

基于服务质量感知的 P2P 流媒体直播点播融合系统^①

冯侦探^{②*} 尤佳丽^{**} 倪 宏^{**} 王劲林^{**}

(* 中国科学院研究生院 北京 100049)

(** 中国科学院声学研究所国家网络新媒体工程技术研究中心 北京 100190)

摘 要 为了提高 P2P 流媒体系统性能,深入分析了 P2P 流媒体系统直播和点播业务的异同,提出采用动态时间坐标系(DTCS)构建直播点播融合的流媒体系统,将系统节点的协作特性与观看内容解耦。此外,充分考虑到节点的异构性与不同频道内容的流行度,提出了融合系统的节点协作模型,通过分布式服务质量感知算法(DQAA)来进行优化分析,充分利用 P2P 流媒体系统中服务能力较强的节点的剩余带宽提高系统性能。从融合系统架构及节点协作策略两个层次描述了直播点播融合的系统特性。实验表明,融合系统能有效提高节点播放质量与系统整体性能。

关键词 P2P 流媒体系统, 融合系统, 直播, 点播, 服务质量感知

0 引 言

在 P2P 流媒体系统中,不同的直播频道和点播影片流行度有很大区别^[1],而且由于网络延时、上行带宽、节点处理器能力、内存等因素,造成网络中的节点特性也有很大不同,这导致 P2P 流媒体系统中的节点资源分布不均。一般而言,流行度较高的频道拥有大量的节点,能提供稳定的 P2P 流媒体服务,节点资源较少的频道服务质量较低^[2]。目前,针对 P2P 流媒体系统的研究工作一般将直播、点播这两种流媒体业务区别对待,分别讨论在不同业务下的覆盖网(overlay)组织及数据调度方式。文献[3, 4]描述了直播多频道间的节点协作关系,但没有考虑如何与点播频道节点进行协作。针对此问题,本文主要从直播点播融合架构和节点协作策略两个层次来进行描述:(1)提出用一种基于动态时间坐标系(dynamic time coordinate system, DTCS)的方式来构建直播点播融合的 P2P 流媒体系统架构,将直播点播融合成统一的 P2P 流媒体系统;(2)从节点协作角度进行建模分析,并提出用分布式服务质量感知算法(distributed Qos-aware algorithm, DQAA)来选择协作节点,充分利用节点的剩余带宽为其他节点提供服务。通过以上两个层次的描述,

使得系统节点通过统一的覆盖网组织起来,不论直播业务还是点播业务,都通过感知服务质量来动态选择邻居节点进行协作,充分利用节点的带宽资源,提高系统性能。

1 相关工作

近年来,P2P 流媒体直播系统得到了广泛研究,数据驱动类型的网状直播系统如 CoolStreaming^[5], GridMedia^[6], PPLive^[7],代表着目前较为成熟的直播系统技术。该类系统的主要特征是通过复杂的数据调度保证服务质量,然而没有详细分析合作节点的服务能力不同对系统带来的影响。典型的点播系统 GridCast^[8]采用有向 Gossip 协议来交换缓存映像表,使处于相邻播放时间的节点进行协作,并采用多视频缓存技术(multiple video cache, MVC)来存储多个点播影片内容,促进节点协作。VoVo^[9]主要从用户行为的角度来分析用户跳转对系统带来的影响,节点通过与邻居节点交换播放跳转信息来预测下一阶段即将播放的内容,根据预测结果进行预取,提高跳转响应。InstantLeap^[10]根据影片时段,将节点划分为不同的群组,每个群组之间通过称为“捷径”(shortcut)的特殊边来进行连接,形成一个密集图

① “十一五”国家科技支撑计划(2008BAH28B04)和国家自然科学基金(60903218)资助项目。

② 男,1983 年生,博士生;研究方向:对等流媒体网络与宽带多媒体通信;联系人,E-mail: spyfeng@gmail.com
(收稿日期:2010-08-13)

(density graph), 因此节点跳转时能通过常数时间获得合作节点。此类点播系统对节点异构性考虑不足, 节点之间的协作通过缓存映像交互, 没有提出主动积极的策略促进系统中不同频道的节点之间的协作。

有些学者将 P2P 流媒体系统所有频道内的节点看做系统的可用资源并进行频道间协作研究。例如, 文献[3]提出的 VUD(view-upload decoupling) 将直播系统中不同频道间的节点带宽预先分配到不同的频道中, 将每个频道的直播数据流划分为不同的子流, 每个节点可以获得不同频道的不同子流, 节点获得某一频道的所有子流后再进行播放, 这样将节点实际观看的内容与服务内容解耦, 减少节点在切换频道带来的播放延时。VUD 在分配节点时需要全局信息以确定节点带宽如何分配, 计算较为复杂。文献[11]采用网络编码技术来解决频道间和频道内节点资源共享问题, 但文献[4]指出文献[11]由于采用网络编码因而有着较大的局限性, 因此提出了一种基于分治策略(divide and conquer, DAC)的算法, 该算法首先将系统划分为不同的频道集合, 每个频道维护一个覆盖网。由于频道间协作的存在, 覆盖网之间存在交叠, DAC 首先将交叠的覆盖网划分为独立的逻辑覆盖网(logic overlay), 之后再解决每个频道的节点协作优化问题, 多个覆盖网的维护导致在处理频道间协作时开销较大。

