

基于P2P的流媒体技术

王艳丽¹, 鲜继清¹, 白洁²

(1. 重庆邮电学院 通信与信息工程学院, 重庆 400065; 2. 重庆邮电学院 计算机学院, 重庆 400065)
(nmccwyl@163.com)

摘要:现有的流媒体系统一般是基于客户/服务器模式和IP组播技术,这限制了用户数量且IP组播需先得到ISP支持,故其发展受到很大限制。而在P2P网络中,每个接收数据的用户同时向外转发数据,这就充分利用了以往忽视的客户机资源。首先指出了P2P流媒体应用中所面临的挑战,并着重从应用层多播树、网络异构性、激励机制等方面阐述了P2P流媒体的研究现状,分析讨论了存在的问题,最后指出进一步的研究方向。

关键词:P2P; 应用层组播; 流媒体; 激励; 异构性

中图分类号:TP393.03 **文献标识码:**A

Media streaming based on P2P

WANG Yan-li¹, XIAN Ji-qing¹, BAI Jie²

(1. School of Communication and Information Engineering, Chongqing University of Posts & Telecommunications, Chongqing 400065, China;
2. School of Computer Science, Chongqing University of Posts & Telecommunications, Chongqing 400065, China)

Abstract: Current media streaming systems are often based on the client-server framework and IP-level multicasting supported by ISP, which limits clients amount to a great extent. In P2P network, each client receives data and forwards it to others at the same time. So large numbers of clients resource ignored before can be effectively used. Firstly challenges in P2P media streaming applications were pointed out, then some recent research results such as application-level multicast tree, heterogeneity of network and incentive mechanism and so on were introduced. Some existent problems and some remained issues for further research were discussed.

Key words: P2P; application-level multicast(ALM); media streaming; incentive; heterogeneity

0 引言

近年来,随着 eMule、Bittorrent^[1]等 P2P 应用软件的出现,对 P2P 网络的研究已广泛展开。P2P 网络中,每个终端也称节点同时具有客户端和服务端功能,每个节点将接收的内容缓存并提供给其他请求节点,这充分挖掘了 Internet 上被忽视的客户机资源,在利用率、扩展性、容错等方面具有巨大潜力。把 P2P 引入流媒体服务,将打破传统的客户/服务器模式,使服务分散化,从而减轻服务器负载并支持更大范围流媒体发布,具有潜在的应用前景。

对面向 Internet 流媒体技术,最简单的解决办法是为每个申请者建立一条发送视频流的链接。但太耗费带宽,又不能支持大量观众实时收看,申请者接收到的将是低质量视频。IP 组播相对而言是一种好的解决方法,但 Internet 中多数 ISP 不支持 IP 组播,造成其发展受限。ALM(即应用层组播)克服了 IP 组播的缺陷^[2]:无需更改网络协议和网络设备的配置,在客户机间复制和转发数据,数据报沿逻辑链路转发,数据路由、复制、转发功能均由客户机完成,客户机间建立一个叠加在 IP 网络上、实现组播业务逻辑功能性网络。

但引入 ALM 的流媒体系统仍面临如下问题:首先,如何构建有效的 ALM 树。其次,如何适应网络异构性。网络异构性指因目前 Internet 接入方式的多样性(如 ADSL、ISDN、无线

接入等),网络上不同链路、节点及终端设备在带宽容量、处理能力等方面不尽相同。最后,如何建立有效的激励机制,以确保节点间的公平性。需要指出:我们称只享受服务而不提供服务的节点为“Free rider”,如何防止“Free rider”的存在是 P2P 流媒体设计的关键,必须保证参与节点都贡献资源,一个 P2P 系统才能正常工作。本文将从这三个方面阐述 P2P 流媒体的研究现状,分析现存的问题,并指出今后的研究方向。

1 应用层多播树

在 P2P 流媒体中,首要问题是将服务器和参与服务的节点组织成 ALM 树。但如何构建有效的 ALM 树仍是个难点。

1.1 单树结构的 ALM 方案

1.1.1 小规模的多源 ALM 方案

该方案多应用于视频会议,例如 ESM^[2]系统。将用户节点组成一个应用层 mesh,周期性检查 mesh 中的连接质量, mesh 上以数据源为根,根据带宽、时延各自构造生成树。可以针对每个源单独优化,每个成员维护一个组成员列表,可靠性高,但开销大,扩展性差。

