

一种基于 BT 的域间 P2P 流量优化方法^①

闫 杨, 黎文伟

(湖南大学 软件学院, 长沙 410082)

摘 要: 随着 P2P 软件的普及, P2P 成为影响 Internet 网络结构的关键应用. 在中国运营商的长途骨干网上, P2P 流量占一半以上, 不仅给 ISP 带来了巨大的压力, 也使网络拥塞现象日益严重. 据此本文以 P2P 常用协议 BT 为对象提出一种优化方案, 通过用户协作与估算模拟底层网络相综合的方法, 在不影响性能下, 引导流量更多地在本地上网发生, 降低域间及主干网络上的 P2P 流量. 测试表明, 该机制不仅能尽量将域间流量本地化达到整体优化的目的, 而且更能增强用户系统的性能.

关键词: BT; 域间; 流量优化

Scheme for Optimizing P2P Traffic Between Networks Based on BitTorrent

YAN Yang, LI Wen-Wei

(Software School, Hunan University, Changsha 410082, China)

Abstract: With the popularity of P2P software, P2P has become critical applications affecting Internet network structure. In the long distance backbone networks of China ISP, P2P traffic accounts for more than half, not only to bring the huge pressure for ISP, also make the network congestion increasingly serious. So this paper puts forward an optimization scheme for agreement BT. With compositing cooperation of users and analog the topology structure of bottom network, it has no impact on the performance. And it guides traffic more happened in local networks, so that reduce the P2P traffic between networks. The tests show that the system can not only as far as possible to achieve the purpose of the whole optimization by traffic localization, and more can enhance the performance of the user system.

Key words: bittorrent; between networks; optimizing traffic

P2P 网络属于叠加在底层通信网络基础设施之上的覆盖网络, 这种网络更利于人们资源交换的自由性. P2P 文件共享技术在给 Internet 用户共享资源带来方便的同时, 也给网络运营商和内容提供商带来了棘手的问题, 即 P2P 文件数据传输占据了大量网络带宽, 特别是骨干网络上带宽的占用, 极大地增加了网络负担, 使网络拥塞现象日益严重.

为应对 P2P 流量对网络带宽消耗带来的挑战, 文献 1 分析了现有的情况, 指出目前优化方案主要分两个方面, 一种是运营商通过加大投入增加带宽来解决问题; 另一种就是通过一系列的技术方法限制或引导 P2P 流量^[1]. P4P^[2]是由美国耶鲁大学网络系统实验室

提出的一种 P2P 流量优化技术架构, 提出通过在 ISP 和 P2P 应用之间进行通信和协作, 达到协调网络拓扑数据, 能够有效选择 Peer, 从而提高网络路由效率, 进行流量优化. 文献 3 是使用 CDN 网络的动态 DNS 重定向技术为客户提供低延迟的镜像服务器, 利用这个信息, 无须另外的辅助设施就可以实现比较优化的节点选择策略^[3].

这些研究方案主要是基于在客户端的改进, 而其优化主要思想就是尽量将流量控制在本地发生. 在此基础上本文提出了建立在协议修改上的优化机制, 可以不需要和运营商合作或者另外搭建缓存设备即可将流量本地化, 达到优化整体网络流量分配的目的.

① 基金项目:国家自然科学基金(61173168);湖南省科技计划(2011FJ3236);东莞市高等院校科技计划(2011108101010)

收稿时间:2012-03-11;收到修改稿时间:2012-04-25

1 优化机制与算法

本文采用的是将用户协作与用户拒绝协作两种方案相结合的机制进行优化。一种是用户愿意合作，可以主动提供其定位信息。另一种是用户不愿意合作，通过一系列算法估算出底层网络信息^[3]。两种方法综合出一个节点选择优先级列表，通过该列表使节点选择尽量控制在域内，以达到降低域间流量占用的目的。

这种方法一方面降低了优化的复杂度，另一方面也可以在用户需求保护隐私的同时不影响优化性能。

1.1 基于用户协作的优化

本小节是基于用户同意协作的基础上进行的，通过修改 BT 协议将用户的定位信息加入种子信息中，制定地理位置的优先级列表来挑选资源种子^[4]。

在 BT 协议中，一次交互至少需要一个 tracker 和 seed。tracker 负责帮助 peers 之间相互建立连接。首先将扩展名为.torrent 的文件放在服务器上，.torrent 文件包含了要共享的文件的信息和一个指向 tracker 的 url。如果有一个 peer 向 tracker 注册后，就可以获取 seed 的信息，与 seed 建立连接。而其他 peer 则接收一个随机 peer 表，与其他下载者建立连接。

