

西南科技大学

硕士学位论文

基于复杂网络的P2P流媒体网络的研究

姓名：李甫

申请学位级别：硕士

专业：计算机应用

指导教师：徐鲁强

20080621

## 摘 要

流媒体正在成为推动未来网络应用的主要动力,传统网络模型已不能满足其对服务质量及网络稳定性的要求,因此,构建适合流媒体传输的 P2P 网络正逐渐成为当今网络模型构建领域的热门话题。本文基于模拟退火算法中的 Metropolis 准则和复杂网络模型,构建出了两种不同类型的 P2P 网络,并根据流媒体传输的特点做了研究与分析,主要工作如下:

(1)基于模拟退火算法中 Metropolis 准则构建出一种结构对等的 P2P 网络,以每个节点的可用带宽与总带宽的比值作为节点转移的参数,按照模拟退火算法在恒温下的 Metropolis 准则进行节点转移,找出理论上最优的节点作为连接对象。

(2)基于复杂网络无标度网络模型,构建一种不对等的 P2P 网络,先部署一定数量的性能较优点,这些节点之间随机建立连接形成树状,后面来的节点按照复杂网络无标度网络模型的偏好性建立连接。

实验表明,通过复杂网络模型无标度方法构建的 P2P 网络具有较好的稳定性与较短的传输延迟,符合流媒体传输的需要,是一种较为合适的 P2P 流媒体传输网络。

**关键词:** P2P 流媒体 模拟退火 复杂网络 幂律分布

## Abstract

While traditional network model fails to meet the requirement of a stable network and a high service quality, streaming media is becoming the major drive for the future network application. Thus, the building of p2p streaming media is becoming a hot topic today in the field of network model. This study, based on the metropolis rule of simulated annealing and complex network model, establishes two different p2p networks. Analysis is also done based on the characteristics of streaming media in this study. The major work is as follows.

(1) Establish a p2p network with the equivalent structure based on the metropolis rule of simulated annealing. The ratio of useful bandwidth and the complete bandwidth is taken as the parameter of peer movement, which is done based on the metropolis rule of simulated annealing at the constant temperature. And then we find the theoretically best peer as the connection target.

(2) Establish a non-equivalent P2P network based on the free-scale model of complex network. First, find some peers with high quality, which are to be connected at random in the form of a tree. And the peers coming later are to be connected based on the preference of complex network free-scale model.

The experiment indicates that the p2p network based on the free-scale model of complex network has a good stability and smaller transmission suspension. It can meet the requirement of streaming media and thus is a suitable transmission network of p2p streaming media.

Key words: p2p streaming media; stimulated annealing; complex network; power-law distribution

## 独 创 性 声 明

本人声明所呈交的论文是我个人在导师指导下进行的研究工作及取得的成果。尽我所知，除了文中特别加以标注和致谢的地方外，论文中不包含其他人已经发表或撰写过的研究成果，也不包含为获得西南科技大学或其它教育机构的学位或证书而使用过的材料。与我一同工作的同志对本研究所做的任何贡献均已在论文中作了明确的说明并表示了谢意。

签 名：李甫

日 期：2008.6.21

## 关于论文使用和授权的说明

本人完全了解西南科技大学有关保留、使用学位论文的规定，即：学校有权保留学位论文的复印件，允许该论文被查阅和借阅；学校可以公布该论文的全部或部分内容，可以采用影印、缩印或其他复制手段保存论文。  
(保密的学位论文在解密后应遵守此规定)

签 名：李甫

导师签名：徐鲁强

日 期：2008.6.21

# 1 绪论

## 1.1 课题背景及意义

多媒体使计算机能够综合处理声音、文字、图像和视频,它以丰富的声、文、图像信息和方便的交互性,极大地改善了人机界面,改变了人们使用计算机的方式。从而为计算机进入人类生活和生产的各个领域打开了方便之门。给人们的工作、生活和娱乐带来了变化。随着网络宽带化的发展趋势。人们不再满足于信息高速公路上仅有文本、图像或声音这一类简单的信息,而越来越希望更直观、更丰富的新一代信息的表现形式,流媒体<sup>[1]</sup>(media streaming 或 streaming media)技术由此便应运而生。

传统的分布式多媒体系统主要使用客户端-服务器模式,服务器以单播的方式和每个客户建立连接。由于流媒体服务具有高带宽、持续时间长等特点,随着客户数目的快速增加,服务器的资源如带宽很快就被消耗完,成为系统瓶颈所在。为了解决系统的可扩展性问题,许多研究提出了相应的解决办法。如 IP 组播技术实现了 Internet 上高效的一对多通信,提高了系统的可扩展性。然而由于 IP 组播存在种种限制,如很难实现可靠性组播和拥塞控制等,IP 组播技术并没有得到广泛的应用。而针对服务器提出的批处理(Batching)技术、补丁(Patching)技术、周期性广播以及流合作、流搭载(PingBack)等技术虽然在降低服务器的带宽需求和客户等待时间方面取得了不错的效果,但仅仅在一定程度上缓解了服务器的压力。另一种方案就是网络边缘部署代理缓存(Proxy caching)或内容分发网络(Content Delivery Networks),媒体服务器将媒体内容以推(Push)的方式存放在代理缓存或 CDN 服务器上,客户请求媒体服务器时,可从代理缓存或 CDN 服务器获得服务,而不必消耗服务器的资源。但这种方案成本较高,只是部分地解决了可扩展性问题,因为此时代理缓存或 CDN<sup>[2]</sup>(Content Delivery Network)服务器很有可能成为系统瓶颈。

近年来,对等网络<sup>[3][4][5]</sup>(Peer-to-Peer Network)的研究在文件共享、应用层组播方面得到了广泛地关注,P2P 技术得以快速地发展,在 P2P 方式下,每个对等实体既是服务器的提供者,又是服务的享用者。Peer 为系统提供有限的计算或存储资源,Peer 之间相互协作为其他 Peer 提供服务,减轻了服务器的负载和网络带宽占用,极大地提高了系统的可扩展性,具有潜在的应用前景。由于 P2P 模式可以在大规模网络应用中有效地提高系统的可扩展

性, 基于 P2P 方式的流媒体传输的研究也逐步引起了人们的重视, 相关技术与原型系统不断出现。

本文的主要工作是先使用模拟退火算法 Metropolis 准则构建 P2P 流媒体网络, 提出了具体的构建思路, 发现通过这种方法形成的 P2P 网络, 节点的负载比较均衡, 节点之间度的差比较小, 但是忽视了实际节点之间的差异, 容易在网络的局部引起振荡, 对于流媒体网络来说, 稳定性反而不好。由于节点的度分布比较散, 那么从叶子节点到源节点的层数增加, 于是跳数加大, 那么网络的延迟较大, 不能满足流媒体实时性高的要求, 因此这种构建方式比较适合于文件共享等 P2P 网络。

通过复杂网络的无标度网络构建方式, 产生的 P2P 网络, 度比较大的节点主要分布在最先随机部署的网络节点上, 而这些度大的节点正是性能较好, 比较稳定的节点, 那么整个网络也比较稳定, 其它节点的离去或出现故障不会引起网络的局部振荡。由于节点间度的差距较大, 叶子节点较多, 那么从叶子节点到源节点的跳数较少, 所以网络的延迟小, 这正满足了流媒体网络实时性高的要求。所以基于复杂网络的 P2P 网络适合于流媒体网络的构建, 对实践有借鉴意义, 为 P2P 流媒体网络的构建寻找新的理论支撑。

## 1.2 传统流媒体技术框架

在一个典型的流媒体应用中, 媒体经过压缩技术存在硬盘上, 流媒体服务器接收客户端的请求, 应用层 Qos 控制层通过 Qos 的要求以及网络状况调整压缩的媒体流的参数(如比特率), 然后流媒体传输层对其进行打包进行连续媒体分发, 客户端接收到媒体流后。通过媒体同步技术进行播放。以上的过程涉及到六个技术方面: 媒体压缩技术, 应用层的 Qos 控制, 数据分发技术, 视频服务器技术, 媒体同步技术和流媒体传输协议。图(1-1)显示了传统典型流媒体系统的技术框架。

传统流媒体服务都是 C/S 模式, 即用户从流媒体服务器点击观看节目, 然后流媒体服务器以单播方式把媒体流推送给用户。这种 C/S 模式加单播方式的缺陷在流媒体业务发展到一定阶段, 用户量上来后就会显现出来。主要有下面两个问题:

流媒体服务器带宽占用大。不同于电台和电视台使用广播形式发送节目, 流媒体业务使用了单播形式, 即一个用户一份流, 即使有两个用户在观看同一个节目。用户数和服务器带宽消耗成正比, 用户越多, 需要带宽也就越多。当用户到达一定规模后, 带宽就会成为业务发展的瓶颈, 这时就需要

投入大量费用购买带宽以满足要求。

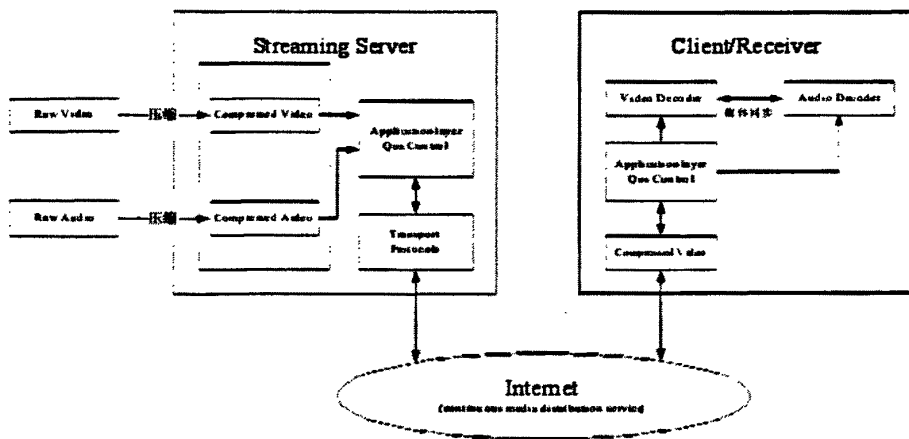


图 1-1 传统的流媒体系统的技术框架

Figure 1-1 Traditional technical frame of steam media system

流媒体服务器处理能力要求高，用户量上来后，流媒体服务器处理能力也会成问题，这时候需要购置服务器以支持更多用户。流媒体负载均衡要求高。为减少骨干网络带宽占用、保证服务质量和就近提供服务，一般流媒体服务都需要部署复杂内容分发系统(CDN)。这样就大大增加了系统投资和管理复杂度。流媒体业务在到达一定阶段后，就需要大规模扩充带宽、服务器和内容分发系统以满足需求，这些举措无疑都会大大增加开销。同时这种用户上来了就扩容，扩容完了再大力发展用户的模式，它无法从根本上解决流媒体业务发展所遭遇到的瓶颈问题，于是我们就需要通过新的 P2P 技术来解决目前存在的这些瓶颈问题。

### 1.3 P2P 流媒体

所谓的 P2P 流媒体，字面上说就是用 P2P 的方式传输流媒体数据。有两个概念组成，P2P 指传输方式，流媒体指传输内容。所以归根结底 P2P 流媒体要解决的问题就是特定目标流媒体的传输，如何用 P2P 的方式优化。P2P 的方式，在于优化资源。何谓资源？带宽是资源，运算力是资源，各个节点的缓冲区也是资源。这些都是可见的，还有一些容易被忽略的资源，比如节点稳定在线时长，各节点间时延、带宽、TTL、物理距离和位置分布(包括相

对位置分布和绝对位置分布), 甚至可以据此绘制出全局路由表, 可以在地图上绘制出节点大概位置。

P2P 流媒体的最终目的就是通过资源的优化调度, 使得数据块在适当的时间被适当的节点持有(或者说在适当的时间到达适当的地点), 达到尽量多的节点能够流畅播放流媒体的效果。

P2P 的流媒体可以细分为三种应用模式: 直播、点播、下载。这三种模式是可以相互转化的。高性能的 P2P 文件下载技术已经很成熟了。点播的定义是可以 Seek。点播的实质就是可以使指定文件偏移位置的数据在一定范围内顺序到达, Seek 的过程就是重新计算偏移量并重定位的过程。点播再进一步, 取消 Seek 功能, 但是对小范围数据块按时到达的要求更加强烈。下载和点播传输的都是静态文件中的数据块, 轮播和现场直播传输的都是实时接入系统的数据片断。很多人对 P2P 流媒体的优化专注在 P2P 传输方面, 以满足以上需求。但是忽略掉了这个主体的另一个方面, 流媒体。P2P 传输可以为了流媒体的特点优化, 流媒体也同样可以为了 P2P 的优化。采用高性能的视频、音频编码算法, 可以让 P2P 的传输压力极大的减轻。索引区块的优化可以让点播的 Seek 过程更加高效。特殊的数据打包格式可以加入跟随流同步的各类信息, 比如字幕、当前播放场景的说明, 甚至超级链接、广告、图片、程序动作等等任何你希望和播放区间同步的内容。把静态文件编码成为在网上传输的数据块集合, 可以优化从读取文件到发送到网络的过程, 如果索引也照此优化, 就可以在 Seek 时直接取得数据块编号。下载、点播、轮播因为可以预先编码, 可以使用计算复杂度高非实时的编码算法, 而现场直播必须使用实时编码, 所以前者理应获得模型更高的质量。考虑 P2P 流媒体, 可能需要更多考虑资源(存储、运算、传输)的成本趋势。