1.1.2 大规模的单源组 ALM 方案

例如 NICE^[3], Zigzag^[4] ALM 树构建方案。两者的思路都是“分层”(Hierarchical)、“分群”(Cluster),成员只和少量固

收稿日期:2004-11-19;修订日期:2005-02-28

作者简介:王艳丽(1979-),女,宁夏石嘴山人,硕士研究生,主要研究方向:网络通信与信息系统;鲜继清(1946-),男,四川西充人,副教授,主要研究方向:网络通信与信息系统;白洁(1979-),男(回族),宁夏石嘴山人,硕士研究生,主要研究方向:网络多媒体。

定数目的节点联系。NICE(如图1)的维护管理具有分布性和自治性,节点的维护负载较轻,且节点的退出只影响局部节点,不影响根节点。缺点是层次越高的节点负载越重,如最高层的节点的度数达到 $O(\log N)$,当系统规模很大时,这会成为系统的瓶颈。

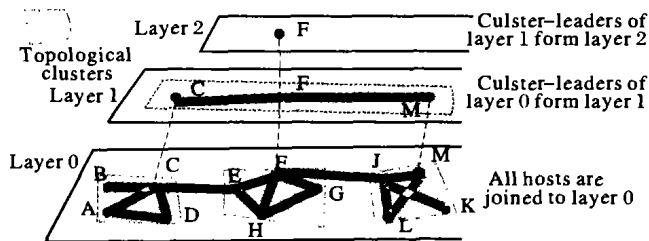


图1 NICE

ZigZag(如图2)与NICE相似,两者在每个节点的平均维护负载都为 $O(k)$,树的高度都为 $O(\log N)$ 。但ZigZag解决了NICE存在的瓶颈问题。其改进点为:Zigzag中Cluster的管理和数据分发由不同节点完成,而NICE将两功能统一在一个节点上。改进后,多播树中节点所带子节点数目最多为 $O(k^2)$,与参与多播树的节点数目无关。

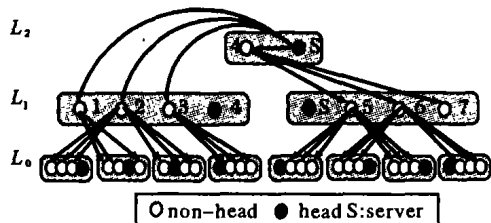


图2 Zigzag

1.2 多树结构的 ALM 方案

单个组播树的解决方案主要缺陷是:组播树中,转发能力较小的内部节点可能需承担很大转发负载,而叶节点却无需承担负载。P2P 流媒体是一种高带宽的应用,单个组播树限制了它的规模。基于多树的方案,以 CoopNet^[5]、SplitStream^[6]为代表,在数据的传输路径上引入了冗余,视频流不再由一个组播树来完成转发,而是先按设定的编码方式被分割成视频段,然后不同视频段由不同生成树完成转发。

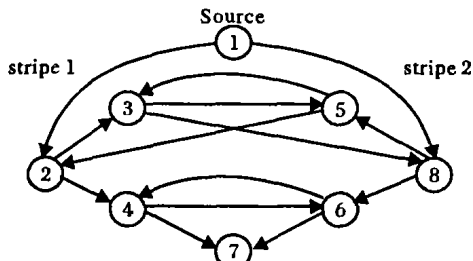


图3 SplitStream

SplitStream的设计目标是利用多树结构平衡负载,SplitStream将视频流分割成多个Stripes,每个Stripe构建自己的组播树,这样可以做到一个节点仅在一个组播树中作为转发节点,在其余组播树中作为叶节点,当一个节点出现死机或网络出现故障时,只影响一小部分节点。图3所显示的是一个SplitStream组播树的例子,源节点1产生两个stripes。对于节点2的stripe1,组播树先将它发送到节点3,4,再发送到节点5,6,7和8。对于节点8的stripe2,组播树先将它发送到节点5和6再发送到节点2,3,4和7。假设当节点3出现问题时,仅影响到节点1,5,8。故通过多个组播树结构传送,可以使每个节点都参与转发视频流,降低了对单个节点带宽要求,减少因数据包丢失对系统的影响,增强了鲁棒性。

SplitStream基于Scribe^[7]协议,Scribe协议的主要作用是把P2P网络中的邻近节点连接在一起,减少了应用层组播的时延,但是每个参与组播的节点必须转发其不感兴趣的视频流。

