

文章编号:1001-9081(2009)02-0507-04

一种适合于区分服务网络的多播路由算法

高 茜¹, 万小燕²

(1. 南京师范大学 数学与计算机学院, 南京 210046; 2. 南京电研电力自动化股份有限公司, 南京 210061)

(gaoqian@njnu.edu.cn)

摘 要:提出一种适合于 DiffServ 网络的 QoS 多播路由算法 PQMRD, 它针对组成员不同类别的 QoS 请求采取不同的路由策略来选取路径, 同时进行分类的接纳控制和资源预留。实验结果表明, PQMRD 缓解了现有多播算法中因所有业务选择相同路径而引起的服务类间的不公平性问题。

关键词:多播; 区分服务; 服务质量; 服务类间的公平性

中图分类号: TP393 **文献标志码:** A

A new multicast routing protocol with QoS constraints

GAO Qian¹, WAN Xiao-yan²

(1. School of Math and Computer Science, Nanjing Normal University, Nanjing Jiangsu 210046, China;

2. Nanjing Dianyan Power Automation Company Limited, Nanjing Jiangsu 210061, China)

Abstract: A QoS multicast routing algorithm for DiffServ networks called PQMRD (Per-class QoS Multicast Routing in DiffServ networks) was proposed. It used different routing strategy to select route in terms of different QoS requirement; in addition, admission control and resource reservation were performed while selecting route. It is proved by the simulation result that inter-class unfairness problem is alleviated.

Key words: multicast; DiffServ; Quality of Service; inter-class fairness

0 引言

随着网络技术与多媒体技术的迅速发展, 涌现了大量的实时多媒体应用, 它们对网络带宽的需求量比较大, 同时对延时、延时抖动等服务质量也有了严格的要求。IETF 提出的区分服务 (DiffServ, DS) 模型^[1]通过合理利用有限的网络资源来保证用户申请的服务质量, 并且以其较好的可扩展性成为了最有可能在下一代网络主干中实施的 QoS 模型, 采用多播传输技术可以实现网络资源的高效利用, 因此二者之间的互补结合便成为一个有效的多媒体应用的 QoS 解决方案。然而, 由于基本结构上的冲突, 二者的简单集成存在着非预留资源子树 (Neglected Reservation Subtree, NRS)、可扩展性、异构性、多点到多点的多播传输策略、服务类间的公平性、以及共享树的流量监控等问题^[2]。

目前有不少学者致力于 DiffServ 网络中多播传输的研究, 可以把他们的研究成果分为两大类。第一类主要侧重于突出 DiffServ 体系结构的优点, 因此要求内部路由器不保留或者减少单个流的状态, 即多播树中的组状态, 因而它们的可扩展性比较好^[3-8]; 第二类解决方案则是从多播的角度出发, 允许内部路由器维护多播组状态, 因此依然存在传统多播中的可扩展性问题^[9-10]。通过对多播请求进行接纳控制和资源预留, 这些方案都能有效解决 NRS 问题。对于多点到多点的多播传输策略问题以及共享树的流量监控问题, 目前的解决方法比较单一, 也不够成熟, 还需要进行进一步的研究^[11-12]。对于服务类间的公平性问题, 目前大部分的解决思路是设计更好的队列管理算法以及调度算法, 这比较适合于单播业务。由于在 DiffServ 模型中, 服务质量的保证是与 IP 路由机制以

及多播路由机制互不相关的, 也就是说, 无论流量属于哪个服务类, 只要它们通过相同的入口路由器和出口路由器, 那么它们所采用的路径就一定是相同的。因此, 当高优先级服务的流量数目很大时, 必然会给低优先级服务流带来负面的影响。对于多播服务来讲, 用户的动态性以及异质性决定了多播树的不断变化, 也就需要不断寻找新用户到多播树的满足 QoS 的路由。现有的 QoS 路由算法并没有考虑到 DiffServ 网络的特点^[13-14], 因此设计适合 DiffServ 网络的路由算法十分必要。目前已有通过改进单播路由算法解决服务类间的公平性问题的研究, 如: DiffServ 域内奖赏服务的逐跳路由算法^[11], 减少了单播业务中奖赏服务流对其他服务流的负面影响。

本文对 DiffServ 网络中的资源分配策略进行改进, 提出了一种适合于 DiffServ 网络的分类的 QoS 多播路由算法 (Per-class QoS Multicast Routing in Diffserv networks, PQMRD), 实验结果表明, 本算法成功地缓解了服务类间的公平性问题。

