

基于功能分组的群体 Web 服务 QoS 动态计算研究

王冠^{1,2}, 马炳先^{1,2}, 相东明^{1,2}

(1. 山东网络环境智能计算技术重点实验室, 山东 济南 250022;

2. 济南大学 信息科学与工程学院, 山东 济南 250022)

摘要:随着服务计算技术的发展及应用,在服务实时组合过程中,服务质量(Quality of Service, QoS)已经成为选择服务的重要依据之一。当前较多的 QoS 动态计算方法提出了多种方案用以解决 QoS 的动态计算问题,但都将 Web 服务视为单一独立个体,首先通过获取独立的参数来测算每个 Web 服务的 QoS,然后在服务组合过程中按照功能需求从具有相同功能的待选服务集合中选取 QoS 最好的服务。为了避免不同服务的功能差异对计算 QoS 的影响,突出相同功能服务各 QoS 因素之间的关联性对计算 QoS 的影响,从而在相同功能的服务中选取相对质量较好的服务,文中提出将 Web 服务群体按照功能分组,在组内实时、动态地计算各服务 QoS 的方法,实验表明,文中方法计算得到的 QoS 具有较好的可靠性和较高的可信度。

关键词: Web 服务;服务质量;群体 Web 服务;功能分组

中图分类号: TP31

文献标识码: A

文章编号: 1673-629X(2013)01-0029-05

doi: 10.3969/j.issn.1673-629X.2013.01.008

Study on Dynamic Calculation Method for Web Service QoS Based on Grouped Function

WANG Guan^{1,2}, MA Bing-xian^{1,2}, XIANG Dong-ming^{1,2}

(1. Shandong Provincial Key Laboratory of Network Based Intelligent Computing, Jinan 250022, China;

2. School of Information Science and Engineering, University of Jinan, Jinan 250022, China)

Abstract: With the developing and application of service computing technology, Quality of Service (QoS) has been becoming a critical factor for services selecting within service composition. Many methods have been put forward to calculate dynamic QoS, while each Web service is regarded as a single individual, independent parameters are obtained to calculate QoS of it, and then the best service is selected based on QoS within service composition. To avoid affection among different functional services on QoS calculation, highlight the correlation among QoS parameters of services that having same functions, and select out the Web service with relative better QoS, in this paper, Web services are grouped according to their functions and the dynamic calculation method of QoS is also proposed for Web services within groups. Experiments show that QoS obtained with the method of it has high credibility and reliability.

Key words: Web service; quality of service; Web services group; grouped function

0 引言

近年来,随着互联网的迅猛发展,Web 服务(Web Service)^[1]以其开放、低耦合、平台无关等特点已成为一种新兴的互联网应用模式与分布式计算模型,得到

人们的广泛关注。Web 服务以 WSDL(Web Services Description Language)^[2]、SOAP(Simple Object Access Protocol)^[3]等平台无关的网络通信协议为用户提供功能服务,构成了未来大规模面向服务的分布式计算基础。

互联网上的服务提供商(Service Provider)众多,可以提供很多相似或相同功能的服务。单个 Web 服务往往只能提供单一、有限的功能,不能满足复杂的用户需求,因此 Web 服务组合技术便应运而生。在 Web 服务组合中,如何进行有效的服务选择一直是学术界和产业界广泛关注的问题之一。由于用户在面对广泛

收稿日期:2012-05-15;修回日期:2012-08-19

基金项目:国家自然科学基金资助项目(60903099);山东省优秀中青年科学家奖励基金(BS2009DX012);山东省高等学校科技计划项目(J09LG14)

作者简介:王冠(1986-),男,硕士,CCF 会员,研究方向为服务计算;马炳先,博士,副教授,硕士生导师,研究方向为 Petri 网、服务计算。

的、服务信息不完全的候选服务的情况下进行服务选择是普遍存在的^[4],因此,当用户进行 Web 服务选择的时候,不应仅仅考虑待选择服务所满足的功能性需求,还应同时考虑待选择服务所能满足的非功能性需求,即服务质量(Quality of Service, QoS)。当用户面对众多功能相似或相同的待选服务时,应同时考虑待选 Web 服务的服务质量,将服务质量纳入服务的查找、选择与组合,对于成功构建基于 Web 服务的应用系统及其质量保障具有重要意义^[5~8]。

