ISSN 1000-9825, CODEN RUXUEW
Journal of Software, Vol.17, No.10, October 2006, pp.2087−2095
DOI: 10.1360/jos172087

E-mail: jos@iscas.ac.cn
http://www.jos.org.cn
Tel/Fax: +86-10-62562563

活跃型用户对 P2P 文件共享系统可用性的影响*

刘翰宇+, 肖明忠, 代亚非, 李晓明

(北京大学 信息科学技术学院 网络实验室,北京 100871)

Impact of Availability in P2P File Sharing System Caused by Active Peers

LIU Han-Yu+, XIAO Ming-Zhong, DAI Ya-Fei, LI Xiao-Ming

(Network Laboratory, School of Electronics Engineering and Computer Science, Peking University, Beijing 100871, China)

+ Corresponding author: Phn: +86-10-62751797, Fax: +86-10-62751799, E-mail: lhy@net.pku.edu.cn, http://net.pku.edu.cn

Liu HY, Xiao MZ, Dai YF, Li XM. Impact of availability in P2P file sharing system caused by active peers. *Journal of Software*, 2006,17(10):2087−2095. http://www.jos.org.cn/1000-9825/17/2087.htm

Abstract: The availability of a P2P (peer-to-peer) file sharing system is heavily affected by the high churn rate of users. When the largest P2P file sharing system Maze in CERNET is developed, the log collected in the system is used to get a better understanding of the users' characteristics, to find the key factor which influences the resource availability, and to instruct the future development of Maze system. In this paper, the concept of P2P file sharing systems' availability is redefined from users' perspective. With the log of Maze, it is the first study to use clustering technique to quantitatively categorize users in a P2P file sharing system. Based on the thorough study of behavior of the active users amounting to 0.77% of the total users and the impact on Maze's availability, this paper concludes that this kind of users plays a similar role to server so that they greatly increase the availability of system and are feasible resource to improve the performance of P2P file sharing systems.

Key words: P2P file sharing system; system availability; active peer; clustering; Maze

摘 要: 对等用户参与 P2P(peer-to-peer)文件共享应用的自由性,影响着该类系统的可用性.作为国内教育网上 Maze 系统的开发者,试图利用收集到的系统日志深入分析 Maze 用户特性,发现影响资源可用性的关键点,以指导 Maze 系统的演进.从用户需求的角度重新定义了 P2P 文件共享系统可用性的概念,并结合 Maze 系统日志,率先采用聚类技术对 P2P 文件共享系统的用户进行了量化分类,且深入研究了占用户总数大约 0.77%的活跃型用户对 Maze 系统可用性的影响.发现活跃型用户具有服务器性质,可大幅提升系统的可用性,是改进 P2P 文件共享系统设计可利用的资源.

关键词: P2P 文件共享系统;系统可用性;活跃型用户;聚类;Maze 系统
中图法分类号: TP311 文献标识码: A

P2P(peer-to-peer)文件共享系统已经成为互联网上的主流应用之一,对等用户参与该类应用的随意性,引起

* Supported by the National Natural Science Foundation of China under Grant No.90412008 (国家自然科学基金); the National Grand Fundamental Research 973 Program of China under Grant No.2004CB318200 (国家重点基础研究发展规划(973)); the Intel Sponsored Research Project (Intel 研究基金)

Received 2005-08-08; Accepted 2005-12-14

了学术界对该类系统可用性问题的广泛关注[1−3].相关工作大都是根据系统运行的日志信息来进行研究,然而,绝大部分研究者并不是使用 P2P 系统的开发维护者,因而不可能全面获知系统的运行信息.本文研究工作所使用的用户下载日志以及用户心跳日志等,均来源于我们自主研发并在中国教育科研网上具有广泛用户群的 Maze 系统[4].

该系统由中心服务程序和用户程序组成.如图 1 所示.其中,①表示用户将共享文件的元数据信息发送到文件收集服务器;②表示索引服务器为文件收集服务器收集到的所有用户共享文件的信息建立索引;③表示用户将查询请求发送到查询服务器;④表示查询服务器从索引服务器中获得结果返回用户;⑤表示用户之间进行文件交换.中心服务程序负责收集用户共享文件信息并生成索引以及提供用户检索接口,同时也收集用户的下载以及心跳等信息以用于对系统的管理与维护,这些功能分别由文件收集服务器、索引服务器、查询服务器以及管理服务器负责执行;用户程序由客户端 UI 以及客户端服务程序两部分组成,UI 主要提供用户搜索和下载操作界面,客户端服务程序独立于 UI,负责与前述中心服务器交换信息并向其他 Maze 用户提供文件下载服务.

