

BitTorrent 带宽模型研究

欧阳荣 雷振明

(北京邮电大学信息工程学院 北京 100876)

摘要 对 BitTorrent(BT)网络中的 Peer 行为进行了细致的分析,并建立了相关的文件片段模型、感兴趣与阻塞模型,对 Peer 对邻居 Peer 的感兴趣概率、被邻居 Peer 及 Seed 阻塞的概率以及选择上传 Peer 的概率进行了分析。利用这些模型及概率表达式,推导出 BT 的带宽模型。

关键词 BitTorrent, 感兴趣, 阻塞, 下载带宽

Research on BitTorrent Bandwidth Model

OU YANG Rong LEI Zhen-Ming

(School of Information Engineering, Beijing University of Posts and Telecommunications, Beijing 100876)

Abstract Peer behavior in BitTorrent (BT) network is analyzed in detail, and the models related with peer behavior including file fragment model, interesting and choking model are presented. The probability of a peer interested in neighbor peer, the probability of a peer being choked by neighbor peer, and the probability of uploading peer being selected are analyzed. Based on these models and probability expression, the BT bandwidth model is evolved.

Keywords BitTorrent, Interest, Choke, Downloading bandwidth

1 引言

近几年,随着 P2P 软件的大量使用,P2P 业务占用了大量的带宽资源。P2P 流量已经占了整个互联网流量的 60% 以上^[1]。这给互联网上的其他业务带来了巨大的冲击。在整个 P2P 流量中,80% 以上是 BitTorrent(BT)流量。目前对 BT 流量的研究大多集中在 BT 的流量测量^[2~4],对 BT 的带宽模型没有作进一步的深入研究。本文对 BT 网络中的 Peer 行为进行了细致的分析,并建立了相关的模型和概率表达式,在此基础上得到 BT 的带宽模型。

2 BT 模型

2.1 BT 简介

BT 系统由 Tracker、Peer 和 Seed 组成,Peer 指正在下载 的节点。Tracker 为每个 Peer 返回一个邻居 Peer 列表,主要是用于帮助 Peer 彼此发现。文件被分割成多个片段。文件的信息及 Tracker 的信息放在一个 torrent 文件中,Peer 下载 torrent 文件从而获知文件信息及 Tracker 地址。Peer 在下载文件片段的同时,也向其他 Peer 上传已有的片段。拥有一个完整文件的节点称为 Seed。Peer 在下载完整文件后,如果不离开系统,就成为 Seed^[5]。

2.2 文件片段模型

BT 将文件切割为固定大小的片段(典型大小为 256k),每个片段又切割为若干个子片段(典型大小为 16k)。假设某个文件一共分为 m 个片段,每个片段又分为 l 个子片段。可将文件的切割抽象为如下矩阵 F 。其中, f_1, f_2, \dots, f_m 为文件片段标记, f_{ij} ($1 \leq i \leq m, 1 \leq j \leq l$) 为文件子片段标记。

$$F = \begin{bmatrix} f_1 \\ f_2 \\ \dots \\ f_m \end{bmatrix} = \begin{bmatrix} f_{11} & f_{12} & \dots & f_{1l} \\ f_{21} & f_{22} & \dots & f_{2l} \\ \dots & \dots & \dots & \dots \\ f_{m1} & f_{m2} & \dots & f_{ml} \end{bmatrix}$$

$$A = \begin{bmatrix} a_1 \\ a_2 \\ \dots \\ a_m \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1l} \\ a_{21} & a_{22} & \dots & a_{2l} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{ml} \end{bmatrix}$$

2.3 感兴趣模型与阻塞模型

我们以任意选定的 Peer A 为考察对象,Tracker 返回 Peer A 一个节点列表,将它分为一个邻居 Peer 列表 {Peer 1, Peer 2, ..., Peer n_p } 和一个邻居 Seed 列表 {Seed 1, Seed 2, ..., Seed n_s }。Peer A 与这些邻居 Peer 建立了 BT 握手连接。图 1 所示的系统是以 Peer A 为中心构成的,我们称它为 BT 的一个子系统。BT 由许多个 BT 子系统构成。在整个生存期,Peer A 会不停地收到邻居 Peer 发来的 have 消息,每个 have 消息都表明对方拥有了某个片段。这样 Peer A 可以知道在整个 BT 子系统中所有文件片段的分布情况。