那么当前 P2P 流媒体的一个热门话题是实时流媒体的分发技术, 即我们建立一个什么样的分发规则, 使流媒体能高效通畅地传到需要它的 Peer, 下面我们讨论一下目前国内外的现状。

## 1.4 国内外研究的现状

针对流媒体高带宽的需求, 采用 P2P 技术是解决流媒体内容分发中服务器链路瓶颈问题的理想模式, 其核心思想就是把流媒体数据采用 P2P 网络进行分发, 在 P2P 系统中各节点地位是平等的, 贡献自己资源给群体, 反过来又使用其他节点提供的资源, 从而达到节点数越多(规模越大)流媒体传输质



量越高。目前普通用户所使用的计算机能力、存储能力和带宽较强大和丰富，并且用户所拥有的巨大的资源通常处于闲置状态，若能将这些闲置的存储空间、带宽资源和处理能力利用起来，就可以在不增加用户使用费用的同时增加巨大的数据分发能力，用户在将自己的计算机资源贡献出来的同时也分享其它用户提供的资源。在 P2P 网络中，流媒体数据使用自己节点的带宽进行分发，节点为了共享内容而合作，数据通信量不会固定在某个特定的地方，不会造成某个特定地方成为性能瓶颈。

如何实现基于 P2P 网络对流媒体内容进行大规模分发是近些年来多媒体领域的热点研究方向。目前基于 P2P 网络的流媒体内容分发技术已引起国内外许多大学和研究机构的纷纷重视并开展研究(如美国的斯坦福大学、普度大学、卡内基梅隆大学、佛罗里达中央大学、微软研究院、国内的香港科技大学、清华大学、国防科技大学、华中科技大学等)，也有些系统已经开始尝试商业化运作。

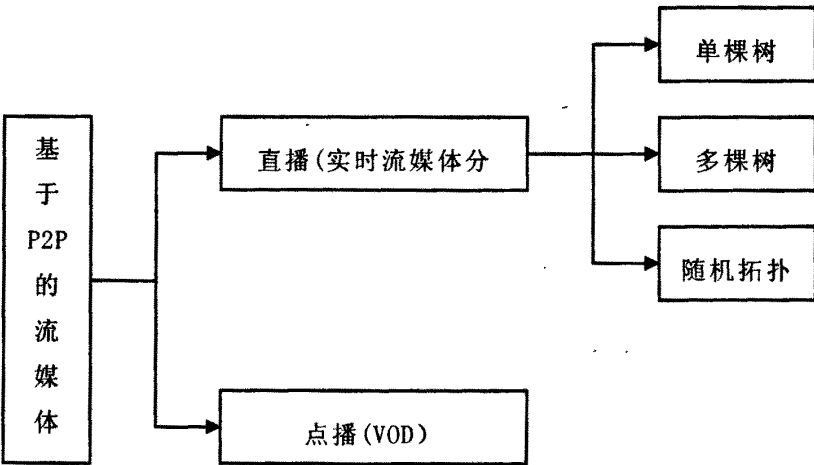


图 1-2 基于 P2P 的流媒体分发技术方案

Figure 1-2 Stream media distribution scheme based on P2P

目前基于 P2P 网络对流媒体分发的应用研究可以划分为两类:基于 P2P 的流媒体直播和基于 P2P 的流媒体点播。前者主要实现对采集到的实时流媒体数据的分发，后者则侧重实现在流媒体文件分发服务中的交互。在基于 P2P 的实时流媒体分发系统中首先需要解决的一个问题是让新加入的节点从哪个或哪些节点获取节目数据，这实际上也是节点逻辑拓扑组织的问题。另

外一个需要解决的问题是当节点退出或失效时对与其相关节点的处理问题，这也就是节点逻辑拓扑调整的问题。在目前基于 P2P 的实时流媒体分发方案中，按照 P2P 网络中节点所构成的网络拓扑划分为三类：基于单棵树的数据分发、基于多棵树的数据分发以及基于随机拓扑的数据分发。

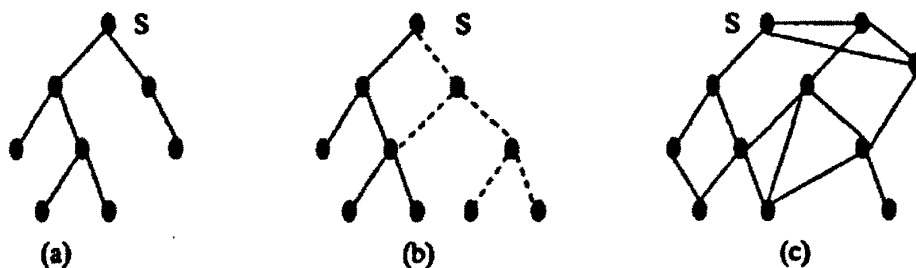


图 1-3 P2P 实时流媒体分发网络拓扑

Figure 1-3 P2P real-time stream media distribution network topology

#### 1.4.1 基于单棵树的数据分发

在基于单棵树数据分发模式中，每棵组播树上的节点共享从源服务器一个频道所流出的数据，每个节点只参与到一棵组播树。研究的焦点主要是如何设计组播树构造协议以达到特定的性能指标要求，此外还包括如何增强系统的容错性，即如何减少节点离开或失效行为对其它节点的影响。

在基于单棵树的 PeerCast<sup>[26]</sup>模型中提出了一种简单组播树构造协议：节点被组织成一个树状结构，树的父节点给子节点提供服务。新节点  $c$  的加入时，试图向根节点  $n$  请求加入，如果节点  $n$  有足够的资源，则向节点  $c$  提供服务，否则节点  $n$  把  $c$  的请求根据一定的策略转发给其某个直接的子节点。子节点根据自己的资源情况判断是否给  $c$  提供服务，以此类推，直到  $c$  找到一个父节点。对于节点的失效情况，通过在所有父子节点之间维持心跳 (heart-beat) 功能，如果在几个周期内收不到父节点的心跳信号，就认为父节点已经失效，并重新向根节点发起加入搜索过程。

在 PeerCast 中，虽然提供了一些策略使组播树尽可能的平衡，但并没有任何理论上的保证。针对这个问题，Tran 等提出了一种基于分层簇思想的组播树构建的 Zig-Zag<sup>[27]</sup>模型，提供了更完整的构造组播树的规则，保证树的深度维持在  $O(\log N)$ ， $N$  为系统中的节点数量。在 Zig-Zag 中，节点之间的

关系有两种,一种是逻辑关系:簇管理结构(Administrative Organization)还有一种是真正负责数据分发物理关系:组播树(Multicast tree)。在 Zig-Zag 中,节点通过某种规则组织成簇(Cluster),簇管理结构通过把所有的节点组织成一个继承体系来成管理系统中所有的节点。

为减少组播树中节点离开或失效时子节点服务被中断的时间, Yang 等提出了一种基于对组播树进行主动重建的中断恢复方案,它让组播树中的每个中间节点预先为每个子节点计算出一将来的父节点(Parent-to-Be),以便当节点退出系统时,服务被中断的子节点就可以直接与其将来的父节点进行联系,从而能够快速重新加入系统。对于如何为每个子节点预先计算将来父节点的问题, Yang 等考虑了每个节点可带子节点个数(度)的约束,把上述问题归结为带度约束的生成树问题(Degree-Constrained Spanning Tree Problem),并提出了相应的算法。

在基于单棵树的数据分发方案中,由于节点的离开或失效都将可能导致其后续子节点传输服务的中断,为维护每个节点上视频播放的连续性, Guo 等提出了一种基于 Patching 流的数据分发模式:在源服务器节点除了用一条频道分发正常的直播数据外,还额外启用多条频道,每条频道用于传输一条 Shifting 流,这些 Shifting 流之间存在固定的发送启动延迟;当有中间节点离开或失效时,在服务被中断的节点上同时发起下述两种类型的操作,一是选择加入某条合适的 Shifting 流,以便从数据被中断处继续获取数据(Patching 流),二是重新加入组播树,以便跟上视频直播的正常进度;当节点服务被中断时所产生的数据“空洞”被 Patching 流填充完毕后,随即释放 Patching 流所占用的数据频道,从而恢复仅从组播树上单个节点获取数据的状态。上述模式中的 Patching 流均直接从源服务器节点获取。针对如何减少节点离开失效或链路故障对节点播放质量的影响, Banerjee 等提出了一种联合随机转发(Randomized Forwarding)技术及 NAK 触发(Triggered NAK)机制的流媒体直播数据分发系统 SRMS。在 SRMS 系统中,每个节点除了给其子节点转发数据外,还从系统中随机挑选出固定数目的节点并以较小的概率向这些节点转发数据;而对任一节点 X 而言,在给另一节点 Y 所转发的每个数据包中还须捎带一些数据元信息,用以表示节点 X 已正确收到了该数据包之前的哪些数据包;节点 Y 在收到该数据包后用该数据包中所附带的元信息与其缓冲区内容进行比较,即可检查出缓冲区中还缺少哪些数据包,从而向节点 X 发送 NAK 消息以便请求传输那些空缺了的数据包。在 SRMS 中以多大的概率向固定数目的节点转发数据是系统设计的一个关

键,如果概率选择过大则可能会导致节点接收到大量的重复报文,从而造成带宽资源的浪费;如果概率选择过小,当节点离开失效或链路发生故障时,子节点可能需要经历较长的等待时间才能有机会发送 NAK 报文,从而使得传输服务被中断的时间也较长。

#### 1.4.2 基于多棵树的数据分发

由于基于单棵组播树的数据分发存在诸如叶节点带宽资源得不到有效利用、中间节点的离开或失效行为对子节点的影响较大、QoS 难以保障等问题,人们因此也提出了基于多棵组播树的数据分发。针对基于单棵组播树的数据分发方案中节点离开或失效行为会中断后续子节点数据传输的问题,Padmanabhan 等首先提出了一种基于多棵组播树的数据分发方案 CoopNet<sup>[28][29]</sup>:它首先在源服务器端对视频节目进行 MDC<sup>[30]</sup>(Multiple Description Coding)编码,产生多条 MDC 子流,这些 MDC 子流在解码时不存在依赖关系;而对每一条 MDC 子流均采用一棵独立的组播树进行分发;新节点在加入系统时可选择同时加入到多棵组播树:对任一节点而言,其接收到的 MDC 子流数量越多,则播放质量越好;当某棵组播树上的父节点离开或失效时,只会导致一条 MDC 子流的数据传输被中断,节点还可以从其它组播树上继续接收其它的 MDC 子流,在这种情形下除了画面质量会有所下降外,并不影响节点对节目数据的连续播放。对于多棵组播树的组织和维护,主要由中心索引服务器集中承担,节点在加入或重新加入时均向中心索引服务器提交请求,以获取多个或单个父节点。此外,针对节点带宽资源异构的问题,Padmanabhan 等还提出了一种基于分层 MDC 编码网的多棵组播树数据分发方案,以解决不同节点对不同播放质量等级的要求。

在基于单棵组播树的数据分发方案中存在如下两方面的矛盾:一方面组播树上的中间节点承担着较重的数据传输负载,另一方面还有大量叶子节点的带宽资源得不到有效利用。针对此问题微软亚洲研究院的 Castro 等提出了另一种基于多棵组播树的数据分发方案 Splitstream<sup>[31]</sup>,它与 CoopNet 一样在源服务器端采用 MDC 编码,并让每棵组播树对应传输一条 MDC 子流,不同之处是在 Splitstream 中,每个节点只在其所参与的一棵组播树中充当中间节点,而在剩余的其它组播树中只充当叶节点,这样就可以减小节点离开或失效对其它节点所造成的影响。

与基于单棵组播树的数据分发方案相比,基于多棵组播树的数据分发方案的优点是可以充分利用系统中每个节点的带宽资源,通过与 MDC 编码结合

可以减少节点离开或失效对其它节点播放质量的影响;其缺点是对组播树的维护和管理比较复杂,此外 MDC 编码本身的效率也比较低,并且其解码过程也比较复杂。

### 1.4.3 基于随机拓扑的数据分发

在基于随机拓扑的数据分发方案中,节点之间所构成的网络逻辑拓扑是随机的,采用这种数据分发方案的典型直播系统有 Coolstreaming<sup>[32]</sup>、GridMedia<sup>[33]</sup>、Anysee<sup>[34]</sup>等,它们的共同特点是节点在加入时均从现有系统中随机挑选出多个节点作为其邻居节点,每个节点均和其邻居节点定时交换其所缓存数据的元信息,并以此为基础在每个节点上定时运行数据调度算法,用于确定从哪个邻居节点获取哪些数据,以最大化节点的播放质量。由于节点同时从多个节点调度获取数据,因此它可以在很大程度上降低节点离开或失效对其它节点的影响。与基于多棵组播树的数据分发方案比较,基于随机拓扑的数据分发方案的最大优势是既不需要在节点间维护复杂的网络逻辑拓扑结构,也不需要源服务器端把数据流切分为多条子流。但是基于随机拓扑的数据分发也有其自身的缺点,需要从多个节点中寻找适合每个数据片断的合适提供者,这样就增加了启动过程中的延时。

## 1.5 本文工作及内容安排

在当前流行的 P2P 网络中,大多构建方式都只注重网络节点的对等性,而忽视那实际网络节点的不均衡性,那么是否可以构建一个 P2P 流媒体网络既满足理想的对等性,也符合实际网络节点的不均衡性呢?答案是肯定的,这正是本文要解决的问题。于是在研究现有 P2P 流媒体网络技术的基础上,结合实际网络的特点,主要研究了在构建 P2P 流媒体网络中,新的节点加入时如何建立连接,主要的内容安排如下:

第一章:查阅了与这一课题相关的大量国内外文献,对 P2P 流媒体网络技术的现状做了简要概论,同时对本论文研究的重点、方法与内容做了论述。

第二章:研究了现有的 P2P 流媒体技术,以及典型的 P2P 流媒体模型。

第三章:介绍了模拟退火算法的基本理论,并论述怎么用模拟退火算法对网络中的节点进行状态转移,在理论上找出网络节点中带宽资源最优的节点。

第四章：首先介绍了复杂网络的基本概念和理论，再论述复杂网络中经典的建模方法，最后分析现实中 P2P 流媒体网络的特点，从理论上得出用复杂网络中无标度网络的建模方法构建 P2P 流媒体网络。

第五章：模拟仿真了两种不同的 P2P 流媒体构建方法，得出相应的数据图表，并对数据图表做出分析。

最后总结了本文所做的工作，并指出待解决的问题和一些尚需进一步研究的内容。

## 2 P2P 流媒体技术的研究

在本章中我们首先研究了 P2P 流媒体的相关理论,然后研究了目前国内外在 P2P 流媒体直播系统领域内的主要成果,并对现有典型的 P2P 流媒体系统模型进行分析。

### 2.1 P2P 流媒体的分类

目前 P2PStreaming 模型主要可以分为两大类:基于树状拓扑的协议及扩展(Tree-based protocol and extensions)和基于网状的协议(Mesh-based protocol)。

#### 2.1.1 基于树状拓扑的协议及扩展的模型

基于树状拓扑协议及扩展的模型把参与P2P流媒体的节点组织成一棵应用层组播树,树的父节点负责为子节点传送数据。因此,这类模型首先构建一个数据的分发树;接着,节点会发现那些树状拓扑中与其不相邻的节点,进而与它们建立一些的控制信息通道,形成控制拓扑结构。这类模型关注的是怎样直接提高这棵数据分发树的效率。这种方式控制相对简单,可以达到很好的数据传输效率。在这类模型中,树的构建与管理主要需要解决如下问题:

(1)树的深度与宽度的平衡:因为节点离根节点越远,则接收到的数据的延时就越大,因此,树的深度应该尽可能短。可是每个节点的输出带宽限制了节点的宽度。树管理的目标就是在深度和宽度之间能够有效的平衡,事实上,当所有节点的深度都为1的时候就退化成了传统的客户端/服务器模型了。

(2)逻辑树的有效性:树的逻辑拓扑应尽可能的符合节点间的物理拓扑。

(3)节点的快速加入和退出:节点的快速加入机制使新加入节点接收服务的延时更小。由于用户行为的不可预期,任何节点都可能在任何时刻离开P2P网络,因此在节点退出后快速的修复树的结构对于服务质量也非常的重要。事实上该约束和上面的两个约束使相互矛盾的,需要做出一些权衡。

(4)扩展性(scalability):树管理的算法在节点数量非常大的时候仍然应该有效。

### 2.1.2 基于网状的协议的模型

基于网状协议的模型中,节点首先分布式地自组织形成一个网状的控制拓扑,一对节点之间可能存在一条或多条连接路径,然后根据协议选取网中边的一个子集计算一棵数据分发树,用以传输数据。这类模型中首先在已有的所有节点之间以分布式算法建立一个较大连通性的覆盖网,网中的每个节点都会和它的邻居节点定期交换状态信息,以达到控制拓扑的优化。重点关注的是怎样提高、优化覆盖网的效率,然后可以利用已有的算法构建数据分发树。这种方式形成的覆盖网连通性较好,可以保证数据分发树足够的冗余备份,使得这种方式的健壮性和可靠性较好。但控制开销太大,导致可靠性问题的出现。

这类模型中有很多是基于Gossip的算法的。基于Gossip的消息散播协议的基本思想是:一个节点随机地选择一组节点发送新产生的消息,收到消息的那些节点使用同样的方式把消息发送出去,直到所有的节点收到消息为止。为了保证这个过程是收敛的,每个消息带有一个TTL,每经过一次转发,TTL就减1。当TTL减少到0时,消息就不再转发了。Gossip协议通过消息的冗余来保证节点崩溃或网络丢包严重的情况下系统的可靠性。基于Gossip协议的P2PStreaming系统并不显式的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,每个节点动态的和其它节点交换缓存信息,并根据缓存信息交换节点之间的媒体数据,在这种系统中,通常需要比较大的缓存,系统的启动延迟相对比较大。

## 2.2 P2P 网络中流数据的传输

P2P流媒体系统在构建P2P覆盖网之后,需要考虑在P2P覆盖网之上选择路径来传输数据。P2P覆盖网中流数据的传输方式存在两种机制:推(PUSH)机制和拉(PULL)机制。

所谓推就是节点主动向另一个节点发送数据,这需要节点之间有一种父与子的关系,父节点依据这种关系主动发送数据给子节点。所谓拉,就是节点首先向另一个节点发出请求,另一节点再根据请求发送数据,这不需要节点之间任何层次性的关系,但是节点需要预先知道对方的含有的数据。一般基于树状拓扑的协议及扩展的结构是典型的推机制,而基于网状的协议的结构则采用拉的机制。

(1) 推(PUSH)机制:



推的机制需要节点之间具有父与子的关系,父节点负责为子节点发送数据。因而父节点的失败时就必须为子节点选择新的父节点,在这种关系调整的时候就可能导致子节点丢失部分数据,因而树结构的机制很难保证质量。此外树结构一般从一个节点请求数据,节点与节点之间带宽一般比较低因而只适合带宽要求低的应用。后来又提出了采用多描述编码(MDC)建立多棵数据传输分发树的机制,通过将单一码流分割为多个码流,对每个码流建立一棵传输树。多树的机制对质量有所改善,但是其算法的复杂度和维护的开销大大增加。

### (2) 拉(PULL)机制:

拉的机制是以存储转发为基础,节点首先获得对方节点所拥有的数据状态信息,然后向其发起请求。拉的机制无需维护树的结构,但是节点之间需要不停的交换缓冲区中数据的状态信息而产生一定的负载。拉的机制可以使节点可以从任何相邻的节点任何时间获取自身需要的数据,因而算法上实现简单而且可以提高自己的下载带宽。

### (3) 推机制和拉机制的比较:

推的机制无须发送数据的状态信息,无须存储转发因而负载较低且延迟较低。当然推的机制需要维护复杂的树结构,而且对动态变化的应对能力比较弱,节点只能从单个节点获取数据,导致采用推的机制质量更差,节点获取数据的能力更差,实现上更为复杂。拉的机制可以随意同时从多个节点获取数据,节点之间完全的对等,其实现也更为简单,更容易处理节点的高动态性,也能获得更高的下载带宽达到更好的质量。拉的机制采用存储转发的思想,需要数据状态的转发,需要在每个节点进行必要的数据缓冲,因而拉的机制导致较长的延时和较多的负载。

## 2.3 现有 P2P 流媒体系统模型

从2000年开始陆续出现了很多基于P2P技术的流媒体系统,例如Stanford大学的PeerCast系统、Microsoft Research的CoopNet系统等等。本节主要对当今典型的基于P2P技术的流媒体系统模型进行了深入地研究,详细分析这这些系统采用的算法和相关机制。

### 2.3.1 PeerCast

PeerCast是美国Stanford大学Peer-to-Peer:研究小组的研究成果。

PeerCast采用一个简单的单根树结构,通过限定每个peer所能服务的节点个数来优化节点间的传输延迟。PeerCast模型通过建立一个以媒体源为根节点的应用层组播数据分发树。PeerCast系统中所有节点在逻辑上组成一棵分发树,系统的数据分发流程均按照该分发树的拓扑结构由上至下分发。

### (1)控制拓扑的构建

节点维护工作可以放在作为根节点的服务器中,或者选择树中其他一部分节点共同维护。每个节点在任何时刻都只与一个节点建立连接,同时向被连接的节点报告自己的存在,然后它从被连接的节点获取数据;当不能获得数据时,它从自己的缓存节点中随机选择节点尝试连接。

### (2)数据传输拓扑的构建

下面主要介绍PeerCast构建组播树的时节点的加入算法和退出算法。

#### a. 节点加入算法

新节点N试图加入时,首先向服务器S请求服务,如果服务器有足够的资源,则服务器向节点N提供服务,否则服务器S把N的请求转发给其某个直接的子节点。子节点根据自己的资源情况判断是否给N提供服务,以此类推,直到N找到一个父节点。在这个模型中,每个节点仅仅维护自己的父节点和直接的子节点的信息。如果被请求节点没有足够的资源,需要把新节点N的请求重定向到(Redirect)其他节点,PeerCast还给出了重定向三种策略:①随机选择;② Round-Robin;③根据物理位置选择(Smart-Placement)。

#### b. 节点退出算法

节点的离开分为正常离开和非正常离开,必须区别对待。在正常离开时,节点通知父节点,并且会重定向子孙节点的数据传输会话。为检测节点的非正常离开,每个节点需要周期性的向他们的父节点和子节点发送心跳包(Heart-Beat),说明自己工作正常。如果某个节点在几个周期内都没有收到子节点正常的消息,则认为子节点已经非正常离开,该节点将回收服务该子节点的资源。同样子节点在几个周期内都没有收到父节点正常的消息,则认为父节点已经非正常离开,该节点以及所有后续节点都需要重新加入网络。

PeerCast在加入和离开,需要不停的建立新的连接才能获得数据,所以节点动态变化对媒体质量有很大的影响。但是对于低码率的应用,PeerCast缓存较多的数据,使得节点在重新连接期间可以正常的播放。PeerCast在数据传输路径上采用推的机制,但是它引入了存储转发的机制使得节点有更多的时间去调整这个树的结构从而达到较好的质量。

### 2.3.2 CoopNet

CoopNet(Cooperative Networking)是微软的一个研究性项目,是Microsoft Research的第一个多树直播视频流媒体网络系统。多树模型通过引入以源节点为根节点的多个分发树,通过多描述编码MDC把流媒体数据分割成 $k$ 个独立的码流,并为每个码流构造一个组播树,形成一片森林。从而每个分发树只分发源的一层数据。多树模型的目标是将同一个节点放在不同分发树的多个位置,这些位置可以随机选择或是采取某种确定性的算法来实现,当前最好的选择算法之一就是使得在这 $N$ 个树中,每个节点只在某一个树中充当内部节点,在其它 $N-1$ 个树中都充当叶子节点。

#### (1)控制拓扑的构建

CoopNet有一个指定的工作站负责管理节点的加入和离开。工作站把组播树的整个结构存储在内存中。当一个节点开始接收现场直播的流媒体时,这个节点与工作站接洽加入的操作。工作站从保存在内存的组播树中找到一个合适的位置,把这个节点的父节点返回给这个节点。

#### (2)数据传输拓扑的构建

CoopNet使用多描述编码(Multiple Description Coding, MDC) [X25]和多路径传输机制保证数据传输的质量。多描述编码将音视频信号编入到多个子集码流中,任意组合的子集码流都可以进行解码。解码之后的质量取决于收到的子集码流数目,子集码流数目越多质量就越好。多描述编码和分层编码是不同的,多描述编码要求描述的每个子集码流都必须是可解码的,而分层编码只要求子集码流的一个嵌套序列可解码。

CoopNet创建多棵数据分发树,每棵数据分发树上传输一个子集码流。如果将视频编码到 $M$ 个子集码流,那么CoopNet创建 $M$ 棵子树。每棵树都是应用层拓扑的生成树,包含所有的节点。同时,每棵树也各不相同,在一棵树中的内部节点在另一棵树中可能为叶子。

CoopNet中使用多描述编码和多路径的传输机制,在一定程度上减小了因为节点动态特性(加入和离开)对其它节点接受质量的影响。同时,CoopNet使用多棵不相交的分发树,使得很多时候节点都能接收到大部分的子集码流,保证一定的流媒体视频质量。但是,CoopNet也有一些缺点。首先,CoopNet需要建立多棵分发树,维护分发树的成本和代价都是很高的。分发树需要不停交换控制信息才能维护,多棵树的维护需要更多的控制信息。一个节点的失败会导致几乎所有分发树的调整,复杂度比较高。其次,需要对音视频信号进行描述编码,制作数据源的负载较大。再者,CoopNet中音视频质量还

是因为部分子集码流的丢失而质量降低。最后，CoopNet难于与现有的系统相融合，不易于使用。CoopNet使用MDC，而现有的系统如Real One Player和Media Player都只支持单描述的编码。

## 2.4 现有系统的分析

在基于树状结构的模型中，(如:PeerCast)虽然取得了网络带宽的有效性，但是组播树中内部节点的离开都不可避免的会影响到后续的子节点，也就是部分的牺牲了系统的稳定性。CoopNet 中采用了多树模型，部署了 MDC，部分地解决了系统稳定性的问题，但却引入了冗余的编码。因此，CoopNet是用网络效率来换取系统的稳定性。再者，这些系统的构建都是基于工程技术人员经验和启发式，缺乏新的理论的支撑，于是本文想寻找一种新的理论，来拟合 P2P 流媒体网络模型，并主要是研究怎么构建。

### 3 模拟退火算法 Metropolis 准则连接的建立

在 P2P 网络中,我们选择什么样的节点才是传输速度最佳的呢?在大多数情况下是网络带宽较大的,但这也不是绝对的,有些情况下网络带宽大的不一定最快,为此在这一章中用模拟退火算法的 Metropolis 准则来解决这一问题。且选择出理论上最佳的节点,作为连接的对象。

