

# 交互式 VOD 视频源非线性智能分段算法

杨 灿 卢正鼎

(华中科技大学电子科学技术系 武汉 430074)

**摘 要** 为了实现视频点播系统的交互性功能,经典的方法是将视频源均匀分割存储,以实现其有限的交互性.从 VOD 系统的实际出发,该文提出了非线性智能分段方法 NLSM.它的核心思想是在一段时间内对一定步长范围内 VCR 交互性操作的到达情况进行统计求和,计算系统总体时频总面积,并按信道数目将各时间段与 VCR 发生频率的乘积把总面积均匀等分,从而实现视频源在时间轴上的非线性分割.针对均匀的周期性广播分段方法 ESM 方案进行比较仿真研究,获得了用户 VCR 请求的响应等待时间的理想结果和最佳的用户等待时间标准偏差、系统拒绝概率和信道吞吐量等方面的优良性能.NLSM 算法对历史信息进行数据挖掘,揭示用户对视频节目内容感兴趣程度的潜在分布规律,具有良好的自适应能力和学习智能.算法性能分析和仿真实验的结果说明该方法在实际应用中是可行且高效的.

**关键词** 视频点播(VOD); VCR; 多点广播; 视频分段; 非线性  
中图法分类号 TP391

## A Non-Linear Algorithm for Video Segmentation on Interactive Video-On-Demand System

YANG Can LU Zheng-Ding

(Department of Electronic Science & Technology, Huazhong University of Science & Technology, Wuhan 430074)

**Abstract** To realize interactive performance of a Video-On-Demand system, a traditional method to segment a video with equal fragments for storage is adopted to implement limit interactive performance of a VODs. According to the reality of VODs, A non-linear segmentation method (NLSM) is adopted in the paper. The key of NLSM is to realize non-linear segmentation of a video in time axis. After computing the total area of time and frequency, NLSM will equally segment the product of interval and frequency of VCR-Requests arrival according to statistic sum of VCR operation in certain step width in a period of time. Comparing with periodical broadcasting by equal segmentation method(ESM) in simulations, we obtained good outcomes in the mean latency of VCR requests and the best performance of standard deviation of user latency, system deflection probability and throughput of a channel. It is found that NLSM can make a data mining in historical information and reveal potential distribution law of users interest in video content. This scheme has good adaptive ability and learning intelligence. By the performance analysis and empirical results, it could be verified that the scheme is feasible and efficient.

**Keywords** Video-on-Demand(VOD); VCR, multicast; video segmentation; non-linear

## 1 引 言

视频点播(VOD)系统对系统资源和网络带宽

的苛刻要求<sup>[1]</sup>,使它在过去十年的发展中碰到很多困难,同时也在国内外受到了广泛的研究<sup>[2]</sup>,尤其是关于它的交互性能的有效实现方法.针对基于多点广播 Multicast<sup>[3,4]</sup>的 VOD 应用,本文主要讨论 Hot

Spot(流行热点)影片智能化分段的问题解决. 无论是集中式还是分布式的体系结构, 视频源的存储和播放方式对系统的影响都是极为重要的.

VOD 系统中大量用户在不同时间选择了同一部影片(这里称之为 Hot Spot 问题), 同时还希望系统提供全 VCR 功能, 即提供快进、倒带、暂停等交互业务, 这些业务必然会产生许多访问冲突和信道拥挤问题.

为了避免使用大型存储器, 一般采用均匀分段传输的技术手段<sup>[5,6]</sup>来满足用户 VCR 请求的需要, 体系结构如图 1 所示.

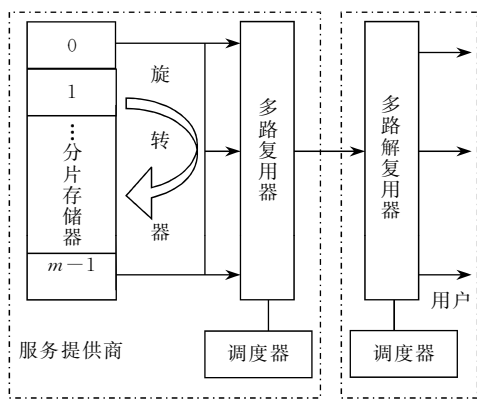


图 1 分段传输体系结构

## 2 非线性分段 NLSM 方法

从图 1 可以看出, 系统采用均匀分段策略(Equal Segment Method, ESM), 从而改进了 VOD 系统的交互性. 但是, 我们认为对不同的影片, 统一采用均匀分段的策略并非最优化的解决方案. 因此我们提出非线性分段的策略(Non-Linear Segment Method, NLSM), 采用一种改进的时频面积等分智能计算技术来解决影片非线性分段的问题. 我们的工作是基于以下的理论前提:

- (1) 观众对一部影片中不同片段的感兴趣程度是不同的, 因此 VCR 控制到达点位分布是非线性的;
- (2) 不同影片 VCR 控制的到达点位分布是不同的, 并且互不相关;
- (3) 不同的观众对同一部影片的兴趣点有相对一致或相对集中的分布, 它不是一个线性分布, 而是一个与影片源相关的非线性分布;
- (4) 统计概率是机器学习与系统优化的重要根据.

### 2.1 数学模型

为了表述我们的方法, 先建立 VOD 中 VCR 控

制的数学模型如图 2 所示, 这里以一个影片的帧序代表时间序列, 我们的分析认为视频流的切换与 VCR 的起点位置无关, 而只与其落点位置有关. 其中

$$\delta(x) = \begin{cases} 1, & x \text{ 有 VCR 请求到达} \\ 0, & x \text{ 无 VCR 请求到达} \end{cases} \quad (1)$$

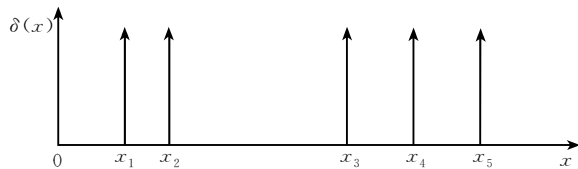


图 2 VOD 中 VCR 控制冲击响应函数模型

求解的基本思路是: 将总长度为  $L$  的影片均匀分割成长度为  $\Delta x$  的子段, 则总共分成  $L/\Delta x$  段, 经过统计获得在  $[x, x+\Delta x]$  区间内的请求数为

$$sum(i) = \sum_{i\Delta x}^{(i+1)\Delta x} \delta(x), \quad i = 0, 1, 2, \dots, [L/\Delta x].$$

总请求数

$$Sum = \sum_0^L \delta(x).$$

然后求得系统时频域的总面积

$$S = \sum_0^{[L/\Delta x]} sum(i) \Delta x = \Delta x \sum_0^{[L/\Delta x]} sum(i).$$

假设系统最多可为一部影片提供  $C$  个信道, 则可将影片分割成  $C$  个子段. 我们将总面积  $S$  均匀分配为  $C$  个面积相等的时频子块, 再在时间轴上求解分割区间——这是本文的核心思想. UEMS 实现了频率高的部分间隔时间短, 频率低的部分间隔时间长, 从而使系统得到优化, 即每个 VCR 控制的平均等待时间减少, 系统总体等待时间缩短. 下面给出具体实现过程和算法性能分析及仿真实验结果.

本节中的有关参数意义如表 1 所示.

表 1 有关符号代表的物理意义

符号	意义	单位
$X$	帧序号	
$T$	影片时间长度	s
$L$	影片总帧数	
$V$	播放速率	frame/s
$\delta(x)$	第 $i$ 帧收到 VCR 的状态	
$C$	系统可为同一影片提供的信道总数	
$\tau$	步长	
$p(i)$	第 $i$ 个步长区间内 VCR 收到的频率	
$s(i)$	第 $i$ 个时频的区间面积	
$S$	一段统计时间内的系统时频总面积	
$S_C$	按 $C$ 个信道划分的时频面积	
$F$	重新分段后的段的序号	
$x_f$	重新分段后的段序号在原影片中的对应帧序号	
$\rho(f)$	新区间 $(x_f, x_{f+1})$ 内 VCR 的到达频率	
$W(f)$	新区间 $(x_f, x_{f+1})$ 内 VCR 的平均等待时间	

## 2.2 算法描述

UEMS 优化分割,可按以下步骤求解.

