

基于云计算的大数据服务资源评价方法

阳小兰^{1,2} 钱 程^{1,2} 朱福喜²

(武昌理工学院信息工程学院 武汉 430223)¹ (武汉大学计算机学院 武汉 430072)²

摘 要 随着大数据服务领域引入云计算技术,需要调动的云服务资源增多且其拓扑结构变得复杂,因此传统基于服务质量(QoS)的加权评价方法无法动态地评价云计算服务资源的有效性和准确性。针对此问题,文中提出了一种基于博弈优化调度的筛选加权评价方法。此方法引入了用户的体验质量(QoE)评价指标,充分考虑了动态调度的业务和时延特性,通过多个指标的博弈,得到加权评价的参数的纳什均衡点。仿真实验结果表明,所提评估方法能够准确地评价云计算服务资源调度的有效性和准确性,并且适合大数据服务业务的拓展。

关键词 博弈优化,加权评价,云计算,资源调度

中图分类号 TP393 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2018.05.051

Evaluation Method of Big Data Service Resources Based on Cloud Computing

YANG Xiao-lan^{1,2} QIAN Cheng^{1,2} ZHU Fu-xi²

(School of Information and Engineering, Wuchang University of Technology, Wuhan 430223, China)¹

(School of Computer, Wuhan University, Wuhan 430072, China)²

Abstract With the introduction of cloud computing technology in the field of big data services, the traditional QoS-based weighted evaluation methods can not evaluate the validity and accuracy of cloud computing service resources dynamically due to the large number of cloud service resources which need to be mobilized and their complex topological structures. In order to solve this problem, this paper proposed a screening weighted evaluation method based on game optimization scheduling. This method introduces user's QoE evaluation index, fully considers the service and delay characteristics of dynamic scheduling, and gets the Nash equilibrium point of weighted evaluation parameters through the game of multiple indexes. The simulation results show that the proposed evaluation method can accurately evaluate the validity and accuracy of cloud computing service resource scheduling and is suitable for the expansion of big data service business.

Keywords Game optimization, Weighted evaluation, Cloud computing, Resource scheduling

1 引言

随着移动互联网的发展,人类的各种活动和需求以及周围环境均可以通过数据形式表示,故通过对数据进行分析可以得到许多以前无法得到的信息,为智能决策、智能分析和智能服务提供条件。大数据技术应运而生。大数据提供的业务具有多样性,但它们有一个统一的特点^[1-3],即需要调用分散数据,且数据量大,消耗的计算资源多。随着通信系统传输能力的增强,计算资源可以跨区域进行调度,云计算服务为大数据服务提供了有力的计算资源保障^[4-6]。

大数据服务可以分为两种形式,即大数据查询服务和大数据分析服务。其中大数据分析服务不但需要调度大量的数据资源,而且需要调度大量的计算资源;但云服务除了满足大数据所需要的数据和计算资源外,还需要满足不同业务所对

应的服务质量(Quality of Service, QoS),然而业务的 QoS 并不能全面地衡量用户的服务感受,因此将其改进为基于用户体验质量(Quality of Experience, QoE)的服务评价体系。在传统的云计算条件下,大数据业务的评价方法是以大数据业务的服务商提供的相关 QoS 记录为依据的,但这种加权评价方法存在两方面的缺点:1)服务商提供的 QoS 数据并不能准确反映用户得到服务的感受;2)简单的加权并不能反映各个节点对大数据服务的贡献度,同时也增加了计算复杂度。因此文献[7]提出了一种减小评价数据规模的加权评价方法,该方法虽然提高了评价效率,但同时也降低了评价的准确性。在大数据服务中,每个加载的任务均可被分解为多个子任务,能够提供子任务的计算和存储资源的云服务很多,但它们存在较大的 QoE 差异。如何确定 QoE 的评价参数,根据所提供服务的不同动态确定云服务的调度方案,以及对云服务调

