

一种基于服务质量预测的 Web 服务选择方法

刘克非,王 红,许作萍

(山东师范大学 信息科学与工程学院,山东 济南 250014)

摘 要:服务质量评估是基于 QoS 服务选择的前提和基础。然而,目前的 Web 服务发现方法在服务质量评估过程中没有充分考虑服务质量的动态变化,不能对服务的表现进行准确的预测。在分析现有相关研究的基础上,提出了一种基于服务质量预测的 Web 服务选择方法,并详细介绍了该服务选择方法的实现和服务质量预测模型的结构。

关键词:Web 服务;服务选择;服务质量;BP 神经网络

中图分类号:TP311

文献标识码:A

文章编号:1673-629X(2007)08-0103-03

A Web Service Selection Mechanism Based on QoS Prediction

LIU Ke-fei, WANG Hong, XU Zuo-ping

(College of Information Science and Engineering, Shandong Normal University, Jinan 250014, China)

Abstract: QoS evaluation plays a key role in QoS-based service selection. However, current QoS-based service selection methods haven't taken the dynamic nature of QoS into consideration and lack mechanisms to provide dynamic service performance prediction effectually. In this paper, propose a Web service selection mechanism which is based on QoS prediction. The architecture of QoS-prediction based service selection framework and the QoS prediction model are discussed in this paper in detail.

Key words: Web service; service selection; QoS; BP NN

0 引言

服务质量(QoS)是 Web 服务性能评价的重要指标,是 Web 服务能否在商业领域成功应用的关键因素。随着网络上 Web 服务的大量增加,特别是当面对多个功能上相同或相似的候选服务时,如何动态地选择出最能满足用户需求的服务是服务发现领域中的重要问题^[1]。对于该问题,在功能匹配基础上,考虑 Web 服务的服务质量是一个有效的解决途径。

文献[2]提出了一种服务发现模型,该模型对 UDDI 进行了扩展,增加了 QoS 属性描述,在服务选择过程中考虑了服务质量。文献[3]提出了一种支持 QoS 约束的 Web 服务发现模型(WSDM-Q),该模型引入了 QoS 量化概念,定义了一组描述服务 QoS 属性及信誉度的分类 tModel。文献[4]在客户端中引入推理机,通过分析其他用户反馈的服务质量信息实现基于 QoS 的服务选择。文献[5]设计了一种基于服务质量的轻

量级 Web 服务描述语言 QWSDL,并引入服务质量相似函数来选择满足特定 QoS 要求的 Web 服务。

基于 QoS 服务选择的前提是对 Web 服务的 QoS 进行准确的评估预测。由于互联网的动态性、不确定性和一些其它原因,Web 服务的服务质量是动态变化的。上面文献提到的方法在服务质量评估过程中没有充分考虑服务质量的动态变化,缺乏对服务质量进行动态评估的机制,不能对服务的表现进行准确的预测。为了克服上述局限性,文中提出一种利用 BP 神经网络对服务质量进行动态预测的方法。这里的预测是对服务质量中反映服务表现的非确定性属性^[6](如执行时间等)进行预测,确定性的属性(如服务价格)通常由服务提供者声明。

1 基于 BP 神经网络的服务质量预测

1.1 BP 神经网络介绍

BP 神经网络^[7](Back Propagation NN)是基于 BP 算法的多层前馈型神经网络,它具有输入层、隐含层(一层或多层)和输出层,其前后层之间各神经元完全连接,每层神经元之间无连接。图 1 是一个典型的三层结构 BP 神经网络。

收稿日期:2006-10-31

基金项目:山东省优秀中青年科学家奖励基金资助项目(03bs009)

作者简介:刘克非(1980-),男,山东济宁人,硕士研究生,主要研究方向为 Web 服务发现、P2P 技术;王 红,博士,教授,主要研究方向为 Web 服务、移动 Agent 等。