1 改进的 DiffServ 网络

本文首先对 DiffServ 网络中的带宽代理 (BandWidth Broker, BB)^[15]进行了扩展, 使其除了完成已有的域内和域间资源分配外, 还能够完成多播组的管理, 包括多播组的登记、多播树的计算、多播流量资源的分配以及多播树信息的发布。

在 DiffServ 中的资源分配是以聚类为单位, 并且每一类的资源分配范围都是以 DS 域的入口路由器为起始点, 而以出口路由器为终止点。对于单播业务来讲, 只要 DS 域的入口流量满足业务流量规范, 就可以保证进入 DS 域流量等于流出 DS 域的流量。这种分配方式便于用户和服务提供商

收稿日期: 2008-08-14; 修回日期: 2008-10-10。

基金项目: 国家自然科学基金资助项目 (60673070); 江苏省普通高校自然科学研究计划资助项目 (08KJB510007)。

作者简介: 高茜 (1975-), 女, 陕西韩城人, 讲师, 博士, 主要研究方向: 高性能网络协议和算法、QoS 路由协议和算法; 万小燕 (1976-), 女, 江苏南京人, 助理工程师, 主要研究方向: 网络路由。

之间的协商,比较简单、灵活。但是对于多播业务来讲,多播数据的传输方式并不是点到点的,而是沿着多播树从一个入口路由器发送到多个出口路由器,如果还是采用边界到边界的服务等级协定(Service Level Agreement, SLA)协商和边界到边界的业务量调节规范(Traffic Conditioning Specification, TCS)来进行资源分配,便会造成网络资源的浪费和接纳控制的失败。因此本文改进了 DiffServ 网络的资源分配策略,进一步细化 TCS,将边界到边界的路径式资源分配策略改为基于链路的分配策略。也就是说,将 TCS 中服务区间的限定改为相邻的两个路由器,即限制在确定的链路上。对每类业务,在每条链路上都有一个与之相对应的业务量简档以及 PHB、处理措施等参数。这个细化的 TCS 并不违背 DiffServ 核心无状态的特性,它是存储在 BB 和边界路由器中,由 BB 根据 SLA 和域内资源情况确定 TCS,并下发至各边界路由器实施执行。细化后的 TCS 为下一节提出的 PQMRD 算法做好了充分的准备。

2 PQMRD 算法

PQMRD 算法的基本思想是:针对组成员的不同类别的 QoS 请求采取不同的路由策略来选取路径,在选路的同时进行分类的接纳控制和资源预留,从而缓解了所有业务选择相同路径而引起的服务类间的不公平性。

PQMRD 算法的设计基于以下几个前提假设:

- 1) 存在一种分类机制,使得各成员可以根据自己的需求申请相应的服务;
- 2) 存在一种统计机制,使得 BB 可以合理地将域内各链路的带宽资源划分给各服务类;
- 3) 存在一种资源预约和资源调度机制,能合理地根据服务等级调度系统资源,保证预约过资源的用户享受相应的服务质量保证;
- 4) 存在一种测量机制,可以综合资源预约、资源调度和其他机制,测量出每个路由器和每条链路的动态状态。

这些假设都是当前未能很好解决的问题,但是不属于本文的研究范围,因此不做深入的讨论。

2.1 算法思想

PQMRD 算法是基于源路由的,它将多播路由层面与多播数据层面分离开来,组成员的加入、离开请求以及相应的 QoS 需求都发送到带宽代理 BB。BB 收到组成员的加入多播组 G 的请求后,根据它的 QoS 需求和 DiffServ 域中该服务类的资源预留情况进行路由计算,并把计算出来路径信息通过扩展的公共开放策略服务(Common Open Policy Service, COPS)^[16]向分发至入口边界路由器。入口边界路由器收到该成员的路径信息后,查询多播组 G 对应的多播树 T,并将新成员的路径信息添加到 T 中。接下来,在入口边界路由器对多播树信息进行编码,并封装到每个多播数据报头中,在内部路由器则根据报头中保存的信息进行路径查询,根据查询结果进行报文复制、重新标记与转发,最终在出口边界路由器对报头解封。组成员需要离开多播组时,同样将离开请求发送到 BB,再由 BB 通知入口边界路由器更新多播树。