1. 设  $\tau$  为固定步长(与 2.1 节中所提  $\Delta x$  相对应),它代表一段连续帧的距离,且  $0 < \tau \ll L/C$ ,将整个影片均匀分为  $k$  个微小片段;

$$k = L/\tau \quad (2)$$

这样影片可以看成是以  $\tau$  为单位的序列  $\{s(i)\}$ , 其中  $i = 0, 1, 2, \dots, k-1$ ;

2. 统计求解每个固定步长  $\tau$  内的

$$p(i) = \sum_{j=x_i}^{x_i+\tau} \delta(j) \quad (3)$$

3. 求解  $s(i)$  所在  $\tau$  区间的矩形面积,取高为  $p(i)$ ,则有  $s(i)$  的面积为  $s(i)$ ,如下式所示,

$$s(i) = \tau p(i) = \tau \sum_{j=x_i}^{x_i+\tau} \delta(j) \quad (4)$$

4. 在时间  $T$  内,系统总面积  $S$  可用下式求得

$$S = \sum_{i=0}^{k-1} s(i) = \tau \sum_{i=0}^{k-1} p(i) = \tau \sum_{i=0}^{k-1} \sum_{j=x_i}^{x_i+\tau} \delta(j) \quad (5)$$

5. 当系统提供的有限信道数为  $C$  时,系统可以在时频二维空间内分为  $C$  等分,每等分的面积为

$$S_C = S/C = \frac{\tau}{C} \sum_{i=0}^{k-1} \sum_{j=x_i}^{x_i+\tau} \delta(j) \quad (6)$$

6. 顺序搜索:在  $i = 0, 1, \dots, k-1$  上求解满足下式的  $d(f)$  个  $\tau$  合并的区间:

$$\tau d(f) \sum_{i=0}^{d(f)} p(i) = S_C \quad (7)$$

这里  $f = 0, 1, \dots, C-1$ ;  $d(f)$  为  $\tau$  的个数;

满足式(7)的  $d(f)$  值及其起始和终止位置为我们求解的结果,且有

$$x_{f+1} - x_f = \tau d(f) \quad (8)$$

$x_{f+1}, x_f$  为它们在影片中的绝对帧序号. 为了更严格起见,我们可以把式(7)写作

$$(x_{f+1} - x_f) \sum_{j=x_f}^{x_{f+1}} \delta(j) = S_C \quad (9)$$

且

$$\sum_{f=1}^C (x_{f+1} - x_f) = L \quad (10)$$

式(9)比式(7)具有更严格的物理意义,也便于编程实现,只是如果利用式(3)的计算结果,可使计算速度更快. 它们计算方法实质上是等效的.

## 2.3 算法分析

以上算法是基于有大量先验样本为基础的,是基于过去事件的统计数据.

2.2 节中  $p(i)$  的统计步长  $\tau$  是一个重要的参数. 它即不能太小,也不能太大. 为了使算法描述表达更清晰,我们在 2.2 节中第 4 步采用影片时间长

度  $T$  为统计时间,在实际系统设计中统计时间可以取  $T$  的整数倍或影片有效播放的时间. 理论上统计时间越长,统计样本越多,系统分割越趋于实际和稳定,同时运算量也将增大.

如果取影片有效播放的时间长度  $nT$  为统计时间,则有式(3)~(9)的表达式依然不变,只是各式的结果发生了变化,当式(9)的解不随统计时间和统计量变化时,系统达到稳定. 我们给出系统在统计时间  $nT$  和  $(n+1)T$  分段结果的误差函数为

$$\sigma = \frac{\sqrt{\sum_{f=0}^{C-1} (|x_{f+1}^{n+1} - x_f^{n+1}| - |x_{f+1}^n - x_f^n|)^2}}{C} \quad (11)$$

当  $\sigma \rightarrow 0$ , 系统达到稳定状态,可以得出其优化的分段方案. 系统稳定状态下,在一个新区间  $(x_f, x_{f+1})$  内,新的 VCR 操作到达频率可记为  $\rho(f)$ , 它满足下式

$$\frac{|x_{f+1} - x_f|}{\tau} \times \rho(f) = \frac{1}{C} \quad (12)$$

有

$$\rho(f) = \frac{\tau}{|x_{f+1} - x_f| \times C} \quad (13)$$

上式的物理意义是很清晰的,我们由此可知:

(1)  $\rho(f)$  与分段后的段长  $(x_{f+1} - x_f)$  成反比;

(2)  $\rho(f)$  与系统可为同一部影片提供的信道总数  $C$  成反比;

(3)  $\rho(f)$  与统计积分的步长  $\tau$  成正比.

