

基于在线模型评估的动态 Web 系统性能保障*

张勇^{1,2+}, 黄涛^{1,2}, 陈宁江³, 金蓓弘^{1,2}

¹(中国科学院 软件研究所 软件工程技术研发中心,北京 100080)

²(中国科学院 软件研究所 计算机科学重点实验室,北京 100080)

³(广西大学 计算机与电子信息学院,广西 南宁 530004)

Online Model Estimating-Based Performance Guarantee Mechanism for Dynamic Web Site

ZHANG Yong^{1,2+}, HUANG Tao^{1,2}, CHEN Ning-Jiang³, JIN Bei-Hong^{1,2}

¹(Technology Center of Software Engineering, Institute of Software, The Chinese Academy of Sciences, Beijing 100080, China)

²(State Key Laboratory of Computer Science, Institute of Software, The Chinese Academy of Sciences, Beijing 100080, China)

³(College of Computer, Electronic, and Information, Guangxi University, Nanning 530004, China)

+ Corresponding author: Phn: +86-10-62630989 ext 208, E-mail: yzhang@otcaix.iscas.ac.cn, <http://otc.iscas.ac.cn>

Zhang Y, Huang T, Chen NJ, Jin BH. Online model estimating-based performance guarantee mechanism for dynamic Web site. *Journal of Software*, 2007,18(7):1660–1671. <http://www.jos.org.cn/1000-9825/18/1660.htm>

Abstract: Dynamic multi-tiered Web system can be affected by many unexpected factors during runtime. At the same time, it has different performance characteristic under different workload pattern condition. Accordingly, different performance models are needed to describe the system. Currently, performance guarantee mechanism based on feedback control theory, which aims to eliminate the effect of unexpected factors, often employs single and fixed system performance model which can not adapt to changing performance characteristics. Under fluctuant and unexpected workload condition of Internet, the method can decrease the precision and stability of performance. To guarantee anticipated response time goal, this paper presents an adaptive performance guarantee mechanism based on online estimation of performance model. Using two complementary transactional Web benchmarks to evaluate the techniques experimentally shows that the mechanism can efficiently reduce the squared departure of performance goal under varying workload.

Key words: dynamic multi-tiered Web site; performance guarantee; adaptive control; online model estimating

摘要: 动态多层 Web 系统在运行时会受到许多不确定性因素的影响.同时,在不同的负载模式下具有不同的性能特性,需要不同的性能模型进行描述.为消除不确定性因素对系统性能的影响,基于反馈控制原理设计的性能保障机制主要采用单一、固定的系统性能模型,对动态 Web 系统变化的性能特征考虑不够.在负载呈波动且具有不可预测特性的 Internet 环境中,这会降低性能目标的精确性和稳定性.采用自适应控制的思想,以满足请求平均响应时间为目标,提出了一种基于在线评估系统性能模型的保障机制,并采用两个不同类型的事务性 Web 测试基准,对所提方

* Supported by the National Natural Science Foundation of China under Grant No.60573126 (国家自然科学基金); the National High-Tech Research and Development Plan of China under Grant Nos.2004AA112010, 2003AA414310 (国家高技术研究发展计划(863)); the National Basic Research Program of China under Grant No.2002CB312005 (国家重点基础研究发展计划(973))

Received 2005-11-14; Accepted 2006-06-06

法进行了评价,结果表明,该方法能够有效减轻变化负载模式下响应时间与预期目标的偏离程度。

关键词: 动态多层 Web 系统;性能保障;自适应控制;在线模型评估

中图法分类号: TP393 文献标识码: A

在负载呈波动且不可预测特性的 Internet 环境中时,准入控制(admission control)是保障 Web 系统性能的常用手段.通过限制进入系统的请求数量来保证系统的性能,其关键之一是如何确定合理的准入控制数量.

Web 系统可以分为两类:静态 Web 系统和动态 Web 系统.静态 Web 系统只提供静态的预先存储的 HTML 页面.而动态 Web 系统,一般采用数据库驱动的方式,基于多层结构实现,包括 Web 服务器、应用服务器及数据库服务器,典型应用如电子商务.动态 Web 系统提供的服务内容需要执行程序动态生成,处理过程复杂^[1].系统中的不确定性因素,包括事务处理、资源竞争、并发访问、组件失效以及随机调度等,会影响动态 Web 系统的性能,造成系统性能与预期目标的偏差^[2,3].为消除这些因素的影响,需要及时调整系统的准入率.

在工程上可采用控制论(control theory)中反馈控制(feedback control)的方法^[4,5]来消除不确定性因素对系统的影响.目前,在保障 Web 系统性能的一些研究中也引入了这种方法^[6-10].其核心思想是:先确定影响输出(预期性能指标)的关键输入(请求准入率),然后确定输入、输出之间定量变化的关系.基于这种关系,通过动态调整输入量来修正输出偏差^[4,5].这里,输入、输出之间的关系可以看作是 Web 系统的一种性能模型.性能模型的精确性直接影响由此确定的准入控制率,进而影响预期的性能目标.

目前,基于此方法的研究^[6-10]一般采用单一而固定的性能模型.这种性能模型的建立,主要基于特定的负载模式或者某种统计意义上的预测规律.基于单一性能模型的保障机制,实质上是使用一种固定不变的规则调整输入输出之间的关系.这对于静态 Web 系统是适用的,而对于性能特征多变的动态 Web 系统来说则不合适.原因如下:静态 Web 系统只提供静态 Web 页面,负载模式单一.静态 Web 请求在服务器端产生的负载,与所请求的静态 Web 页面的文件大小呈线形关系^[1].动态 Web 系统在不同的负载模式下具有较大的性能差异,比如:只访问 Web 层或应用服务器层的负载模式与访问后台数据层的负载模式相比;以读数据库操作为主的负载模式和以写数据库操作为主的负载模式相比;从缓存中访问数据和从磁盘中访问数据相比等^[3,11].为有效反映这种性能差异,需要在不同负载模式下使用不同的性能模型.否则,会降低性能目标的精确性,造成实际控制目标与预期目标的偏移.

本文提出了在不同负载模式下,在线评估动态 Web 系统性能模型的思想,并基于自适应控制(adaptive control)的相关理论^[12,13]提出了一种基于在线性能模型评估的保障机制.其基本思想是:根据在线测量的系统输入输出数据,以采样周期(sampling period)为单位,运用统计学方法,动态评估适应当前负载模式的性能模型,并据此建立了一套准入控制机制.文中以请求平均响应时间这个性能指标为例进行了研究.通过 Web 基准测试,表明该方法在波动的负载条件下,能够有效减轻响应时间与预期目标的偏离.

