

# 基于评价分类的可信 QoS 服务选择方法

夏薪棋,刘茜萍

(南京邮电大学 计算机学院,江苏 南京 210003)

**摘要:**随着众多的服务参与到服务计算中,服务选择变得更加重要。现有的服务选择方法往往仅进行可信用户的甄别,可信用户的所有评价均直接用于服务选择,然而,可信用户的一些评价习惯往往会造成服务评价的一些偏离,影响了服务选择的结果。为此,提出一种基于用户评价分类的可信 QoS 服务选择方法,在分析用户评价习惯的基础上将用户评价详细分类,进而筛选恶意用户,同时,标记出可信用户的偏离评价,并进一步计算偏离评价数据的偏离度以修正可信用户的偏离评价数据,使偏离评价数据的参考价值更高。然后基于相似度的计算与量化以确定每个服务下各可信用户的参考权重,进而计算得到各服务的 QoS 评分值并完成服务选择。该方法的有效性通过一个具体实例得到了验证。

**关键词:**用户评价习惯;评价分类;评价修正;可信 QoS;服务选择

**中图分类号:**TP301

**文献标识码:**A

**文章编号:**1673-629X(2018)08-0114-05

doi:10.3969/j.issn.1673-629X.2018.08.024

## A Trusted QoS Selection Method Based on Evaluation Classification

XIA Xin-qi, LIU Xi-ping

(School of Computer Science, Nanjing University of Posts and Telecommunications, Nanjing 210003, China)

**Abstract:** As more and more users and services participate in service computing, service selection becomes increasingly important. Existing service selection methods are often conducted only for trusted users, and all of evaluation of trusted users are used directly in service selection. However, some of the evaluation habits of trusted users tend to cause some deviations from service evaluations that affect the outcome of service selection. For these reasons, we propose a trusted QoS service selection method based on user evaluation classification. On the basis of analyzing the user's evaluation habits, the user evaluation is classified in detail, then screen out a malicious user comments and mark the deviation of the trusted user. The deviation from the evaluation data is calculated to correct the deviation evaluation data of the trusted user so that the reference value deviated from the evaluation data is higher. Based on the similarity of service benchmark evaluation vectors and trusted user evaluation vectors the reference weights of each trusted user are calculated, and the final QoS evaluation scores of each candidate service are obtained for the target users. The validity and feasibility of this method are verified by a specific example.

**Key words:** user evaluation habits; evaluation classification; evaluation correction; trusted QoS; service selection

## 0 引言

随着科技的发展与网络的兴起,计算机软件的相服务逐渐增多<sup>[1]</sup>。然而,满足相同功能需求的服务也越来越多。为了使得用户有更好的体验,一些服务商家不仅需要考虑服务本身的功能需求,还需要考虑服务的非功能需求。可信问题是当前服务选择面临的一个重要问题,不能解决可信问题,就无法从相似的服务集中选择出更好的服务。一个服务的好坏可以使用 QoS (quality of service)<sup>[2-4]</sup>来衡量。历史用户给出的反馈信息可以用来计算服务的 QoS,现实中这些用户

反馈的可信性是难以保证的。某些用户给出的虚假好评或者恶意差评影响了 QoS 计算的准确性<sup>[5-7]</sup>。除此之外,用户的个人主观评价习惯对评价数据也有一定的影响。以往的研究往往认为可信的实体做出的服务评价是完全合理可用的<sup>[8-12]</sup>,但实际并非如此,考虑到用户的评价习惯对评价数据的影响,可信用户给出的主观评价可能与实际情况存在一定偏离,对这些偏离评价进行甄别并修正,将可以准确有效地选择出更适用户需求的服务。

文中提出一种基于评价分类的可信 QoS 服务选

收稿日期:2017-08-30

修回日期:2018-01-05

网络出版时间:2018-04-28

基金项目:国家自然科学基金(61402241)

作者简介:夏薪棋(1989-),男,硕士研究生,研究方向为服务计算;刘茜萍,副教授,硕导,研究方向为服务计算、工作流。

网络出版地址: <http://kns.cnki.net/kcms/detail/61.1450.TP.20180427.1630.030.html>

择方法。基于历史用户对服务的反馈评价,充分考虑用户评价习惯对评价的影响,将评价的服务属性详细分类为中立评价、宽容评价、挑剔评价、恶意评价,并进一步对可信用户的偏离评价数据进行修正,以减少用户评价习惯对服务评价的影响。这里偏离评价主要包括可信用户的宽容评价和挑剔评价。

1 相关研究

近年来,服务选择的相关研究取得了一定成果。熊润群等<sup>[13]</sup>设计了三维 QoS 模型,并结合层次分析法设计出了 QoS 偏好感知算法与副本选择算法,进而获取 QoS 排序。唐朝刚<sup>[14]</sup>对 QoS 的服务选择问题进行了研究,设计出了 top-k 的服务选择查询技术。

