

文章编号: 2095-2163(2019)01-0199-04

中图分类号: TP393

文献标志码: A

# Web 服务 QoS 预测与主动推荐方法综述

闫红丹, 杨怀洲

(西安石油大学 计算机学院, 西安 710065)

**摘要:** 随着 Web 服务数量的迅猛增加,如何及时有效地根据服务质量(Quality of Service, QoS)推荐优质服务已成为研究热点。在实际应用中,由于网络环境、地理位置和服务运行环境等存在差异,导致 QoS 的不同并表现出强烈的不稳定性,仅通过挖掘 QoS 数据中近邻信息难以保证预测的准确度。协同过滤算法作为现今在解决 QoS 预测准确度方面的一个重要方法,就非常值得研究。本文针对基于 Web 服务 QoS 预测的主动服务推荐方法及其存在的问题进行了分析和总结。总结了当下 QoS 预测方法的研究趋势和发展前景。

**关键词:** Web 服务; QoS 预测; 协同过滤; 服务推荐

## QoS prediction and active recommendation for Web services overview

YAN Hongdan, YANG Huaizhou

(School of Computer Science, Xi'an Shiyou University, Xi'an 710065, China)

**[Abstract]** With the rapid increase in the number of Web services, how to recommend quality services according to the quality of service (Quality of Service, QoS) in a timely and effective manner has become a hot research topic. In practical application, due to the difference of network environment, geographical location and service running environment, QoS is different, which shows strong instability. It is difficult to guarantee the accuracy of prediction only by mining neighbor information in QoS data. As an important method to solve the problem of QoS prediction accuracy, collaborative filtering algorithm is worth studying. In this paper, an active service recommendation method based on Web service QoS prediction and its existing problems are analyzed and explored. Subsequently, the research trend and development prospect of current QoS prediction methods are summarized.

**[Key words]** Web services; QoS prediction; collaborative filtering; service recommendation

## 0 引言

近年来,Internet 的迅猛发展使其成为全球信息传递与共享的巨大资源库。越来越多的网络环境下的 Web 应用系统被建立起来,Web 服务也日益丰富。根据近几年数字统计显示,网络上有由 7 739 个提供者提供有 28 606 个可用的 Web 服务<sup>[1]</sup>。那么如何从海量的 Web 服务中快速有效地帮用户推荐一个优质的服务,就非常重要了。

Web 服务作为独立的,自描述的模块化应用程序,是一种松散耦合的软件系统。旨在支持网络上机器到机器的自动交互,已经普遍部署并且可以在 Web 上使用。这让使用者查找、选择和调用性能较好的 Web 服务变得很难。QoS 通常被用于描述 Web 服务的非功能特性<sup>[2]</sup>,QoS 的服务特性有性能(执行时间、响应时间、吞吐量等)、可信度(可用性、

可靠性、准确性、稳定性、完整性等)和用户满意度等。QoS 是动态发现、查询、选择和主动推荐服务的基础<sup>[3]</sup>。但是获取真实有效的与用户相关的 QoS 属性的值是一项具有挑战性的任务,这就需要对 QoS 进行预测。

预测较为准确的 QoS,就可以向用户推荐一些水平较好且功能相同的服务。因此,如何将用户偏好和 QoS 预测相结合并准确地向用户推荐优质的服务,是 Web 服务研究的热点和难点,不仅具有重要的理论意义,还具有重大的实用价值。

## 1 Web 服务 QoS 预测方法

为了更好地给用户推荐一个优质的 Web 服务,协同过滤(collaboration filtering, CF)算法<sup>[4]</sup>作为用于 Web 服务 QoS 预测与主动服务推荐的一个重要方法,国内外学者都对其进行了研究。协同过滤最

基金项目: 西安市科技计划项目(201805038YD16CG22(2))。

作者简介: 闫红丹(1993-),女,硕士研究生,主要研究方向: Web 服务推荐、QoS 预测; 杨怀洲(1970-),男,博士,副教授,主要研究方向: 物联网、服务计算。