CoopNet类似于SplitStream,它利用MDC(多描述编码),将视频流编码成多个独立的Descriptions,再针对每个Description构建ALM树。与SplitStream的区别在于:CoopNet基于客户/服务器模型,当服务器未超过负载时,服务器直接向请求客户提供服务,当服务器超过负载时,服务器指定已享受服务的客户进行服务。

基于多树结构组播通过采用MDC编码和多树结构在传输数据和传送路径中引入冗余,在所有参与的节点中平衡了负载,且减少了某个节点失败及数据包丢失对整个系统的影响。但其缺点也很明显:采用多树结构,需同时维护多个组播树,造成开销过大,还需保证多路传送时的数据同步,且MDC编码效率低。

2 网络异构性

Internet的异构性是制约P2P流媒体发展的一个重要因素,一种好的解决方案应支持大范围的流媒体发布,应最大化系统的吞吐量且公平分发视频数据到每个接收者。目前的P2P流媒体技术对网络异构性的解决方案主要有以下几种。

2.1 复制流

发送源通过不同的组播地址为同一视频服务请求开辟多个不同质量的组播流,接收者选择一个最能与自身接收能力相匹配的组播流,同时每个流可根据相应接收者的反馈,在一定范围内调整流速率。文献[8]中将视频源分别编码成高质量视频(300Kbps)、低质量视频(100Kbps)和声音(20Kbps)来满足不同带宽节点的需求,如图4所示,在节点接收过程中,通过丢包率动态地控制视频流速率。该方法的缺点是接收者能力会随网络状况的变化而变化,会在不同流之间跳跃。此外,复制流技术所产生的冗余信息增加了网络负担。

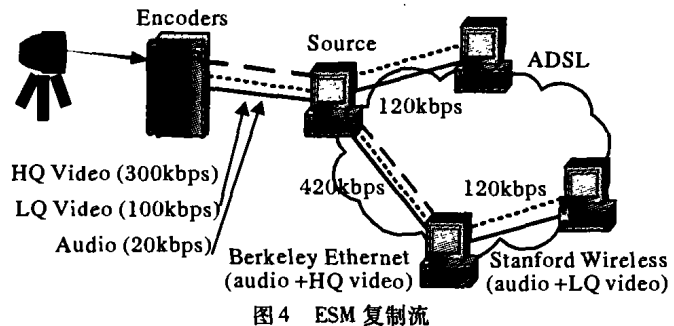


图4 ESM复制流

2.2 分层编码

此方法将视频信息分成一个基本层和若干个增强层,并分别在不同的组播地址空间组播。基本层提供最基本的服务质量,可独立完成解码,增强层提高了服务质量。文献[9]提出基于分层编码和multiple-sender方式传递分层视频信息,针对多个发送端向同一接收端传递数据,提出多个发送端之间进行带宽和数据分配的方案,以解决网络的异构性。

文献[9]假设:

1) $\{H_0, H_1, H_2, \dots, H_N\}$ 为参与应用层组播的节点, H_0 代表服务器,对于任一节点 H_k ,假设 H_k 的输入/输出带宽为 I_k/O_k ;

2) 假设视频流分层为 $\{l_0, l_1, l_2, \dots, l_L\}$, l_0 为基层,其他为增强层,且每层所具有的速率分别为 $\{r_0, r_1, r_2, \dots, r_L\}$;

3) A_k 表示节点 H_k 缓存内可用层数量;

4) Q_k 表示用户 H_k 接受视频流的层数量, Q_k^m 表示节点 H_k 从 h_m 接收视频流的层数量。

并给出计算节点 H_k 接收能力的算法:

```

UnconstrainedSenders ( $I_k, H_1, \dots, H_M$ )
 $Q_k \leftarrow 0$ 
 $P \leftarrow \phi$ 
 $m = 1$ 
repeat
 $Q_k^m \leftarrow \max(n \mid \sum_{i=Q_k^m}^{Q_k+n} r_i \leq Q_m, \sum_{i=0}^{Q_k+n} r_i \leq I_k)$ 
enque( $P, H_m$ )
 $Q_k \leftarrow Q_k + Q_k^m$ 
 $m \leftarrow m + 1$ 
until  $\sum_{i=0}^{Q_k} r_i > I_k$  or  $m > M$ 

```

该算法输入参数为:

1) I_k : 节点 H_k 的输入带宽; 2) $S = \{H_0, H_1, H_2, \dots, H_M\}$: 可为节点 H_k 提供服务的节点集合, 其中节点按 $A_1 \leq A_2 \leq \dots \leq A_M$ 排序。算法中 P 指存放可提供服务节点的集合。用户节点的层分配策略如下:

```

index  $\leftarrow 0$ 
repeat
 $H_m \leftarrow \text{deque}(P)$ 
allocate layers  $\sum_{i=\text{index}+Q_k^m-1}^{\text{index}+Q_k^m-1} l_k$  to  $H_m$ 
 $O_m \leftarrow O_m - O_k^m$ 
index  $\leftarrow \text{index} + Q_k^m$ 
until  $P = \phi$ 

```

该算法通过分层编码且在每层中采用不同的速率解决网络异构性; 并在满足异构性的同时, 追求节点带宽利用率最大化以降低服务器带宽消耗。但文献[9]中层解码必须是累积的, 也就是层 K 要解码必须同时有前面 1 至 $K-1$ 层。接收者至少要加入基本层, 并根据自身条件尽可能同时接收多个增强层。在文献[5]、[6]中采用了非累积层编码 MDC, 增加了灵活性, 消除了各层之间的优先级, 但非累积层编码效率低, 维护开销大。

3 激励机制

2000 年, 对 P2P 文件共享系统 Gnutella^[11] 24 小时的运行活动跟踪记录, 发现系统 70% 的节点未共享本身资源, 1% 的节点承担了系统 47% 的下载负载, 25% 的节点承担了系统 99% 的下载负载^[12]。激励机制是保证 P2P 流媒体系统正常运行的关键, 但它也是目前 P2P 流媒体未能很好解决的问题, 对于 P2P 流媒体激励机制的经验多数来自 P2P 文件下载系统。目前, P2P 流媒体的激励机制主要有两类。

3.1 物品交换(Barter)/征税(taxation)

Barter 是一种 Bit-for-Bit 的激励机制, 时下流行的 Bittorrent 应用了这种激励机制。Bit-for-Bit 指把共享数据划分成固定大小的数据块, 每个节点以自己拥有的数据块作为与别人交换的筹码, 节点下载数据块的同时必须提供上传。这种物品交换的方法消除了系统的“Free rider”。Barter 虽简单易实现, 但它严格地限制了资源薄弱节点的接收带宽, 故并不适用于 P2P 流媒体系统。

文献[13]的线性征税激励机制是在 Barter 基础上提出的, 它在 P2P 流媒体系统的激励机制中引入税收的概念, 综合考虑网络中节点的带宽差异, 提出资源充裕的节点共享富余的资源, 并帮助资源匮乏的节点。为了使所有节点享受资源达到最大化, 并实现资源的再分配, 发布者拟订、发布并强

制实施“征税”方案, 参与系统的所有节点必须履行此方案, 促进系统资源的再分配。方案在接收和上传之间提供一种直接映射:

$$f = \max[t * (r - G), 0]$$

其中 f 为上传带宽, r 为接收带宽, t 为边际税率 ($t \geq 1$), G 为补贴 (demogrant)。 G 和 t 由发布者设置, 节点负责选择 f 和 r , 每个节点接收到的资源至少为 G 。为保证 P2P 流媒体系统的正常运行, 必须保证共享资源大于等于接收资源, 即 $\sum f_i \geq \sum r_i$ 。当 $t = 1$ 且 $G = 0$ 时, 就回到了“Bit-for-Bit”的激励机制。如果节点为“Free rider” ($f = 0$), 它只能接收到系统给予的补贴 ($r = G$), 这样就杜绝了“Free rider”节点只享受服务不提供服务的行为。

3.2 互惠机制(Reciprocity)

互惠机制有直接和间接之分。前者以循环机制为代表, 后者以信誉机制为代表。

P2P 流媒体系统中节点角色的不对称性决定了直接互惠机制的不适用。

间接互惠机制鼓励节点合作以获取信誉值, 信誉高的节点将得到优先对待。文献[14]提出了基于积分的激励机制, 它建立在如下观点之上:

- 1) 节点间增进合作促进系统全部节点平均服务质量的提高;
- 2) 当一个节点上传过多时, cost (即节点耗费成本) 将影响其自身和远端接收节点的服务质量;
- 3) 服务质量很大程度上依赖于选择什么样的节点提供服务。