然而以上这些研究没有考虑 QoS 描述信息的可信性。对此,姚建华等<sup>[15]</sup>使用多种服务的非功能属性计算服务的 QoS,并使用聚类方法提高服务的 QoS 可信性。刘昕民等<sup>[16]</sup>使用 D-S 证据理论计算服务的 QoS 可信度。杨丹榕<sup>[17]</sup>结合贝叶斯理论和粗糙集理论对 QoS 的属性进行了可信度评估和权重计算,并通过 QoS 度量模型选择出优质的服务。

这些传统的服务选择方法,多数都是基于历史用户的反馈评价剔除恶意用户,然后直接使用可信用户的评价数据进行 QoS 的计算,进而以 QoS 值的高低完成服务选择。直接将可信用户的反馈数据用来计算综合 QoS 具有不合理性。因为现实生活中确实存在这种情况,某些用户对服务的不同属性具有不同的评判标准,这是由他们的评价习惯造成的。有些用户或许对服务的某种属性要求很挑剔或者很宽容。总的来说,这些用户是可信的,因为他们并没有恶意的评价,而是由于个人的评价习惯不同导致评分的不合理。如果不对这些用户的评价进行详细的分类,那么直接使用剔除后所得的可信用户数据也有失公正合理性。

为了保证数据具有更好的合理性,需要充分考虑用户的评价习惯对数据的影响。对此,文中提出了一种基于评价分类的可信 QoS 服务选择方法,分析用户的历史评价并将其详细分类为中立评价、宽容评价、挑剔评价和恶意评价。剔除恶意用户并筛选出值得信赖的可信用户评价,并对得到的用户数据做进一步的修正处理,以削弱用户的主观个人评价习惯对 QoS 数据正确性的影响,从而选择出更合理更优质的服务。

2 基于用户评价分类的可信用户获取

历史用户的反馈评价是描述候选服务质量的重要依据。基于历史用户的反馈将评价详细分类,从中甄别出具有参考价值的可信用户反馈评价。

2.1 历史用户评价信息描述

一个服务应该具有多个非功能属性,历史用户在使用服务后都会对服务的各非功能属性给出具体评分值,即历史用户的反馈评价。

相关定义如下:

定义 1(  $S,U,A,P$  ):  $U$  表示历史用户集,  $U = \{u_1, \dots, u_i, \dots, u_m\}$ ;  $S$  表示所有用户使用过的服务集合,  $S = \{s_1, \dots, s_j, \dots, s_n\}$ ;  $A$  表示服务的属性集合,  $A = \{a_1, \dots, a_k, \dots, a_l\}$ ;  $P$  表示历史用户的具体评分值,  $P = \{P_{i,j} \mid P_{i,j} \text{ 为用户 } u_i \text{ 关于服务 } s_j \text{ 的评价向量}, P_{i,j} = \langle p_{i,j,1}, \dots, p_{i,j,k}, \dots, p_{i,j,l} \rangle, i = 1, 2, \dots, m, j = 1, 2, \dots, n, k = 1, 2, \dots, l, 1 \leq p_{i,j,k} \leq 10\}$ ,  $p_{i,j,k}$  表示  $u_i$  对  $s_j$  的  $a_k$  的评分值,  $p_{i,j,k}$  越大表明用户对服务越满意。

2.2 历史用户评价分类及可信用户获取

如果某个历史用户给出的反馈评价向量明显偏离该服务的总体评价向量,那么这个反馈评价向量就有可能是虚假好评或者恶意差评。考虑到用户的评价习惯对数据的影响,有的用户对某种属性是偏宽容的,那么他给的评分可能较总体评分高,相反地,用户对某种属性是偏挑剔的,那么他给的评分可能较总体评分低。这里将用户的评价分为四类:中立评价、挑剔评价、宽容评价、恶意评价,其具体计算过程如下:

首先计算每个服务的每种属性的一个具体平均评分值,即所有用户对某个服务的某种属性的评分值总和除以用户的个数之和。相关定义如下:

定义 2(  $e$  ):  $e_j^k$  表示服务  $s_j$  的属性  $a_k$  的平均值,其中  $j = 1, 2, \dots, n, k = 1, 2, \dots, l, 1 \leq e_j^k \leq 10$ 。

$e_j^k$  的计算公式如下:

$$e_j^k = \frac{\sum_{i \in (1,m)} p_{i,j,k}}{m}$$

(1)

若有  $n$  个服务,对于某属性  $a_k$  就有  $n$  个平均值,即构成一个  $n$  维向量,这可作为衡量用户对属性  $a_k$  的评价可信与否的基准向量  $E^k$ 。对用户  $u_i$  来说,在属性  $a_k$  上也会有对应  $n$  个评分值,同样构成一个  $n$  维向量  $F_i^k$ ,具体定义如下:

