

基于蜂拥的 P2P 流媒体系统可扩展性分析

段晓阳 韩志杰 王冠男

(河南大学计算机与信息工程学院 开封 475004)

摘 要 P2P 流媒体直播系统在互联网上显示出了巨大的潜在吸引力,但此类系统的大规模部署严重依赖于它们处理高动态变化的效率,特别是在蜂拥时期。其主要原因是 P2P 流媒体直播系统的扩展在很大程度上取决于流媒体应用的时间需求。在分析和实验的基础上,提出了系统规模与时间约束的内在联系及其制约因素,构建了一个通用的 P2P 流媒体系统模型,来集中分析节点在蜂拥时期加入系统的过程。首先说明了简单使用“需求供给”概念模型来描述系统的规模是不够的,然后利用类似 Gossip 协议的随机伙伴选择机制,推出系统规模随时间变化的上限,在 Matlab R2010a 平台上比较分析了一些关键性因素的变化对系统扩展的影响。

关键词 P2P 流媒体,蜂拥,可扩展性,随机伙伴选择机制

中图分类号 TP393 **文献标识码** A

Analysis on Scalability of P2P Streaming System Based on Flash Crowd

DUAN Xiao-yang HAN Zhi-jie WANG Guan-nan

(School of Computer and Information Engineering, Henan University, Kaifeng 475004, China)

Abstract Peer-to-Peer(P2P) live streaming systems have recently shown great potential attractive on the Internet. However, large-scale deployment of such systems relies too much on their efficiency in dealing with high dynamic changes, especially in the flash crowd period. The main reason is that the expansion of a P2P live streaming system largely depends on the of the time demand of streaming media applications. Based on the proposed analysis and experiment, we found the inherent relation between the system scale and the time, as well as the constraints. We constructed a generic model for P2P streaming system to concentrate on analyzing the process that the nodes added to the system in the flash crowd. The paper first argued that simply using the "supply-demand" concept model to describe the system scale is not enough, and then introduced the upper bound of system scale over time with a random partner selection strategy like Gossip protocol. Finally we showed the impact on system scalability by critical factors clearly through a comparative analysis on Matlab R2010a platform.

Keywords P2P streaming system, Flash crowd, Scalability, Random partner selection strategy

1 引言

近年来, P2P(Peer-to-Peer)流媒体直播应用在互联网上越来越受到大家的青睐,然而大规模地部署高效、可靠的流媒体系统仍然是个可望而不可及的目标^[1]。虽然覆盖网络的组织与维护对于 P2P 流媒体直播系统的性能有着重要的影响^[2],但是节点的动态行为和传输延迟对系统的可扩展性能的影响却更加明显。其中一个关键的因素就是用户对系统具有严格的时间要求,它严重限制了系统规模的扩展,这也是 P2P 文件共享系统中常见的问题,比如大家熟悉的 Bittorrent。在 Bittorrent 中,用户为了成功地获得自己需要的文件可以等几个小时甚至几天,但在流媒体直播系统中显然是不现实的。因此,要保证高质量、大规模的 P2P 流媒体应用,深入了解时间需求以及它对 P2P 流媒体直播系统的影响就

显得非常必要了。

通过对 Coolstreaming 的研究^[3]可知,当 P2P 流媒体直播系统的规模达到一个合理的规模后,其性能就会维持在一个较高的稳定水平。但流媒体的性能会受到系统动态变化的严重影响,这是流媒体系统所独有的特性,特别是在 flash crowd 时期^[4],当成千上万个用户在一个集中的时间内同时加入一个直播系统中时, P2P 流媒体的效率将严重依赖于用户之间的流媒体资源的共享效率。在 flash crowd 期间,由于有大量用户几乎同时到达系统,因此会有许多用户不能获得足够的流媒体资源,这样用户要想成功加入系统就会等待很长时间,于是系统的整体启动延迟就会延长。

文献[5-7]已经对 P2P 流媒体系统模型进行了初步研究。文献[8]对时间限制与系统规模的关系进行了建模分析,但其模型不够直观明晰,在其基础上,以时间段为计算单位,组织

