属性集， I 是对象集 G 与属性集 M 之间的关系。举例说明，如果存在 $g \in G, m \in M, gIm$ 表示对象 g 具有属性 m ，也表示属性 m 对应于对象 g 。为了能够利用形式概念分析来分析数据，数据集总是表示成一个 n 行 m 列的表格（即上文所示的形式背景），对象作为行元素，属性作为列元素。如果对象 g 具

有属性 m ，则在表格中的一个“x”表示 gIm 。

2.2 概念格

定义：设 H 是一个集合，如果存在 H 上的一个关系 R ，对于 $\forall x, y, z \in H$ 满足如下条件：（1）非完全性： $\exists x$ 与 y 不可比较；（2）自反性： xRx ；（3）反对称性： $xRy, yRx \Rightarrow x=y$ ；（4）传递性： $xRy, yRz \Rightarrow xRz$ 。则称 R 是 H 上的一个偏序关系，表示为“ \leq ”，具有这种偏序关系的集合称为偏序集，表示为 $\langle H, \leq \rangle$ 。偏序集 (H, \leq) 加上它所具有的属性构成一种新的数据结构概念格[6]。其中元素为概念的完备格，如若 $(A1, B1)$ ， $(A2, B2)$ 是某个背景 (G, M, I) 上的两个概念，而且 $A1A2$ （也等价于 $B1B2$ ），则我们称 $(A1, B1)$ 是 $(A2, B2)$ 的子概念， $(A2, B2)$ 也是 $(A1, B1)$ 的超概念。并记作 $(A1, B1) \leq (A2, B2)$ ，关系成为概念的层次序，简称为序。 (G, M, I) 的所有概念用这种序组成的集合可以用 $B(G, M, I)$ 表示，称它为背景 (G, M, I) 上的概念格。

形式概念分析从一个形式背景中提取所有形式概念，并揭示所提取概念中的继承关系（超概念和子概念），然后将它们组织进概念格。对应的概念格如图 1。

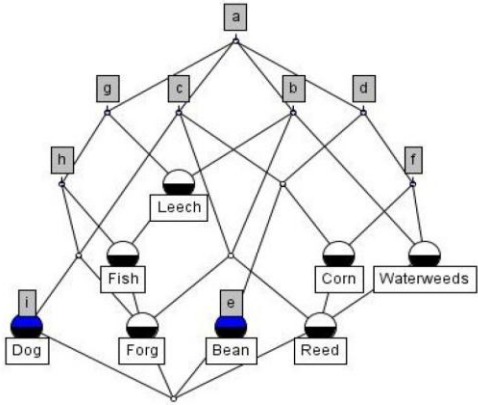


图 1 表 1 所示形式背景对应的概念格

2.3 QoS 无因次化

服务质量是几种服务品质或属性的结合。服务质量这一属性有许多特性，如一些是积极的，而其他的不是；服务质量的值是连续的。为了简化利用 QoS 进行服务选择，并且使复杂的数据建模问题合理化，这里使

用图 2 中的公式来将 QoS 的值进行无因次化。

$$q_v = \begin{cases} \frac{q_v - q_{\min}}{q_{\max} - q_{\min}} & q \text{ is positive attributes} \\ \frac{q_v - q_{\max}}{q_{\min} - q_{\max}} & q \text{ is negative attributes} \end{cases}$$

图 2 QoS 值无因次化公式

这里利用一种名为 BoxPlot[7]的统计技术将收敛值聚集在一起,以便通过形式概念分析构建服务的概念格。它采用给定的一组数值作为输入,并产生一到五个对应的值级别: $L=\{\text{VeryBad}, \text{Bad}, \text{Medium}, \text{Good}, \text{VeryGood}\}$ 这个方法应用于每个 QoS 属性。

3 方法

所用方法分为两个步骤: (a) 在服务选择之前的语义功能匹配; (b) 基于形式概念分析的分类技术对概念格中的候选服务进行分类。最后一步使得能够以更少的时间和精力最佳地匹配指定的 QoS 要求的服务。