针对以上问题, 本文深入分析了直播点播的特点, 从架构上将直播点播进行融合, 系统节点资源皆为系统可用资源, 并提出了分布式服务质量感知算法(DQAA), 促进系统中的节点协作。

2 直播点播融合

2.1 理想直播系统与虚拟直播频道

定义 1 对直播系统频道, 设节点 X, Y 缓存的片段集合分别为 $B(X), B(Y)$, 若对任意时刻节点均有 $B(X) \cap B(Y) \neq \emptyset$, 则系统称为理想直播系统。

在理想直播系统中, 任意两个节点之间由于所需片段存在交集, 其中一个节点拥有该数据片段后能为另外一个节点服务, 因此, 理想直播系统中任意两个节点可以互相作为合作节点。在直播系统中, 节点不能进行跳转操作, 节点具有较强的同步特性, 若带宽足够, 随着播放的进行, 节点之间会一直保持合作关系。点播系统由于点播节点能进行跳转操

作, 节点拥有异步特性, 因而与理想直播系统不同, 节点跳转后需要查找新的合作节点。理想直播系统的主要特点是节点之间具有高度的协作特性。为便于描述如何对直播点播系统进行融合, 本文所提到的直播系统均认为是理想直播系统。

设 P2P 流媒体系统中有 M 个直播频道 $l_i (i = 1, 2, \dots, M)$, N 个点播频道 $c_i (i = 1, 2, \dots, N)$, 点播影片长度为 J_i , 考虑恒定码率情况下, 将其均匀分为大小为 T_i 的 K_i 个片段, 则 $K_i = J_i/T_i (i = 1, 2, \dots, N)$, 节点缓存大小为 B_s 。

定义 2 对于点播频道 $c_i (i = 1, 2, \dots, N)$, 若 $T_i < 2B_s$, 则对播放点属于 K_i 的任意节点 X, Y , 均有 $B(X) \cap B(Y) \neq \emptyset$, 即满足理想直播系统条件。考虑某一虚拟节点不断循环播放 T_i 片段拥有的内容 K_i , 由于处于该片段内的节点都能像理想直播系统一样相互协作, 称该虚拟播放点循环播放的片段为虚拟直播频道(virtual live channel, VLC)。

根据定义 2, 点播频道 c_i 由 K_i 个虚拟直播频道构成, 即 $c_i = \bigcup_{j=1}^{K_i} c_{ij}$ 。考虑到直播频道, 可统一描述为 $C = \{c_{ij}\}$, 其中 $(i = 1, 2, \dots, M, M+1, \dots, M+N)$, $(j = 1, 2, \dots, K_i), \forall i \in [1, M], j, j' \in [1, K_i]$ 有 $c_{ij} = c_{ij'}$, 即直播频道拥有 $\max\{K_i\}$ 个与自身完全相同的虚拟直播频道。

2.2 动态时间坐标系(DTCS)

有了虚拟直播频道的概念, 本文通过引入动态时间坐标系的概念, 结合虚拟直播频道的特性将不断变化的内容请求转化为相对静止的节点合作关系。

设系统有一全局基准时间为 T_b , 对每一点播频道 c_i 有一虚拟节点 V_i 从基准时间 T_b 开始播放, 当前时间为 T_c , 则可以计算出虚拟节点播放的位置:

$$O_i = (T_c - T_b) \bmod J_i \quad (1)$$

O_i 称为频道 c_i 的动态时间坐标系原点。若已知用户实际播放点为 P , 则可以求出动态时间坐标系下的虚拟直播频道号:

$$j = [(O_j - P + J_i)/T_i] \bmod K_i \quad (2)$$

根据定义 2, 虚拟直播频道可看做理想直播系统, 因此, 无论直播业务还是点播业务, 节点加入 P2P 系统即是寻找特定虚拟直播频道的合作节点, 根据式(2)可以计算出要加入的虚拟直播频道号。因此, 通过虚拟直播频道, 将点播业务中的跳转操作转化为虚拟频道之间的切换, 将节点不断获得不同的内容的行为转化为相对静态的合作节点关系, 大

大降低了不断请求动态数据内容带来的查找合作节点的开销。此外,通过本节分析可见,融合的流媒体系统由虚拟直播频道构成,因而可以充分利用直播系统的相关数据调度技术来实现点播系统,提高系统维护性。

2.3 融合系统覆盖网组织

有了虚拟直播频道的概念以后,关键问题转化为如何将节点组织起来。本文采用分布式哈希表(distributed Hash table, DHT)的方式进行组织,具体实现方式为改进后的 Kademlia^[12]协议。每个节点加入系统后发布信息 $\langle key, value \rangle$, 其中 $key = Hash(c_{ij})$ 表示节点的内容信息; $value = IP$ 表示哪些节点拥有内容 c_{ij} , 节点切换直播频道和点播中的跳转操作对应为在不同的频道 c_{ij} 之间切换。若平均每个点播频道有 $\bar{K}$ 个虚拟直播频道,则系统总频道数 $|C| = M + N\bar{K}$, 节点查找效率为 $O(\log^{(M+N\bar{K})})$ 。为加快用户的跳转操作响应,处于同一点播频道中的节点会通过 DHT 网络定期主动获得本频道其他虚拟直播频道的节点信息,这样节点进行跳转操作时能通过近似 $O(1)$ 的时间复杂度获得目标虚拟直播频道的邻居节点信息。