定义 3(  $E,F$  ):  $E^k$  表示由  $n$  个属性平均值构成的  $n$  维向量,  $E^k = \langle e_1^k, \dots, e_j^k, \dots, e_l^k \rangle$ , 其中  $j = 1, 2, \dots, n, k = 1, 2, \dots, l$ ;  $F_i^k$  表示  $u_i$  对所有服务  $S$  的属性  $a_k$  的评分值所构成的向量,  $F_i^k = \langle p_{i,1,k}, \dots, p_{i,j,k}, \dots, p_{i,n,k} \rangle$ , 其中  $i = 1, 2, \dots, m, j = 1, 2, \dots, n, k = 1, 2, \dots, l$ 。

用户消费了多个服务,每个服务都具有相同的多个属性,同一类型的属性上多个服务的不同评价就可以组成一个向量  $F_i^k$ ,将该向量与基准向量  $E^k$  做差异计算。这里的差异计算,不是向量的整体一次性比较,而是两个向量相对应元素间的差异计算。计算结果仍

是一个  $n$  维向量,相关定义如下:

定义 4(  $H, h$  ):  $H_i^k$  表示用户对  $F_i^k$  与  $E^k$  差异计算的结果,  $H_i^k = \langle h_{i,1}^k, \dots, h_{i,j}^k, \dots, h_{i,n}^k \rangle$ , 其中  $i = 1, 2, \dots, m, j = 1, 2, \dots, n, k = 1, 2, \dots, l$ 。

$H_i^k, h_{i,j}^k$  的计算公式如下:

$$H_i^k = F_i^k - E^k \quad (2)$$

$$h_{i,j}^k = p_{i,j,k} - e_j^k \quad (3)$$

然后再将向量  $H_i^k$  中的每个元素与阈值做差异比较,这里设置两个阈值,相关定义如下:

定义 5(  $\alpha, \beta$  ):  $\alpha$  表示较小阈值,  $\beta$  表示较大阈值,且  $\alpha \geq 0, \beta \geq 0$ 。

评价分类定义如表 1 所示。

表 1 评价分类定义表

评价	定义
中立评价	$(\forall j) h_{i,j}^k \in \langle -\alpha, \alpha \rangle$
宽容评价	$((\forall j) h_{i,j}^k \in \langle -\beta, \alpha \rangle) \wedge ((\exists j) h_{i,j}^k \in \langle \alpha, \beta \rangle)$
挑剔评价	$((\forall j) h_{i,j}^k \in \langle -\beta, \alpha \rangle) \wedge ((\exists j) h_{i,j}^k \in \langle -\beta, -\alpha \rangle)$
恶意评价	$((\exists h_{i,j}^k)((h_{i,j}^k \geq \beta) \vee (h_{i,j}^k \leq -\beta)) \vee ((\alpha \leq h_{i,j}^k \leq \beta) \wedge (-\beta \leq h_{i,j}^k \leq -\alpha)))$

对所有属性都进行类似的处理,进而得到用户对各属性的评价类型。如果存在某个属性上的恶意评价,那么此用户的评价参考价值较小,此用户将被认定是不可信用用户,因而将这些用户从历史用户集中剔除,剩余的用户便是可信用用户。

然而,考虑到用户的主观评价习惯对评分值的影响,即便是可信用用户,其评分也可能出现偏离的情况,需要对用户的偏离评价数据进行修正,具体方法见下节。

### 3 基于可信用用户评价数据修正的服务选择

由于可信用用户主观的评价习惯使得他们的评分值与实际客观值可能有所偏离,在使用用户评价进行服务选择之前需要修正可信用用户的偏离评价数据。

#### 3.1 修正可信用用户的偏离评价

如果某个用户对服务的某种属性是偏宽容的,那么这个用户给出的评分值相对正常评分值将会偏高,这样就和服务的真实评价有所偏离,同样地,某个用户对服务的某种属性是偏挑剔的,评分值就会偏低。上一节基于恶意评价筛选了恶意用户,从而得到了可信用用户集,对每一个可信用用户而言,需要遍历所有评价数据并对其偏离评价数据进行修正。这里给出一个偏离度的概念,偏离度就是宽容评价或挑剔评价偏离中立评价的幅度。修正数据的目的是缩小偏离度,使得偏离评价与合理评价相接近,从而使可信用用户的数据参

考价值更高。偏离度定义如下:

定义 6(  $d$  ):  $d_i^k$  表示用户  $u_i$  对某种属性  $a_k$  所给评价的偏离度,即宽容评价或挑剔评价偏离中立评价的幅度,  $1 \leq d_i^k \leq 10, i = 1, 2, \dots, m, k = 1, 2, \dots, l$ 。