源文件和 tracker 的响应都采用一种简单有效可扩展的格式，被称为 bencoding，它可以包含字符串和整数。由于对不需要的字典关键字可以忽略，所以这种格式具有可扩展性，其他选项可以方便的加进来。用户下载的.torrent 文件就是“Metainfo file”，里面存储有关于下载内容的 announce 地址、长度、大小等内容。客户端根据 Metafile 指出的“announce”地址，对 tracker 提出申请，Tracker 将返回一个“text/plain”文档，含有 Bencode 编码的字典，其中一个字段是“peers”，一个含有字典的列表，每一个字典含有如下内容：

表 1 原 peers 列表字段

Peer id	ip	port
---------	----	------

其中，peer id(字符串)：peer 的唯一名称。ip(字符串)：peer 的 IPV4 或 IPV6 地址，或是 DNS 名。Port(整数)：peer 的端口。

因为本文采用的思想是将域间流量本地化，基于用户同意协作方面，就可将用户信息收集，扩展部分列表字段，从而确定 peer 的选择优先级，尽量将 peer 选在控制在局域网或者同一个 ISP 内。

在一个 peer 向 tracker 注册的时候，其 ip 字段添加

一些内容，如表 2 所示：

表 2 peers 列表扩展字段

ip: 原 ip 地址	as: 所在 AS 号	ad: 用户地址	isp: 运营商
-------------	-------------	----------	----------

其中，用户地址的设定编码如下：取省的简称的第一个拼音字母作为该省的标示符，后面是市的拼音简写。ISP 的设定编码如下：同样取运营商拼音简写的首字母组合而成。

然后在 peer 列表节点选取时，不是以往的随机挑选一些节点提供给用户，而是有个筛选过程：根据字段的排列优先级。具体可归类为以下列表：

表 3 基于用户协作的节点分类

AS 号相同	一级节点
AD 码相同	二级节点
AD 码第一个字母相同 +ISP 相同	三级节点
AD 不同，ISP 相同	四级节点
ISP 不同	五级节点

这是初步的优先级列表，主要存于 tracker 服务器上，并在返回节点列表的同时作出筛选。在现实生活中，用户的协作性并不高，这时就需要根据小节 1.2 中的算法更新个人优先级列表。

1.2 基于用户拒绝协作的优化

以上是建立在用户协作基础上的，真正实施起来有一定的困难性。本小节就是针对这一情况进行优化。

在整个网络中，我们知道 IP 地址一定局部区域连续，大部分区域不连续^[5]，这就使得直接通过 IP 地址判断用户的物理拓扑比较困难^[6]。

根据 BT 协议，每一次一个 peer 请求资源，tracker 都会返回一个随机列表，然后 peer 与该列表中的其他 peer 进行握手连接。但是，这个随机列表中的 peer 可能是和该 peer 位于同一域，也可能跨域。首先，假设一个节点 peer P 与 tracker 返回的随机列表进行连接。

$P \rightarrow (\text{peer1}, \text{peer2}, \dots, \text{peer n})$

通过两次握手连接之后，peer1 从其中一个 peer 开始下载文件片段。其中，与多个 peer 连接是同时进行的。但是由于地理位置，运营商及其他的不同条件，peer1 与其他 peer 连接的时间也不一样。我们设定一个参数为 T，T1 为连接上 peer1 的时间，依次类推。

$P \rightarrow ((\text{peer1}, T1), (\text{peer2}, T2), (\text{peer3}, T3), \dots)$

因为环境的不同，单按照时间来断定误差性比较

大, 所以引出一个值 B , 即为连接上节点的下载速率.

$P \rightarrow ((peer1, T1, B1), (peer2, T2, B2), (peer3, T3, B3), \dots)$

根据多次试验, 在局域网的两个节点互相传递文件的速率远远大于其他跨网的节点, 但是, 也会有一些环境因素, 单从下载速率判断也会有一定的误差性. 所以引入参数值 $K(K=T/B)$, 建立一个本地 peer 列表 S .

K 越大, 则该节点的可取性越低, 反之, K 越小, 该节点的可取性越大. K 的取值可通过实验确定范围. 此时根 K 值以从小到大排序, P 的 S 列表更新为:

$P: S(p1, p2, \dots, pn), K_{p1} < K_{p2} < \dots < K_{pn}$

这个列表是直接存储在本地的. 每次 P 从 tracker 处得到随机列表, 便可根据该算法更新本地的 S 内容. 这样每一个节点都有属于自己的 S 列表^[7].

