

# 移动计算环境下的 QoS 技术<sup>\*</sup>

张知讷 徐明

(长沙国防科技大学国家并行与分布处理国家科技重点实验室 长沙410073)

**摘要** QoS管理和控制是移动计算中的一个重要研究领域。该领域的主要研究内容有环境识别、适应性管理、资源管理和资源预留等。本文论述了移动计算环境下 QoS 研究中的一些基本概念、主要问题、研究内容和方法,并对该领域最近的研究工作进行了综述。

**关键词** 移动计算,服务质量,无缝跨区,环境识别,适应性管理,资源预留

## A Survey of Quality of Service Mechanisms in Mobile Computing

ZHANG Zhi-Jiao XU Ming

(National Laboratory on Parallel & Distributed Processing, National University of Defence Technology, Changsha 410073)

**Abstract** QoS management and control have become an important part in mobile computing. This includes context awareness, management adaptivity, resource management and reservation. This paper is a survey of recent works about QoS research in mobile computing. The general conceptions, main problems, mechanics and approaches of QoS specification and management are also discussed.

**Keywords** Mobile computing, Quality of service, Seamless handover, Context awareness, Management adaptivity, Resource reservation

### 1. 移动计算概述

近年来,伴随着计算机技术日新月异的发展,移动通信和计算也不断进步。蜂窝通信、无线局域网、卫星通信、蓝牙技术等无线通信手段越来越深入人们的生产和生活。随着移动设备性能的不提高和计算理论的不进步,移动计算成为一个新兴的研究方向,前景非常广阔。该技术将计算机技术与移动通信技术相结合,使得人们可以随时随地享受信息服务。

移动计算的难点就在于它的移动所带来的不可预测性以及蓄电技术的限制所带来的能源有限性问题。由于这些原因,移动系统从物理层到应用层都要解决固定系统所没有的问题。

用户需求永远是推动技术发展的动力,今天,除手机外,手持 PDA、笔记本电脑等移动设备已经被用户大量的使用。简单的对声音和文本符号的无线传输已经远远不能满足用户需求。移动用户也希望像能在固定环境下那样进行 Web 访问、视频点播等享受实时以及多媒体服务。移动计算技术也正是在这样的背景下发展起来的。

移动计算与传统的移动通信的区别在于<sup>[1]</sup>:移动通信研究的范围主要是物理层和 MAC 子层,其主要研究目标是如何更好地实现在无线的环境下完成机器与机器之间的通信;而移动计算通常研究移动网络的 MAC 层以上的部分,以及相关移动应用软件,即研究如何在移动环境下实现各种各样的应用。

移动性带来了许多难题,导致移动网络从网络层一直到应用层都需要进行修改。例如,在网络过载时,传输层的 TCP 协议会丢弃一些数据报文,而移动计算环境下报文的丢失往往是由通信环境太差引起的。还有 TCP 的“三次握手”机制在带宽有限的无线通信环境下也需要进行修改或干脆采用其它

机制。

移动计算系统可以再细分为两种情况,一是游牧系统(Nomadic systems),它的特征是虽然通信单元(移动主机)是可以移动的,但是它在每次通信时位置却是固定的,要与固定网络相连才能进行通信,即它在移动时不进行通信;另一种是移动系统(Mobile system),它的特征则是移动主机边移动边通信。虽然这两种系统有所差别,但是它们需要解决的许多问题却是相同的,而且很多系统既有游牧系统的特征,又有移动系统的特征,所以通常将这两种情况放在一起研究,而将它们统称为移动系统。

目前,许多研究机构或公司都在发展移动计算系统,包括 Xerox 的 Ubiquitous Computing 系统<sup>[26]</sup>、伯克利的 Infopad<sup>[27]</sup>、CMU 的 Coda<sup>[28]</sup>等等。