考虑到节点的异构性,为便于在寻找合作节点时找到服务能力较强的节点,本文在 DHT 网络中添加超级频道来存储服务能力较强的节点,节点动态统计服务能力,按照服务能力级别加入超级频道中,超级频道中的节点可以为任意频道内的节点服务。

如图 1 所示,4 个虚拟直播频道 VLC1-VLC4 的节点存储在 DHT 环上,每个虚拟直播频道内的节点直接相互协作,如 VLC1 中的节点 A、B 和 C。虚拟频道之间的切换需要通过 DHT 网络重新获得合作

节点。节点 C、H 和 F 为服务能力较强的节点,按照下节描述的策略加入超级频道,可以为其他频道内的节点服务。

3 融合系统节点协作

上节描述了直播点播融合的 P2P 流媒体系统通过 DHT 的方式进行组织,形成了统一的覆盖网结构,便于查找合作节点。下面从如何协作的角度来更深层次地分析讨论节点之间的合作特性。

3.1 模型分析

设 P2P 流媒体网络构成一个无向图 $G = (V, E)$, 其中每一个节点 $v_i \in V$ 表示网络中的节点,图 G 中的边 $e_{ij} = (v_i, v_j) \in E$ 表示 v_i 和 v_j 属于邻居关系, $e_{ij} = e_{ji}$ 。系统节点分布在不同的频道, $V(c_{ij})$ 表示频道 c_{ij} 的节点集合, $n_{ij} = |V(c_{ij})|$ 表示该频道节点数目,播放码率为 r_{ij} 。系统中所有节点数目为 n 。在本文中假设网络传输能力无限,节点拥有足够的下行带宽获得所需数据,但拥有有限上行带宽为其他节点服务。

定义 3 设节点 p 固有上行带宽为 U_p , 对频道 c_{ij} 的使用带宽为 u_p^{ij} , 则节点 p 对频道 c_{ij} 的带宽使用系数 $\mu_p^{ij} = \frac{u_p^{ij}}{U_p}$, 节点 p 的上行带宽利用率为 $\mu_p = \frac{\sum_{ij} u_p^{ij}}{U_p}$, 其中 $\mu_p \in [0, 1]$ 。

定义 4 频道 c_{ij} 播放质量为 $\varphi_{ij} = \frac{\sum_{p \in V} U_p \cdot \mu_p^{ij}}{n_{ij} \cdot r_{ij}}$ 。

$\varphi_{ij} = 1$ 时,频道 c_{ij} 中的节点理论上刚好满足带宽需求,仅须向数据源服务器获得一路数据,通过有效的数据调度策略便能达到流畅播放; $\varphi_{ij} < 1$ 时,表明系统中没有足够的节点为频道 c_{ij} 节点服务,需要向服务器获得额外数据。为充分利用整个系统中的节点上行带宽,以降低服务器负载为目标,可以建立如下模型:

$$\max_{\mu_p^{ij}} \min_{ij} \varphi_{ij} = \frac{\sum_{p \in V} U_p \cdot \mu_p^{ij}}{n_{ij} \cdot r_{ij}} \quad (3)$$

s. t.

$$(i) \mu_p = \sum_{i=M+N, j=K} \mu_p^{ij} \leq 1.$$

(ii) $\forall p \notin V(c_{ij})$, 满足 $U_p \cdot (1 - \mu_p) > 2\alpha \cdot r_{ij}$, 其中 $\alpha \in (0, 1]$ 。

