

学校编码: 10384

分类号_____密级_____

学号: X200343045

UDC _____

廈門大學

硕士学位论文

基于 P2P 的流媒体分发关键技术研究

Research on Key Technology of Streaming Media
Dissemination Based on P2P Network

王爽

指导教师姓名: 李绍滋教授

专业名称: 计算机应用

论文提交日期: 2006 年 12 月

论文答辩时间: 2007 年 月

学位授予日期: 2007 年 月

答辩委员会主席: _____

评阅人: _____

2007 年 月

厦门大学学位论文原创性声明

兹提交的学位论文，是本人在导师指导下独立完成的研究成果。本人在论文写作中参考的其他个人或集体的研究成果，均在文中以明确方式标明。本人依法享有和承担由此论文产生的权利和责任。

声明人（签名）：

年 月 日

厦门大学学位论文著作权使用声明

本人完全了解厦门大学有关保留、使用学位论文的规定。厦门大学有权保留并向国家主管部门或其指定机构送交论文的纸质版和电子版，有权将学位论文用于非赢利目的的少量复制并允许论文进入学校图书馆被查阅，有权将学位论文的内容编入有关数据库进行检索，有权将学位论文的标题和摘要汇编出版。保密的学位论文在解密后适用本规定。

本学位论文属于

- 1、保密（ ），在 年解密后适用本授权书。
- 2、不保密（ ）。

（请在以上相应括号内打“√”）

作者签名：

日期： 年 月 日

导师签名：

日期： 年 月 日

摘 要

随着网络宽带的普及，流媒体服务也变得日渐流行。目前很大一部分的 Internet 通信量是传送流媒体数据。许多流媒体应用，如新闻直播、视频点播、网络电视等都是将流媒体数据从一个源端分发到大量的客户端。

P2P 流媒体模型与传统的流媒体模型有极大的不同：传统的基于客户机/服务器的结构模式中，所有的管理集中于服务器一方来完成；P2P 模式中，终端用户既可作客户机也可作服务器，加入的节点以对等的方式分享资源、平衡负载。

当前关于 P2P 流媒体分发的研究主要集中于应用层。许多覆盖网架构的算法也是采用建造一棵树——组播树的方法，数据就在树的节点间传递。然而组播树的构建较复杂、扩展性不强、对动态变化的应用层覆盖网适应性较差。

为了解决上述问题，提出了 gossip 算法。在 gossip 算法中，每个对等节点每轮和若干个随机选择的伙伴交换信息。gossip 算法不仅具有内在的扩展性，同时还具有易配置、鲁棒性强、从失效中恢复力强等特点，更适用于 P2P 这样连接可靠性较差的网络。

本文主要研究基于 P2P 的流媒体分发。首先概括介绍了流媒体技术和 P2P 技术，接着提出了源驱动和数据驱动的概念并讨论了几种典型的组播树模型。最后在 gossip 算法的基础上，设计了一个数据驱动的 P2P 流媒体分发模型——DGM，在 DGM 中，采用幂律和小世界理论组织 P2P 覆盖网，用局部视图管理节点。仿真实验结果表明：DGM 有效的成员管理、伙伴管理和调度策略实现了流媒体数据实时、持续的分发，在用户动态变化的环境中表现较优。

关键词： 流媒体分发；对等网络；Gossip 算法

Abstract

With the widespread penetration of broadband accesses, streaming media services are getting increasingly popular among users and have contributed to a significant amount of present Internet traffic. Many streaming media applications, such as news live broadcast, video on-demand and IPTV are all techniques which distribute streaming media data from a source end to a large population of user ends.

The P2P streaming media model differs greatly from traditional one. Traditional model relies on the client-server architecture, thus all the management activities are handled centrally on the server side. However, in P2P model, the end-user act as both a client and a server, which means each participating peer can share resources and balance the loads in a peer style.

Currently, most researches on streaming media dissemination over P2P network have been focused on application-level. Many overlay construction algorithms also resort to building a tree structure, namely multicast tree, and the data relaying among these nodes of the tree. However, the multicast trees are relatively complex and hard to scale. Besides, their match abilities with the dynamic application level overlay are poor.