Web 服务组合中,动态服务组合策略是一种在系统运行期进行的服务组合方法,根据用户需求实时地组合 Web 服务,并及时地判定服务组合质量。使用动态服务组合策略的系统往往是一些提供复杂动态服务功能的系统,这种系统需要实时根据功能需求及各个 Web 服务的状态从众多 Web 服务中选择一些服务进行服务组合^[9,10]。许多学者将动态服务组合的研究重点放在语义、组合策略和 QoS 上^[11,12]。

目前,对于 Web 服务的 QoS 动态计算,通常以 Web 服务的历史 QoS 算术平均值来近似服务的 QoS^[7],并假定在服务的选择与评价过程中 QoS 不变。事实上,由于 Web 服务运行环境的变化及任务类型、任务粒度的不同,使 Web 服务的 QoS 具有很强的动态性。文献[13]提出了一种基于用户之间及服务之间的相似度,结合用户的历史经验对用户未使用过的服务的 QoS 进行预测。文献[14]提出了基于免疫系统的多信号机制实时计算 QoS 的模型,该计算模型只给出了服务信誉度的计算方法,没有给出其他重要 QoS 指标的计算方法。上述工作促进了 Web 服务 QoS 动态计算的研究,但将所有的已知服务视为单一集合,再计算每个 Web 服务的 QoS,并没有对是否提供相似或相同功能的服务明确加以区分。

Web 服务的功能多样,服务组合过程是一个根据功能需求,从待选服务集合中选择 QoS 较好服务的过程,待选集合往往是满足某一功能需求的原子服务集合。影响服务质量的因素中,有许多因素在相同功能的服务之间具有关联性、相对性。

服务组合过程中选取 Web 服务是基于服务功能的,因此 Web 服务是功能相关的。QoS 描述的是 Web 服务的质量,因此 QoS 间接反应了对服务功能的相关评价,QoS 与服务功能是相互关联的。传统的 QoS 测算方法,将原子服务视为单一个体,获得原子服务的独立 QoS,在接下来的服务组合过程中,再依照功能将服务分组,然后再选择 QoS 较好的服务,用于构建基于服务的应用系统。这种方法人为地将 QoS 的计算过程与服务功能分组过程分离,由此计算得到的 QoS 已经不再包含服务功能信息,不能反映功能组状态。

假定 QoS 由原子服务性能和原子服务可靠性复合而成。在服务组合过程中,希望能够根据 QoS 选择到能满足功能需求的最优原子服务,而最优的原子服务往往满足以下两种情况:

(1)当提供相同功能的原子服务较多时,原子服务的选择余地较大,QoS 评分应该偏向服务性能;

(2)当提供同功能的原子服务较少时,原子服务的选择余地较小,一旦被选择的原子服务不能正常提供功能时,可替换的原子服务数目较少,则 QoS 评分会偏向服务的可靠性。如果将原子服务视为单一个体测算 QoS,得到的 QoS 将无法反映相同功能的服务数量信息,因而无法根据相关功能的服务数量来动态地调整服务性能和可靠性在 QoS 计算结果中所占的比例,最终依照 QoS 进行服务选择就难以得到最优服务。

综上所述,只有在测算原子服务 QoS 之前,先按照功能将原子服务分组,然后在组内获取各个原子服务间的相对参数,最终给出各个服务在组内的 QoS 评分,才能准确有效地反映服务的真实质量,才能真正为服务组合选择出服务质量较高的服务。

1 Web 服务组的动态服务 QoS 计算

Web 服务组合过程,针对某一功能,需要从待选服务集合中选取一个 QoS 较好的 Web 服务参与服务组合。因此,需要按服务提供的功能对 Web 服务分组,即功能相同或相似的服务被分为一组。对 Web 服务分组后,单个分组内的样本数减少,有利于提高 QoS 的计算效率。此外,对 Web 服务分组后,不同组间的 Web 服务不会相互影响,有利于提高单一功能服务相应的 QoS 的稳定性,便于进行 Web 服务组合。因此,将具有相同功能的服务归为一组,称之为 Web 服务组。

影响 QoS 的因素有多种,大致可以分为客观因素和主观因素两种。客观因素是指与 Web 服务功能相关的因素,是决定 Web 服务质量的本质的因素,是内因,如服务的运行效率、服务的可靠性等。主观因素是指与 Web 服务功能不相关,却又能影响 Web 服务质量评分的因素,是外因,如 Web 服务的价格(一般由服务提供商给出,有严格的主观性)、Web 服务提供商给出的服务质量评分等。Web 服务的评分应能综合反映主观和客观因素,但应以客观因素为主体,辅以主观因素。由于客观因素的客观性,故应可以在任何一个 Web 服务应用系统上实时、客观地测量。我们选取原子服务的性能和可靠性作为主要的客观因素来评价原子服务。

