第36卷 第15期 计 算 机 工 程 2010年8月
Vol.36 No.15 Computer Engineering August 2010
·博士论文· 文章编号：1000—3428(2010)15—009—03 文献标识码：A 中图分类号：TP311

基于网络编码的P2P-VoD数据调度算法

胡玉琦，王 勇，蔡璐玮
(燕山大学信息科学与工程学院，秦皇岛 066004)

摘 要：针对基于网络编码的P2P-VoD数据调度中对Peer节点异构性和网络动态性考虑不足的问题，提出一种基于自适应随机线性网络编码的P2P-VoD数据调度算法。Peer节点根据自身缓存、链路可用带宽和视频块的播放时限自适应调整网络编码窗口大小。基于NS2设计P2P-VoD仿真通用平台进行实验。实验结果表明，该算法具有节点的自适应异构性和网络动态性，能提高Peer节点的视频播放连续度。
关键词：网络编码；视频点播；数据调度；NS2模拟器

Data Scheduling Algorithm of P2P-VoD Based on Network Coding

HU Yu-qi, WANG Yong, CAI Lu-wei
(School of Information Science and Engineering, Yanshan University, Qinhuangdao 066004)

【Abstract】Aiming at the problem of the lack consideration of Peer node heterogeneity and the dynamics of networks in data scheduling of P2P-VoD based on network coding, a data scheduling algorithm of P2P-VoD based on adaptive random network coding is proposed. Network coding window is adaptively adjusted by the Peer node according to its cache, the link available bandwidth and the deadline of playable block. The simulation is conducted on the general P2P-VoD simulation platform based on NS2 design. Experimental results show that the node heterogeneity and the dynamics of networks are adaptive, the playback continuity index in the peer node is enhanced.
【Key words】network coding; Video-on-Demand(VoD); data scheduling; NS2 simulator

1 概述

P2P-VoD系统通过利用网络上普通主机节点的资源来提供流媒体数据服务，是一种扩展性好、性价比高的视频点播服务体系。由于节点动态、网络异构、单点瓶颈等P2P网络特性，加上用户对媒体的实时性要求，P2P-VoD系统服务质量很难保障。Peer节点间高性能的数据调度算法是保障Peer节点服务质量的关键技术之一。

网络编码是一种融合编码和路由的信息交换技术，在传统存储转发的路由方法基础上，通过允许对接收的多个数据包进行编码信息融合，增加单次传输的信息量提高网络整体性能。将网络编码理论应用于P2P流媒体系统进行数据调度[1-2]，从而发挥网络编码提高网络吞吐量、减小传输延迟、增强网络鲁棒性等优越性，提升系统Peer节点服务质量。

现有基于网络编码的P2P-VoD系统数据调度研究有：Redcarpet[3]在服务器端将视频文件分割成块，由一定数量的连续块组成一个视频段，在段内使用网络编码解决块调度问题。网络编码窗口固定，系统强加了一个启动缓冲时间，这个时间至少为一个视频段的长度；DNC[4]提出了一种具有截至时间意识的网络编码数据调度算法，解决了基于网络编码的P2P-VoD系统数据调度中视频块的延时问题。DNC仅考虑Peer节点播放块时限来确定网络编码窗口大小。以上现有的2种调度策略提高了Peer节点的播放质量和网络资源的利用率。但是，这2种策略都没有考虑Peer节点及链路的异构性。在实际情况中，链路的可用带宽和节点的缓存情况直接影响视频流稳定连续的传输和Peer节点视频的流畅播放。

本文在现有研究基础上提出了一种基于自适应随机线性网络编码的P2P-VoD数据调度算法——ARLNC，设计了基于NS2的P2P-VoD系统仿真平台进行仿真实验。Peer节点在综合考虑自身缓存、链路可用带宽和即将播放视频块时限的情况下，动态调整网络编码窗口大小，自适应网络动态性和Peer节点异构性，提高了Peer节点播放质量。

2 ARLNC数据调度算法设计实现

在ARLNC调度算法中，假设Peer节点已和系统中M个节点建立了邻居伙伴关系，所有的块都能够被无差错接收。规定所有的Peer节点必须从头开始观看视频，不支持各种VCR操作。由于P2P网络的分布式特性，调度算法中采用随机线性网络编码。随机线性网络编码是一种实用的分布式网络编码方式，容易实现，计算开销较小。

由于流媒体系统特点，将网络编码理论应用于P2P流媒体系统时，必须在源服务器端对视频流进行预处理：将文件按固定大小进行块划分，由连续的若干个块组成段，在段内使用网络编码理论进行调度。定义连续的组成段的块的个数为网络编码窗口大小(CW)。Peer节点只有从邻居节点收到CW个打包的数据块，并且这CW个数据块所对应的系数线性无关时，Peer节点才能完整解码整段数据。