固体退火是先将固体加热至熔化,再徐徐冷却使之凝固成规整晶体的热力学过程,属于热力学与统计物理研究的范畴。

热力学与统计物理所研究的对象,通常称为热力学系统,是指在给定范围内,由大量微观粒子所组成的宏观物体。如气体、液体、固体、等离子体等。对同一研究对象,热力学与统计物理从不同角度加以研究。热力学从经验总结出的定律出发,找出系统宏观量之间的联系以及宏观量变化的规律;统计物理学从物质的微观结构出发,把宏观量作为相应微观量的统计平均值来计算,可以从理论上计算某些宏观量及其涨落,因此更能反映热运动的本质。

#### 3.1 固体退火过程的物理图像

在加热固体时,固体粒子的热运动不断增强,随着温度的升高,粒子与其平衡位置的偏离越来越大。当温度升至熔解温度后,固体的规则性被彻底破坏,固体熔解为液体,粒子排列从较有序的结晶态转变为无序的液态,这个过程称为熔解。熔解过程的目的是消除系统中原先可能存在的非均匀状态,使随后进行的冷却过程以某一平衡态为始点。熔解过程与系统的熵增过程相联系,系统能量也随温度升高而增大。

冷却时,液体粒子的热运动渐渐减弱,随着温度的徐徐降低,粒子运动渐趋有序。当温度降至结晶温度后,粒子运动变为围绕晶体格点的微小振动,液体凝固成固体的晶态,这个过程称为退火。退火过程之所以必须“徐徐”进行,是为了使系统在每一温度下都达到平衡态,最终达到固体的基态。退火过程中系统的熵值不断减小,系统能量也随温度降低趋于最小值。冷却时若急剧降低温度,则将引起淬火效应,即固体只能冷凝为非均匀的亚稳态,系统能量也不会达到最小值。

退火过程中系统在每一温度下达到平衡态的过程,可以用封闭系统的等温过程来描述。根据 *Baltamann* 有序性原理,退火过程遵循应用于热平衡封闭系统的热力学定律—自由能减少定律:

“对于与周围环境交换热量而温度保持不变的封闭系统,系统状态的自发变化总是朝着自由能减少的方向进行,当自由能达到最小值时,系统达到平衡态”。

系统的自由能  $F = E - TS$ , 其中  $E$  是系统的内能,  $T$  是系统温度,  $S$  是系统的熵。设  $i$  和  $j$  是恒温下系统的两个状态, 即

$$F_i = E - TS_i \text{ 和 } F_j = E - TS_j$$

而

$$\Delta F = F_j - F_i = (E_j - E_i) - T(S_j - S_i) = \Delta E - T\Delta S$$

若系统状态由  $i$  自发变化到  $j$ , 则应有  $\Delta F < 0$ 。显然, 能量减少 ( $\Delta E < 0$ ) 与熵增加 ( $\Delta S > 0$ ) 有利于自发变化。因此任一恒定温度下, 系统状态从非平衡态自发变化到平衡态, 都是能量和熵竞争的结果, 温度决定着这两个因素的相对权重。在高温下, 熵占统治地位, 有利于变化的方向就是熵增加的方向, 因而显出粒子的无序状态。而低温对应于低熵, 低温下能量占优势, 能量减少的方向有利于自发变化, 因而得到有序(低熵)和低能的晶体结构。在这种结构内部, 每个拉子都被它与相邻粒子间相互作用的势能“囚禁”着。

系统在恒定温度下的自由能  $F$  是“热力学势”的例子。热力学势  $F$  所取的极小值确定吸引中心态, 系统将会自发地趋向这些态。一旦达到吸引中心态, 系统只能在它附近涨落。

在平衡态下出现的平衡结构可以看作是大量微观拉子活动的统计抵偿的结果, 这些平衡结构都是确立在分子水平上, 作用于大约  $10^{-3}$  厘米数量级(这是分子中原子直径的数量级)范围内分子间的相互作用, 使得晶体结构稳定, 并赋予它宏观性质。

## 3.2 统计物理的基本假设

### 3.2.1 系统的宏观状态与微观状态

系统的宏观状态是指系统的热力学状态。处于平衡态的系统, 用体积、温度、压强、总能量等宏观量来描述。系统的微观状态是指系统的动力学状态。若系统由  $N$  个粒子组成, 每个粒子的自由度为  $s$ , 则系统的微观状态可用  $N_s$  个广义坐标与  $N_s$  个广义动量组成的  $2N_s$  维的相空间来描述, 相空间中的一点代表系统的一个动力学状态, 也即系统的一个微观状态。

平稳态是系统的一种宏观状态, 可以对应各种不同的微观状态。系统平

衡时, 宏观状态已经确定, 但组成系统的粒子的动力学状态却在不断变化, 使得系统的微观状态不断变化, 而这一系列不同的微观状态都属于同一宏观状态.例如, 系统第  $i$  个粒子的状态与第  $j$  个粒子的状态发生交换. 对系统来说, 这是两种不同的微观状态, 在相空间中以不同的代表点来代表, 但它们又属于同一种宏观状态。

### 3.2.2 系综

统计物理学首先要解决的是:具有一定宏观状态的系统, 处于某一微观状态或某一微观状态附近的几率是多少?

Gibbs 引进的系综概念可以形象地表示系统处于某一微观状态或某一微观状态附近的几率.系综是大量系统的集合, 这些系统具有相同的粒子数, 相同的化学性质和相同的宏观状态, 其微观状态则按各自的统计规律分布着.如果真实系统存在  $M$  个微观状态, 则该系综便由  $M$  个系统组成, 真实系统的每个微观状态都与系综中的一个系统相对应, 对系综中所有系统取平均(称为系综平均)就能正确给出真实系统的时间平均值.因此, 只要将系综中所有系统的代表点画到相空间中, 则代表点的密度分布就形象地表示了真实系统处于某一微观状态附近的几率。

### 3.2.3 微正则系综与等几率原理

微正则系综由这样一些系统组成, 它们具有确定的粒子数  $N$  和确定的体积  $V$ , 系统与外界的联系很微弱, 使得系统的能量只能在  $E \rightarrow E + \Delta E$  之间变动, 因而系统基本上是孤立的.微正则系综描述了一孤立系统的统计性质:系统的状态只能处于限定的能量区间  $E \rightarrow E + \Delta E$  之内, 处在这个能量区间的任一微观状态出现的几率相等, 而能量区间外的任一微观状态都是不可能出现的, 这就是统计物理的基本假设—等几率原理:

对于处在平衡态的孤立系统, 系统处于每一微观状态的几率是相等的。

L3 正则系综的分布函数。

为了描述固体退火过程的统计性质, 设想固体与一热源接触而处于平衡状态,  $I$  为所研究的系统,  $II$  为热源. 系统  $I$  可以与热源交换能量, 但系统  $I$  的拉子数和体积不变。

正则系综由这样一些系统组成, 它们具有确定的拉子数  $N$  和确定的体积  $V$ , 系统与一热源(恒温)相接触而平衡.正则系综是能量可变, 但拉子数守恒的系统的统计分布.利用微正则系综的等几率分布可以导出正则系综的几率

分布。

设  $E_1$  代表系统 I 的能量,  $E_2$  代表热源 II 的能量, 且  $E_1 \ll E_2$ 。将 I 和 II 合在一起成为 (I + II) 大系统, 这大系统是一个孤立系统, 其能量为:

$$E \rightarrow E + \Delta E, \quad \Delta E \rightarrow 0$$

$$\text{而 } E = E_1 + E_2$$

这里忽略了系统 I 与热源 II 之间的相互作用能里。(I + II) 大系统既然是一孤立系统, 则其统计性质就可以用微正则系综来描述, 即处于任一微观状态的几率相等。

系统 I 处在能最为  $E_i$  的一个微观状态  $i$  上时, 热源 II 可以处在能量为  $E - E_i \rightarrow E - E_i + \Delta E$  的所有微观态上, 这些微观态以  $Q_2(E - E_i)$  来表示, 则大系统的微观态数是

$$1 \times Q_2(E - E_i) = Q_2(E - E_i)$$

若系统 I 的  $E_i$  能量的简并度(相同能量的状态数, 称为该能量的简并度)为  $g(E_i)$ , 则大系统 (I + II) 的总微观态数  $Q(E)$  为:

$$Q(E) = \sum_i g(E_i) Q_2(E - E_i) \quad (3-1)$$

由等几率原理可知, 系统 I 处在  $E_i$  某一微观态  $i$  的几率为:

$$P_i = \frac{Q_2(E - E_i)}{Q(E)} \quad (3-2)$$

其中  $Q(E)$  是一常数。由于  $E_i \ll E$ , 可以将  $\ln Q_2$  展开为  $E_i$  的 Taylor 级数:

$$\ln Q_2(E - E_i) = \ln Q_2(E) - E_i \left( \frac{\partial \ln Q_2}{\partial E_2} \right)_{E_1=E} \quad (3-3)$$

其中  $Q_2(E)$  为热源作为一孤立系统的状态数。按熵的定义, 热源的熵  $S_2$  为:

$$S_2(E) = k \ln Q_2(E)$$

其中  $k$  为 Boltzmann 常数, 而

$$\left( \frac{\partial S_2}{\partial E_2} \right)_V = \frac{1}{T}$$

于是

$$\left( \frac{\partial \ln Q_2}{\partial E_2} \right) = \frac{1}{kT}$$

(3-3) 式子变为



$$\ln Q_2(E - E_i) = \frac{S_2(E)}{k} - \frac{E_i}{kT}$$

即有

$$Q_2(E - E_i) = \exp\left(\frac{S_2(E)}{k}\right) \exp\left(\frac{-E_i}{kT}\right)$$

代入(3-2)式得

$$P_i = A \exp\left(\frac{-E_i}{kT}\right) \quad (3-4)$$

其中  $A$  是与  $E_i$  无关的常数。取系统 I 中所有可能状态的几率总和，称为几率的归一化，即  $\sum_i P_i = 1$ 。 $\sum_i$  表示对系统 I 所有微观态求和。于是

$$A \sum_i \exp\left(\frac{-E_i}{kT}\right) = 1$$

令

$$Z = \sum_i \exp\left(\frac{-E_i}{kT}\right)$$

为系统的配分函数，得

$$A = \frac{1}{Z}$$

代入(3-4)式，得

$$P_i = \frac{1}{Z} \exp\left(\frac{-E_i}{kT}\right) \quad (3-5)$$

其中  $\exp\left(\frac{-E_i}{kT}\right)$  称为 Boltzmann 因子， $T$  是绝对温度， $k$  是 Boltzmann 常数。

Metropolis (3-5) 式的分布称为 Gibbs 正则分布。这种分布给出温度  $T$  时固体处于能量为  $E_i$  的微观态  $i$  的几率。显然，固体处于能最较低的微观态的几率较大。在温度降低时，那些能量相比最低的微观态最有可能出现，当温度趋于零时，固体只能处于能最为最小值的基态上。

### 3.3 Metropolis 准则

固体在恒定温度下达到热平衡的过程可以用 Monte Carlo 方法进行模拟。Monte Carlo 方法的特点是算法简单，但必须大量采样才能得到比较精

确的结果, 因而计算量很大。

从物理系统倾向于能  $t$  较低的状态, 而热运动又妨碍它准确落入最低态的物理图像出发, 采样时着重取那些有重要贡献的状态, 则可以较快地达到较好的结果。

1953 年, *Metropolis* 等提出重要性采样法。他们用下述方法产生固体的状态序列:

先给定以粒子相对位置表征的初始状态  $i$ , 作为固体的当前状态, 该状态的能量是  $E_i$ 。然后用摄动装置使随机选取的某个粒子的位移随机地产生一微小变化, 得到一个新状态  $j$ , 新状态的能量是  $E_j$ , 如果  $E_j < E_i$ , 则该新状态就作为“重要”状态, 如果  $E_j > E_i$  则考虑到热运动的影响, 该新状态是否“重要”状态, 要依据固体处于该状态的几率来判断。固体处于状态  $i$  和  $j$  的几率的比值等于相应 *Boltzmann* 因子的比值, 即

$$\gamma = \exp\left(\frac{E_i - E_j}{\kappa T}\right) \quad (3-6)$$

$\gamma$  是个小于 1 的数。用随机数发生器产生一个  $[0,1)$  区间的随机数  $\xi$ , 若  $\gamma > \xi$ , 则新状态  $j$  作为重要状态, 否则舍去。

若新状态  $j$  是重要状态, 就以  $j$  取代  $i$  成为当前状态, 否则仍以  $i$  为当前状态。再重复以上新状态的产生过程。在大量迁移(固体状态的变换称为迁移)后, 系统趋于能量较低的平衡状态, 固体状态的概率分布趋于式(3-5)的 *Gibbs* 正则分布。

由(3-6)式可知, 高温下可接受与当前状态能差较大的新状态为重要状态, 而在低温下只能接受与当前状态能差较小的新状态为重要状态, 这与不同温度下热运动的影响完全一致。在温度趋于零时, 就不能接受任一  $E_j > E_i$  的新状态  $j$  了。