To deal with the problems mentioned above, the gossip-based algorithms are proposed. In gossip-based algorithms, each peer exchanges information with randomly-chosen communication partners in each round. In addition to their inherent scalability, gossip-based algorithms are easy to deploy, robust, and resilient to failure, thus making them ideally suited for P2P networks with potentially poor link-reliability.

This thesis emphasizes on streaming media dissemination based on P2P network. Firstly, the background knowledge of streaming media technology and P2P technology were briefly introduced. Then, the concepts of source-driven and data-driven were proposed and some representative models of tree-based were discussed respectively. Furthermore, based on the theory of gossip algorithm, a data-driven streaming media dissemination model (DGM) over P2P overlay network was designed. We adopted the ideas of power law and small world in DGM to organize the P2P overlay network and managed the peers in a partial view. Simulation results showed that in DGM, the efficient membership and partnership management scheme, together with scheduling scheme that achieved real-time and continuous distribution of streaming media data which performed much better in dynamic user environments.

Keywords: Streaming media dissemination; Peer-to-peer(P2P); Gossip-based algorithm

目 录

第一章 绪论	1
1.1 论文的研究背景	1
1.1.1 P2P 流媒体技术的由来	1
1.1.2 P2P 流媒体技术发展历程及现状	2
1.1.3 P2P 流媒体研究和应用面临的挑战	3
1.2 论文的选题依据	4
1.3 论文的主要内容和组织	7
第二章 流媒体技术和 P2P 技术	9
2.1 流媒体技术	9
2.1.1 流媒体的概念和系统组成.....	9
2.1.2 流媒体的传输原理	9
2.1.3 流媒体的数据压缩技术	10
2.1.4 支持流式传输的网络协议.....	11
2.1.5 常用的流媒体文件格式	12
2.1.6 流媒体的传输方式	13
2.1.7 流媒体技术的应用	14
2.2 Peer to peer 技术	14
2.2.1 P2P 的定义和 P2P 覆盖网	14
2.2.2 P2P 的特点.....	15
2.2.3 P2P 的拓扑结构分类.....	16
2.2.4 P2P 技术面临的问题和挑战	17
2.2.5 P2P 的应用.....	19
2.3 本章小结	19
第三章 源驱动的组播树模型	21
3.1 源驱动和数据驱动的概念	21
3.2 典型的源驱动的组播树模型	21
3.2.1 PeerCast.....	21
3.2.2 NICE.....	23
3.2.3 Zigzag.....	24

3.2.4	SplitStream	29
3.2.5	CoopNet	32
3.3	本章小结	32
第四章	数据驱动的 Gossip 算法模型	34
4.1	相关概念	34
4.1.1	幂律	34
4.1.2	小世界现象	34
4.1.3	Gossip 算法	36
4.2	DGM 系统模型	38
4.2.1	覆盖网架构	38
4.2.2	系统结构	39
4.2.3	节点的加入与管理	40
4.2.4	伙伴的管理和优化	42
4.2.5	缓冲映射和数据交换	42
4.2.6	调度管理和算法	45
4.3	本章小结	46
第五章	实验	47
5.1	网络仿真器 NS-2	47
5.1.1	NS-2 介绍	47
5.1.2	NS-2 网络仿真实例	48
5.2	实验方案和参数设计	50
5.3	实验结果与分析	50
5.3.1	静态环境下的实验结果与分析	50
5.3.2	动态环境下的实验结果与分析	52
5.3.3	对比实验结果与分析	53
5.4	本章小结	55
第六章	结束语	56
	参考文献	57
	硕士期间发表的论文	60
	致谢	61
	附录	62