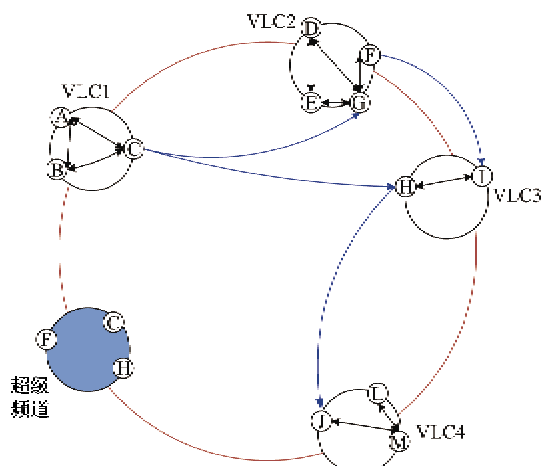


图1 融合系统覆盖网示意图

(iii) $\forall \varphi_{(ij)}' > 1$, 在优化过程中满足 $\varphi_{(ij)}' > 1$ 。

该式的含义是指,充分利用节点的带宽资源,最大化播放质量最低的虚拟直播频道的下界,提高系统的整体性能。条件(i)是节点 p 的上行带宽限制。条件(ii)说明对于 $p \notin V(c_{ij})$, 节点 p 没有频道 c_{ij} 的内容,因此节点 p 需要首先加入频道 c_{ij} 获取所需数据,然后才能为其他节点服务,设节点 p 获取数据消耗的带宽为 $\alpha \cdot r_{ij}$, 其上行带宽至少要大于 $2\alpha \cdot r_{ij}$ 。条件(iii)表明在优化过程中要保证原有服务频道的服务质量,避免出现盲目优化带来的系统波动。式(3)所描述的优化问题,是时间相关的,该优化过程随着节点的加入和离开,需要不断进行调整。文献[13]指出,该问题属于 NP-hard 问题,本文采用启发式算法来进行优化。

协作模型数学符号含义见表 1。

表 1 协作模型数学符号含义表

符号	含义描述
$V(c_{ij})$	频道 c_{ij} 节点集合
r_{ij}	频道 c_{ij} 码率
n_{ij}	频道 c_{ij} 节点数目
U_p	节点 p 上行带宽
u_p^{ij}	节点 p 对频道 c_{ij} 上行带宽
μ_p	节点 p 上行带宽使用率
φ_{ij}	频道 c_{ij} 播放质量
$\bar{r}_{ij}$	频道 c_{ij} 平均播放率
$\bar{R}_p^{ij}$	节点 p 在频道 c_{ij} 的平均播放速率
S_{ij}	频道 c_{ij} 抽样节点集合
$\gamma_p(t)$	节点 p 上行带宽利用率
$\theta_p(t)$	节点 p 冗余能力系数

3.2 分布式服务质量感知算法 (DQAA)

根据上节的模型描述,本节通过分布式服务质量感知算法来对模型进行优化。服务质量是指节点对 P2P 系统所提供的服务的满意度,这里采用定义 4 所描述的虚拟直播频道的播放质量 φ_{ij} 来进行定量描述, φ_{ij} 越大,服务质量越好。然而,单个节点很难获得频道的播放质量,本文基于 Metropolis-Hastings^[14] 算法构造无偏随机行走对节点进行抽样,根据抽样节点来估计服务质量,进而引出如何进行优化。本节首先对节点的服务能力进行量化描述,给出了节点冗余能力系数的定义,用于度量节点的服务能力,接着描述算法的详细步骤。

3.2.1 服务质量感知

设节点 p 获得频道 c_{ij} 数据的带宽为 r_p^{ij} , 则有

$$\sum_{p \in V} U_p \cdot \mu_p^{ij} = \sum_{p \in V(c_{ij})} r_p^{ij} \quad (4)$$

根据式(3),(4),有

$$\begin{aligned} \varphi_{ij} &= \frac{\sum_{p \in V} U_p \cdot \mu_p^{ij}}{n_{ij} \cdot r_{ij}} = \frac{\sum_{p \in V(c_{ij})} r_p^{ij}}{n_{ij} \cdot r_{ij}} \\ &= \frac{\sum_{p \in V(c_{ij})} r_p^{ij} / n_{ij}}{r_{ij}} = \frac{\bar{r}_{ij}}{r_{ij}} \end{aligned} \quad (5)$$

$\bar{r}_{ij}$ 表示频道 c_{ij} 中节点平均播放速率,节点 p 按照周

期 T 统计平均播放速率 $\bar{r}_p^{ij} = \frac{\int_t^{t+T} r_p^{ij}(t) dt}{T}$, 为提高精

确度,将节点 p 的邻居节点平均播放速率也作为参考。设 $q \in NBR(p)$, 节点 p 的平均播放速率为:

$$\bar{R}_p^{ij} = \frac{\sum_{q \in NBR(p)} \bar{r}_q^{ij} + \bar{r}_p^{ij}}{|NBR(p)| + 1} \quad (6)$$

若节点 p 为抽样点,设频道 c_{ij} 的抽样点集合为

S_{ij} , 抽样点的平均播放速率 $(\bar{r}_{ij})' = \frac{\sum_{p \in S_{ij}} \bar{R}_p^{ij}}{|S_{ij}|}$, 进

而有

$$\varphi_{ij} = \frac{\bar{r}_{ij}}{r_{ij}} \approx \varphi'_{ij} = \frac{(\bar{r}_{ij})'}{r_{ij}} \quad (7)$$

本文用抽样节点计算所得值 φ'_{ij} 来估算实际值 φ_{ij} 。文献[15]介绍了多种随机行走方式。设节点 p, q 是网络中的节点,表达式

$$P(p, q) = \begin{cases} \frac{1}{\text{degree}(p)} & q \in NBR(p) \\ 0 & \text{other} \end{cases} \quad (8)$$

为随机行走的转移矩阵。文献[15]指出,若 $\vec{p}_0$ 表示节点 p_0 的当前转移概率矢量,经过 k 步随机行走后对 $x \in V$ 有

$$\pi(x) = \lim_{k \rightarrow \infty} (\vec{p}_0 \cdot P^k)(x) = \frac{\text{degree}(x)}{2|E|} \quad (9)$$