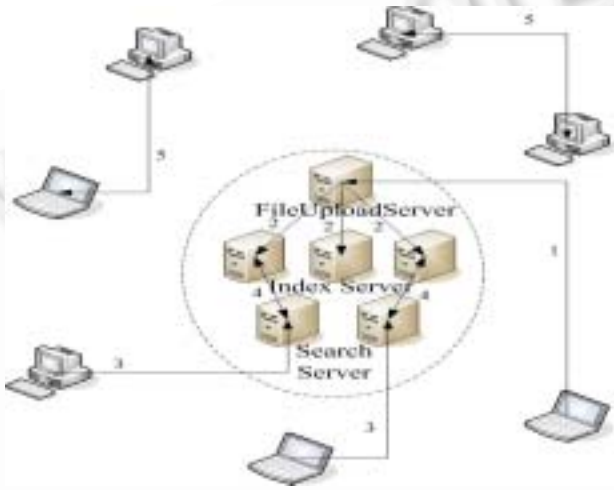

Fig.1 Architecture of Maze

图 1 Maze 系统结构

1.1 相关工作

文献[1−3,5,6]采用模拟特定 P2P 软件客户端探测系统中其他节点的方法来获取系统的运行情况,这不仅不能获得系统全面的运行信息,同时还会给系统带来额外的负担,容易被误认为是攻击行为.文献[7,8]则是通过在边缘路由器上获得网络的所有数据信息,从中过滤出 P2P 软件的数据流进行分析,这种方法要求对网络有较高的管理权限,同时该网络中必须有足够多的 P2P 用户与外界用户通信.虽然该方法能够通过较长时间的分析得出 P2P 网络的一些宏观规律,但是它无法分析系统中节点的在线时间、共享文件等信息.

在文献[5]中分别获得了 Napster 系统 4 天和 Gnutella 系统 8 天的数据,得出 P2P 系统中节点在传输速度、延迟、在线时间、共享数据等各方面都存在很大的异构性的结论,并定性地将系统中的节点分为两类:服务器型节点和客户型节点.文献[6]对 Gnutella 系统的每个用户共享出来的资源进行了分析,希望发现用户间共享文件的相似性.特别地,文献[6]根据共享文件的名字的相似度对用户进行简单分类,并得到了 5 个类别.

文献[1−3]分析了 P2P 系统中用户的可用性;文献[1]分析的是 BitTorrent 系统,认为为了提高系统的可用性,必须要将中心分布化,但是由于系统中缺乏高可用性节点,需要相应的激励机制;同时,文献[1]还指出中心节点可以有效地保证数据的完整性(即没有虚假文件),这样,少量管理员即可维护系统中新加入的文件,反映出 P2P 系统中高可用性与数据完整性之间存在一定的矛盾;文献[2]测量了 Napster 和 Gnutella 一个月左右的数据,该文指出用户可用性在一天的不同时段呈现出明显的不同,系统中的用户可用时间非常短,故 P2P 系统中采用缓存

可极大地提高系统性能;文献[3]对 Overnet 系统进行了 7 天的测试,该文重点研究了 P2P 系统中节点的可用性,指出其他相关研究使用 IP 标识节点会使结果严重失真,并强调 P2P 系统的可用性受时间影响严重;而后又通过实测数据利用条件概率说明,在 P2P 系统中节点的可用性相互独立.

1.2 本文的贡献

现有工作由于存在测试方法的局限性,无法完整地获得 P2P 文件共享系统的运行信息.而本文作者以开发维护者的身份全面分析了 Maze 系统 121 天的日志.表 1 给出了 2004 年 12 月 1 日~2005 年 3 月 31 日共计 121 天的用户下载日志的部分统计信息.该日志记录 Maze 系统用户每次下载结束后所下载文件的 MD5 值、大小、类型、文件提供者、下载开始时间、结束时间等信息.

