

TCP-HY: 优化 TCP 在混合 IP 网络环境中的效率

徐航宇, 罗 键

(厦门大学 信息科学与技术学院, 福建 厦门 361005)

摘 要:在混合网络环境下, 由于数据包传输的非时序性以及传输过程中相对较高的误码率, TCP 的传输效率极大地被降低了。对此, 提出了一种新的 TCP 改进版本: TCP-HY, 它能够有效解决数据包延迟以及高误码率的问题。试验表明, TCP-HY 在混合网络环境中能够大大提高 TCP 的效率, 充分利用有效带宽, 是一种优秀的解决方案。

关键词: TCP-HY; 混合网络; 带宽估计; 拥塞控制

中图分类号: TP393

文献标识码: A

文章编号: 1673-629X(2007)11-0004-04

TCP-HY: Improving TCP Performance in Hybrid IP Network

XU Hang-yu, LUO Jian

(School of Information Science and Technology, Xiamen University, Xiamen 361005, China)

Abstract: In hybrid network environment, since the package often disordered during transferring and the channel error can not be ignored. The efficiency of TCP protocol is dramatically depressed. In this paper, propose TCP-HY, a set of simple modify to the TCP protocol to improve the efficiency in hybrid network environment. Evaluate TCP-HY through simulation to show that it offers significantly better performance when the packages disorder ratio and the bit error ratio can not be ignored. It is a good solution in hybrid network environment.

Key words: TCP-HY; hybrid network; bandwidth estimate; congestion control

0 引 言

随着通信技术的发展, 无线通信技术日益成熟, 人们对通信的方式和质量提出了更高的要求, 传统的有线网络或无线通信网已经不能满足人们的需要。将有线网络和无线通信网络无缝融合, 构造一个全 IP 有线-无线混合网络将是通信网络的一个主要发展方向。然而, TCP 协议在混合网络中的效率十分低下, 远远不如其在传统有线网络中的效率。其主要原因是由于 TCP 过于简单的拥塞控制算法^[1]造成的: 当 TCP 发送端连续收到一个数据包的 3 个重复 ACK 时, 就认为该数据包在网络中由于网络的拥塞而被丢弃了, 随之马上启动拥塞避免算法, 将发送窗口和快启动门限设置为当前发送窗口的一半, 降低发送速度以缓解网络拥塞状况。由于在传统有线网络中, 数据传输中的误码率(BER)十分低, 并且数据包乱序的情况并不严重, 发送端连续收到重复的 ACK 往往意味着该数据包因网

络发生拥塞而被丢弃。因此, 拥塞避免算法在传统有线网络中运行十分有效。然而, 在现今的网络中, 数据包乱序的情况更加频繁^[2], 数据发送端单单等待 3 个重复 ACK(DUPACKS)就启动拥塞避免算法已经逐渐不能适用。并且, 在混合网络中, 由于无线网络的数据误码率远远高于有线网络, 即使在链路层运用本地重传的方法也不能保证数据误码率足够低^[3], 而且本地重传也将大大增加 ACK 乱序的概率。这些都导致了 TCP 在混合网络环境中错误地将数据包传输错误或乱序当作网络拥塞来处理, 不必要地降低了传输速度, 导致传输效率低下^[4]。

针对有线-无线混合网络中较高的数据包乱序传输率和较高的 BER 的情况, 文中提出了一种新的 TCP 解决方案——TCP-HY。

1 TCP-HY

在有线-无线混合网络中, 如何有效地区分 DUPACKS 产生的原因并采取相应的措施是提高 TCP 传输效率同时保证其友善度的最根本问题。TCP-HY 能够在发送端收到 DUPACKS 之后分析其产生的原因, 从而采取相应的措施, 从根本上很好地解决了这个

收稿日期: 2007-02-27

作者简介: 徐航宇(1982-), 男, 福建建阳人, 硕士研究生, 研究方向为复杂网络优化; 罗 键, 教授, 博士生导师, 研究方向为系统控制与优化、复杂网络优化。

