

DOI:10.3876/j.issn.1000-1980.2015.06.014

云计算中服务质量预测数据的信心建模

张雪洁^{1,2}, 王志坚¹, 张伟建³

(1. 河海大学计算机与信息学院, 江苏 南京 210098; 2. 南京航空航天大学计算机科学与技术学院, 江苏 南京 210016;
3. 河海大学远程与继续教育学院, 江苏 南京 210098)

摘要: 为处理服务质量(quality of service, QoS)预测所用数据的不确定性,增加预测结果的信心值,使预测的 QoS 值更可信,建立了量化 QoS 预测中信心的概率模型。构建模型过程中,考虑了预测所用的 QoS 数据项数量、数据的波动情况(数据偏差)以及数据随时间的衰减情况。数据项数量表明参与预测的数据多少对预测结果可信度的影响程度;数据偏差表明服务的实际 QoS 值和预期值的一致程度;数据衰减程度表明随时间变化,数据对预测结果的影响程度。仿真试验表明,该信心模型能够准确有效地帮助用户选择满足其需求的服务。

关键词: 云计算; QoS 预测; 信心建模; 数据量; 数据波动; 数据衰减; 服务选择

中图分类号: TP391 **文献标志码:** A **文章编号:** 1000-1980(2015)06-0588-06

Confidence model of QoS prediction data in cloud computing

ZHANG Xuejie^{1,2}, WANG Zhijian¹, ZHANG Weijiang³

(1. College of Computer and Information, Hohai University, Nanjing 210098, China;

2. College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China;

3. College of Distance Learning and Continuing Education, Hohai University, Nanjing 210098, China)

Abstract: In order to handle the uncertainty of data used in quality of service (QoS) prediction, increase the confidence value of prediction results, and make the QoS prediction more reliable, a probability model for quantifying the confidence in QoS prediction was built. In the process of building the model, the number of QoS data items used in prediction, the data fluctuation (data deviation), and the data decay over time were considered. The results show that the number of data reflects the impact of the number of QoS data used in prediction on the reliability of the prediction result, the data deviation reflects the consistency degree of the actual value and predicted value of QoS in service, and the data decay reflects the impact of data on the prediction result over time. The simulated test indicates that the confidence model can help consumers effectively select services based on their requirements.

Key words: QoS prediction; confidence modeling; number of data; data fluctuation; data decay; services selection

服务质量(quality of service, QoS)预测可以广义地描述为使用服务的历史数据确定用户从提供商获得的可能的服务质量。随着云计算的广泛应用,网络上功能相同的服务越来越多,想用这些服务的用户希望用到质量最好、最符合自己需求的服务。对用户来说,评估所有服务是不现实的,所以需要借助预测 QoS 的技术和工具,预测 QoS 的能力就显得尤为重要。

在云计算环境中,由于受网络拥堵、资源限制或质量管理水平等因素的影响^[1-4],服务提供商的行为不总是期望的那么稳定,所以预测的性能就会有波动。这些不确定性引起的服务性能的变化会反映在收集的

收稿日期: 2014-12-10

基金项目: “十二五”国家科技支撑计划(2013BAB05B00, 2013BAB06B04); 江苏水利科技项目(2013025); 河海大学淮安研究院开放基金(2014502512)

作者简介: 张雪洁(1979—),女,辽宁铁岭人,博士研究生,工程师,主要从事服务计算、Web 服务质量评估与推荐研究。E-mail: xuejie_zh@hhu.edu.cn

QoS 数据中。根据变化很大的数据预测 QoS 很可能是不可靠的,而根据少量数据预测 QoS 也不可信。此外,数据收集的时间因素也会影响预测的可信度,根据新近数据预测的结果更有说服力。所以,需要对 QoS 预测结果附加一个信心值,这样预测的 QoS 更可信,能为用户选择服务提供更可靠的指示。

