

# 大型视频组播系统整体效用层次化建模与求解

王 晖 沙基昌 姜志宏 张 鑫

(国防科技大学信息系统与管理学院 长沙 410073)

**摘 要** 为了有效地分析和评价大型视频组播系统的自适应方法与技术,文中首先抽象描述了大型视频组播系统整体效用最大化问题,针对该问题集中式求解存在的单点失效、反馈内爆和自适应控制延迟大等缺陷,该文提出了基于瓶颈链路的层次化分解方法,建立了大型视频组播系统整体效用层次化模型,给出了一个基于层次化代理的分布优化求解方法 OptUtility 算法,可获得大型分层视频组播系统中通过各个层次代理的各个会话的优化带宽、优化的分层数量和各层次优化的发送速率. 通过对该算法的计算复杂度的分析与对比实验,证明该算法对于解决大型视频组播系统整体效用优化问题具有较低的计算时间复杂度,容易实现,且实时性能好.

**关键词** 大型视频组播;整体效用;层次化代理;分布求解;性能优化

**中图法分类号** TP391

## Hierarchical Total Utility Modeling and Solution for Large-Scale Video Multicast System

WANG Hui SHA Ji-Chang JIANG Zhi-Hong ZHANG Xin

(College of Information System and Management, National University of Defense Technology, Changsha 410073)

**Abstract** To effectively analyze and evaluate the adaptive methods for large-scale video multicast system, this paper first abstractly describe the maximum problem of the total utility for large-scale video multicast system. Since the centralized solution for the maximum problem has disadvantages of single-node failure, feedback implosion, and bad real-time performance, a hierarchical decomposition method based on bottleneck link is presented in this paper, and the hierarchical total utility model for large-scale video multicast is also given. Then, OptUtility, an extendable, hierarchical proxy-based distributed solution, is proposed, by which we can figure out optimally the bandwidth allocation of each session, the layering number, and the rate of each layer in all hierarchical proxies. Analysis of computational complexity and experimental results prove the proposed distributed method to be of low complexity, easy in implementation, and well performed in real time.

**Keywords** large-scale video multicast; total utility; hierarchical proxy; distributed solution; optimal performance

## 1 引 言

自适应视频组播技术是目前基于 IP 的分布视

频会议、网络实时教学以及视频流媒体应用的主流支撑技术. 随着视频组播系统规模的逐步扩大,受接收者状态信息反馈到发送端延迟的影响,传统的基于接收者或基于发送者的分层视频组播端点自适应

方法难以快速适应组播成员和网络拓扑的动态变化和扩展,无法满足大型视频组播系统的服务质量要求.因此,人们提出了多种基于代理的应用自适应方法,比较有代表性的是 AS1<sup>[1]</sup>和 LVMR<sup>[2]</sup>.AS1 是一种基于代理的主动服务框架,它引入了 RTP 媒体网关的概念,利用转换编码技术实现了多种视频源编码算法在大型异构组播网络中的速率自适应控制问题. AS1 从分布系统服务的角度研究了媒体网关的定位和发现问题,但对媒体网关如何优化选择与定位缺乏具体的理论分析与评价. LVMR 是一种层次化的视频组播速率控制方法,它结合了基于代理的集中控制方法与端点分布控制方法的优点以解决大型视频组播网络拥塞和异构的自适应问题,但 LVMR 没有就代理的拓扑位置优化选择问题进行研究. 要对基于代理的应用自适应方法进行分析与评价,就必须从整体效用的角度探讨视频组播系统的自适应方法与协议的优化设计问题<sup>[3]</sup>. 在视频组播系统整体效用建模与优化方面,文献[4]针对单一的分层视频组播会话中异构性能接收者的分组优化问题,通过引入接收者效用函数,将会话内接收者分组优化问题形成一个最大化接收者效用之和的动态规划问题,并提出了一个最优分组算法 Optimal-Partition. 但这一方法的局限性在于只考虑了小型视频组播系统中会话内接收者的优化分组问题. 为了解决具有多个并发分层视频组播会话的动态分层与带宽分配的优化问题,文献[5]针对无线视频组播应用环境,研究了基于单一瓶颈链路的小型视频组播系统整体效用模型,并提出了一个 OLBA 算法 (Optimal Layering and Bandwidth Allocation),将在共享带宽约束下多个视频组播会话组成的系统整体效用最大化问题分解为会话间带宽最优分配问题和会话内动态最优分层及层次速率最优分配问题. 但 OLBA 算法没有对大型视频组播系统整体效用建模和优化问题进行深入研究.

本文归纳了具有并发的多个视频组播会话的大型视频组播系统整体效用最大化问题. 由于整体效用最大化问题是带约束的非线性规划问题,集中式的递推求解方法要求所有接收者性能等信息都发送给集中决策者,容易造成“反馈内爆”、单点失败、集中计算量较大、无法满足实时性要求等问题,本文提出了基于瓶颈链路的层次化分解方法,建立了大型视频组播系统整体效用层次化模型,给出了一个基于层次化代理的分布优化求解方法 OptUtility 算法,并对该算法的计算复杂度进行了分析与评估.

## 2 大型视频组播系统整体效用最大化问题

### 2.1 大型视频组播系统抽象描述

针对大型视频组播系统的共性,本文作如下抽象描述:

(1) 在某个自适应周期,大型视频组播系统可表示为一个四元组  $(S, V, R, U)$ . 其中,  $S$  表示网络中并发的视频组播会话集合,  $V$  表示视频组播系统中存在的瓶颈链路的集合,  $R$  表示加入视频组播会话的所有接收者集合,  $U$  表示该视频组播应用系统的整体效用.