通过计算  $F_i^k$  与  $E^k$  的欧几里得距离来实现偏离度的量化。

$d_i^k$  计算如下:

$$d_i^k = \sqrt{\frac{\sum_{j=1}^n (p_{i,j,k} - e_j^k)^2}{n}} \quad (4)$$

利用式 1~3 计算用户的评价类型,从而完成可信用户的筛选,利用式 4 计算出偏离评价的偏离度,这些偏离评价数据影响了服务选择质量。因此,需要修正这些偏离评价,可信用用户的原始偏离评价数据在不同的情况下减去或者加上偏离度,就得到了相对合理的可信评价数据。

修正结果定义如下:

定义 7(  $G, g$  ):  $G_i^k$  表示对原始属性评分向量  $F_i^k$  修正后的向量,  $G_i^k = \langle g_{i,1}^k, \dots, g_{i,j}^k, \dots, g_{i,n}^k \rangle$ , 其中  $i = 1, 2, \dots, m, j = 1, 2, \dots, n, k = 1, 2, \dots, l$ 。

下面给出修正偏离评价的计算公式:

$$g_{i,j}^k = p_{i,j,k} \pm d_i^k \quad (5)$$

比如某个用户  $u_i$  对某种属性  $a_k$  是宽容的,通过上述公式求出宽容度  $d_i^k$ ,然后在原始评价向量  $F_i^k$  的基础上减去宽容度  $d_i^k$  得到向量  $G_i^k$ ,这样得到的修正数据更合理。

#### 3.2 基于修正后评价数据的服务选择方法

基于上一节修正后的评价数据计算各服务的 QoS 具体值。计算历史用户对各服务的评价向量与相对应的服务基准向量之间的相似度,历史用户的评价向量与服务基准向量的相似度越小说明该用户对该服务反馈评价的参考价值越小,然后将得到的相似度以权重的形式进行量化,最后基于历史用户的反馈评价与量化后的权重计算得到各服务的 QoS 具体值,从而完成服务选择的目标。

相关定义如下所示:

定义 8(  $C$  ):  $C_j$  表示服务  $s_j$  的基准向量,即服务属性的平均值组成的向量,  $C_j = \langle e_j^1, \dots, e_j^k, \dots, e_j^l \rangle$ 。

向量间相似度的计算方法有很多种。其中,欧几里得相似方法能更好地表达用户历史评价向量与基准向量的相似程度。文中采用欧几里得相似方法来计算向量间的相似度。计算历史用户  $u_i$  的评价向量  $P_{i,j}$  与  $C_j$  的相似度,计算公式如下:

$$\text{sim}(P_{i,j}, C_j) = \frac{1}{1 + \sqrt{\sum_{k=1}^l \sum_{j=1}^n (p_{i,j,k} - e_j^k)^2}} \quad (6)$$

计算出用户评价向量  $\boldsymbol{P}_{i,j}$  与基准向量  $\boldsymbol{C}_j$  的相似度后,将相似度转换成参考权重,然后再计算候选服务最终 QoS 评分值。计算候选服务的 QoS 评分值  $qos_i$  的公式如下:

$$qos_i = \sum_{i=1}^m (\sum_{k=1}^l R_{i,j} * \text{sim}(\boldsymbol{P}_{i,j}, \boldsymbol{C}_j)) \quad (7)$$

最后,计算出每个服务的 QoS 值,最大 QoS 值对应的服务则是最适合目标用户的服务。具体算法如下所示:

算法:基于用户评价分类的可信 QoS 服务选择方法。

输入:  $s_j, a_k, p_{i,j,k}, \alpha$

输出: QoS(服务综合质量)值最高的服务  $s$

1. for each  $s_j$
2. for each  $a_k$
3.  $e_j^k = \text{getAverage}(p_{i,j,k})$
4. for each  $a_k$
5.  $E^k = \text{getAttributeVector}(e_j^k)$
6.  $C_j = \text{getAllAttr}(E^k)$

7. for each  $u_j$
8. for each  $a_k$
9.  $\boldsymbol{H}_i^k = \text{evalClassfy}(e_j^k, a)$  //用户评价分类
10.  $u_i^{\wedge} = \text{eliniSpiteUser}(u_i)$  //获得可信用户集
11. for each  $u_i^{\wedge}$
12. for each  $a_k$
13.  $d_i^k = \text{divergence}(\boldsymbol{H}_i^k, \boldsymbol{E}^k)$  //偏离度计算
14.  $\boldsymbol{G}_i^k = \text{amend}(\boldsymbol{H}_i^k, d_i^k)$  //修正偏离评价数据
15.  $\text{similarity}(R_{i,j}, \boldsymbol{C}_j)$  //相似度计算并做权重量化
16. for each  $s_j$
17.  $\text{calculateQos}(s_j)$  //服务选择
18.  $s = \text{sort}(s_j)$