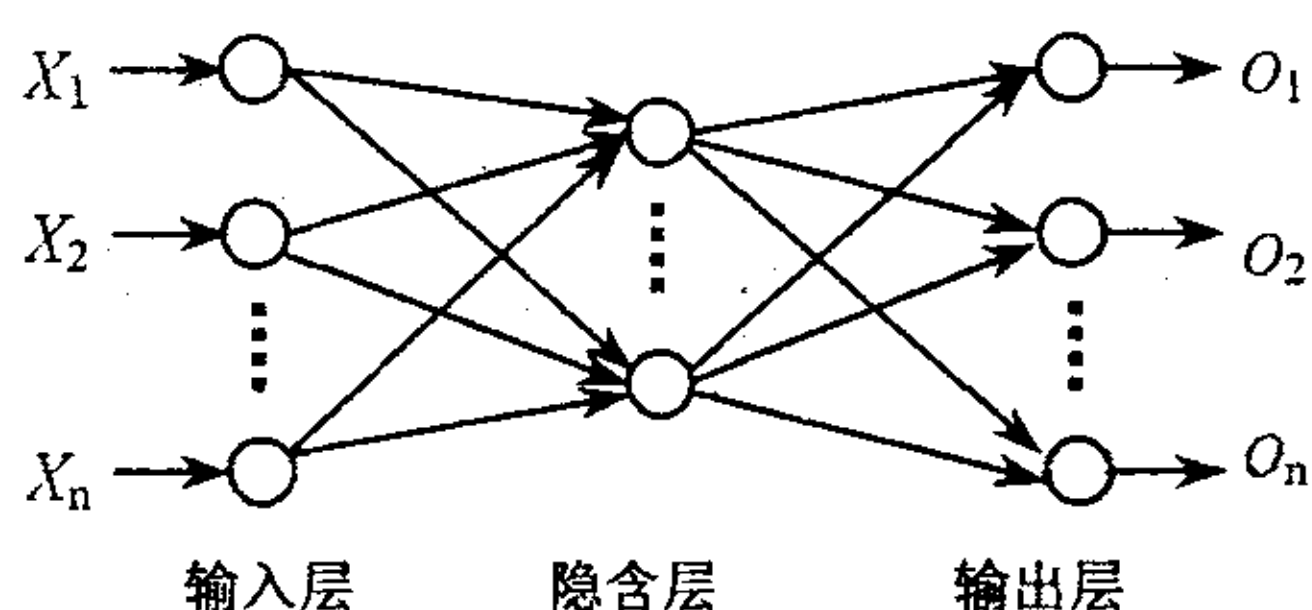


图 1 三层结构 BP 神经网络

BP 算法是一种有教师的学习方法,其学习过程分为两个阶段:第一阶段(正向传播),给定输入信息通过输入层经隐含层计算实际输出值;第二阶段(误差反传),如果输出信号与期望信号有差别,即存在误差,则逐层递归地计算各层的误差,并根据各层误差调节各层的权值。

BP 算法是建立在梯度下降法基础上的。设一 BP 网络,第 \$k\$ 层的第 \$j\$ 个神经元的输入为 \$I_j^k\$,输出为 \$O_j^k\$, \$k-1\$ 层的第 \$i\$ 个神经元与 \$k\$ 层的第 \$j\$ 个神经元的权连接为 \$W_{ji}^{k,k-1}\$,则

$$I_j^k = \sum_i W_{ji}^{k,k-1} O_i^{k-1}$$

$$O_j^k = f(I_j^k)$$

对于任一输入模式,若网络的整体误差为 \$E\$,输出层第 \$j\$ 个神经元的期望输出 \$d_j\$,根据梯度下降法得:

$$\Delta W_{ji}^{k,k-1} = -\eta \frac{\partial E}{\partial W_{ji}^{k,k-1}}, \eta > 0 \text{ 为学习系数,限于篇幅,}$$

文中不继续推导,直接给出结论:

$$\Delta W_{ji}^{k,k-1} = -\eta \delta_j^k O_i^{k-1}$$

$$\text{其中 } \delta_j^k = \begin{cases} (O_j^k - d_j) f'(I_j^k), & \text{当 } k \text{ 为输出层} \\ (\sum_m \delta_m^{k+1} W_{mj}^{k+1,k}) f'(I_j^k), & \text{当 } k \text{ 为中间层} \end{cases}$$

1.2 QoS 预测

BP 神经网络本质上是一种输入到输出的映射,它通过学习大量样本,使网络具有某种输入与输出间的映射能力,而不需要任何输入输出间精确的数学表达式。为了对某一 Web 服务的服务质量进行预测(主要指服务表现方面),我们收集其他用户使用该服务后的反馈信息作为训练样本训练 BP 网络,然后把训练好的 BP 网络用于以后服务发现过程中对该 Web 服务的服务质量、服务表现进行评估预测。在下一节给出的基于 QoS 预测的服务选择框架中,为每个 Web 服务构造一个 BP 网络用于服务质量预测。下面给出了 QoS 预测模型结构(predict model schema),我们将在仿真实验中给出一个包含具体输入输出参数的服务质量预测模型。

```
//QoS predict model schema
```

```
<QoS prediction>
```

```
<QoS_prediction_input>
```

```
<! A set of QoS input parameters>
```

```
</QoS_prediction_input>
```

```
<QoS_prediction_output>
```

```
<! A set of QoS output Parameters>
```

```
</QoS_prediction_output>
```

```
</QoS_prediction>
```