基于此, 该激励机制通过间接互惠^[10] 鼓励合作。选择节点质量的差异性体现了积分的不同: 贡献大的节点允许它有灵活的节点选择权, 相应地能得到高质量的视频服务, 而对于“Free rider”则限制其节点选择范围, 其只能收到低质量的视频 (如图 5)。

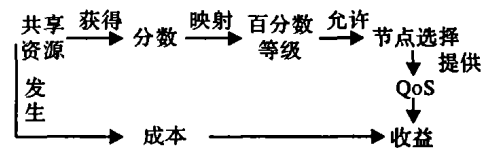


图5 间接互惠机制

假设节点贡献资源 x_i , 可获得积分 S_i , 并映射为相应的等级 R_i 。规定节点只能选择等于或低于本身等级的节点作为服务供应节点, 故等级不同的节点可选择供应节点的范围不同。节点等级越高, 可供选择的供应节点可靠性越高, 服务质量 Q_i 越有保障。 Q_i 使节点获得收益 U_i , 节点 i 上传资源耗费的成本 C_i 又使节点收益 U_i 降低。故 U_i 、 Q_i 、 C_i 之间的关系为:

$$U_i(x_i) = a_i Q_i(x_i) - b_i C(x_i)$$

a_i 、 b_i 分别定义了节点 i 的视频流质量和共享成本值。上传字节数 $S = B_{\text{out}}$, 上传资源所耗费节点的成本考虑了存储所需成本及带宽因素:

$$C_i = c_L L_i M + c_T T_i D_i$$

其中, c_L 是单位存储成本, L_i 是节点上传资源块数量。 M 是资源块平均大小, c_T 是单位传输成本, T_i 是 I 的传送速率, D_i 是 i 的传输时长。积分 S 的累积分布函数为:

$$F(S) = \sum_{i=S_{\text{low}}}^{S_{\text{high}}} f(i)$$

这里 f 为积分 S 的概率密度函数, 积分的评定依赖于累积

分布函数 $F(S)$ 及参与系统的节点数量。最后,质量函数 Q_i 需考虑的特性为:

- 1) 节点积分递增时,质量函数非递减;
- 2) 质量无限接近 Q_{MAX} , Q_{MAX} 为节点接收视频质量由 N_s 个最好的供应节点提供;
- 3) 质量函数的初始值为非负数 Q_{BE} , 其中 $Q_{BE} = Q(S_i = 0) \geq 0$, Q_{BE} 为节点接收视频质量由 0 个最好的供应节点提供。

基于这 3 个特性,质量函数被定义为:

$$Q_i(R_i) = \begin{cases} Q_{MAX}, & R_i \geq \frac{N_s}{G_i} \\ \frac{R_i G_i}{N_i} Q_{MAX} + \left(1 - \frac{R_i G_i}{N_i}\right) Q_{BE}, & \text{otherwise} \end{cases}$$

其中, R_i 为节点 i 的等级, G 为可供选择的高质量服务节点个数, $R_i G$ 为高质量服务供应节点数。

需指出,线性征税激励机制的前提是共享和接收同步,故适用于 P2P 流媒体直播系统;互惠激励机制不考虑共享和接收是否同步的因素,因此主要应用于 P2P 视频点播系统。

4 结语

基于 P2P 的流媒体服务体系目前还处于研究的初步阶段,还有许多问题需要解决。

1) 多播树协议的问题。迄今为止,多数系统仅用时延或负载平衡等一两种属性作为度量值,并未综合考虑节点可靠性、网络带宽、网络时延和数据包丢失率及整个算法的维护开销。还有一个被忽略的重要属性是:网络中存在大量的防火墙和 NAT。目前多数计算机与 Internet 连接存在防火墙,不能直接连接,故还需深入研究。

2) 如何有效地利用下层网络资源的问题。由于应用层网络叠加于 IP 网络之上,这就需要一个测量网络实时状况的有效方法,利用测量到的信息构造适应网络变换的 ALM 树。目前对网络测量已有一些解决方案,但如何将其集成到 P2P 流媒体系统中仍未得到解决。

3) 应用层组播的安全问题。应用层组播的安全性研究包括组播机制和信息传递两个方面。组播是一个开放环境,为防止恶意用户的侵入和参与,必须提供针对发送者和组成员的身份鉴别技术。安全性是流媒体组播得以大规模应用的前提。