此式表明仅以邻居节点的度作为转移概率矩阵会导致随机行走倾向于停留在度较大的节点,即抽样节点为系统中邻居节点数目较多的节点,因此不能正确反映出系统的实际特性。文献[14]提供了一种已知随机行走的静止概率分布函数构造转移概率矩阵的算法,将节点之间的转移关系看做马尔科夫链蒙特卡罗 (MCMC) 过程^[16],若静止概率分布为 $\pi(x)$, 根据 Metropolis-Hastings 可得转移概率矩阵:

$$Q(p, q) = \begin{cases} P(p, q) \min\left(\frac{\pi(q)P(q, p)}{\pi(p)P(p, q)}, 1\right), & p \neq q \\ 1 - \sum_{p \neq q} Q(p, q), & \text{其他} \end{cases} \quad (10)$$

为保证抽样的公平性,令 $\pi(p) = \pi(q)$, 即当 $p \neq q$ 时构造如下转移概率矩阵:

$$\begin{aligned} Q(p, q) &= P(p, q) \min\left(\frac{P(q, p)}{P(p, q)}, 1\right) \\ &= \frac{1}{\text{Degree}(p)} \cdot \min\left(\frac{\text{Degree}(p)}{\text{Degree}(q)}, 1\right) \end{aligned} \quad (11)$$

采用式(10)、(11)提供的转移概率矩阵能均匀地获得某频道的抽样节点,更准确地反映实际情况。

3.2.2 算法描述

为便于描述算法,本节首先给出富余能力系数定义。根据定义3,节点在 t 时刻的上行带宽利用率为 $\mu_p(t)$, 节点 p 以周期 T 统计平均上行带宽利用

率,在 $t + T$ 时刻 $\bar{\mu}_p(t + T) = \frac{\int_t^{t+T} \mu_p(t) dt}{T}$ 。在统计节

点上行带宽使用率时,最近的统计值越能代表节点当前的带宽使用情况,为此,对最新时刻统计数据设置较大的权重。设最近的 η 次统计权重分别为 $w_i = \chi^i, \chi \in (1, 2], i = 1, 2, \dots, \eta$, 则节点在 t 时刻的平均上行带宽利用率为 $\gamma_p(t) = \frac{\sum_i \bar{\mu}_p(t - iT) \cdot w_i}{\sum_i w_i}$, 进而定义富余能力系数:

$$\theta_p(t) = U_p \cdot \frac{1 - \gamma_p(t)}{\gamma_p(t) + \varepsilon} \quad (12)$$

根据式(12), $\theta_{ij} = 0$ 表示频道内 c_{ij} 的节点上行带宽已经用尽,不能为其他节点提供服务。富余能力系数越高,表明该频道内的节点越能提供其他节点服务。 $\varepsilon \in (0, 1]$ 为扰动因子。

分布式服务质量感知算法基于节点的无偏随机行走选取抽样节点来感知频道服务质量,同时节点动态统计富余能力系数,通过调配拥有剩余带宽的节点至服务质量较低的频道来提高系统整体性能。详细描述如下:

(1) 节点 p 加入系统后按照周期 T 产生 β 个步长为 TTL 的随机行走报文,按照式(10)计算下一步的转移概率,同时按照式(12)来计算节点富余能力系数。若节点 p 已经存在超级频道中,则更新节点富余能力系数,否则,若大于一定门限 τ , 则节点加入超级频道。富余能力系数的更新通过网络中的消息报文来传递,并进行定时维护更新。

(2) 设随机报文 TTL 减为 0 时停留的节点为 z , 则节点 z 为抽样点,根据式(6)计算平均播放速率,节点 z 将该值通过报文消息反馈给节点 p 。

(3) 节点 p 收到 h 个抽样点结果时根据式(7)估算频道 c_{ij} 服务质量,若低于一定门限 σ , 则节点 p 会选择超级频道中的节点作为合作节点来提高频道 c_{ij} 的服务质量。

(4) 若超级频道中的节点 f 收到节点 p 请求协作请求报文时,根据式(3)的条件(ii)判断是否能加入到请求节点所在频道。

(5) 超级频道中的节点 f 会根据式(7)中服务质量计算值来判断式(3)中的条件(iii),确保节点 f 所在频道的服务质量。

该算法是针对式(3)的一种启发式算法,通过节点局部优化来达到系统整体性能的提升,不需要全局信息,节点在运行该算法仅需要邻居节点信息,且随着时间的推移而不断调整。

4 仿真实验与性能分析

本文基于仿真器 p2pstrmsim^[6] 进行了二次开发,加入了多频道的支持和覆盖网的构建模块。p2pstrmsim 是一个基于离散事件的 P2P 直播系统仿真器,主要仿真数据驱动类型的直播系统。仿真底层拓扑与延时参数设置采用 Meridian^[17] 提供的实际网络拓扑数据,每次仿真节点数据为 1000 个节点,节点泊松过程加入,在线时长根据 GridMedia^[6] 系统的 Trace 文件,按照上行带宽将节点分为 3 类,如表 2 所示。