本文第 1 节介绍相关背景,包括动态 Web 系统及基于反馈原理的准入控制机制.第 2 节说明动态 Web 系统变化的性能特征及基于单一性能模型机制的不足.第 3 节阐述如何基于在线模型评估建立准入控制机制.第 4 节通过模拟实验说明方法的有效性.第 5 节比较相关工作.最后总结全文.

1 背景介绍

动态 Web 应用一般在典型的 3 层结构中实现,如图 1 所示.包括前端的 Web 服务器、中间层的应用服务器和后台的数据库服务器.Web 服务器处理静态 Web 请求,并将所有需动态处理的请求转发给应用服务器,应用服务器负责表示逻辑和业务逻辑的处理,在处理过程中往往需要与后台数据库服务器交互^[1].鉴于动态 Web 系统内部处理的复杂性,我们将其看作一个“黑箱”系统进行处理,只关注影响请求响应时间的关键输入.

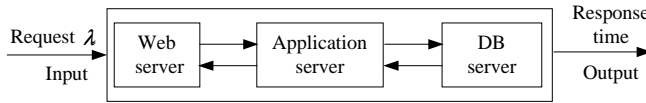


Fig.1 Typical architecture of dynamic Web site

图 1 动态 Web 系统的典型结构

请求负载是影响 Internet 环境中 Web 系统响应时间的关键因素^[14,15].对于动态 Web 应用来说,尽管不同请求处理过程和资源消耗差异较大,但是从整体上看,短时间内其负载模式不会发生剧烈变化.在相邻的时间间隔内,请求数量的变化仍可看作负载变化的主要因素.通过限制访问 Web 站点的请求数量,可以实现对响应时间目标的有效保障^[16,17].关键在于如何确定一个合理的请求准入数量.

本文中,我们将请求平均响应时间 $r(k)$ 和请求准入数量 $\lambda(k)$ 之间的函数关系称为动态 Web 系统的性能模型,形如 $r(k)=f(\lambda(k))$.不同的函数关系,比如 $f_1(\lambda(k)),f_2(\lambda(k))$,表示不同的性能模型.直观上理解,为满足预期响应时间目标 R_{ref} ,只要确定 $f(\lambda(k))$,然后根据 $R_{ref}=f(\lambda(k))$ 求得具体 λ 即可.但实际上,动态 Web 系统在运行过程中会受到许多不确定性因素的影响,基于 $R_{ref}=f(\lambda(k))$ 确定的 λ 并不能完全满足预期目标 R_{ref} ,不确定因素会使响应时间产生偏差和波动.

为消除不确定性因素的影响,可以基于反馈控制原理对 $\lambda(k)$ 进行调整^[6,7],如图 2 所示.在每个采样周期,监控器 (monitor)在线测量平均响应时间 $r(k)$ 与预期目标 R_{ref} 进行比较,得到误差 $\Delta r(k)=r(k)-R_{ref}$.反馈控制器 (controller)根据性能模型 $r(k)=f(\lambda(k))$ 和误差 $\Delta r(k)$,产生调整量 $\Delta\lambda$,期望该调整量能够消除原来的误差.假如当前准入率为 λ ,那么下一个周期,准入控制器 Gate 将根据 $\lambda+\Delta\lambda$ 执行准入控制.

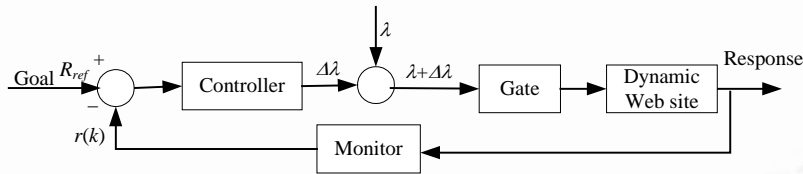


Fig.2 Performance guarantee based on feedback control

图 2 基于反馈控制的性能保障

2 基于单一性能模型机制的不足

首先通过实验说明,动态 Web 系统在不同的负载模式下具有不同的性能特征.然后说明,由于此原因,基于单一性能模型的反馈机制会造成性能目标的偏离.

利用两个不同类型的事务型 Web 测试基准“TPC-W”^[18]和“拍卖站点(auction site)”^[11],在不同的负载模式下,测量其吞吐量及对应平均响应时间,结果见表 1、表 2.关于实验环境及测试基准的详细介绍,请参看本文第 4 节“系统实验与结果”.

Table 1 TPC-W: Test data under different workload
表 1 TPC-W:不同负载模式下的性能数据

Workload	Browsing (95% read)	Shopping (80% read)	Ordering (50% read)
Response time	5.63	7.37	9.26
Peak throughput	195	247	372

Table 2 Auction site: Test data under different workload
表 2 拍卖站点:不同负载模式下的性能数据

Workload	Browsing (100% read)	Bidding (85% read)
Response time	3.25	4.71
Peak throughput	3 170	3 889

TPC-W 从浏览模式(browsing)改变到购买模式(shopping)和订购模式(ordering)时,读、写操作的改变量分别为 15%和 45%,平均响应时间的改变量分别为 31%和 64%,峰值吞吐量的改变量分别为 27%和 90%.拍卖站点从浏览(browsing)模式改变到投标(bidding)模式时,读、写操作的改变量为 15%,对应的平均响应时间和吞吐量的

改变量分别为 45% 和 23%。可以看出,当负载模式发生变化时,动态 Web 系统的性能特征呈现出了明显的差异,反映性能特征的性能模型也将会发生变化。

假如保障机制按照性能模型 $r(k)=f_1(\lambda(k))$ 来设计,该模型在负载模式 *Workload1* 下得到;在 k 时刻,产生了性能误差 Δr ,反馈机制将基于 $r(k)=f_1(\lambda(k))$ 进行调整,调整量为 $\Delta\lambda_1$ 。

(1) 如果当前负载模式等于或近似于 *Workload1*,性能特征能以 $r(k)=f_1(\lambda(k))$ 表示,那么,调整量 $\Delta\lambda_1$ 产生的修正效果为 $f_1(\Delta\lambda_1)=\Delta r$,能够修正误差。