Contents

Chapter 1 Introduction	1
1.1 Research background	1
1.1.1 Origin of P2P streaming media	1
1.1.2 History and status of P2P streaming media	2
1.1.3 Problems and challenges of P2P streaming media	3
1.2 Science basis of this thesis	4
1.3 Main contents and organization of this thesis	7
Chapter 2 Streaming media technology and P2P technology	9
2.1 Streaming media technology	9
2.1.1 Concept and system organization	9
2.1.2 Transmission principle	9
2.1.3 Data compression technology	10
2.1.4 Streaming transmission protocols	11
2.1.5 File formats	12
2.1.6 Transmission modes	13
2.1.7 Applications of streaming media	14
2.2 Peer to peer technology	14
2.2.1 Definition of P2P and P2P overlay	14
2.2.2 Characteristics of P2P	15
2.2.3 Classifications of P2P topological structure	16
2.2.4 Problems and challenges of P2P	17
2.2.5 Applications of P2P	19
2.3 Summary	19
Chapter 3 Source-driven multicast tree models	21
3.1 Concepts of source-driven and data-driven	21
3.2 Representative Source-driven multicast tree models	21
3.2.1 PeerCast	21
3.2.2 NICE	23
3.2.3 Zigzag	24

3.2.4	SplitStream	29
3.2.5	CoopNet	32
3.3	Summary	32
Chapter 4	Data-driven model based on gossip-algorithm	34
4.1	Related theories	34
4.1.1	Power law	34
4.1.2	Small world	34
4.1.3	Gossip-based algorithm	36
4.2	Design and optimization of DGM	38
4.2.1	P2P overlay architecture	38
4.2.2	Modular architecture of DGM	39
4.2.3	Peer join and membership management	40
4.2.4	Partnership management and partner refinement	42
4.2.5	Buffer map representation and data exchange	42
4.2.6	Schedulling management and algorithm	45
4.3	Summary	46
Chapter 5	Experiments	47
5.1	Network simulator	47
5.1.1	NS-2	47
5.1.2	Simulation instances	48
5.2	Experiment scheme and parameters design	50
5.3	Design and results analysis of experiments	50
5.3.1	Performance under stable environment	50
5.3.2	Performance under dynamic environment	52
5.3.3	Comparison between DGM and Zigzag	53
5.4	Summary	55
Chapter 6	Conclusion	56
References	57
Researches	60
Acknowledgement	61
Appendix	62

厦门大学博硕士学位论文摘要库

第一章 绪论

本章首先概述了 P2P 流媒体技术、P2P 流媒体的发展历程及面临的主要挑战，然后对论文的选题依据进行阐述，最后介绍了论文的主要内容及组织。

1.1 论文的研究背景

1.1.1 P2P 流媒体技术的由来

计算机技术特别是网络技术的飞速发展，深刻地改变着人们的工作、生活和娱乐方式。随着单机处理能力的增强和宽带网络的普及，人们不再满足于传统的网页浏览、文件下载、聊天等 Internet 呆板的表现形式，流媒体(Streaming Media 或 Media Streaming) 以其特有的娱乐性和交互性逐渐成为 Internet 应用的一个热点。

1995 年 4 月，美国西雅图的一个名为 Progressive Networks 的小公司，在全美广播者联合会上，推出了一款名为 RealAudio 的软件，该软件可以通过一种我们后来称为“流”的方式，实现音频在 Internet 上的实时传送，普遍认为这就是流媒体技术的起源。经过十余年的发展，现在流媒体技术已是 Internet 上传送音、视频数据的重要技术之一。

传统的流媒体服务系统一般是基于 C/S 模式，服务器以单播的方式和每个用户建立连接，由于流媒体具有数据存储量大、带宽占用高、持续服务时间长、高 QoS 等特点，随着用户数量的增加，服务器的带宽很快被消耗完，这种情况下的音、视频播放经常出现断断续续，需要不断进行缓冲的情况。为了解决系统的可扩展性，研究者提出了多种解决方案：比如设置服务器集群或提高单个服务器带宽，但这需要巨大的资金支持；采用 IP 组播技术可实现高效的一对多的通信，提高了系统的可扩展性，但 IP 组播存在种种限制，很难实现可靠性组播和拥塞控制等，所以也没有得到广泛的应用。

P2P (Peer to peer) 技术的出现为流媒体服务技术的发展提供了新的契机。P2P 技术受到广泛关注源于 1999 年 Napster^[18] 引起的音乐文件下载热潮。Napster 采用了与传统的 C/S 完全不同的方式，即 P2P 方式提供音乐文件的共享、交换和下