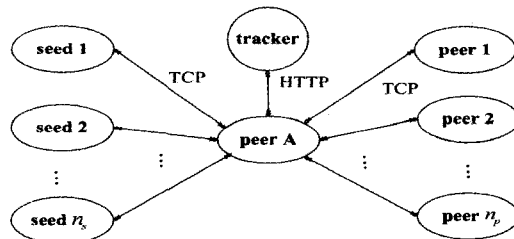


图 1 BT 子系统

定义 1 片段 f_j 在 BT 子系统中的拥有度:

$$x_j = \frac{\sum_{i=1}^{n_p} a_j^{(i)} + n_s}{\sum_{i=1}^{n_p} \sum_{j=1}^m a_j^{(i)} + mn_s}$$

定义 2 如果在 Peer A 尚未下载的所有片段中, x_j 值最小的片段 f_j 在 Peer i , 则称 Peer A 对 Peer i 感兴趣。

Peer A 对 Peer i ($1 \leq i \leq n_p$) 感兴趣的概率为:

$$p_{int}^{(Ai)} = \frac{1}{n_p + 1} (1 - P\{\text{peer } i \text{ 没有 peer A 所需要的片段}\})$$

$$= \frac{1}{n_p + 1} \left(1 - \frac{\binom{m - m_i}{m_A - m_i}}{\binom{m}{m_A}}\right)$$

m_A, m_i 分别为 Peer A 与 Peer i 所拥有的片段数。

Peer A 被邻居 Peer i ($1 \leq i \leq n_p$) 阻塞的原因有两点: ① Peer i 从 Peer A 得到的下载速率不够高, 没有排在前 D_{up} 位; ② Peer A 没有被 Peer i 的“优化非阻塞”机制所随机选中。设 Peer A 被任一个 Peer i ($1 \leq i \leq n_p$) 阻塞的概率是一样的, 为:

$$p_{ic}^{(Ai)} = \frac{D_{up}}{n_p} (1 - p_{uc}^{(Ai)}) + (1 - \frac{D_{up}}{n_p}) \frac{1}{n_p}$$

$$= \frac{D_{up} n_p (1 - p_{uc}^{(Ai)}) + n_p - D_{up}}{n_p^2}$$

其中 $p_{uc}^{(Ai)}$ 为 Peer A 向 Peer i 上传的速率排在 Peer i 前 D_{up} 位的概率, 它与 Peer i 的具体的网络参数相关。

设 Peer A 被邻居 Seed i ($1 \leq i \leq n_s$) 阻塞的概率是一样的, 为 $p_{ic}^{(Ai)}$ 。同理, 可以得到

$$p_{ic}^{(Ai)} = \frac{D_{up}}{n_s} (1 - p_{uc}^{(Ai)}) + (1 - \frac{D_{up}}{n_s}) \frac{1}{n_s}$$

$$= \frac{D_{up} n_s (1 - p_{uc}^{(Ai)}) + n_s - D_{up}}{n_s^2}$$

其中 $p_{uc}^{(Ai)}$ 为 Seed i 向 Peer A 上传的速率排在 Seed i 前 N_{up} 位的概率。

Peer A 选择 Peer i ($1 \leq i \leq n_p$) 为上传 Peer 的概率为 $p_{uc}^{(Ai)}$, 但只有当 Peer i ($1 \leq i \leq n_p$) 不被 Peer A 阻塞时, Peer A 向 Peer i 的上传才发生, 其概率为

$$p_{up}^{(Ai)} = p_{uc}^{(Ai)} p_{int}^{(Ai)} (1 - p_{ic}^{(Ai)})$$

$$= p_{uc}^{(Ai)} \left(1 - \frac{D_{up} n_p (1 - p_{uc}^{(Ai)}) + n_p - D_{up}}{n_p^2}\right)$$

$$= \frac{1}{n_p + 1} \left(1 - \frac{\binom{m - m_i}{m_A - m_i}}{\binom{m}{m_A}}\right)$$