(2) 假设大型视频组播系统是由  $M$  个视频组播会话组成的,一个视频组播会话  $S_m$  可用一个四元组  $(s_m, V_m, LR_m, U_m)$  来表示,其中,  $s_m$  表示视频组播会话发送源,  $V_m$  表示该视频组播会话经过的瓶颈链路集合,  $LR_m$  表示该视频组播会话的接收者集合,  $U_m$  表示该视频组播会话的效用.

(3) 假设以大型累积分层视频组播系统为研究对象,视频组播会话  $S_m$  的发送源  $s_m$  将发送视频流分成  $K_m$  个组. 若发送源第  $k$  个组以 RTP 会话  $G_{m,k}$  进行发送,则  $G_{m,k}$  可以用四元组  $(g_{m,k}, V_{m,k}, LR_{m,k}, U_{m,k})$  表示,其中,  $g_{m,k}$  表示加入视频组播会话  $S_m$  中第  $k$  个分组的累积发送速率,  $LR_{m,k}$  表示加入第  $m$  个视频组播会话最高为第  $k$  个分组 RTP 会话的接收者集合,  $U_{m,k}$  表示加入视频组播会话  $S_m$  中第  $k$  个分组的累积效用.

(4) 假设  $L$  表示大型视频组播系统在某个自适应周期所共享的瓶颈链路数量,  $c_l$  表示第  $l$  个瓶颈链路的带宽,并假设所有带宽都以信道为单位,带宽大小取有限的离散值,  $LS_l$  表示经过第  $l$  个瓶颈链路的所有视频组播会话中所有分组 RTP 会话的集合.

(5) 假设一个接收者同时只能加入一个视频组播会话,第  $n$  个接收者  $R_n$  的接收性能  $r_n$  用 TCP 友好的吞吐量模型<sup>[9]</sup>进行测算.

(6) 用  $u(r', r) = \frac{\text{opt}Q(r')}{\text{opt}Q(r)}$  表示接收者的相对效

用函数,其中  $r'$  表示接收者实际接收的速率,  $r$  表示接收者期望接收的速率,即接收者的接收性能,函数  $\text{opt}Q(r)$  表示当接收速率为  $r$  时,经率失真优化的视频接收质量.

### 2.2 大型视频组播系统整体效用最大化问题

根据以上抽象描述,可以完成以下的推导:

### (1) RTP 会话组播组效用

RTP 会话组播组效用是指加入该 RTP 会话的所有接收者的效用之和,初始化  $U_{m,k} = 0, 1 \leq m \leq M, 1 \leq k \leq K_m$ , 则  $U_{m,k}$  的计算公式如式(1)所示:

$$U_{m,k} = \sum_{R_n \in LR_{m,k}} u(g_{m,k}, r_n) \quad (1)$$

式(1)必须满足约束条件:  $g_{m,k} \in (0, r_n]$ .

### (2) 视频组播会话效用

一个视频组播会话  $S_m$  的效用表示如式(2)所示:

$$U_m = \sum_{k=1}^{K_m} U_{m,k} \quad (2)$$

### (3) 视频组播系统整体效用

视频组播应用系统整体效用表示如式(3)所示:

$$U = \sum_{m=1}^M U_m \quad (3)$$

### (4) 视频组播应用系统整体效用最大化问题

根据以上推导,可以将视频组播系统整体效用最大化问题描述如式(4)所示:

$$P1. \quad \max \sum_{m=1}^M U_m \quad (4)$$

P1 问题必须满足以下约束条件:

$$\sum_{m=1}^M \max_{G_{m,k} \in LS_l} g_{m,k} \leq c_l, \quad 1 \leq l \leq L \quad (5)$$

约束条件保证整个系统不处于拥塞状态,在得到系统整体最大效用的同时,可求得相应的每个视频组播会话的最优分配带宽、每个会话中最优的分组数以及每个分组的最优发送速率. 对于问题 P1, 本文假设以下条件成立:

① 假设条件 1: 存在  $\bar{k}$ , 使得对于所有  $l \in L$ , 可满足  $\sum_{m=1}^M \max_{G_{m,k} \in LS_l} g_{m,k} \leq c_l$ .

② 假设条件 2: 接收者效用函数是严格凹函数, 即对于任意  $g_{m,k}$ , 接收者效用函数  $u(g_{m,k}, r_n)$  是随  $g_{m,k}$  递增, 两次连续可导且严格凹的.

以上假设条件 1 意味着 P1 问题有可行解, 假设条件 2 意味着 P1 问题的解是唯一的.

由于接收者效用函数是非线性的, 约束条件式(5)中取最大值函数也是非线性, 因此, P1 问题是一个带约束条件的非线性规划问题. 要求解最大化问题 P1, 所有接收者的状态信息和瓶颈链路信息都要聚集到一个集中计算节点, 由它完成 P1 的求解, 并把计算结果发送到各个视频组播会话的发送源和各个接收者. 由于 P1 是非线性的, 包含多个变量, 计算复杂度很大. 而且, 将所有的接收者状态信息和瓶颈链路信息聚集到一个集中节点, 一是要选择一

个能完成这项复杂计算的网络节点或专门在网络上设置一个计算节点, 这将加大系统自适应机制的复杂性和容易导致“单点失败”; 二是要解决“反馈内爆”问题, 三是随着组播规模的扩大, 从接收者状态和网络状态发生变化到集中计算决策后采取自适应控制, 整个延迟必然增大, 达不到实时视频组播应用的要求. 因此, 需对 P1 问题寻求一个分布且可扩展的求解方法.

## 3 大型分层视频组播系统整体效用层次化建模

通过研究, 可以发现视频组播会话的所有接收者的状态虽然异构, 但是也呈现一定的聚类特性, 这些特性可以概括为包丢失相关性和速率分布相关性.