本文受国家自然科学基金青年基金(61103195),中国博士后科学基金面上项目(20100480048),河南省科技攻关项目基金(072102210044, 102102210489)资助。

段晓阳(1985—),男,硕士生,主要研究方向为 P2P 网络、无线传感器网络、云计算, E-mail: henudxy@163.com; 韩志杰(1979—),男,博士,主要研究方向为 P2P 网络、无线传感器网络、云计算等。

数列进行建模分析,利用随机伙伴选择机制,并对所有影响因子进行仿真分析。本文首先列出 P2P 流媒体系统规模上限的变化与时间变化的关系,即在 P2P 流媒体直播系统所能提供服务的最大节点数随时间的动态变化过程。然后主要对直播系统在 flash crowd 时期进行建模分析,找出影响系统规模随时间变化关系的一些关键因素,特别是节点的上载容量和系统的初始规模。基于我们的分析研究和仿真结果,提出了 P2P 流媒体直播系统满足具有严格时间限制的大规模用户的基本特征。

本文第 2 节通过模型分析在 flash crowd 时期的启动过程,并计算在该时期系统规模随时间段推移时的上限值,推出影响该上限值的一些关键因素;第 3 节主要根据模型的理论结果进行比较分析;最后总结全文。

2 系统建模

2.1 前提及假设

在 P2P 流媒体直播过程中有两个主要的组成部分,即流媒体数据定位和流媒体数据传输,二者都与重叠结构有着密切的关系。已经提出的结构模型有树/多树结构和 Gossip 网状结构,目前应用得比较广泛的是 Gossip 结构,如 Coolstreaming 和 PPlive。因此,我们的分析研究主要集中于 Gossip 结构。在该结构中,新到达的节点从系统中随机地选择一些稳定节点作为自己的伙伴,为自己提供流媒体资源。我们假设流媒体要正常播放需要的最低速率为 $R = xr$,其中 r 表示对应于一个带宽单位的比特率, R 表示 x 个单位的带宽资源需求。

在直播过程中,一个新节点达到系统后需要先在索引节点进行登记,然后从索引节点处获得系统中的稳定节点列表,再从列表中选择部分节点来为自己提供流媒体服务。虽然文献[3]说明了互联网上有充足的资源提供给 P2P 流媒体应用,但是仅仅拥有资源是远远不够的。因此我们认为供应相对于需求是过剩的,即总的节点上载带宽不小于总的下载带宽,其实这也是 P2P 流媒体直播系统规模能够扩展的必要条件。在这种情况下,如果有充足的时间,所有新节点最终都能够加入系统。

2.2 系统规模随时间的变化关系

系统模型中的关键符号及相关定义如表 1 所列。根据对现存 P2P 流媒体系统的测试,我们假设节点的上载能力是流媒体系统扩展的唯一瓶颈。用 u_i 表示第 i 个节点的上载能力。在分析之前,我们认为所有节点的平均上载能力为 u 。由于流媒体正常播放的速率为 R ,因此可以定义节点的平均相对剩余上载能力为 $h = \frac{u-R}{R}$ 。

表 1

变量	定义
M	初始系统规模
N	Flash Crowd 时期到达的新节点数
R	保证流媒体正常播放所需速率(=xr)
u	所有节点的平均带宽
h	所有节点的平均相对上载能力(=(u-R)/R)
k	一个新节点的资源提供者数量
S(n)	在第 n 个时间间隔后的系统规模
U_s	服务器上载能力
u_s	服务器的相对上载能力(= U_s/R)

在新节点到达之前,系统拥有 M 个节点,即系统的初始规模;另外还有总上载能力为 U_s 的流媒体服务器。于是,可以计算出初始系统可以服务的新节点的最大数量 $g = \frac{U_s + M(u-R)}{R}$,其定义为初始系统的剩余力。

