TCP 连接往返时延的被动测量算法实验验证

周 波[1],张代远[1,2]
(1. 南京邮电大学 计算机学院,江苏 南京 210003;
2. 江苏省无线传感网高技术研究重点实验室,江苏 南京 210003)

摘 要:引用了一种应用于宽带网络环境的被动式环回时间(RTT)测量算法,用于对传输控制协议(TCP)的报文环回时间进行估计。该算法通过估计同一轮次报文的发送间隔来挑选相邻 2 个发送轮次之间的间隙,进而估算出 TCP 的报文环回时间,该方法的测量对象和测量结果更加具有网络管理意义,能够在不同的 TCP 行为模式中获得更多的测量采样并显著提高测量成功率。文中通过设计实验,设计程序分析所得的测量值与实际所得的 RTT 比较,验证该算法的有效性。

关键词:传输控制协议;环回时间;被动测量;参数估计
中图分类号:TP31 **文献标识码**:A **文章编号**:1673-629X(2012)09-0083-04

A Passive TCP Connection Round-trip Time Measurement Algorithm Experimental Verification

ZHOU Bo[1], ZHANG Dai-yuan[1,2]
(1. College of Computer, Nanjing University of Posts and Telecommunications, Nanjing 210003, China;
2. Jiangsu High Technology Research Key Laboratory for Wireless Sensor Networks, Nanjing 210003, China)

Abstract: Passive round trip time (RTT) estimate algorithm called PRE (passive RTT estimate) is proposed for the traffic monitor devices to estimate RTT of TCP connections in the broadband network. PRE algorithm can find out sending rounds of the monitored TCP flow according to the estimator of package arriving interval (PAI) and estimates RTT by measuring the length of the sending round, which approximately equals to RTT. This method could obtain more measurement samples from different TCP activity models, thus observably improve the measurement success rate. In this paper, design an experiment and program analyzing the measured values and actual RTT values to verify the effectiveness of this algorithm.

Key words: transmission control protocol; round trip time; passive measurement; parameter estimation

0 引 言

往返时延(RTT)被定义为从源端发出一个包到接收到通信对端确认 ACK 包所经历的时长,是影响 TCP 性能和表征网络运行状况的重要参数。作为网络性能的重要指标,路径 RTT 的测量一直受到重视。被动测量 RTT 因为其巨大的应用背景,所以引起了中外广大学者的关注和研究。Hao Jiang 等人在文献[1]中提出 SYN ACK 方法,通过监测 TCP 三次握手建立连接的过程来估计 RTT,用监测到的最后一个 SYN 和第一个 ACK 之间的时间间隔来计算 RTT。Bryan Veal 等人在文献[2]中提出基于时间戳匹配的 RTT 测量方法。文中是根据张轶博博士在文献[3]中提出一种基于参数估计的被动式 TCP 环回时间测量算法,进行实验验证算法的有效性。

收稿日期:2012-01-16;**修回日期**:2012-04-27
基金项目:江苏省高校优势学科建设工程资助项目(yx002001)
作者简介:周 波(1987-),男,江苏连云港人,硕士,研究方向为网络测量、神经网络;张代远,教授,博士,研究方向为神经网络、人工智能等。

1 被动式环回时间测量算法

首先明确几个概念:同轮次包发送间隔(PSI, Packet Sending Interval)、轮次间隙(RG, Round gap)、到达间隔(AI, Arriving Interval)以及同轮次包到达间隔(PAI, Packet Arriving Interval)。PSI 定义为相邻的且同轮次的 2 个发送包之间的间隔;RG 定义为前一发送轮次最后一个包与后一轮次第一个包之间的间隔;AI 定义为测量点连续接收的 2 个包的到达间隔;PAI 定义为同轮次且相邻的 2 个包到达测量点的时间间隔。在 TCP 拥塞控制机制的作用下,RTT 源端会连续发送多个包,一直到到达发送窗口大小限制,接下来必须收到目标端的确认报文后才能继续发送[4]。这样 TCP 发送行为呈现出多个轮次,在高速链路中包发送间隔远小于轮次间隔,可以将测量到的较大的包到达