### (1) 包丢失相关性

由于存在瓶颈链路, 多个视频组播会话通过瓶颈链路时容易造成丢包, 一旦因拥塞造成丢包, 则瓶颈链路的下游的任何接收者也将出现丢包现象, 这称为包丢失相关性.

### (2) 速率分布相关性

由于视频组播会话的接收者都采用标准的网络设备和接口, 如以太网接口、无线局域网接口以及 ADSL 接口等. 因此, 接收者的接收速率呈现出聚类的分布.

### 3.1 大型分层视频组播系统的层次化分解与描述

根据以上视频组播会话接收者的包丢失相关性, 可以定位瓶颈链路和测算瓶颈带宽, 为不改变路由器的结构和路由算法, 在每个瓶颈链路上流最靠近该链路的接收者中选择一个接收者作为通过该瓶颈链路的视频组播会话的代理, 由它负责收集其下游接收者的状态信息, 并根据其通过的一个瓶颈链路, 利用速率分布相关性, 计算并决策出它下游局部范围内最优的会话带宽分配和各会话内各分组最优的速率. 每个瓶颈链路的代理对下相当于新的组播会话发送者, 对上更是更上层次的组播会话的接收者. 由此, 构成视频组播应用系统的层次化结构. 针对大型视频组播系统的层次化分解, 本文作如下抽象描述:

(1) 在某个自适应周期, 假设大型视频组播系统从组播树根结点往下根据经过的瓶颈链路将系统分解为第 1, 第 2, ..., 第  $H$  层, 假设其中第  $h$  层包含  $A_h$  个瓶颈链路,  $M_{h,j}$  表示通过第  $h$  层第  $j$  个瓶颈链

路的视频组播会话数量,  $S^{h,j}$  表示通过第  $h$  层第  $j$  个瓶颈链路的视频组播会话集合。

(2) 假设  $S_m^{h,j}$  表示通过第  $h$  层第  $j$  个瓶颈链路的第  $m$  个视频组播会话,  $S_m^{h,j}$  分成  $K_m^{h,j}$  个分组进行分层组播,  $G_{m,k}^{h,j}$  表示视频组播会话  $S_m^{h,j}$  中第  $k$  个分组的 RTP 会话,  $g_{m,k}^{h,j}$  表示  $S_m^{h,j}$  中第  $k$  个分组的累积发送速率。

(3) 假设  $c_{h,j}$  表示第  $h$  层第  $j$  个瓶颈链路的带宽,  $LR_{m,k}^{h,j}$  表示加入通过第  $h$  层第  $j$  个瓶颈链路的第  $m$  个视频组播会话最高为第  $k$  个分组 RTP 会话的接收者集合。

(4) 假设  $U_{m,k}^{h,j}$  表示通过第  $h$  层第  $j$  个瓶颈链路的第  $m$  个视频组播会话第  $k$  个分组的累积效用,  $U_m^{h,j}$  表示通过第  $h$  层第  $j$  个瓶颈链路的第  $m$  个视频组播会话的效用,  $U_{h,j}$  表示通过第  $h$  层第  $j$  个瓶颈链路的新视频组播系统的效用,  $U_h$  表示第  $h$  层的效用。

除以上假设外, 其它变量沿用 2.1 节中的描述。

### 3.2 大型视频组播系统整体效用的层次化建模

根据以上基于代理的层次化结构描述, 本文对大型视频组播系统的整体效用进行了层次化建模, 具体方法如下:

(1) 通过第  $h$  层第  $j$  个瓶颈链路的第  $m$  个视频组播会话第  $k$  个分组 RTP 会话组播效用计算如式(6)所示:

$$U_{m,k}^{h,j} = \sum_{R_n \in LR_{m,k}^{h,j}} u(g_{m,k}^{h,j}, r_n) \quad (6)$$

(2) 通过第  $h$  层第  $j$  个瓶颈链路的第  $m$  个视频组播会话的效用计算如式(7)所示:

$$U_m^{h,j} = \sum_{k=1}^{K_m^{h,j}} U_{m,k}^{h,j} \quad (7)$$

(3) 通过第  $h$  层第  $j$  个瓶颈链路的新视频组播系统整体效用计算如式(8)所示:

$$U_{h,j} = \sum_{m=1}^{M_{h,j}} U_m^{h,j} \quad (8)$$

约束条件:  $\sum_{m=1}^{M_{h,j}} G_{m,k}^{h,j} \in LR_{m,k}^{h,j} \max g_{m,k}^{h,j} \leq c_{h,j}$ 。

(4) 第  $h$  层的整体效用计算如式(9)所示:

$$U_h = \sum_{j=1}^{A_h} U_{h,j} = \sum_{j=1}^{A_h} \sum_{m=1}^{M_{h,j}} U_m^{h,j} \quad (9)$$

约束条件:  $\sum_{m=1}^{M_{h,j}} G_{m,k}^{h,j} \in LR_{m,k}^{h,j} \max g_{m,k}^{h,j} \leq c_{h,j}, \forall j, 1 \leq j \leq A_h$ 。

(5) 层次化结构视频组播系统整体效用计算如式(10)所示:

$$U = \sum_{h=1}^H U_h \quad (10)$$

(6) 层次化结构视频组播系统整体效用最大化问题可以表示如下:

$$P2. \quad \max \left( \sum_{h=1}^H U_h \right) \quad (11)$$

约束条件:

$$\sum_{m=1}^{M_{h,j}} G_{m,k}^{h,j} \in LR_{m,k}^{h,j} \max g_{m,k}^{h,j} \leq c_{h,j}, \quad \forall h, 1 \leq h \leq H, \quad \forall j, 1 \leq j \leq A_h \quad (12)$$

$$g_{m,K_m^{h,j}}^{h,j} \leq g_{m,K_m^{h-1}, \pi_m(j)}^{h-1, \pi_m(j)}, \quad \forall h, 1 \leq h \leq H,$$

$$\forall j, 1 \leq j \leq A_h, \quad \forall m, 1 \leq m \leq M_{h,j}, \quad \pi_m(j) \neq \emptyset \quad (13)$$