当 P 与随机列表 $peer1$ 连接时, 两个节点可以互相匹配各自的 S 列表, 这里会设定一个匹配值 $N(0 < N < 100)$, 如果 N 处于一定范围内, 可预想这两个节点处于同一网间. 可将对方节点加入自己的 S 列表中, 并处于高优先级位置, 等待下次下载资源, 可首先连接对方节点. 具体优先级设定在 2.3 小节中提出.

1.3 综合两种优化算法实现

在 1.2 中, 每个用户本地都会储存属于自己的 S 列表, 并根据每次下载或者上传时不断的更新列表内容. 当两个 peer 的 S 列表匹配度达到 N 的时候, 对方 peer 就会加入自己的 S 列表中, 并根据匹配度处于一定的优先级位置上. 引入临界值 $\mu > 0$, 参数 $k > 0$, 当 $N < \mu$ 的时候, 两个节点完全不同, 当 $N \geq 3k\mu$ 的时候, 两个节点近似在同一级子网内. 表 3 可更新如下:

表 4 综合节点优先级列表

AS 号相同/ $N \geq 3k\mu$	一级节点
AD 码相同/ $3k\mu > N \geq 2k\mu$	二级节点
AD 码第一个字母相同+ISP 相同/ $2k\mu > N \geq k\mu$	三级节点
AD 不同, ISP 相同/ $k\mu > N > \mu$	四级节点
ISP 不同/ $0 < N < \mu$	五级节点

具体的优化流程思想见图 1. 首先, 客户 peer 向服务器请求资源, 服务器根据用户协作返回一定的节点列表, 若此时还未有用户加入则返回随机列表. 然后, 该 peer 与列表中的节点连接进行资源下载, 与此同时,

与连接的节点交换彼此的 personal S 列表. 在更新列表的时候, 该 peer 会从本地存储的个人 S 列表中挑选优先级高的节点连接进行资源下载. 其中, 列表的更新和资源的下载可以同时进行.

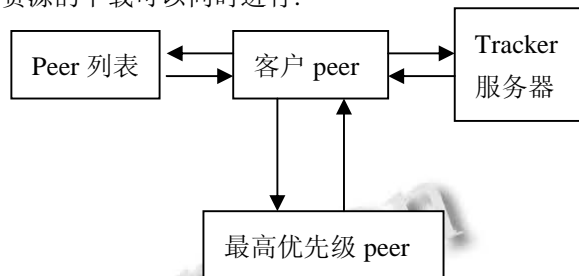


图 1 算法连接图

2 优化机制建模与评估

首先直观上了解该算法, 在采用了安置优先级的节点列表进行选择后, 由于在匹配度 N 比较高或者物理位置比较靠近的节点之间优先传输文件资源^[8], 可最大限度减少对上一级网络线路的使用, 特别在用户已经有相当完整程度的 S 列表的时候, 效果更加明显.

2.1 优化方案的理论推导评估

构建一个简单网络模型, 为了降低复杂度, 可认为同一网络中节点有相同的带宽, 设定环境理想, 时延根据跨域跳数决定.

此时, 当两个节点处于同一级子网时, 它们的跳数为 1, 设定它们优先级为 1. 处于相同一级子网不同二级子网时, 跨越的跳数为 3, 设定它们优先级为 2. 依次类推, 每次 $n-1$ 级子网相同, n 级子网不同的时候, 跨越的跳数为 $2n-1$, 优先级为 n . 而常见的 TTL 就是以路由跳数计算时延.

由小节 2.2 的算法可知, $T=2n-1$

$$K=T/B \tag{1}$$

在各个节点有相同的下载速率的时候, 处于相同子网的下载速率大于外网的下载速率, 近似成等差数列. 设定 1 级节点的速率为 B , 最大连接数是 M , 则引入参数值 x , 使得 n 级节点速率为

$$B_n=B-(n-1)x \tag{2}$$

将公式(2)代入公式(1), 得出结果如下:

$$K=(2n-1)/(B-(n-1)x) \tag{3}$$

最后由公式推导出优先级 n , 结果如下:

$$n = \frac{2B+x}{2x+\frac{1}{K}} + 1 \tag{4}$$

由公式(4)可看出,当 K 越大的时候表示节点跨越的跳数越多,两个节点所处的网络位置差距越大,基本验证了小节 2.2 中算法的可行性。

2.2 优化方案的实验评估

BT 实验工具选择 Vuze4.7.0.2(基于 java 平台,可添加插件),实验环境选定在校园网内外,忽略部分误差影响。实验发现,由于 BT 节点反馈的随机性,peer 并不容易全部连接比较接近的节点,反而有可能部分连接到上一级网络的的节点^[9]。不仅仅下载速率受到影响,并会产生一定的跨域流量。实验在设定 $\mu=10$, $k=2$ 时,平衡机制达到最优。