为了量化 QoS 预测中的信心,笔者建立了考虑 3 个可靠性度量的信心模型,并展示信心模型能够有效帮助用户根据他们的期望选择服务。3 个可靠性度量为:(a) 预测中需要用到的所有 QoS 数据项数量;(b) 数据项中数据值的变化情况;(c) 数据随时间的衰减情况。

1 信心模型相关工作

目前,许多信心模型已被提出,特别是在信任和信誉系统,这些模型的主要目标是在 QoS 预测时处理用户行为的一些不确定性。Sabater 等^[5]介绍了 2 种方法(评价的数量、评价值的偏差)用于计算信任值的可靠性,这 2 种方法得到了 Keung 等^[6]、Huynh 等^[7]、Dossari 等^[8]的认同。Mui 等^[9]用 Chernoff 范围确定获得某一等级的信心需要的最小样本大小,Zhang 等^[10]用该方法计算代理的经验是否足够可靠,即经验的可靠程度,代理经验由样本大小度量。上述研究已表明在产生信心值时考虑数据特征的重要性,而数据特征由数据项数量和数值的偏差来度量。但上述信心模型不能直接应用到笔者的工作中,原因有 2 个:(a) 这些模型主要是针对离散数据,而本文处理的是连续数据,即监测数据;(b) 预测结果的可靠性是“静止地”推导出来的,忽略了用户的期望,而期望对个性化预测很重要,是选择满足用户需求服务的不可少因素。

为确定一个指定服务的预测是否可信,必须根据用户需要的质量“动态地”做出预测。为此,笔者提出一个考虑预测数据的大小、数据的波动以及数据随时间衰减的信心模型。这样,可以“动态地”推断出预测结果的可信度。该模型能适当地改进已有模型的不足,有效处理 QoS 预测中存在的一些不确定性。

2 信心模型的建立

基于 Huynh 等^[7]、Zhang 等^[10]的研究,本文信心模型用 3 个度量(数据大小、数据偏差和数据衰减程度)计算对某个 QoS 属性预测的信心。数据大小表明用于预测的数据集支持 QoS 预测方法进行预测的强度,即参与预测的数据多少对预测结果可信度的影响程度。数据偏差表明服务的实际 QoS 值和预期值的一致程度。数据衰减程度表明随时间变化,数据对预测结果的影响程度。

2.1 数据大小度量

数据大小度量是基于预测时所用数据项数量进行的。预测时所用的每个数据项都是一个服务过去服务质量的证据。拥有证据越多,对评估的信心越高,即随着数据项数量的增加,可信度增加。本文采用式(1)形式化表示度量方法,根据数据项数量计算得到可信度,式(1)表明随着数据项数量 n 的增加,可靠程度增加,直到达到定义的阈值(需要的最小数据项数量) m :

$$C_{\omega} = \begin{cases} \frac{n}{m} & n < m \\ 1 & n \geq m \end{cases} \tag{1}$$

其中

$$m = -\frac{1}{2\varepsilon^2} \ln \frac{1-\lambda}{2}$$

式中: C_{ω} ——数据规模的度量值(当 n 从 0 增加到 m 时其从 0 增加到 1,当 n 超过 m 时其值保持 1, m 根据边界定理^[9]计算得到); ε ——用户可以接受的最大误差等级; λ ——需要的信心等级。

λ 越大, ε 越小, m 值越大,说明需要的数据项数量越多。例如,如果设 $\lambda = 0.99$ 、 $\varepsilon = 0.1$,则需要的最小数据项数 $m = 1\,060$,这意味着要对预测结果有信心,就要尽可能多地使用 QoS 数据项记录。

2.2 数据偏差度量

数据偏差度量是基于预测中所用数据的变化进行的。服务选择是根据历史性能数据预测出的 QoS 值来决定的,选择用于预测的数据时考虑数据的偏差很重要。因为分布不同的数据可能计算得到同样的平均值,但它们变化不同,如,一个服务的某个 QoS 属性值(如,响应时间)变化比较大,另一个服务的响应时间相对稳定,但二者的平均值可能相同。而数据变化越大,计算出的平均值越不可靠,也就是说,根据波动大的数据预测出的结果可靠性低。