Table 1 Summary of log information

表 **1** 日志信息

Log duration	121 days
# of active users	473 935
# of transfer log	102 136 695
# of transfer unique files	12 991 876
Total transfer size	1 540.51 TB

在 Maze 系统日志的基础上,根据 12 个 P2P 用户属性,本文率先采用聚类技术将 P2P 文件共享系统用户以量化的方式分为 4 类:NAT 型用户、活跃型用户、客户端用户以及惰性用户.其中,全面分析了占用户总数仅 0.77%的活跃型用户的在线情况和网络接入带宽等特征,发现这一类用户具有服务器特性.进一步地,从用户需求的角度重新定义了 P2P 系统可用性概念,且考察了这一类用户及其所共享文件对系统可用性的贡献,数字证明了它们是提高 P2P 文件共享系统可用性的重要资源.

本文首先给出对等用户量化分类的详细描述.接着,深入研究占用户总数仅 0.77%的活跃型用户的服务器特性.然后从用户可用性和文件可用性两方面考察这一类用户对系统可用性的贡献.最后,总结全文并给出活跃型用户对改进 P2P 文件共享系统设计的简要讨论.

2 对等用户量化分类

相关研究表明,用户在上传、下载、共享以及在线等属性上具有很强的异构性,得出 P2P 系统用户非对等的结论.但这些工作并未对用户进行定量的分类,或者仅是根据用户的某一属性来对 P2P 用户进行分类,如文献[6]中的共享文件相似度.针对现有研究工作的不足,本文首次在 P2P 用户分类中采用聚类这种量化分析方法来进行分类.我们采用统计工具 SAS 的 fastclus 过程进行聚类,表 2 列出了所选择参与聚类的 12 个用户属性.

Table 2 Attributes used for clustering

表 **2** 参与聚类属性

The online time of non-NAT server ($T1$)	The online time of non-NAT client ($T2$)	The online time of NAT server ($T3$)	The online time of NAT client ($T4$)
The size of shared files ($S1$)	The number of shared files ($S2$)	The size of download files ($D1$)	The number of download files ($D2$)
The download time ($D3$)	The size of upload files ($U1$)	The size of upload files ($U2$)	The upload time ($U3$)

为了确定用户分类的数目,我们首先将聚类个数设定为 2~12,并考察在这些情况下的 R^2 统计量、伪 F 统计量及 CCC 统计量,其中

$$R^2 = 1 - \frac{P_G}{T},$$

其中,P_G 为分类数为 G 个类时的总类内离差平方和;T 为所有变量的总离差平方和.R^2 越大,说明分为 G 个类时,每个类内的离差平方和都比较小,即分 G 个类是合适的.

$$F = \frac{(T - P_G)/(G-1)}{P_G/(n-G)}.$$

对伪 F 统计量而言,如果分 G 个类是合理的,则类内离差平方和(分母)应该较小,类间平方和(分子)相对应该较大.即应取伪 F 统计量较大而类数较少的聚类水平;CCC 也是一种考察聚类效果的统计量,CCC 较大的聚类水平是较好的.

图 2 给出了在聚类水平为 2~12 情况下的 3 个统计量的图形.观察图形可以看到,随着类数的增加,R^2 总是增大;而伪 F 统计量建议分为两类或 4 类;CCC 统计量建议分成 4 类,所以 Maze 用户可以分为 4 类.

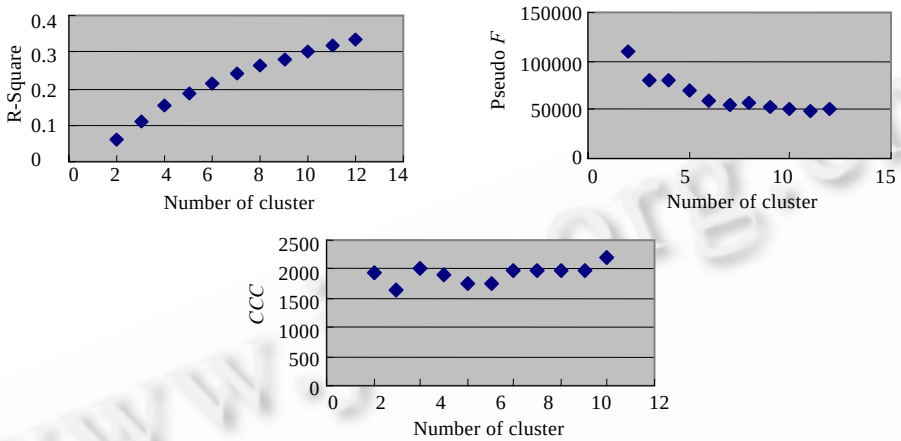

Fig.2 The statistics (cluster number between 2 and 12)

图 2 聚类水平为 2~12 情况下的 3 个统计量

表 3 给出了聚类后每类的数量及它们在各属性上的均值.注意:由于聚类时我们通过将各维数据标准化来消除各维度上量纲不同带来的影响,因此,这里的均值是原始值标准化后的均值.