3 BT 带宽模型

3.1 Peer 模型与 Seed 模型

我们考虑如图 2 所示的网络。AS 为广域网中的一个自治域网络。称路由器组以内(含路由器)的部分为内网, 以外的为外网。称内网的每个交换机及其它所包含的 Peer 为一个域。设 AS 网络中一共有 s 个交换机。以下载或上传同一个文件的网内外的所有 Peer 及 Seed 为一个系统进行考察。

假设 Peer 在下载完文件之前不离开系统。Peer 的到达可以认为是泊松过程, 到达率为 λ_p , 它的状态转换如图 3。其中, α 为未下载完就离开 BT 系统的离开率; β 为下载完成后就离开系统的离开率; λ_s 为下载完成后不离开系统的离开率, 也就是 Seed 的到达率。Peer 和 Seed 都可以用 M/G/ ∞ 排队模型来描述。在 t 时刻, 系统的 Peer 数为: $N_p = \lambda_p t - \alpha t - \beta t$

$-\lambda_s t$ 。设在系统中, 内网的 Peer 所占的比例为 p , 外网的 Peer 所占的比例为 $1-p$ 。设内网的 Peer 在各个域内的比例分别是 $q_1, q_2, \dots, q_s, \sum_{j=1}^s q_j = 1$ 。在时刻 t , 在域 j 内的 Peer 数为: $N_{m-j} = p q_j N_p$, 在内网的 Peer 数为: $N_m = p N_p$, 在外网的 Peer 数为: $N_{out} = (1-p) N_p$ 。

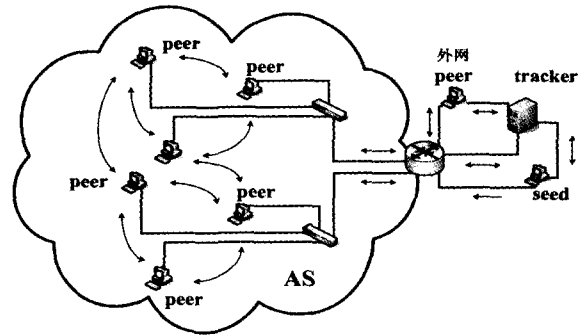


图 2 AS 网络内外的 BT 下载

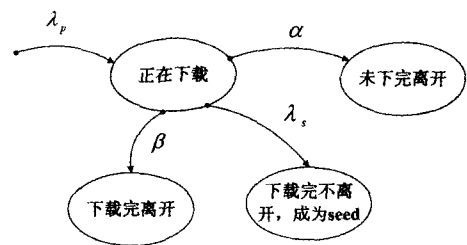


图 3 Peer 的状态转换

在 AS 网络中任意选择一个 Peer A, 设 ξ_1, ξ_2, ξ_3 分别是在 Peer A 的邻居 Peer 列表中域内、域外(但在内网中, 以下简称域外)、外网的 Peer 的角标集, $\xi_1, \xi_2, \xi_3 \subset \{1, 2, \dots, n_p\}$ 。设 η_1, η_2, η_3 分别是在 Peer A 的邻居 Seed 列表中域内、域外、外网的 Seed 的角标集, $\eta_1, \eta_2, \eta_3 \subset \{1, 2, \dots, n_s\}$ 。设域内、域外、网外 Peer 和 Seed 对 Peer A 的每上传连接的平均速率为 $\overline{b_{up1}}, \overline{b_{up2}}, \overline{b_{up3}}$ 。

Seed 也是按照泊松过程到达的, 到达率为 λ_s , 离开率为 u_s 。在 t 时刻, 系统的 Seed 数为: $N_s = \lambda_s t - u_s t$ 。Seed 在域内、域外、外网的比例与 Peer 完全相同。则在时刻 t , 在域 j 内的 Seed 数为: $S_{m-j} = p q_j N_s$, 在内网的 Seed 数为: $S_m = p N_s$, 在外网的 Seed 数为: $S_{out} = (1-p) N_s$ 。