### 2. 移动计算环境下的 QoS 技术

#### 2.1 QoS 概述

一般地讲,QoS 就是指在网络系统或软件系统中满足一系列的用户服务质量标准或服务质量保证。如进行金融业务的用户要求网络传输极为可靠,不能出现错误;而在线观看视频节目的用户则不能容忍由于受数据传输率的影响而使画面长时间的静止;在一些收费服务中,一些优先级别高的用户希望能享受更高级别的服务等等。具体可以将这些用户的服务请求量化为数据分组延迟、分组丢失率、延迟抖动、带宽和吞吐量等 QoS 参数。

目前,QoS 技术的研究进展得非常迅速。在主机、路由器、交换机等各个硬件层次和操作系统、中间件、应用程序等各软件层次上都采用了多种技术来提供 QoS 保证。在需要可靠传输的非同步系统、视频、音频等实时系统、移动系统甚至 Web 系统中,有关 QoS 技术的研究都在展开。为了给网络应用提

<sup>\*</sup> 本文受国家自然科学基金资助,项目编号:60073002。张知讷 硕士研究生,主要研究方向为移动计算技术。徐明 博士生导师,教授,主要研究方向为移动计算技术。

供性能保证和区分服务, IETF 还提出了两种 QoS 体系结构, 即综合服务(IntServ)和区分服务(DiffServ)。

固定环境下 QoS 就是一个有待解决的问题。移动环境下更是如此。除了前面提到的 QoS 参数外, 移动计算环境还包括分组丢弃方式(loss profile)和无缝连接可能性(probability of seamless communication)<sup>[3]</sup>以及数据传输率降低系数(rate reduction factor)<sup>[5]</sup>等新的参数。

## 2.2 移动环境对 QoS 的影响

移动环境下的 QoS 技术研究有很多问题需要解决, 而且相互之间影响比较大。比如供电问题, 目前蓄电技术还有待发展, 像手机这样的微型移动通信设备待机时间也就是以小时计, 而像 PDA、笔记本电脑这些移动计算设备在移动环境下能够连续工作的时间更短。这样, 应用的开发者不得不考虑如何让这些设备在工作时更能省电。除了电路的功耗要尽可能小以外, 还要让设备在未使用时进入“休眠”状态。

还有移动设备的跨区(handover)的问题。目前大多数移动系统都包含有固定部分(Ad hoc 除外), 这些固定部分也组成一个网络, 固定部分有诸如基站、移动支持节点等固定节点, 移动节点依托它们与固定节点通信或是相互之间进行通信。这些固定节点每一个负责管理和支持一定区域内的移动节点。蜂窝系统就是这样的典型例子: 每个通信单元中心有一个基站, 它发射的无线电波覆盖一定的地理范围, 每个单元有六个其它单元与之相邻, 当移动设备进入某一个单元时, 它只能与该单元区域内的基站进行通信, 通过该基站与本单元或其它单元区域中的固定或移动节点相通信。这样, 当一个移动节点从一个区域离开进入另一个区域时, 其通信必然受到影响。最起码, 它要采用一定的机制断开与旧通信单元的连接, 还要在新的通信单元上注册, 这就是“跨区”。为了使跨区时移动节点上的应用不受到影响, “无缝跨区”就自然而然地成为一个问题, 它是指跨区对用户是透明的, 即用户跨区时感觉不到自己接受服务的变化。实现无缝跨区就要制定相应的跨区协议, 它的工作包括基站对移动节点的身份验证、在旧区上注销和在新区上登记、资源的重新分配、更新移动节点相关的参数, 等等<sup>[10]</sup>。实现无缝跨区有时候需要用到环境信息, 而像资源预留这样的技术还必须要考虑跨区问题。

移动网络完成跨区的同时, 移动主机也就完成了迁移。整个跨区过程可以分为迁移和跨区准备、迁移开始和跨区开始、迁移结束和跨区结束等阶段<sup>[10]</sup>。移动网络 N 中的移动主机 M 从基站 BS<sub>1</sub> 的服务区域迁移到基站 BS<sub>2</sub> 的服务区域的过程如图 1 所示。