间隔认定为轮次间隔。将轮次间隔加上本轮前面的包到达间隔就得到了 RTT。这个方法的主要问题在于当观测的 TCP 流传输中出现停顿时,测量点难以将其与轮次间隔时间区分[5]。PRE 算法(Passive RTT Estimate)的重要目标是,从测量点获得的到达间隔中分离出轮次间隔,进而得到发送轮次并计算出其时长[6]。由于测量点测量到的到达间隔由同轮次包到达间隔和轮次间隔组成的,要准确地区别轮次间隔就必须对同轮次包到达间隔与轮次间隔进行区分。PRE 算法对轮次间隔进行遴选:因同轮次包到达间隔必然服从高斯分布[7],由 3δ 原则制定判决门限 $Tth = E[PAI] + 3\sqrt{D[PAI]}$,如到达间隔 AI 大于门限则判为 RG;反之判为 PAI。算法采用 Tth 作为判决门限,$\alpha \in (1,2)$;同时,为了避免将传输超时误判为 RG,算法将 RTT 估值的 2 倍作为 RG 的上限,设定最终的 RG 判决准则如式(1),其中 Tn 为检测到的某个 RG,RTTE 为已经获得的环回时间估计值[3]:

$$\alpha Tth < Tn < 2RTTE \quad (1)$$

PRE 算法的实现流程如图 1 所示。

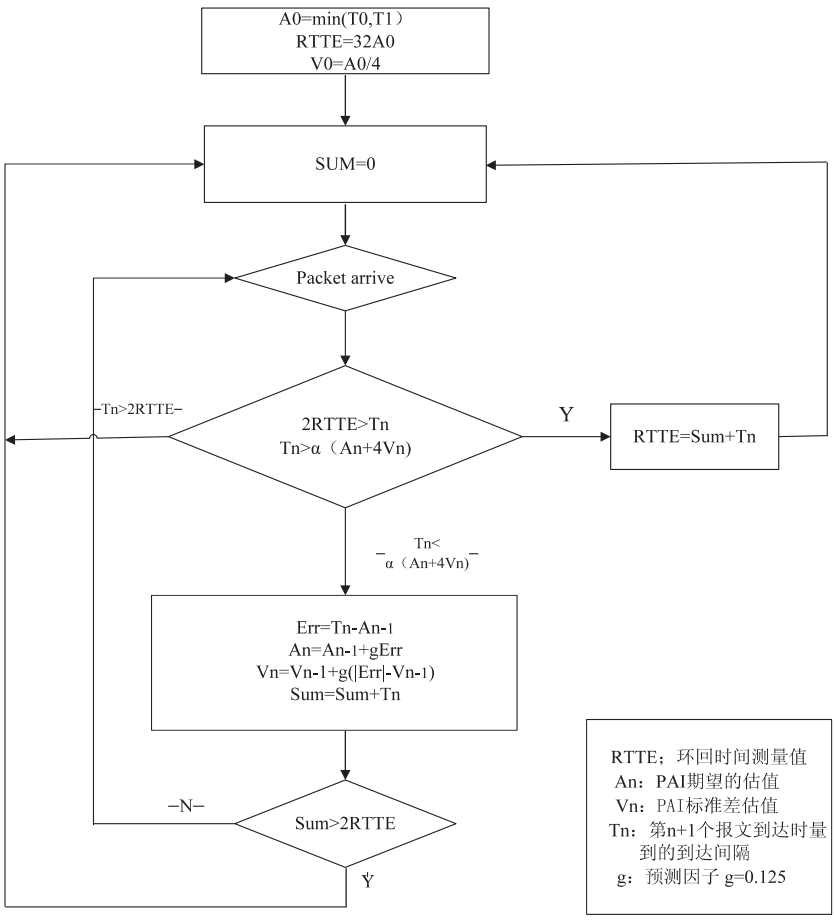

图 1 PRE 算法流程图

2 实验步骤

本节要根据上一节所描述的算法具体实现被动测量 RTT 的方法。首先完成的工作是网络抓包,使用的软件是 Wireshark,该软件是一款功能强大的网络抓包分析软件,可以获取网络数据并可以对抓包数据进行分析。通过这款软件获取被动测量 RTT 所需要的实验数据,为以后设计程序进行实验提供数据。同时,需要连接到一个 FTP 网站,并且选择一个文件进行下载,目的是使得此时网络拥塞,已得到更多的轮次间隔,满足实验要求。同时在自己的主机上每隔 2 秒不断进行 Ping 连接上述的 FTP 网站,获得主动测量 RTT(RTTM)的数据,并将这些数据保存到文档中。最后按照第一节提到的算法设计 C 语言程序对抓包数据进行分析得到被动测量 RTT,并将这些数据与通过 Ping 命令得到的主动测量 RTT 进行比较。进而分析上节提到的 PRE 算法有效性。