QoS 预测结果加上其他 QoS 参数(如服务价格等)得到一个服务完整的 QoS 信息,然后根据用户的个性化需求(比如有的用户关注服务价格而有的用户更关注服务表现)计算出一个总的 QoS 值,用户根据每个候选服务的这个 QoS 值选择出最满意的服务^[5]。

2 基于 QoS 预测的服务选择框架

如图 2 所示,基于 QoS 预测的服务选择框架主要包括三个部分:服务管理中心(Service Management Center, SMC), QoS 预测代理(QoS Prediction Broker, QoS PB)和 QoS 预测服务(QoS Predict Service, QoS PS)。图中每个虚线框中包含一个 Service 和一个 QoS PS,这意味着每个 Web 服务需要一个特定的 QoS PS 为其提供服务质量预测。图中 Service1 到 Service N 代表候选服务。

服务管理中心(SMC)维持一个列表文件,该列表记录每个 service 和其 QoS PS 的对应关系。列表中的每个记录包含三个元素:service URL, QoS PS URL 和 service state。文中定义三种服务状态(service state):活动(Active),挂起(Suspend)和永久失效(Dead),并由 SMC 监视服务状态的变化。

QoS 预测服务(QoS PS)用来完成服务质量的预测,其作为一个 Web 服务来实现。如图 2 所示,每个 QoS PS 对应一个 Web 服务,为该 Web 服务提供服务质量预测。每个 QoS PS 包含一个训练样本集和一个 BP 神经网络,这里的训练样本集是 Web 服务的用户反馈信息。随着用户对该服务反馈信息的增加, QoS PS 定期更新训练样本集,并用最新的样本集重新训练 BP 网络。BP 网络从 QoS PB 获取服务质量预测的输入信息,然后将预测结果返回给 QoS PB。

