大规模网络中 BitTorrent 流行为分析*

陈亮[1,2] 龚俭[1,2]

[1](东南大学计算机学院，江苏 南京 210096)

[2](江苏省计算机网络技术重点实验室，江苏 南京 210096)

摘　要：针对 Peer-to-Peer(P2P)应用流量已成为 Internet 带宽的最大占用者的现状，本文在使用特征串方法准确采集国内最流行的 P2P 应用——BitTorrent 应用流量的基础上，深入研究了 BitTorrent 应用的流长、流持续时间、流速以及结点传输的流量、连接数等测度的分布情况，指出其流长、持续时间均服从 Weibull 分布，流速较一般 TCP 流速慢，并且 BitTorrent 网络呈现很强的不平衡性，同时分析了各分布中的异常情况。

关键词：大规模网络；P2P；BitTorrent；流行为；测度

中图分类号：TP393　　**文献标识码**：A

1 引言

在过去的几年中，Peer-to-Peer（P2P）应用的发展十分迅速。P2P 流量早已超过 Web 服务的流量，成为 Internet 上最大的带宽占用者[1]。P2P 模型异于传统 C/S 架构模型的特点必然使得其流量特征与传统应用层协议的流量特征不同，这些不同导致当前 Internet 流量特征较过去有很大的变化。因此，深入分析 P2P 流量特征对分析 Internet 总体流量特征、进行流量规划以及区分服务都有着重要的意义。

早期的 P2P 流量行为分析主要针对于每个端用户的行为特征[2][3]，例如共享的文件数、在线时间和传输的数据量，以及底层网络的一些特征如延迟等。这些研究均关注于用户在 P2P 网络中表现的行为，没有针对 P2P 系统在传输文件时的流特征（如流长、流速等）进行探讨。在 2004 年 PAM 会议上，Kurt Tutschku 使用被动测量的方法对 eDonkey 应用协议的流量特征进行了较详细的分析[4]，以统计的方法给出了流长、流到达间隔等测度以及用户地域的分布，但是其只给出了测度分布的曲线，既没有得出曲线的分布函数，也没有对其进行任何讨论。同年的 PAM 会议上，M. Izal 等人使用主动测量的方法研究了 BitTorrent 协议的流量行为[5]，但是其给出的测度大多为均值，且没有进行深入的分析。2005 年，L. Plissonneau 等人在前人的基础上全面分析了四种 P2P 协议的行为特征[1]，使用统计的方法得出了传输数据量、持续时间等的分布函数，并给出了流结束方式、结点地域分布、主机连接数等的统计分布。但是该研究仍存在两个缺陷：一是其在给出测度的分布函数时过于粗糙，并没有给出分布的参数以及任何数学证明；二是虽然其讨论了测度的分布情况，但并没有对产生该情况的原因做进一步的分析。本文将在这些方面做更深入的研究。另一方面，国内做 P2P 行为特征的研究尚未成为热点，据作者了解，国内只有文献[6]对 P2P 流量行为进行了八点分析，但是除了流的长相关性分析以外，对其它的特征都只给出了统计数值，没有进行深入分析。

本文针对国内的特点，选取国内使用量最大的 P2P 应用——BitTorrent 协议[7]（下简称 BT）作为代表，深入研究了其流长、持续时间、流速、主机连接数、传输数据量等流量特征。与前期研究相比，本研究在以下几个方面有所不同：1、前期所有使用被动测量获得 P2P 流量的研究均是基于端口区分 P2P 流量和其它流量，而大多数 P2P 应用为了避免识别，均使用随机端口进行通信[8]，这就造成样本数据的不可信，从而导致分析结果的不可信。而本文采用基于内容检查的方法识别 BT 流量，样本完全可信。2、本文是对前期研究的深入，在得到流长、持续时间等测度的累积分布曲线基础上，进一步分析得到这些测度的分布函数，并使用 Kolmogorov 检验函数的合理性。3、本文在讨论测度分布的基础上，更进一步的分析产生这些