4) 数据放置策略。P2P 流媒体将资源存储由服务器推向了拥有强大缓存的终端用户。但由于终端用户在处理能

力、存储和带宽上的差异性,采用什么策略在不同的终端放置数据,以达到加快系统响应速度、增强系统的可靠性和扩展性是值得进一步研究的。

5) 对无线用户的支持。无线局域网技术的蓬勃发展在为移动用户带来高带宽的同时,也使对移动用户的流媒体服务成为可能。但因无线网络高度动态性和无线信道的误码率等因素的限制,如何在这个特殊的网络中构造 ALM 树,如何进行拥塞控制和差错控制还有待进一步研究。

参考文献:

- [1] <http://bitconjurer.org/BitTorrent> [EB/OL], 2001-04.
- [2] CHU Y-H, RAO SG, ZHANG H. A case for end system multicast [A]. ACM SIGMETRICS [C], 2000.
- [3] BANERJEE S, BHATTACHARJEE B, KOMMAREDDY C. Scalable application layer multicast [A]. ACM SIGCOMM [C], 2002.
- [4] TRAN DA, HUA KA, DO TT. Zigzag: An efficient peer-to-peer scheme for media streaming [A]. Submitted to IEEE INFOCOM [C], 2003.
- [5] CHOU PA, PADMANABHAN VN, WANG HJ. Resilient peer-to-peer streaming [R]. Microsoft Research, 2003.
- [6] CASTRO M, DRUSCHEL P, KERMARREC A, et al. Splitstream: High-bandwidth multicast in cooperative environments [A]. 19th ACM Symposium on Operating Systems Principles [C], 2003.
- [7] CASTRO M, DRUSCHEL P, KERMARREC A-M, et al. SCRIBE: A Large-scale and Decentralized Application-level Multicast Infrastructure [A]. IEEE JSAC [C], 2002, 20(8): 100-110.
- [8] EUGENE NG TS, RAO SG, SRIPANIDKULCHAI K, et al. Early Experience with an Internet Broadcast System Based on Overlay Multicast [R]. USENIX Annual Technical Conference, 2004.
- [9] CUI Y, NAHRSTEDT K. Layered peer-to-peer streaming [A]. NOSSDAV [C], 2003.
- [10] NOWAK M, SIGMUND K. Evolution of indirect reciprocity by image scoring [J]. Nature, 1998, 393: 573-577.
- [11] <http://Gnutella.wego.com> [EB/OL], 2005.
- [12] ADAR E, HUBERMAN BA. Free riding on Gnutella [J]. First Monday, 2000, 5(10).
- [13] CHU Y-H, CHUANG J, ZHANG H. A Case for Taxation in Peer-to-Peer Streaming Broadcast [A]. ACM SIGCOMM04 Workshop on Practice and Theory of Incentives in Networked Systems (PINS) [C], 2004.
- [14] HABIB A, CHUANG J. Incentive Mechanism for Peer-to-Peer Media Streaming [A]. 12th IEEE International Workshop on Quality of Service [C], 2004.

(上接第 1263 页)

- [4] LIN JS, LIU M, HUANG NF. The shortest path computation in MOSPF protocol using an annealed Hopfield neural network with a new cooling schedule [J]. Information Sciences, 2000, (129): 17-30.
- [5] ZHANG S, LIU Z. A new dynamic routing algorithm based on chaotic neural networks [J]. Journal of china institution of communication, 2001, (12): 1-7.
- [6] XIANG F, JUNZHOU L, JIEYI W, et al. QoS routing based on genetic algorithm. Computer Communications [J], 1999, (22): 1394-1399.
- [7] WANG Z, SHI B. Bandwidth-delay-constrained least-cost multicast routing based on heuristic genetic algorithm [J]. Computer Commu-

nications, 2001, (24): 685-692.

- [8] HAGHIGHATAB AT, FAEZB K, DEGHAN M, et al. GA-Based heuristic algorithms for QoS based multicast routing [J]. Knowledge-Based Systems, 2003, (16): 305-312.
- [9] RAVIKUMAR CP, BAJPAI R. Source-based delay-bounded multicasting in multimedia networks [J]. Computer Communications, 1998, (21): 126-132.
- [10] SALAMA HF, REEVES DS, VINIOTIS Y. Evaluation of multicast routing algorithms for real-time communication on high-speed networks [J]. IEEE JSAC, 1997, (15): 332-345.