到稿日期:2017-08-26 返修日期:2017-11-26 本文受湖北省自然科学基金(2014CFB356),国家自然科学基金(61272277)资助。

阳小兰(1978—),女,硕士,副教授,CCF 会员,主要研究领域为大数据技术、数据挖掘,E-mail:254958769@qq.com(通信作者);钱 程(1980—),男,硕士,副教授,主要研究领域为大数据技术、数据挖掘;朱福喜(1957—),男,教授,博士生导师,主要研究领域为大数据技术、数据挖掘、人工智能。

度的评估方法的设计,均会影响大数据业务的执行。根据大数据业务的特点,即需要调动大量的数据资源和计算资源,传统的调度和评价方式无法适应云计算环境动态的网络指标,且没有考虑各个云计算节点的分配博弈问题。博弈论^[8-10]通过多个控制处理单元来使用强化学习的马尔科夫过程,但由于执行时间有限,多个云计算控制处理单元无法进行完全的信息交互^[11],因此博弈策略采用非完整信息的非合作博弈方法,针对不同的 QoE 指标的相互博弈,最终达到纳什均衡点,此纳什均衡点可以得到最优的评估参数。由于各个 QoE 参数均以各方利益最大化为目标,不需要考虑参数间的相互联系,因此降低了算法的复杂度,从而可以提高算法的收敛速度^[12]。对于评估算法,需要兼顾云计算调度的动态特性,也需要考虑 QoE 在动态网络中的统计分布特性,以不断修正加权系数^[13-15]。

针对以上评价问题,本文提出了一种基于博弈论的加权评价方法。此方法对影响 QoE 的指标建立博弈模型,根据博弈论理论的纳什均衡理论,各个节点的控制不具有协作性,故需要对节点的指标建立非协作博弈模型,从而确定各个 QoE 指标的权值,达到快速、准确评估大数据服务 QoE 的目的。

随着移动边缘计算(Mobile Edge Computing, MEC)技术的发展,数据存储和计算调度的粒度逐渐减小,计算和存储效率更高,QoE 评价指标的分布会发生变化。由于本文采用统计性修正评价参数,以体现网络的动态特性,因此本文所提方法同样适用于云计算和移动边缘计算共存的调度网络,也说明了所提方法对后续技术的兼容性。

2 QoE 评价指标

传统云计算服务质量均以服务价格、可靠性以及时间延迟等指标作为评价标准,统称为云计算的 QoS。由于 QoS 并没有考虑最终接受服务的用户的体验,因此本节提出用户的 QoE 评价体系。此评价体系不仅需要考虑传统互联网中云计算的相关指标,还要充分考虑网络结构的变化,特别是针对移动互联网云计算服务,需要考虑其容错性和大面积网络故障等稳定性因素以及服务连接效率等相关因素。根据用户体验,同时考虑到大数据调度大量存储资源和计算资源的特点,QoE 评价体系总体分为业务特点评价指标、网络技术评价指标以及服务影响效果评价指标 3 类,本节将针对这 3 类评价指标进行分析,并将其细化为一个评价指标体系。

2.1 业务功能评价指标

在业务特点评价指标中,根据对业务功能的不同影响,其可以分为三大类:1)所提供业务的共同协作性;2)所提供业务的覆盖范围;3)所提供业务的鲁棒性。

对于所提供业务的协作性,主要从以下几个方面进行考虑。

1)所提供的数据信息的协作性。信息的协作性,从网络方面也体现了信息调用算法的合理性。信息调用算法越合理,其调用效率越高,协作性越强,云计算中各个存储节点与终端的信息协作效率也就越高。需要注意的是,由于移动互联网技术和移动边缘计算技术的大面积应用数据信息的协作性会大幅提高。此指标可以用符号 C_i 表示,一般用于衡量协