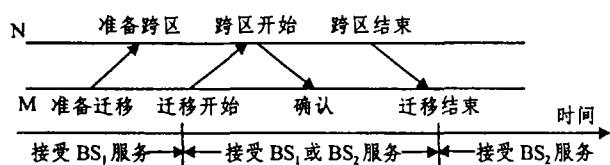


图1 跨区和迁移过程

其中在重叠服务区域可以由 BS<sub>1</sub> 或 BS<sub>2</sub> 为 M 提供服务, 或是两者都为 M 提供服务, 而由 M 进行选择接受谁的服务。跨区协议要对整个跨区过程的各时间段移动网络和移动节点所做的工作进行明确的规定, 才能实现跨区的无缝性。

由于移动环境的特性, 很难为移动用户提供服务质量的“硬保证”, 这一点与固定网络中的 QoS 是不同的。在环境发生变化时, 应用要考虑改变对相应的 QoS 参数的要求, 有时需要在用户可忍受的情况下降低为其提供的服务质量等级<sup>[4]</sup>。

硬件的发展对 QoS 技术的影响也非常大。有些以前需要解决的 QoS 问题在硬件技术发展后可能就不再需要研究, 例如, 高度依赖通信带宽的应用在带宽较低的时代需要采用许多技术来保证 QoS, 而在无线通信技术发展使数据传输率大幅提高后这些技术可能就不再需要了; 同时, 硬件技术的发展可能还会给 QoS 技术提出新的要求, 例如, 狭小的手机屏幕只需要很小的分辨率就可以了, 而对于笔记本电脑这样大的屏幕, 低分辨率往往是用户不能忍受的。当然, 有些问题不光是靠硬件的发展就能解决的, 而必须要采用一定的软件技术才能实现 QoS 的保障。

## 3. 主要研究内容

目前关于移动计算环境下 QoS 管理和控制技术方面的研究主要有环境识别、适应性管理以及资源管理和资源预留等<sup>[2]</sup>。

### 3.1 环境识别

与环境识别技术相关的研究已经有很长时间, 例如移动定位、识别设备类型、甚至当前的天气情况等等, 但是随着移动应用的不断发展, 环境的概念被赋予更多新的内容。

这里的环境是指能够对移动应用产生影响的所有因素。包括网络连接特性、不同连接方式的代价、带宽、当前位置、硬件和软件特性以及用户习惯等等<sup>[2]</sup>。在固定环境下, 通常要求环境能够适应系统, 而在移动计算中, 对环境提出要求是不现实的, 相反, 系统应该适应环境, 因此首先必须要有一定的环境识别技术。移动设备的千差万别要求移动应用能够识别目前的硬件特性和操作系统, 需要位置信息的应用还要求能够知道目前移动设备的位置, 等等。

环境识别可以说是 QoS 中的一项关键技术, 它往往是系统提供服务质量保证的前提。如移动计算中最常见的跨区问题, 移动单元需要知道自己当前受到哪个(些)固定节点的服务, 要能够判断是否在某一个固定节点的服务区域之内。对于用户来说通信费用往往是需要首先考虑的问题, 因此系统应该能根据当前的通信环境和计费方法来选择不同的通信方式, 等等。

位置识别几乎是所有具有 QoS 管理和控制能力的移动应用都要考虑的问题。室外环境下可以采用 GPS 进行定位, 室内环境下除了采用 ping 的办法来获得到某一固定节点的距离外, 发射源还可以同时发射无线电波和超声波, 移动主机通过比较接收到的超声波和无线电波, 就可以计算出到发射源之间的距离<sup>[11]</sup>。

对于当前时间、温度、湿度、光线强度、通信带宽、天气状况等基本的物理环境信息, 可以通过计时器、探测器、传感器、光电转换器等设备以及通过连接固定网络中含有该信息的服务器加以获取。

对于用户行为这样较高层次的环境需要复杂的摄像跟踪、参考用户的行动计划表、采用人工智能技术等方法来获取。而像有些环境如用户当前的情绪则几乎无法获得<sup>[25]</sup>。