(2) 如果当前负载模式为 *Workload2*,假设其对应性能模型为 $r(k)=f_2(\lambda(k))$ 。如果反馈机制仍然使用性能模型 $r(k)=f_1(\lambda(k))$,那么调整量仍为 $\Delta\lambda_1$;由于当前实际性能特征为 $r(k)=f_2(\lambda(k))$, $\Delta\lambda_1$ 产生的修正效果将为 $f_2(\Delta\lambda_1)$,因此产生了与预期修正目标 $f_1(\Delta\lambda_1)$ 的偏差,值为 $|f_1(\Delta\lambda_1)-f_2(\Delta\lambda_1)|$ 。

3 基于在线评估性能模型的机制

为了消除上述偏差,准入控制机制需要使用适应当前负载模式的性能模型。本文借鉴自适应控制的思想^[12,13],提出了一种基于在线评估性能模型的准入控制机制,如图 3 所示。在每个采样周期,模型评估器(model estimator)利用测得的输入输出数据 $r(k)$ 和 $\lambda(k)$,在线评估动态 Web 系统当前性能模型。基于最新模型,利用反馈控制原理,确定下一周期请求准入率。下面先说明如何动态评估系统性能模型。

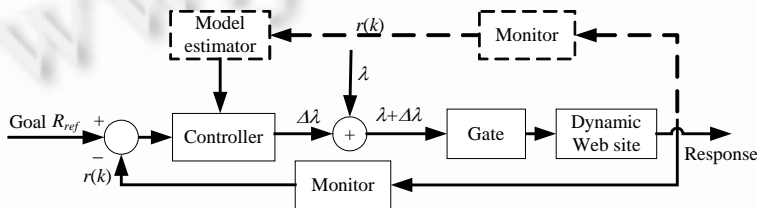


Fig.3 Adaptive guarantee mechanism based on online model estimating

图 3 基于在线模型评估的自适应性能保障机制

3.1 系统性能模型及在线评估过程

系统性能模型可以看作当前响应时间与过去响应时间及请求数量的函数^[6,7]。对动态 Web 系统来说,尽管难以确定一个适用于长期运行过程的函数,但是在较短的时间间隔内,系统输入和输出的改变量都比较小,因此可以采用反馈控制中常用的近似化处理方法,采用线形模型来表示这种关系^[4,5,12,13]。这种处理办法也适用于 Web 系统的准入控制^[6-9]。我们借助自适应控制中一个广泛使用的线形模型^[4,12,19],如式(1)所示,描述动态 Web 系统的性能。实践已经表明,该模型在针对离散系统的反馈控制中具有良好的效果^[19]。

$$r(k) = \sum_{i=1}^N A_i r(k-i) + \sum_{i=0}^{N-1} B_i \lambda(k-i-1) \quad (1)$$

其中, $r(k)$ 表示系统当前响应时间, $r(k-i)$ 表示历史阶段的响应时间, $\lambda(k-i)$ 表示历史阶段的请求输入量。我们需要将式(1)这个一般性的表达式确定为适合描述动态 Web 系统性能的式子。待定因子包括模型阶数(order) N 和模型参数 A_i, B_i ^[12]。模型阶数为 n ,表示系统当前状态与过去 n 个采样周期的状态及输入相关,模型参数 A_i, B_i 决定了因变量随自变量的变化程度。不同的阶数和参数代表不同的性能模型。性能模型评估的任务就是确定 N 和 A_i, B_i 。

3.1.1 确定动态 Web 系统性能模型阶数

性能模型阶数反映当前状态受历史状态及输入影响的程度,即模型阶数为 n 表示系统当前状态与过去 n 个采样周期的状态及输入相关。在动态 Web 系统的准入控制中,如果请求延迟比采样周期短得多,则一个采样周期以前的请求对系统当前状态的影响将会很小,一阶模型就可以描述。如果请求延迟长于采样周期,则需要高阶模型描述。下面我们确定模型阶数,采用先直观分析、再统计验证的方法。

对于非过载的动态 Web 系统来说,请求需要在较短时间内得到响应(几秒或十几秒),过长则会失去意义,故

可以不考虑很早以前请求对当前状态的影响.另一方面,可以认为负载模式在短时间内不会发生剧烈变化,这个时间周期相对于请求的响应时间要长得多.因此,我们可以推定,只要选取合适的采样周期,较低阶数的模型即可.对于这个假定,将通过统计学中拟合(fitting)的方法予以验证.因此,先假设阶数为 1,2 和 3,然后再通过实验从中选择一个合适的值.

利用动态 Web 基准 TPC-W,在本文第 4 节“系统实验与结果”所述的“实验环境”中,测量得到一组响应时间和请求数量值,然后利用最小二乘法(least-squares regression)在假定的阶数下对模型进行数据拟合,看拟合效果是否满足要求.考虑到读数据库操作的 Web 请求和写数据库操作的 Web 请求,两者处理过程和资源消耗的差异比较显著,选择全部是读操作和全部是写操作的负载模式进行实验.因为按照控制论相关知识,在极端模式下获得的数据能够更好地捕捉系统的动态行为^[4,11].

按照统计学相关理论^[20],决定系数(coefficient of determination, R^2)是模型拟合的重要度量指标.表 3 列出了假定系统性能模型为 1 阶时,在不同采样周期下的 R^2 值.对于采用线性方程近似表示的系统,当 $R^2 > 0.8$ 时就可以被认为是一个好的拟合^[4,12].从表 3 可以看出,在合适的采样周期下(12s 和 14s)已经具有了较好的拟合效果(>0.8).此外,我们也分别将模型假定为 2 阶和 3 阶进行了实验,其拟合效果比 1 阶模型分别改进约 6%和 9%.考虑到一阶模型在处理上的方便且已满足要求,因此在合适的采样周期下,可以选用一阶模型描述动态 Web 系统的性能.故可将式(1)简化为

$$r(k)=a_{k-1} \cdot r(k-1)+b_{k-1} \cdot \lambda(k-1) \quad (2)$$

Table 3 Fitting metric values under First-order model

表 3 假定性能模型为 1 阶时的数据拟合值

Sampling period (s)	4	6	8	10	12	14
All-Read	0.35	0.42	0.59	0.77	0.82	0.86
All-Write	0.47	0.51	0.57	0.68	0.80	0.84

实践经验也表明,如文献[19],在受控的计算机系统中,模型阶数一般比较低且保持不变,可由离线方式评估得到.这样,在确定模型阶数以后,评估性能模型的任务主要是动态确定式(2)中的参数 a 和 b .