不失一般性,在整个过程中我们认为时间是以时间段(T)变化的。现在开始计算系统规模随时间变化的基本上限。用 a_i 表示在第 i 个时间间隔里能够成功加入系统的新节点数量,其中 i 从 flash crowd 时期的开始计起。这 a_i 个节点在第 i 个时间间隔完成启动,并将在第 $i+1$ 个时间间隔开始为新的节点服务。假设节点在 flash crowd 时期一旦到达将不会离开系统,且节点的相对上载能力能够得到完全利用,则我们可得出下列关系:

(1)在第一个时间间隔,由系统初始剩余力来为新的节点服务,所以可以保证 g 个节点成功启动,即 $a_1 = g$;而在以后的每个时间间隔中它都可以保证 g 个新节点成功启动,并加入流媒体系统。

(2)在第二个时间间隔中,除了系统的初始剩余力,还将有 a_1 个节点加入系统为新节点服务,它们可以保证能完成启动的新节点的数目为 $a_2 = g + a_1 h$ 。

(3)对于第 n 个时间间隔,有 $a_n = h \sum_{i=1}^{n-1} a_i + g$,可以看出关系 $a_n - a_{n-1} = h a_{n-1}$,于是可以得出 $a_n = g(1+h)^{n-1}$ 。

根据以上关系可得系统规模 $S(n)$ 满足下式:

$$S(n) = M + a_1 + a_2 + \dots + a_n = M + \frac{g((1+h)^n - 1)}{h} \quad (1)$$

又有系统规模不可能大于系统的所有节点(包括已存在的和新到达的),于是有

$$S(n) = M + N \quad (2)$$

由式(1)和式(2)可得系统规模在第 n 个时间间隔结束时的基本上限,即

$$S(n) = M + \frac{g((1+h)^n - 1)}{h} \quad (3)$$

式(3)已经把 P2P 流媒体系统规模的上限描述得很好了,但是我们要注意到该式的前提假设为节点的相对上载能力能够得到完全利用,这显然是不实际的。新的节点可能通过竞争得到部分媒体资源,但这些资源又不能达到流媒体播放的需求,因此,式(3)实际上是不够准确的。现在采用文献[8]中类似于 Gossip 协议的随机伙伴选择机制,在以上计算的基础上重新计算系统规模。

仍然用 a_i 表示在第 i 个时间间隔内加入到系统的新节点数目, S_i 表示系统从开始到 i 时间间隔结束加入到系统的新节点总数。则在使用随机伙伴选择机制的 P2P 流媒体系统中,新的节点到达系统之后将在稳定节点中独立地、随机地选择 k 个作为自己的伙伴,则在第 n 个时间间隔中一个稳定节点被 $q(n, k, u_s)$ 个新节点选为伙伴。假设流媒体服务器在每个时间间隔随机地为 $u_s = U_s/R$ 个新节点保证流媒体服务,则第 n 个时间间隔内仍有 $N - S_{n-1} - u_s$ 个新节点需要选择伙伴。随机变量 $q(n, k, u_s)$ 满足二项分布 $(N - S_{n-1} - u_s, k/(M + S_{n-1}))$,而且期望为

$$E[q(n, k, u_s)] = \frac{k(N - S_{n-1} - u_s)}{S_{n-1} + M} \quad (4)$$

用 $E[q(n, k, u_s)]$ 近似等于每个已存在节点的 $q(n, k,$

u_s),从而一个新节点从已存在节点获得1个单位的上载资源的概率为 $p(n, k, h, u_s) \approx \frac{h}{E[q(n, k, u_s)]} = \frac{h(S_{n-1}+M)}{k(N-S_{n-1}-u_s)}$, 则一个新节点从已稳定节点获得的资源数满足参数为 $(k, \frac{h(S_{n-1}+M)}{k(N-S_{n-1}-u_s)})$ 的二项分布,相应的概率函数为 $C_k^i p(n, k, h, u_s)^i (1-p(n, k, h, u_s))^{k-i}$. 用 p_i 简单地表示 $\sum_{j=x}^k C_k^j p(i, k, h, u_s)^j (1-p(i, k, h, u_s))^{k-j}$, 则我们可以重新得到以下关系式:

(1)在第一个时间间隔里,由初始系统的剩余力提供流媒体服务,于是有 $a_1 = u_s + (N - u_s)p_1$.