环境识别技术与硬件的发展结合得相当紧密, 以上的各种获取环境的方法都存在一定的误差, 尤其是对于与用户相关的很多环境目前还无法识别。

### 3.2 适应性管理

移动环境和固定环境的一个重要差别是移动环境由于受到自然因素和移动性的影响, 其通信能力会时常变化, 而且目前移动设备没有统一的标准, 各种移动设备软件和硬件环境差别很大。在环境发生变化时, 系统应该能够采用一定的方法来消除这些影响, 例如采用缓存技术来消除延迟抖动, 等等。

适应性管理对于有些应用来说非常重要,例如在更换连接时系统可能要变换计费方式。

在通信能力经常波动的移动环境下,具有适应能力的系统通常的作法是在满足一个最低服务质量的前提下,采用尽力而为(best-effort)方式进行数据传输,为用户提供服务。

S. W. Lu 和 K. W. Lee 等人提出了适应性服务的概念<sup>[12]</sup>,并提出了适应性服务应该具有的三个特征:一是提出 QoS 边界的概念,二是增加与移动相关的 QoS 参数,三是增加更多的服务分类等级。

Dan Chalmers 和 Morris Sloman 等人则提出了一种分层的体系结构来进行适应性的管理和控制,该系统由三个组件组成,分别是环境管理器、资源管理器和 QoS 管理器。环境管理器负责获取并描述当前的环境信息,资源管理器负责底层资源的管理、分配和运转,当前者得知环境发生变化时,会告知后者如何操作,比如进入不同性质的网络时需要后者用不同的方法访问资源。QoS 管理器则通过与资源管理器协商获取需要的资源,并为移动应用进行资源预留。

Bobby Vandalore 和 Raj Jain 等人提出了一种反馈机制来实现适应性管理:移动应用定期地向具有 QoS 管理和控制能力的高层控制节点(如 MSS)发送一种特殊的反馈报文,该反馈报文包括从物理层到应用层的各层反馈信息,高层控制节点的相关模块根据反馈信息进行适应性管理和控制<sup>[4]</sup>。其反馈报文的格式如图2所示。

反馈分组头
应用层反馈
网络层反馈
数据链路层反馈
物理层反馈
反馈分组尾

图2 反馈分组格式

除了基于通信条件的适应性研究外,基于通信费用的适应性研究也是一项重要的研究内容,一个典型的例子是在按通信时间计费的系统中,进行文件传输的应用应该在数据传输率较高的时候打开连接接收或发送数据,而在数据传输率很低的情况下关闭连接,进入休眠状态。

### 3.3 资源管理和资源预留

移动环境千差万别,不同的移动网络通信能力大不相同,稳定和可靠的传输在移动环境下往往得不到保证。在资源稀缺的情况下,对资源的管理和访问控制就十分必要。尤其在跨区通信时,需要预留资源来保证 QoS。资源预留的思想在固定网络上就已经提出并且被实现,资源预留协议(RSVP)就是由 USC ISI 和 PARC 提出并写入 RFC 的标准协议,该协议定义了 Path, Resv 等一些消息,提出了预留路径等概念,并使用软状态对预留路径进行维护,以及采用隧道技术等方法来保证在网络中进行可靠的资源预留<sup>[19]</sup>。

A. K. Talukdar 和 B. R. Badrinath 等人基于 RSVP 设计了一种在综合服务分组交换网络的移动计算环境下为无线用户的实时应用提供 QoS 管理和控制的协议,称为 MRSVP<sup>[7]</sup>。该协议最大的改动之处是它增加了主动预留和被动预留的概念,它在每一个移动发送者可能进入的区域和每一个移动接收者可能进入的区域之间建立资源预留路径。发送者和接收者之间的资源预留路径称为主动预留路径,其它为被动预留路径,如图3所示。