下面比较分析我们提出的非线性分段方法与均匀分段方法的使用效率.

按照均匀分段方法,一部影片被均匀分割成  $C$  段,每一段的长度为  $L/C$ ,这一分段方案实际上是认为一个随机的 VCR 控制到达每一段的概率是一常数  $\rho_0 = 1/C$ ,这显然是与实际的 VCR 分布不相符合的,因此它对系统的影响也是很大的.

采用我们提出的非线性分段方法远远优越于均匀分段方案,下面的实验仿真得到了良好的结果.

## 3 仿真实验与性能比较

### 3.1 仿真实验

下面来比较 NLSM, ESM 对一个给定的典型的 VCR 达到进行视频源子段分割的结果.

如图 3 和图 4 所示曲线为实际 VODs 中常见的 VCR 分布情况. 图 3 和图 4 是分别采用 ESM, NLSM 将其分割成 16 个子段的结果. 这些图中,横坐标表

示视频节目的播放时间长度为 90min, 纵坐标表示在相应的视频子段位置的 VCR 到达的实际历史统计值。

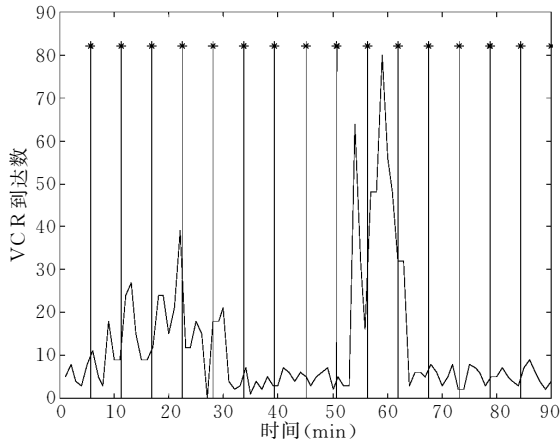


图 3 ESM 分割结果 (带星号线条表示分割位置)

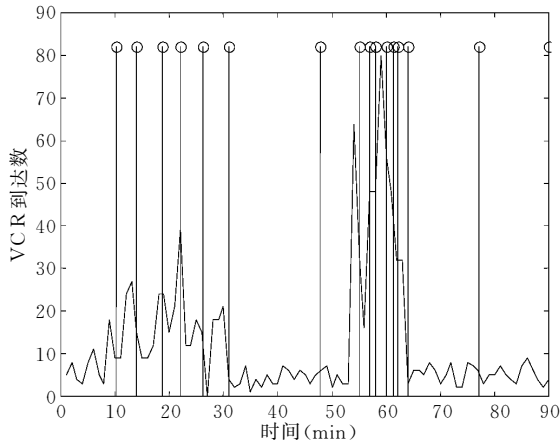


图 4 NLSM 分割结果 (带圆圈线条表示分割位置)

从图 3 和图 4 可以清楚地看出: NLSM 分割的结果是相当良好的, 因为该方案的基本思想充分考虑用户在一个视频节目各子段内容的兴趣程度, 通过历史统计的 VCR 到达分布进行自学习的智能化分割, 这种分割相对 ESM 是随用户 VCR 到达分布在时间和空间上的变化而变化的, 因而它是一种完全自适应的动态分割方案. 它的基本设计思想中就蕴涵着用户对不同视频节目的不同子段的兴趣的潜在分布特征, 因此得到较为完美的分割结果。

### 3.2 性能比较

为了更进一步地分析研究用户 VCR 请求到达率与系统提供的信道数对分割性能的影响, 在以上分段实验相似的环境下, 我们得到了用户 VCR 请求的平均等待时间及其标准偏差、系统拒绝概率和信道利用率随到达率和信道数的变化关系. 本小节