通讯作者: 杨怀洲 Email: 172050528@stumail.xsyu.edu.cn

收稿日期: 2018-11-02

先由 Goldberg 在 1992 年提出<sup>[6]</sup>, CF 是通过收集其它类似用户或 Web 服务的历史 QoS 信息来预测当前用户的 QoS 值的方法。众所周知协同过滤算法根据实现算法的推荐策略不同分为基于邻域 (Memory-based) 的 CF 方法和基于模型 (Model-based) 的 CF 方法。接下来介绍几种具有研究价值的协同过滤算法。

### 1.1 基于邻域 (Memory-based) 的协同过滤算法

基于邻域的协同过滤算法依据系统中已有的用户 QoS 信息,在内存中通过一定的启发式策略实现活动用户对目标项目的 QoS 信息预测,并通过相似度较高的一部分邻居用户或者 Web 服务来帮助预测 QoS 值并发布建议。Memory-based CF 分为 User-based CF 和 Item-based CF。

#### 1.1.1 基于用户 (User-based) 的协同过滤算法

基于用户的 (User-based) 协同过滤算法是基于这样一个假设:如果一些用户对某一类服务项的 QoS 信息(如,响应时间)比较接近,则对其它类似服务项的响应时间也比较接近。相似用户对某一 item 的 QoS 信息相似,即先计算用户相似性,然后找到对 item  $i$  预测过的用户,找到最相似  $top - k$  个用户进行预测。User-based 协同过滤推荐算法的核心就是通过相似性度量方法计算出最近邻居集合,并将最近邻的 QoS 信息结果作为推荐预测结果返回给用户。

目前主要有 3 种度量用户间相似性的方法,分别是:余弦相似性、修正的余弦相似性以及皮尔逊相关系数 (Pearson correlation coefficient, PCC)<sup>[4]</sup>。

(1) 余弦相似性 (Cosine)。余弦相似性度量方法是通过计算向量间的余弦夹角来度量用户间的相似性。其实现简单、计算速度快,但未能体现出用户 QoS 信息的特征。

(2) 修正的余弦相似性 (Adjusted Cosine)。余弦相似性未考虑到用户 QoS 信息(如,吞吐量)尺度问题。修正的余弦相似性通过减去用户对服务项的平均吞吐量,修正的余弦相似性度量方法改善了以上问题,更多地体现了用户的相关性而不是相似性。

(3) 皮尔森 (Pearson) 相关系数。PCC 不仅考虑了服务项的平均 QoS 信息,还可以有效地防止某用户总是倾向给比另外一个用户更高的 QoS 值,而二者的分分之差又始终保持一致,那么则认为二者具有很好的相似性。

#### 1.1.2 基于服务项 (Item-based) 的协同过滤算法

基于服务项的 (Item-based) 协同过滤是先通过

计算 item 的相似性,然后根据 item 相似值寻找某一用户下最接近预测 item 的  $top - k$  个 items 进行预测。Item-based 协同过滤算法的关键步骤仍然是计算项目之间的相似性,并选出最相似的服务项,这一点与 User-based 协同过滤类似。

众多学者对基于协同过滤的 QoS 预测方法进行了大量的研究。基于用户和服务的协同过滤算法在很大程度上依赖于历史 Web 服务的调用信息,每个用户每次会从众多的候选对象中调用一个或者多个 Web 服务,导致可用 QoS 数据矩阵高度稀疏的,而且一些历史值没有实时更新,甚至过时。因此,在稀疏和冷启动的场景下得到的预测结果不理想。同时,协同过滤算法往往会考虑  $top - k$  个相似用户或服务,在 Web 服务非常多时,往往会使得参与计算的样本集太小,导致结果不全面。但是如果扩大样本集,即  $K$  值设定较大,那么算法的效率就会受到影响。此外,因为 QoS 的特性较多,上述方法也不支持多维度的 QoS 预测。

由于上述单一使用 User-based 或 Item-based 对 QoS 预测准确度的不足,Z. Zheng 等人<sup>[4]</sup>结合使用这 2 种算法,其不需要实际的 Web 服务调用,通过分析来自类似用户的 QoS 信息,为用户发现合适的 Web 服务就可以预测目标 Web 服务的相关用户的 QoS 信息。在此文中,Z. Zheng 等人针对原有用于计算相似度的 PCC 存在未考虑重叠记录项的数量对相似性计算的影响,设计出了一个具有显著加权的用于避免出现高估相似度(即出现相似偶然性)的现象的一个自适应的相似度计算方法。比之前的用于计算相似性的方法有很大的改善。