在 Web 服务信息不完全的情况下,计算 QoS 的常

用方法是使用 Web 服务的历史 QoS 的算术平均值来近似该 Web 服务的 QoS^[15]。这种方法虽然实现简单,但并没有对 Web 服务的功能加以区分,导致其并不能动态有效地反映 Web 服务的 QoS。

在评价 Web 服务组内原子服务时,组内单原子服务的可靠性对其 QoS 评分的影响力应由 Web 服务组的规模和组内所有原子服务的可靠性共同决定,组内服务普遍可靠且组的规模较大时,单个服务的可靠性的重要性就会随之降低,可靠性对其 QoS 评分的影响力随之降低;反之,个体服务的可靠性的重要性就要更加明显。原子服务的性能也是个相对概念,是对组内各服务之间相对性能的衡量,服务组合是功能的组合,单独考察单一原子服务的性能没有实际意义,要将服务放在服务组内,考察原子服务在组内相对于组内其他原子服务的性能。因为,用户选择服务时首先要考虑的是服务的可用性,即可靠性,脱离了服务的可靠性来谈论服务的性能是没有意义的,所以,组内单原子服务的性能对其 QoS 评分的影响力还应该结合服务的可靠性。用于计算 Web 服务组内原子服务 $i(0 \leq i \leq n)$ 的 QoS_i 的公式,如下所示:

$$QoS_i = \frac{t_i}{\sum_{j=1}^n t_j} \times \left(1 - \frac{1}{n}\right) \times \frac{\sum_{j=1}^n S_j}{\sum_{j=1}^n C_j} + \left(1 - \left(1 - \frac{1}{n}\right) \times \frac{\sum_{j=1}^n S_j}{\sum_{j=1}^n C_j}\right) \times \frac{S_i}{C_i} \quad (1)$$

其中, t_i 为 Web 服务组内第 i 个服务的累计调用总时间; n 为 Web 服务组规模, $n \geq 1$; S_i 为 Web 服务组内第 i 个服务的成功调用次数; C_i 为 Web 服务组内第 i 个服务的总调用次数。

在基于服务计算技术的应用系统中,通过实时采集式(1)中涉及的多种参数,如服务执行时间,服务调用次数等,根据式(1)便可动态计算得到 Web 服务组内各个原子服务的 QoS。

2 实验分析

文中将通过四组实验分别说明 Web 服务组内原子服务的 QoS 计算公式在多种情况下的表现。

2.1 实验一

实验方案:本实验中,我们定义了一个服务组,称之为 B_{11} ,并向其中注册一个(提供三个整数相加功能的)原子服务,设为“原子服务 X”,后称之为“X”。X 每次执行时间会增加一个常量。对 X 调用 25 次(调用序号从 0 至 24),并记录每次调用后 X 的 QoS,如图 1 所示。

结果分析:通过图 1 可以看出,随着调用次数的增加,X 的 QoS 先随之上升,后稳定于 1.0。本组实验表明,当服务组内只有一个原子服务时,使用式(1)计算得到的原子服务 QoS 并不会受到服务性能的影响。这与“在服务使用者面对待选择服务组内只有一个原子服务时,如果服务使用者希望使用 Web 服务组内原子服务提供的功能,则只有唯一选择”的实际情况相吻合。同时证明了,使用式(1)计算得到的原子服务 QoS 具有稳定性。

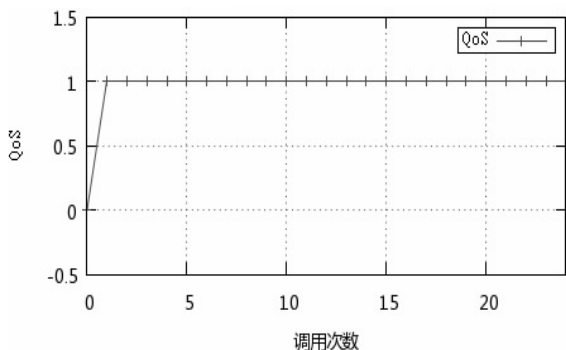


图 1 实验一

2.2 实验二

实验方案:本实验中,定义了服务组,称之为 B_{12} ,并向其中注册 20 个(提供三个整数相加功能的)原子服务(序号从 0 至 19)。对 B_{12} 做 20 次迭代调用,每次迭代过程中,群内调用失败的原子服务的数量会以每次 5% 的幅度增加,如图 2(a) 所示。同时,在每次迭代过程中记录序号为 0 的原子服务的 QoS,结果如图 2(b) 所示。