中采用的符号和性能参数定义如表 2 所示。

表 2 符号意义

符号	意义	单位
$N$	微分区间的数目	
$Pk(i)$	每一微分区间的历史统计数目	
$ML$	每个 VCR 请求的平均等待时间	min
$L(i)$	VCR 请求 $i$ 的等待时间	min
$SN$	历史统计的 VCR 请求的总数 $SN = \sum_{i=1}^N Pk(i)$	
$P(i)$	每一微分区间的统计到达概率 $p(i) = Pk(i)/SN$	
$STD$	VCR 请求等待时间的标准偏差	min

平均等待时间:

$$ML = \sum_{i=1}^N L(i)p(i) \quad (14)$$

等待时间的标准偏差:

$$STD = \sqrt{\frac{\sum_{i=1}^N [(L(i) - ML)p(i)]^2}{N}} \quad (15)$$

在本文的研究中定义超过某一预先设定的等待时间阈值时, 发出 VCR 请求的用户将放弃等待, 系统中用户放弃等待的数量占系统接收到的请求总数的比值为系统拒绝概率  $DP$ , 这里取时间阈值为平均等待时间  $WL$ .

拒绝概率:

$$DP = \sum_{i=1}^N p(i), \quad L(i) > WL \quad (16)$$

未超过等待时间阈值的用户请求被系统接受, 平均每个信道中系统接受的请求数占系统接收到的请求总数的比值为系统吞吐量或系统信道利用率

$$CUP = \sum_{i=1}^N p(i) = 1 - DP \quad (17)$$

下面图 5 和图 6 给出平均等待时间随到达率和信道数的变化; 图 7, 图 8 给出其标准偏差随到达率和信道数的变化; 图 9, 图 10 给出系统拒绝概率随到达率和信道数的变化; 图 11, 图 12 给出信道利用率随到达率和信道数的变化关系。

从图 5 和图 6 可以发现, ESM 与 NLSM 的平均等待时间是十分接近的, 从概率的意义上可以认

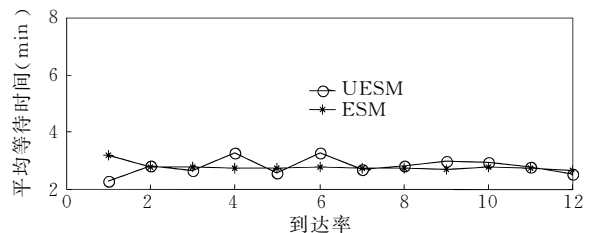


图 5 VCR 到达平均等待时间随到达率的变化 ( $C=16$ )

为它们相等,这与本文 NLSM 基于时频面积等分的基本思想是一致的。NLSM 的优点主要不是表现在其总体平均等待时间上。

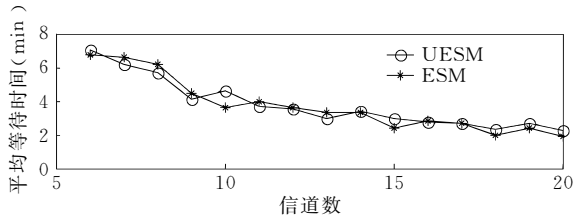


图 6 VCR 到达平均等待时间随信道的变化 ( $\lambda=5$ )

标准偏差是 NLSM 优越于 ESM 的重要特性,它代表用户个体等待时间偏离其平均值的大小;显然 STD 值越大,系统性能越差,即有较多的用户等待时间与其平均值偏离较大;STD 值越小,系统性能越好,即只有较少的用户等待时间与其平均值偏离较大。图 7 与图 8 反映 NLSM 较 ESM 在该项指标上具有更好的性能。

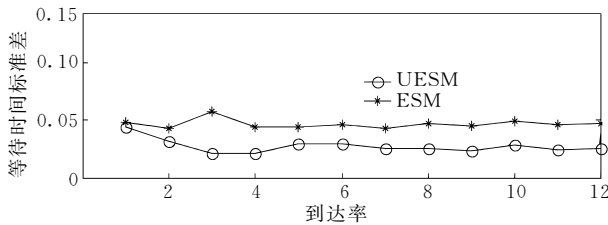


图 7 等待时间标准偏差随到达率  $\lambda$  的变化 ( $C=16$ )