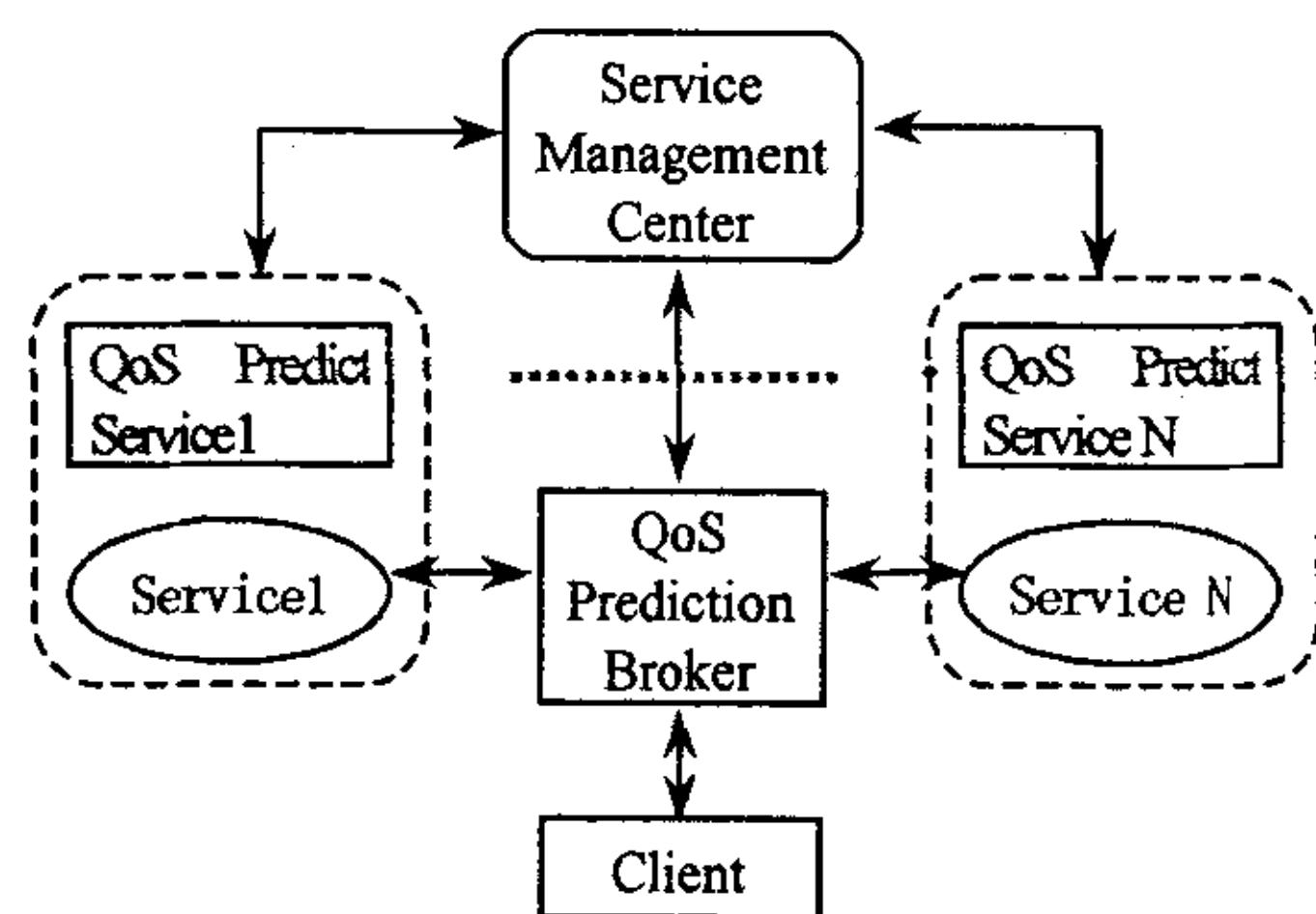


图 2 基于 QoS 预测的服务选择框架

QoS 预测代理(QoS PB)收集服务质量预测所需

的必要信息提供给 QoS PS, 并从各 QoS PS 接收候选服务的 服务质量预测结果。在此基础上, 结合用户个性化需求, 计算出最能满足需求的服务返回给用户。同时, QoS PB 监视服务的实际执行情况, 并把这些实际执行信息反馈给 QoS PS, 以便更新其训练样本集重新训练 BP 网络。

下面给出一个简单的例子来说明基于服务质量预测的服务选择过程。假设, 当一个用户查询 UDDI 并返回了多个功能相同的候选服务时:

- 1) 用户的 QoS 预测代理首先查询 SMC, 获得各候选服务的状态, 并查询每个活动的 (Active) 候选服务对应的 QoS PS。
- 2) 分析每个 QoS PS 的服务质量预测模型和输入接口以便为每个 QoS PS 提供必要的输入信息。
- 3) QoS PS 把输入信息提交给各自的 BP 神经网络进行服务质量预测。
- 4) 每个 QoS PS 的预测结果返回给 QoS 预测代理, 然后结合用户个性化需求, 计算出最能满足需求的服务返回给用户。
- 5) 最终的 Web 服务选定后, QoS 预测代理监视服务的实际执行情况, 并把这些实际执行信息反馈给 QoS PS, 以便更新其训练样本集重新训练 BP 网络。

3 仿真实验

1.2 节给出了 QoS 预测模型的结构 (predict model schema), 下面结合一个具体的 QoS 预测模型来说明基于 BP 神经网络的服务质量预测。我们将根据给出的服务质量预测模型构造 BP 神经网络, 然后利用一组训练样本训练该 BP 网络, 在 BP 网络训练好后, 用新的输入来检验该 BP 网络的预测效果。

下面首先给出一种服务质量预测模型:
predict model = (predict_input, predict_output);
predict_input = (Avail, Relia, Bandw, RTime);
predict_output = (EDu, TranS)。

可用性 (Availability, Avail) 用来衡量当前 Web 服务的系统负载能力, 可由 QoS PB 访问服务提供者获得。设 Web 服务的服务器并发访问的最大会话数为 λ_{max} , 当前可用会话数为 λ , 则当前服务的可用性为:

$$Avail = \frac{\lambda}{\lambda_{max}} \times 100\%$$

可靠性 (Reliability, Relia) 表示 Web 服务正常运行的概率, 定义为成功执行次数与执行总次的比率, 即:
$$Relia = \frac{s}{s + f}$$
, 其中 s 表示成功执行次数, f 为失败执行次数。