* Supported by the National Grand Fundamental Research 973 Program Foundation of China under Grant No. 2003CB314804, 国家重点基础研究发展规划(973); the Key Project of Chinese Ministry of Education Foundation of China under Grant No. 105084, 教育部科学技术重点研究项目; the Jiangsu Province Key Laboratory of Network and Information Security under Grant No.BM2003201, 江苏省网络与信息安全重点实验室

作者简介：陈亮(1981.－),男,江苏南京人,博士生,主要研究领域为网络行为学. Corresponding author: E-mail: lchen@njnet.edu.cn, Phn: +86-25-83794000 ext 304; 龚俭(1957－),男,教授,博士生导师, CCF 高级会员,主要研究领域为网络安全,网络行为学.

现象的原因。4、前期所有的研究均没有对流速进行分析，本文将其纳入研究范围。

本文所使用的实验 TRACE 采集自江苏省教育网边界路由到国家主干路由之间，采用华东（北）地区网络中心设计开发的高速网络采集系统——Watcher，使用分光的方式采集。采集器采用 Intel XEON(TM) 2.40GHz*2 的处理器和 2GB 物理内存的硬件配置，操作系统平台采用 Linux 2.4.26 内核，基于 Ring-Buffer 零拷贝技术，支持 3 条 GE 信道，丢包率小于 0.5%，各采集器之间采用 NTP 对时。由于实验需要，本文只对信道上的 TCP 流量进行采集。TRACE 的具体描述见表 1。

表 1 实验数据描述

开始时间	持续时间	可用带宽	采集长度	流数（M）	bps(M)	pps(K)
2005-11-10 00:00	24 hours	1G*2*3(6Gb)	60B	814.9	2054.5	419. 1

全文组织如下：第 2 部分首先描述了 BT 流量的总体情况，第 3 部分分析了 BT 协议在传输文件时的流特征，第 4 部分分析了 BT 端用户的行为特征，第 5 部分是对全文的总结。

2 BT 流量总体情况

由于 P2P 应用用于维护网络，查询文件和应答的流量只占 P2P 总流量的不到 5%[1][4]，因此本文重点讨论 BT 应用在实际传输文件时所造成的流量。文章使用匹配特征串的方法识别 TRACE 中的 BT 流量[9]。流量的总体描述见表 2。

表 2 BT 流量总体信息

	流数	报文数	字节数	平均流长（pkt）	平均流长（byte）	平均报长（byte）
TCP	814.9M	34.5G	21.7TB	43.4	26.6K	627.5
BT	19.5M	19.0G	13.4TB	999.8	686.2K	702.7
比例（BT/TCP）	0.024	0.55	0.62	23.0	25.8	1.1

表 2 中列出了 BT 流量和 TRACE 中 TCP 流量的总体信息，以及它们之间的比例关系。由表中可见，虽然 BT 流数只占 TCP 总流数的 2.4%，且二者的平均报长基本相同，但由于其流长比 TCP 多一个数量级（见“平均流长”的两列），其传输的字节数占 TCP 总字节数的 62%，即 BT 应用的流量已经占据了网络总流量的约 60%。造成这个现象的原因有二：一是 BT 应用传输的多为多媒体数据等大文件，使得 BT 流较一般 TCP 流长；二是由于攻击、路由循环等原因存在于网络上的大量单报文流极大的减少了 TCP 平均流长。这样的流量统计结果和文献[1]中的结论非常相似，但是和国内先前的研究结果相差很大（文献[6]的研究结果表明 P2P 流量只占网络总流量的 1%）。我们分析造成这样差距的原因有二：首先因为当时 P2P 应用并没有像现在一样的流行；更重要的是先前研究采用基于端口识别 P2P 流量的方法，而这样的方法是不可信的，因此造成了统计结果的较大差异。

图 1 显示了一天时间内 BT 流量的变化情况。从图中可以看出人的行为对 BT 流量的影响：随着人们的活跃程度，参与到 BT 网络中的结点数发生变化，导致了 BT 流量在一天之中的有约 10 倍的变化量（6:00 最低，21:00 最高）。