3.1 功能匹配

在这部分,将介绍一种功能匹配的方法。这种方法可用于筛选不符合用户需求的服务。输入、输出、前提条件和结果形成了服务的功能信息,因此功能匹配主要是 IOPE 匹配。功能匹配的算法如图 3 所示。

Algorithm 1. Functionality Matchmaking Algorithm
Input: candidate service set(CS), service request(R)
Output: select service set(SS)
1: for each s in CS
2: if |R.inputs|!=0
3: uI=availParam(s.inputs, R.inputs);
4: end if
5: if |R.outputs|!=0
6: uO=availParam(s.outputs, R.outputs);
7: end if
8: if |R.precondition|!=0
9: uP=availParam(s.precondition, R.precondition);
10: end if
11: if |R.result|!=0
12: uR=availParam(s.result, R.result);
13: end if
14: if uI==1 & uO==1 & uP==1 & uR==1
15: Calculate s and R input, output, precondition, result similarity by semantic distance, respectively, and sum them as Sum
16: if Sum>= the Threshold
17: let s into SS;
18: end if
19: end if

图 3 功能匹配算法

算法 1 是功能匹配的算法。首先,它判断候选服务 s 是否需要计算相似性,如果它不满足可用性,跳转到 CS 中的下一项服务。然后,计算 s 和用户请求的语义的相似度,如果候选服务的相似度大于阈值,让它进入 SS。算法 1 保证了精度,同时也通过可用性参数提高了功能匹配的效率。

3.2 基于形式概念分析的服务选择

在这部分,为找到服务选择的近最优解决方案而设计了一种基于形式概念分析的服务选择算法。这一算法的中心思想是构建候选服务的概念格,在概念中,候选服务和用户 QoS 需求作为对象, QoS 等级作为属性,通过这种方式查找哪一种候选服务在概念格中最接近用户需求。这里将服务选择分为四种情况讨论。

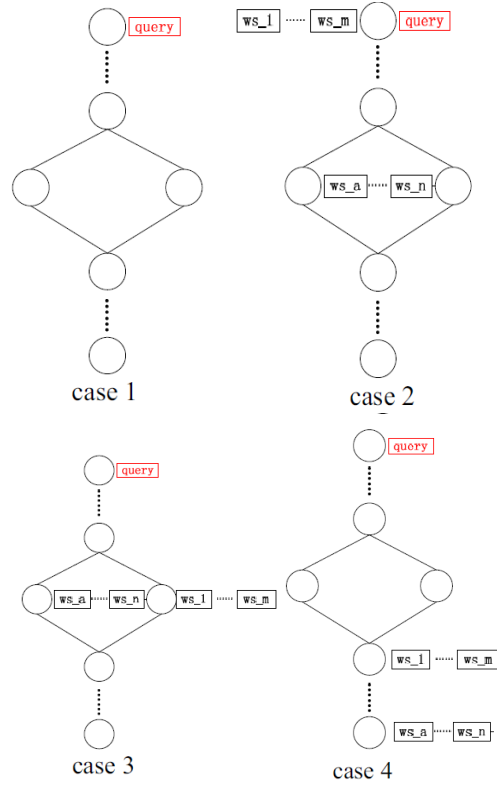


图 4 服务选择情况

情况 1 中在所查询概念下无匹配服务,向用户返回空值;情况 2 中概念格中有一些符合查询概念的候选服务,且 QoS 属性是无因次化的,所以子概念意味着更好的 QoS。为了找到符合用户要求的服务,如果在查询概念下有一些服务,基于形式概念分析的服务选择将选择这些服务中功能匹配得分最高的服务。在情况 3 中,因为没有符