上述方法,没有考虑 QoS 数据缺失甚至没有的情况。Z. Zheng 等人<sup>[5]</sup>针对此情况进行了研究。通过给 PCC 增加一个参数,克服了计算用户或项目相似性精度的下降,并提出了一种有效的缺失数据预测算法。该算法考虑了用户信息和项目信息,分别为用户和项目设置了相似性阈值,预测算法将决定是否预测丢失的数据。改进了协同过滤算法,并且对数据的稀疏性具有更强的鲁棒性。

### 1.2 基于模型 (Model-based) 的协同过滤算法

为了避免 Memory-based 中 QoS 矩阵稀疏性和在处理大量数据时的时效性等问题,提出了基于模型的协同过滤技术。该方法是利用历史数据得到一个模型,模型的建立可以使用各种机器学习的方法,再用此模型进行预测。基于模型的协同过滤算法是一种线下学习,然后进行线上预测。这里主要介绍

了基于聚类的 CF<sup>[7-11]</sup> 和基于矩阵分解的方法<sup>[12-13]</sup>。

### 1.2.1 基于聚类的协同过滤算法

聚类是数据分析中常用的一种技术。聚类的主要任务是将数据聚类成不同的组,并且同一类中的数据比其它类别中的数据更相似。Z. Zheng 等人<sup>[7]</sup>为了解决数据稀疏的问题,提出了一种新的基于聚类的 QoS 预测方法。通过向框架中设置一组固定的地标(计算机),地标可以周期性地监视可用的 Web 服务,以丰富 QoS 数据,从而更准确地预测 QoS。Marin Silic 等人<sup>[8]</sup>利用 K 均值聚类算法来聚合以前可用的调用数据对 QoS 预测的可靠性问题进行了研究,提出了一个用于原子 Web 服务可靠性预测的 CLUS 模型,CLUS 利用以前调用中收集的数据预测正在进行的服务调用的可靠性。Marin Silic 等人通过考虑调用内容的用户、服务和环境参数来提高当前状态预测模型的准确性,以解决与调用内容相关的计算性能的可伸缩性问题。K-均值聚类是一种迭代算法,在簇间移动项目直至达到所需的集合<sup>[10]</sup>。使用 KMC 减少搜索空间<sup>[11]</sup>。标准的 KMC 方法生成  $k$  个聚类,每个簇都由具有相似偏好的客户组成,在该方法中,分别选择任意  $k$  个客户作为  $k$  簇的初始中心点。然后,将每个客户分配到集群中,使客户与集群中心之间的相似性最大化。A. Suresh Poobathy 等人<sup>[9]</sup>对基于 K-均值聚类算法提出了精化的 K Means Clustering (KMC) 算法。KMC 是一种流行的基于数据划分的聚类算法。用户首先给出了聚类的个数及其对初始条件的敏感性,其次给出了线性可分聚类。并采用了遗传算法,提高了聚类质量。由此可见,基于聚类的协同过滤算法在处理数据稀疏和预测的可靠性方面有突出的优势。

### 1.2.2 基于矩阵分解的 QoS 预测算法

Wei Lo 等人<sup>[12]</sup>针对历史记录中存在许多缺失的 QoS 值和为了避免昂贵的 Web 服务调用等问题,提出了一种扩展矩阵因式分解(extended Matrix Factorization, EMF)的方法。本框架结合关系正则化进行缺失的 QoS 值预测,首先为了更准确地收集人群的智慧,研究者在用户端和服务端使用不同的相似性度量来识别邻域,然后在邻域内系统地设计了 2 个新的关系正则化项。最后,将这 2 个术语合并成一个统一的 MF 框架。