根据目前 DiffServ 中典型的两类服务:奖赏服务(Premium Services, PS)和确保服务(Assured Service, AS),将网络带宽分按照服务类别分为 EF、AF、BE 和共享四个区间。其中 EF 表示服务采用 EF PHB,它保证任何时候接受此服务的流的离开速率大于等于设定速率,并且与其他 PHB 共存时,EF 总是优先级最高的,因此它用来实现 PS 服务。而 AF 表示服务采用 AF PHB,它可以保证该服务的流在任何时候

都能获得预约的最小带宽,用来实现 AS 服务。BE 区间的带宽用于尽力而为的服务。共享区间则用于在前三类的带宽资源不满足用户要求时提供的少量补充资源。这里需要指出的是,对于带宽区间的划分只是一种逻辑上的概念,目的是在选路时能够考虑到链路上已有的业务流情况。物理上对资源的预留是通过配置基于链路的 TCS 来实现的。

PQMRD 算法根据每个成员的 QoS 请求在对应的区间按照不同的路由策略计算路由。

路由策略 1 对于申请 PS 服务的用户,采用满足带宽约束和延迟约束的最小延迟路由策略。

路由策略 2 对于申请 AS 服务的用户,采用满足带宽约束的最小代价路由策略。

路由策略 3 对于 BE 用户,采用最大带宽路由策略。

算法流程如图 1 所示。对于每一个新成员的加入请求,首先根据其 QoS 请求判断它属于哪个服务类别,如果属于 PS 类,则根据多播树上各节点已经预留的服务等级信息选取 k 个 EF 等级的树上节点组成一个候选加入节点集合,这 k 个节点中包括本域中多播数据的入口节点,其余节点为到该入口节点延迟最小的 $k-1$ 个节点;然后根据路由策略 1 分别计算候选集合中的树上节点到新成员之间的路径,也就是满足带宽约束的最小延迟路径。在这里采用改进的 Dijkstra 算法,首先对网络中的 EF 区间剩余带宽不满足带宽要求的链路延迟进行预处理,将它设为 $+\infty$;然后在处理后的网络上使用 Dijkstra 算法计算最小延迟路径;最后,在这 k 条路径中选择延迟最小的路径作为候选路径,并对该路径进行资源预留,如果预留成功,则算法结束,否则将预留结果反馈给新成员,并重新协商 QoS,重新进行路由计算。

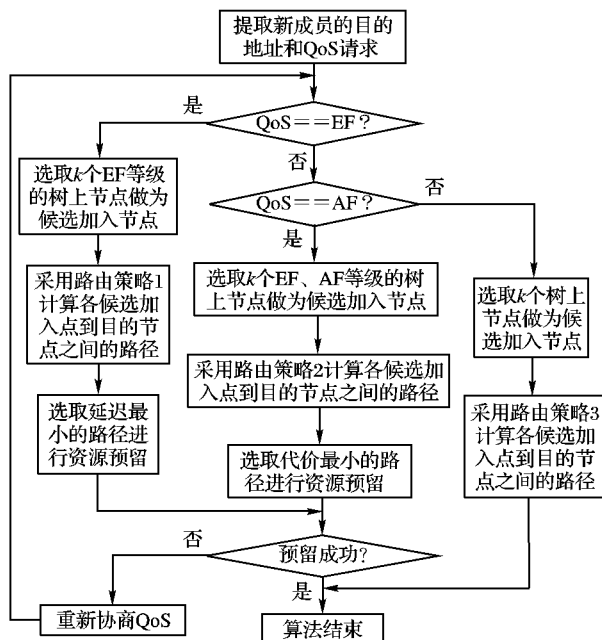


图1 PQMRD 算法流程

对申请 AS 类服务的新成员的路由计算与 EF 类新成员的类似。不同的是,在进行候选加入节点集合选择时,除了树上服务等级为 AF 的节点外,服务等级为 EF 的节点也可以选择。另外,在计算候选节点到新成员的路径时采用路由策略 2,即满足带宽约束的最小代价路由策略。该策略使用的也是改进的 Dijkstra 算法。首先对网络中的 AF 区间剩余带宽不满足带宽要求的链路的代价进行预处理,将它设为 $+\infty$;然后在处理后的网络上使用 Dijkstra 算法计算最小代价路径。

对于没有服务质量要求的新成员将它们归到 BE 类,算法将采用路由策略 3 在网络中 BE 区间上计算 k 个树上节点到新成员的最大瓶颈带宽路径,并从此 k 条路径中选择瓶颈带宽最大的路径作为候选路径。