上述接受新状态的准则称为 *Metropolis* 准则, 相应的算法称为 *Metropolis* 算法, 这种算法的计算量显著减少。

### 3.4 基于模拟退火算法 Metropolis 准则的连接建立

#### 3.4.1 模拟退火算法的特点:

模拟退火算法具有十分顽强的全局搜索性,这是因为比起普通的优化搜索方法,它采用了许多独特的方法和技术,归纳起来主要有以下两个方面。

(1)在模拟退火算法中,基本不用搜索空间的知识或者其它的辅助信息(启发式搜索算法需用),而只是定义邻域结构,在其邻域结构内选取相邻解,再用目标函数进行评估。

(2)模拟退火算法不是采用确定性规则,二是采用概率的变迁来指导它的搜索方向的。它所采用的概率仅仅是作为一种工具来引导其搜索过程朝着搜索空间的更优化的解区域移动。因此,虽然看起来它是一种盲目的搜索方法,但实际上有着明确的搜索方向。

以往的排料系统中,遗传算法应用的很多,相对也较成熟,模拟退火算法在自动排料系统的研究处于发展阶段,有研究的空间;还有相对于遗传算法,模拟退火算法无需进行编码的操作,减少了计算的复杂性。

#### 3.4.2 模拟退火算法实现的技术问题

##### (1)解的形式和邻域结构

在组合优化问题中,解的表现形式直接决定于邻域的结构。例如,在排料问题中,解的形式和邻域结构都是指零件所能达到的任正数二,当温度  $t$  时,算法停止,表示已经到达最低温度。

##### (2)循环总数控制法

即总的温度下降次数为一定值  $K$ ,当温度迭代次数达到  $K$  时,停止运算。整个算法的总迭代步数为一定数,它表示各温度时马氏链迭代数(内循环)的总和为一个给定的数,这样很容易计算算法的复杂性,但需要合理分配内循环的长度和温度下降的次数。

(3)基于不进规则的控制法在一个温度和给定的迭代次数内没有改进当前的局部最优解,则停止运算。模拟退火的一个基本思想是跳出局部最优解,则在低的温度跳出最优解的可能也比较小。由此产生上面的停止原则。

#### 3.4.3 模拟退火算法参数实际选择

理论上是用一个马尔可夫链描述模拟退火算法的变化过程,因此具有全局最优性。它有诸多的参数需要调整,如起始温度、温度下降的方案。这样

需要人为地调整,人为的因素,如对问题的了解,参数和规则的搭配等造成计算结果的差异。解决这一矛盾的方法主要通过大量的数值模拟计算,从中选择比较好的参数搭配。

#### 3.4.4 基于模拟退火算法 Metropolis 准则的连接建立

在本章上面的部分介绍了模拟退火算法,现在本文作者提出基于模拟退火算法的连接建立的模型如下:

由于模拟退火算法逃离局部最优,而达到全局最优的特点,我们则可以选出在 P2P 流媒体网络选择出节点性能最好的作为新加入节点的连接对象。现在的问题是评价 P2P 流媒体网络节点的性能,我们采用什么参数作为判断标准?

如果能够测出网络节点的综合性能,即既考虑节点的带宽也考虑节点主机自身的性能那样最好,但在实际网络中测出网络的带宽或可用带宽比较好测,而要测出节点的主机性能比较难,于是本文认为用节点主机的可用带宽与节点总带宽的比值最为参数较好。

这样,当有一个新的节点要求加入 P2P 流媒体网络时, P2P 流媒体网络中已有的节点就作为转移的状态,每个节点的可用带宽与总带宽的比值最为转移的根据,一般的情况下总是从比值小的转移到比值大的节点上。

上面讲了以什么来作为转移个条件,但是在有的时候比值大的节点不一定比比值小的节点的快,因为有可能该节点的网络在某出现瓶颈或主机自身性能的差别。为此当一个比值大的和比值小的节点时,对这两个节点做归一化处理,这两个比值之和为分母,比值小的最为分子,这样得到一个小于 1 的数;再生成一个小于 1 的随机数,如果生成的随机数比刚才产生的值小,就从上面定义的比值大的转移到比值小的节点上,描述如下:

设 P2P 流媒体中  $i$  的可用带宽与总带宽的比值为  $P_i$ , 则从节点  $i$  转移到节点  $j$  的条件是:

- (1) 如果  $P_j > P_i$ , 则从  $i$  节点转移到  $j$  节点。
- (2) 如果  $P_j \leq P_i$ , 且  $P_j > \text{random}(0,1)$ , 则从  $i$  节点转移到  $j$  节点。
- (3) 其它的则不考虑  $j$  节点, 继续判断其他节点。

这样就得到通过模拟退火算法的 Metropolis 准则建立 P2P 流媒体网。

### 3.5 本章小结

在本章中, 通过对模拟退火算法以及其中的 Metropolis 准则的介绍,

得到了一种基于模拟退火算法 Metropolis 准则的 P2P 流媒体网络连接的准则，对新加入节点来说，这种准则能够选择出理论上最优的连接。

## 14 复杂网络及其建模

在现实的 P2P 流媒体网络中,大都是数网络并不是完全对等的,其中有源节点,有事先随机部署的节点,其中这些节点的综合性能比较好。于是在构建 P2P 流媒体网络时,怎么样建立才能满足实际网络节点的不均衡性呢?下面来介绍一下复杂网络的基本概念,以及复杂网络的建模方式。

在实际的 P2P 流媒体网络中,大多是混合型 P2P 流媒体网络,整个网络结构看似对等的点到点,实际上是有些节点的性能和网络带宽资源占有很大的优势,这些占优势的节点就是 P2P 中的强管理节点,那么在这种情况下,我们对连接节点的分配不可能完全“平等”,那么怎么分配才能是真正的平等使网络资源充分利用呢?为此我们认为 P2P 网络是复杂网络的概念,本文想通过复杂网络的建模方式,使我们的连接满足网络拓扑以及资源分配的实际情况,即下来本文则阐述一下复杂网络以及它的建模方式。

复杂网络的研究近年来受到科学和工程各个领域研究人员的广泛关注,已经成为近年来一个研究热点。网络由一些基本单元(通常我们称之为节点或顶点)和它们之间的连接(通常我们称之为边或连接)所组成。图 4-1 至图 4-2 是一些典型的复杂网络图。

传统上对网络的研究一般采用数学上的图论的方法,但经典图论所考虑的一般是规则图。在上个世纪 50 年代末 60 年代初,两位匈牙利数学家 Paul Erdos 和 Alfred Renyi (ER)在图论领域做出了一个重要突破,随机图理论。他们用随机图来描述网络的拓扑结构,这为复杂网络的研究奠定了一个数学理论基础。用随机图理论对复杂网络的研究持续了 40 年,直到今天仍有人在做这方面的研究。但随机图理论是否能准确地描述现实世界中的网络呢?大量的观察结果告诉我们:现实世界中的网络既非完全规则亦非完全随机。ER 随机图理论虽然不能准确地描述现实世界中的网络,但却能统治这个领域这么多年,主要还是由于过去我们缺乏对现实世界中网络数据的分析能力和对现实世界网络拓扑结构的准确认识。

但近年来事情发生了变化,由于科学技术领域,特别是 IT 领域的高速发展,使我们获得的可供我们刻画现实世界网络特征的数据越来越多,加之超级计算机的发展为我们提供了强大的计算和数据分析能力,以及学科之间的相互交叉和融合趋势在不断加强,使得人们有能力在对各种不同类型网络的数据分析的基础上,揭示复杂动力网络的一些共有的特征和性质,从而激发起人们以理论、仿真和实际数据验证三种手段研究复杂网络的浓厚兴

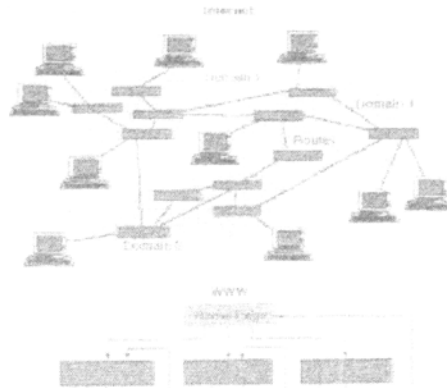


图 4-1 Internet 和 WWW 网络结构示意图 [14]

Figure 4-1 The sketch map of Internet & www network [14]

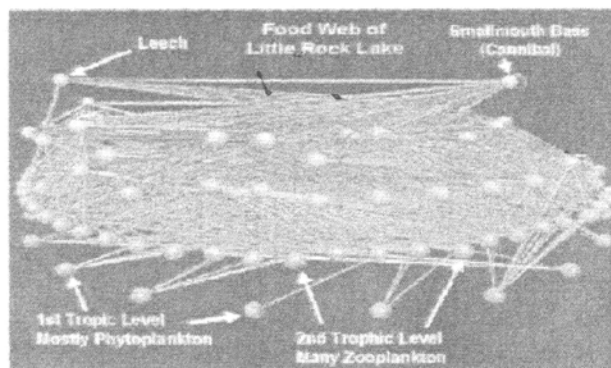


图 4-2 清水湖中各个种群间扑杀猎食交互网络 [19]

Figure 4-2 The food chain network among different races in clear lake [19]

趣。以此为契机，经过在这一研究方向上不懈的探索和努力，最近在复杂网络研究领域取得了两项重要的发现：大多数复杂网络都具有小世界 (small-world) 效应和标度无关 (scale-free) 特性。小世界的概念从字面上的意思是说尽管网络很大，但通常在网络的任意两个节点间都有一条相对于网络规模来讲是很短的路径。我们可能经常遇到这样的情形，当你遇到一个陌生人的时候，经过交谈你却发现你们有共同的好朋友（他是你朋友的朋友），于是双方都会发出这样的惊叹“世界真小”。小世界的另一个代名词是所谓的“六度分离” (six degree of separation)。1967 年社会心理学

家 Stanley Milgram 断言在大多数情况下两个美国人之间一般存在一条不超过 6 的“认识”路长(即通过一个人与另一个人之间互相认识来构成的路径,在任意两个美国人之间该路径长度一般不超过 6)。1998 年,为了描述从规则格子到随机图之间的转变,Watts 和 Strogatz 提出了小世界网络的概念。

在复杂网络中还有另一项重要的发现。Barabasi 和 Albert 率先发现在很多大规模的复杂网络中节点的度呈现一种幂律(power-law)分布。该类复杂网络中的一些基本概念尽管近年来提出了许多刻画和度量网络的概念。但其中有三个概念发挥了重要作用,它们是:平均路长、聚类系数和度分布。事实上最初 Watts 和 Strogatz 就是试图构造一种具有小的平均路长(与随机网络类似)和较大的聚类系数(与规则网络类似)的网络,循此途径最后得到了小世界网络模型。标度无关网络的提出是由于观察到许多大规模网络的节点的度呈现一种幂律分布。所以说这三个概念在复杂网络的奠基性研究中起到了至关重要的作用。

#### 4.1 复杂网络中的一些基本概念

尽管近年来提出了许多刻画和度量网络的概念。但其中有三个概念发挥了重要作用,它们是:平均路长、聚类系数和度分布。事实上最初 Watts 和 Strogatz 就是试图构造一种具有小的平均路长(与随机网络类似)和较大的聚类系数(与规则网络类似)的网络,循此途径最后得到了小世界网络模型。标度无关网络的提出是由于观察到许多大规模网络的节点的度呈现一种幂律分布。所以说这三个概念在复杂网络的奠基性研究中起到了至关重要的作用。并且在后来的复杂网络研究中它们也扮演了重要的角色。因此我们下面先介绍一下这些概念。

##### (1) 平均路长(average path length)

在一个网络中,两个节点  $i$  和  $j$  之间的距离  $d_{ij}$  定义为连接  $i$  和  $j$  的最短路径所包含的边的数目。网络中任意两个节点间的距离的最大值称为网络的真径  $D$ 。一个网络的平均路长  $L$  定义为网络中任意两个点间距离的平均值。一个有趣的发现是在很多大规模复杂网络中,平均路长相对都很小,即使在节点间连接很稀疏的网络中情况也是如此。这个“小”也就是所谓的小世界效应,也就是小世界网络名称的来源。

##### (2) 聚类系数(clustering coefficient)



在朋友网络中，经常会遇到这样的现象，你朋友的朋友还是你的朋友，或者换句话说，你的两个朋友，他们之间也是朋友。这种特性被称为聚类特性，而聚类系数就是用来衡量大家的朋友之间互相还是朋友的比率有多大的。更确切地说，聚类系数可以由下面的方式定义。假设节点  $i$  连接着  $k_i$  条边，并通过他们连接到  $k_i$  个其它的节点。这些节点都是节点  $i$  的邻接节点。易知最多可能存在有  $k_i(k_i-1)/2$  条边在这些邻接节点之间，此时这些邻接节点之间是全部互相连接的。而实际在这些节点之间存在的边的数目  $E_i \leq [k_i(k_i-1)/2]$ 。节点  $i$  的聚类系数就定义为  $E_i/[k_i(k_i-1)/2]$  的比值。整个  $P \ll 1, 1/N$  网络的聚类系数  $C$  定义为网络中所有节点的聚类系数的均值，即

$$C = \frac{1}{n} \sum_i C_i \quad (4-1)$$

在具有  $N$  个节点连接概率为  $P$  的随机网络中 (通常  $P \ll 1, P \gg 1/N$ )， $C \approx P$ ；

在全局耦合网络中  $C=1$ 。而在实际网络中  $C$  通常远大于  $1/N$ ，但却小于 1。所以从这个意义上讲，随机网络和全局耦合网络都不能很好地描述现实世界的网络。