问题。

TCP-HY 的主要原理如下: TCP-HY 将错误恢复的责任交给无线网络的本地链路层进行处理^[3], 从而把关注的焦点放在网络传输层的拥塞状况以及相对较低的 BER 的情况上。TCP-HY 发送端通过接收的 ACKS 通过估算网络的带宽^[5], 记为 $BWE_CURRENT$ 。当 TCP 发送端收到第一个 DUPACKS 时, 将此时网络带宽值记录为 $BWE_DUPACKS$ 并暂时保存下来。并等待 λ (一般将 λ 设置为一个 RTT 时间 (在动态网络情况下, 稍微延长响应时间是安全且公平的^[6]); 在时间 λ 内, TCP-HY 发送端保持发送窗口大小不变, 根据 pipe 原理继续发送数据; 同时, 根据接收的 ACKS 继续更新网络带宽 $BWE_CURRENT$ 。若在 λ 内发送端收到新的 ACKS, 则返回常规 TCP 发送状态; 否则, 若 λ 超时发送端还未接收到新的 ACKS, 则比较 $BWE_DUPACKS$ 与 $BWE_CURRENT$ 以决定如何采取进一步的处理措施。

若 $BWE_DUPACKS < BWE_CURRENT * \eta$ (η 为设置的约束参数, $0 < \eta \leq 1$), 则表示在时间 λ 内, 网络带宽并没有降低, 即此时网络并未处于拥塞状态。产生 DUPACKS 的原因可能是数据包在传输中发生了错误而并没有被本地链路层有效恢复, 或者该数据包由于非拥塞的原因发生了严重的延迟。接下来, 计算在时间 λ 内网络带宽增加的程度 BWE_INC , 根据计算所得的 BWE_INC , 更新当前合适的发送窗口 $SSTHRESH$ 的值以及慢启动门限 $CWND$, 相关程序代码如下:

```
IF (BWE_DUPACKS < BWE_CURRENT * η)
{
    BWE_INC = (BWE_CURRENT - BWE_DUPACKS) / BWE_DUPACKS;
    SSTHRESH = INT(SSTHRESH * (1 + BWE_INC));
    if (SSTHRESH > WINDOW()) {
        SSTHRESH = WINDOW();
    }
    ELSE IF (SSTHRESH < SSTHRESHBWE) {
        SSTHRESH = SSTHRESHBWE;
    }
    CWND = SSTHRESH;
}
```

其中, $WINDOW()$ 为当前发送窗口大小, $SSTHRESHBWE$ 为仅由 $BWE_CURRENT$ 直接计算出的发送窗口的大小。将 $SSTHRESH$ 限制于 $[SSTHRESHBWE, WINDOW())$ 内是为了避免 TCP-HY 对其他 TCP 版本过于具有侵略性或过于保守。

若 $BWE_DUPACKS \geq BWE_CURRENT * \eta$,

则表示在该时间段 λ 内网络带宽接近或低于收到第一个 DUPACKS 时的网络带宽, 说明此时网络拥塞情况与之前类似或有恶化。此时, 启动拥塞避免算法, 将发送窗口和慢启动门限减少为当前窗口的一半:

```
ELSE IF (BWE_DUPACKS ≥ BWE_CURRENT * η) {
    SLOWDOWN(CLOSE_SSTHRESHHALF | CLOSE_CWND_HALF);
}
```

$SLOWDOWN(CLOSE_SSTHRESHHALF | CLOSE_CWND_HALF)$ 的作用是将发送窗口和慢启动门限减少为当前窗口的一半。

由于考虑到数据恢复的效率问题, 本地链路层一般只执行一次错误重传, 在一次错误重传发挥作用的情况下, 若无线网络误码率低于 10% 时, 则经本地重传恢复后的网络误码率即可低于 1%, 本地链路重传虽不能完全修复数据传输错误, 但可以有效地将误码率降低。降低后的误码率与 TCP-HY 有效的拥塞控制策略结合起来, 可以大大提高在复杂网络环境下的传输效率。

2 仿真分析

使用 NS2^[7] 对 TCP-HY 进行仿真分析, 网络结构如图 1 所示。图 1 中 $Scr(i)$ 为发送端, $Sink(i)$ 为数据接收端, 发送端与 R1 直接相连, 网络带宽设为 100M 链路延迟设为 1ms, 接收端与 R2 直接相连, 同样地, 网络带宽设为 100M 链路延迟设为 1ms。R1 与 R2 之间的链路 L1 为网络瓶颈, 网络带宽设为 10M, 链路延迟设为 50ms。发送端向接收端发送的数据包大小为 1000bytes。R1 与 R2 使用 DropTail 缓冲区管理, 缓冲区大小设定为 50 个数据包。