为计算可靠性,笔者把服务提供的 QoS 值看作伯努利(Bernoulli)试验集合,交付结果分为:成功的交付(交付的服务满足用户需求的质量)或不成功的交付(不能交付满足用户需求质量的服务),然后用 Beta 分布建模^[11-12]:

$$C_{\theta} = \frac{\alpha}{\alpha + \beta} \tag{2}$$

其中

$$\alpha = s + 1 \qquad \beta = f + 1$$

式中: C_{θ} ——数据偏差的度量值; s ——监测的成功交付的服务数量; f ——不成功的服务数量。

根据拉普拉斯继承规则^[13], α 和 β 的比率可以确定 $[0,1]$ 区间的分布峰值, α 越大分布模式越接近 1。采用式(3)来判断一个服务对一个 QoS 属性的历史交付是否成功:

$$S(A_i) = \begin{cases} 1 & d(A_i) \in [p_i - \theta, p_i + \theta] \\ 0 & \text{其他} \end{cases} \tag{3}$$

式中: $S(A_i)$ ——记录一个服务提供的服务质量是否满足用户需求的参数; $d(A_i)$ ——QoS 数据集 D 中服务属性 A_i 的质量值,是监测到的服务实际提供的质量值; p_i ——对 QoS 属性 A_i 预测的结果,根据用户需求计算得到; θ ——用户可接受的质量等级波动值。

如果 $d(A_i)$ 在 p_i 指定范围内,则认为是成功的交付,记 $S(A_i) = 1$;否则,认为不成功,记 $S(A_i) = 0$ 。范围 $[p_i - \theta, p_i + \theta]$ 称为信心范围。基于式(3)可得

$$s = \sum_{S \in D} (S(A_i) = 1) \qquad f = \sum_{S \in D} (S(A_i) = 0) \tag{4}$$

选择不同的 θ 对 C_{θ} 有直接影响,同样的数据大小, θ 值越大,则会产生更多认为是成功的交付,因此对 QoS 预测结果的信心越大、 C_{θ} 值越大。所以,用户接受服务提供商提供的服务 QoS 值波动大时要选择大的 ε 值。

与已有方法^[9-12]相比,笔者提出的信心模型中一个服务是否交付成功是动态确定的,因为 p_i 是根据用户需求计算的,所以 $d(A_i)$ 是否在信心范围内受用户需求的影响,也就是说,判断一个服务是否成功是由用户的需求影响的。例如,设 $\theta = 0.1$,假设服务的交付值和需求的质量等级足够近才算满足用户需求,即 $[p_i - \theta, p_i + \theta]$ 是一个小范围。如果一个服务 S 对属性 A_i 的交付质量是 0.5,那么对于对 A_i 的期望是 0.5 的用户是成功的,但对于需求是 0.9 的用户就是不成功的。

2.3 数据衰减度量

监测的 QoS 数据在时间上是连续的,数据中包含的用于预测的信息会随着时间的推移而发生变化。在实际应用中,距离预测时间点最近的数据所蕴含的信息往往要比历史数据有价值。因此,根据不同时间段的数据预测出的结果其可靠性不同。为此,对预测所用数据在预测中的权重进行度量,为预测的可信度做出指示。

应用时间衰减函数(time decay function, TDF)来度量逐步衰减的数据对预测的支持程度^[14-15]。监测的 QoS 数据对预测的支持度随时间 t 逐步衰减,衰减函数为

$$f(t) = 2^{-\mu t} \qquad 0 < \mu \leq 1 \tag{5}$$

式中: t ——预测时间点与数据监测时间点的时间差; μ ——数据在一个时间单元内的衰减比率,即衰减因子。

μ 是算法中的重要参数,它反映了历史数据对预测结果的影响程度,有助于提高预测结果的可信度。当 μ 减小时,衰减速度变慢,同一时间点的 QoS 数据对预测结果的影响程度也相应提高。如果 μ 太小,数据衰减不明显,数据衰减对预测的影响就很小,即久远的数据和新近的数据效果差不多,所以 μ 太小不会影响预测结果的可信度。 μ 太大,数据对预测的影响迅速衰减,同样会影响预测结果的可靠性。所以,应该在取值范围内尽可能大地设置 μ 。