由于 PQMRD 对各类服务请求的路径计算是限制在该类的剩余带宽区域中的,因此各服务类间互不影响。但是,严格限制区间会产生这样一种情况,在计算路由时可能会由于发现只欠缺少量资源而不能接纳某个用户,这样会导致网络资源的浪费和用户接纳率的降低。因此在划分带宽区间时预留出一个共享区间,它的作用是弥补各类资源的少量不足,减少资源的浪费,提高用户的接纳率。共享区间通常不是很大,占整个链路资源的 10% 左右。

2.2 相关问题

在现实网络中,单播业务总是与多播业务共存,因此需要考虑 PQMRD 算法对多播业务的影响。PQMRD 算法在计算路由时将选路操作限制在一定区间,然而,在路由器分配资源时,对单播流和多播流的调度处理操作是相同的,也就是说,只要同属于一个聚类,它们就会排在同一个队列里,采用同样的调度策略调度报文。所以必须在划分区间时考虑到单播流量,可以有两种方法来考虑单播流:一种方法是直接将单播流与多播流的带宽分隔开来,也就是说 PQMRD 算法只在多播流的带宽区域中寻路;另外一种方法是采用与 PQMRD 算法类似的单播路由算法,使得单播路由的选择也需要根据服务类别限制在相应的带宽区域中。本文采用后一种方法。

对单播流也采取源路由方式,在入口边界路由器利用 IP 报头中的源路由选项字段,将事先计算好的路径信息放置其中,中间路由器根据各分组中的路由信息来完成选路功能,并不依赖路由器的路由表,出口路由器负责将该分组中的路径域删除。

3 算法分析

3.1 正确性分析

在算法的计算过程中,无论组成员申请的是哪个服务等级,算法都从已有树上选择 k 个树上节点作为该成员加入多播树的候选加入节点。而这 k 个节点中包含多播树的入口节点,因此,算法一定能依据路由策略找到满足 QoS 条件的路径。例如,对于申请 PS 服务类的组成员,算法首先会对网络中不满足带宽约束的链路的延迟参数进行预处理,并在处理后的网络中以延迟作为参数计算最小延迟路径,如果该路径不能满足延迟约束条件,则显然不存在满足带宽和延迟约束的路径。否则,算法将从多个加入点计算出的满足条件的路径中选取最小延迟的路径加入多播树。另外,因为在选择候选加入点时已经考虑了它们预留的服务等级,所以最后生成的树一定是满足服务质量等级的多播树。

3.2 公平性分析

算法对不同服务类的路由计算是在不同的带宽区间上进行的,因此可以保证高服务等级的业务集中在某条链路上抢占低服务等级业务的带宽资源。当然,仅仅在选路上区分路由是不够的,服务类间的公平性维护还需要资源调度机制的支持。

3.3 时间复杂性

PQMRD 算法的核心是分类计算策略,虽然对申请不同类

服务的组成员采用的是不同的路由策略,但是它们的核心都是 Dijkstra 算法,每个用户的加入的运行时间为 $O(|V|^2|E|)$ 。如果同时有 m 个用户申请加入多播组,则整个算法的运行复杂度为 $O(m|V|^2|E|)$ 。

4 仿真分析

PQMRD 的仿真实验在 NS^[17] 进仿真平台下进行。NS 仿真器中已经实现了基本的 DiffServ 网络仿真模块,其中包括流的区分功能,能够对输入流进行分类;同时,在 IP 层增加了多个排队机制,使得可以根据报头灵活地选用不同的排队调度算法。本文在此基础上进行了扩展;扩展了单播路由协议 OSPF 以实现 QoS 状态的扩散;添加了 BB 代理和 PQMRD 路由代理;修改了接纳控制代理 ADC 和 Policy 的实现。仿真实验中的网络拓扑是依旧是根据 B. Waxman 提出的模型^[18] 生成的,取 $\alpha = 0.2$, $\beta = 0.4$,随机拓扑图的节点的平均度数为 4。生成了一个拥有 30 个边界节点,70 个内部节点的 DiffServ 域。其中,链路可用带宽分布在 $[0.56, 2]$ Mbps,而链路延迟由网络节点之间的笛卡儿距离决定。

在实验中选择边界节点作为多播组的源节点和成员节点,它们是随机选择的。多播业务所需要的带宽为 $[0, 2]$ Mbps 间的随机数,多播业务所需的延迟上限在 $[0.1, 1]$ s 间变化。给 EF、AF 和 BE 的流量分配的带宽区域分别为 40%、30% 和 20%,剩下 10% 则为共享流量,各链路的 TCS 也照此进行配置,根据各类流的实际流量与 TCS 决定它的 DSCP,即它所对应的 PHB。