为了实验准确性的评判,需要准确的 RTT 时间,作为 PRE 算法是否有效的判定标准,需要得到主动测量 RTT 的值。这里就用到 Windows 系统里的 Ping 命令,Ping 是用来测试两个主机之间的连通性的命令。Ping 使用了 ICMP 回送请求与回送回答报文。Ping 是应用层直接使用网络层 ICMP 的一个例子,它没有通过运输层的 TCP 或 UDP。使用 Ping 命令时,PC 机一连发出四个 ICMP 回送请求报文。如果 FTP 工作站正常工作而且相应这个 ICMP 回送请求报文(有的主机为了防止恶意攻击就不理睬外界发送过来的这种报文)[8],那么它就发回 ICMP 回送回答报文。由于往返 ICMP 报文都有时间戳,因此很容易得出往返时间。最后显示出的是统计结果:发送到那个机器(IP 地址)、发送的、收到的和丢失的分组数(当不给出分组丢失的原因)。往返时间的最小值、最大值和平均值。为了将 Ping 命令得到数据保存在硬盘上,使用 Visual C++编写这样一个程序,它每隔两秒定时 Ping 此 FTP 网站,在 DOS 命令下打开该可执行 exe 文件,并通过 DOS 命令将这个程序调用 Ping 命令得到的数据保存到一个 txt 文件中。通过这种方式,就得到了准确的用主动测量方法得到的 RTT 的值。图 2 为在实验时 DOS 窗口的截图。

在使用 Ping 命令获得主动测试得到的 RTT 的同时,就要开始用 Wireshark 进行抓包。为了以后处理数

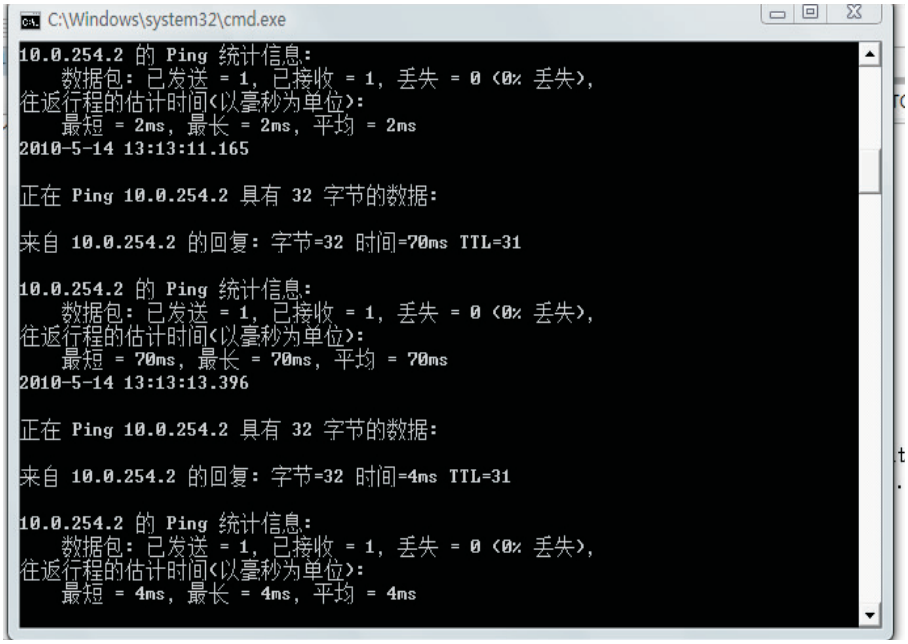

图 2　使用 DOS 命令获得主动测试 RTT 的工作环境

据变得简单一些,就需要对网络数据包进行过滤,只对符合条件的数据包进行抓取。在用 Wireshark 抓包过滤条件不强,可能在抓到的自己发送到 FTP 网站的包中,有些不是基于 ftp 协议,如可能是 http 等协议的包,在用程序进行读文件时,误把这些包当作实验所需的数据进行计算,这些原因都会对最后的结果产生影响。可以使用 Wireshark 的过滤器功能,过滤掉不需要的包,在通过 Wireshark 官方网站查询过滤语法,根据语法规则写出 Filter String:src host 192. 168. 1. 232 and des host 202. 133. 24. 25 and tcp。符合这个过滤语句的包就是被动测量 PRE 算法实验所需要的数据。写好过滤条件以后,就可以开始用 Wireshark 进行抓包。