3.2 BT 带宽模型

分析 Peer 的下载带宽和上传带宽是类似的, 我们只分析 Peer 的下载带宽。

3.2.1 Peer A 与交换机之间链路的下载带宽

$$B_{p-d} = B_{p-d-1} + B_{p-d-2}$$

其中, $B_{p-d-1} = \sum_{i \in \xi_1} p_{uc}^{(Ai)} \overline{b_{up1}} + \sum_{i \in \xi_2} p_{uc}^{(Ai)} \overline{b_{up2}} + \sum_{i \in \xi_3} p_{uc}^{(Ai)} \overline{b_{up3}}$

$$= \frac{1}{n_p + 1} \left(\sum_{i \in \xi_1} p_{uc}^{(Ai)} (1 - \frac{D_{up} n_p (1 - p_{uc}^{(Ai)}) + n_p - D_{up}}{n_p^2}) \right.$$

$$\left. (1 - \frac{\binom{m - m_i}{m_A - m_i}}{\binom{m}{m_A}}) \overline{b_{up1}} + \sum_{i \in \xi_2} p_{uc}^{(Ai)} \right.$$

$$\left. (1 - \frac{D_{up} n_p (1 - p_{uc}^{(Ai)}) + n_p - D_{up}}{n_p^2}) \right)$$

$$\left(1 - \frac{\binom{m-m_i}{m_A-m_i}}{\binom{m}{m_A}}\right) \overline{b_{up2}} + \sum_{i \in \xi_3} p_{ic}^{(A)} \left(1 - \frac{\binom{m-m_i}{m_A-m_i}}{\binom{m}{m_A}}\right) \overline{b_{up3}}$$

$$\frac{D_{up} n_p (1 - p_{ic}^{(A)}) + n_p - D_{up}}{n_p^2}$$

$$\left(-1 \frac{\binom{m-m_i}{m_A-m_i}}{\binom{m}{m_A}}\right) \overline{b_{up3}}$$

$$B_{p-d-2} = \sum_{i \in \eta_1} (1 - p_{ic}^{(A)}) \overline{b_{up1}} + \sum_{i \in \eta_2} (1 - p_{ic}^{(A)}) \overline{b_{up2}} + \sum_{i \in \eta_3} (1 - p_{ic}^{(A)}) \overline{b_{up3}}$$

$$= \sum_{i \in \eta_1} \left(1 - \frac{D_{up} n_s (1 - p_{ic}^{(A)}) + n_s - D_{up}}{n_s^2}\right) \overline{b_{up1}} + \sum_{i \in \eta_2} \left(1 - \frac{D_{up} n_s (1 - p_{ic}^{(A)}) + n_s - D_{up}}{n_s^2}\right) \overline{b_{up2}} + \sum_{i \in \eta_3} \left(1 - \frac{D_{up} n_s (1 - p_{ic}^{(A)}) + n_s - D_{up}}{n_s^2}\right) \overline{b_{up3}}$$

3.2.2 在域 j 出口的下载带宽

$$B_{s-d-j} = B_{s-d-j-1} + B_{s-d-j-2}$$

$$\text{其中, } B_{s-d-j-1} = \sum_{\gamma=1}^{N_{in-j}} \sum_{i \in \xi_2} p_{up}^{(\gamma)} \overline{b_{up2}} + \sum_{\gamma=1}^{N_{in-j}} \sum_{i \in \xi_3} p_{up}^{(\gamma)} \overline{b_{up3}}$$

$$B_{s-d-j-2} = \sum_{\gamma=1}^{N_{in-j}} \sum_{i \in \eta_2} (1 - p_{ic}^{(\gamma)}) \overline{b_{up2}} + \sum_{\gamma=1}^{N_{in-j}} \sum_{i \in \eta_3} (1 - p_{ic}^{(\gamma)}) \overline{b_{up3}}$$

3.2.3 在路由器出口的下载带宽

(上接第 54 页)

数,节点将以竞争失败结束本次发送。由于节点退避时间和数据发送的随机性,数据包发送延迟时间是不确定的。为了统计网络中各节点的平均延迟时间,现将网络中不同网络负载情况下系统工作节点的平均延迟数据统计绘制成延迟曲线,如图 7 所示。图 7 表明,在不同的网络实际负载流量下,网络的平均延迟时间小于 8ms,低于协议规定的延迟标准:15~30ms。因此,PM-WPAN 的数据延迟性能完全满足生理监护的需要。