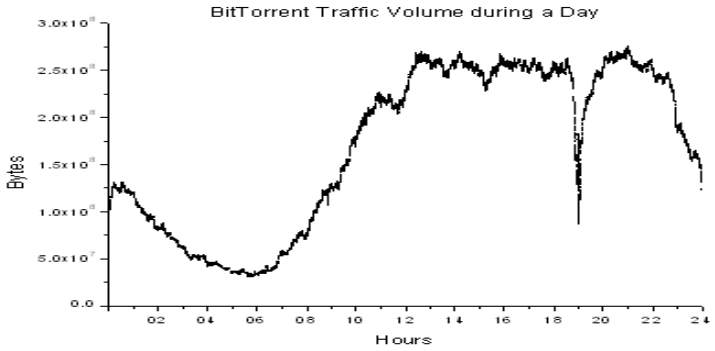


图 1 BitTorrent 流量在一天中随时间的分布

3 BT 流特征分析

本章将研究 BT 应用在传输文件时的流长、持续时间、流速的分布，并对其中的异常情况分析原因。

3.1 BT流流长分析

流长是指流内的报文数量或字节数。参照国际上大多数对流长的定义，本文采用流内报文数量作为流长的定义。BT 流长特征是指不同长度的 BT 流在网络中数量分布状况。

图 2 所示为实验所得的 BT 流长分布曲线，为了便于显示，坐标采用了 log-log 方式。其中图(a)是 BT 流长分布曲线，从总体上看 BT 流长服从重尾分布的特征，随着流长的增加，流的数量呈指数减少；但是数据同时也表明 BT 流并不是严格服从均一的重尾分布，存在与分布特征相违背的地方。如在流长约为 6-7、30、120 时均存在流数量突然增加的现象，特别在流长为 6-7 时突发现象很明显。我们对流长为 6-7 流进行检查分析后发现，这些流绝大多数为完整的 TCP 流，且都是 BT 网络中的一个结点向另一个结点发起连接而对方 BT 应用协议拒绝该连接所造成的。对于 TCP 协议而言，一个完整的交互需要至少 6 个以上的报文，且 TCP 第三次握手的报文和结束时的 FIN 报文都可以携带数据[10]。流长为 6-7 的流正是由于将发起方的 BT 应用协议握手信息或携带在 TCP 第三次握手的报文或作为单独报文，并且接收方的拒绝信息携带在 FIN 报文中所造成的。对流长为 30、120 左右的 BT 流分析发现，在这些发生数量突变的流中绝大部分有如下特点：1、这些流基本都未接收到任何应答；2、持续时间极短，一般持续时间为 10-100 毫秒之间。进一步分析这些流中的报文发现，属于这些 BT 流的报文的 TTL 值均具有递减的趋势，分别从 30 和 120 左右递减为 1，这是由路由表配置错误所导致路由循环的典型现象。由于本文所定义的流为双向流，所以每次报文经过观测节点都被观测到。进一步的证据来自于目前普遍使用的 Windows 操作系统中传递 IP 报文的 TTL 初始值，如表 3 所示，可以发现操作系统 TTL 值的设置和这些流流长度关系十分密切。因此，如果除去路由循环对流长所造成的影响，BT 流长分布的平滑性将更好。