带宽 (Bandwidth, Bandw) 衡量用户的网络状况。
请求时间 (Request time, RTime) 表示用户请求特定服务的时间。

执行时间 (Execution duration, EDu): 用户从提交服务请求到获得结果的时间, 包括服务时间和往返通信时间。训练样本中 EDu = 1 表示服务请求超时 (见表 1)。

执行状态 (Transaction state, TranS) 表示 Web 服务是否成功执行, 1 代表成功, 0 代表失败。

表 1 训练样本集

| | predict_input | | | | predict_output | |
|----|---------------|-------|-------|-------|----------------|-------|
| | Avail | Relia | Bandw | RTime | EDu | TranS |
| 1 | 0.60 | 0.65 | 0.912 | 0.6 | 0.0821 | 1 |
| 2 | 0.67 | 0.66 | 0.627 | 0.8 | 0.2811 | 1 |
| 3 | 0.82 | 0.72 | 0.827 | 0.8 | 0.11891 | 1 |
| 4 | 0.23 | 0.66 | 0.420 | 0.8 | 0.44841 | 0 |
| 5 | 0.67 | 0.72 | 0.399 | 0.6 | 0.04979 | 1 |
| 6 | 0.78 | 0.71 | 0.659 | 0.8 | 0.25471 | 1 |
| 7 | 0.75 | 0.31 | 0.670 | 0.6 | 1 | 0 |
| 8 | 0.85 | 0.69 | 0.861 | 0.8 | 0.0916 | 1 |
| 9 | 0.55 | 0.68 | 0.883 | 0.3 | 0.13791 | 1 |
| 10 | 0.83 | 0.78 | 0.138 | 0.8 | 0.67681 | 1 |
| 11 | 0.71 | 0.79 | 0.406 | 0.6 | 0.4926 | 1 |
| 12 | 0.70 | 0.74 | 0.722 | 0.6 | 0.23581 | 1 |
| 13 | 0.50 | 0.73 | 0.264 | 0.3 | 0.63891 | 1 |
| 14 | 0.80 | 0.67 | 0.397 | 0.8 | 0.46741 | 1 |
| 15 | 0.20 | 0.73 | 0.522 | 0.8 | 0.3663 | 0 |
| 16 | 0.75 | 0.79 | 0.433 | 0.8 | 0.4379 | 1 |
| 17 | 0.63 | 0.27 | 0.532 | 0.6 | 1 | 0 |
| 18 | 0.69 | 0.73 | 0.628 | 0.6 | 0.3126 | 1 |

根据该 Web 服务的服务质量预测模型, 选择 4-6-2 结构的 BP 网络, 激活函数选用 Sigmoid 函数 $f(x) = 1/(1 + \exp(-x))$, 学习系数 $k = 0.02 (0 < k < 1)$, 初始权值取 0~1 之间的随机数。表 1 给出了实验中的训练样本集, 该样本集是原始数据经过归一化处理后的数据。对于请求时间 (RTime) 的归一化处理, 为了简单起见, 实验中将一天 24 小时划分为三个时段: 0.8 代表 0:00~8:00, 0.6 代表 18:00~0:00, 0.3 代表 8:00-18:00。当系统误差 $C = 0.003$ 时, 网络训练 152219 次。表 2 给出了部分样本预测值与实际值的比较, 可以看出, 利用 BP 神经网络能够较好地对服务质量进行预测评估。

表 2 部分预测值与实际值比较

| | EDu | | TranS | |
|---|--------|--------|-------|-----|
| | 预测值 | 实际值 | 预测值 | 实际值 |
| 1 | 0.7020 | 0.6968 | 0.999 | 1 |
| 2 | 0.5179 | 0.5168 | 0.999 | 1 |
| 3 | 0.9986 | 1 | 0.001 | 0 |
| 4 | 0.3049 | 0.3042 | 0.999 | 1 |
| 5 | 0.2396 | 0.2412 | 0.006 | 0 |
| 6 | 0.6519 | 0.6589 | 0.999 | 1 |

(下转第 109 页)

(2)数据库操作:将 DataSet 中的解析结果作为参数,构建 SQL 命令,通过 Command 对象对数据源执行命令。

(3)填充 DataSet:将查询结果生成数据集。

(4)关闭数据连接。

(5)在 DataSet 上进行所需要的操作:DataSet 是任意数目的表在内存中的表示形式,以及这些表 and 任何约束之间的关系,因此在断开与数据库的连接之后仍可独立地对数据结果进行访问。