### (3) 度分布 (degree distribution)

刻画一个节点的特性的可能最简单同时也是最重要的概念就是度。一个节点  $i$  的度  $k_i$ ，定义为它的连接 (边) 的数目。也就是说，一个节点的度越大，那么他在网络中的重要性越高。一个网络中节点度的分布可以用一个分布函数  $P(k)$  来刻画， $P(k)$  定义为一个随机选择的节点恰好具有  $k$  条边的概率。易知，全局耦合网络的度分布很简单，就是一个 Delta 函数，因为在全局耦合网络中所有节点都和其它所有节点连接，每个节点所连接的边数都相等。随机网络和小世界网络的度为泊松分布。而通过对许多实际的大规模网络的统计分析发现，大多数实际网络的度呈一种幂律分布，也就是分布函数的形式为  $P(k) \sim k^{-\tau}$ 。服从这种分布的网络我们称之为无标度网络。我们可以通过比较美国的高速公路网络和航空线路网络来直观地感受度分布为 Poisson 分布和幂律分布的网络的差别。

## 4.2 复杂网络的模型

### 4.2.1 规则网络模型

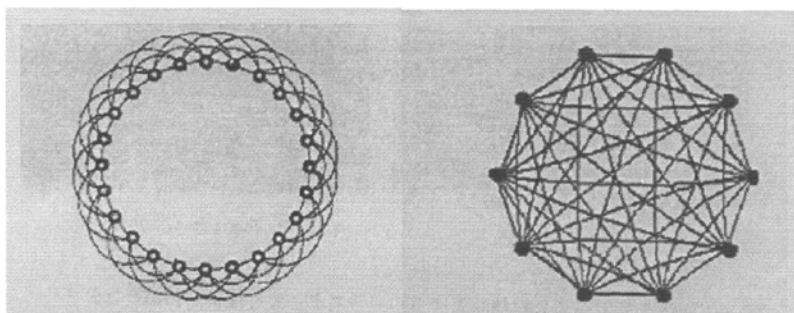


图 4-3 最近邻网络

图 4-4 全局耦合网络

Figure 4-3 Nearest Neighbor network

Figure 4-4 Global-coupling network

两种最常被研究的规则网络模型是最近邻网络和全局耦合网络

(1)最近邻网络如图 4-3 所示的是一个最近邻网络模型，它是经常被研究的一个网络模型，通常也被称为格子(lattice)。易知，当网络的节点数目很大时，最近邻网络的平均路长也很大。当节点数  $N \rightarrow \infty$  时，平均路长也趋于无穷大。所以最近邻网络不具有小世界效应。

(2)全局耦合网络模型如图 4-4 所示。在该网络中，每个节点都和其它所有节点连接，网络的聚类系数为 1。在一个具有  $N$  个节点的网络中共有  $N(N-1)/2$  条边。但在实际网络中，边是比较稀疏的，一般网络具有的边的数目是  $N$  的量级而不是  $N^2$  的量级。

### 4.2.2 随机图模型 ( $p=1$ )

如前所述，ER 随机图理论 40 多年前由 Erdős 和 Renyi 提出。假设网络有  $N$  个节点，我们以概率  $P$  来连接一对随机选定的节点。这样就生成了一个具有  $N$  个节点和大约  $pN(N-1)/2$  条边的随机图。图 4-7 示出了随机图网络的演化过程。随机图理论研究的主要问题是确定随机图产生某种特定性质的概率值。随机图网络的平均度是  $\langle k \rangle = p(N-1)$ ，度的分布呈 Poisson 分布。平均路长  $L \sim \ln N / \langle k \rangle$  较小，具有典型的小世界特性。但 ER 随机图的聚类系数  $C \sim 1/N$ ，而实际网络的聚类系数通常远大于这个数值。

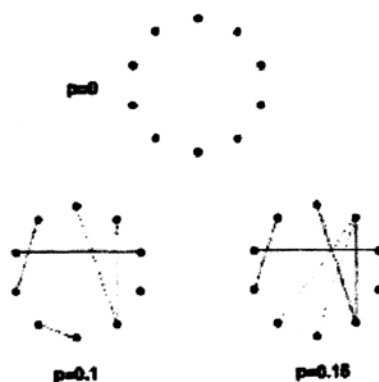


图 4-5 ER 随机图的演化过程

Figure 4-5 The evolution process of ER random graph

### 4.2.3 小世界网络模型

通过上面的分析我们看到在规则的最近邻网络中不具有小世界特性。在全局藕合网络中具有较大的聚类系数和小世界特性，但边连接过于稠密。在随机图网络中具有小的平均路长，但却不具有实际网络的聚类特性，聚类系数过小。而实际网络通常是既具有某种规则性，又有一些随机性。例如，人们通常都认识自己的邻居和住在附近的人，也有可能认识住得很远处的某些人。为了描述从规则格子到随机网络之间的转变，1998 年 Watts 和 Strogatz 引入了一种很有趣的所谓的小世界网络模型，我们称之为 WS 小世界模型。WS 小世界网络模型可以通过如下的算法形成，如图 4-6：

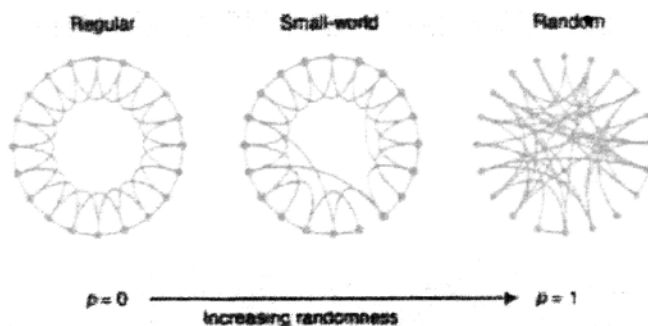


图 4-6 小世界网络的演化过程

Figure 4-6 The evolution process of small-world network

(1)从规则开始:从一个具有  $N$  个节点, 每个节点具有  $k$  个近邻( $k$  为偶数, 每边  $k/2$  个)的最近邻网络开始。为了使网络连接具有稀疏性, 要满足  $N \gg k$ 。

(2)随机化:以概率  $P$  “重新连接”(rewire)每条边, 重新连接在这里的意思是把一条边的一端从一个节点转移到在网络中随机选取的另一个节点上, 并保证节点没有自连接以及两个节点间没有重复连接。这样网络中就产生了  $pNk/2$  条“长距离连接”(或称“非局部连接”、“捷径”)。通过变动  $p$ , 我们可以产生从规则格子( $p=0$ )到随机网络( $p=1$ )的转变。

我们可以把聚类系数  $C(p)$  和平均路长  $L(p)$  看成是概率  $P$  的函数。规则的最近邻网络( $P=0$ )是高聚类的, 聚类系数  $C(0) \approx (k-1)/k$ , 但具有长的平均路长( $L(0) \approx N/2k \gg 1$ )。研究发现, 对于一个较小的重新连接概率  $P$ , 聚类系数改变很小, 而平均路长却减小很快。

注意到在 WS 小世界网络模型中, 存在通过重新连接而把整个网络断开成不连接的几部分的可能性。为了克服这个问题使得研究更为方便, Newman 和 Watts 在对这个小世界网络模型做了一点修改, 提出了 NW 小世界网络模型。在这个模型中, 不断开原有的连接, 而是以概率  $P$  在随机选定的一对网络节点间增加新的连接。同样, 也要保证节点没有自连接以及两个节点间没有重复连接。当  $p=0$  时, NW 网络就是原来的规则最近邻网络; 当  $p=1$  时, 网络变成规则的全局藕合网络。研究发现, 在  $P$  较小时, NW 模型具有和 WS 模型类似的特性。

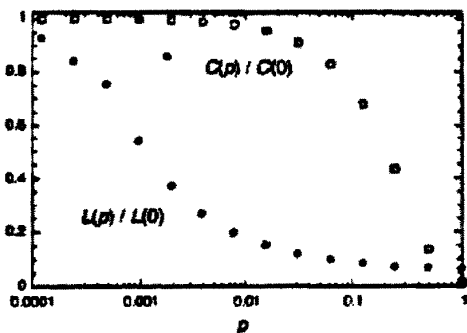


图 4-7 WS 小世界网络中  $C(p)$  与  $L(p)$  随  $p$  的变化 [22]

Figure 4-7 The change of  $C(p)$  &  $L(p)$  along with  $p$  in WS small-world network [22]

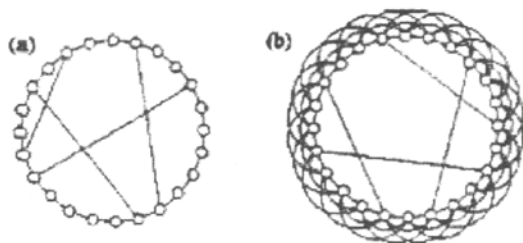


图 4-8 具有四条长距离连接的 NW 小世界网络模型, 其中初始的最近邻网络中的近邻数目为 (a)  $k=2$ , (b)  $k=6$  [23]

Figure 4-8 The small-world network model with 4 long distance connection, the initial number of nearest neighbors in nearest neighbor network is: (a)  $k=2$ , (b)  $k=6$  [23]

#### 4.2.4 标度网络模型

上述网络模型中规则网络的度分布为 Delta 函数, 因为所有的网络节点都有相同的连接数。ER 随机图网络和小世界网络的度分布都是 Poisson 分布。该分布在度的均值处有一个峰值, 在两侧呈指数衰减, 因此这样的网络也被叫做指数网络(exponential networks)。而最近研究发现, 实际的很多大规模复杂网络的度分布都服从幂律(power-law)分布。在图 4-11 中我们给出了一个典型的具有幂律度分布的网络模型。

为了解释这种幂律分布, Barabasi 和 Albert (BA) 提出了无标度网络模型。BA 无标度网络模型的演化生成主要有两个要点: 增长((growth)和偏好性连接(preferential attachment)。他们认为其它大多数复杂网络模型都忽略了这两点。首先, 大多数网络都是开放的, 不断有新的节点加入的, 例如, 在 WWW 中不断有新的网页加入, 在科研合作网络中不断的有新的研究者加入到某领域的研究中。但规则网络、ER 随机图网络和小世界网络等网络模型却是静态的网络模型, 他们考虑的都是具有固定不变节点数目的网络。另外, ER 随机图网络和小世界网络模型中考虑加入连接和重新连接是等概率一致分布的。但现实中很多网络却并非如此, 例如, 在 WWW 中, 新加入的网页或网站更愿意连接到已经存在的点击率高的著名站点上, 在科研合作网中, 新加入的研究者更愿意与已经在该领域有影响力的已成名的研究者合作。这种偏好性会导致所谓的“富有者会更富(richest richer)”的现象。BA 无标度网络模型包含了增长和偏好性连接, 在 BA 网络模型当中, 不断地

有新的节点加入,新加入的节点连接到具有连接数越大的节点的的概率越高。BA 无标度网络模型的生成算法为阵:

(1)增长:初始时为  $m_0$  个节点,在每步增加一个新的节点,该节点连接到网络中  $m \leq m_0$  个已存在的节点上。

(2)偏好性连接:当在网络中选择新增加的节点所要连接的  $m$  个节点时,我们根据概率  $\Pi$  来选择,假设第  $i$  个节点此时的度为  $k_i$ ,那么该节点被选中的概率定义为

$$\Pi(k_i) = \frac{k_i}{\sum_j k_j} \quad (4-2)$$

这样经过  $t$  步之后,就生成了一个具有  $N=t+m_0$  个节点和  $mt$  条边的网络。直观上我们可以知道,在这样的网络中,有少数节点会有很高的度(即有很多个节点连接到它上面),而大多数节点的度都比较低,这 and 实际网络的情况比较类似,比如,在 WWW 中象 Yahoo 和 CNN 这样的网站有很多连接,但大多数网站的连接数都比较少。数值仿真发现,该网络模型生成的网络的度分布确实为幂律分布。

### 4.3 基于复杂网络的无标度网络模型的 P2P 流媒体网络建模

在现实的 P2P 网络中,大多都是混合型的 P2P 网络模型,即就是说网络中的节点不是绝对的对等,有一部分是强管理节点。比如实际中的直播系统或则大的 P2P 应用系统,他们都是在各大地区放有自己的强管理节点,这些强管理节点不仅计算机性能好而且网络带宽大也比较流畅,那么这些强管理节点连接的流媒体用户就应该比较多,按照复杂网络的无标度模型,即就是这些节点的度大。反过来,为什么这些节点的度大呢?就是因为这些节点的性能好,网络流畅,所以它们的度大。从某种意义上讲,网络节点的度的大小一定程度上反映本身的性能以及网络流畅成都和网络带宽的大小。

于是本文作者认为建立一个无标度的复杂网络模型更能与实际中的网络资源相符,且这样的 P2P 流媒体网络模型不但能与实际的网络相符,更能保证 P2P 流媒体网络的健壮性。在此,将通过建立无标度的复杂网络模型的方法在这些节点中选择一个节点使他既考虑网络带宽又考虑网络的流畅程度以及节点的自身性能。

无标度 P2P 流媒体网络模型的算法过程如下:

(1)最先设置随机地部署些网络带宽和主机性能的比较好的节点,这些

节点的连接完全随机。

(2) 按照复杂网络的无标度网络建模方式，对后面来的节点按照偏好性连接建立连接，即那些度大的节点被选中的概率较大。

(3) 偏好性连接使用轮盘赌的方法建立。

通过上面的方法，就可以从一个源接点开始逐步建立 P2P 流媒体网络，此方法主要是尊重“富有者更富有的”的原则，从而实现网络节点间真正的对等。

5 系统的模拟仿真分析设计

在前面的章节中，我们已经阐述了基于模拟退火算法中 Metropolis 准则构建 P2P 网络的理论，以及基于复杂网络模型构建 P2P 网络所采用的建模方法的选择，接下来，本章将对两种建模理论建立的 P2P 网络进行仿真分析，首先通过模拟退火算法中的 Metropolis 准则构建，其次通过复杂网络模型的无标度建模方法构建。我们采用 MATLAB7.0 作为仿真平台，并以网络稳定性及网络延迟作为指标，比较两种 P2P 网络的性能。