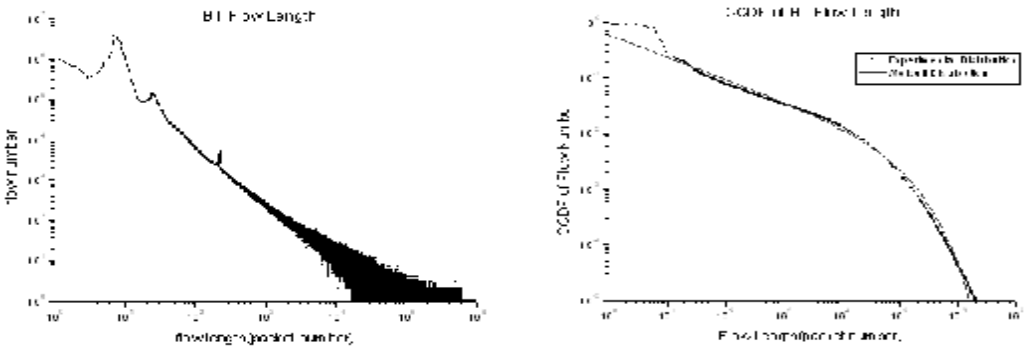
表 3 Windows 操作系统的 TTL 初始值

操作系统类型	TTL 初始值
Windows 95/98	32
Windows NT/2000/XP	128

图 2(b)中的黑色点线为流长的 CCDF 曲线。可以看出，流长小于 10 个报文的流占总流数的 70%以上，流长小于 100 个报文的流数超过总流数的 90%。然而，一个简单的统计可以知道，这不到 10%的流所造成的流量占 BT 总流量的 98%，我们称这些流为“巨流”。我们分析 BT 应用中短流数量多的原因有二：一是因为很多用户试图连接其他已经下线的结点；二是由于下载速度不理想等原因造成的用户取消连接。

从图 2(b)中还可以看出，当流长大于 10 之后，流长分布的 CCDF 曲线呈一条光滑的曲线。通过拟合得到流长大于 10 的 BT 流长分布模型：$\mathrm{f(x)}=\frac{\mathrm{k}}{\mathrm{s}^{\mathrm{k}}}\cdot x^{k-1}\cdot e^{-\left(\frac{x}{s}\right)^{k}}$，其中，$k=0.6, s=10^{5}$。

将 f(x)通过比例 $m=10^{3}$ 放大成 $m\cdot f(x)$ 即得到图 2(b)中灰色实线。即当流长大于 10 的情况下，除去比例关系 m，BT 流长分布基本服从 Weibull 分布。使用 Kolmogorov 拟合检验[11]评价拟合的结果：假设样本服从 Weibull 分布，$\mathrm{F(x)}=\mathrm{P}(X\le x)=1-\exp\{-(sx)^{k}\}$，随机取 $n=10217$ 个样本，算得 $D_n=0.000098$；取置信度 $a=0.05$，查表得 $D_{n,a}=1.36/\sqrt{n}=0.01345$。因 $D_n<D_{n,a}$，接受假设，可得 BT 流长分布和 F(x)拟合得很好，BT 流长分布服从 Weibull 分布。



(a) 流长分布曲线　　(b) 流长的 CCDF 曲线

图 2 BT 流长分布

3.2 BT流持续时间分析

图 3(a)所示为对 TRACE 统计得到的 BT 持续时间分布曲线，其依然服从重尾特征，但是突发现象较 BT 流长分布发生的范围更广。经过分析可知在持续时间为 5 秒左右的突发情况是由于上文所说的连接拒绝所导致的，但是对其后十几秒至三百秒之间的若干突发情况并没有找到实际的证据说明其原因，对此需要进行进一步的研究。

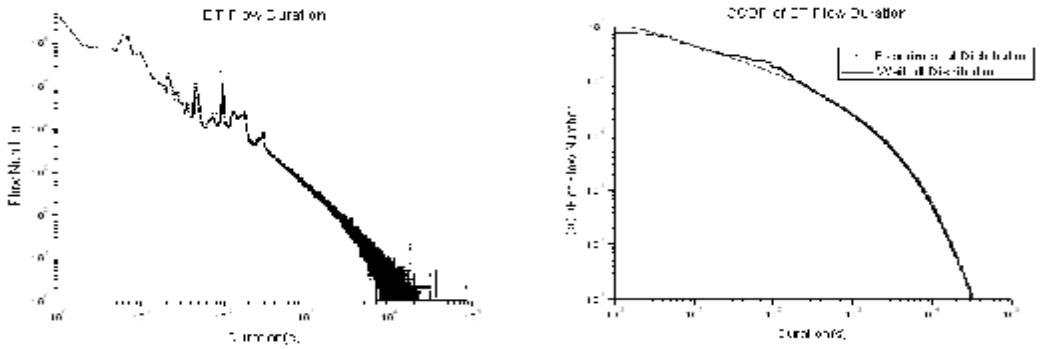
图 3(b)中的黑色点线为流持续时间的 CCDF 曲线。由图可见，持续时间小于 10 秒的流占总流数的 60%，小于 100 秒的流占总流数的 90%以上。然而统计发现，类似于 BT 流长分布，流持续时间大于 100 秒的这 10%的流所造成的流量占总流量的 98%，我们称这些流为“长流”。由图 4 中流持续时间与流长的关系点图可以看出，BT 流的持续时间越长，流内报文数就越多。从图中还可以得出，“长流”所对应的流内报文数基本上大于 100，这些流对应于 3.1 节中所述的“巨流”。也就是说，在 BT 网络中存在着一些流，这些流具有较长的持续时间和较多的报文数，虽然这些流的数量只占 BT 网络中总流数的不到 10%，但却占据了总流量的 98%以上。如果能够实时地发现这些流，就可以很好的遏制或者优化 BT 应用协议。