设用户预测所用的最少数据项数量为 m ,第 m 个数据的记录时间和预测时间差记为 t_m ,用户所能接受的最小衰减值为 $f_{\min}$,理想状态这个值出现在第 m 个数据,那么 $f_{\min} = 2^{-\mu t_m}$,即 $f_{\min} \leq f(t) \leq 1$ 。所以, μ 的取值范围为

$$0 \leq \mu \leq \frac{-\log_2 f_{\min}}{t_m} \tag{6}$$

总的的数据衰减度量值为

$$C_f = \frac{\sum_{t_i \in [t_1, t_n]} N(t_i)f(t_i)}{\sum_{t_i \in [t_1, t_n]} N(t_i)}$$

(7)

式中: $N(t_i)$ ——预测数据中距离预测时间点 t_i 时间的数据项数量。

2.4 总体信心值

数据规模度量 C_ω 表示 QoS 预测值受数据集支持的程度,数据偏差度量 C_θ 表示监测的服务 QoS 值和平均值的一致程度,数据衰减度量 C_f 表示服务的 QoS 值随时间的衰减程度。根据这 3 个度量,计算总的信心值为

$$C = C_\omega C_\theta C_f$$

(8)

3 仿 真 试 验

试验的目的是验证信心模型是否有助于提高服务选择的准确性。

3.1 试验数据

试验模拟 6 个服务: S_1 、 S_2 、 S_3 、 S_4 、 S_5 和 S_6 ,6 个服务的测试数据见表 1,数据分布如图 1 所示。6 个服务的 QoS 值的模拟值为正态分布,平均值相同,都为 0.5,但数据的大小、数据的波动(受标准偏差 σ 控制)和数据收集的时间跨度不同。创建的 6 个服务模拟了服务的不同行为。

表 1 6 个服务的测试数据

Table 1 Test data of six services

服务	数据项数/条	时间跨度/s	数据说明
S_1	1000	0~1000	均值为 0.5,波动幅度为 0.25,表示数据一直波动比较大
S_2	1000	0~1000	均值为 0.5,波动幅度为 0.05,表示数据一直波动比较小
S_3	1000	0~1000	前 500 条数据均值为 0.5,波动幅度为 0.05,波动比较小;后 500 条数据均值为 0.5,波动幅度为 0.25,波动比较大
S_4	1000	0~1000	前 500 条均值为 0.5,波动幅度为 0.25,数据波动比较大;后 500 条均值为 0.5,波动幅度为 0.05,数据波动比较小
S_5	500	0~1000	均值为 0.5,波动幅度为 0.05,数据一直波动比较小,数据量比较少
S_6	500	0~500	均值为 0.5,波动幅度为 0.05,数据一直波动比较小,数据量比较少,数据时间跨度较小