Zheng 等人<sup>[13]</sup>提出了一种自适应矩阵分解(adaptive matrix factorization, AMF)方法,其利用

不同用户观测到的历史 QoS 数据来准确估计候选服务的 QoS 值的 QoS 预测问题,同时消除了目标用户对额外服务调用的需求。其支持及时准确的自适应决策,可以高效地预测 QoS 来获得组件服务的实时 QoS 信息。AMF 为了适应 QoS 随时间的变化对候选服务进行在线 QoS 预测,利用数据转换、在线学习和自适应加权等新技术,对传统的矩阵分解模型进行了显著的扩展。与现有方法相比,AMF 不仅在准确度上得到了很大地提高,而且保证了高效率 and 鲁棒性,对于实现最佳运行时服务适配至关重要。

基于 Model-based 的 CF 在离线阶段进行数据的预处理;在运行阶段,只有 learned model 才能用于预测,如果在系统中添加新的用户项时,模型要定期的更新和重构。运用大量的技术,模型的建立和更新计算量相当大。

### 1.3 混合算法

上述方法对 QoS 预测在 Memory-based 方面存在的数据稀疏性和数据集过大等问题利用 Model-based 方法进行了开创性的研究,但是并未考虑 Memory-based 方法中用户偏好和 QoS 特性的多样性问题。Karta<sup>[14]</sup>提出了结合基于规则和内容的协同过滤的多维度推荐方法,从而避免和弥补了各推荐技术的弱点。

为解决用户评分数据极端稀疏的问题, Sarwar BM 等人<sup>[15]</sup>通过奇异值分解(singular value decomposition)减少了项目的空间维数。但维数的降低会导致 QoS 信息损失,特别是在项目空间维数很高的情况下,推荐效果难以保证。张卫光等人<sup>[16]</sup>利用云模型在定性知识表示以及定性、定量知识转换时的桥梁作用,提出了一种基于云模型的用户相似度比较方法,在一定程度上克服了用户评分数据极端稀疏的负面影响。

针对不可信用户提供的不可靠数据对推荐质量的影响, Massa 等人<sup>[17]</sup>提出了基于信任的协同过滤技术,即根据用户之间信任关系推荐服务:若用户 A 信任用户 B,那么用户 B 向其推荐的服务一定满足 A 的需求。但是并未解决如何在成千上万的用户中寻找可信用户的问题<sup>[18]</sup>。

现有的用于 Web 服务推荐的方法,在个别方面(如:敷衍性评分、评价指标等)还不能兼顾,还不够成熟。随着 Web 服务的使用和存在,设计有效的 Web 服务推荐的新方法正变得越来越重要,现有的 Web 服务发现和推荐方法要么集中在消亡的 UDDI 注册中心<sup>[19]</sup>,要么是以关键字为主的 Web 服务搜

索引引擎,这些方法存在推荐性能不足和对用户输入的依赖程度大的缺点。为了向用户推荐出一个更优质的 Web 服务,还需要人们更深入的研究。

## 2 结束语

随着用户对服务质量的重视和要求的提高,Web 服务的 QoS 预测还有很多方面需要进行更深入的研究,主要体现在如下几个方面。

(1)对于推荐系统存在冷启动(没有用户行为数据)问题。

(2)项目评分矩阵存在诸多的敷衍性评分(如:用户对不感兴趣的敷衍性评分),不可信用户提供的不可靠数据会误导相似度计算,影响推荐质量。

(3)现有的评分指标还不够健全,没有统一的用于评价众多推荐算法优良的指标。

(4)现有的预测算法的算法复杂度是  $O(mn + n^2)$ <sup>[3-4]</sup>(指改推荐系统有  $m$  个训练用户和  $n$  个 Web 服务项),整体的计算时间复杂度与特征的个数呈线性关系,如何降低算法的复杂度,在未来非常有必要研究。