在链路 L1 上附加一个错误模型来模拟网络延迟以及网络错误情况, 具体设定依照不同试验内容而定。发送端依照不同情况使用 TCP-Sack^[8], TCP-Dcr^[9], TCP-WestWood^[5] 以及 TCP-HY 协议, 接收端统一使用 TCP-Sink/Sack1^[8] 协议。

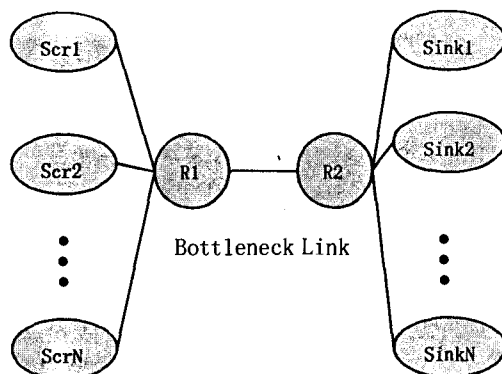


图 1 网络结构

2.1 不同延迟率的情况下各种 TCP 解决方案比较

令网络中被延迟数据包的延迟时间呈泊松分布,其分布平均值为 30ms,标准差为 10ms,当数据包的延迟率不同时,各种 TCP 方案的效率如图 2 所示。从图中不难看出,TCP-DCR 与 TCP-HY 对网络中出现延迟的数据包的情况处理的很好,其效率远远高于 TCP-SACK 与 TCP-WESTWOOD。在延迟率不高的情况下,TCP-HY 的效率略高于 TCP-DCR。

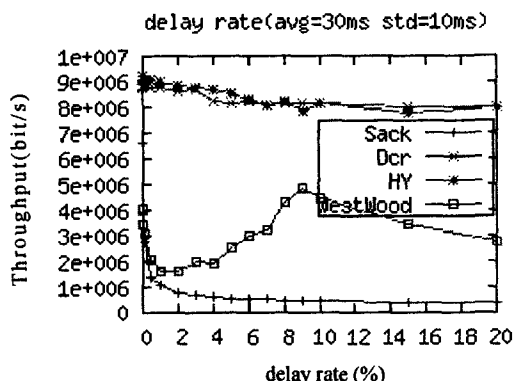


图 2 不同延迟率时各种 TCP 性能比较

2.2 不同延迟时间的情况下各种 TCP 解决方案比较

令网络中数据包的延迟率不变,使其恒定为 5%,在不同延迟时间的情况下,各种 TCP 方案的效率如图 3 所示。TCP-HY 与 TCP-DCR 在不高的网络延迟的情况下(低于 40ms)表现十分好,当网络延迟高于 50ms,其性能下降比较明显,网络延迟高于 100ms 时其性能直线下降,到达和 TCP-SACK 与 TCP-WESTWOOD 相同的水平。而 TCP-SACK 与 TCP-WESTWOOD 当网络中数据包延迟高于 5ms 时,其性能就直线下降到很低的水平,效率很低。

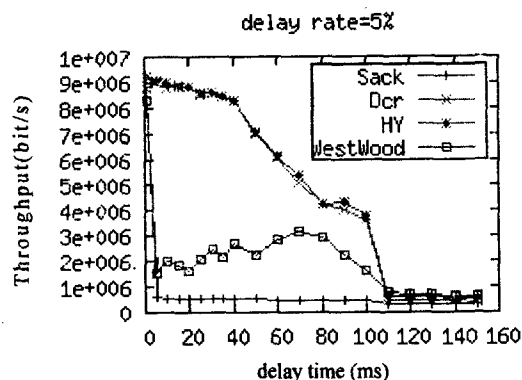


图 3 不同延迟时间时各种 TCP 性能比较

2.3 不同误码率的情况下各种 TCP 解决方案比较

当本地链路层重传不能完全消除 BER。各种 TCP 方案在不同 BER 的情况下的传输效率如图 4 所示。其中 TCP-DCR 与 TCP-SACK 效率基本相同,效率最低。而 TCP-HY 与 TCP-WESTWOOD 的表现要好一些,当 BER 比较低时(低于 0.05%),TCP-