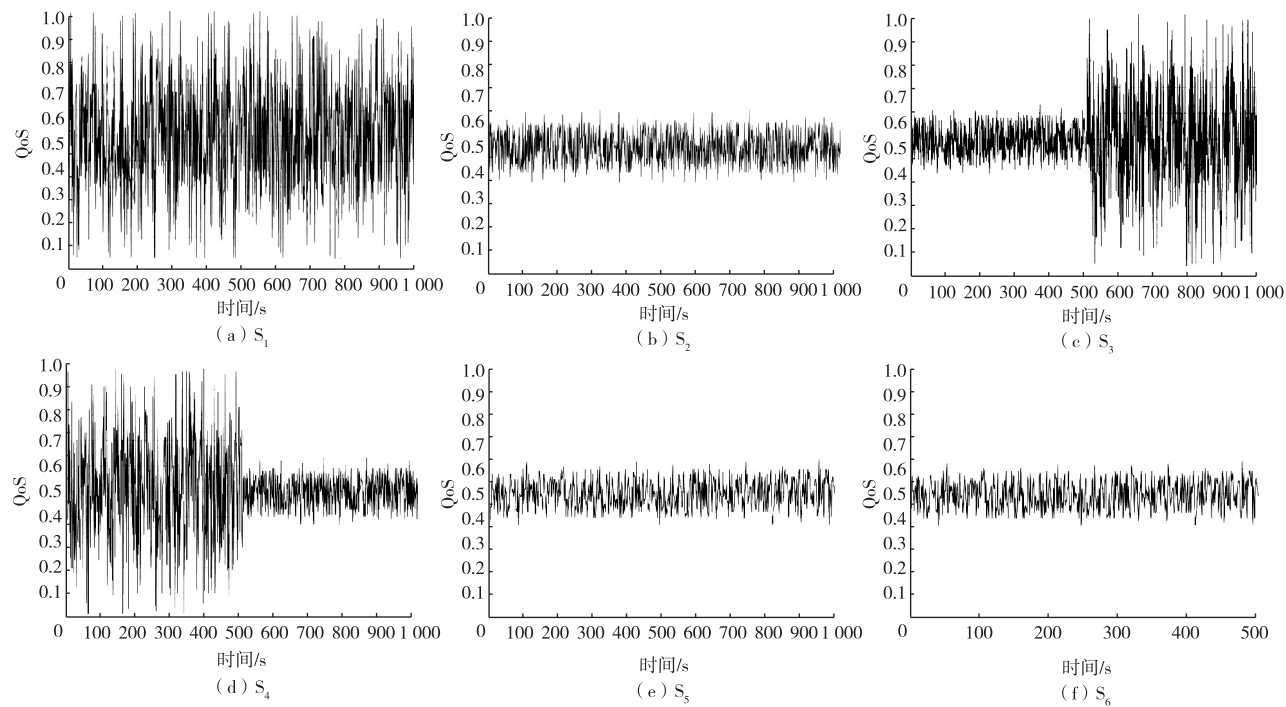


图 1 服务数据分布

Fig. 1 Distribution of services data

3.2 信心在服务选择中的作用

QoS 预测的目的是帮助用户选择满足他们需求的服务。利用信心模型情况下的服务选择情况来验证本文提出的信心模型是否能提高服务选择的准确性。本文提出的信心模型可以和任意 QoS 预测方法结合,帮助提高预测的可靠度。

图 2 为不考虑信心模型情况下的服务选择。由于 6 个服务的历史 QoS 数据的平均相似,预测的 QoS 值也相似,所以对服务做出随机选择。虽然 S_1 偏离服务的需求等级很多,但被选择了 3 次,有 33.3%的时间选择它,大于平均时间。而和用户需求最一致的服务 S_2 只被选择 1 次。

考虑信心模型的方法刚开始数据量小的时候在服务 S_4 和 S_6 之间随机选择,之后一直稳定地选择服务 S_2 。由式(8)分别计算 6 个服务的信心值,6 个服务信心值随时间/QoS 数据量的变化情况如图 3 所示。从图 3 可知服务的行为表现可以分为 2 个阶段,第一阶段(时间在 0~500 s 之间), S_2 、 S_4 、 S_5 和 S_6 的行为一直很接近,它们的预测信心值也比较接近,所以 QoS 预测方法选择有更高信心值的服务。第二阶段(时间在 500~1 000 s 之间), S_4 、 S_5 和 S_6 的预测信心值开始下降或保持不变。 S_5 和 S_6 用于预测的数据量比较少,没有更多的数据用于预测,所以预测信心值不会增加。 S_5 的数据跨度比较大,所以数据衰减比较大,其预测的信心值开始下降。 S_4 在这个阶段用于预测的数据波动比较大,所以预测的信心值较上一个阶段急速下降,之后稳定在 0.25 左右。只有 S_2 的预测信心值一直稳步上升,所以将选择 S_2 。