表 2 节点类型参数

类型	上行带宽	下行带宽	节点比例
1	1Mbps	3Mbps	0.1
2	384kbps	1.5Mbps	0.4
3	128kbps	768kbps	0.5

根据本文第 3 节所述,点播系统能转化为直播问题来处理,因此仿真中不同频道为直播频道,其中频道流行度遵循 Zipf 分布^[1],仿真频道数目为 10,每个频道仅有一个源节点,其上行带宽为 600kbps,频道播放码率均为 300kbps,采用 Pull-Random 策略作为数据调度算法。节点加入频道后先进行缓冲,当缓冲区达到 1500 个连续数据报文时,开始运行分布式服务质量感知算法(DAQQ),随机行走的步长 $TTL = 10$, $\beta = 3$,每隔 10s 会对系统进行一次统计分析。本文与传统的无频道间协作(Trad)及 VUD^[3] 中的多频道协作策略进行对比分析。

图 2 描述了各频道的平均播放质量。图中左侧频道为流行度较低的频道,频道中节点数目较少,无频道间协作策略情况下,平均播放质量约为 0.6。采用 DQAA 后频道 1-5 的平均播放质量提高了 25% 左右,这主要是因为流行度较高的频道中存在冗余能力系数较高的节点, DQAA 充分利用节点的剩余带宽,促进了频道之间的节点协作。VUD 与 DQAA 出发点一致,都是充分利用节点的服务能力,最大化利用流行度较高的频道中的节点上行带宽,但 VUD 在选择合作节点时按照节点的绝对上行带宽实现进行预分配,从流行度较高的频道中随机选择节点作为协作节点,该节点虽处于流行度较高的频道,节点自身的上行带宽可能剩余较少,无法及时为节点提供数据,造成重复请求。DQAA 感知实际播放质量,从超级频道中有目的地选择冗余能力系数较高的节点作为协作节点,同时动态地更新节点的冗余能力系数,避免了选择到虽处于流行度较高的频道,但上行带宽无法提供服务的节点,提高了协作的可靠性。

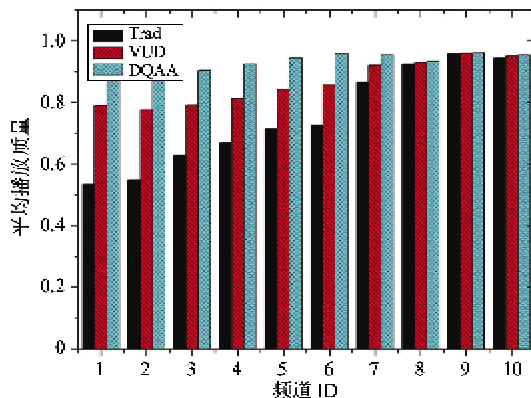


图 2 平均播放质量

以频道 3 为例,图 3 描述了在不同仿真时刻的平均播放质量。在最开始统计时由于节点还处于不断加入过程,且上行带宽较大的节点首先加入,能及时获得所需数据,因此平均播放质量在 0.85 以上,之后大量上行带宽较低的节点加入造成平均播放质量有所下降。从图 3 中还可以看出采用 VUD 策略在进行节点优化平均播放质量波动较大,这也正好说明了 VUD 在选择节点时具有一定的盲目性,造成数据报文丢失和节点重新选择,而 DQAA 能更好地促进节点的协作。

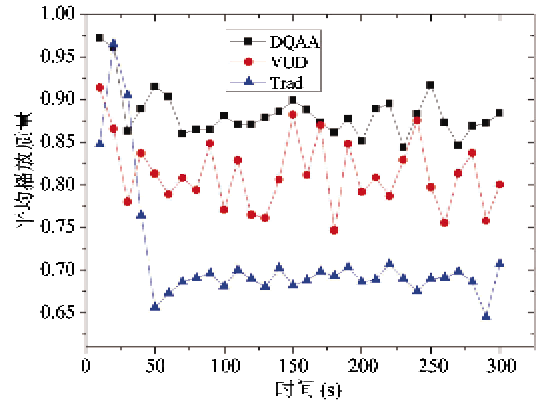


图 3 平均播放质量

图 4 和图 5 是频道 3 中节点的平均报文延时和平均跳数。P2P 直播系统中,平均报文延时是一项重要参数,含义是指数据报文从服务器发出时刻,到系统中的节点收到该报文之间的平均延时,延时越小说明系统中节点能更快地获得最新的数据。平均报文跳数与平均报文延时是互洽的一组数据,是指数据报文经过几次转发最终达到目的节点。在链路延时一定的情况下,一般跳数越少延时越小。仿真环境中,节点之间的平均延时约为 150ms,根据图 5