5.1 基于 Metropolis 准则构建的 P2P 网络

设 P2P 流媒体中节点  $i$  的可用带宽与总带宽的比值为  $P_i$ ，则从节点  $i$  转移到节点  $j$  的条件是：

- (1) 如果  $P_j > P_i$ ，则从  $i$  节点转移到  $j$  节点；
- (2) 如果  $P_j \leq P_i$ ，且  $P_j / (P_i + P_j) > \text{random}(0,1)$ ，则从  $i$  节点转移到  $j$  节点；
- (3) 如果  $P_j \leq P_i$ ，且  $P_j / (P_i + P_j) \leq \text{random}(0,1)$ ，则放弃  $j$ ， $j=j+1$ ，返回(1)；

假设所有的节点的带宽都为 100M，每个节点未进入 P2P 流媒体网络之前的已用带宽是在 1M-15M 之间随机生成。其中节点的度是指该节点下游的节点个数，以下是不同规模的 P2P 流媒体网络的度分布图：

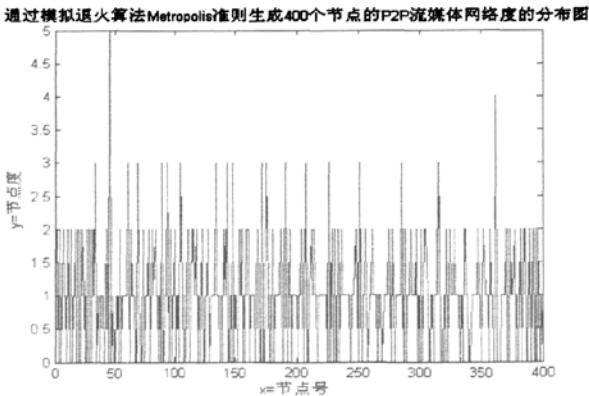


图 5-1 生成 400 个节点的 P2P 网络的度分布图

Figure5-1 The scale distribution of network with 400 nodes



通过模拟退火算法Metropolis准则生成1000个节点的P2P流媒体网络度的分布图

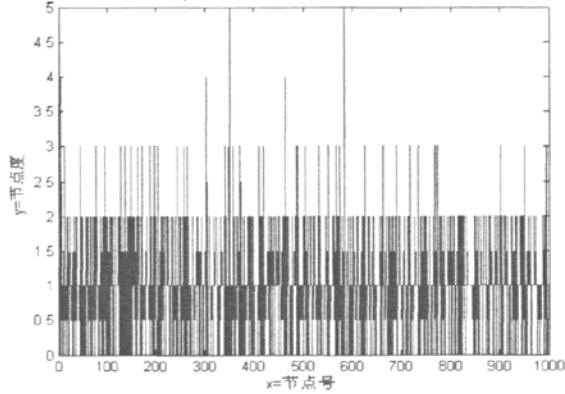


图 5-2 生成 1000 个节点的 P2P 网络的度分布图

Figure5-2 The scale distribution of network with 1000 nodes

通过模拟退火算法Metropolis准则生成2000个节点的P2P流媒体网络度的分布图

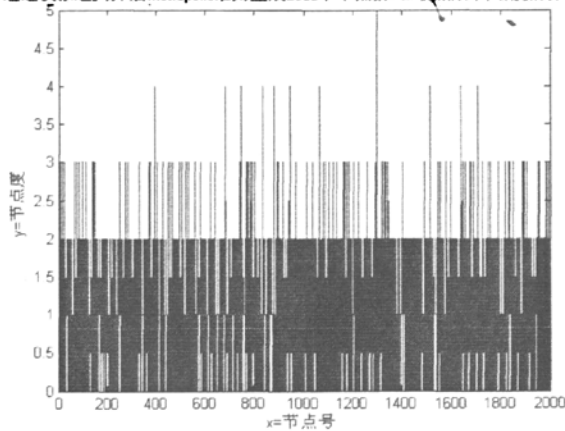


图 5-3 生成 2000 个节点的 P2P 网络的度分布图

Figure5-3 The scale distribution of network with 2000 nodes

图 5-1 至 5-3 展示了基于模拟退火算法中 Metropolis 准则构建的不同规模的 P2P 网络的度分布图。为了便于定量研究，我们同时给出所构建的以上三种不同规模网络中节点度的统计表：

表 5-1 通过模拟退火算法 Metropolis 准则生成 400 个节点的 P2P 流媒体网络的度分布表

Table5-1 The scale distribution of network with 400 nodes

度	节点数	度	节点数	度	节点数
0	118	2	76	4	1
1	187	3	17	5	1

表 5-2 通过模拟退火算法 Metropolis 准则生成 1000 个节点的 P2P 流媒体网络的度分布表

Table5-2 The scale distribution of network with 1000 nodes

度	节点数	度	节点数	度	节点数
0	293	2	201	4	3
1	464	3	37	5	2

表 5-2 通过模拟退火算法 Metropolis 准则生成 2000 个节点的 P2P 流媒体网络的度分布表

度	节点数	度	节点数	度	节点数
0	603	2	396	4	10
1	904	3	86	5	1

联合上图以及表我们可以看出，基于 Metropolis 准则构建的 P2P 网络，其单个节点的度比较小，节点之间的度差也较小，整个网络的负载比较均衡，对整体网络资源的利用比较充分。但是，对于流媒体应用而言，基于 Metropolis 准则的 P2P 网络存在很多局限性，包括：

首先，基于 Metropolis 准则构建的网络对于所有节点都一视同仁，我们无法保证带宽和性能都较优的节点的度一定大于一般的非较优节点。因此对于非较优节点，当其带宽和性能不能达到要求时，就有可能在局部造成频繁的重复性连接，引起网络局部振荡，稳定性较差。

而且，此类网络中，多数节点到源节点的跳数都较大，这会导致较大的网络延迟。流媒体网络具有较高的实时性要求，网络延迟是用户体验的重要

指标,较大的网络延迟容易导致用户的不满。

另外,在现实的网络中,节点之间的性能存在着不对等性,那么用这种方法所构建的 P2P 网络看似满足了 P2P 所期望的对等性,却在实质上造成了网络节点之间的不均衡性,从整体上降低了网络的效率。

综上所述,基于 Metropolis 原则构建的 P2P 网络无法满足流媒体网络在实际运行时的种种需求,不适合构建 P2P 流媒体网络。

## 5.2 基于复杂网络模型无标度方法构建的 P2P 网络

基于复杂网络模型无标度方法,我们可以如下构建一个 P2P 网络:

(1) 针对给定数目的,网络带宽和主机性能的比较好的节点集(较优节点集)进行随机部署,这些节点间的连接也是随机建立的;

(2) 根据复杂网络模型无标度方法的建模方式,任意新进节点均按照偏好性与其他节点连接建立连接;就是说,有较大连接度的节点被选中与新进节点连接的概率较大;

为了满足新进节点的偏好性,我们采用轮盘赌算法确定与之连接的节点。设网络中全部已有节点的连接度的和为  $sum$ ,我们将整个轮盘分为  $sum$  份,并从网络中第一个已有节点开始,标出每个节点在轮盘上所占的区间,区间大小即为该节点度的大小。之后,生成一个其值在  $[1, sum]$  间的随机数,并选择该值对应的节点作为与新进节点建立连接的对象。(最后统计的度是节点下游节点的个数)

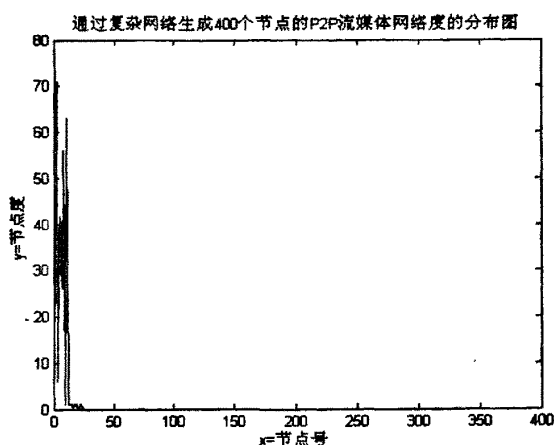


图 5-4 生成 400 个节点的 P2P 网络的度分布图

Figure5-4 The scale distribution of network with 400 nodes

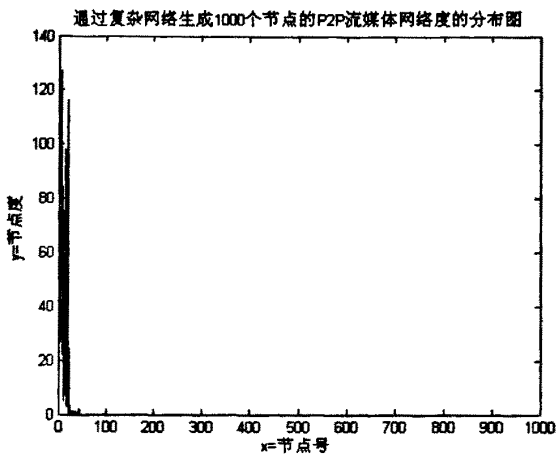


图 5-5 生成 1000 个节点的 P2P 网络的度分布图

Figure5-5 The scale distribution of network with 1000 nodes

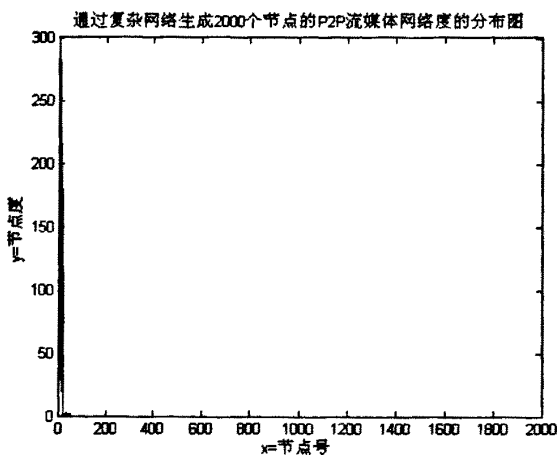


图 5-6 生成 2000 个节点的 P2P 网络的度分布图

Figure5-6 The scale distribution of network with 2000 nodes

图 5-4 至 5-6 展示了基于复杂网络模型无标度方法构建的不同规模的 P2P 网络的度分布图。为了便于定量研究，我们同时给出所构建的以上三种不同规模网络中节点度的统计表：

表 5-4 初始 30 个节点的度分布

Table5-4 The distribution of the initial peers scale

节点号	节点度	节点号	节点度
1	3	14	1
2	3	15	1
3	1	17	1
4	1	20	1
5	3		
6	5		
7	3		
8	1		
9	2		
11	1		

表 5-5 生成 400 个节点的网络度分布

Table5-5 The scale distribution of the network  
of constructing 400 peers

节点号	节点度	节点号	节点度
1	23	11	63
2	71	12	1
3	6	13	1
4	37	14	1
5	42	15	1
6	31	16	1
7	26	17	1
8	56	18	1
9	33	19	1
10	1	20	1

表 5-6 生成 1000 个节点的网络度分布  
Table6-6 The scale distribution of the network  
of constructing 1000 peers

节点号	节点度	节点号	节点度	节点号	节点度
1	89	11	75	21	3
2	69	12	9	22	1
3	27	13	7	23	3
4	37	14	46	24	1
5	55	15	98	25	1
6	127	16	32	26	1
7	42	17	33	27	1
8	28	18	3	28	1
9	39	19	116	29	2
10	5	20	46	30	2

表 5-7 生成 2000 个节点的网络度分布  
Table5-7 The scale distribution of the network  
of constructing 2000 peers

节点号	节点度	节点号	节点度	节点号	节点度
1	161	12	29	30	1
2	38	13	215	32	1
3	121	14	20	34	1
4	7	15	179	35	2
5	117	16	88	41	1
6	117	17	79	45	1
7	101	18	22	47	1
8	282	19	114	51	1
9	108	20	1		
10	44	21	2		
11	79	22	2		

联合上图以及表我们可以看出, 基于复杂网络模型无标度方法构建的 P2P 网络, 其节点度的分布具有极大的不均衡性, 占节点总数很少的较优节点集承载了大多数的网络连接, 从而决定了网络的健壮性; 网络负载不均衡, 整体资源的利用效率似乎较低。但是, 在流媒体的应用中, 与一般的估计相反, 这种倾斜性恰恰构成了它的优点。这是因为:

首先, 此类网络中, 预先部署的较优节点集仅占网络节点总数的很小一部分, 我们可以通过控制、调整较优节点集中节点的网络带宽及主机性能, 方便地控制整体网络的稳定程度。

其次, 此类网络遵照的是“富有者更富有”的原则, 网络中节点之间度差较大, 较优节点集承载了大多数的网络连接, 叶子节点到源节点的跳数较少, 因此网络延迟较小, 易于满足流媒体网络较高的实时性要求。

综上可知, 基于复杂网络模型标度无关方法构建的 P2P 网络, 能够较好地满足流媒体对于网络稳定性与实时性的要求, 适合于实际 P2P 流媒体网络的构建。

### 5.3 实际系统设计描述

上面给出了在理想状态下基于复杂网络的无标度网络构建 P2P 流媒体网络模型, 但在现实系统中, 即使性能好的节点它能承载的流媒体用户数量也是有限的, 但 P2P 流媒体系统扩大到一定的时候, 就会出现一些不能承载负荷的节点, 而这些节点大多是出现在先进入的节点中, 于是本文给出此算法在现实中的设计方案描述如下:

首先, 在流媒体的源端保存一张节点信息表, 表中记录所有访问该流媒体的节点信息, 节点的信息包括(节点的序号, 节点进入时间, 节点的承载能力), 而且表的排序按照进入时间的先后排序, 为什么要按照节点进入的先后顺序排序呢? 因为这样便于我们统计哪个节点在现时间长, 如果在线时间长的被选中的概率大, 那么整个系统会比较稳定。另外根据基于复杂网络无标度方法构建的 P2P 流媒体网络模型, 现进入的节点被选中的概率较大, 那么按照节点在线时间长短排序, 构建基于复杂网络的 P2P 流媒体模型, 不仅遵照“富有者更富有”的原则, 而且使在线时间长的节点被选中的概率增大。

其次, 在进行偏好性连接选择时, 那些已不能承载新节点加入的节点不能作为被选择的对象。

### (1)新节点加入的算法

当有节点 N 提出请求加入到服务 S 时, 服务 S 调用复杂网络的无标度网络构建方法, 在节点信息表中选择出被连接的对象 X, 则 N 与 X 建立连接, 同时把 N 的信息存在节点信息表中。

### (2)节点退出的算法

当有节点 N 退出服务 S 时, 节点 N 向服务 S 发出退出请求, 然后服务 S 把节点信息表中 N 和 N 的下游节点的记录删除, 节点信息表重新排序, 节点 N 的下游节点按照新节点加入的算法重新加入。

上面描述了新节点如何加入以及节点的退出和重构, 但是当某些节点它既在现但它的网络速度在某些地方出现瓶颈, 这时流媒体的播放质量出现严重问题, 这个节点和它下游的节点是请求退出还是只有该节点退出呢?