与不考虑信心的试验相比,这里一直没有选择 S_1 、 S_3 和 S_5 。因为 S_1 的 QoS 值一直比较波动,意味着 S_1 的 QoS 平均值不太可靠。 S_3 在第一阶段的波动比较大, S_5 的数据衰减一直比较大,所以都没有被选择。试验表明,在云计算环境中把信心作为 QoS 预测的一部分,能更好地帮助用户选择满足其需求的服务。

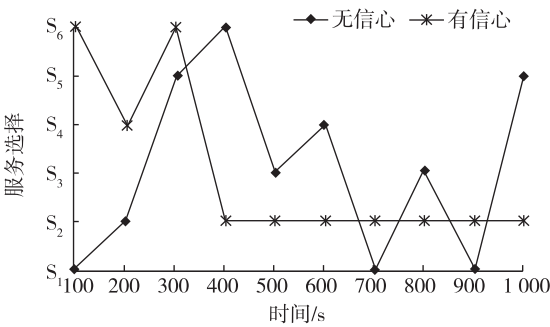


图 2 有无信心模型的服务选择比较
Fig. 2 Comparison between service selections with and without confidence model

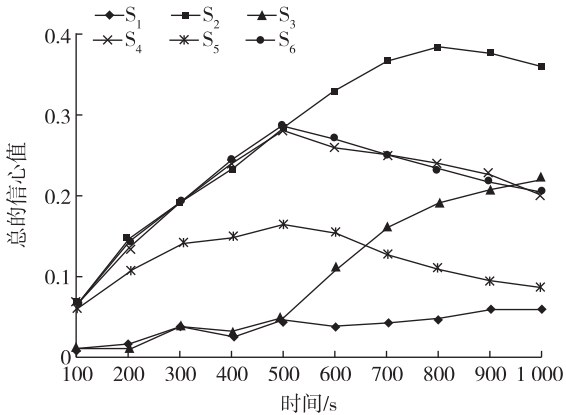


图 3 6 个服务的信心值
Fig. 3 Confidence values of six services

4 结 语

为量化 QoS 预测中的信心,提出了一个概率模型,该模型集成了 3 个可靠性度量:预测中所用 QoS 数据项的数量、数据偏差和数据衰减。该模型中,数据大小度量表明数据集或过去的证据对 QoS 预测方法推导出预测的支持程度,数据偏差度量表明监测的服务 QoS 值和平均值的一致程度,数据衰减度量表明服务的 QoS 值随时间的衰减程度。最后,通过试验验证了在预测过程增加信心模型能够更准确地选择满足用户需求的服务,增加用户的总体效用。目前的信心模型中只考虑了监测的 QoS 数据,接下来可以进一步扩展模型,加入预测服务时考虑的其他因素,如服务提供商的信誉等。此外,还可以利用衰减函数去度量随时间推移服务提供商的行为变化。

参考文献:

[1] SAHAI A, OUYANG J, MACHIRAJU V, et al. Specifying and guaranteeing quality of service for web services through real time measurement and adaptive control [R]. Palo-Alto: EServices. Software Research Department, HP Laboratories, 2001.

[2] KIM D, LEE S, HAN S, et al. Improving web services performance using priority allocation method [C]//Anonymous. Proceedings of International Conference on Next Generation Web Services Practices. Los Alamitos: IEEE Computer Society, 2005: 201-206.

[3] JURETA I J, HERSSENS C, Stéphane F, et al. A comprehensive quality model for service-oriented systems [J]. Software Quality Journal, 2009, 17: 65-98.