(2)在第二个时间间隔里,将要加入系统的新节点数(既有在第一时间间隔里没有加入系统的节点,也有在第二时间间隔里到达的新节点)为 $N - u_s - a_1$. 于是可得第二个时间间隔里加入系统的新节点数为 $a_2 = u_s + (N - u_s - a_1)p_2$.

(3)第 n 个时间间隔里,有 $a_n = u_s + (N - u_s - S_{n-1})p_n$.

显然有 $S_n = S_{n-1} + u_s + (N - u_s - S_{n-1})p_n$, 因此,在第 n 个时间间隔结束时,可得系统的近似规模满足:

$$S(n) \approx M + S_n = S(n-1) + u_s + (M + N - u_s - S(n-1))p_n \quad (5)$$

式中, $p_n = \sum_{j=x}^k C_k^j p(n, k, h, u_s)^j (1-p(n, k, h, u_s))^{k-j}$, 而 $p(n, k, h, u_s) = \frac{hS(n-1)}{k(M+N-S(n-1)-u_s)}$.

在此系统模型的基础上,我们得出一个与之相吻合的理论框架。它充分显示了 P2P 流媒体系统在 Flash Crowd 期间系统规模随时间变化的基本关系,以及一些关键因素影响的了解,包括 k, h, M, N 和 u_s 。我们将在下一部分对各个因素对系统规模与时间关系的影响进行仿真分析。

3 结果分析

在该部分,我们将利用上面模型推出的理论结果(见式(5))来分析 P2P 流媒体直播系统在 Flash Crowd 时期的系统规模随时间变化的关系,以及一些关键因素的变化所产生的影响。在此,我们在 Matlab R2010a 平台上对所有影响因素进行比较分析,分析某一因素时,固定其他因素,与文献[8]相比,其更加直观,更加全面。具体分析如下:

(1)图1主要分析不同的服务器相对上载能力 u_s 对系统规模扩展的影响。其他因素取值为 $M=1500, N=10000, h=x=5, k=20$ 。从中我们可以看出,在刚开始的时间段里,系统规模增加相对缓慢,这主要是大量新的节点从数量有限的已存节点竞争媒体资源所导致的。而随着新节点加入系统,系统的整体相对上载能力与系统的下载能力的比率增大,于是系统规模增加的速度逐渐增大。另外, u_s 不同,则系统规模扩展的速度也不同,适当增加系统初始服务器的上载力,可以提高系统规模的扩展速度。

(2)图2则显示了在随机伙伴选择机制中选择伙伴数 k 的变化也影响到系统规模的变化。其他因素取值为 $M=1500, N=10000, h=x=5, u_s=0$ 。随着 k 值的增大,系统规模的扩展速度也明显增加。同样,对于给定规模的系统,它的启动延迟也随 k 的增加而减少。

(3)图3主要分析因素 h 的变化对系统规模变化的影响。其他因素取值为 $M=1500, N=10000, x=5, u_s=0, k=20$ 。显然,当节点相传上载能力提高时,将会保证更多节点对媒体资源的需求,所以 h 的增加会明显提高系统的扩展速度。

(4)图4展示了 flash crowd 时期到达的新的节点数 N 变化时,系统规模随时间的曲线变化。其他因素取值为 $M=1500, k=20, h=x=5, u_s=0$ 。显然,当 N 增大时,系统扩展速度随之减小。

(5)图5则说明了系统初始规模 M 对系统规模扩展的影响。其他因素取值为 $N=10000, k=20, h=x=5, u_s=0$ 。可以发现,适当增加系统的初始规模 M ,可以有效提高系统的扩展速度。