3.1.2 在线评估性能模型参数

动态确定模型参数的目的,在于使 a, b 值代入式(2)后所构成的模型能够体现当前负载模式的性能特征.因此,我们动态评估参数可采用一种方法:在采样周期 k ,计算按照性能模型预测的响应时间与当前实际测量值之间的误差.如果误差为 0,说明原来的模型能够适应当前负载模式,不需要更新模型的参数;否则,可以利用误差求解新的模型参数,具体可根据统计学中的递归最小二乘法(recursive least-squares,简称 RLS)来实现^[12,13].为了书写方便,将式(2)改写为 $r(k)=I(k)\phi(k-1)$,其中, $I(k)=[b_{k-1}a_{k-1}]$, $\phi(k-1)=[\lambda^T(k-1)r^T(k-1)]^T$.动态评估参数的算法,以伪码形式描述如下:

Params_estimating ($r(k), \hat{I}(k-1)$)

/* $r(k)$:测量的响应时间; $\hat{I}(k-1)$: $k-1$ 周期确定的参数值,可认为是对 k 周期参数值的估计,故表示为 $\hat{I}(k-1)$;*/

Var: $\varepsilon(k); p(k)$; /* $p(k)$ 表示协方差矩阵(covariance matrix)*/

$\varepsilon(k)=r(k)-\hat{I}(k-1)\phi(k-1)$; /*实测值 $r(k)$ 与按性能模型预测值 $\hat{I}(k-1)\phi(k-1)$ 之间的误差*/

if $\varepsilon(k)=0$ then $\hat{I}(k):=\hat{I}(k-1)$; /*模型能够适应当前负载模式,参数不需要更新;*/

else /*按下面的等式计算参数 $\hat{I}(k)$ 的新值*/

$$p(k-1) = p(k-2) - \frac{p(k-2)\phi(k-1)\phi^T(k-1)p(k-2)}{1 + \phi^T(k-1)p(k-2)\phi(k-1)};$$

$$\hat{I}(k) = \hat{I}(k-1) + \frac{\varepsilon(k)\phi^T(k-1)p(k-2)}{1 + \phi^T(k-1)p(k-2)\phi(k-1)};$$

```

return  $\hat{I}(k)$ ;
}
    
```

3.1.3 根据性能模型确定请求准入率

确定性能模型的目的在于确定请求准入率 $\lambda(k)$.将式(2)中的 k 提前一个时间单元,可得到

$$\lambda(k) = \frac{1}{b}r(k+1) - \frac{a}{b}r(k) \tag{3}$$

如果在周期 k 以 $\lambda(k)$ 控制动态 Web 系统的请求,那么, $r(k+1)$ 就表示在 $k+1$ 周期的响应时间,而这正是我们预期要达到的目标 R_{ref} .因此,式(3)可进一步表示为

$$\lambda(k) = \frac{1}{b}R_{ref} - \frac{a}{b}r(k) \tag{4}$$

式(4)中,不同的 a, b 值将得到不同的 $\lambda(k)$,相应的系统响应时间也会不同,适应变化负载模式的动态性能模型即体现于此.在控制论中,式(4)被称为系统的控制律^[4,5],在反馈控制器中得以实现.

3.2 准入控制机制的实现

准入控制机制的结构如图 4 所示,整体上实现为一个轻量级 Web 代理.轻量级 Web 代理比常规 HTTP 代理处理速度快、消耗资源少^[21].Web 代理拦截所有访问动态 Web 系统的 HTTP 请求,统计请求数量并执行准入控制.如果当前实际请求数量 λ 低于 $\lambda(k)$,则不进行准入控制,将所有请求直接转发给后面的动态 Web 系统;否则,按照 $\lambda(k)$ 执行准入控制,限制进入动态 Web 系统的请求.在策略上,采取基于会话的准入控制,优先接收已接受会话的请求,已接受会话的请求比新到达会话的请求往往具有更高的价值^[22].Web 代理将接受的请求转发给后面的 Web 服务器,进入动态 Web 系统.处理完毕,响应结果再经过 Web 代理返回客户端.

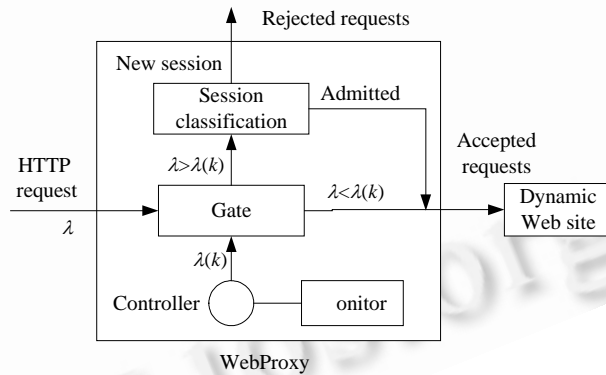


Fig.4 Realization of admission control mechanism

图 4 准入控制机制的实现

通过记录请求到达时刻和得到响应的时刻,监控器可以计算请求平均响应时间;反馈控制器实现模型评估并计算 $\lambda(k)$;准入控制机制 Gate 采用令牌环(token)方法,根据动态确定的 $\lambda(k)$ 生成同等数量的令牌.当新的请求到达时,如果存在可用令牌,则接受该请求并消耗掉一个令牌,否则请求被拒绝.

计算请求准入数量.准入控制机制在运行以前,需要初始化与应用相关的参数:(1) 采样周期.采样周期长短受动态 Web 请求延迟时间的影响,而请求延迟时间又与具体应用相关.在我们的实验中,取平均响应时间的 3 倍,可以获得较好的拟合效果;(2) 预期响应时间目标 R_{ref} ;(3) 为算法估计一个初始准入率 $\lambda(0)$.

从每个采样周期,可采用如下算法求得 $\lambda(k)$:

```

Admission_level(sampling_period, r(k), R_ref) /*r(k):在线测量的响应时间,sampling_period:采样周期长度*/
{
    Initial  $\hat{I}(0) = 1$ ; /*因为参数评估中会修正  $\hat{I}(k)$ ,不妨设初值  $\hat{I}(0) = 1$ */
    Params_estimating(r(k),  $\hat{I}(k-1)$ ); /*利用第 3.1.2 节中的算法,动态确定模型参数 a,b*/
}
    
```

```

λ(k) = 1/b * R_ref - a/b * r(k); /*计算新的请求准入率*/
return λ(k);
}
    
```

4 系统实验与结果

4.1 实验方法和评价指标

下面通过实验测试所提机制对变化负载模式的适应能力以及在线模型评估所造成的系统开销.