基于测量的测试标准,分别从以下两个图表体现了本文优化机制的优越性。X 轴为测试参数, Y 轴为概率分布。

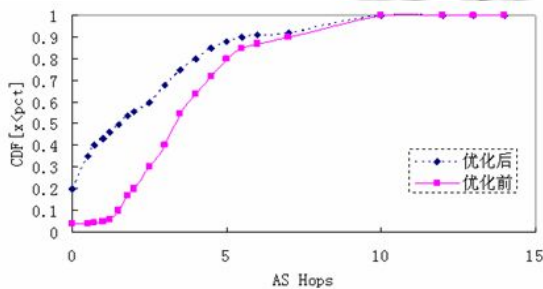


图 3 平均 AS 跳数

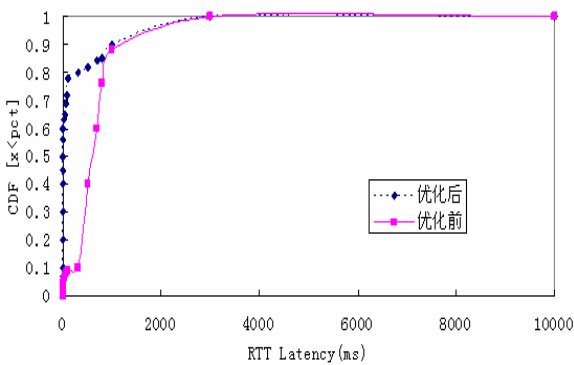


图 4 平均 ICMP ping 往返时延概率分布

由图 3 和图 4 可看出,经过优化后,节点连接之间经过的跳数比原来随机的改进很多,这样不仅节省了

跨域的带宽占用,而且根据往返时间可以看出,在域内的节点会有更好的资源下载速度,优化后的节点可靠性也提高了。

3 结语

本文主要的思想在于将流量尽可能的控制在本地,以减少骨干网络上带宽的占用。这一思想不仅仅是对于骨干网上流量堵塞的“解放”,而且由于节点优先级列表对于资源的优化选择,增加节点的可靠性^[10],更是增强了用户的使用性能。但是本文只是基于 BT 协议作出优化,对于其他协议及其他下载工具的优化研究还需改进。

参考文献

- 1 张国强,唐明董,程苏琦,张国清,宋海滨,曹蓟光,杨景.P2P 流量优化.中国科学:信息科学,2012,42(1):1-19.
- 2 Xie HY, Krishnamurthy A, Silberschatz A, et al. P4P: Explicit communications for cooperative control between P2P and network providers. P4PWG, 2007: 32-56.
- 3 David RC, Bustamante FE. Taming the Torrent: A Practical Approach to Reducing Cross-ISP Traffic in Peer-to-Peer Systems. SIGCOMM, 2008: 363-374.
- 4 刘华春.P2P 网络的分类以及关键技术分析.微计算机信息, 2008,24(3-3):112-114.
- 5 欧阳荣,苗卉,雷振明.一种减少网间 P2P 流量的 Peer 选择算法.计算机工程,2008,38(8):108-110.
- 6 王波,李志蜀.基于“伪 IP”匹配的 P2P Peer 选择算法.云南大学学报(自然科学版),2008,(S2):257-261.
- 7 李峰.校园网中 P2P 流量监管与控制.网络安全技术与应用,2008,(3):115-117.
- 8 傅建明,孙芳琪,刘力,张焕国.P2P 中一种使用缓存和测量的节点选择模.电子学报,2006,(8):1545-1548.
- 9 郭振滨,裘正定.P2P 资源流行度测量与建模.通信学报, 2009,30(3):66-74.
- 10 郭涛,周旭,王治平,唐晖.基于网络测量的 P2P 跨域流量优化机制.计算机应用,2010,30(4):889-891.