作性的指标为信息调度的时延指标,因此其应该为一个时间物理量,以毫秒为单位。

2)所提供的应用协作性。应用协作性是用户侧体验的重要指标,为一个综合性指数,主要关系到计算资源调度的效率、云平台代码执行的效率以及服务流程的优化程度。协作性指标是这 3 项指标的加权和,用 C_s 表示。考虑到移动云计算平台的特性,必须思考移动终端接入的方式,需要考查移动终端接收结果对整个流程的影响,因此其没有所对应的物理量单位,应该表示为:

$$C_s = E_p \cdot E_c \cdot O_f \quad (1)$$

其中, E_p 代表计算资源调度的效率; E_c 代表云平台代码执行的效率; O_f 代表服务流程的优化程度。

3)所提供服务的云平台协作性。云平台的协作性反映了云平台的硬件部署以及软件资源部署是否全面,并发传输的机制是否高效;云计算的关键数据核实用户的通信保障是否可靠,分布式计算任务的分配是否公平高效,以及云平台服务的共同协作和限制机制是否完善。云平台的协作性从宏观上可以分为硬件性能和软件性能,硬件性能主要通过存储速度和计算速度来衡量,软件性能主要分为平台软件 and 用户通信两部分。所提供服务的云平台的协作性指标 C_c 可以表示为:

$$C_c = H_c \cdot S_c \cdot T_c \quad (2)$$

其中, H_c 表示硬件协作性能指标, S_c 表示用于提供云服务的软件的协作性能指标, T_c 表示用于数据存储调用的通信协作性能指标。 H_c 所对应的物理量指标应该以时间为单位描述; S_c 所对应的物理量应以正常运行的概率来进行衡量; T_c 主要是衡量通信的物理指标,以频谱利用率和误码率为度量,因此可以表示为:

$$T_c = SE \cdot EP \quad (3)$$

其中, SE 代表通信的平均频谱利用率, EP 代表通信的平均误码率。

2.2 网络技术评价指标

网络技术评价指标主要从网络响应的相关参数、网络连接的相关参数以及系统稳定的相关参数 3 方面进行评价。

1)网络响应的相关参数。其可以是用户对云计算平台为大数据提供服务的响应时间的满意程度。用户的满意程度是一个相对概念,因此衡量用户对服务响应时间的满意程度的关键指标是最优响应时间和平均响应时间的比值,可以表示为:

$$R_s = \frac{T_b}{T_a} \quad (4)$$

其中, T_b 代表所对应服务的最优响应时间; T_a 代表所对应服务的平均响应时间。最终的满意程度由用户的平均值决定,而对于用户的平均值,需要根据平均响应时间来衡量,故第 i 个用户的满意程度为:

$$S_i = \frac{T_i}{T_a} S_t + R_s S_m \quad (5)$$

其中, S_t 代表服务满意程度的最高分值, S_m 代表服务满意程度的基础分值。

2)网络连接的相关参数。其主要分为 3 个指标:①云计算网络发生中断的概率 P_c ;②云计算网络的连接时延超过该服务规定时延的概率 P_r ;③云计算服务有效的网络通路连接

概率 P_L 。因此,网络连接的相关参数可以由一个指标来衡量,即综合性网络故障概率 P_N ,可由式(6)计算得到:

$$P_N = P_r \cdot (1 - P_L) \cdot P_N \quad (6)$$

3) 云服务系统稳定的相关参数。其可以根据整个云数据平台分为3个部分:①云平台服务可持续利用率 P_{sc} ;②云平台故障修复性能指标 D_{sc} ;③云平台维护的难易程度 M_{sc} 。这3个指标能够反映用户对云服务系统的稳定程度的感受。

2.3 用户操作体验指标

用户操作体验指标用于衡量大数据业务在云计算平台上是否便于用户操作,主要体现在4个方面。

1) 用户操作过程体验指标,可从3个方面对其进行衡量:①整个服务操作的简单程度,分为是否容易学习和是否容易操作;②操作达到用户的目标结果的概率;③系统服务的信息安全性能。

2) 信息安全性能指标,可以分为用户信息的安全性、用户需要数据的安全性、用户权限与其他等级权限的完备性。安全性能指标只能通过信息泄露概率来衡量,故记为 P_{leak} ;用户数据安全性主要是网络受攻击后再被侵入的概率,记为 P_{break} ;在用户权限设置方面,需要统计无权限设置概率,统计的概率记为 P_{blank} 。可以通过式(7)来衡量用户对大数据云平台服务的信息安全的感受:

$$P_s = (1 - P_{leak}) \cdot (1 - P_{break}) \cdot (1 - P_{blank}) \quad (7)$$

3) 软件平台的人机交互体验程度,包括软件平台的层次性、个性化设置选项的丰富度以及界面的友好程度。其很难用一个指标来衡量:软件平台的层次性可以根据软件平台目录层次的个数及性能来衡量;个性化设置选项的丰富程度可以根据需要选择的项数决定;界面的友好程度需要通过用户评测的方法进行统计。