4 实例分析

以餐厅服务类为例,这里有 10 个候选服务以及 10 个消费过这些服务的用户,选取服务态度、价格、上餐速度三个指标作为服务的属性,反馈评价向量表如表 2 所示。

表 2 反馈评价信息表

	$S_1$	$S_2$	$S_3$	$S_4$	$S_5$	$S_6$	$S_7$	$S_8$	$S_9$	$S_{10}$
$u_1$	8,8,8	8,6,8	7,6,9	8,7,9	8,8,9	8,7,9	8,6,9	8,6,9	6,7,8	7,8,8
$u_2$	7,8,7	6,7,8	6,7,6	6,7,6	6,7,8	6,8,7	7,8,8	7,6,8	6,6,9	6,6,8
$u_3$	7,6,8	6,7,8	5,7,8	5,6,7	6,8,7	5,8,6	5,7,7	6,7,6	6,6,8	5,8,7
$u_4$	7,7,9	8,5,7	8,6,7	9,6,8	7,5,8	6,6,6	6,5,6	8,5,7	5,5,6	6,5,6
$u_5$	8,7,8	7,8,7	8,8,8	8,7,7	8,8,7	6,6,6	7,6,7	9,7,6	8,5,8	7,7,7
$u_6$	8,6,7	6,7,8	8,8,7	7,7,8	7,8,6	6,7,8	7,7,8	8,8,7	6,8,7	9,8,7
$u_7$	8,8,6	9,6,7	8,6,6	8,7,5	9,7,7	9,7,9	9,8,6	8,6,9	8,7,8	8,8,8
$u_8$	7,8,7	6,8,88	6,7,4	7,7,6	7,9,5	7,8,8	8,8,6	7,8,8	7,8,7	9,3,8
$u_9$	7,8,7	6,8,8	6,7,6	7,7,6	7,7,9	7,6,8	8,7,7	7,6,7	7,7,8	9,6,8
$u_{10}$	9,7,7	9,7,8	7,5,9	7,6,8	5,9,5	5,8,7	9,6,4	4,5,7	9,9,6	8,3,7

4.1 历史用户评价分类及恶意用户剔除

Step1:利用式 1 计算各服务属性的平均值,得到属性的均值向量,这里以属性 1 为例:

$$e_1^1 = \frac{8 + 7 + 7 + 7 + 7 + 8 + 8 + 8 + 8 + 7 + 9}{10} = 7.7$$

依次求取  $e_2^1$  到  $e_{10}^1$  的平均值为 $\langle 7.1, 7.0, 7.4, 6.7, 6.7, 7.5, 6.9, 6.6, 7.3 \rangle$ ,则服务属性 1 的均值向量  $\boldsymbol{E}^1$  为 $\langle 7.7, 7.1, 7.0, 7.4, 6.7, 6.7, 7.5, 6.9, 6.6, 7.3 \rangle$ ,类似地,可以得到  $\boldsymbol{E}^2$  为 $\langle 7.3, 6.9, 6.7, 6.7, 7.6, 7.1, 6.8, 6.6, 6.8, 6.2 \rangle$ ,  $\boldsymbol{E}^3$  为 $\langle 7.4, 7.7, 7.0, 7.0, 7.1, 7.5, 6.7, 7.4, 7.5, 7.4 \rangle$ 。

Step2:利用式 2、3,对用户初始向量  $\boldsymbol{F}_i^k$  与对应的基准向量  $\boldsymbol{E}^k$  做差异计算,将评价属性分类。