测试基准.本文采用两个不同类型的测试基准:一个是基于事务性 Web 测试基准 TPC-W 规范实现的在线书店(online bookstore)^[18],共包含 14 个业务交互,其中 8 个涉及数据库读、写操作,6 个只涉及读操作.按不同比例组合业务交互,可构成不同负载模式,常用的几种模式是浏览(browsing)、购买(shopping)和订购(ordering),见表 4.第 2 个是拍卖站点测试基准(auction site,简称拍卖基准)^[11],模拟一个实际拍卖站点的功能,它包括 26 个业务交互,主要负载模式包括浏览(browsing)和投标(bidding),见表 5.

Table 4 Workload mix of online bookstore benchmark

表 4 在线书店测试基准负载模式

Mix	Browsing	Shopping	Ordering
Read operation (%)	95	80	50
Write operation (%)	5	20	50

Table 5 Workload mix of auction site benchmark

表 5 拍卖站点测试基准负载模式

Mix	Browsing	Bidding
Read operation (%)	100	85
Write operation (%)	-	15

这两个基准的侧重点不同,在线书店包含复杂数据库操作,侧重数据库层;拍卖站点数据库操作相对简单,主要是生成动态 Web 内容,侧重前端 Web 服务器层.

客户负载.客户端采用负载生成器,模拟产生会话类型的负载,会话过程中发出一系列 Web 请求,两个请求之间会等待一定的思考时间(think time),负载的轻重可通过改变会话数量来调整.表 6 的负载用于驱动在线书店,表 7 的负载用于驱动拍卖基准,表中各阶段负载数量变化保持均匀.为充分发挥准入控制机制的作用,表中各阶段负载都超过了各自系统的容量.实验中,在线书店响应时间目标设为 4s,采样周期设为 12s;拍卖基准响应时间目标设为 2s,采样周期设为 6s.

Table 6 Changing input workload for online bookstore

表 6 用于在线书店的测试负载

	Period1	Period2	Period3	Period4
Browsing	0→400	400→0	0	0
Shopping	0	0→400	400→0	0
Ordering	0	0	0→400	400→0

Table 7 Changing input workload for auction site

表 7 用于拍卖站点的测试负载

	Period1	Period2	Period3
Browsing	0→4000	4000→0	0
Bidding	0	0→4000	4000→0

实验环境.如图 5 所示,一台客户机运行负载生成器;Web 代理单独运行在一台机器上;服务器端实现为 3 层结构,一台机器运行 Web 服务器,一台机器运行应用服务器,采用中国科学院软件研究所开发的 OnceAS2.0^[23],后端一台 PC Server 运行数据库服务器,所有机器通过 100Mbps 以太网互联.服务器机均运行 Redhat Linux7.1 操作系统.

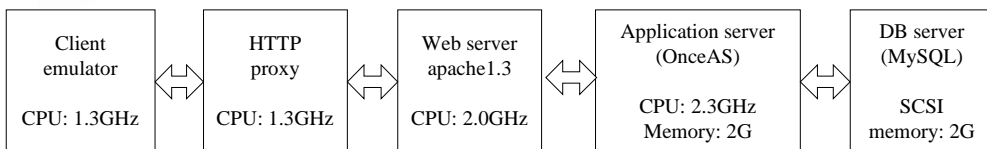


Fig.5 The testing environment

图 5 实验环境

4.2 实验结果

4.2.1 对变化负载模式的适应能力

先使用在线书店测试基准,得到浏览模式和订购模式下的性能模型.在表 6 的负载条件下,基于在线模型与基于浏览模式性能模型的对比结果,包括响应时间及请求准入率,如图 6(a)所示;与基于订购模式性能模型的对比结果如图 6(b)所示.利用拍卖站点也进行了类似实验,在表 7 负载条件下,基于在线模型与单一模型的比较结果,如图 7(a)(浏览模式)和图 7(b)(投标模式)所示.