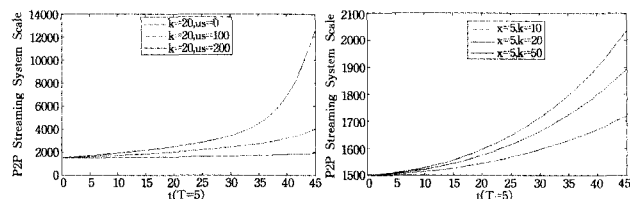


图1

图2

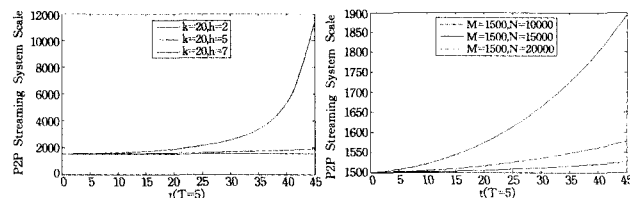


图3

图4

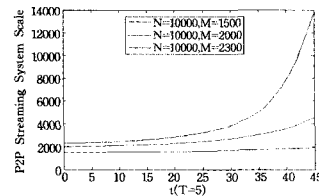


图5

结束语 本文针对 P2P 流媒体直播系统的特征,对系统在 flash crowd 时期规模扩展进行建模分析。根据模型的模式结果,在 Matlab R2010a 平台上进行比较分析,深入地了解各关键性因素对系统规模的扩展速度的影响。但是模型中仍有很多不足之处,如节点获得流媒体资源时的等待时间是有限的,这里并没有考虑到节点的动态离开;P2P 流媒体系统中的节点到达是具有很大动态性的,该模型并没有很好地体现出来。我们的下一步工作就是将这些未考虑全面的因素加入模型,从而完善模型,提出更准确的分析方案,为大规模部署 P2P 流媒体系统提供良好的理论依据。

参考文献

- [1] Liu J C, Rao S, Li B, et al. Opportunities and Challenges of Peer-to-Peer Internet Video Broadcast[J]. (invited) Proceedings of the IEEE, Special Issue on Recent Advances in Distributed Multimedia Communications, 2007
- [2] 彭凯, 武娟, 等. 基于 P2P 的流媒体直播技术与展望[J]. 计算机科学, 2009(1): 10-15

- [3] Li B, Xie S S, Keung G Y, et al. An Empirical Study of the Cool-streaming+ System[J]. IEEE Journal on Selected Areas in Communications, Special Issue on Advances in Peer-to-Peer Streaming System, 2007, 25(9): 1-13
- [4] Rubenstein D, Sahu S. Can unstructured P2P protocols survive flash crowds? [J]. IEEE Transactions on networking, 2005, 13(3): 501-512
- [5] Zhou Y, Chiu J C S. A Simple Model for Analyzing P 2 P Streaming Protocols[J]. IEEE International Conference on Network

Protocols(ICNP), 2007: 226-235

- [6] Kumar R, Liu Y, Ross K. Stochastic Fluid Theory for P 2 P Streaming Systems[C]//26th IEEE International Conference on Computer Communications(INFOCOM), May 2007: 919-927
- [7] Liu Y. On the Minimum Delay Peer-to-Peer Video Streaming: how Realtime can it be? [C]//Proc. of ACM Multimedia. 2007
- [8] Liu F, Li B, Zhong L, et al. Understanding the flash crowd in P2P live video streaming systems [C]//Proc of the International Packet Video Workshop, Seattle, WA, USA, 2009

(上接第 124 页)

5 个节点被分成了 1 类, 即 $(u_1, u_2, u_3, u_4, u_5)$ 。从结果可以看到, λ 取不同值时对分类结果的影响。

4 算法时间复杂度分析

该算法分为 4 个步骤完成, 下面分析每个部分的时间复杂度。

在 2.2 节中由式(2)看出, 实际上是对一个 $n \times 2$ 的矩阵中的个数进行标准化处理, 时间复杂度为 $O(n)$ 。在 2.3 节中, 模糊矩阵 $R_{n \times n}$ 中的每一项 r_{ij} 均通过式(3)来计算, 其时间复杂度为 $O(n^2)$ 。在 2.4 节中, 通过 k 次 $n \times n$ 的矩阵相乘得到模糊等价矩阵 R^* , 其时间复杂度为 $O(n^3)$ 。在 2.5 节中, 将矩阵 R^* 中的每一项与 λ 进行比较, 得到矩阵 $R^*(\lambda)$, 其算法时间复杂度为 $O(n^2)$ 。