属性与数据  $\boldsymbol{F}_1^1 - \boldsymbol{E}^1 = \langle 0.3, 0.9, 0, 0.6, 1.3,$

$1.3, 0.5, 1.1, -0.6, -0.3 \rangle$ ,依次求取  $\boldsymbol{H}_2^1, \dots, \boldsymbol{H}_{10}^1, \boldsymbol{H}_2^1 = \langle -0.7, -1.1, -1, -1.4, -0.7, -0.7, -0.5, 0.1, -0.6, -1.3 \rangle$ ,  $\boldsymbol{H}_3^1 = \langle -0.7, -1.1, -2, -2.4, -0.7, -1.7, -2.5, -0.9, -0.6, -2.3 \rangle$ ,  $\boldsymbol{H}_4^1 = \langle -0.7, 0.9, 1, 1.6, 0.3, -0.7, -1.5, 1.1, -1.6, -1.3 \rangle$ ,  $\boldsymbol{H}_5^1 = \langle 0.3, -0.1, 1, 0.6, 1.3, -0.7, -0.5, 2.1, 1.4, -0.3 \rangle$ ,  $\boldsymbol{H}_6^1 = \langle 0.3, -1.1, 1, -0.4, 0.3, -0.7, -0.5, 1.1, -0.6, 1.7 \rangle$ ,  $\boldsymbol{H}_7^1 = \langle 0.3, 1.9, 1, 0.6, 2.3, 2.3, 1.5, 1.1, 1.4, 0.7 \rangle$ ,  $\boldsymbol{H}_8^1 = \langle 0.3, -1.1, 0, 1.6, -2.7, 2.3, 1.5, -2.9, -1.6, 0.7 \rangle$ ,  $\boldsymbol{H}_9^1 = \langle -0.7, -1.1, -1, -0.4, 0.3, 0.3, 0.5, 0.1, 0.4, 1.7 \rangle$ ,  $\boldsymbol{H}_{10}^1 = \langle 1.3, 1.9, 0, -0.4, -1.7, 1.5, -2.9, 2.4, 0.7 \rangle$ 。假定阈值设置为  $\alpha = 2, \beta = 3$ ,结合第三节属性评价分类规则,  $\boldsymbol{H}_1^1, \boldsymbol{H}_2^1, \boldsymbol{H}_4^1, \boldsymbol{H}_5^1, \boldsymbol{H}_6^1, \boldsymbol{H}_9^1$ ,这些向量的所



有元素全部都在  $-2 \sim 2$  之间,所以这些向量所对应的用户评价属性:  $F_1^1, F_4^1, F_5^1, F_6^1, F_9^1$  是中立评价;  $F_3^1$  的元素都在  $-3 \sim -2$  之间,且存在元素在  $-3 \sim -2$  之间,所以  $F_3^1$  是挑剔评价;  $F_7^1$  的元素在  $-2 \sim 3$  之间,且存在元素在  $2 \sim 3$  之间,所以  $F_7^1$  是宽容评价;  $F_8^1$  中有的元素大于 2,有的元素小于  $-2$ ,所以  $F_8^1$  是恶意评价,同理,  $F_{10}^1$  是恶意评价。因为  $u_8, u_{10}$ , 存在恶意评价,所以  $u_8, u_{10}$  是不可信用户,可以直接将其剔除。

属性二:经过上段的筛选,现有 8 个可信用户。类似属性 1 的计算方法,可得到  $F_1^2, F_2^2, F_3^2, F_5^2, F_6^2, F_7^2, F_9^2$  是中立评价;  $F_4^2$  是挑剔评价。

属性三:类似属性 1 的计算方法,可得到  $F_2^3, F_3^3, F_4^3, F_5^3, F_6^3, F_7^3, F_9^3$  是中立评价;  $F_1^3$  是宽容评价。

#### 4.2 修正偏离评价数据

只需要修正宽容评价以及挑剔评价的数据。按上节分类结果,需要修正用户 1 的属性 3,用户 3 的属性 1,用户 4 的属性 2,用户 7 的属性 1。利用式 3 求出偏离度  $d$  (保留一位有效数字)。

用户 1 属性 3:

$$d_1^3 = [(0.25 + 0.16 + 3.61 + 4 + 1.96 + 2.25 + 3.61 + 2.56 + 1.44 + 2.56)/10]^{1/2} = 1.5$$

修正后数据  $G_1^3 = F_1^3 - d_1^3 = \langle 6.5, 6.5, 7.5, 7.5, 7.5, 7.5, 7.5, 6.5, 6.5 \rangle$ 。

用户 3 属性 1:

$$d_3^1 = [(0.25 + 1 + 4 + 1.69 + 8.41 + 2.56 + 0.25 + 4.41 + 5.29 + 4.52)/10]^{1/2} = 1.8$$

修正后数据  $G_3^1 = F_3^1 - d_3^1 = \langle 8.8, 7.8, 6.8, 6.8, 7.8, 6.8, 6.8, 7.8, 7.8, 6.8 \rangle$ 。

用户 4 属性 2:  $d_4^2 = 1.5$ ,  $G_4^2 = \langle 8.5, 6.5, 7.5, 7.5, 6.5, 7.5, 6.5, 6.5, 6.5, 6.5 \rangle$ ; 用户 7 属性 1:  $d_7^1 = 1.4$ ,  $G_7^1 = \langle 6.6, 7.6, 6.6, 6.6, 6.6, 7.6, 7.6, 7.6, 6.6, 6.6 \rangle$ 。

#### 4.3 基于修正数据的服务选择方法

剔除了两个恶意用户后,这里还剩下 8 个用户,基于这 8 个用户计算服务的 QoS 综合评分值。计算服务  $s_1$  各属性的均值所构成的基准向量为:  $e_1^1 = \frac{8 + 7 + 8.8 + 7 + 8 + 8 + 6.6 + 7}{8} = 7.6, e_1^2 = 7.4,$