载。Napster为各个用户提供了一种平台，允许用户发布自己拥有的音乐文件并允许用户搜索自己感兴趣的音乐，一旦定位到音乐文件存放的主机位置，交换就在用户之间直接进行，Napster本身并不参与文件的交换。由于利用P2P技术实现了信息的直接交互和共享，Napster在短短的一年多的时间，就吸引了上千万的用户，成为当时成长最快的站点。

P2P技术在文件共享等大规模网络应用方面的成功促使流媒体的研究者将P2P技术应用于流媒体服务系统中。流媒体服务最重要的两种资源就是服务器能力和网络资源，采用了P2P技术，可以合理地使用用户计算机空闲的资源和带宽提供部分服务，每个节点在从别的节点接收媒体流的同时还可以向其他节点发送媒体流，系统的处理能力随着用户数量的增多而不断增加。以中国目前用户最多的提供视频点播的PPLive^[4]网站为例，每一个流媒体源分发服务程序只占用5%左右的CPU负载、20 MB的内存和10 Mbps 的网络带宽。一台接入互联网的100 Mbps 的普通PC服务器可以同时提供5~10 路视频节目的直播，每一路节目均可以支持百万用户同时收视。可以说，P2P技术在现有的服务器处理能力和带宽资源实现了大规模网络流媒体播放，使广大互联网用户进入了廉价视频新时代。

1.1.2 P2P 流媒体技术发展历程及现状

流媒体的兴起及对传统视听媒体产生的影响，就如同互联网对印刷媒体的影响一样。利用P2P 技术实现流媒体点播和直播的系统Webcast出现于1998年。Webcast利用一棵二叉组播树在用户之间进行实时多媒体数据的传输和共享。

2000年出现第一套P2P视频直播系统的原型——ESM系统，该系统采用用户网状结构互连构造最优媒体数据组播树的方法在用户间传播实时的多媒体内容。由于算法限制，这套系统只能供几千人同时在线，但已经标志着P2P流媒体直播系统进入了系统发展期。

此后，Stanford University、Purdue University、Maryland University 等许多大学，Microsoft Research等一些研究机构纷纷开展基于P2P的流媒体服务体系的研究。其中典型的系统有提供音频广播的Peercast系统和P2PRadio系统，应用层组播协议有Coopnet^[34]/Splitstream^[33]协议、Overcast协议、NICE^[30]协议。这些研究都为P2P流媒体直播打下了坚实的理论基础。

2004年6月BBC采用PeerCast技术开始进行其“柔性电视”(Flexible TV)试验。英国的宽带用户可以下载安装“BBC互联网媒体播放器”，用它下载收看前一周和后一周的BBC的全部节目。当时网络视频流量首次超过音频流量。所以2004年6月被认为是互联网历史上的分水岭^[1]。

2004年5月欧洲杯期间，香港科技大学张欣研博士开发的CoolStreaming原型系统在开放性测试平台Planetlab^[28]上试用成功。该系统是第一次真正将组播协议应用在P2P流直播系统当中，标志P2P直播技术进入准商业运作阶段。在CoolStreaming成功的鼓舞下，国内一些高校和科研单位，如清华大学、华中科技大学等纷纷投身于P2P流媒体直播或点播的研究。

目前我国P2P流媒体技术和应用发展迅速，已经开发出实用的系统和软件，如GridMedia^[2]、AnySee^[3]等；诞生了多个专门性的视频网站：如PPLive^[4]、QQ直播^[5]、沸点网络电视^[6]等。它们使用各自开发的软件提供P2P流媒体视频业务，掀起了一股P2P网络电视浪潮。

除了网络电视点播，许多网站还开展新闻、大型体育赛事或娱乐节目的直播业务，例如新浪等网站对“神舟六号”升空和返回过程进行的视频直播，PPLive对“超女”比赛进行的直播，中央电视台对春节晚会进行的直播等。

现在P2P流媒体的发展方向主要体现为网络电视、电影、音乐、游戏、动漫等多种媒体的结合。这类的网站有：PPStream^[8]、MySee^[9]、猫眼网络电视^[10]等。