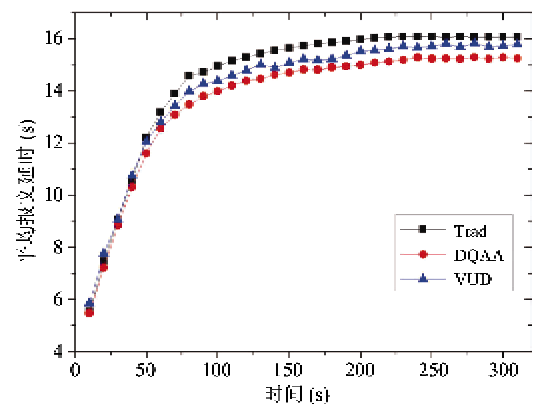


图 4 平均报文延时

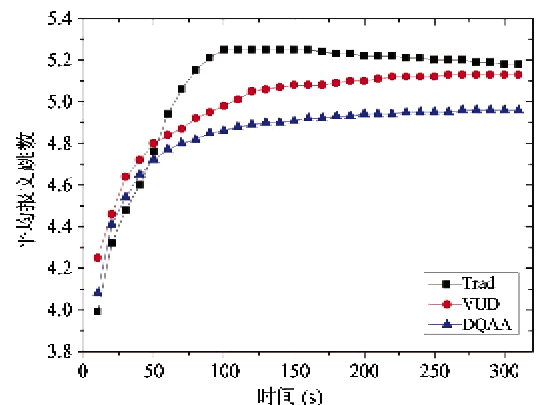


图 5 平均报文跳数

所示, DQAA 在 200s 后的稳定状态平均报文跳数比无频道间协作小 0.4 跳, 因此对应图 4 平均报文延时略小 0.6s 左右。流行度较低的频道有其他频道协作节点的加入, 本质可以看做是合作节点的增加。新增的节点服务能力越强, 能服务的节点越多, 数据分发树的高度越扁平, 报文延时和跳数也会降低。由于冗余能力系数较高的节点的加入, 在节点缓冲大小一定的情况下增加了该频道的上行带宽, 冗余能力系数较高的节点能主动服务较多的其他节点, 因此在一定程度上也降低了报文延时和平均跳数。

图 6 是系统节点冗余能力系数累计函数分布图, 横坐标采用对数刻度。根据 3.2.2 节描述, 冗余能力系数越接近于 0, 表明节点的上行带宽使用率越高。设定仿真中 $\varepsilon = 0.1$, 定期对系统的冗余能力系数进行统计, 根据图 6 所示, 在不同时刻 (100s, 200s, 300s) DQAA 约有 77% 的节点冗余能力系数趋近于 0, 表示节点的上行带宽得到充分利用, 而传统策略约为 40%。冗余能力系数 128 以下的节点数目 DQAA 约占总节点数目的 93%, 而传统策略下所占比例约为 50%。因此, 从冗余能力系数的分布图也能看出 DQAA 在促进节点协作和充分利用冗余节点能力方面有着重要作用。

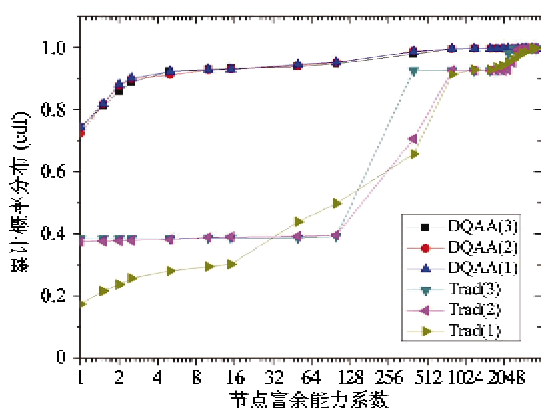


图6 节点冗余能力系数

图 7 反映了 DQAA 对服务质量的感知的准确度, 根据式(7)计算误差值 $\Delta = \varphi_{ij} - \varphi'_{ij}$, 随机行走抽样节点集合 $|S_{ij}| = 15$ 。对系统节点每隔 100s 进行 3 次统计分析, 根据图 7 显示通过无偏随机行走获得采样节点的平均播放质量 φ'_{ij} 与该频道实际平均播放质量 φ_{ij} 的误差值在 0.15 以内, 说明 DQAA 能较为准确地感知频道服务质量。

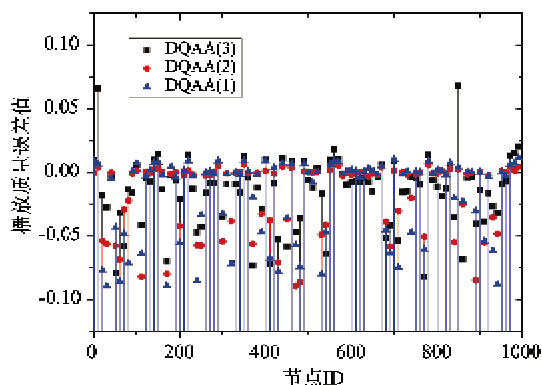


图7 频道服务质量感知误差

5 结论

针对 P2P 流媒体系统的主流业务直播与点播, 本文提出了一种融合架构来进行统一描述, 将点播系统问题转化直播系统问题进行描述, 采用 DHT 的方式将系统所有节点组织起来, 并探讨了如何促进不同频道间的节点协作, 充分利用节点剩余带宽, 提高系统性能的分布式算法。仿真实验表明该算法能有效提高系统性能。