综上所述可以得到, 模糊分类算法的时间复杂度为 $O(n^3)$ 。

5 算法仿真

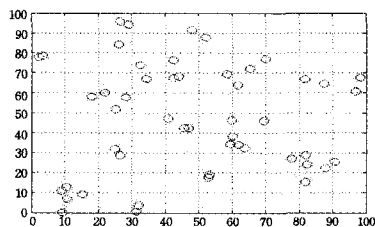


图 4 50 个节点非均匀的分布在 100m * 100m 的区域里

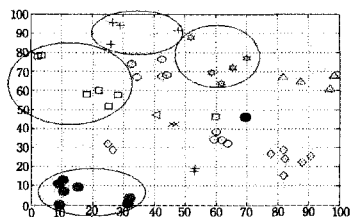


图 5 利用 k -均值聚类算法得到的分区结果, 椭圆标示区域显示边缘节点没有独立分出

利用 MATLAB 进行算法仿真, 在 100m * 100m 的区域中随机放置 50 个无线传感器节点(见图 4), 这些节点非均匀分布。利用 k -均值聚类和基于模糊等价矩阵聚类两种算法进行节点分区仿真, 结果如图 5、图 6 所示: 同颜色、同形状

点属于同一类, 两个图中的节点都分成了 14 个类。从图中可以看出, 基于模糊等价矩阵聚类分区算法成功地将边缘节点(离聚类中心较远的节点)分为单独的区域。

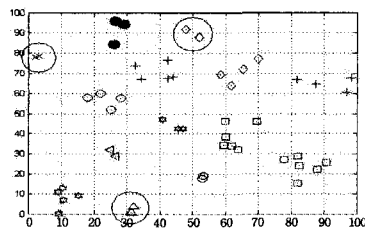


图 6 利用基于模糊等价矩阵聚类的分区结果, 椭圆圈出区域显示边缘节点被独立分出

结束语 本文的创新之处在于, 将模糊集理论中的模糊分类思想用于无线传感器节点分区, 为后续的节能路由算法设计提供了一种新的平台。从算法仿真的实验结果可以看出, 该算法确实可以将区域中的边缘节点分成独立的区域, 弥补了 k -均值聚类分区算法的不足。在分区算法中为无线传感网络路由算法的设计做了铺垫, 如何设计出更节能的路由算法是以后工作的方向和重点。

参考文献

- [1] Heinzelman W R, Chandrakasan A, Balakrishnan H. Energy-Efficient communication protocol for wireless micro sensor networks[C]//Proceedings of the Hawaii International Conference on System Sciences, San Francisco: IEEE Computer Society, 2000: 3005-3014
- [2] Manjeshwar A, Grawal D P. TEEN: A Protocol for Enhanced Efficiency in Wireless Sensor Networks [C]//Proceeding of the 15th Parallel and Distributed Processing Symp. San Francisco: IEEE Computer Society, 2001: 2009
- [3] Younis O, Fahmy S. HEED: A Hybrid Energy-efficient Distributed Clustering Approach for Adhoc Sensor Networks[J]. IEEE Trans. on Mobile Computing, 2004, 3(4): 660
- [4] 王佳英, 马洪连, 丁男. 无线传感器网络聚类改进层次路由协议[J]. 计算机工程与应用, 2010, 46(25): 91-93
- [5] 贾百川, 刘胜全, 黄蓬国. 基于轮班工作机制的 WSN 多跳聚类协议[J]. 计算机工程与应用, 2010, 46(25): 90-93
- [6] 陈水利, 李敬功, 王向公. 模糊集理论及其应用[M]. 2009: 59-122