同样对 BT 流持续时间的 CCDF 函数进行曲线拟合，得到 BT 流持续时间的分布模型：

$\mathrm{f(x)}=\frac{\mathrm{k}}{\mathrm{s}^{\mathrm{k}}}\cdot x^{k-1}\cdot e^{-\left(\frac{x}{s}\right)^{k}}$，其中，$k=0.58, s=10^{3}$。

图 3(b)中灰色实线为 $m\cdot f(x), m=120$。使用 Kolmogorov 拟合检验评价结果：随机取 $n=10217$ 个样本，得 $D_n=0.000068$；取 $a=0.05$，$D_{n,a}=0.01345$。因 $D_n<D_{n,a}$，接受假设，即 BT 流持续时间分布和 $\mathrm{F(x)}=1-\exp\{-(sx)^{k}\}$，$\mathrm{k}=0.58,\mathrm{s}=10^{3}$ 拟合得很好，BT 流持续时间也服从 Weibull 分布。

从以上分析都可以看出，对于 BT 应用协议，流内报文数和流持续时间呈现出很强的一致性。因此，今后对 BT 协议的行为分析不必同时选取这两个测度，择一即可。



(a) 流持续时间分布曲线　　(b) 流持续时间的 CCDF 曲线

图 3 BT 流持续时间分布

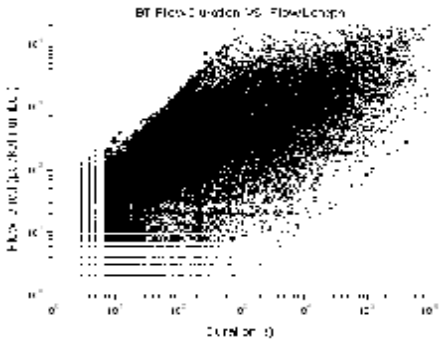


图 4 BT 流持续时间与流长关系

3.3 BT流流速分析

由图 5 所示的 BT 流速 CCDF 曲线可见，约有 95%的 BT 流流速小于 10pps，流速的均值为约 3.3pps。然而对整体 TCP 流统计得到其流速均值为 6.8pps，高于 BT 流速的均值，这说明了 BT 连接中存在较多的空闲时间，违背了一般长流多为快流的概念。造成这种现象的主要原因是 BT 协议规范中的“choke/unchoke”策略[7]：简单的说，BT 网络中的每个端结点在同一时间只会服务 4 个具有最大上传/下载速率的结点，且同时保持其它结点的连接，但不传输任何文件数据，评价其速率。并且，每隔 10 秒，结点根据评价结果重新选择当前具有最大上传/下载速率的结点进行服务；每隔 30 秒，结点对所有结点开放服务，以全面评价其速率。由于“choke/unchoke”策略的原因，结点之间的很多连接处于空闲状态，造成了 BT 流流速较低。

图 6 为 BT 流速与流长的关系点图。由 3.2 节的分析可知，该图也反映了 BT 流速与流持续时间的关系。总体上看，流长越长的流具有越快的流速。造成这种现象的原因是各用户网络带宽的不同。首先，高带宽的用户更愿意下载更大的文件，另一方面，高带宽意味着较少的拥塞，对应着较快的流速。因此，较快的流速就对应着较大的文件，即较长的流。另外，高带宽对应的较快的流速也使得该下载任务在“choke/unchoke”策略中更容易被选中，因此连接中具有较少的空闲时间，使得平均流速更快。