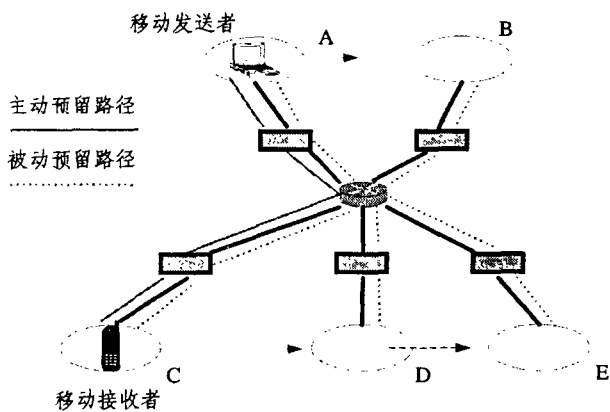


图3 主动预留与被动预留

执行该协议的一个很大难题是系统需要预测移动节点的活动范围。I. Mahadevan 和 K. M. Sivalingam 则对移动环境下的资源预留技术在蜂窝网络上进行了试验<sup>[5]</sup>。由于每个通信单元相邻的通信单元数是固定的(6个),所以资源预留路径仅在发送者与接收者当前所在的通信单元(主动预留)以及它相邻的通信单元(被动预留)之间建立。如果能够预测移动节点的移动方向,则可以进一步提高协议的执行效率。当移动主机迁移时,需要重新建立预留路径,如图4所示。

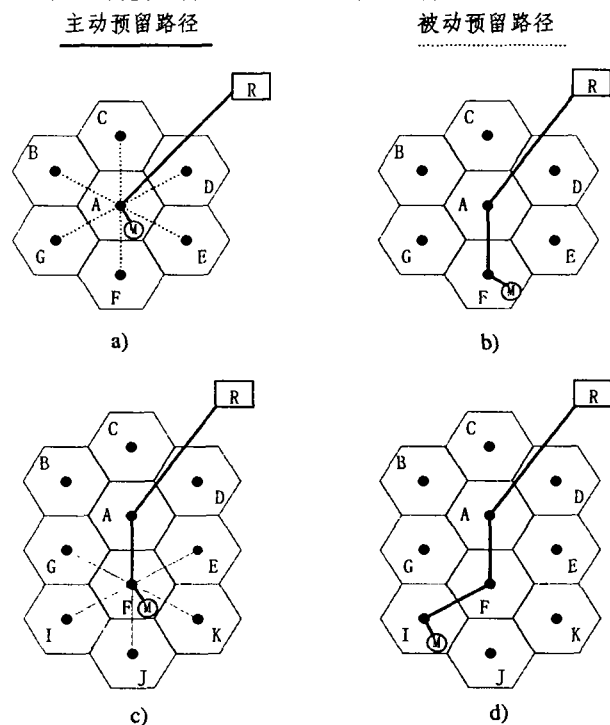


图4 资源预留路径建立方法

其中 A~G 及 I, J, K 为蜂窝通信单元, M 为移动节点, R 为具有路由功能的上层节点,一端可以与这些蜂窝单元中的基站直接建立连接,另一端与固定网络相连(未画出)。M 每一次进行迁移时系统都要重新建立资源预留路径。从 R 到 M 的路由路径上,原来的被动预留路径变为主动预留路径。

由于每一次建立新的资源预留路径都要以前面的预留路径为基础,这种方法会导致资源预留路径的不断增长,后来他们对预留路径的建立方法做了进一步的改进,提出了 QoS 区域的概念。一个 QoS 区域由若干个通信区域组成,在建立被动预留路径时,如果移动节点当前所在区域的基站发现一个相邻区域的基站与自己属于同一个 QoS 区域,则和原

来一样在自己和它之间建立预留路径;如果该基站与自己分属不同的 QoS 区域,则在该基站与上层节点之间直接建立预留路径,如图5<sup>[9]</sup>;