分析发现，编码窗口大小的选择是影响基于网络编码P2P-VoD系统性能的重要因素：编码窗口过小，不能充分发挥网络编码带来的优势；编码窗口过大，因节点需要较长时间来收集足够多的信息才能解码，造成播放不够流畅。编码窗口大小的选择需要考虑整个网络的拓扑结构、节点数，节点缓存和带宽等情况。这里首先给出Peer节点在考虑自身缓

基金项目：国家自然科学基金资助项目(60273078)
作者简介：胡玉琦(1964－)，女，副教授、博士，主研方向：对等网流媒体技术；王 勇、蔡璐玮，硕士研究生
收稿日期：2009-12-23 **E-mail：**wang_yong_260@126.com

存、链路可用带宽和即将播放块时限的情况下，自适应计算编码窗口大小 CW 的大小，然后给出了 Peer 节点数据调度算法伪代码实现。

2.1 Peer节点自适应编码窗口计算

Peer 节点加入 P2P-VoD 系统，初始编码窗口计算：

$$CW_{\text{init}}=B_{\min}(t_{\text{next}}-t_{\text{now}})N_{\text{nei}}$$

其中，CW_{init} 为初始编码窗口大小；$B_{\min}$ 为 Peer 节点与其所有邻居节点间可用带宽的最小值；t_{next} 为 Peer 节点即将播放的下一个块的播放时限；t_{now} 为 Peer 节点现在播放的块的时限；N_{nei} 为 Peer 邻居节点中具有 Peer 节点即将播放块的节点个数。

Peer 节点在前一次编码窗口大小的基础上不断自适应调整编码窗口大小：

$$CW=\beta\{[P_{\text{cachemax}}M/S_{\text{play}}-T_{\min}]CW_{\text{front}}/T\}+(1-\beta)CW_{\text{front}}$$

其中，P_{cachemax} 为 Peer 节点调度块缓存的最大值；M 为每个块的大小；S_{play} 为码率；CW_{front} 为前一个段的编码窗口大小；T 为下载完毕并解码上一个视频段所用的时间；$T_{\min}$ 是播放缓存内的最小预留时间，选取这一冗余量是为了尽量保证整个播放期间缓存内至少有可供播放 $T_{\min}$ 时间的内容，从而在发生下载困难等情况时为处理赢得时间，而不因缓存中没有可供播放的内容而迫使播放线程中断播放。在启动播放线程前，$T_{\min}$ 取值为 0，直到缓存填充完毕。其余情况 $T_{\min}$ 取值为播放缓存的 1/4 左右。另外，为了限制编码窗口 CW 分布过于分散，引入了一个加权因子 β，取值为 0.5。

2.2 Peer节点数据调度算法

数据调度算法独立运行在每个 Peer 节点上。对于 Peer 节点的邻居节点(假设有 M 个)，数据调度算法的任务是决定该 Peer 节点应该从哪些邻居节点获取哪些数据块，以便每个数据块在播放之前都能被网络解码；如果不能被及时解码，则使不能被及时解码的数据块个数最少。下面给出调度算法的主要过程：

输入

$\{1\leqslant i\leqslant M\}$：运行调度算法的节点 P 的邻居节点集合

Block_size ：每个数据块的大小

$Band(P,P_i)$：P 与 P_i 之间的可用网络带宽，以数据块大小为单位计

$rank(seg,P)$：P 上属于 seg 的系数向量所组成矩阵的秩数

$rank(seg,P,P_i)$：联合 P 及 P_i 上属于 seg 的系数向量所组成矩阵的秩数

算法过程：

```
For each peer join in P2P-VoD do
Calculate CW;
  If rank(seg)<CW then
For i=1 to M do
If rank(seg,P,Pi)>rank(seg, P) then
      Block_number=min{band(P,Pi),rank(seg,P,Pi)-rank (seg,P)};
      Request and set block_number of data blocks from Pi that
Belongs to seg;
      Band(P,Pi)=band(P, Pi)-block_number;
      If rank(seg, P) reaches CW, then break;
     End if
   End   for i
  End if
End for peer
```