(5)由于实际环境中服务 QoS 值巨幅变化,因此如何使预测方法适应 QoS 值的动态变化是下一步研究的重点。

## 参考文献

- [1] YAO L, SHENG Q Z SEGEV A. Recommending Web services via combining collaborative filtering with content-based features [C]//2013 IEEE 20<sup>th</sup> International Conference on Web Services(ICWS). Santa Clara, California; IEEE, 2013;42-49.
- [2] ZENG Liangzhao, BENATALLAH B, NGU A, et al. QoS-aware middleware for web services composition[J]. IEEE Transactions on Software Engineering, 2004, 30(5) :311-327.
- [3] YANG Huaizhou, ZHANG Liumei. A collaborative filtering QoS prediction and service recommendation method for internet of things[J]. Computer Engineering, 2016, 42(11) : 1-7.
- [4] CHEN Xi, ZHENGZibin, LYU M R. QoS-aware web service recommendation via collaborative filtering [M]//BOUGUETTAYA A, et al. Web Services Foundations. New York; Springer Science+Business Media, 2014;563-588.
- [5] MAHao, KING I, LYU M R. Effective missing data prediction for collaborative filtering[C]//30<sup>th</sup> Annual International ACM SIGIR Conference on Research & Development on Information Retrieval. Amsterdam; ACM, 2007; 39-46.
- [6] GOLDBERG D, NICHOLS D, OKI B M, et al. Using collaborative filtering to weave an information tapestry [J]. Communications of the ACM, 1992,35(12) :61-70.
- [7] ZHUJieming, KANG Yu, ZHENG Zibin, et al. A clustering-based QoS prediction approach for Web service recommendation [C]// 2002 IEEE 15<sup>th</sup> International Symposium on Object Component Service-oriented Real-time Distributed Computing Workshops.Washington, DC, USA; IEEE, 2012; 93-98.
- [8] SILICM, DELAC G, SRBLJIC S. Prediction of atomic web services reliability based on K-means clustering[C]// Proceedings of the 2013 9<sup>th</sup> Joint meeting on Foundations of Software Engineering. Saint Petersburg, Russia; ACM,2013;70-80.
- [9] POOBATHYA S, MANMEGALAI C. Model-based collaborative filtering using refined K-Means and genetic algorithm [J]. International Journal of Scientific Engineering and Technology, 2013,2(6) : 505-510.
- [10] CHITTUV, SUMATHI N. A modified genetic algorithm initializing K-Means clustering [J]. Global Journal of Computer Science and Technology, 2011,1(2).
- [11] KIM T H, RYUY S, PARK S I, et al. An improved recommendation algorithm in collaborative filtering [M]//BAUKNECHT K, TJOA A M, QUIRCHMAYR G. E-commerce and web technologies. EC-Web 2002. Lecture Notes in Computer Science. Berlin/ Heidelberg;Springer, 2002,2455;254-261.
- [12] WEI Lo, YIN Jianwei, DENG shuigang, et al. An extended matrix factorization approach for QoS prediction in service selection [C]//IEEE Ninth International Conference on Services Computing. Honolulu, Hawaii, USA;dblp,2012: 162-169.
- [13] ZHUJieming, HE Pinjia, ZHENG Zibin, et al. Online QoS prediction for runtime service adaptation via adaptive matrix factorization[J]. IEEE Transactions on Parallel and Distributed Systems, 2017,28(10) :2911-2924.
- [14] KaRTAK. An investigation on personalized collaborative filtering for web service selection[R]. Australia;The university of Western Australia, 2005.
- [15] SARWARB, KARYPIS G, KONSTAN J, et al. Application of dimensionality reduction in a recommender system-A case study [C]//WebKDD-2000 Workshop.Boston, MA; ACM 2000;1-12.
- [16]ZHANG Guangwei, LI Deyi, LI Peng, et al. A collaborative filtering recommendation algorithm based on cloud model [J]. Journal of Software, 2007,18(10) :2403-2411.
- [17] MASSAP, BHATTACHARJEE B. Using trust in recommender systems; An experimental analysis[M]//JENSEN C, POSLAD S, DIMITRAKOS T. Trust Management. iTrust 2004. Lecture Notes in Computer Science, Berlin/ Heidelberg; Springer, 2004, 2995;221-235.
- [18] RAN Shuping. A framework for discovering web services with desired quality of service attributes [C]//Proceedings of the International Conference on Web Services. Las Vegas, Nevada, USA;dblp,2003: 208-213.
- [19]GUINARD D, TRIFA V, KARNOUSKOS S, et al. Interacting with the SOA-based Internet of Things: Discovery, query, selection, and on-demand provisioning of web services[J]. IEEE Transactions on Services Computing, 2010, 3(3) :223-235.