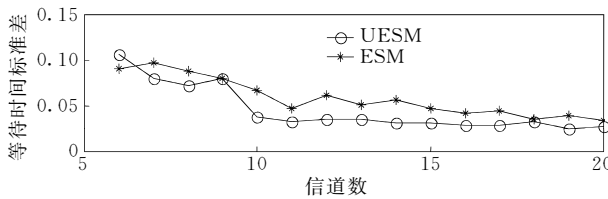


图 8 等待时间标准偏差随信道数的变化 ( $\lambda=5$ )

从图 9 和图 10 可以发现,ESM 与 NLSM 的系统拒绝概率差距较大,NLSM 具有较小拒绝概率,

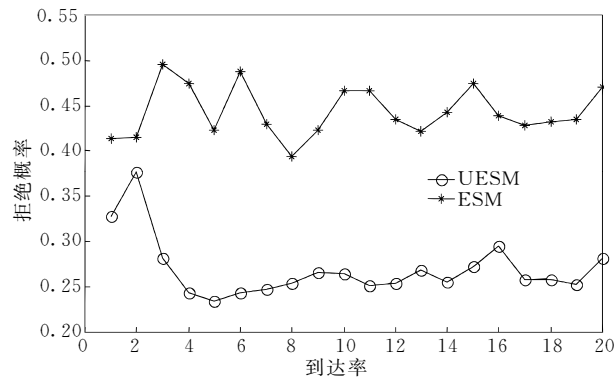


图 9 用户拒绝概率与到达率的关系 ( $C=16$ )

显然,拒绝概率越小系统性能越好,最小的拒绝概率是系统设计优化的主要目标之一,因此 NLSM 方法的优越性能已经清晰可见。

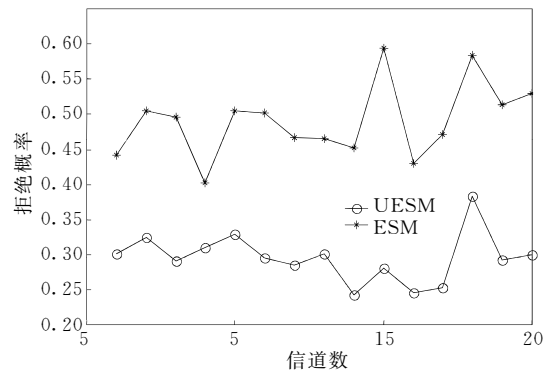


图 10 用户拒绝概率与信道数的关系 ( $\lambda=5$ )

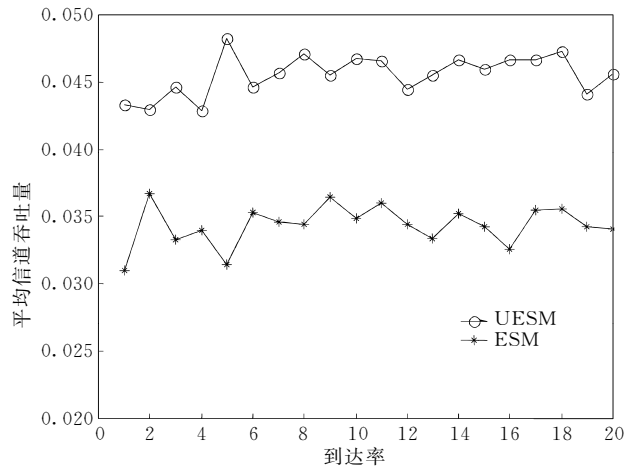


图 11 信道吞吐量与到达率的关系 ( $C=16$ )

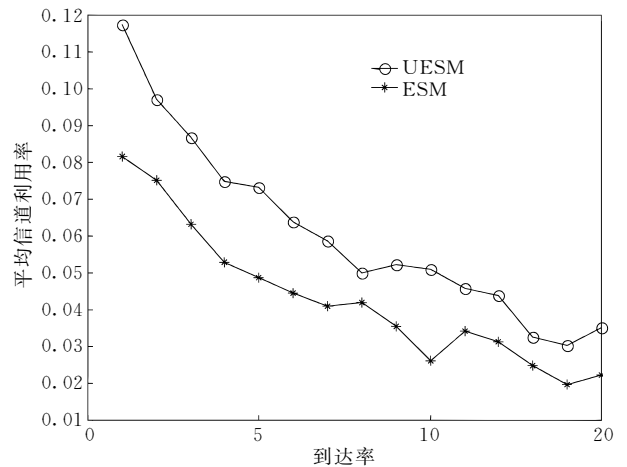


图 12 每信道吞吐量与信道数的关系 ( $\lambda=5$ )