$e_1^3 = 7.1$ 。所以候选服务  $s_1$  的基准向量  $C_1 = \langle e_1^1, e_1^2, e_1^3 \rangle = \langle 7.6, 7.4, 7.1 \rangle$ , 然后再计算各可信用户对服务  $s_1$  的评价向量与服务  $s_1$  的基准向量  $C_j$  之间的相似度。

$$e_1^1 \text{ 与基准评价向量的相似度为: } \text{sim}(P_{1,1}, C_1) = \frac{1}{1 + \sqrt{\frac{(8 - 7.6)^2 + (8 - 7.4)^2 + (6.5 - 7.1)^2}{3}}} =$$

0.6, 依次求取其他用户相似度并求和得:  $0.6 + 0.7 + 0.5 + 0.4 + 0.6 + 0.5 + 0.6 = 4.4$ 。计算得到服务  $s_1$

的最终 QoS 评分值:  $\text{qos}_1 = \frac{0.6}{4.4} * 22.5 + \frac{0.7}{4.4} * 22 + \frac{0.5}{4.4} * 22.8 + \frac{0.4}{4.4} * 24.5 + \frac{0.6}{4.4} * 23 + \frac{0.5}{4.4} * 21 + \frac{0.5}{4.4} * 20.6 + \frac{0.6}{4.4} * 22 = 22.2$ 。类似地,分别对候选服务  $s_1 \cdots s_{10}$  进行计算,得出其对应的最终 QoS 评分值:  $\text{qos}_2 = 24.6, \text{qos}_3 = 24.2, \text{qos}_4 = 21.1, \text{qos}_5 = 22.3, \text{qos}_6 = 27.3, \text{qos}_7 = 21.1, \text{qos}_8 = 21.5, \text{qos}_9 = 20.9, \text{qos}_{10} = 21.6$ 。将上述计算得到的  $\text{qos}_i$  降序排序,  $\text{qos}_2$  排在第一,所以将  $\text{qos}_2$  对应的服务  $s_2$  推荐给目标用户。

## 5 结束语

文中提出了基于评价分类的可信 QoS 服务选择的方法,充分考虑了用户个人评价习惯对服务评价的影响。首先基于历史用户对服务的反馈评价,将用户的评价详细分类。然后在评价分类的基础上剔除恶意用户,并进一步修正可信用户的非中立评价数据。最后在这些修正后数据的基础上,完成服务选择。通过实例验证表明,基于用户评价分类的可信 QoS 服务选择方法可以选择出优质的服务。

目前没有考虑用户对服务属性的喜好权重,今后的研究方向将进一步探讨如何结合机器学习来确定用户对服务属性的个性化权重。

#### 参考文献:

- [1] 林 闯,陈 莹,黄霁崑,等. 服务计算中服务质量的多目标优化模型与求解研究[J]. 计算机学报,2015,38(10): 1907-1923.
- [2] 刘飘悦. 基于 QoS 的 Web 服务可信性研究[D]. 杭州:浙江理工大学,2014.
- [3] LI Yang, XU Guoyan. A service selection algorithm based on the trust of data provenance[C]//12th international symposium on distributed computing and applications to business, engineering & science. Kingston upon Thames, Surrey, UK: IEEE,2013.
- [4] 王尚广,孙其博,杨放春. Web 服务选择中信誉度评估方法[J]. 软件学报,2012,23(6): 1350-1367.
- [5] 张龙昌,杨艳红. 动态 QoS 数据驱动的可靠 Web 服务选择[J]. 电子与信息学报,2016,38(6): 1368-1376.
- [6] 李金忠,夏洁武,唐卫东,等. 基于 QoS 的 Web 服务选择算法综述[J]. 计算机应用研究,2010,27(10): 3622-3627.
- [7] LUO Yihang, FAN Yushun, GUO Xiaohui. Trust assessment of business service workflow based on complex trust network[J]. 高技术通讯:英文版,2015,21(1): 22-30.
- [8] 贺兴亚,王海艳,杨文彬. 一种 QoS 可信增强的服务选择方法[J]. 武汉大学学报:理学版,2013,59(5): 443-448.
- [9] HUSSAIN S, WANG Zhaoshun, TOURE I K. An approach

workers[C]//Proceedings of the 30th annual conference on advances in cryptography. Santa Barbara, CA, USA: [ s. n. ], 2010:465-482.

[ 6 ] BACKES M, FIORE D, REISCHUK R M. Verifiable delegation of computation on outsourced data[C]//Proceedings of the 2013 ACM SIGSAC conference on computer & communications security. Berlin, Germany: ACM, 2013:863-874.