1.1.3 P2P 流媒体研究和应用面临的挑战

尽管满足P2P流媒体服务的基础环节如服务器性能、网络带宽、视频编码、传输等技术取得飞速发展，但流媒体有其独特性质：如数据存储量大、带宽占用高、持续服务时间长、高QoS要求等。另外，P2P流媒体系统与P2P文件共享系统也有着极大的不同，主要表现在对等节点间数据共享模式上：传统文件共享系统是“下载后运行”(Play after downloaded)模式，而流媒体系统采用“边下载边运行”(Play while downloading)。具体来说，在一个P2P流媒体系统中，一个对等节点拥有一个特定的媒体文件(或文件的一部分)，并向对此文件感兴趣的其他节点提供媒体数据。与此同时，请求数据的节点在下载媒体数据的过程中回放并存储这个媒体的数据，并成为可以为其他节点提供流媒体数据上传服务

的节点。而这些节点与传统的服务器相比，具有动态和异构的特性，主要表现为：

- (1) 对等网系统中的自主节点可能随时加入或退出系统；
- (2) 数据发送节点随时可能加入或停止流媒体服务会话；
- (3) 节点的网络连接可能随时断开；
- (4) 发送节点的上传带宽可能随时改变；
- (5) 数据接收节点与多个数据发送节点间的连接具有不同的端到端带宽、丢包率和失效率。

因此，在 P2P 流媒体服务体系中，如何在充分利用众多节点资源的同时并确保服务质量，面临着许多挑战，可以概括为：

- (1) 服务搜索。即在系统中如何搜索定位合适的服务节点。
- (2) 数据分配。即如何组织数据流，如何在节点间进行数据分配传输。
- (3) 容错。即针对节点的动态性和节点间网络带宽的动态性，如何保证节点的服务质量。
- (4) 激励。如何对系统中节点的行为和兴趣进行统计建模，确立一个合理的激励机制，鼓励参与的客户能最大化地共享资源。
- (5) 安全。在互不信任的节点间提供服务无疑需要一个有效的认证或授权等安全机制。包括如何评估共享资源的真实性和可靠性，如何提供数字证书、对等节点认证、安全存储等。

1.2 论文的选题依据

流媒体技术是近几年刚兴起的一项技术，正处于蓬勃发展的阶段，介入其中进行研究是一项十分有意义的工作。但流媒体技术是个繁杂的系统工程，其关键技术包括流媒体编解码技术、流媒体的网络传输和控制协议、调度和存储技术、代理服务器及缓存技术、拥塞控制技术等多项技术和多个研究领域，要对所有内容都进行细致研究不太现实，本文研究的重点主要是基于P2P覆盖网的流媒体分发。

目前流媒体的分发技术，有基于传统C/S的CDN (Content Delivery Network, 内容分发网)和IP组播。它们虽然能够减缓服务器和网络的负载，但CDN需要在大量范围内布置大量的代理服务器，代价太高，而IP组播实现方面的复杂性以及在拥

塞控制、可靠性管理方面的不足使其在近几年内难以广泛实施。所以，多数研究又转向围绕应用层的组播展开，出现了许多可扩展的组管理和消息传播协议。应用层组播相比于IP组播，具备下列优点：不需要路由器的支持，可以在现有的网络基础上渐进部署；比IP组播更灵活，能够适应上层应用不同的散发要求。

基于应用层组播的协议通常用组播树来实现，即把参与服务的所有节点组织成一棵组播树，树的根节点为源服务器，父节点负责为子节点传送数据。组播树的构建与管理需要解决如下问题：

(1) 树的深度与宽度的平衡。由于组播树中一般节点的服务能力有限，其所带子节点的数目不能太多；另外，节点与源节点服务器之间的路径不能太长，否则流的延迟太大。树管理的目标就是在深度和宽度之间能够有效地平衡。