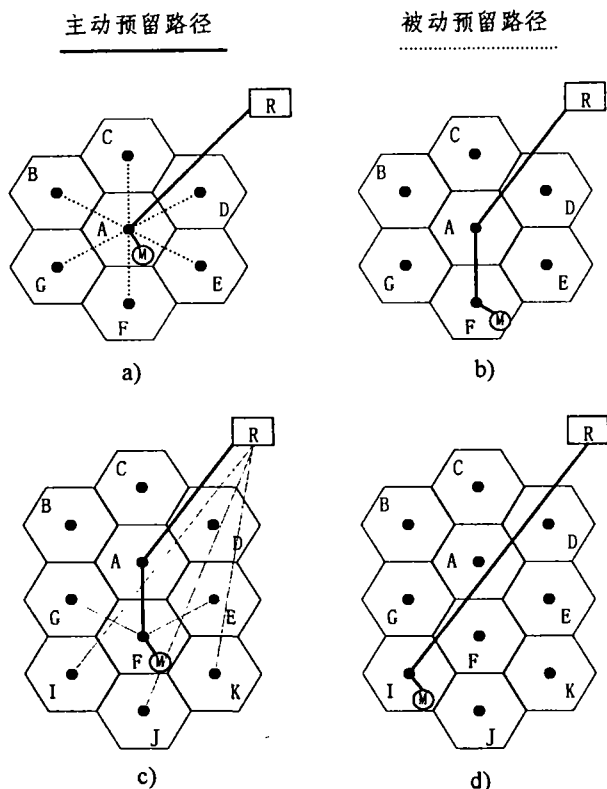


图5 改进的预留路径建立方法

其中 A~G 同属一个 QoS 区域,而 I、J、K 则属于别的 QoS 区域。这样,当 M 从 F 迁移到 I 时,直接在 R 与 I 之间建立资源预留路径。

这种方法在一定程度上解决了资源预留路径不断增长的问题,但是如果一个移动节点仅在一个 QoS 区域中不断移动(如将一个学校的范围划分为一个 QoS 区域,移动用户总是在该范围内的不同蜂窝单元间来回移动),预留路径还是会增加到令人难以忍受的程度。解决这个问题的办法,除了变更划分 QoS 区域的粒度,可以设置预留路径的最大存在时间来改进。

以上这些研究内容并不是相互独立的,例如实现适应性管理和资源预留往往要用到环境信息,因此,在设计实现能进行 QoS 管理和控制的移动计算系统时,这些问题都要进行考虑。

**结束语** 目前,关于移动计算领域的研究仍然很不成熟,其中的很多课题都有待研究,QoS 便是其中之一。本文对新兴的移动计算环境下的 QoS 管理和控制技术的一些基本概念、存在的主要问题、基本实现方法和主要研究内容进行了综述。

当然,移动计算环境下的 QoS 技术的研究范围非常广阔,有些问题任何移动系统都不得不考虑,例如无缝跨区问题,以及如何省电进行通信的问题。而像设计易于操作的系统,设计良好的人机界面以及容错技术等方面的问题也与 QoS 技术密切相关。

## 参考文献

- 1 Imielinski T, Korth H F. Mobile Computing. Kluwer Academic Publishers, 1996
- 2 Chalmers D, Sloman M. A Survey of Quality of Service in Mobile Computing Environments, Imperial College, London, TR., 1999
- 3 Singh S. Quality of Service Measures in Mobile Computing. The