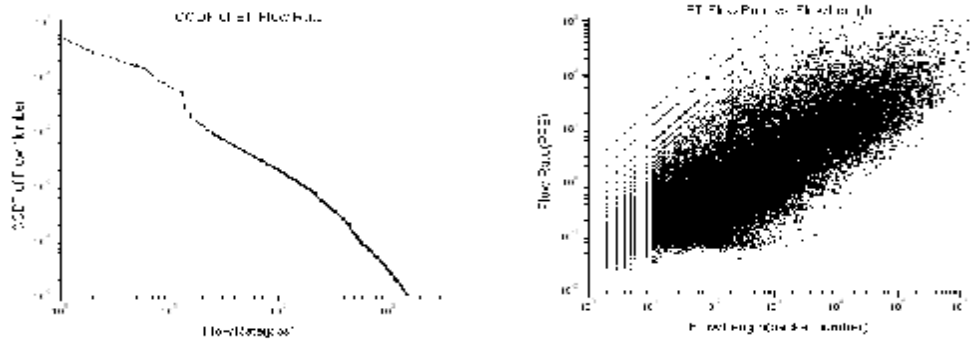


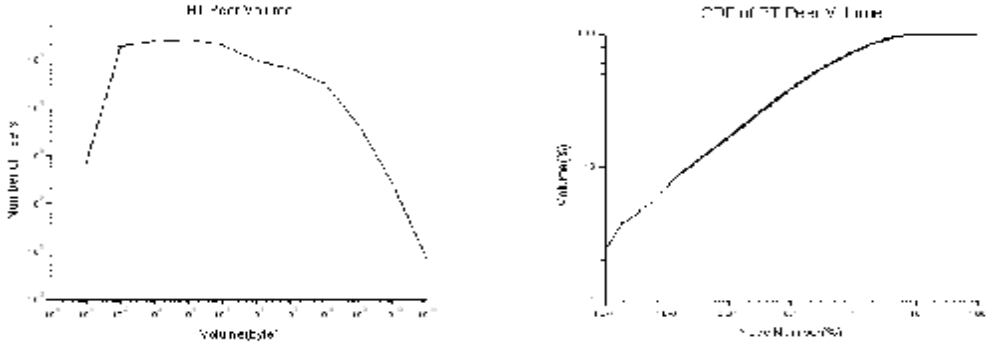
图 5 BT 流速 CCDF 曲线　　图 6 BT 流速与流长关系

4 BT 端结点行为分析

全面分析 BT 协议的行为特征，除了要分析其在传输数据时的流特征外，还需要对 BT 网络中端结点的行为进行研究。所谓端结点，是指参与流量传输 Peer。本文将每个独立的 IP 地址作为端结点的代表，不同的 IP 地址被认为是不同的端结点。在本文的 TRACE 中，共有 1076673 个不同的端结点构成整个 BT 网络。本章将分析 BT 网络中端结点传输数据量和连接数的分布，以及二者之间的关系。

4.1 BT端结点流量分析

图 5 所示为一天中 BT 端结点流量的分布情况。由图(a)可见，端结点的流量大多在 1KB 到几百 MB 之间。总所周知，BT 的典型应用是传输多媒体文件，因此，结点的典型流量应该在十几 MB 至几百 MB 之间，我们称该流量区间为正常区间。对正常区间的之外的结点分析发现：小于此流量区间的结点的流大多都具有较短的流长，本文 3.1 节分析短流数量较多的原因也是造成结点流量小于正常区间的原因；大于正常区间的结点具有较长的流长。图 5(b)为对结点流量按从大到小排序后的 CDF 曲线，其横坐标为结点数量的百分比。对比图(a)和图(b)可见，即使流量大于正常区间的结点很少，但是约 1%的结点所造成的流量就占据了总流量的 80%，而约 10%的结点所造成的流量超过了总流量的 90%，BT 网络表现出很强的不平衡性。并且从下一节的分析可知，流量大于正常区间的结点具有很高的连接度，因此，这些结点或者是在传输很大的文件，或者是作为“种子”存在 BT 网络中。