首先对 PQMRD 和 DSMCast^[4] 以及直接在 DiffServ 中采用 PIM-SM 路由的方案进行了链路资源利用率的比较,其中平均链路利用率 = 所有链路利用率之和/所有链路的个数。如图 2 所示,在多播组规模比较小时,这三种算法的链路利用率基本相同,但是随着多播组规模的增加,PQMRD 的平均链路利用率明显高于其余两种方案,这是因为,DSMcast 和 PIM-SM 都采用最短路径路由,流量都集中在最短路径的链路上,而 PQMRD 对各类流量采用不同的路由策略,从而扩大了寻路的范围,使得流量可以分布到各条链路上。由此可见 PQMRD 方案可以更加有效地利用网络资源。

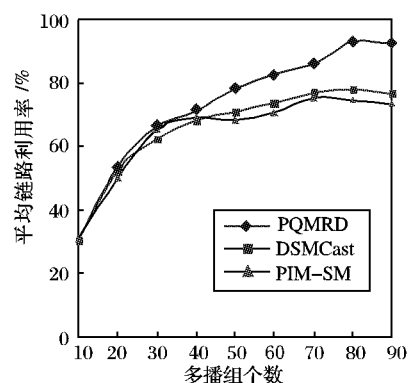


图2 平均链路利用率的比较

图3中比较了EF业务量对AF用户的影响,其中用户接纳率 = 接纳用户数/请求用户数。从图中可以看出,在EF多播业务比较少时,PQMRD和DSMCast的用户接纳率都为100%,当EF业务量增加,PQMRD的AF用户接纳率依然能保持一个较高的水平,但是DSMCast的AF用户接纳率的下

降趋势却很快。这是因为在 PQMRD 中对 AF 业务采取和 EF 不同的路由策略,并且进行分类的资源预留。在 EF 多播业务达到 35 个以上,两者的用户接纳率又趋于相同,这是因为此时 PQMRD 中 AF 业务也已经达到饱和。对于 PIM-SM,由于它本身没有接纳控制,所以不具有可比性。

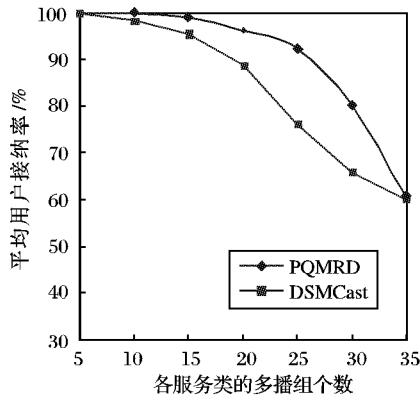


图3 AF用户平均接纳率的比较

图4和图5分别描述了EF业务量对AF和BE业务的平均报文丢弃率的影响,其中报文丢弃率 = (源节点发送的报文数 - 成员收到的报文数) / 源节点发送的报文数。从图中可以看出,PQMRD中EF对其他两类流的影响比较小,DSMCast次之,而在PIM-SM中影响最大。这是因为在PQMRD算法将各类流的路由分散在不同的链路上并且对每条链路分别进行资源分配,也在一定程度上减少了高服务等级的流对低服务等级的影响,从而缓解了服务类间的公平性问题。