从式(17)可以发现系统信道吞吐量或信道利用率跟系统拒绝概率是直接相关的,图 11 反映系统平均信道吞吐量与到达率没有明显的变化,这主要是因为公式中采用的都是归一化概率计算的结果。图

12 指出在到达率不变的情况下,随着系统信道数的增加,平均信道利用率降低,这是显然的.从图中曲线可以清楚地看出在相同信道数的情况下,系统平均吞吐量 NLSM 大于 ESM.

平均信道利用率和系统拒绝率一样是重要的系统性能参数,它反映系统平均每一信道的通信能力,从图 11 和图 12 中我们得到 NLSM 在系统吞吐量方面的表现,再次有力地证实了它在视频源分段方案中的优良性能.

## 4 结束语

在实际的交互式 IVOD 应用中,如何实现一部影片的非线性分段一直是一个很难有明确解答的问题.一般人们均保守地采用均匀分段的 ESM 方案,但是这实际上对信道资源和用户请求响应时间造成很大的浪费;相比之下,本文提出的时频面积均分的 NLSM 算法在以下几方面具有明显的优越性:

(1) 基于大量用户点播信息的统计结果进行学习和分类,因而具有普遍意义;

(2) 时频面积均分法保证 VCR 密集度高的段分辨率高;密集度低的段分辨率低;

(3) 在不增加系统平均等待时间的基础上,缩小用户 VCR 到达等待时间的标准偏差,保障绝大多数用户的等待时间不超过给定的等待时间阈值;

(4) 大大降低系统的拒绝概率;

(5) 有效保障系统的吞吐量,提高平均信道的利用率;

(6) 大大降低绝大多数用户 VCR 请求等待的时间;

(7) 适应性强,与系统底层无关,可以在任何支持 Multicast 的 VOD 广播系统中应用.

本文提出的视频源智能化非线性分段 NLSM 算法,是在深入研究用户 VCR 请求在特定视频节目上分布的普遍规律的基础上提出的,它能自适应地挖掘出用户对于特定视频节目上视频子段的兴趣集中程度的潜在分布.因而,NLSM 在不增加系统平均等待时间的前提下,获得了用户等待时间标准偏差、系统拒绝率和平均信道利用率等方面的优良性能.

## 参 考 文 献

- 1 Li O K, Liao W J, Qiu X X, Wong W M: Performance model of interactive Video-on-Demand systems. *IEEE Journal of Selected Areas in Communications*, 1996, 14(6): 1099~1109
- 2 Aetal V. Distributed multimedia applications and quality of service: A survey. *IEEE Multimedia Journal*, 1995, 3(8): 87~95
- 3 Armitage G J. IP multicasting over ATM networks. *IEEE Journal of Selected Areas in Communications*, 1997, 15(3): 445~457
- 4 Almeroth K C, Ammar M H. The use of multicast delivery to provide a scalable and interactive Video-On-Demand service. *IEEE Journal of Selected Areas in Communications*, 1996, 14(4): 1111~1122
- 5 Zhou Hui, Zhang Lian-Zhong, Chen Li-Qiang. VOD: The Soul of Information High Speed Way. Beijing: China Ocean Press, 1998(in Chinese)  
(周 晖,张连仲,陈利强. 信息高速公路上的灵魂——视频点播 VOD. 北京:海洋出版社,1998)
- 6 Zhang L, Fu H. Dynamic bandwidth allocation and buffer dimensioning for supporting Video-on-Demand services in virtual private networks. *Journal of Computer Communications*, 2000, 23: 1410~1424



**YANG Can**, born in 1971, Ph. D. . Since Aug. 2002, he has been working at the postdoc of Computer Science and Technology in HUST. His research interests include multimedia communications and application software on Internet.

**LU Zheng-Ding**, professor, Ph. D. supervisor. His research area is software engineering, database and so on.