基于以上分析,对 QoE 的衡量指标体系进行总结,结果如表1所列,其评价基于云平台的大数据服务指标。

表1 基于云平台的大数据服务指标 QoE 体系

Table 1 Big data service index QoE system based on cloud platform

指标名称	具体衡量指标
业务功能 评价指标	数据信息的协作性指标
	应用协作性指标
	云平台协作性指标
网络技术 评价指标	网络响应相关参数指标
	网络连接相关参数指标
	云服务系统稳定相关参数指标
用户体验 评价指标	用户操作过程体验指标
	信息安全性能指标
	人机交互体验程度指标

3 基于 QoE 多方博弈算法

3.1 基于 QoE 多方博弈模型

在多方博弈模型中,需要建立多方博弈参数集合,其表示为 $G = \{C_i, C_s, C_c, \dots, S_i, P_N, P_{sc}, D_{sc}, M_{sc}, P_s\}$,表示 QoE 的衡量指标,每个元素均可以在第2节中找到。由于每一个指标均对最后的评价结果有影响,因此需要将每个指标在具体的云计算平台上映射到博弈空间中,其映射规则为 $f: S_i^n \rightarrow X_i$,其中 S_i^n 代表第 i 个用户的指标状态,根据映射规则可以映射到第 i 个集合指标, n 是第 n 个时刻的指标状态。

多个指标之间存在着相互制约和相互影响的关系,随着时间的推移,可以得到博弈集合空间的观测数据符合有限的马尔科夫离散状态,故每个 QoE 指标参数均是一个动态变化的量。因此,在 QoE 的指标体系中,不同的加权系数也是动态变化的,故这是一个动态博弈优化问题,其优化流程如图1所示。

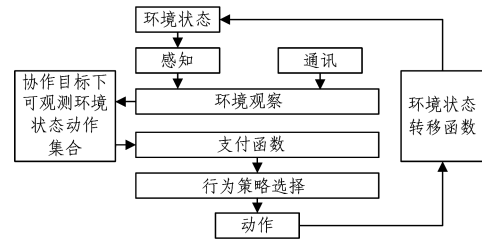


图1 动态博弈的优化流程

Fig. 1 Optimization process of dynamic game

动态博弈优化的核心策略为追踪博弈策略,得到的加权系数是一种最优组合集合,最优加权组合系数集合一旦形成,每一个参数所对应的加权系数将无法通过单独改变一个指标来增加整体的目标收益。通过以上分析可以看出,需要应用博弈理论来确定动态博弈优化中的追踪策略,从而达到最优的纳什均衡点。

对于追踪博弈策略,需要找到每个状态参数所对应的加权,根据此加权值改变函数关系,设每个加权值组成的集合为:

$$\lambda = \{\lambda_1, \lambda_2, \lambda_3, \dots, \lambda_n\} \quad (8)$$

其中, λ_i 表示第 i 个加权系数根据平台的动态变化而调整的数值,调整函数可以表示为:

$$\lambda_i \rightarrow \beta_i \quad (9)$$

其中, $i=1, 2, \dots, N$, β_i 为调整后的系数,系数调整只由两个方向组成,即增大和减小,因此较为简单。

由于 λ 是一个整体,因此每个元素均随着 λ 的变动而变动;由于系数集合的元素不断变化,其组成的状态空间也不同,因此需要进行状态映射,映射之后再行状态迁移函数的变换。状态迁移函数为:

$$\lambda \otimes G \rightarrow \beta \quad (10)$$

其中, $\beta = \{\beta_1, \beta_2, \dots, \beta_n\}$, $\otimes$ 表示向量的张量运算,从而张成了变化的空间。此数据需要映射到目标空间,对目标空间进行测量,从而得到目标函数最优时所对应的 β 张量空间。为了多样化 QoE 的目标函数,本节不针对某种特定的目标函数进行优化,而是将目标函数定义为:

$$\beta \otimes G \rightarrow D \quad (11)$$

对张成的空间进行优化,这个空间 D 是多个目标函数,其组成的集合为:

$$D = \{D_1, D_2, \dots, D_n\} \quad (12)$$

式(12)的集合方式更适合多目标优化问题,根据多个目标能够针对不同服务需求的用户体验进行修改,因此所建立的基于博弈的 QoE 评价目标具有鲁棒性。

3.2 基于博弈的系数优化算法

根据 3.1 节所示的博弈模型,可以得到优化的目标 D ,通过目标 D 可以得到多目标集合,通过优化 β ,得到最优的多目

标可行解,其求解过程如下。

Step1 确定 QoE 的优化目标,若已经确定,则转向 Step4;

Step2 如果 QoE 的优化目标中缺少衡量指标,则需要收集并计算缺少的衡量指标,转向 Step1;

Step3 如果确定在 QoE 的衡量体系中,某个指标不需要参与博弈,则去除该指标,转向 Step16;

Step4 初始化参数指标的状态,设置其时隙为 n ;

Step5 计算当前参数指标状态下的纳什均衡解;

Step6 各个加权系数根据纳什均衡解以及优化的目标空间来确定自己的变化方向;

Step7 各个参数指标观测其他参数指标的变化趋向,计算优化目标函数;

Step8 若满足目标函数优化,则转向 Step12;

Step9 各参数指标根据限制条件确定是否参与评价体系,若确定不参与,则转向 Step16;

Step10 根据参数变化的历史记录,确定各个参数向各方向变化的概率,并根据各方向概率的大小确定下个时隙各个参数的变化方向;

Step11 根据其他参数的变化方向,选择此参数的最佳变化方向,并转向 Step7;

Step12 判断优化的张量空间是否达到最优,若达到最优,则转向 Step15;

Step13 判断是否在规定的博弈时间内完成了优化任务,若未完成,则记为失败,转向 Step17;

Step14 $n=n+1$,下一个时隙按照 Step5 进行优化;

Step15 博弈结束,各个优化参数确定其对应加权系数的数值,转向 Step17;

Step16 进行博弈后的补偿和惩罚,博弈中各个 QoE 指标相互支付补偿;

Step17 更新各个参数的加权系数;

Step18 结束。

在 Step5 中,需要进一步说明纳什均衡解的求解方式。在求解纳什均衡的过程中,首先要确定 QoE 评价函数,其定义为:

$$V_i^k = \frac{1}{1+N} \times \beta \otimes G \quad (13)$$

根据此评价函数,可以得到各个参数指标所对应的加权参数变化获得的收益值,收益矩阵如下所示:

$$V_n = \begin{bmatrix} V_1^1 & V_1^2 & \cdots & V_1^n \\ V_2^1 & V_2^2 & \cdots & V_2^n \\ V_3^1 & V_3^2 & \cdots & V_3^n \\ V_n^1 & V_n^2 & \cdots & V_n^n \end{bmatrix} \quad (14)$$

根据矩阵博弈论,可以很容易求得此矩阵的最优纳什均衡解。

4 实验仿真结果

本节从时间复杂度的角度来验证所提出的云环境下大数

据服务的可信组合评估方法的可行性。随机布置 12000 个计算数据资源的节点,这些节点服从泊松独立点分布,而其间的连接和通信传输速率通过随机图的方式确定。云平台环境基于两台服务器,服务器的配置为 48 GB 内存,八核主频为 3.2 GHz 的中央处理器,并且使用不同的大数据业务对云平台的数据和计算资源进行调度。根据调度结果,采用所提 QoE 评价体系来评估考查的大数据业务,大数据业务主要分为数据查找业务和数据分析业务。在选择大数据服务方面,本文以执行最多的数据作为分析对象,根据 QoE 的评价体系,对各个评价指标进行了统计和记录,如表 2 所列。

表 2 基于云平台的大数据服务指标 QoE 体系的实验数据

Table 2 Experimental data of big data service index QoE system based on cloud platform