Table 3 Mean for each classes of clustering

表 3 聚类得到各类均值

Class	Ⅰ	Ⅱ	Ⅲ	Ⅳ
Quantity	12 136	3 636	34 195	423 968
$T1$	0.003 827	408 316.7	196 575.7	1 924.758
$T2$	0.014 296	1 464 949	619 193.4	2 104.649
$T3$	276 714.2	0.007 92	0.006 298	0.014 984
$T4$	796 548.7	0.034 415	0.026 012	0.051 109
$S1$	4.57E+09	1.29E+10	4.94E+09	2.56E+09
$S2$	3371.824	1 360.7	4 065.957	15.623 62
$D1$	1.22E+10	3.2E+10	1.52E+10	1.29E+08
$D2$	398.074 3	919.284 5	693.077 5	49.990 31
$D3$	75 574.59	81 814.6	61 645.24	922.554 9
$U1$	1.82E+09	1.31E+11	2.35E+09	1 898.578
$U2$	279.686 3	5 237.015	74.650 45	9.16E-05
$U3$	16 782.96	616 817.1	10 722.98	0.006 809

为了进一步分析各类用户的特征,我们对用户的原始数据进行了因子分析.由于前 3 个因子的累计特征值超过特征值总和的 100%,故最后结果中只选 3 个因子.表 4 列出了方差最大正交旋转后得到的因子载荷.而后我们计算了所有用户的因子得分,为了得到 4 类用户的不同特征,我们求出了各类用户因子得分的均值,见表 5.

观察表 4 可以看到,下载相关 3 个变量 $D1$,$D2$,$D3$ 及在线时间 $T1$,$T2$,$T3$,$T4$ 这 4 个变量对第 1 个因子的载荷值较大,说明第 1 个因子反映了用户使用 Maze 的情况及用户可用性的情况;从第 2 个因子的载荷系数来看,上传相关 3 个变量 $U1$,$U2$,$U3$ 具有很大的值,说明第 2 个因子表现了用户上传情况,即用户作为服务器的情况;第 3 个因子的载荷系数在 NAT 外的在线时间是负值,而在 NAT 内的在线时间具有很大的正值,说明第 3 个因子反映了用户位于 NAT 内的特性.

第 1 类用户在第 3 个因子上得分很高,说明这类用户多处于 NAT 内.且其在第 1 个因子上也有较高的得分,

说明他们在线时间较长,是 NAT 内使用系统较多的用户.但其在第 2 个因子的得分为负值,说明 NAT 内的用户不易表现出服务器的特性.这类用户占聚类总用户数的 2.56%,我们称其为 NAT 用户.

第 2 类用户在第 2 个因子的得分极高,说明这是系统中具有服务器特性的那一群用户.他们在第 1 个因子的得分也很高,表明这类用户是系统中的活跃人员,在线时间长,上传、下载都很多,应该是我们进行 P2P 系统分析建模应该着重关注的用户群.另外,需要注意到这类用户在人数上是最少的一类,仅为 3636 人,占我们分析的聚类总用户数的 0.77%,我们称其为活跃型用户.

第 3 类用户在第 1 个因子的得分较高,而在第 2 个因子和第 3 个因子的得分较低或为负值,说明这类用户表现出很强的客户性:参与系统的活动比较多,但是多为下载活动,上传很少.且这类用户多处于 NAT 以外,占聚类总用户数的 7.22%,我们称其为客户端用户.

第 4 类用户在 3 个因子上的得分都为负值,对系统的参与,无论是上传还是下载都极少,表现出极强的惰性,我们称其为惰性用户.但是,这类用户人数巨大,占聚类总用户数的 89.46%.