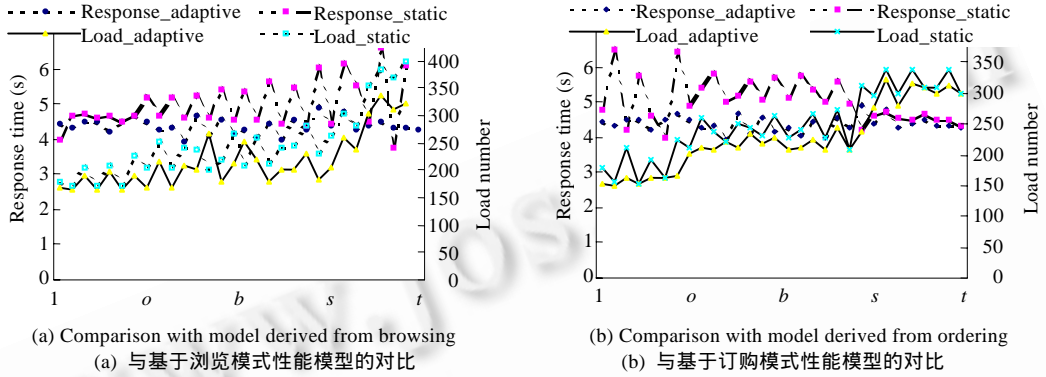


Fig.6 Online bookstore

图 6 在线书店

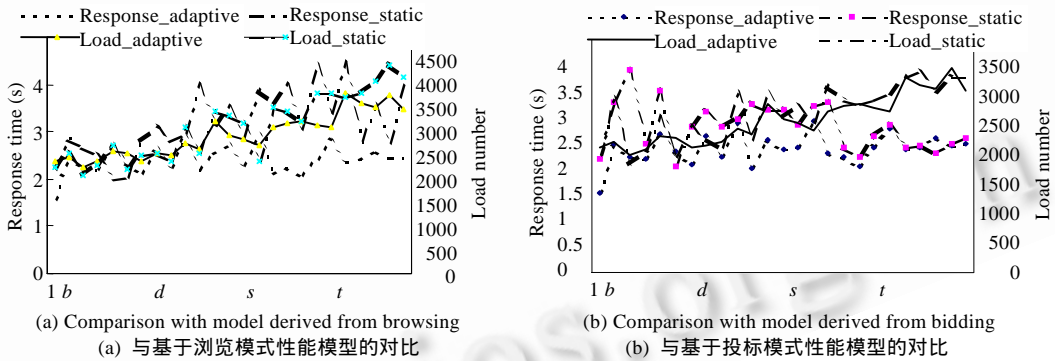


Fig.7 Auction site

图 7 拍卖基准

从图 6 和图 7 可以看出:(1) 采用单一性能模型的机制,只有在与性能模型比较符合的负载模式下,响应时间才保持得比较好,在其他阶段波动都比较大;采用在线模型的机制,响应时间一直比较稳定;(2) 采用在线模型和单一模型,得到的请求准入率是不同的,相应造成各自响应的的时间也不同;(3) 无论哪种机制,在各个阶段请求准入率都不断发生变化,并非一个固定值,这主要是受随机因素的影响,因而需要不断调整准入率来消除这些因素的影响.

为了进一步说明响应时间与预期目标的偏离程度,表 8、表 9 给出了上述实验中,基于不同性能模型机制下响应时间的方差 S^2 (mean square).

可以看出,基于在线模型的机制能够明显减轻偏差.例如,在线书店采用在线模型的方差为 0.58,比基于订购模式的 2.23,改进 $(2.23-0.58)/2.23 \approx 74\%$;拍卖基准采用在线模型的方差为 0.51,比基于浏览模式的 1.37,改进 $(1.37-0.51)/1.37 \approx 63\%$.

Table 8 S^2 of response time of online bookstore**表 8** 在线书店响应时间在不同负载条件下的方差

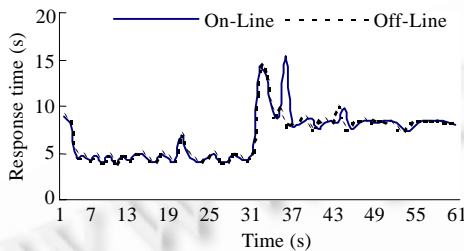
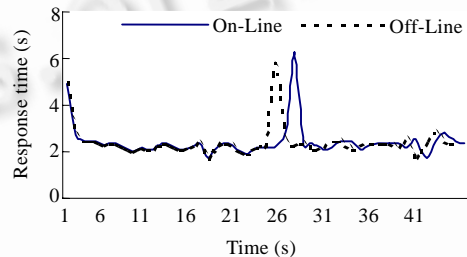
S^2	Period1	Period2	Period3	Period4	Whole
Browsing	1.64	0.56	3.88	5.79	2.96
Ordering	0.76	3.12	4.15	0.82	2.23
Adaptive	0.48	0.62	0.67	0.51	0.58

Table 9 S^2 of response time of auction site**表 9** 拍卖基准响应时间在不同负载条件下的方差

S^2	Period1	Period2	Period3	Whole
Browsing	0.44	2.87	0.49	1.37
Bidding	4.65	0.38	3.61	2.83
Adaptive	0.47	0.55	0.52	0.51

4.2.2 系统开销

在 Proxy 中测量采用不同机制的请求响应时间,考察动态评估性能模型所造成的相对开销.在线书店测试过程中,由 400 个浏览模式的客户变化为 400 个订购模式的客户,在采用已知性能模型及动态计算性能模型两种情况下,响应时间如图 8 所示.利用拍卖站点的测试,由 4 000 个浏览模式的客户变化为 4 000 个投标模式的客户,在采用已知性能模型及动态计算性能模型两种情况下,响应时间如图 9 所示.

**Fig.8** Overhead under online bookstore benchmark**图 8** 利用在线书店测试得到的系统开销**Fig.9** Overhead under auction site benchmark**图 9** 利用拍卖基准测试得到的系统开销

总体上看,动态评估模型所花费的系统开销比较小.例如,在线书店在两种情况下达到稳定的时间分别为 2.6s 和 3s,开销为 $(3-2.6)/2.6=15\%$;拍卖基准在两种情况下达到稳定的时间分别为 2.2s 和 2.4s,开销为 $(2.4-2.2)/2.2\approx 10\%$.从本实验还可以看出,性能目标在较短时间内(约 3s)就可达到稳定.

5 相关工作比较

Web 站点的处理能力受到系统整体资源的制约,请求产生的负载超过了处理能力,系统就会造成过载,性能降低.准入控制是保证系统性能的重要手段.目前,Web 站点准入控制的研究从方法上可分为基于静态约束的方法、基于策略的方法、基于控制论方法以及基于服务降级的方法;根据研究对象,可分为静态 Web 系统和动态 Web 系统.目前,针对静态 Web 系统的研究较多,对动态 Web 系统尤其是数据库驱动的事务型动态 Web 系统的研究较少.

静态约束的方法.在基于线程/进程的 Web 服务器中(比如 Apache 服务器),采用限制连接数量的方式.这种限制一般采用静态方式事先指定.在动态 Web 系统中,由于不同请求类型处理代价差异较大,再加上 Internet 环境中请求的不可预测性和波动性,采用静态约束方式难以确定合适的“限制值”.如果以静态请求为主,连接数量可能达到几千;而如果以请求动态页面为主,可能最多只能支持到几百.设置太低,系统资源得不到充分利用;设置过高,又会导致系统过载.因此,约束值难以静态确定.

基于控制论反馈控制的方法.这本质上是一种动态方法,基本思想是基于在线监测某些指标来动态调整请求准入的数量.文献[6,7]针对 Web 请求的不可预测性,利用反馈控制,通过调整和维护 CPU 的利用率来保证响应时间目标.在设计中,将 Web 服务器的性能模型看作一个静态增益(static gain)函数;文献[8]利用系统辨识方法和线性控制理论为 Apache 服务器设计了一种调度算法,通过动态控制分配给每类请求的服务器进程,保证不同服务类别之间的相对延迟;文献[9]中利用一个多输入、多输出(multi input multi output)的比例积分控制器(proportion-integral controller),通过动态调整进程数量和连接空闲超时这两个参数,以满足 Apache 服务器的资源利用目标;文献[24]根据一个描述 Web 服务器中客户和服务器之间交互关系的性能模型,通过动态调度会话

中的请求,以获得最大的获益(reward)目标.以上几个研究主要使用了单一模型机制,不适用于动态 Web 系统.