其中,  $\pi_m(j)$  表示通过第  $h$  层第  $j$  个瓶颈链路的第  $m$  个会话所对应的第  $h-1$  层中所通过的瓶颈链路。

P2 问题是在对原始的 P1 问题进行层次化分解后形成的最大化问题, 是对原 P1 问题中目标函数的接收者效用进行分组求和, 并求最优解。约束条件式(12)表示采用累积分层组播方法, 经过每个层次的每个瓶颈链路的视频组播会话的最大分组速率之和不能大于该瓶颈链路的总带宽。约束条件式(13)表示经过每个层次的每个瓶颈链路的视频组播会话的最大分组速率不能大于该会话经过其上层的瓶颈链路的最大分组速率。

## 4 OptUtility: 一个基于层次化代理的分布优化求解方法

以上将原始的 P1 问题转化成 P2 问题, 旨在获得分布且可扩展的求解方法。求解 P2 问题, 即求从第 1 层到第  $H$  层各个层次整体效用之和的最大化问题, 这个问题可进一步表示为如下形式:

$$\begin{aligned} \max \left( \sum_{h=1}^H U_h \right) &= \max \left( \sum_{h=1}^H \sum_{j=1}^{A_h} U_{h,j} \right) \\ &= \max \left( \sum_{h=1}^H \sum_{j=1}^{A_h} \sum_{m=1}^{M_{h,j}} U_m^{h,j} \right) \end{aligned} \quad (14)$$

假设  $U_m^{h,j}$  表示当第  $h$  层第  $j$  个瓶颈链路分配给第  $m$  个视频组播会话的带宽为  $c_m^{h,j}$  时的会话效用, 则式(14)即表示对经过所有瓶颈链路的视频组播会话进行带宽的最优化分配, 使得所有会话的效用之和最大化。这样, 结合式(12)和式(13)的约束条件和式(14), P2 问题可以表示成如下 P3 问题:

$$P3. \quad \max \left( \sum_{h=1}^H \sum_{j=1}^{A_h} \sum_{m=1}^{M_{h,j}} U_m^{h,j}(c_m^{h,j}) \right) \quad (15)$$

约束条件:

$$\sum_{m=1}^{M_{h,j}} c_m^{h,j} \leq c_{h,j}, \quad \forall h, 1 \leq h \leq H, \quad \forall j, 1 \leq j \leq A_h \quad (16)$$

$$c_m^{h,j} \leq c_m^{h-1, \pi_m(j)}, \quad \forall h, 1 \leq h \leq H, \\ \forall j, 1 \leq j \leq A_h, \quad \forall m, 1 \leq m \leq M_{h,j}, \pi_m(j) \neq \emptyset \quad (17)$$

P3 问题对 P2 问题进行简化求解, 即暂时不考虑各个会话内的最优分组问题, 先考虑经过每个瓶颈链路的多个并发视频组播会话间的带宽最优分配问题。

P3 问题是一个典型的动态规划问题, 同时也是带约束的非线性规划问题。若采用集中求解的方法, 可以用递推的方法进行求解, 但这与本文寻求的分布求解方法不相适应。为了得到分布求解方法, 可以利用拉格朗日乘子将 P3 问题表示成一个可分布求解的对偶问题<sup>[6]</sup>, 并用次梯度方法和最优近似解方法对该分布对偶问题进行近似求解。文献[6]采用以上方法对多速率组播会话速率控制的最优化问题进行了研究, 其研究表明: 该方法虽具有理论上的可行性, 但在具体实现时, 各个分布对偶问题所采用的对偶变量的初始值和迭代步长的精确度难以保证, 从而造成最优近似解的求解收敛时间和精度难以保证。考虑到理论上整体效用最优解的证明难度以及实现上的综合可行性, 本文提出了一个基于层次化代理的分布优化求解算法——OptUtility 算法来逐步优化视频组播系统整体效用。

#### 4.1 OptUtility 算法的基本思想

所谓“基于层次化代理的分布优化求解”即由相邻两个层次的代理进行优化求解。以第  $h$  层和第  $h+1$  层为例, 第  $h$  层第  $j$  个瓶颈链路的代理根据当前已知的通过该瓶颈链路的视频组播会话数量  $M_{h,j}$ 、各个会话通过该瓶颈链路前所占用的带宽  $c_m^{h-1, \pi_m(j)}$  ( $1 \leq m \leq M_{h,j}$ )、瓶颈链路的带宽  $c_{h,j}$  以及通过该瓶颈链路的各个会话在第  $h+1$  层中的代理所代理的会话接收者性能等信息进行局部最优化求解。

假设用  $\tilde{U}_m^{h,j}(c_m^{h,j})$  表示当给视频组播会话  $S_m^{h,j}$  分配的带宽为  $c_m^{h,j}$  时, 会话  $S_m^{h,j}$  获得的最大效用, 则通过第  $h$  层第  $j$  个瓶颈链路的新视频组播系统整体效用的最大化问题可重新定义如下:

$$U_{h,j}^* = \max \sum_{m=1}^{M_{h,j}} \tilde{U}_m^{h,j}(c_m^{h,j}) \quad (18)$$