指标类别	具体指标名称	取值
业务功能评价指标	数据信息的协作性指标	10 ms
	应用协作性指标	0.857
	云平台协作性指标	0.763
网络技术评价指标	网络响应相关参数指标	82
	网络连接相关参数指标	0.92
	云服务系统稳定相关参数指标	86%
用户体验评价指标	用户操作过程体验指标	95
	信息安全性能指标	88
	人机交互体验程度指标	92

本文获取了各个指标值服从的概率分布,其中,网络连接的相关指标服从 $[0.78, 1]$ 的均匀分布,而其他指标均服从正态分布。图 2 给出了基于 QoS 中最常用的 Benchmark 最优组合评价方案与本文所提方案在时间消耗指标的对比。从图中可以看出,所提方案在评价指标体系较复杂的情况下,执行时间较少,说明所提方法的计算复杂度较小。因此,所提方法更适用于大数据服务的评价。

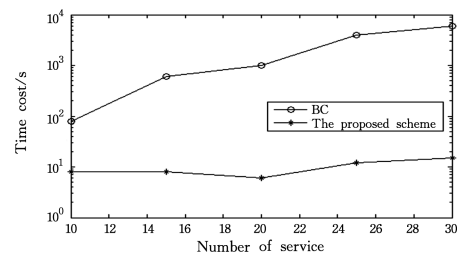


图 2 服务数量与时间消耗的关系

Fig. 2 Relationship between number of service and time cost

同时从图 2 中也可以得到,随着服务数量的增加,本文提出的评价方法所花费的时间基本保持稳定;而对比方案 BC 的时间波动较大,尤其是在服务数量增加的情况下时间消耗较多。

文献[12]提出了一种衡量组合评估方案的指标——相对评估值,其定义为:

$$V_{OP} = \frac{\bar{V}_n}{V_n} \quad (15)$$

其中, V_n 代表通过穷搜法得到的最优组合值, $\bar{V}_n$ 代表各种组合评估的平均最优组合值。

图 3(a)和图 3(b)给出了不同时间下,数据查询业务和数

据分析业务的 QoE 的相对评估值。可以看出,随着参考历史记录的增加,偏离程度有所不同,当历史记录为 15 和 20 时,评估值达到最优。

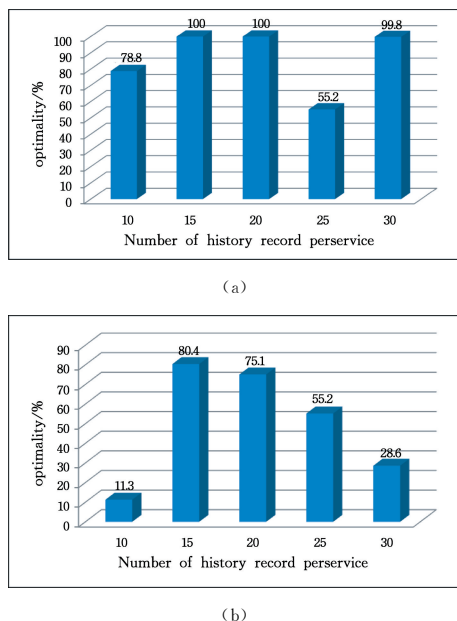


图3 所提组合方法的相对最优值

Fig. 3 Optimality of proposed method

此外,本文没有将所提方法与其他方法在 QoE 评估结果上的性能进行比较,这是因为本文提出的是全新的 QoE 评价体系,其他 QoS 方法的性能必定差于本文方案;并且,本文所提 QoE 方案更加科学地描述了用户的感受,能更加全面和直观地提高服务质量,也适应于 MEC 和移动互联网等新技术,以及新的网络拓扑结构。

结束语 本文根据大数据服务的特点及所调用的云平台数据资源多和计算资源分散等特点,提出了一种基于 QoE 的评价体系;根据评价体系,采用评价指标组合加权的形式分析用户的评价感受,并采用博弈论的方法对组合加权值进行求解,最终得到了纳什均衡解。此方法的计算复杂度低,评价效果好,为大数据服务的评价提供了有益参考。