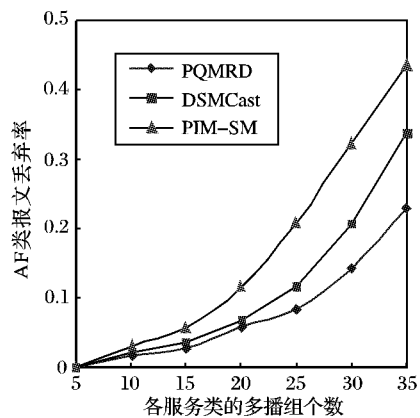


图4 AF类报文丢弃率的比较

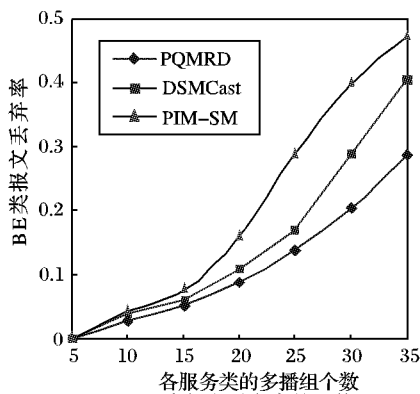


图5 BE类报文丢弃率的比较

5 结语

在网络飞速发展的今天,在主干网络中实现 QoS 多播成

为实时多媒体应用发展的巨大推动力。由于 DiffServ 网络体系结构与 IP 多播模型存在着基本矛盾,DiffServ 网络中的多播传输存在着 NRS、可扩展性、多播组的异构性以及服务类间的公平性等问题,目前大多数的研究集中在前几个问题上,对服务类间的公平性问题的研究比较少。本文在前人研究的基础上,提出了一种适合于 DiffServ 网络的多播路由 PQMRD,它在选路时结合各链路的动态资源分配情况,区分对待不同服务等级的业务,在选路的同时进行分类的接纳控制和资源预留。分析与仿真结果证明,PQMRD 算法可以更加有效地利用网络资源,有效地缓解了服务类间的公平性问题。接下来的工作将对 PQMRD 的域间路由进行仿真实验研究。

参考文献:

- [1] RFC 2475. An architecture for differentiated services[S]. 1998.
- [2] 高茜,罗军舟. 区分服务网络中 IP 多播: 问题与解决方案[J]. 计算机研究与发展, 2005, 42(5): 823 - 629.
- [3] STRIEGEL A, BOUABDALLAH A, BETTAHARET H, et al. EBM: Edge-based multicasting in diffServ networks[C]//Proceeding of 5th International Workshop on Network Group Communications (NGC). Palo Alto: ACM Press, 2003: 3 - 4.
- [4] STRIEGEL A, MANIMARAN G. A Scalable approach for diffServ multicasting[C]// Proceeding of IEEE International Conference on Communications. Helsinki Finland: IEEE Press, 2001: 2327 - 2331.
- [5] LI ZHI, MOHAPATRA P. QoS-aware multicasting in diffServ domains[C]// Proceeding of IEEE Global Telecommunications Conference. New York: ACM Press, 2002: 2118 - 2122.
- [6] TONG S, CHANG C C. Harmonic DiffServ: Scalable support of IP multicast with Qos heterogeneity in DiffServ backbone networks[J]. Computer Communications, 2006, 29(10): 1780 - 1797.
- [7] CUI JUNHONG, LAO LI, FALOUTSOS M, et al. AQoS: Scalable QoS multicast provisioning in diffServ networks[J]. Computer Networks, 2006, 50(1): 80 - 105.
- [8] YANG BAIJIAN, MOHAPATRA P. Multicasting in differentiated service domains[C]// Proceeding of IEEE Global Telecommunications Conference. New York: IEEE Press, 2002: 2074 - 2078.
- [9] RFC 3754, IP Multicast in Differentiated Services (DS) networks [S]. IETF, 2004.
- [10] BIANCHI G, BLEFARI-MELAZZI N, BONAFEDE B, et al. QUASIMODO: Quality of service-aware multicasting over DiffServ and overlay networks[J]. IEEE Network, 2003, 17(1): 38 - 45.
- [11] WANG J, NAHRSTEDT K. Hop-by-Hop routing algorithms for premium-class traffic in diffServ networks[C]// Proceeding of IEEE Conference on Computer Communications. New York: IEEE Press, 2002: 705 - 714.
- [12] SU HENGCHI, HWANG R H. Multicast provision in a differentiated services network[C]// Proceeding of 15th International Conference on Information Networking. Beppu City, Oita Japan: IEEE Press, 2001: 189 - 196.
- [13] 高茜,李勇,罗军舟. 一种新的 QoS 约束的多播路由协议[J]. 计算机学报, 2003, 26(11): 1441 - 1449.
- [14] 高茜,罗军舟. 基于 Tabu 搜索的 QoS 多播路由快速优化算法[J]. 软件学报, 2004, 15(12): 1877 - 1884.
- [15] NICHOLS K, JACOBSON V, ZHANG L. A two-bit differentiated services architecture for the Internet[EB/OL]. [2008 - 06 - 22]. <http://irl.cs.ucla.edu/papers/twobit.pdf>.
- [16] RFC3084. COPS Usage for Policy Provisioning[S]. IETF, 2001.
- [17] UCB/LBNL/VINT. Network Simulator ns-2[EB/OL]. [2008 - 06 - 22] <http://www.isi.edu/nsnam/ns/ns-documentation.html>.
- [18] WAXMAN B. Routing of Multipoint Connections[J]. IEEE Journal Selected Areas in Communications, 1988, 6(9): 1617 - 1622.