结果分析:通过图 2(b) 可以看出,使用式(1)计算得到的组内原子服务的 QoS 会随着组内产生失败调用的原子服务所占比例的增加而降低。本组实验表明,随着组内不可靠原子服务数的增加,在外部看来,整个服务组都会变得不可靠。由此可知,文中提出的式(1)能正确反映,Web 服务组内原子服务普遍不可靠且群体规模较大时,组内原子服务本身的可靠性会变得尤为重要,对原子服务的 QoS 影响力度增加,而性能指标对整个 QoS 的影响力就会降低,如图 2(c) 所示。

2.3 实验三

实验方案:本实验中,定义了一个服务组,称之为 B_{13} 。首先向 B_{13} 注册一个(提供三个整数相加功能的)原子服务,并成功调用一次。随后,向 B_{13} 继续注册 29 个(提供三个整数相加功能的)原子服务,对每个新加入 B_{13} 的原子服务做一次成功调用,并在每次注册并成功调用原子服务后测算当前组内所有原子服务的可靠性贡献度(可靠性因子对 QoS 结果的影响力),结果如图 3 所示。

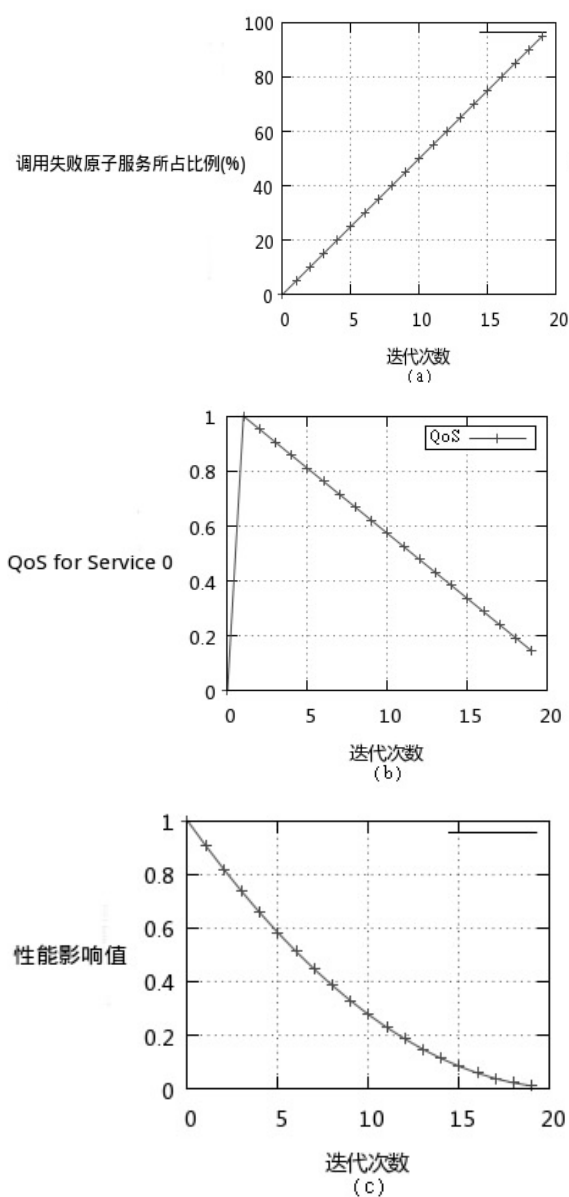


图 2 实验二

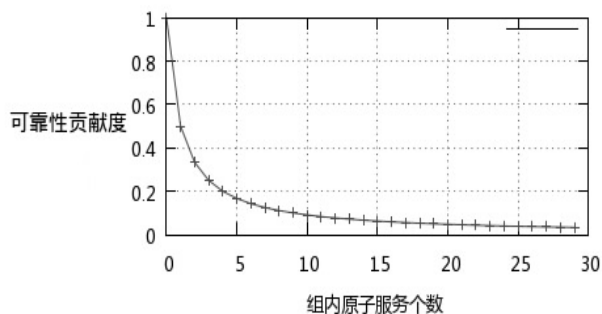


图 3 实验三

结果分析:通过图 3 可以看出,随着 B_{13} 内原子服务数量的增加,可靠性贡献度随之降低。本组实验结果表明,式(1)能正确反映:如果 Web 服务组内原子服务普遍可靠且群体规模增大,那么原子服务的可靠性对 QoS 计算的贡献度就减小。根据式(1)计算得到的组内 QoS 满足,在服务组合过程实践中,提供同一功能