参考文献

- [1] YANG C, HUANG Q, LI Z, et al. Big Data and cloud computing: innovation opportunities and challenges[J]. International Journal of Digital Earth, 2017, 10(1): 13-53.
- [2] MANJUNATH R, AKSHATHA A, BALAJI S. Reverse engineering in Big Data using Cloud computing and Open Stack virtual machine[C] // International Conference on Applied and Theoretical Computing and Communication Technology. IEEE, 2017: 848-853.
- [3] LI J, HUANG L, ZHOU Y, et al. Computation Partitioning for Mobile Cloud Computing in a Big Data Environment[J]. IEEE Transactions on Industrial Informatics, 2017, 13(4): 2009-2018.
- [4] ZHOU C, JIANG H, CHEN Y, et al. User interest acquisition by adding home and work related contexts on mobile big data analysis[C] // 2016 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS). IEEE, 2016: 201-206.
- [5] OTHMAN M, MADANI S A, KHAN S U. A survey of mobile cloud computing application models[J]. IEEE Communications Surveys & Tutorials, 2014, 16(1): 393-413.
- [6] YANG C, HUANG Q, LI Z, et al. Big Data and cloud computing: innovation opportunities and challenges[J]. International Journal of Digital Earth, 2017, 10(1): 13-53.
- [7] ENCK W, GILBERT P, HAN S, et al. TaintDroid: an information-flow tracking system for realtime privacy monitoring on smartphones[J]. ACM Transactions on Computer Systems (TOCS), 2014, 32(2): 1-29.
- [8] ZHANG Y, REN S, LIU Y, et al. A big data analytics architecture for cleaner manufacturing and maintenance processes of complex products [J]. Journal of Cleaner Production, 2017, 142(2): 626-641.
- [9] CHANG R M, KAUFFMAN R J, KWON Y O. Understanding the paradigm shift to computational social science in the presence of big data[J]. Decision Support Systems, 2014, 63(3): 67-80.
- [10] KLAUSER F R, ALBRECHTSLUND A. From self-tracking to smart urban infrastructures: Towards an interdisciplinary research agenda on Big Data[J]. Surveillance & Society, 2014, 12(2): 273-286.
- [11] SIVARAJAH U, KAMAL M M, IRANI Z, et al. Critical analysis of Big Data challenges and analytical methods[J]. Journal of Business Research, 2017, 70(C): 263-286.
- [12] HAKIRI A, GOKHALE A, BERTHOUE P, et al. Software-defined networking: Challenges and research opportunities for future internet[J]. Computer Networks, 2014, 75: 453-471.
- [13] QIAO S J, HAN N, ZHANG K F, et al. Algorithm for detecting overlapping communities from complex network big data[J]. Journal of Software, 2017, 28(3): 631-647. (in Chinese)
乔少杰, 韩楠, 张凯峰, 等. 复杂网络大数据中重叠社区检测算法[J]. 软件学报, 2017, 28(3): 631-647.
- [14] SONG J, SUN Z Z, MAO K M, et al. Research advance on mapreduce based big data processing platforms and algorithms[J]. Journal of Software, 2017, 28(3): 514-543. (in Chinese)
宋杰, 孙宗哲, 毛克明, 等. MapReduce 大数据处理平台与算法研究进展[J]. 软件学报, 2017, 28(3): 514-543.
- [15] ZHAI J H, ZHANG M Y, WANG T T, et al. K-Nearest neighbor algorithm based on hash technology and MapReduce[J]. Computer Science, 2017, 44(7): 210-214. (in Chinese)
翟俊海, 张明阳, 王婷婷, 等. 基于哈希技术和 MapReduce 的大数据集 K-近邻算法[J]. 计算机科学, 2017, 44(7): 210-214.
- [16] ZHAO J B, LIU Y X, SONG L B, et al. Application of Big Data Processing and Monitoring Technologies in Landslide Monitoring and Warning System[J]. Journal of Chongqing University of Technology (Natural Science), 2018, 32(2): 182-190. (in Chinese)
赵久彬, 刘元雪, 宋林波, 等. 大数据关键技术在滑坡监测预警系统中的应用[J]. 重庆理工大学学报(自然科学), 2018, 32(2): 182-190.