HY 表现较好,当 BER 较高(高于 0.05%)时,TCP-WESTWOOD 的表现要强一些。

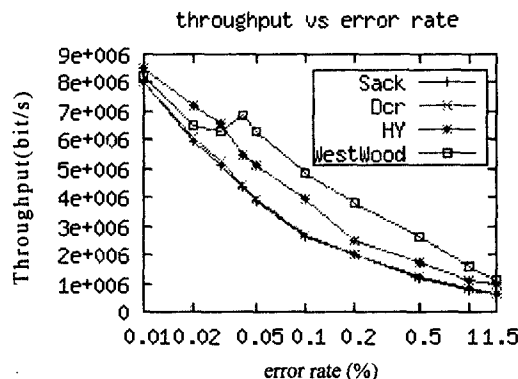


图 4 不同误码率时各种 TCP 性能比较

2.4 各种 TCP 解决方案的友善度比较

除了传输效率外,友善度是衡量 TCP 方案是否优秀的另一项重要指标,由于网络上的主机数量巨大,不可能只使用一种解决方案,必然要考虑多种方案共存时它们之间的相互影响。若一种 TCP 方案十分高效,但对其它的 TCP 方案效率造成了很大的抑制,称这种方案带有侵略性,或友善度不高,不适合部署应用。

令网络瓶颈带宽为 10M,链路延迟为 50ms,存在 20 个发送端同时发送数据,其中 10 个发送端应用 TCP-SACK 方案,另外 10 个发送端分别应用 TCP-DCR, TCP-WESTWOOD 或 TCP-HY。以 TCP-SACK 为比较对象,其它不同的解决方案与之共存时,在不同的 BER 的情况下各自的平均效率如表 1 所示。其中,第一列为 BER,分别取值为 0, 0.01, 0.05, 0.1, 0.5; 平均传输速度单位为 M bit/s。

表 1 不同方案共存时友善度比较

	Sack/DCR		Sack/Westwood		Sack/HY	
0	0.490	0.428	0.318	0.610	0.455	0.456
0.01	0.478	0.435	0.320	0.605	0.463	0.450
0.05	0.479	0.438	0.341	0.593	0.457	0.464
0.1	0.480	0.437	0.352	0.581	0.451	0.472
0.5	0.458	0.482	0.361	0.578	0.438	0.507

由表 1 不难看出,TCP-DCR 与 TCP-SACK 在不同 BER 时的平均传输速度稳定(在 SACK=DCR 基准线附近),说明 TCP-DCR 对 TCP-SACK 有着良好的友善度。而 TCP-WESTWOOD 在不同 BER 的情况下对带宽的占用率大大高于与之同时工作的 TCP-SACK,压制了 TCP-SACK 的效率,因此 TCP-WESTWOOD 的带宽占用策略具有侵略性,尤其是在网络 BER 很低的情况下其友善度不好。TCP-HY 在 BER 比较低时,具有很好的友善度,而当 BER 较高时,其友善度稍微有所降低。然而在本地链路层发挥

作用的情况下,网络整体 BER 一般可以控制在 0.5% 以下,所以说,TCP-HY 的友善度还是完全可以接受的。

3 小 结

提出了一种新的 TCP 方案——TCP-HY。它通过延长发送端的拥塞响应时间和比较在这段时间内估算的网络带宽,从而评估网络的拥塞状况,进而根据不同的评估状况做出不同的响应策略。从而使得做出的策略更加有针对性更加准确。这样避免了传统 TCP 协议盲目地采用拥塞避免算法而导致的传输效率的不必要的降低,有效提高 TCP 的传输效率。并且通过一系列的仿真试验表明,在有线-无线混合网络中,无论是在网络状况良好的环境下,还是在网络状况不理想的环境下,TCP-HY 均保持着较高的传输效率,并且对其它 TCP 方案表现出了良好的友善度。由此,有理由相信,TCP-HY 在有线-无线混合网络环境中,是一种优秀的解决方案。

参考文献:

- [1] Steven W R. TCP/IP Illustrated Volume 1: The Protocols [M]. New York, USA: Addison-Wesley, 1994.
- [2] Bennett J, Partridge C, Shectman N. Packet Reordering is Not Pathological Network Behavior [J]. IEEE/ACM Transaction on Networking, 1999, 7(6): 789-798.