3 P2P-VoD系统通用仿真平台设计

由于现有 P2P 网络仿真器缺乏通用性和可扩展性，为完成仿真实验，在通用仿真器 NS2 基础上，设计了 P2P-VoD 系统的通用仿真平台。

NS2 并非专有的 P2P 模拟软件，首先在 NS2 经典模型基础上，构造 Socket 和接口模块使其支持 P2P 应用，设计 Peer 管理、数据调度、缓存管理等 P2P-VoD 应用模块完成仿真平台整体框架。仿真平台框架各个模块间的关系如图 1 所示。为了实现上述框架中模块功能，仿真平台在 NS2 原有系统类 Agent, Application, FullTcpAgent, TclObject 以及 Node 上进行了扩展，设计类实现各个模块功能。由于篇幅限制，这里不对各个模块具体实现做介绍。最后，将 P2P-VoD 系统主要模块实现类 PeerManager、DataScheduler 和 PeerApplication 利用 NS2 中的 TclCL 机制与 Otcl 对象关联起来为 Otcl 编程提供接口。

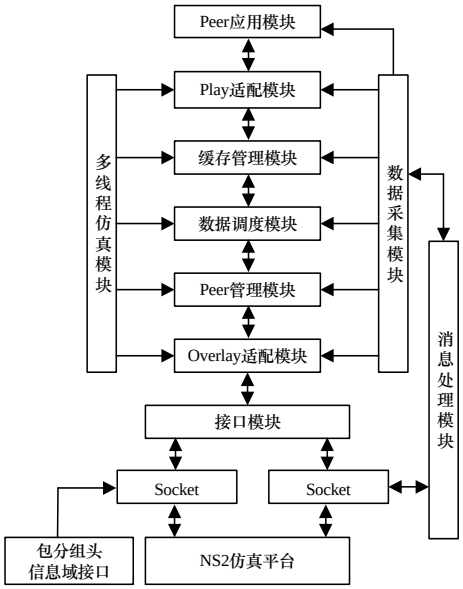


图 1　基于 NS2 的 P2P-VoD 系统仿真平台架构

对于基于网络编码的 P2P-VoD 数据调度算法，主要是通过在数据调度模块中添加随机线性编解码模块实现。模块架构如图 2 所示。

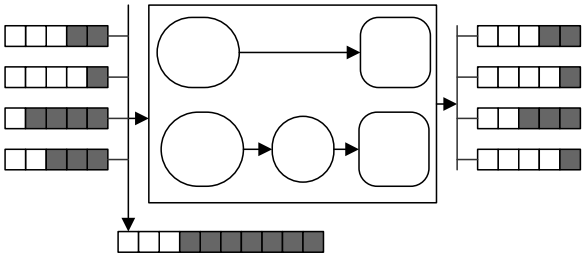


图 2　随机线性编解码模块

4 仿真实验及性能分析

4.1 实验设计

为验证设计的基于自适应网络编码的 P2P-VoD 数据调度算法 ARLNC 的有效性，设计仿真实验与基于网络编码的 P2P-VoD 系统 Redcarpet[3]和 DNC[4]中数据调度算法进行仿真比较。

仿真实验操作系统采用 Linux Redhat 9.0，在 Linux 操作系统中安装 NS2 仿真软件，将用 C++编写好的仿真框架类程序和数据调度算法 ARLNC, Redcarpet, DNC 的实现程序放到

/ns-allinone-2.33/ns-2.33 目录中，修改 Makefile 文件，对 NS2 进行重新编译成功后，再编写 TCL 脚本程序生成仿真场景和控制仿真过程，最后提取仿真数据进行分析。

实验中以 GT-ITM(Georgia Tech Internetwork Topology Models)生成一个具有 500 个节点的网络拓扑，并随机选取了 200 个节点作为仿真对象。设定节点带宽在一定范围内(2 Mb/s~8 Mb/s)随机分布，服务器带宽为 10 Mb/s。系统内消息传递 *TTL* 值为 5。节点动态缓存大小在 60 s~120 s 缓存时间之间随机分布。流媒体传输速率为 500 Kb/s，数据块大小为 256 KB。仿真实验中，分别选取媒体文件大小为 128 MB、256 MB、512 MB 和 1 024 MB 进行多次试验。

4.2 性能分析

仿真实验中选取网络性能指标节点播放连续度对仿真实验结果进行分析。

播放连续度用于评价节点的播放质量，它为 Peer 节点在一段时间内得到播放(即在其播放期限之前达到 Peer 节点)的数据块个数与应该播放的数据块个数的比值。播放连续度的值界于 0~1 之间，播放连续度的值越大，表示节点的播放质量越好。节点播放连续度比较如图 3 所示。