未来研究方向: 融合系统中如何充分利用频道间的相关性加速合作节点的查找; 通过数据调度来分析节点之间的相对服务能力, 分析合作节点之间的相似特性, 研究将服务能力相近的节点聚类的算法。

参考文献

- [1] Cheng H, Jin L, Keith W R. Can internet video-on-demand be profitable? *SIGCOMM Comput Commun Rev*, 2007, 37(4): 133-144
- [2] Xiaojun H, Yong L, Ross K W. Inferring network-wide quality in P2P live streaming systems. *Selected Areas in Communications, IEEE Journal on*, 2007, 25(9): 1640-1654
- [3] Di W, Chao L, Yong L, et al. View-upload decoupling: A redesign of multi-channel P2P video systems. In: *Proceedings of the IEEE International Conference on Computer Communications*, Rio, Brazil, 2009. 2726-2730
- [4] Miao W, Lisong X, Ramamurthy B. A flexible divide-and-conquer protocol for multi-view peer-to-peer live streaming. In: *Proceedings of the IEEE 9th International Conference on Peer-to-Peer Computing*, Seattle, USA, 2009. 291-300
- [5] Susu X, Bo L, Keung G Y, et al. Coolstreaming: Design, theory, and practice. *Multimedia, IEEE Transac-*

- tions on, 2007, 9(8): 1661-1671
- [6] Meng Z, Qian Z, Lifeng S, et al. Understanding the power of pull-based streaming protocol: Can we do better? *Selected Areas in Communications, IEEE Journal on*, 2007, 25(9): 1678-1694
 - [7] Long V, Indranil G, Jin L, et al. Measurement and modeling of a large-scale overlay for multimedia streaming, In: *Proceedings of the 4th International Conference on Heterogeneous Networking for Quality, Reliability, Security and Robustness*, Vancouver, Canada, 2007. 1-7
 - [8] Bin C, Lex S, Hai J, et al. GridCast: Improving peer sharing for P2P VoD. *ACM Trans Multimedia Comput Commun Appl* 1551-6857, 2008, 4(4): 1-31
 - [9] Yuan H, Student M, Yunhao L, et al. VOVO: VCR-oriented video-on-demand in large-scale peer-to-peer networks. 2009
 - [10] Xuanjia Q, Chuan W, Xiaola L, et al. InstantLeap: fast neighbor discovery in P2P VoD streaming, In: *Proceedings of the 18th International Workshop on Network and Operating Systems Support for Digital Audio and Video*, Williamsburg, USA, 2009. 19-24
 - [11] Chuan W, Baochun L, Zongpeng L. Dynamic bandwidth auctions in multioverlay P2P streaming with network coding. *Parallel and Distributed Systems, IEEE Transactions on*, 2008, 19(6): 806-820
 - [12] Maymounkov P, Mazières D. Kademlia: A peer-to-peer information system based on the XOR metric. In: *Proceedings of the 33rd Peer-To-Peer Systems: First International Workshop*, Cambridge, USA, 2002. 53-65
 - [13] Chakrabarty D, Mehta A, Vazirani V. Design is as easy as optimization, In: *Proceedings of the 33rd International Colloquium on Automata, Languages and Programming*, Venice Italy, 2006. 477-488
 - [14] Chih S, Greenberg E. Understanding the Metropolis-Hastings algorithm. *The American Statistician*, 1995, 49(4): 327-335
 - [15] Lovász L. Random walks on graphs: A survey. *Combinatorics: Paul Erdos is Eighty*, 1993, 2(1): 1-46
 - [16] Hastings W K. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 1970, 57: 97-109
 - [17] Bernard W, Aleksanders S, Emin G, et al. Meridian: a lightweight network location service without virtual coordinates. In: *Proceedings of the Conference on Applications, Technologies, Architectures, and Protocols for Computer Communications*, Philadelphia, USA, 2005. 85-96

A P2P media streaming system integrated with live and VoD based on QoS-aware

Feng Zhentan^{*,**}, You Jiali^{**}, Ni Hong^{**}, Wang Jinlin^{**}

(^{*} Graduate University of Chinese Academy of Sciences, Beijing 100049)

(^{**} National Network New Media Engineering Research Center,
Institute of Acoustics, Chinese Academy of Sciences, Beijing 100190)

Abstract

To improve the performance of a P2P streaming media system, the paper deeply analyzes the similarities and differences between a P2P live system and a P2P VoD (Video-on-Demand) system, and based on this, presents a method to build a P2P media streaming system integrated with live and VoD based on the Dynamic Time Coordinate System (DTCS). The integrated system decouples the collaboration feature with the particular content watched by users. In addition, taking consideration of the heterogeneity among peers and the popularity of different channels, the paper proposes a peer collaboration model for the integrated system. Based on this model, the paper addresses the Distributed QoS-Aware Algorithm (DQAA) to optimize the system, which makes full use of the peers' surplus bandwidth to improve system performance. The paper describes the features of the integrated system from two aspects: overlay construction and cooperation strategy. The simulation results demonstrate that the integrated system can significantly improve the playback quality of peers and system performance.

Key words: P2P media streaming system, integrated system, live, VoD, QoS-Aware