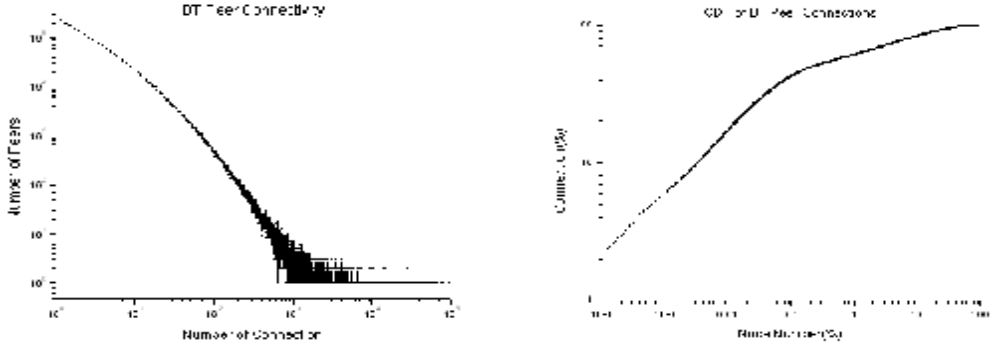
(a)端结点流量分布曲线　　(b) 端结点流量 CDF 曲线

图 5 BT 端结点流量分布

4.2 BT端结点连接数分析

图 6 为对 TRACE 统计得到的一天内 BT 结点连接数分布曲线。由图可见，结点连接数的分布有重尾分布的特征，1%的结点的连接数约占总连接数的 80%，10%的结点的连接数超过总连接数的 90%。

进一步分析端结点流量和连接数的关系：连接数排名在前 1%内的结点中，有约 75%的结点流量超过 500MB。流量超过 1GB 的结点的连接数排名全都在前 1%内。对比流量前 1%和连接数前 1%的结点发现，有超过 87%的结点相同。即，连接数高的结点贡献了绝大部分的流量。这些结点或作为 BT 网络中的“种子”或连接了很多其它结点以下载很大的文件。因此，对这些结点进行有效的控制将可以在很大程度上控制整个 BT 网络的流量，这对当前使用端口或特征串识别并阻止所有的 BT 流量是一个很好的改进。



(a) 端结点连接数分布曲线　　(b) 端结点连接数 CDF 曲线

图 6 BT 端结点连接数分布

5 总结

本文针对目前国内使用最普遍的 P2P 应用——BitTorrent 应用协议，使用基于特征串的方法在江苏省教育网边界和国家主干路由之间收集了一天的 BT 流量。并以此 TRACE 为对象，分析了 BT 流量的行为特征，结果显示：1、BT 流量已经占用超过 60%的网络带宽；2、BT 流长分布、流持续时间分布均服从 Weibull 分布模型，并且二者的特征相似，可以作为同一测度研究；3、BT 流平均流速慢于其它的 TCP 流，但是随着流长的增长流速也变快；4、BT 网络中的结点有很强的不平衡性，只有极少数结点有很高的连接数，并且传输了很大的流量，绝大部分结点都只贡献了很少的连接数和较少的流量。文章还指出，可以使用流长、结点流量或结点连接数等测度找出 BT 网络中为数极少的关键流或关键结点并加以适当控制，以达到控制整个 BT 网络流量的目的，从而改进传统的基于端口识别并阻止所有 BT 流量的恶性方法。

本文还尚不能对 BT 流持续时间分布中几个流数突发的异常情况作出合理的解释，这是今后需要进一步研究的内容。另外，未来的研究可以向两个方向发展：1、对 BT 或者其它 P2P 协议流量进行长时间的监测，研究其测度变化，这对网络控制和规划有着很好的指导意义；2、可以分析 BT 以外的应用程序流量行为特征，将传统的网络层和传输层流测度扩展至应用层流测度研究。

References:

[1] Louis Plissonneau, Jean-Laurent Costeux, Patrick Brown. Analysis of Peer-to-Peer Traffic on ADSL[J]. In PAM 2005, volume 3431 of LNCS Springer: 69--82.