约束条件为  $1 \leq c_m^{h,j} \leq \min\{c_{h,j} - M_{h,j} + 1\}$ ,

$$c_m^{h-1, \pi_m(j)}, \sum_{m=1}^{M_{h,j}} c_m^{h,j} \leq c_{h,j}.$$

$$\text{假设用 } V^*(n, i) = \max_{n=\sum_{k=1}^i n_k} \sum_{k=1}^i \tilde{U}_k^{h,j}(n_k), \quad n=1, 2, \dots, c_{h,j}, i=1, 2, \dots, M_{h,j}$$

表示将带宽  $n$  分配给  $m$  个视频组播会话的最大整体效用, 则式(18)可用以下递推公式计算:

$$U_{h,j} = V^*(c_{h,j}, M_{h,j}) \quad (19)$$

式(19)递推计算过程的核心思想是将有限而离散的瓶颈链路带宽逐个分配给多个视频组播会话, 然后, 各个视频组播会话根据分配的带宽, 在分组数量上限给定的情况下, 求解最优分组数量和各个分组发送速率, 以达到视频组播系统整体效用最大化。其核心算法包括会话间带宽最优分配算法和会话内最优分组算法。

#### 4.2 会话间带宽最优分配算法——

##### BandwidthOptimalAllocation 算法

会话间带宽最优分配算法的主要流程如图 1 描述。设视频组播会话  $S_m^{h,j}$  的接收者中最大性能为  $r_m^{h,j}$ ,  $1 \leq m \leq M_{h,j}$ 。

```

BandwidthOptimalAllocation
( $c_{h,j}, M_{h,j}, \{r_1^{h,j}, r_2^{h,j}, \dots, r_{M_{h,j}}^{h,j}\}$ )
1. for  $k=1$  to  $M_{h,j}$ 
2.    $c_k^{h,j} = 0$ 
3.   for  $n=1$  to  $c_{h,j}$ 
4.      $V^*(n, k) = 0$ 
5.   for  $i=1$  to  $M_{h,j}$ 
6.     if  $i=1$  then
7.       for  $n=1$  to  $\min(c_{h,j}, r_1^{h,j})$ 
8.          $V^*(n, i) = U_i^*(n)$ 
9.          $c_i^{h,j} = n$ 
10.    else
11.      for  $n=i$  to  $\min(c_{h,j}, r_1^{h,j} + \dots + r_i^{h,j})$ 
12.        for  $l=1$  to  $\min(r_i^{h,j}, n-i+1)$ 
13.          if  $V^*(n-l, i-1) < 0$  then
14.             $v = V^*(n-l, i-1) + U_i^*(l)$ 
15.            if  $V^*(n, i) < v$  then
16.               $V^*(n, i) = v$ 
17.               $c_i^{h,j} = c_i^{h,j} + l$ 

```

图 1 会话间带宽最优分配算法主要流程

算法说明如下:

(1) 当  $i=1$  时,  $V^*(n, i)$ ,  $1 \leq n \leq \min(c_{h,j}, r_1^{h,j})$  表示分别分配给第一个视频组播会话 1 个, 2 个, ..., 直到  $c_{h,j}$  个带宽时所对应的会话最大效用。这是以下递推计算的收敛点。

(2) 当  $i>1$  时, 可采用以下递推计算公式来求解:

$$V^*(n, i) = \max_{1 \leq l \leq \min(r_i^{h,j}, n-i+1)} \{V^*(n-l, i-1) + \tilde{U}_i^{h,j}(l)\},$$

其中,  $1 \leq i \leq M_{h,j}$ ,  $1 \leq n \leq \min(c_{h,j}, r_1^{h,j})$ 。这个递推公

式表示的是当得到将  $n-l$  个带宽分配给  $i-1$  个会话的最优方案后,将剩余的  $l$  个带宽分配给第  $i$  个会话的最优方案所对应的最大整体效用。

在会话间带宽最优分配算法中,要求解当分配给会话  $S_m^{h,j}$  瓶颈链路共享带宽为  $c_m^{h,j}$  时,会话内最优分组所达到的最大效用  $\tilde{U}_m^{h,j}(c_m^{h,j})$ 。假设通过第  $h$  层第  $j$  个瓶颈链路的第  $m$  个会话的接收者中存在一个接收者成为该会话通过第  $h+1$  层某个瓶颈链路的代理,其代理的会话当前的最优分组数为  $\hat{K}_{m,j}^{h,j}$ ,且各组的接收者性能和数量可以表示为  $\{(\hat{r}_{m,1}^{h,j}, \hat{d}_{m,1}^{h,j}), (\hat{r}_{m,2}^{h,j}, \hat{d}_{m,2}^{h,j}), \dots, (\hat{r}_{m,\hat{K}_{m,j}^{h,j}}^{h,j}, \hat{d}_{m,\hat{K}_{m,j}^{h,j}}^{h,j})\}$ 。这样,会话  $S_m^{h,j}$  的整体效用  $\tilde{U}_m^{h,j}(c_m^{h,j})$  等于通过第  $h$  层第  $j$  个瓶颈链路的第  $m$  个会话的接收者效用加上该会话在第  $h$  层的代理所代理的经过优化分组的接收者效用之和。由此假设,会话  $S_m^{h,j}$  的最大效用  $\tilde{U}_m^{h,j}(c_m^{h,j})$  可以用以下会话内最优分组算法求解。