(上接第 3 页)

面进行凸分解,若处理碰撞精度要求高的相交问题,可以以分解后的凸块集为基础,判定凸块之间的相交,从而简化精度碰撞预处理工作。

文中采用的 OBB 以矩形为包围盒,由于凸块形状各异,有些包围盒并不具有高包含度^[10],因此,优化包围盒形状,使其高度包含凸块,以进一步提高碰撞检测效率成为以后工作的重点。

参考文献:

- [1] 汪成为,高文,王行仁. 灵境(虚拟现实)技术的理论、实现及应用[M]. 北京:清华大学出版社,1996.
- [2] 范昭伟,万华根,高曙明. 基于图象的快速碰撞检测算法[J]. 计算机辅助设计与图形学学报, 2002, 14(9): 805-809.
- [3] Chung K, Wang W. Quick elimination of non-interference polytopes in virtual environments[C]//3rd European Workshop on Virtual Environments. London, UK: Springer-Verlag, 1996: 171-178.
- [4] 魏迎梅. 虚拟环境中碰撞检测问题的研究[D]. 长沙:国防

- [3] Eckhardt D A, Steenkiste P. Improving Wireless LAN Performance via Adaptive Local Error Control[C]//Sixth IEEE International Conference on Network Protocols (ICNP'98). Washington DC, USA: IEEE Computer Society, 1998.
- [4] Balakrishnan H, Padmanabhan V, Seshan S, et al. A Comparison of Mechanisms for Improving TCP Performance over Wireless Links[J]. IEEE/ACM Transactions on Networking, 1997, 5(6): 756-769.
- [5] Mascolo S, Casetti C, Gerla M, et al. TCP westwood: Bandwidth estimation for enhanced transport over wireless links [C]//Proceedings of the 7th annual international conference on mobile computing and networking. Rome, Italy: [s. n.], 2001: 287-297.
- [6] Bansal D, Balakrishnan H, Floyd S, et al. Dynamic Behavior of Slowly Responsive Congestion Control Algorithms [C]//Sigcomm 2001. San Diego, CA: [s. n.], 2001.
- [7] NS2 Network Simulator. A collaboration between researchers at UC Berkeley, LBL, USC/ISI, and Xerox PARC [EB/OL]. 2003-02-28. <http://www.isi.edu/nsnam/>.
- [8] Wierman A, Osogami T, Ols'en J. A unified framework for modeling TCP-Vegas, TCP-SACK, and TCP-Reno [C]//in Proc IEEE/ACM MASCOTS. Orlando, Florida, USA: [s. n.], 2003: 269-278.
- [9] Bhandarkar S, Sadry N E, Reddy A L N, et al. TCP-DCR: A Novel Protocol for Tolerating Wireless Channel Errors [J]. IEEE Transactions on Mobile Computing, 2005, 4(5): 517-529.

科技大学, 2000.

- [5] 王兆其, 赵沁平, 汪成为. 面向对象碰撞检测方法及其在分布式虚拟环境中的应用[J]. 计算机学报, 1998, 21(10): 990-994.
- [6] 吴明华, 余永翔, 周济. 采用空间分割技术的八叉树干涉检测算法[J]. 计算机学报, 1997, 20(9): 849-854.
- [7] Cameron S. Collision detection by four-dimensional intersection testing[J]. IEEE Trans Robot Autom, 1990, 6(3): 291-302.
- [8] Van den Bergen G. Efficient collision detection of polytopes in virtual environments [C]//Proceedings of the ACM Symposium on Virtual Reality Software and Technology. [s. l.]: [s. n.], 1996: 125-131.
- [9] Ballard D H. Strip trees: A hierarchical representation for curves[J]. ACM Communication, 1981, 24(5): 310-321.
- [10] Adelson S J, Hodges L F. Generating exact ray-traced animation frames by reprojection[J]. IEEE Computer Graphics and Applications, 1995, 15(3): 43-52.
- [11] Goldsmith J, Salraon J. Automatic creation of object hierarchies for my tracing [J]. IEEE Computer Graphics and Applications, 1987, 7(1): 14-20.