合查询概念的服务且只有一些服务符合查询概念的子概念，所以给予形式概念分析的服务选择法将如同情况 2，选择功能匹配得分最高的服务。在情况 4 中，基于形式概念分析的服务选择首先找到底部的候选服务，然后从中选择功能匹配得分最高的候选服务。给予形式概念分析的服务选择算法如图 5 所示。

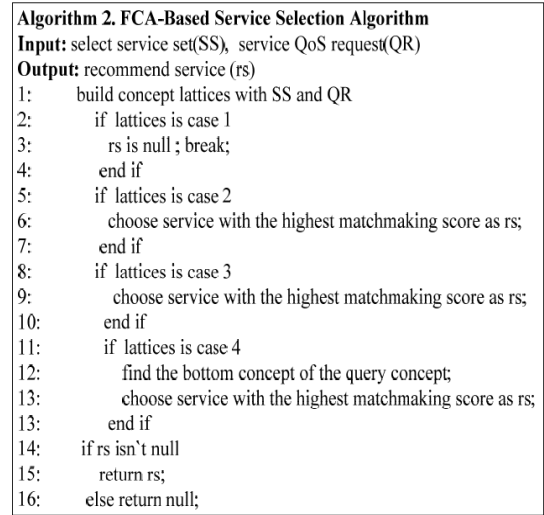


图 5 基于形式概念分析的服务选择算法

4 实验和评估

在这部分，将通过一个实验来评估服务发现算法。实验基于 OWLS-TC，选择价格(P)和反应时间(RT)作为实验 QoS 属性。实验中我们通过一些实验者测试了基于形式概念分析的服务选择算法，以验证基于形式概念分析算法的服务选择满意度，结果如表 2 所示。

表 2 用户满意度数据

	vol A	vol B	vol C	Avg
eduction	0.92	0.91	0.93	0.920
medical care	0.87	0.88	0.86	0.870
food	0.94	0.97	0.92	0.943
travel	0.95	0.96	0.91	0.940
communication	0.97	0.92	0.93	0.940
economy	0.95	0.97	0.96	0.960
weapon	0.82	0.79	0.83	0.813
geography	0.91	0.92	0.92	0.917
simulation	0.82	0.85	0.87	0.847
Avg	0.906	0.908	0.903	

三位志愿者的满意度数据表明，基于形式概念分析的服务选择算法总能选出最适

合用户的服务。

5 结论

在这篇论文中，我们提出了基于形式概念分析的云服务发现方法。这种方法共两个主要步骤：功能匹配和通过可用性参数和语义距离相似性过滤不相关的服务，最后基于形式概念分析来选择服务。我们利用 OWLS-TC 验证了这一方法，并证实我们的方法不仅能够满足用户的特定需求，而且能够提高为用户选择服务的效率。

参 考 文 献

[1] M. Bohm, S. Leimeiste, C. Riedl, H. Krcmar, in Cloud Computing Technologies Business Models Opportunities and Challenges, S. Murugesan, Ed.

[2] Menasce, D.A. QoS issues in Web services. IEEE Internet Computing 6, 72-75.

[3] X. Wang, T. Vitvar, M. Kerrigan, I. Toma, A. Dan, W.Lamersdorf, Eds. Exchange Organizational Behavior Teaching Journal 4294, 390-401.

[4] Rao, J. & Su, X. A Survey of Automated Web Service Composition Methods. Semantic Web Services and Web Process Composition 3387, 43-54.

[5] D Mello, D.A. & Ananthanarayana, V. A review of Quality of Service (QoS) driven dynamic Web service selection techniques. 2010 5th International Conference on Industrial and Information Systems 201-206 (2010).

[6] Azmeh, Z., Huchard, M. & Tibermacine, C. Using Concept Lattices to Support Web Service Compositions with Backup Services. 2010 Fifth International Conference on Internet and Web Applications and Services 363-368 (2010).

[7] Royeen, C.B. The boxplot: a screening test for research data. The American journal of occupational therapy official publication of the American Occupational Therapy Association 40, 569-571 (1986).