Table 4 Factor pattern

表 **4** 因子载荷

Attributes \ Factor	Factor1	Factor2	Factor3
*T*1	0.615 14	0.396 5	–0.155 2
*T*2	0.384 84	0.397 98	–0.170 01
*T*3	0.149 33	0.015 24	0.653 85
*T*4	0.001 48	0.033 64	0.549 35
*S*1	0.178 15	0.230 27	0.123 62
*S*2	0.033 51	0.065 19	0.075 59
*D*1	0.638 74	0.127 55	0.119 39
*D*2	0.824 61	0.192 59	0.166 18
*D*3	0.793 72	0.114 07	0.302 74
*U*1	0.122 42	0.806 9	0.079 46
*U*2	0.154 77	0.673 11	0.038 28
*U*3	0.189 26	0.861 08	0.125 46

Table 5 Factor score for each classes

表 **5** 各类用户因子得分

Class \ Factor	Factor1	Factor2	Factor3
Ⅰ	1.233 9	–0.203 5	2.950 9
Ⅱ	2.399 3	7.335 9	0.331 4
Ⅲ	2.018 7	0.202 5	–0.208 9
Ⅳ	–0.218 7	–0.073 4	–0.070 5

3 深入研究活跃型用户

3.1 基本统计信息

如前所述,该类用户数占参与聚类用户总数的 0.77%,其共享的不重复文件数为 10 693 653,占系统中不重复文件数的 3.6%,共享大小为 59.8T,占系统中总文件大小的 5.64%,而其上传量占系统总上传量的 38.2%,下载量占总系统的 7.3%.

特别地,所有用户在 121 天中下载的文件有 88.44%可以在该类用户中找到,且与全部用户共享的近 3 亿个文件仅 1.7%被下载所不同的是,该类用户共享的 1 000 多万文件中有 19.3%的文件被下载过.

3.2 带 宽

图 3 显示了活跃型用户的上传带宽,可以看到他们的上传带宽都非常大,89.11%的用户上传带宽大于 1Mbps,且大部分都集中在 1Mbps~4Mbps 之间,充分验证了我们认为活跃型用户具有服务器性质的结论.

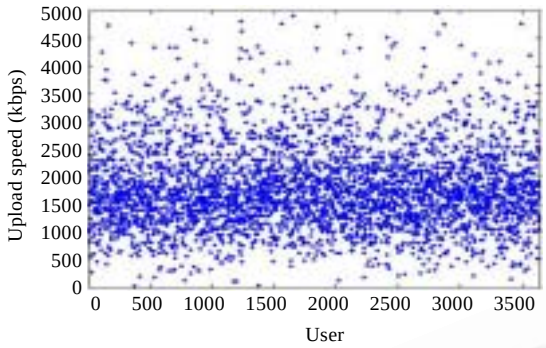

Fig.3 The upload bandwidth of active user

图 3 活跃型用户的上传带宽

3.3 在线情况

活跃型用户体现出极大的服务器特性,他们参与系统的情况反映出系统的服务能力,为此,我们详细研究了活跃型用户在线的情况*.由于用户在线受时间的影响很大,凌晨在线人数较少而傍晚在线人数较多,且其以天为单位呈现出强烈的周期性,为此,我们分析了 3 月份的日志,得到该月活跃型用户每天 0 点的在线人数,以及该日内每小时上、下线的人数.

图 4 显示了 3 月份活跃型用户每天 0 点的在线人数.可以发现,平时约 1/3 的该类用户 0 点在线,而周末在线人数为平时在线人数的 1.5 倍.我们采用柯尔莫哥洛夫-斯米尔诺夫(K-S)检验法对工作日数据进行正态分布检验,得到双侧 p 值为 0.917,即活跃型用户平时 0 点在线人数服从正态分布,其均值为 1 190,即占该类总用户数的 33%,方差为 64,占该类用户的 1.8%.

为了得到活跃型用户的在线规律,我们进一步研究了 3 月 18 日~3 月 24 日这 7 天的日志,得到每天从 0 点起每隔 6 分钟活跃型用户的在线人数.图 5 示例了工作日(3 月 18 日)与非工作日(3 月 19 日)的活跃型用户的在线曲线.可以看到,二者之间差别仅在凌晨,至上午 7 点后,二者趋于一致.另外,我们还可以发现:活跃型用户在线人数呈现出明显的 3 个阶段:午夜 0 点~上午 7 点为用户减少阶段,在这段时间内用户人数迅速下降,在 6 点左右达到一天中人数的最低点;上午 7 点~下午 3 点为用户迅速增加阶段;下午 3 点~晚上 11 点为用户缓慢增加阶段,经过上午的用户迅速增加后,在这个阶段人数增加放缓,至晚上 22 点左右,用户人数到达一天中的最高点,而 23 点用户人数明显减少,这是由于很多校园内用户受学校熄灯制度的影响.