如果该节点和它下游的节点一起全部退出, 则给服务 S 造成较大的开销, 而且对该节点的下流节点的稳定性评价不公平, 于是就只有该节点请求退出重新建立连接。

## 5.4 本章小结

本章通过对两种不同的 P2P 网络构建方法进行仿真, 考察了不同构建方法实现的 P2P 网络在应用于流媒体传输时的性能差异。结果表明, 与基于模拟退火算法中 Metropolis 准则构建的 P2P 网络相比, 基于复杂网络模型无标度方法构建的 P2P 网络, 其稳定性和实时性都较优, 更适宜进行流媒体传输。



## 结 论

P2P 流媒体网络是网络模型构建和流媒体传输方面的热门话题，而现有的 P2P 流媒体网络的构建大多依赖于工程技术人员经验和启发式，于是本文想通过复杂网络无标度网络模型的构建方法，拟合出现实中的 P2P 流媒体网络，为 P2P 流媒体网络寻找一种新的理论支撑。

(1) 基于模拟退火算法 Metropolis 准则构建对等的 P2P 流媒体网络，这种模型，看似负载较均衡，但对于实时性要求高的网络来说，稳定性能较差，容易在局部引起网络振荡。由于网络节点的度分布比较均衡，那么实际造成网络节点到源节点层数大，则网络延迟较大。所以这种构建 P2P 网络的方式适合于那些实时性要求不高的 P2P 网络，比如文件共享网络等等。

(2) 基于复杂网络的无标度网络的构建方式，即遵照“富有者更富有”的原则，所构建出不对等的 P2P 网络的性能，对于实际中预先部署的网络节点的依赖性较强，由于实际中预先部署的网络节点的性能比较稳定，所以整个 P2P 网络也比较稳定，而且这种构建方式形成的 P2P 流媒体网络容易通过控制小部分节点而达到对整个网络的控制。度大的节点分布在先进入网络的节点的概率较大，那么从叶子节点到源节点的层数较小，这样网络传输的延迟也就较小，这正适合象流媒体这样实时性和稳定性要求较高的网络。

由于时间的原因，本文存在如下不足。

(1) 本文只是对构建 P2P 流媒体网络做了研究，但当网络中有节点离开或出现故障，它的下线节点“如何快速重新建立连接？”这有待研究。

(2) 对于网络节点如何对流媒体进行缓冲，以及缓冲管理等都需要继续研究。

## 致 谢

首先向我的导师徐鲁强副教授表示由衷的谢意和诚挚的敬意！本论文的研究工作是在导师徐鲁强老师的悉心指导下完成的。从论文选题、文献查找、思路探索到论文组织，都凝聚了导师的心血。授人以鱼不如授人以渔，导师在方法上的指导，课题进行中的不断鼓励和充分信任使我的分析问题和解决问题的能力获得进一步提高。深深感谢导师在三年时间里给予我学业上的教导和生活上的关怀，也将勉励我今后刻苦学习和积极工作。

感谢刘自伟研究员、张晖副教授等老师，他们严谨的治学态度及教导，使我终身受益；特别是张辉老师提示我用模拟退火算法做连接选择判断，刘自伟老师指导我怎么做规则库和推理机；有了他们的帮助和支持才使我能写出这篇文章。同时也感谢所有在我的研究工作中给予大力帮助的各位老师。

感谢朱林华同学对我学习上的帮助。罗婷婷、刘薇、毕振颇、戴敏、钱勇、何臻祥、李燕、刘渊、李秀燕等同学给了我一贯的支持和帮助，与他们共同生活和学习是我最难忘的愉快时光。

感谢我的家人和朋友，他们给我提供的无私的关心和支持，为我顺利完成学业提供了坚强的后盾！

感谢计算机学院老师们三年来对我的热情关心和指导，使我在计算机专业的学习受益良多。

向所有关心和支持我的师长、亲友和同学致以最诚挚的谢意！

## 参考文献

- [1] 高宗敏. 新媒体技术—有线电视技术[J]. 流媒体技术, 2005, 9:105-135
- [2] Kangasharju, Jussi. A T. Internet Content Distribution [J]. PhD thesis, 2002, 4:88-102
- [3] The Napster Homepage <http://www.napster.com/>, 2006. 7. 9
- [4] The Gnutella Homepage <http://gnutella.wego.com/>, 2006. 5. 7
- [5] Singh M P. Peering at Peer-to-Peer Computing [A]. IEEE InternetComputing, 2001, 10-20
- [6] D. J. Watts Princeton Small Worlds: The Dynamics of Networks between Order and Randomness [A]. University Press, Princeton, NJ, 1999:34-40
- [7] D. J. Watts, Six Degrees: The Science of a Connected Age [A]. Norton, New York, 2003:45-56
- [8] A.-L. Barabasi, Linked: The New Science of Networks [A]. Perseus, Cambridge, 2002:81-87
- [9] M. Buchanan. Nexus: Small Worlds and the Ground Breaking Science of Networks [A]. Norton New York, 2002, 90-97
- [10] S. H. Strogatz. SYNC: The Emerging Science of Spontaneous Order [A]. New York: Hyperion, 2003:78-89
- [11] S. H. Strogatz. Exploring complex networks [J]. Nature, 2001. 410:268-276
- [12] R. Albert, A. L. Barabasi, Statistical mechanics of complex networks [J]. Reviews of Modern Physics, 2002. 74:47-97
- [13] S. N. Dorogovtsev, J. F. F. Mendes, Evolution of networks [J]. Advances in Physics, 2002. 51:1079-1187
- [14] X. F. Wang, Complex networks: Topology, dynamics and synchronization [J]. Int. J. Bifurcation and Chaos, 2002. 12:885-916
- [15] M. E. J. Newman, The structure and function of complex networks [J]. SIAM Review, 2003. 45:167-256
- [16] X. F. Wang, G. Chen, Complex networks: small-world, scale-free and beyond [J]. IEEE Circuits and Systems Magazine, 2003. 1:6-20
- [17] Stephen E. D, Deborah. E, Dino. Fetal. An Architecture for Wide-area Multicast Routing [J]. IEEE/ACM Transaction on Networking, 1996. 4(2):12-18

- 
- [18] 贾娟, 曲晶, 伊鹏, 江斌强一种新型的联合单播/组播的两级交换结构 [J]. 计算机应用研究, 2007. 1:34-42
- [19] N.D. Martinez, Artifacts or attributes? Effects of resolution on the Little Rock Lake food web [J]. Ecological Monographs, 1991. 61:367-392
- [20] R.Ferrer i Cancho, C. Janssen, R.V. Sole, Topology of technology graphs: Small world patterns in electronic circuits [J]. Phys. Rev. E, 2001. 64: 46-119
- [21] A.-L. Barabasi et.al, Scale-free networks: Structure and properties, <http://www.nd.edu/networks/>, (2007. 3. 5)
- [22] D.J. Watts, S.H. Strogatz, Collective dynamics of "small-world" networks [J]. Nature, 1998. 393:440-442
- [23] M.E.J. Newman, D.J. Watts, Scaling and percolation in the small-world network model [J]. Phys. Rev. E, 1999. 60:7332-7342
- [24] H. Jeong, S.P.Mason, A.-L.Barabasi, et. al, Lethality and centrality in protein networks [J]. Nature, 2001. 411:41-42
- [25] A.-L. Barabasi, R. Albert, H. Jeong, Mean-field theory for scale-free networks [J]. Physica A, 1999. 272:173-187
- [26] Project Website <http://www.peercast.org>. PeerCast, 2006. 8. 6
- [27] D. A. Tran, K. A. Hua and S. Sheu. ZIGZAG: An Efficient Peer-to-Peer Scheme for Media Streaming[A]. In Proceedings of IEEE INFOCOM[C]. San Francisco, CA, USA, 2003
- [28] V N. Padmanabhan, H. d. Wang, P.A. Chou and K. Sripanidkulchai. Distributing Media streaming Workshop Content Using Cooperative Networking[A]. In Proceedings of International Network and Operating Systems Support for Digital and Video[C]. Miami Beach, FL, USA, 2002
- [29] V N. Padmanabhan, P. A. Chou. Resilient Peer-to-Peer streaming[A]. In Proceedings of IEEE International Conference on Network Protocols[C]. Atlanta, GA, USA. November 2003. 37-46
- [30] V K. Goyal. Multiple Description Coding[J]. Compression Meets the Network. IEEE Signal Processing Magazine, 2001, 18(5):74-93
- [31] M. Castro, P. Druschel, A.-M. Kermarrec, A. Nandi, A. Rowstron and A. Singh. High-bandwidth content distribution in a cooperative environment[A]. In Proceedings of Second International Workshop on Peer-to-Peer

System[C]. Berkeley, CA, 2003

- [32] M. Castro, PDruschel, A. Nandi, A. Rowstron and A. Singh. High-bandwidth multicast in a cooperative environment[A]. In Proceedings of 19th ACM Symposium on Operating Systems Principles[C]. Lake Bolton, New York. 2003.
- [33] Gridmedia Project Website <http://www.gridmedia.com.cn>. 2006. 9. 3
- [34] Anysee Project Website <http://grid.hust.edu.cn/anysee>. 2006. 10. 9
- [35] 詹雪峰. 流媒体系统同步机制和缓冲机制的研究与应用[D]. 成都: 电子科技大学, 2006. 27-41
- [36] 鲍广华. 多对多流媒体系统的同步研究[D]. 合肥: 合肥工业大学, 2003. 12-28
- [37] 李慧霸. P2P 流媒体服务中索引技术的研究与实现[D]. 长沙: 国防科学技术大学, 2005. 13-25
- [38] 江红霞. 基 JXTA 的 P2P 虚拟学习社区的设计研究[D]. 长春: 东北师范大学, 2006. 24-30
- [39] 雷雨. 流媒体系统的研究及实现[D]. 吉林: 吉林大学, 2006. 13-34
- [40] 谢勇均. P2P 视频直播传输系统的研究与实现[D]. 北京: 中国科学院研究生院, 2006. 8-17
- [41] 田阳. PVM 集群技术在流媒体服务器中的应用[D]. 北京: 清华大学, 2004. 81-85
- [42] 陈洲. P2P 流媒体广播系统的设计[D]. 上海: 华东师范大学, 2006. 12-31
- [43] 李晓燕. 基于树优先应用层组播技术的研究与实现[D]. 北京: 首都经济贸易大学, 2005. 24-32
- [44] 逢栋. 面向流媒体的应用层组播原型系统的研究与实现[D]. 南京: 东南大学, 2006. 34-48
- [45] 林众. 基于 MPEG-4 的流媒体技术开发与应用研究[D]. 大连: 大连理工大学, 2003. 18-26
- [46] 黎明. 基于对等网的广域对象存储系统研究[D]. 北京: 清华大学, 2004. 31-55
- [47] 张朕. 基于对等网络架构的流媒体在线直播系统的设计和实现[D]. 上海: 华东师范大学, 2006. 24-32
- [48] 刘昌玉. 应用层网络中多约束的组播路由算法研究[D]. 长沙: 湖南大学, 2006. 25-32
- [49] 李刚生. 基于组播的校园网络教学系统的设计与实现[D]. 长春: 东北师范大学, 2006. 10-15
- [50] 段翰聪. P2P 流媒体分发技术研究[D]. 成都: 电子科技大学, 2007. 18-30

## 攻读硕士学位期间发表的学术论文及研究成果

论文

李甫，徐鲁强. 基于复杂网络的 P2P 流媒体网络的研究[J]. 内江科技, 2008. 2