#### 4.3 会话内最优分组算法——OptimalGrouping 算法

首先将通过第  $h$  层第  $j$  个瓶颈链路的第  $m$  个会话的接收者和该会话在第  $h$  层的代理所代理的经过优化分组的接收者的性能进行排序和聚类,假设经过排序和聚类后,会话的所有接收者性能序列为  $\{r_1, r_2, \dots, r_{N_m}\}$ ,分组数量为  $K_m$ ,由于受会话分配带宽的限制,接收者性能最大只能为  $r_{N'_m} \leq \min(c_m^{h,j}, c_m^{h-1, \pi_m(j)}, r_{N_m})$ ,则调用如图 2 所示的会话内最优分组算法  $OptimalGrouping(K_m, \{1, 2, \dots, N'_m\})$  得到最优的分组,即可求得此时的最大会话效用。

```

OptimalGrouping( $K, \{1, 2, \dots, N\}$ )
1. for  $i=1$  to  $N$ 
2.    $V^*(i, 1) = U^*(\{1, 2, \dots, i\})$ 
3.    $p(i, 1) = -1$ 
4. for  $m=2$  to  $K$ 
5.   for  $i=m$  to  $N$ 
6.     if  $m \neq K$  then
7.        $V^*(i, m) = 0$ 
8.        $p(i, m) = i$ 
9.     else
10.       $i = N$ 
11.       $V^*(i, m) = 0$ 
12.       $p(i, m) = i$ 
13.      for  $j=m-1$  to  $i-1$ 
14.         $v = V^*(j, m-1) + U^*(\{j+1, \dots, i\})$ 
15.        if  $V^*(i, m) < v$  then
16.           $V^*(i, m) = v$ 
17.           $p(i, m) = j$ 
18.       $j = N$ 
19. for  $m=K$  to  $1$ 
20.    $i = j$ 
21.    $j = p(i, m)$ 
22.    $G_m = \{j+1, \dots, i\}$ 

```

图 2 会话内最优分组算法主要流程

算法说明如下:

(1) 假设用接收者序列编号  $\{1, 2, \dots, i, \dots, n\}$  对应接收者序列  $\{\bar{r}'_1, \bar{r}'_2, \dots, \bar{r}'_i, \dots, \bar{r}'_n\}$ , 其中  $i$  对应  $\bar{r}'_i$ , 对  $i$  的分组即是对  $\bar{r}'_i$  的分组,  $V^*(n, k)$  表示将  $\{1, 2, \dots, i, \dots, n\}$  分成  $k$  个组的最大会话效用,  $U^*(G)$  表示分组  $G$  的最大效用, 则有

$$V^*(n, k) =$$

$$\max_{k-1 \leq j < n} [V^*(j, k-1) + U^*(\{j+1, \dots, n\})].$$

(2) 要求解图 2 算法, 必须明确组最大效用  $U^*(\{j+1, \dots, i\})$  的求解方法。当分组  $\{j+1, \dots, i\}$  的发送速率等于  $\bar{r}'_{j+1}$  时, 分组 RTP 会话效用达到最大, 已知接收者效用函数为  $u(r', r) = \frac{optQ(r')}{optQ(r)}$ ,

这时的分组 RTP 会话效用的计算公式如下:

$$U^*(\{j+1, \dots, i\}) =$$

$$\sum_{k=j+1}^i d_{m,k} u(\bar{r}'_{j+1}, \bar{r}'_k) = \sum_{k=j+1}^i d_{m,k} \frac{optQ(\bar{r}'_{j+1})}{optQ(\bar{r}'_k)}.$$

由于  $\bar{r}'_{j+1}$  和  $\bar{r}'_k$  取分布在正整数域的离散值, 若已知率失真优化的视频质量函数  $optQ(r)$ , 可以通过事先的计算得到该函数的计算关系表, 已知速率  $r$ , 通过查表可以快速得到其对应的视频质量。

## 5 OptUtility 算法评价

### 5.1 OptimalGrouping 算法复杂性分析

OptimalGrouping 算法复杂性包括计算时间和空间复杂性, 由于视频组播应用对算法的实时性要求较高, 在此主要分析计算时间, 计算时间复杂度主要由以下几个部分组成:

(1)  $N$  个接收者的性能排序时间复杂度

考虑  $N$  会较大(一般大于 50), 本文采用快速排序方法, 时间复杂度为  $O(N \log N)$ 。

(2)  $N$  个接收者的性能聚类时间复杂度

聚类时间复杂度为  $O(N)$ , 假设经过聚类后, 不同接收者性能数量为  $N'$ ,  $1 \leq N' \leq N$ 。

(3) 会话内最优分组算法时间复杂度

效用函数的计算可以通过数组查表完成, 计算时间复杂度可忽略, 则图 2 所示 OptimalGrouping 算法主要计算流程时间复杂度约为  $O(N' + KN'^2 + K)$ 。考虑有分组负荷问题,  $K$  具体为多少时比较合适呢? 文献[4]通过均匀分布、正态分布以及双模分布模拟 300 个接收者接收性能的分布情况, 并使用接收者间公平性函数 IRF 作为接收者效用函数进行了数值仿真, 仿真结果说明: 当  $K=4$  或 5 时, 会

话效用已可以达到最大会话效用(对应于  $K=N'$ ) 的 80%,再增加层次,会话效用的增幅不大. 因此,  $K=4$  或 5 是比较理想的取值.

综合以上时间复杂度,若  $N'>50$  且  $N'$  不远小于  $N$ ,则 OptimalGrouping 算法的时间复杂度为  $O(KN'^2)$ ;否则,OptimalGrouping 算法的时间复杂度为  $O(N\log N+KN'^2)$ .

(4) OptimalGrouping 算法与 OptimalPartiton 算法的对比

表 1 OptimalGrouping 算法与 OptimalPartiton 算法的对比

算法	对比情形	
	$N'$ 与 $N$ 相当时, 时间复杂度	$N'$ 远小于 $N$ 时, 时间复杂度
OptimalGrouping 算法	$O(KN'^2)$	$O(N\log N+KN'^2)$
OptimalPartiton 算法	$O(KN^2)$	$O(N\log N+KN^2)$
评价	两者相当	OptimalGrouping 算法比 OptimalPartiton 算法明显优越

5.2 OptiUtility 算法复杂性分析

OptiUtility 算法的计算时间复杂度主要包括会话间带宽最优分配算法和会话内最优分组算法的联合计算时间复杂度.