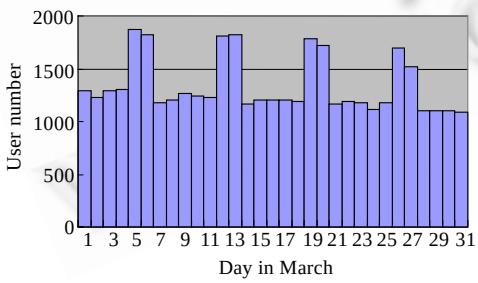

Fig.4 The number of active users which are online at 0:00 in March

图 4 3 月份活跃型用户每天 0 点在线人数

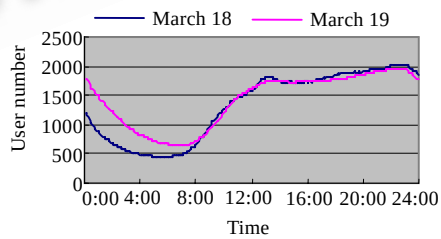

Fig.5 The online number of active users

图 5 活跃型用户的在线曲线

我们对 3 月 18 日~3 月 24 日的活跃型用户一天中在线人数随时间的变化进行了多项式拟合,从图 5 可以

* 这里的在线时间特指服务器程序的在线时间,未包括客户端程序在线时间.

看到曲线较为复杂,如果直接进行多项式拟合结果十分复杂.而根据上面的分析,我们可以按用户在线人数变化的 3 个阶段分段进行拟合.表 6 给出了时间段采用 2~5 次的多项式对上述 7 天数据的拟合而得到的 R^2 值.可以看到,在前两段,我们采用 3 次多项式即可得到较好的效果;而第 3 段由于有 23 点用户的突然减少过程,使得需用 5 次多项式拟合结果方能令人满意.需要说明的是,这里,我们没有专门区别工作日与分工作日,供拟合的数据点取值为上述 7 天在相应时刻在线人数的均值.式(1)列出了得到的拟合结果.

$$N(t)=\begin{cases}1309.161986-1309.161986t+60.919309t^2-3.370888t^3, & 0\le t<7\\ 1836.777575-800.589652\,t+116.222050t^2-4.254690t^3, & 7\le t<15\\ 1181823.614472-320219.915747t+34560.633544t^2- & \\ 1854.914282\,t^3+49.517389\,t^4-0.526033t^5, & 15\le t<24\end{cases} \tag{1}$$

Table 6 The value of R^2 for various degree polynomial fit

表 **6** 不同多项式拟合得到的 R^2 值

Degree \ Time in a day	0:00~7:00	7:00~15:00	15:00~23:00
2	0.986 4	0.986 4	0.455 6
3	0.996 4	0.996 8	0.691 6
4	0.999 1	0.997 8	0.869 0
5	0.999 4	0.998 0	0.962 1

4 P2P 系统中对象的可用性

4.1 可用性定义

通常,用户可用性(availability)定义为:在时间段$[t_1,t_2]$内,用户 u 的可用性为用户在线时间长度 T_o 与该段时间间隔ΔT 的比值,即有

$$A_{\Delta T}(u)=\frac{T_o}{t_2-t_1}=\frac{T_o}{\Delta T} \tag{2}$$

然而,该定义是从系统角度出发的.从前面的分析可以看到:在 P2P 系统中,用户使用系统以天为单位呈现出明显的周期性.例如:具有相同资源的两个用户分别在凌晨 6 点和晚上 10 点在线一个小时,他们的可用性对其他用户而言显然是不同的.在这样的系统中,我们认为用户可用性的定义应该从用户的角度出发,即可用性应该反映用户的使用需求,而下载请求正好反映了 P2P 系统中用户的使用需求,所以 P2P 系统用户可用性应如下定义:

定义 **1**. 时间段$[t_0,t_n]$为一时间周期段,将该时间段分为若干子段$[t_0,t_1],\ldots,[t_{n-1},t_n]$,若这 n 个子时间段中用户下载请求比例为 $d_0:d_1:d_2:\ldots:d_{n-1}$,则在时间段$[t_0,t_n]$中,用户 u 的可用性 A_p 定义为其在 n 个子时间段中原始可用性 $A_i(u)$的加权和,且第 i 段的权值 $p_i=d_i/(d_0+d_1+d_2+\ldots+d_{n-1})$,即

$$A_p(u)=\sum_{i=0}^{n-1}p_iA_i(u)=\sum_{i=0}^{n-1}\frac{d_i}{\sum_{j=0}^{n-1}d_j}A_i(u) \tag{3}$$

相应地,某一类用户的可用性可如下定义:

定义 **2**. 若用户类 C 包括 $u_1,u_2,u_3,\ldots,u_k$ 共 k 个用户,则时间段$[t_0,t_n]$内,用户类 C 的可用性为所有 k 个用户可用性的算术平均和,即

$$A_P(C)=\frac{\sum_{i=1}^{k}A_p(u_i)}{k} \tag{4}$$

4.2 活跃型用户的可用性分析

为了计算活跃型用户在定义 1、定义 2 下的可用性,我们仍以小时为界,将一天分成 24 小时,然后计算测试期内每小时用户下载请求到达的总数量,从而得到了表 7 中各段的加权值.而后,我们以 3 月份的日志为例,从

中分别算出了活跃型用户3月份在两种定义下的可用性 A_o^C 为0.359 4及 A_p^C 为0.433 7.可见:活跃型用户的可用性明显高于整个系统用户的可用性.另一方面,就用户角度而言,活跃型用户的实际可用性比单纯从系统角度看要高20.67%.

Table 7 Weight of each hour in a day

表7 各小时加权值

Hour	d_0	d_1	d_2	d_3	d_4	d_5	d_6	d_7	d_8	d_9	d_{10}	d_{11}
Weight	0.045	0.036	0.024	0.015	0.009	0.006	0.004	0.005	0.014	0.030	0.044	0.049
Hour	d_{12}	d_{13}	d_{14}	d_{15}	d_{16}	d_{17}	d_{18}	d_{19}	d_{20}	d_{21}	d_{22}	d_{23}
Weight	0.053	0.063	0.063	0.064	0.065	0.061	0.059	0.061	0.060	0.059	0.057	0.053

事实上,我们在上一节已经拟合得到活跃型用户在线数量随时间变化的函数,因此,我们还可以通过如下定义求得活跃型用户的可用性,且下面的定义与定义2等价.

定义3. 若 $F(x)$ 为用户类 C 在线人数随时间的变化曲线,则该类用户的在时间段 $[t_0,t_n]$ 的可用性定义为该用户类在 $[t_0,t_n]$ 的每个子时间段 $[t_0,t_1],\ldots,[t_{n-1},t_n]$ 平均在线人数 N_t 与该类用户总人数 N 比值的加权和,每段权值与定义1中各段权值相同,即

$$A_P(C)=\sum_{i=0}^{n-1}p_i\frac{\overline{N_{\forall t}}}{N}=\sum_{i=0}^{n-1}p_i\frac{\int_i^{i+1}N(t)\mathrm{d}t}{N} \tag{5}$$

我们以小时为段,将式(1)带入式(5)可以算出 $A_P(C)$ 为0.429 5,与实际测得值仅差0.97%.由于式(1)是我们通过一周数据拟合得到的结果,这里,我们验证了式(1)是能在整体上反映系统状态的普适公式.

4.3 活跃型用户中文件可用性

事实上,式(2)和式(3)同样可以表示系统中文件的可用性.另外,在上一节,我们对活跃型用户的可用性进行了详细分析,并得到在我们从用户角度出发定义的可用性下,该类用户可用性可以达到0.433 7.相关研究[3]显示出用户的可用性相互独立,因此,文件的可用性还可以表示为

$$A_p^f=\sum_{i=1}^{n}\binom{n}{i}(A_p(C))^i(1-A_p(C))^{n-i} \tag{6}$$

其中,n 为文件副本数.为了验证我们前面的工作及相关工作[3]中的结论,我们随机选择了活跃型用户副本数在4~8之间的文件各10个,对它们在3月份的可用性进行了实际计算.首先,我们找到文件对应的用户集合;然后,通过消除这些用户重叠的在线时间,得到该文件的实际在线时间,并代入式(3)求出其在P2P系统中的加权可用性,表8给出了结果.

Table 8 The availability of file with different replicas' number

表8 不同副本数的文件可用性

File / Replica No.	File1	File2	File3	File4	File5	File6	File7	File8	File9	File10	Theoretic value
4	0.9196	0.8480	0.7512	0.8432	0.9042	0.9847	0.6383	0.9881	0.9992	0.9959	0.8972
5	0.9915	0.9476	0.8687	0.9119	0.9780	0.9901	0.9105	0.9287	0.9057	0.9514	0.9418
6	0.9773	0.9805	0.9996	0.9854	0.9002	0.9032	0.9798	0.9403	0.9236	0.9421	0.9670
7	0.9617	0.9850	0.9049	0.9541	0.94389	0.9348	0.9385	0.9779	0.9874	0.9779	0.9813
8	0.9889	0.9511	0.9955	0.9567	0.9350	0.9624	0.9846	0.9679	0.9927	0.9871	0.9894

通过表8的数据可以看到:在活跃型用户中,较少的文件副本数就能保证较高的可用性,这一点与理论计算值一致.