结束语 由于无线生理监测系统可以在不影响被测者工作生活的情况下,全面准确地监测人体自然状态下的生理信号参数,提高医学诊断的准确性,而日益受到人们的重视。基于 IEEE802.15.4 的低速率个域网由于其本身在传输范围、功耗、传输速率和价格等方面的性能,非常适合用来构建人体生理监测系统的无线传输网络。本文探讨了基本人体生理传感器的数据传输特性和放置部位,由此设计了无线传输网络工作频段、拓扑结构、节点功能及传输机制。人体生理监测需要稳定可靠的网络传输,为了确切地提高网络效率,把握网络性能,通过网络仿真实验,对网络吞吐量、丢包率和传输平均延迟时间等性能进行了评估,从而对实际系统的开发和设计起指导作用。

参考文献

1 Townsend K, Haslett J W, Tsang T, et al. Recent advances and

$$B_{r-d} = B_{r-d-1} + B_{r-d-2}$$

$$\text{其中, } B_{r-d-1} = \sum_{\gamma=1}^{N_{in}} \sum_{i \in \xi_3} p_{up}^{(\gamma)} \overline{b_{up3}}$$

$$B_{r-d-2} = \sum_{\gamma=1}^{N_{in}} \sum_{i \in \eta_3} (1 - p_{ic}^{(\gamma)}) \overline{b_{up3}}$$

结论 随着 BT 类型的 P2P 软件的不断广泛使用, BT 流量占用了大量的带宽资源,研究 BT 的带宽模型就显得相当重要。本文对 BT 网络中的 Peer 行为进行了细致的分析,并建立了相关的文件片段模型、感兴趣与阻塞模型,对邻居 Peer 之间的感兴趣概率、阻塞概率以及上传的发生概率做了分析。利用这些模型及概率表达式,推导出 BT 的带宽模型。

参考文献

1 Parker A. The true picture of Peer-to-Peer file sharing[EB/OL]. 2004. <http://www.cachelogic.com>

2 Saroiu S, Gummadi P K, Gribble S D. A Measurement Study of Peer-to-Peer File Sharing Systems. Multimedia Computing and Networking 2002 (MMCN '02)

3 Sen S, Wang J. Analyzing Peer-to-Peer Traffic Across Large Networks. ACM / IEEE Transactions on Networking, 2004, 12(2): 137~150

4 李江涛, 雷振明. P2P 网络性能测度及监测系统模型[J]. 北京邮电大学学报, 2006, 29(3): 17~21

5 Cohen B. BitTorrent Protocol Specification [EB/OL]. <http://www.bittorrent.com/protocol.html>

future trends in low power wireless systems for medical applications. In: 5th IEEE Int. Workshop on System-on-Chip for Real-Time Applications, Banff, Alberta, June 2004. 476~481

2 Ooi P, Culjak G, Lawrence E. Wireless and wearable overview: stages of growth theory in medical technology applications. In: Proceedings of the International Conference on Mobile Business (ICMB'05), July 2005. 528~536

3 IEEE 802.15.4, Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for Low-Rate Wireless Personal Area Networks (LR-WPANs). New York, NY: IEEE, October 2003

4 Lee J S. An experiment on performance study of IEEE 802.15.4 wireless networks. In: Proc. IEEE International Conference on Emerging Technologies and Factory Automation, Catania, Italy, Sept. 2005, 2: 451~458

5 Golmie N, Cypher D, Rébala O. Performance analysis of low rate wireless technologies for medical applications. Computer Communications, 2005, 28(10): 1255~1275

6 Timmons N F, Scanlon W G. Analysis of the performance of IEEE 802.15.4 for medical sensor body area networking. In: Proc. IEEE Int Conf Sensor and Ad Hoc Communications and Networks (secon'04), Santa Clara, CA, October 2004. 16~24

7 Jaleleddine S M S, Hutchens C G, Strattan R D, et al. ECG data compression techniques - A unified approach. IEEE Transactions on Biomedical Engineering, 1990, 37(4): 329~343