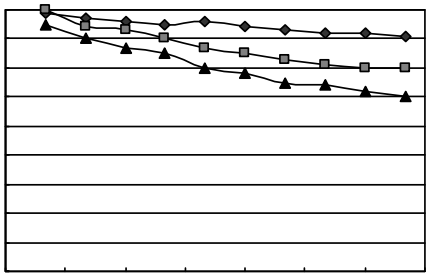


图 3 节点播放连续度比较

由图 3 可见，随着系统内仿真节点数目增多，3 种调度算法节点的播放连续度都呈下降趋势。ARLNC 中节点的播放连续度始终维持在 0.95 左右。各个节点动态自适应调整编码窗口大小，优化了节点间的协作，适应了特定的异构的网络拓扑环境，获得了高的播放连续度。另外，在网络编码窗口自适应计算中加入的加权因子，限制了网络编码窗口大小的过于分散，有助于 Peer 节点对视频块下载，提高了播放连续度。在 DNC 中，仅仅依靠将播放块的时限调整为 Peer 节点播放块的编码窗口大小，没有考虑网络异构性和节点异构性的情况，播放连续度维持在 0.85 左右。另外，编码窗口大小过于分散，给视频块的并行下载带来负面影响。在 Redcarpet 中，编码窗口固定，节点播放连续度大约维持在 0.8 左右。没有考虑 Peer 节点即将播放块时限，这样某些节点块的延迟很大，影响 Peer 节点播放性能。

5 结束语

本文提出了一种基于自适应随机网络编码的 P2P-VoD 数据调度算法，Peer 节点综合考虑自身缓存、链路可用带宽和播放视频块时限，动态调整网络编码窗口大小，自适应网络动态性和 Peer 节点异构性。仿真结果表明，应用该算法进行数据调度，Peer 节点视频播放的服务质量明显提高。

在以后工作中，将进一步优化网络编码窗口的自适应计算，使得自适应算法达到更加满意的效果。比如参数 β 选取问题，在算法中取值为 0.5，有待于通过大量实验寻找其最优化的具体表达式。另外，非法恶意节点加入系统后，通过分发虚假编码信息，可以致使系统中的节点解码错误，从而无法播放。所以，安全性问题需要进一步研究。

参考文献

[1] Wang Me, Li Baochun. Lava: A Reality Check of Network Coding in Peer-to-Peer Live Streaming[C]//Proc. of the 27th IEEE International Conference on Computer Communications. Anchorage, AK, USA: IEEE Press, 2007: 1082-1092.

[2] Nguyen K, Nguyen T, Cheung S. Peer-to-Peer Streaming with Hierarchical Network Coding[C]//Proc. of ICME'07. Beijing, China: [s. n.], 2007: 396-399.

[3] Annapureddy S, Guha S, Gkantsidis C. Exploring VoD in P2P Swarming Systems[C]//Proc. of the 26th IEEE International Conference on Computer Communications. [S. 1.]: IEEE Press, 2007: 2571-2575.

[4] Chi Huicheng, Zhang Qian. Deadline-aware Network Coding for Video on Demand Service over P2P Networks[J]. Journal of Zhejiang University: Science A, 2006, 7(5): 755-763.

编辑 索书志

（上接第 8 页）

4 结束语

本文提出基于 FARIMA 流量预测的抽样方法，实验结果表明，该方法能够较好地体现原始数据的流量行为统计特征，为流识别和异常流检测等应用提供了依据。由于网络行为的复杂性，在实际应用中出现了许多针对不同的应用特点设计出来的抽样方法，如 Choi B Y 等人提出的在存储内存限制下的自适应报文抽样方法，Duffield N G 等人提出的用来寻找报文流路径的基于数据报文内容的抽样技术。因此，下一步工作将更深入地研究对比各种抽样方法，验证方法的有效性。

参考文献

[1] Claffy K C, Polyzos G C, Braun H W. Application of Sampling Methodologies to Network Traffic Characterization[J]. ACM SIGCOMM Computer Communication Review, 1993, 23(4): 194-203.

[2] Anselm T, Jayasuriya P P. Efficient Admission Control Based on Predicted Traffic Characteristics[C]//Proc. of the 18th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications. Athens, Greece: [s. n.], 2007.

[3] 曾宪玮, 裴昌幸, 朱畅华, 等. 基于 XScale 和嵌入式 Linux 的网络测量仪[J]. 计算机工程, 2007, 33(16): 253-255.

[4] Harmantzis F, Hatzinako D. Heavy Network Traffic Modeling and Simulation Using Stable FARIMA Processes[C]//Proc. of the 19th International Teletraffic Congress. Beijing, China: [s. n.], 2005.

[5] Ilow J. Parameter Estimation in FARIMA Processes with Applications to Network Traffic Modeling[C]//Proc. of the 10th IEEE Workshop on Statistical Signal and Array Processing. Snowbird, USA: [s. n.], 2000.

编辑 索书志