(1) 会话内最优分组算法时间复杂度

不失一般性,假设分配给第  $m$  个视频组播会话的带宽为  $c_m$ ,该会话的所有接收者数量为  $N_m$ ,将分成  $K_m$ 组,该会话中最大的接收性能为  $r_{N_m}, 1\leq m\leq M_{h,j}$ ,由于受分配给会话的带宽的限制,接收者性能最大只能为  $r_{N'_m}\leq \min(c_m, r_{N_m}), 1\leq N'_m\leq N_m$ ,这样,该会话内最优分组算法的计算时间复杂度为:若  $N'_m>50$  且  $N'_m$  不远小于  $N_m$ ,则时间复杂度为  $O(K_mN'^2_m)$ ;否则,时间复杂度为  $O(N_m\log N_m+K_mN'^2_m)$ .

(2) 会话间带宽最优分配算法时间复杂度

假设用  $r_N$ 表示  $M_{h,j}$ 个视频组播会话中接收者性能最大值的平均值,则 BandwidthOptimalAllocation 算法的计算时间复杂度为  $O(M_{h,j}\cdot c_{h,j}\cdot r_N)$ .

(3) OptiUtility 算法与 OLBA 算法的对比

① 算法总体对比

OLBA 算法只考虑了通过一个瓶颈链路的小型视频组播系统整体效用最大化问题,而 OptiUtility 算法将一个大型视频组播系统分解为多个相互约束的层次化结构,是一个基于层次化代理的分布优化求解算法. 相比之下,OptiUtility 算法更具通用性.

② 会话间带宽最优分配算法

OptiUtility 的会话带宽最优分配算法与 OLBA 的会话间带宽最优分配算法基本相同,但本文提出的算法对 OLBA 的算法进行了改进,主要改进之处

文献[4]中提出的 OptimalPartiton 算法也要求对接收者性能进行排序,由于它没有进一步进行聚类,因此,其计算时间复杂度为  $O(KN^2)$ . 两个算法对比情况如表 1 所示. 通过比较可知,当  $N'$  远小于  $N$  时,OptimalGrouping 算法比 OptimalPartiton 算法的计算复杂度低很多,实际上,由于接收者性能的分布聚类特性随着  $N$  的增大将更加突出, $N'$  远小于  $N$  的概率就很大.

有以下两点:

(i) 在 OLBA 算法中假设已知视频组播应用系统的系统带宽,系统带宽的定义比较含糊,而本文提出的算法明确就是某个瓶颈链路带宽的最优分配问题,而且是针对整个视频组播应用系统的局部问题提出来的,目标十分明确.

(ii) 在 OLBA 算法中考虑了当系统带宽超过所有视频会话最大接收者性能之和的情况的处理,本文认为这种情况是多余的,因为为保证整个视频组播应用系统的稳定性,自适应性控制一般只在系统某些瓶颈链路出现严重拥塞现象时进行,在这种情况下,瓶颈链路的带宽肯定小于通过该瓶颈链路的所有视频会话最大接收者性能之和. 因此,OptiUtility 算法比 OLBA 算法要少考虑这种情况的处理,计算复杂度得到优化.

③ 会话内最优分组算法

OptiUtility 的会话内最优分组算法是假设分组数有固定上限的情况下的最优分组方案求解,而 OLBA 的会话内最优分组算法考虑的是最大分组数等于各个会话最大接收者性能时的最优方案求解,它没有考虑分组数量增大给整体效用带来的负效用,因此,OLBA 的求解复杂度很大. 而且,由于没有像 OptiUtility 算法的排序与聚类过程,OLBA 算法采取穷举的方法进行求解,其计算复杂度为  $O[(K_j)^3\cdot E]$ ,其中,  $K_j$  表示该会话中接收者接收性能的最大值,  $E=\sum_{k=m}^{K_j} M_{j,k} \Delta(k, m, i, l)$  表示的是在  $l-1$  层的基础上增加第  $l$  层后的会话效用差,其中  $M_{j,k}$  表示的是具有接收性能为  $k$  的接收者数量.

OLBA 的算法计算十分复杂. 为改进计算时间复杂度,文献[5]中提出了将  $E$  提前计算的算法,用空间复杂度换取时间复杂度. 这样考虑的前提是必须事先就知道所有接收者性能,并计算好存储起来,而对于接收者性能变化的实时计算将无能为力,相比之下,本文提出的算法可以不用事先知道会话所有接收者的性能,而一旦获得这些参数,可以进行实时计算.

④ 算法实时性对比实验

为了更好地说明 OptiUtility 算法的计算实时性和优化效果,本文采用与文献[5]同样的实验方

法,用应用感知的公平性指标作为接收者效用函数,用 C++ 作为编程语言,在配备 256MB 内存的奔腾 III 900MHz 的 PC 机实现算法,在与文献[5]同样的实验配置方案下,运行时间对比结果如表 2 所示,其中 OLBA 算法的运行时间结果来自文献[5]. 从对比结果来看,无论是会话间带宽分配算法还是会话内最优分组算法,OptiUtility 算法的计算复杂度都比 OLBA 算法优越,且都满足实时视频组播要求. 其中,  $(c, M, r^{\max})$  分别表示瓶颈链路带宽、通过瓶颈链路的会话数和所有会话接收者接收性能最大值.