(6)根据需要可将 DataSet 的变化更新到数据库,或自定义 SQL 语句,直接通过 Command 更新数据库。

若需要向请求方返回操作结果,则将结果信息按指定格式封装为 XML,由数据发送模块发送到相应站点。

4) 发送模块:它与接收模块一起实现与外部数据的联系,不同的是发送模块负责将信息封装成 XML 数据,通过网络套接字发送到指定站点。

5 结 语

该系统中各站点每个站点都可以拥有相互独立的 DBMS,故各站点既可以自主管理自己的内部数据,又可以通过网络传输实现对外部异构数据的有条件共享。通信双方通过 ADO.NET 对 XML 数据进行组织和解析,并根据解析情况执行相应的业务处理,可方便地实现 XML 数据与后台管理数据库的关系型结构数据间的相互转换,从而使系统具有较大的灵活性。系统中还可以 XML 格式传递辅助系统任务的配置信

息,以保证相应站点数据的同步更新,便于数据任务的统一管理。该系统在结构上便于向上扩展,实现更大范围的数据集成与共享。目前,该方法已在云南某钢厂得到成功应用。

笔者创新点:利用 ADO.NET 作为数据访问和数据转换的渠道,将传统数据库技术与 XML 相结合,有效地实现异构数据库的交互。同时提出将数据访问、传输与转换的方法应用于系统辅助信息,在企业应用中有助于数据任务的统一管理,更加灵活地实现企业间异构数据库数据的共享。

参考文献:

- [1] 黄乐丹. XML 技术与数据库技术的比较及应用[J]. 温州职业技术学院学报, 2005, 5(1): 46 - 48.
- [2] Wildermuth S. ADO.NET 实用指南——面向 Internet 世界的数据库访问技术[M]. 周靖译. 北京:清华大学出版社, 2003.
- [3] Beauchemin B. ADO.NET 本质论[M]. 周靖译. 北京:清华大学出版社, 2003.
- [4] Esposito D, Wintellect. ADO.NET 中的 XML 特性[EB/OL]. 2004 - 01 - 01. <http://www.microsoft.com/china/MSDN/library/data/dataAccess/XMLFeaturesinADO.NET.aspx?mfr=true>.
- [5] 朱敏,王开建,苏博. 基于 XML 的企业网络数据集成模型研究[J]. 微计算机信息, 2005(8): 21 - 23.
- [6] Dick K. XML: 管理者指南——构筑 XML 的企业级应用[M]. 邓尚贤译. 北京:清华大学出版社, 2003.
- [7] 尹立宏. Visual C#.NET 应用编程 150 例[M]. 北京:电子工业出版社, 2003.

(上接第 105 页)

4 结束语

由于互联网的动态性、不确定性和一些其它原因,导致 Web 服务的服务质量不断变化,这给基于 QoS 的服务发现提出了挑战。文中结合 BP 神经网络提出了一种基于服务质量预测的 Web 服务选择方法,并详细介绍了基于 QoS 预测的服务选择框架和服务质量预测模型。下一步工作将继续完善服务质量预测模型,优化 BP 网络的训练算法,研究更合理有效的训练样本集更新策略。

参考文献:

- [1] Vu L H, Hauswirth M, Aberer K. QoS - based Service Selection and Ranking with Trust and Reputation Management [C]//In: Proceedings of the International Conference on Cooperative Information Systems (CoopIS 2005). Agia Napa,

Cyprus: [s. n.], 2005: 466 - 483.

- [2] Ran S. A model for web services discovery with QoS[J]. ACM SIGecom Exchanges, 2003, 4(1): 1 - 10.
- [3] 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.
- [4] Day J, Deters R. Selecting the best web service[C]//In: Proceedings of the 14th Annual IBM Centers for Advanced Studies Conference (CASCON). Canada: [s. n.], 2004: 293 - 307.
- [5] 胡建强,邹鹏,王怀民,等. Web 服务描述语言 QWSDL 和服务匹配模型研究[J]. 计算机学报, 2005, 28(4): 505 - 513.
- [6] Liu Y, Ngu A H, Zeng L Z. QoS computation and policy in dynamic web service selection[C]//In: Proceedings of the 13th international World Wide Web conference. New York, USA: [s. n.], 2004: 66 - 73.
- [7] 袁曾任. 人工神经网络及其应用[M]. 北京:清华大学出版社, 1999: 66 - 68.