(2) 逻辑树的有效性。树的逻辑拓扑应尽可能地符合节点间的物理拓扑。

(3) 节点的快速加入和退出。由于用户行为的不可预期，任何节点都可能随时离开P2P网络，因此在节点退出后如何快速地修复树的结构对于服务质量也非常的重要。

(4) 扩展性 (scalability)。如何保持树管理的算法在节点数量非常大的时候仍然有效。

事实上，维持一个稳定的树状结构十分困难，并且在组播树中，处于最下层的节点只享受资源，而不提供服务，致使网络资源并未完全得到利用。

Gossip算法(gossip-based algorithm)为问题的解决提供了新的思路。基于Gossip算法的系统并不显式地构造节点之间的拓扑结构，而是通过Gossip协议，每个节点维护系统中部分其他节点的视图 (View)，动态地和其他节点交换缓存信息，并根据缓存信息交换节点之间的媒体数据。

用Gossip算法构建的覆盖网相比于组播树的覆盖网，有下列优势：

(1) 非中心化和动态性。信息转发过程中，子节点的选择完全是随机的；

(2) 对失效的恢复力强。因为转发节点选择的随机性，某个节点故障不会影响信息在整个网络的发送；

(3) 可靠性强。一条信息一次被随机地分发给多个接收者，这种信息的冗余保证了传送的可靠性；

(4) 扩展性强。文献^[7]证明，在具有N个节点的网络中，信息只需 $O(\log N)$

轮就可传遍整个网络。

(5) 控制开销低。因为它并不需要显式地维护一个固定的结构。

基于 Gossip 算法的上述优点, 它特别适用于 P2P、无线和传感器网络等动态性比较强、连接可靠性差的随机网络中^[7]。目前有许多研究和应用也围绕着这些方面展开, 如 P2P 系统中节点的失效侦测^[46]、资源发现和监控^[47]、数据库复制^[43]、无线和传感器网络^[56, 57]等。对流媒体分发技术进行了研究的文献有^[45, 48, 49, 53, 58]等。

文献^[49]主要基于 Gossip 算法信息分发过程中组员维护 (membership maintenance)、网络认知 (network awareness)、缓冲管理 (buffer management) 和信息过滤 (message filtering) 的问题进行了理论上的探讨。关于组员维护中, 文章提出节点之间如何互相发现及每个节点需要和多少个节点保持联系。在网络认知中, 提出如何组织节点, 以反映网络真实的拓扑结构, 以便信息分发更为有效; 在缓冲管理中, 提出当缓冲器满时, 用什么策略释放缓冲器的空间; 在信息过滤中, 提出节点如何接收或转发自己或别的节点感兴趣的信息, 避免存储或转发不感兴趣的信息, 从而减少传输的信息流。

文献^[48]提出了一个 SCAMP (Scalable Membership Protocol 可扩展的成员协议)。协议包括基本的组员管理和维护图平衡机制。在基本组员管理协议中, 它详细定义了订阅 (subscription) 和发送订阅 (forwarded subscription) 的概念和算法, 提出用心跳机制 (heartbeat mechanism) 解决孤立节点的恢复问题。在维护图平衡机制部分中, 它用间接机制 (indirection mechanism) 规定了一个节点发送订阅和停止订阅的规则, 用租约机制 (lease mechanism) 规定了每个订阅都有一个生命期 (lifetime), 当订阅生命期结束, 保留在节点中的信息就会被删除, 需要重新订阅 (resubscription), 文章认为原始的订阅通常被集中于几个有限的源节点中, 租约机制有利于将订阅信息在各个节点重新分发, 实现非中心化的操作。该协议也提出了部分视图的概念。

文献^[58]构建了 lpbcast (Lightweight Probabilistic Broadcast) 的模型, 它是基于 Gossip 广播算法和基于 Gossip 的组员管理, 描述了信息的 Gossip 接收和分发算法, 订阅 (subscribing) 和退订 (unsubscribing) 协议, 为了避免退订者信息被永久留在系统中, 时间戳 (timestamp) 被加在退订信息上, 过了

Degree papers are in the "[Xiamen University Electronic Theses and Dissertations Database](#)". Full texts are available in the following ways:

1. If your library is a CALIS member libraries, please log on <http://etd.calis.edu.cn/> and submit requests online, or consult the interlibrary loan department in your library.
2. For users of non-CALIS member libraries, please mail to etd@xmu.edu.cn for delivery details.

厦门大学博硕士学位论文摘要库