表 2 OptiUtility 算法与 OLBA 算法计算时间复杂度对比

实验配置 ( $c, M, r^{\max}$ )	OptiUtility 算法运行时间/ms			OLBA 算法运行时间/ms		
	会话内分组	会话间分配	联合分配	会话内分组	会话间分配	联合分配
(64,8,12)	0.6	1.1	6.9	0.9	1.2	8.4
(128,10,15)	1.0	2.8	12.8	1.4	2.9	16.9
(512,20,30)	1.6	6.9	36.5	2.1	7.0	49.0

6 结 论

本文针对大型视频组播系统整体效用最大化问题,提出了基于瓶颈链路的层次化分解方法,建立了大型视频组播系统整体效用层次化模型. 考虑到相邻两个层次与同一视频组播会话相关的瓶颈链路带宽之间的约束性,完全分布的求解方法难以实现,且无法满足实时视频组播应用要求的问题,本文提出了一个由相邻两个层次进行“基于层次化代理的分布优化求解方法”OptiUtility 算法. 通过计算时间复杂度分析和优化对比实验,OptiUtility 算法对于解决大型视频组播系统整体效用优化问题具有通用性,且计算时间复杂度更优,实时性能更好.

参 考 文 献

[1] Amir E, Mccanne S, Katz R H. An active service framework and its application to real-time multimedia transcoding//Proceedings of the SIGCOMM'98. Vancouver, BC Canada, 1998: 178-189

[2] Li X, Paul S, Pancha P, Ammar M H. Layered multicast with retransmission (LVMR): Evaluation of error recovery schemes//Proceedings of the NOSSDAV'97. Washington University, St Louis, 1997

[3] Wang Hui. On optimization for large-scale adaptive video multicast in heterogeneous environments [Ph. D. dissertation]. Changsha: National University of Defense Technology, 2005(in Chinese)

(王晖. 异构环境下大型自适应视频组播方法优化研究[博士学位论文]. 长沙:国防科学技术大学, 2005)

[4] Yang Y, Kim M, Lam S. Optimal partitioning of multicast receivers//Proceedings of the ICNP'00. Osaka, Japan, 2000: 129-140

[5] Liu J, Li B, Hou Y T, Chlamtac I. Dynamic layering and bandwidth allocation for multi-session video broadcasting with general utility functions//Proceedings of the IEEE INFOCOM 2003. San Francisco, CA, USA, 2003

[6] Koushik Kar, Saswati Sarkar, Leandros Tassioulas. Optimization based rate control for multirate multicast sessions//Proceedings of the IEEE Infocom. Anchorage, Alaska, USA, 2001: 123-132

[7] Wang Hui, Zhang Jun, Jiang Zhi-Hong et al. ALM: A novel adaptive layer mapping algorithm for intra-session bandwidth allocation in layered video multicasts//Proceedings of the 8th IASTED International Conference on Internet and Multimedia Systems and Applications. Kauai, Hawaii, USA, 2004: 124-130

[8] Liu J, Li B, Zhang Y-Q. An end-to-end adaptation protocol for layered video multicast using optimal rate allocation. IEEE Transactions on Multimedia, 2004, 6(1): 87-102

[9] Christos Bouras, Apostolos Gkamas. SRAMT: A hybrid sender and receiver-based adaptation scheme for TCP friendly multicast transmission. Computer Networks, 2005, 47(4): 551-575

[10] Liu J, Li B, Zhang Y-Q. Adaptive video multicast over the Internet. IEEE Multimedia, 2003, 10(1): 22-31

[11] Li B, Liu J. Multirate video multicast over the Internet: An overview. IEEE Network, Special Issue on Multicasting: An Enabling Technology, 2003, 17(1): 24-29





**SHAN Ji-Chang**, born in 1945, professor, Ph. D. su-

**WANG Hui**, born in 1968, Ph. D. , professor, Ph. D. supervisor. His re- search interests include information sys- tem engineering, multimedia networks.

**Background**

With the popularization of access to broadband net- works, the demand for network video services is on the great increase in recent years. Video multicast, the particular net- work video service apt at multipoint data delivery, has been widely applied to distance learning, distributed video confer- ence, digital video surveillance and real-time video distribu- tion. In accordance with the enlarging of video multicast ap- plication scale, the capability of video multicast networks and receivers varies greatly. Video multicast has suffered criti- cism on its performance, such as the complexity of manage- ment, the difficulty of Quality of Service (QoS) guarantee, etc. In order to solve these problems, the academic and in- dustrial communities have concerned on adaptive video multi- cast (AVM) for years, but there are still many theoretical and practical problems left to be resolved, especially for large-scale video multicast applications. This paper studies the optimization for large-scale AVM to make it apply in het-

pervisor. His main research interests focus on analysis of de- fense systems.

**JIANG Zhi-Hong**, born in 1976, Ph. D. candidate, lec- turer. His research interests include information system en- gineering, multimedia networks.

**ZHANG Xin**, born in 1978, Ph. D. , lecturer. His main research interests focus on process of multimedia informa- tion.

erogeneous environments.

In this paper, the authors first classify and analyze sev- eral typical AVM schemes. It is found that most of these schemes take small-scale video multicast applications as the backgrounds, and they only improve the local performance metrics. To extend the results, a total system utility model for AVM systems is presented in this thesis. And an extend- able, hierarchically distributed solution is also proposed, by which we can figure out optimally the bandwidth, the layer number, and the layer rates for each session. The proposed method can handle the problems of feedback implosion, sin- gle-node failure, and bad real-time performance caused by centralized adaptive control and decision in large-scale AVM applications. Experimental results and analysis of computa- tional complexity prove the proposed distributed method to be of low complexity, easy in implementation, and well per- formed in real time.