的服务越多,选择空间就越大,小部分服务可靠性降低对功能组产生的影响越小,组内 QoS 变动越小。

2.4 实验四

实验方案:本实验中,定义了一个服务组,称之为 B_{14} 。首先向 B_{14} 注册 30 个(提供三个整数相加功能的)原子服务,并对每个原子服务调用 5 次,并记录每个原子服务的 QoS 得分。然后对前 20% 的原子服务进行失败调用,并记录每个原子服务的 QoS 得分。

本实验中,群内基于平均值的 QoS 通过式(2)可获得:

$$QoS_{avg} = \frac{P + R}{2} \quad (2)$$

其中, P 是 Web 服务组内原子服务的平均性能, R 是 Web 服务组内原子服务的平均可靠性。

基于文中所提出的群内服务 QoS 动态计算式(1)所得到的 QoS 评分如图 4 曲线一、二所示,基于式(2)得到的 QoS 评分如图 4 曲线三、四所示。

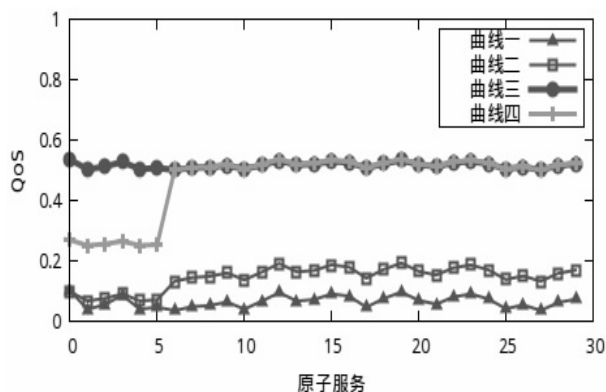


图 4 实验四

结果分析:曲线一和曲线二,反映了随着原子服务调用次数的增加,使用式(1)计算得到的原子服务 QoS 会整体提高,但因为一些失败调用,导致前 20% 的原子服务的 QoS 评分提升速度减慢。曲线三和曲线四反映了,基于均值的计算式(2)计算得出的 QoS 在发生一系列失败调用后,对应原子服务的 QoS 会急速下降。

基于以上对比,说明基于均值的 QoS 算法对于调用失败的原子服务过于敏感。使用式(1)计算所得到的 QoS 结果如曲线一、二所示,反映了随着调用次数的增加,Web 服务组内所有服务平均可靠性有所提升, QoS 评分更可靠,而基于均值的计算式(2)计算得出的 QoS 如曲线三、四所示,没有反映这一变化。

本实验表明,在发生服务调用失败的情况下,文中提出的组内 QoS 计算方法比基于均值的计算方法所得到的 QoS 更可靠。

3 结束语

文中提出了服务群体按功能分组的概念,将 QoS

评分放在 Web 服务组内完成,并区分了影响服务质量的客观因素和主观因素,并根据客观影响因素,提出了 Web 服务组内原子服务的动态 QoS 计算方法。实验证明,组内 QoS 动态计算方法能够实时有效地获得组内原子服务的 QoS,避免了不同功能组内原子服务状态的变化对其他功能组的影响。下一步,将尝试把主观影响因素整合到现有的动态 QoS 计算公式内,进一步完善 QoS 的度量方法,建立服务质量的度量指标体系以及服务组合质量的动态保障机制。

参考文献:

[1] Web Services Activity[EB/OL]. 2002. <http://www.w3.org/2002/ws>.

[2] WSDL. Web Services Description Language. 1. 1. [EB/OL]. 2001-03. <http://www.w3.org/TR/wsdl>.

[3] W3C. SOAP Specification[EB/OL]. 2007. <http://www.w3.org/TR/soap/>.

[4] 岳昆,王晓玲,周傲英. Web 服务核心支撑技术:研究综述[J]. 软件学报,2004,15(3):428-442.

[5] 杨胜文,史美林. 一种支持 QoS 约束的 Web 服务发现模型[J]. 计算机学报,2005,28(4):589-594.