[2] Stefan Saroiu, P. Krishna Gummadi, Steven D. Gribble. A Measurement Study of Peer-to-Peer File Sharing Systems[C]. In Proc. of MMCN, Jan 2002: 156-170.

[3] Subhabrata Sen, Jia Wang. Analyzing Peer-to-Peer Traffic across Large Networks[C]. IEEE/ACM Transactions on Networking. NJ: IEEE Press, 2004: 219-232.

[4] Kurt Tutschku. A Measurement-based Traffic Profile of the eDonkey Filesharing Service[C]. In Proc. of the 5th PAM, Antibes Juan-les-Pins, France, April 2004: 12-21.

[5] M. Izal, G. Urvoy-Keller, E.W. Biersack, P.A. Felber, A. Al Hamra, and L.Garces-Erice. Dissecting BitTorrent: Five Months in a Torrent‘s Lifetime. In Proc. of the 5th PAM, Antibes Juan-les-Pins, France, April 2004: 1-11.

[6] 张云飞, 雷连虹, 陈常嘉. Internet 中 Peer-to-Peer 应用流量测量与分析[J]. 铁道学报. 2004, 26(5): 55-60.

[7] BitTorrent [EB/OL]. http://www.bittorrent.com/protocol.html.

[8] Thomas Karagiannis, Andre Broido, Nevil Brownlee, kc claffy, Michalis Faloutsos. Is P2P dying or just hiding?[C]. In Proc. of IEEE Communications Society, Globecom 2004, Nov.-Dec.2004, 3(29):1532-1538.

[9] 陈亮, 龚俭, 徐选. 基于特征串的应用层协议识别[J]. 计算机工程与应用. 2006, 42(24): 16-19.

[10] RFC 793. Transmission Control Protocol[S].

[11] 叶慈南, 曹伟丽. 应用数理统计[M]. 北京. 机械工业出版社. 2004.

Analysis of BitTorrent Flow Behavior on Large-Scale Networks

Chen Liang[1,2], Gong Jian[1,2]

[1](College of Computer Science, Southeast University, Nanjing 210096, China)

[2](Jiangsu Province Key Laboratory of Computer Networking Technology, Nanjing 210096, China)

Abstract: The Peer-to-Peer (P2P) file sharing applications have become the major traffic sources in the Internet. This paper captures the traffic of BitTorrent application, which is the most popular P2P application in China, based on its characteristics. Based on this trace, analysis shows that the distribution of BitTorrent flow length and flow duration can be approximated by a Weibull distribution, flow rate is less than average TCP flow rate, the distribution of traffic volume and connectivity of each peer indicates the imbalance of BitTorrent network.

Abnormal conditions in each distribution are also analyzed.

Keywords: large-scale network; P2P; BitTorrent; traffic behavior; traffic measurement

关于文章所述工作的背景：

文章主要受资助于国家重点基础研究发展计划（973 计划）下的网络动态行为和传输控制理论。如何建立系统、科学与本质地刻画用户及其流量动态特性以及网络自身行为特征的模型是深刻认识和把握当前乃至新一代互联网的关键所在。该理论深入研究“网络动态行为及其可控性”科学问题，解决未知的网络行为与确定的传输控制目标之间的矛盾。它必将对网络行为状态的描述与预测，新协议的设计、开发与应用，新一代互联网络的规划、管理与控制，以及构建安全、可信赖的网络基础设施起到至关重要的指导作用。由于当前 P2P，特别是 BitTorrent 流量在网络流量中占据了最大的地位，故对其行为的研究必将对 Internet 整体流量行为的研究，乃至控制和管理有着积极的指导意义。

声明：

稿件内容属于作者的科研成果，数据真实；署名无争议；引用他人成果已注明出处；未公开发表过。

陈 亮

2007-07-20