[ 7 ] 任艳丽, 丁 宁, 王天银, 等. 可完全验证的双线性对运算外包算法[J]. 中国科学信息科学, 2016, 46(7):855-869.

[ 8 ] 蒋铁金, 任艳丽. 基于单个服务器的双线性对运算外包算法[J]. 计算机应用, 2016, 36(7):1866-1869.

[ 9 ] 黄春水, 任艳丽, 蔡建兴. 可验证模指数批计算外包方案[J]. 西安电子科技大学学报:自然科学版, 2016, 43(4):135-140.

[ 10 ] FIORE D, GENNARO R. Publicly verifiable delegation of large polynomials and matrix computations with applications [ C ] // Proceedings of the 2012 ACM conference on computer and communications security. Raleigh, North Carolina, USA: ACM, 2012:501-512.

[ 11 ] ZHANG Liangfeng, SAFAVINAINI R. Private outsourcing of polynomial evaluation and matrix multiplication using multilinear maps[C]//International conference on cryptography and network security. [ s. l. ]: [ s. n. ], 2013:329-348.

[ 12 ] 任晓霞, 黄宏宇. 安全高效的大矩阵行列式计算云外包协议[J]. 计算机工程与应用, 2014, 50(10):82-86.

[ 13 ] 申银杰. 可验证的安全矩阵行列式计算云外包协议[J]. 计算机与现代化, 2015, 21(5):103-106.

[ 14 ] LI Hongwei, ZHANG Shenmin, LUAN T H, et al. Enabling efficient publicly verifiable outsourcing computation for matrix multiplication[C]//International telecommunication networks and applications conference. Sydney, NSW, Australia: IEEE, 2015:44-50.

[ 15 ] LEI Xinyu, LIAO Xiaofeng, HUANG Tingwen, et al. Outsourcing large matrix inversion computation to a public cloud [ J ]. IEEE Transactions on Cloud Computing, 2013, 1(1):1.

[ 16 ] WANG Cong, REN Kui, WANG Jia, et al. Harnessing the cloud for securely outsourcing large-scale systems of linear equations[J]. IEEE Transactions on Parallel & Distributed Systems, 2013, 24(6):1172-1181.

[ 17 ] ZHANG Jian, YANG Yang, WANG Zhibo. Outsourcing large-scale systems of linear matrix equations in cloud computing[C]//IEEE international conference on parallel and distributed systems. Wuhan, China: IEEE, 2016:438-447.

[ 18 ] SHENG Gang, TANG Chunming, GAO Wei, et al. MD-VC<sub>Matrix</sub>: an efficient scheme for publicly verifiable computation of outsourced matrix multiplication [ C ] // International conference on network and system security. [ s. l. ]: [ s. n. ], 2016:349-362.

[ 19 ] LINDELL Y, PINKAS B. Secure multiparty computation for privacy-preserving data mining [ J ]. Journal of Privacy & Confidentiality, 2008, 25(6):761-766.

+++++

(上接第 118 页)

for QoS measurement and web service selection sureness [ J ]. 高技术通讯:英文版, 2013, 19(3):283-289.

[ 10 ] HU Chunhua, CHEN Xiaohong, LIANG Ximing. Dynamic services selection algorithm in Web services composition supporting cross-enterprises collaboration[J]. 中南大学学报:英文版, 2009, 16(2):269-274.

[ 11 ] KIM Y, DOH K G. Quantitative trust management to support QoS-aware service selection in service-oriented environments[C]//International conference on parallel and distributed systems. Seoul, South Korea: IEEE, 2013:504-509.

[ 12 ] MEHDI M, BOUGUILA N, BENTAHAR J. A QoS-based trust approach for service selection and composition via Bayesian networks [ C ] // International conference on web services. Santa Clara, CA, USA: IEEE, 2013:211-218.

[ 13 ] 熊润群, 罗军舟, 宋爱波, 等. 云计算环境下 QoS 偏好感知的副本选择策略[J]. 通信学报, 2011, 32(7):93-102.

[ 14 ] 唐朝刚. 基于 QoS 的服务选择问题的研究[D]. 合肥:中国科学技术大学, 2012.

[ 15 ] 姚建华, 吴加敏, 牛温佳, 等. 基于可信 QoS 聚类的遥感服务发现机制[J]. 计算机应用, 2013, 33(2):587-591.

[ 16 ] 刘昕民, 桂卫华, 潘迪宏, 等. 一种基于 D-S 证据理论的 QoS 可信度评估方法[J]. 哈尔滨工业大学学报, 2013, 45(3):96-101.

[ 17 ] 杨丹榕. 基于可信度计算与 QoS 相结合的 Web 服务组合优化问题研究[D]. 南京:东南大学, 2016.