在实际中,一般采用线性模型来近似表示具有非线性特性的系统,为了使近似更加精确,文献[25,26]采用控制论中前馈(feed forward control)控制和反馈控制相结合的方式,通过初步预测系统工作点,使反馈控制器尽可能工作在平衡点附近,以使近似更加精确.文中利用排队论对 Web 系统进行了分析,并采用一种线性确定性模型(linear deterministic model)来描述系统.这种方法,一是在整个控制过程中只使用了固定不变的系统模型;二是采用确定性模型分析排队论系统的做法,文献[27]曾经提出异议.

面对日渐复杂的系统,一些基于反馈控制原理的性能保障机制,采用了私有(ad hoc)控制器^[10,28].这种方法的优点是能够克服设计控制器时复杂的建模和参数调整过程,缺点是无法对系统进行形式化的分析.

在其他相关领域也有少数几项利用自适应反馈机制进行控制的研究.包括:文献[29]针对系统缓存问题,研究了缓存大小和命中率之间的关系;文献[30]针对外部存储系统,研究了请求和延迟之间的关系等.

基于策略的方法,典型代表是文献[22]提出的基于会话(session)的准入控制策略,提出了基于会话而不是单个请求或者连接执行准入控制,意识到了会话作为负载管理单元的重要性.另外一种常用方法是基于优先级的准入控制,而这往往是特定于应用的.

基于降低服务质量的方法.这种方法与常规的准入控制有些不同,通过降低提供给客户的服务,减少每个请求所需要的工作量来减少系统的负载.比如降低图形图像的质量或者降低数据的新鲜度等.然而,并非所有的服务都适用于服务降级的思想.采用这种方式的前提是客户愿意接受更低的服务质量,特别是对于以电子商务为代表的动态 Web 系统,客户可能并不愿意接受降低数据新鲜度的方式.

此外,文献[31]还提出了一种称作 SEDA(staged-event-driven-architecture)的 Internet 服务范型,并基于此设计了一种负载控制机制.在 SEDA 中,Internet 服务被分解为由事件驱动的阶段集合,这些阶段使用请求队列连接起来,每个阶段都能够控制来自于请求队列的请求准入率.这种方法的主要缺点是基于特有的服务范型,与现有 Web 系统的实现模式不一致,不具有通用性.

与以上工作比较,本文的研究工作主要特点在于:从研究对象上看,针对动态 Web 系统进行研究,以动态确定请求准入率来维持预期响应时间指标;从采用的方法上看,属于反馈控制的方法,但与以往研究的不同在于,本文充分考虑了动态 Web 系统不同负载模式具有不同的性能特征,提出了基于在线评估性能模型的准入控制方法,通过分析和拟合验证,提出了在一个采样周期内可以采用一阶线性模型进行近似处理,并给出了动态确定模型参数的方法.从而克服了单一性能模型不能适应动态 Web 系统多变性能特征的不足,有效降低了预期性能目标的偏差;在实验方法上,选用了事务性 Web 测试基准进行测试,而没有简化对动态请求的负载处理,如文献[2]将动态页面的处理时间简单地设置为一个固定值;在准入控制机制的实现上,结合了面向会话准入控制策略的优点;最后,本文所提方法无须修改操作系统和 Web 应用,在动态多层 Web 系统中易于实现.

6 结束语

本文将自适应的反馈控制机制引入到动态 Web 系统的性能保障中.通过在线测量系统的输入/输出值,动态评估系统性能模型,并据此建立性能保障机制.这种方法适应了动态 Web 系统在不同负载模式下的不同性能特征,克服了基于单一性能模型的保障机制不能适应变化负载模式的问题.模拟实验表明,该方法以较小的系统开销,较大幅度地改进了变化负载模式下动态 Web 系统性能目标的偏离问题.性能保障过程中无须人工干预和调整,所有的步骤都可以在一个采样周期内完成,而且所提方法无须修改 Web 应用和底层系统.

本文所提方法需要一个约束条件,即假定在一个较短的采样周期内,动态 Web 系统的负载模式不会发生大的变化;以采样周期为单位,通过控制请求总量就可以对性能产生较大影响.这个约束并不会对控制效果产生大的负面影响,因为系统的性能模型是动态计算的,即便是系统的负载模式在一个采样周期内发生了较大变化,在下一个采样周期内,系统的模型很快就会得到修正.

下一步的研究工作主要包含 3 个方面:(1) 进一步确定动态 Web 系统性能指标与具体资源保障之间的关系;(2) 研究在大规模应用环境下,基于集群结构实现的动态 Web 系统的性能保障问题;(3) 研究有区别服务与

准入控制机制的结合.

References:

- [1] Cardellini V, Casalicchio E, Colajanni M, Yu PS. The state of the art in locally distributed Web-server systems. *ACM Computing Surveys*, 2002,34(2):263–311.
- [2] Chen X, Mohapatra P, Chen H. An admission control scheme for predictable server response time for Web accesses. In: *Proc. of the 10th World Wide Web Conf.* New York: ACM Press, 2001. 545–554. <http://portal.acm.org/citation.cfm?id=372156&dl=>
- [3] Vilella D, Pradhan P, Rubenstein D. Provisioning servers in the application tier for e-commerce systems. In: *Proc. of the 12th Int'l Workshop on Quality of Service.* New York: IEEE Computer Society Press, 2004. 57–66. <http://portal.acm.org/citation.cfm?id=1189740.1189747>
- [4] Franklin GF, Powell JD, Workman M. *Digital Control of Dynamic Systems*. 3rd ed., Boston: Addison-Wesley Publishing Company, 1998.
- [5] Dai ZD. *The Basis of Automatic Control Theory*. Beijing: Tsinghua University Press, 1991 (in Chinese).
- [6] Abdelzaher TF, Lu C. Modeling and performance control of Internet servers. In: *Proc. of the 39th IEEE Conf. on Decision and Control (CDC)*. New York: IEEE Computer Society Press, 2000. 2234–2239. <http://citeseer.ist.psu.edu/abdelzaher00modeling.html>
- [7] Abdelzaher T, Shin KG, Bhatti N. Performance guarantees for Web server end-systems: A control-theoretical approach. *IEEE Trans. on Parallel and Distributed Systems*, 2002,13(1):80–96.
- [8] Lu CY, Abdelzaher TF, Stankovic JA, Son SH. A feedback control approach for guaranteeing relative delays in Web servers. In: Ho JM, ed. *Proc. of the IEEE Real-Time Technology and Application Symp.* Taipei: IEEE Computer Society, 2001. 51–62.
- [9] Gandhi N, Hellerstein JL, Diao YX, Parekh SS, Tilbury D. Using MIMO feedback control to enforce policies for interrelated metrics with application to the apache Web server. In: *Proc. of the NOMS 2002 IEEE/IFIP Network Operations and Management Symp.* New York: IEEE Computer Society Press, 2002. 219–234. http://ieeexplore.ieee.org/xpl/freeabs_all.jsp?arnumber=1015566
- [10] Kamra A, Misra V, Nahum E. Yaksha: A controller for managing the performance of 3-tiered Web sites. In: *Proc. of the 12th IEEE Int'l Workshop on Quality of Service (IWQoS)*. IEEE Computer Society Press, 2004. 47–56. http://ieeexplore.ieee.org/xpl/freeabs_all.jsp?arnumber=1309356
- [11] Cecchet E, Chanda A, Elnikety S, Marguerite J, Zwaenepoel W. Performance comparison of middleware architectures for generating dynamic Web content. In: Endler M, Schmidt DC, eds. *Proc. of the ACM/IFIP/USENIX Int'l Middleware Conf. (Middleware 2003)*. LNCS 2672, Springer-Verlag, 2003. 242–261.
- [12] Astrom KJ, Wittenmark B. *Adaptive Control*. 2nd ed. In: *Series in Electrical Engineering: Control Engineering*. Boston: Addison-Wesley Publishing Company, 1995.
- [13] Xie XM, Ding F. *Adaptive Control System*. Beijing: Tsinghua University Press, 2002 (in Chinese).
- [14] Arlitt M, Krishnamurthy D, Rolia J. Characterizing the capability of a large Web-based shopping system. *ACM Trans. on Internet Technology*, 2001,1(1):44–69.
- [15] Kant K, Mohapatra P. Scalable Internet servers: Issues and challenges. *ACM SIGMETRICS Performance Evaluation Review*, 2000, 28(2):5–8.
- [16] Elnikety S, Nahum E, Tracey J, Zwaenepoel W. A method for transparent admission control and request scheduling in e-commerce Web sites. In: Fledman SI, ed. *Proc. of the Int'l World Wide Web Conf.* New York: ACM Press, 2004. 276–286.
- [17] Urgaonkar B, Shenory P. Cataclysm: Policing extreme overloads in Internet applications. In: *Proc. of the Int'l World Wide Web Conf.* New York: ACM Press, 2005. 740–749. <http://portal.acm.org/citation.cfm?id=1060745.1060852>
- [18] Menasce DA. TPC-W: A benchmark for e-commerce. *IEEE Internet Computing*, 2002,6(3):83–87.
- [19] Hellerstein J, Diao YX, Parekh S, Tilbury D. *Feedback Control of Computing Systems*. New York: Wiley-IEEE Press, 2004.
- [20] Hellerstein J, Zhang F, Shahabuddin P. A statistical approach to predictive detection. *Computer Networks*, 2001,35(1):77–95.
- [21] Kaes RJ, Young S. Tinyproxy. 2002. <http://tinyproxy.sourceforge.net/>
- [22] Cherkasova L, Phaal P. Session-Based admission control: A mechanism for peak load management of commercial Web sites. *IEEE Trans. on Computers*, 2002,51(6):669–685.

- [23] <http://www.once.com.cn>. 2005.
- [24] Carlstrom J, Rom R. Application-Aware admission control and scheduling in Web servers. In: Proc. of the IEEE INFOCOM 2002. New York: IEEE Press, 2002. 506–515. http://ieeexplore.ieee.org/xpl/freeabs_all.jsp?arnumber=1019295
- [25] Lu Y, Abdelzaher T, Lu C, Sha L, Liu X. Feedback control with queuing theoretic prediction for relative delay guarantees in Web servers. In: Proc. of the 9th IEEE Real Time and Embedded Technology and Applications Symp. Washington: IEEE Computer Society Press, 2003. 208–218. <http://portal.acm.org/citation.cfm?id=827266.828543>
- [26] Henriksson D, Lu Y, Abdelzaher T. Improved prediction for Web server delay control. In: Proc. of the 16th Euromicro Conf. on Real Time Systems (ECRTS 2004). Washington: IEEE Computer Society Press, 2004. 61–68. <http://portal.acm.org/citation.cfm?id=1009383.1009824>
- [27] Jr Stidham S. Optimal control of admission to a queuing system. IEEE Trans. on Automatic Control, 1985,30(8):705–713.
- [28] Katabi D, Handley M, Rohrs C. Congestion control for high bandwidth-delay product networks. In: Proc. of the ACM SIGCOMM 2002. New York: ACM Press, 2002. 89–102. <http://www.sigcomm.org/sigcomm2002/papers/xcp.html>
- [29] Lu Y, Abdelzaher T, Lu C, Tao G. An adaptive control framework for QoS guarantees and its application to differentiated caching services. In: Proc. of the Int'l Workshop on Quality of Service. New York: IEEE Computer Society Press, 2002. 23–32. http://ieeexplore.ieee.org/xpl/freeabs_all.jsp?arnumber=1006571
- [30] Karlsson M, Karamanolis C, Zhu X. Triage: Performance isolation and differentiation for storage systems. In: Proc. of the Int'l Workshop on Quality of Service 2004. New York: IEEE Computer Society Press, 2004. 67–74. http://ieeexplore.ieee.org/xpl/freeabs_all.jsp?arnumber=1309358
- [31] Welsh M, Culler D. Adaptive overload control for busy Internet servers. In: Proc. of the 4th USENIX Conf. on Internet Technologies and Systems (USITS 2003). <http://www.usenix.org/publications/library/proceedings/index.html#2003>

附中文参考文献:

- [5] 戴忠达.自动控制理论基础.北京:清华大学出版社,1991.
- [13] 谢新民,定锋.自适应控制系统.北京:清华大学出版社,2002.



张勇(1975 -),男,山东宁阳人,博士生,主要研究领域为网络分布计算,软件工程.



黄涛(1965 -),男,博士,研究员,博士生导师,主要研究领域为网络分布计算,软件工程.



陈宁江(1975 -),男,博士,副教授,主要研究领域为网络分布计算,软件工程.



金蓓弘(1967 -),女,博士,研究员,CCF 高级会员,主要研究领域为网络分布计算,软件工程.