- Journal of Computer Communications, 1996, 19
- 4 Vandalore B, Jain R, Fahmy S, Dixit, AQuaFWiN: Adaptive QoS Framework for Multimedia in Wireless Networks and its Comparison with other QoS Frameworks. In: 24th Conf. on Local Computer Networks, Oct. 1999
- 5 Mahadevan I, Sivalingam K M. An Experimental Architecture for providing QoS guarantees in Mobile Networks using RSVP. IEEE, 1998
- 6 Campbell A T. A Quality of Service Architecture: [A thesis submitted for the degree of Doctor of Philosophy]. Computing Department, Lancaster University, Jan. 1996
- 7 Talukdar A K, Badrinath B R, Acharya A. MRSVP: A Resource Reservation Protocol for an Integrated Services Network with Mobile Hosts. The Journal of Wireless Networks, 2001, 7(1)
- 8 Kunz T. Adaptive Mobile Applications. University of Waterloo, Aug. 1997
- 9 Mahadevan I, Sivalingam K M. An Architecture for QoS Guarantees and Routing in Wireless/Mobile Networks. ACM, 1998
- 10 Endler M, Nagamuta V. General Approaches for Implementing Seamless Handover. ACM 1-58113-511-4/02/0010, 2002
- 11 Priyantha N B, Chakraborty A, Balakrishnan H. The Cricket location-support system. In: Proc. of the Sixth Annual ACM Intl. Conf. on Mobile Computing and Networking, Boston, MA, Aug. 2000
- 12 Lu S W, Lee K W, Bharghavan V. Adaptive Service in Mobile Computing Environments. In: Proc. IFIP IWQoS'97, 1997
- 13 Badrinath B R, Talukdar A. IPv6 + Mobile IP + MRSVP = Internet Cellular Phone?. The Journal of Wireless Networks, 1999, 5
- 14 Chalmers D, Sloman M. QoS and Context Awareness for Mobile Computing. Imperial College, London SW72BZ, U.K. The Journal of Lecture Notes in Computer Science, 1999, 1707
- 15 Fallmyr T. Stabell-kulo T. QoS applied to security in mobile computing. University of Tromso, NORWAY, June 1997
- 16 Belof J. Networks and Mobile Computing. The OpenGroup Conference, San Jose CA, Feb. 2001
- 17 Bechler M, Ritter H, Schiller J H. Quality of Service in Mobile and Wireless Networks: The Need for Proactive and Adaptive Applications. IEEE 0-7695-0493-0/00, 2000
- 18 Snow A, Varshney U, Malloy A D. Reliability and Survivability of Wireless and Mobile Networks. IEEE 0018-3162/00, 2000
- 19 Resource Reservation Protocol, Cisco Internetworking Technologies Overview Handbook, Chapter 43. <http://www.cisco.com/univercd/cc/td/doc/cisintwk/ito-doc/rsvp.htm>, June 1999
- 20 Das S K, Jayaram R, Kakani N K, Sen S K. A call admission and control scheme for quality-of-service provisioning in next generation wireless networks. The Journal of Wireless Networks, 2000, 6
- 21 Zaruba G V, Chlamtac I, Das S K. A Prioritized Real-Time Wireless Call Degradation Framework. In: Proc. of the 3rd ACM Intl. Workshop on Modeling analysis and simulation of wireless and mobile systems, 2000
- 22 Geihs K. Analysis of Adaptation Strategies for Mobile QoS-Aware Applications. In: Proc. of the 5th ACM Intl. Workshop on Modeling analysis and simulation of wireless and mobile systems, Sep. 2002
- 23 Boukerche A, Das S K, Fabbri A. SWiMNet: A Scalable Parallel Simulation Testbed for Wireless and Mobile Networks. Wireless Networks, 2001, 7: 467~486
- 24 De S, Das S K, Wu H Y, Qiao C M. Trigger-Based Distributed QoS Routing in Mobile Ad Hoc Networks. ACM SIGMOBILE Mobile Computing and Communications Review, 2002, 6(3)
- 25 Chen G L, Kotz D. A Survey of Context-Aware Mobile Computing Research. Dept. of Computer Science, Dartmouth College, Nov. 2000
- 26 Abowd G D, Mynatt E D. Charting Past, Present, and Future Research in Ubiquitous Computing. ACM Transactions on Computer-Human Interaction, 2000, 7(1)
- 27 Brodersen R. InfoPad - An Experiment in System Level Design and Integration. In: Proc. of the 34th annual conf. on Design automation conf. June 1997
- 28 Braam P J. The Coda Distributed File System, Linux Journal, # 50, June 1998
- 29 单志广, 林闯. 服务保障走入 Web—Web QoS 成为热点技术. 计算机世界报, 2002(41) B1, B2, B3
- 30 林闯, 等. 多媒体信息网络 QoS 的控制. 软件学报, 1999, (10) 10: 1016~1024
- 31 张信明, 陈国良, 刘峰. Internet QoS 控制技术综述. 计算机科学, 2002, 29(3)