[4] WOHLSTA E, TAI S, Mikalsen T, et al. Glueqos: middleware to sweeten quality-of-service policy interactions [C]//Anonymous. Proceedings of the 26th International Conference on Software Engineering. Washington DC: IEEE Computer Society, 2004:189-199.

[5] SABATER J, SIERRA C. REGRET: reputation in gregarious societies [C]//Anonymous. 5th International Conference on Autonomous Agents and Multiagent systems. Hakodate: ACM Press, 2006: 194-195.

[6] KEUNG S N L C, GRIFFITHS N. Using recency and relevance to assess trust and reputation[C]//Anonymous. Proceedings of AISB 2008 Symposium on Behaviour Regulation in Multi-agent Systems. Aberdeen: The Society for the Study of Artificial Intelligence, 2008:13-18.

[7] HUYNH T D, JENNINGS N R, SHADBOLT N. R fire: an integrated trust and reputation model for open multi-agent systems [C]//Anonymous.16th European Conference on Artificial Intelligence, Valencia: IOS Press, 2004: 18-22.

[8] DOSSARI A H, SHAO, J H. Modelling confidence for quality of service assessment in cloud computing [C]//Anonymous. Proceedings of CONF-IRM. Natal: Berkeley Electronic Press, 2013: 48-59.

[9] MUI L, MOHTASHEMI M, HALBERSTADT A. A Computational model of trust and reputation[C]//Anonymous.35th Annual Hawaii International Conference on System Sciences. Hawaii: IEEE Computer Society, 2002: 2431-2439.

[10] ZHANG J, COHEN R. A personalized approach to address unfair ratings in multiagent reputation systems[C]//Anonymous. AAMAS Workshop on Trust in Agent Societies. Hakodate:Springer. 2006: 89-98.

[11] TEACY W T L, PATEL J, JENNINGS N R, et.al. Coping with inaccurate reputation sources: experimental analysis of a probabilistic trust model[C]//Anonymous. 4th International Joint Conference on Autonomous Agents and Multiagent Systems. Utrecht: Association for Computing Machinery, 2005: 25-32.

[12] SANG A J, ISMAIL R. The Beta Reputation System[R]. Brisbane: Queensland University of Technology, 2002.

[13] RISTAD E S. A natural law of succession; Research Report CS-TR-495-95[R]. Princeton: Princeton University, 1995.

[14] 李国徽,陈辉. 挖掘数据流任意滑动时间窗口内频繁模式[J]. 软件学报,2008, 19(10): 2585-2596. (LI Guohui, CHEN Hui. Mining the frequent patterns in an arbitrary sliding window over online data streams[J]. Journal of Software, 2008, 19 (10): 2585-2596. (in Chinese))

[15] 吴枫, 仲研, 吴泉源. 基于时间衰减模型的数据流频繁模式挖掘[J]. 自动化学报, 2010, 36(5): 674-684. (WU Feng, ZHONG Yan, WU Quanyuan. Mining frequent patterns over data stream under the time decaying model[J]. Acta Automatica Sinica, 2010, 36(5): 674-684. (in Chinese))

· 简讯 ·

河海大学杰出校友陆佑楣院士获世界工程组织联合会优秀工程奖奖章

近日,在世界工程组织联合会(WFEO)举办的2015年世界工程师大会(WEC 2015)上,河海大学杰出校友、我国著名的水利水电工程专家、中国大坝协会荣誉理事长、河海大学博士生导师陆佑楣院士被授予优秀工程奖(Medal of Engineering Excellence)奖章。这是中国专家首次获此殊荣。

优秀工程奖又称工程成就奖(the Engineering Achievement Award),由世界工程组织联合会设置于1989年,每两年颁发一次,旨在增强全球公众对工程的实践、理论和社会贡献的关注。该奖项主要颁发给具备丰富职业经验,在实践、理论和社会地位方面有突出成就、并具有较强国际影响力的工程师。

(本刊编辑部供稿)