5 小结及启示

本文基于对实用P2P文件共享系统Maze中用户具有强烈异构性的认识,在P2P用户分析领域首次引入聚类分析方法,得到了一类活跃型用户并对其进行了详细分析.

可以看到:这类特殊的用户在P2P系统中具有非常重要的作用,我们可以基于他们对P2P系统的设计、实

现进行优化.例如:Napster 的中心检索存在单点瓶颈,而类似 Gnutella 的分布式检索效果往往较差,考虑到整个系统被下载过的资源 88.44%可以在活跃型用户中找到,我们可以结合在活跃型用户中广播查询与中心检索.由于广播查询失效概率仅为 11.56%,故可以有效地解决中心负载问题.同时,相对于其他用户,活跃型用户具有高带宽、高可用性,他们表现出的这种天然的服务器性质,使得我们可以以他们作为超级节点,使其充分发挥作用,以提升系统的整体性能.另一方面,该类用户数还不到 1%,反映出系统中大部分用户仍然缺乏贡献精神,需要我们采用更加有效的激励机制,使得系统能够健壮发展,不至于由于这类用户的离去而面临崩溃.

另外,本文还借助加权思想提出了一种在 P2P 系统中衡量用户和文件可用性全新定义.与传统定义相比,本文提出的加权可用性定义可以有效地反映系统中用户的需求,从而更好地反映 P2P 系统中用户的特点.

致谢 向给予有益建议与帮助的同课题组成员杨懋、高乾、彭宇、侯潇潇等同学表示感谢.

References:

[1] Pouwelse JP, Garbacki P, Epema DHJ, Sips HJ. The Bittorrent P2P file-sharing system: Measurements and analysis. In: Castro M, van Renesse R, eds. Peer-to-Peer Systems IV, 4th Int'l Workshop, IPTPS 2005. LNCS 3640, Ithaca: Springer-Verlag, 2005. 205−216.

[2] Chu J, Labonte K, Levine B. Availability and locality measurements of peer-to-peer file systems. In: Firoiu V, Zhang ZL, eds. Proc. of the SPIE Vol. 4868, Scalability and Traffic Control in IP Networks II. Boston: SPIE, 2002. 310−321.

[3] Bhagwan R, Savage S, Voelker GM. Understanding availability. In: Kaashoek F, Stoica I, eds. Peer-to-Peer Systems II, the 2nd Int'l Workshop, IPTPS 2003. LNCS 2735, Berkeley: Springer-Verlag, 2003. 256−267.

[4] The Maze Web site. http://maze.pku.edu.cn

[5] Saroiu S, Gummadi PK, Gribble SD. A measurement study of peer-to-peer file sharing systems. In: Kienzle MG, Shenoy PJ, eds. Proc. of the SPIE Vol. 4673, Multimedia Computing and Networking 2002. San Jose: SPIE, 2001. 156−170.

[6] Makosiej P, Sakaryan G, Unger H. Measurement study of shared content and user request structure in peer-to-peer Gnutella network. In: Unger H, ed. Design, Analysis and Simulation of Distributed Systems 2004, Advanced Simulation Technologies Conf. Arlington: The Society for Modeling and Simulation International, 2004. 115−124.

[7] Sen S, Wang J. Analyzing peer-to-peer traffic across large networks. IEEE/ACM Trans. on Networking, 2004,12(2):219−232.

[8] Saroiu S, Gummadi KP, Dunn RJ, Gribble SD, Levy HM. An analysis of Internet content delivery systems. In: Culler D, Druschel P, eds. Proc. of the 5th Symp. on Operating Systems Design and Implementation (OSDI 2002). Boston: USENIX Association, 2002. 315−327.

刘翰宇(1979－),男,四川成都人,硕士,主要研究领域为 P2P 计算技术.

代亚非(1958－),女,博士,教授,博士生导师,主要研究领域为 P2P 计算技术,语义 Web.

肖明忠(1970－),男,博士,主要研究领域为 P2P 计算技术.

李晓明(1957－),男,博士,教授,博士生导师,CCF 高级会员,主要研究领域为信息检索,P2P 计算技术.