[6] Comuzzi M, Pernici B. A framework for QoS-based Web service contracting[J]. ACM Trans. on Web, 2009, 3(3):1252-1254.

[7] Lin Chia-Feng, Sheu Ruey-Kai, Chang Yue-Shan, et al. A relaxable service selection algorithm for QoS-based web serv-

ice composition[J]. ELSEVIER Information and Software Technology, 2011, 53(12):1370-1381.

[8] Stephen S, Yau Yinyin. QoS-based Service Ranking and Selection for Service-based Systems[C]//2011 IEEE International Conference on Services Computing. [s. l.]:[s. n.], 2011.

[9] Yang S J H, Hsieh J S F, Lan B C W, et al. Composition and Evaluation of Trustworthy Web Services[J]. International Journal of Web and Grid Services, 2006, 2(1):5-24.

[10] Vambenepe W, Thompson C, Talwar V, et al. Dealing with Scale and Adaptation of Global Web Services Management[C]//Proceedings of the IEEE International Conference on Web Services(ICWS'05). Orlando:[s. n.], 2005:339-346.

[11] Ran S. A Model for Web Services Discovery with QoS[J]. ACM SIGEcom Exchanges, 2003, 1(4):1-10.

[12] Wang Xia, Vitvar T, Kerrigan M, et al. A QoS-aware Selection Model for Semantic Web Services[C]//Proceedings of the 4th International Conference on Service-oriented Computing. [s. l.]:[s. n.], 2006:390-401.

[13] 邵凌霄,周立,赵俊峰,等. 一种 Web Service 的服务质量预测方法[J]. 软件学报,2009,20(8):2062-2073.

[14] 黄景文,胡志华. Web 服务 QoS 的免疫多信号预测模型研究[J]. 广西大学学报(自然科学版), 2009, 34(4):535-539.

[15] Papazoglou M P, Paolo T, Schahram D, et al. Service-oriented Computing: State of the Art and Research Challenges[J]. IEEE Computer, 2007, 40(11):38-45.

(上接第 28 页)

4 结束语

人们对数据的应用从简单的数据查询,到从海量数据中挖掘知识并提供关联的服务,这就是数据挖掘的作用。文中针对基于关联算法的数据挖掘系统进行了分析和设计,并实现了一个简单的但是完整的银行业务数据挖掘系统,对客户账户情况、贷款情况等进行信息的预测,促进了基于关联的数据挖掘系统的应用研究。

参考文献:

[1] 安淑芝. 数据仓库与数据挖掘[M]. 北京:清华大学出版社,2005.

[2] 赵裕啸,倪志伟,王园园,等. SQL Server 2005 数据挖掘技术在证券客户忠诚度的应用[J]. 计算机技术与发展, 2010, 20(2):229-232.

[3] 陈艳,张燕平. 数据挖掘技术在保险客户信用评估的应用[J]. 计算机技术与发展, 2008, 18(5):179-181.

[4] 宋宝莉,覃征. 分布式全局频繁项目集的快速挖掘方法

[J]. 西安交通大学学报, 2006, 40(8):923-927.

[5] 唐瑜,王勇. 挖掘最大频繁项集的优化方法[J]. 计算机工程与应用, 2006, 42(31):171-173.

[6] 邵峰晶,于忠清. 数据挖掘原理与算法[M]. 北京:中国水利水电出版社, 2003:123-135.

[7] 黄进,尹治本. 关联规则挖掘的 Apriori 算法的改进[J]. 电子科技大学学报, 2003, 32(1):76-79.

[8] 朱孝宇,王理东,汪光阳. 一种改进的 Apriori 挖掘关联规则算法[J]. 计算机技术与发展, 2006, 16(12):89-90.

[9] 李绪成,王保保. 挖掘关联规则中的 Apriori 算法的一种改进[J]. 计算机工程, 2002, 28(7):104-105.

[10] Han Jiawei, Kamber M. 数据挖掘概念与技术[M]. 范明, 孟小峰译. 北京:北京机械工业出版社, 2001:149-180.

[11] 宋余庆,朱玉全,孙志挥,等. 基于 FP-tree 的最大频繁项目集挖掘及更新算法[J]. 软件学报, 2003, 14(9):1586-1592.

[12] 凌绪雄,王社国,李洋,等. 无项头表的 FP-Growth 算法[J]. 计算机应用, 2011, 31(5):1391-1394.

[13] Microsoft 关联算法技术参考[EB/OL]. 2008. <http://msdn.microsoft.com/zh-cn/library/cc280428.aspx>.