PRE 算法初始值确定是使用前 3 个报文,这三个报文就是 TCP 建立连接时的三次握手阶段,理论根据是 SYN ACK 方法。SYN ACK(SA)方法是通过观测 TCP 握手阶段的 SYN 和 ACK 包的时间差来估计 TCP 流开始时端到端的静态 RTT;只能测量到 TCP 连接建立时 RTT,不能提供持续的数据[1]。因为这三个数据是非常重要的,在第一节介绍 PRE 算法时,RTT 的初始值就是依靠这三个包确定的,所以这是以后测量的基础数据。如果在初试化阶段,抓包并没有抓到这三个包或者前三个包不是 3 次握手阶段的包,则对以后的分析带来灾难性的后果,产生较大的误差,实验出现误差就表明抓包的步骤可能存在问题,准确的步骤应该先打开 Wireshark,然后才与 FTP 网站连接,只有这样才可以准确的抓到前三个包,才能对以后的计算不带来太大的误差。

在抓包的过程中,为了获得准确的数据,需要在这段时间里不断的打开网页和去其它网站下载,这样做的原因是保证网络有一定程度的拥塞,获得更多的轮次间隔[9]。TCP 工作原理详细叙述,拥塞控制就是防止过多的数据注入到网络中,这样就可以使网络中的路由器或链路不致过载。拥塞控制所要做的都有一个前提,就是网络能够承受现有的网络负荷。当网络拥塞时,发送方的发送窗口没有收到接收方的确认号,且此时发送窗口已满,可用窗口已减小到零,因此必须停止发送。这就是流量控制(Flow Control)原理,目的是让发送方的发送速率不要太快,要让接收方来得及接收。这时,就产生了轮次间隔。因此,为了获得更多的轮次间隔,便于以后的分析,使得实验更具有有效性,非常有必要使网络保持一定程度的拥塞[10]。

持续一段时间以后,结束实验,下面要做的就是保存数据。使用 Wireshark 抓包的数据以. cap 为格式保存分析数据,并将数据放到可执行程序所在的文件夹里,为下一步使用程序分析文件做好数据准备。另外,通过 Ping 命令得到的主动测量的 RTT 数据,还需要进一步提取,因为在这个数据文件中有很多不需要的信息,对于实验和验证结果来说需要的只是发包的时刻和平均 RTT 时间。此时就需要对数据进行更进一步的提取,这里,使用 UltraEdit 软件打开该 txt 文件,使用 UltraEdit 软件的替换功能,用如下的一个正则表达式获得主动测量 RTTM 实验所需要的数据信息:“Ping . * ? ”。这样,该 txt 文件就只有实验验证所需要的数据:发送 ICMP 报文时刻,和该时刻 ICMP 报文所测量的 RTT 的平均值[11]。现在,就获得了 TCP 连接往返时延实验所需要的所有数据。在下一节就可以设计 Matlab 程序分析这些这些数据,得到坐标比较图,比较 PRE 算法和 RTTM 算法得到的 RTT 的误差范围。以验证文中第一节所提出的 PRE 算法的有效性。

3　处理数据

根据第一节所提供的 PRE 算法,使用 C 语言编写实现该算法,该程序的主要功能是读取在实验中抓到的包,并根据 PRE 算法的算法流程图(图 1),分析出轮次间隔。并将通过程序分析得到的每一轮次发包的

时刻和轮次间隔写入一个 txt 文件中。上一节,已经得到了主动测量得到的数据,并且保存在一个 txt 文件中。这时,实验所需要的数据就全包括在这两个 txt 文件中了,下面的工作就要通过 Matlab 软件画出主动测量和被动测量 PRE 算法获得的 RTT 结果比较的坐标图。设计 Matlab 程序读取上述所得到的两个 txt 文件,对比相同时间点上得到的主动测量和 PRE 测量得到的环回时间的值。坐标图的第一例输入 Ping 命令得到的 RTT,第二列输入使用 PRE 算法分析得到的 RTT。将对应列一命名为 RTTM,对应列二命名为 PRE,x 轴命名为 t/s,y 轴命名为环回时间 RTT/ms,图表命名为发送端 PRE 与 RTTM 处理结果比较坐标图。这样就获得了一个比较坐标图,图 3 就为经过上述分析、设计所得到的比较图。

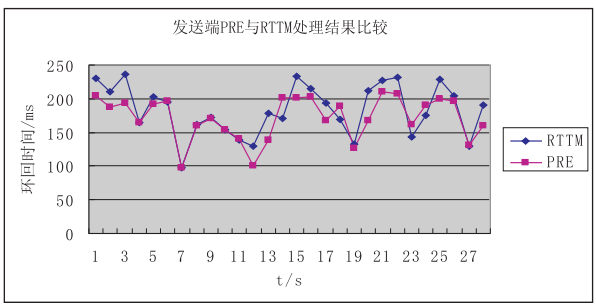

图 3 测量结果坐标比较图

4 结束语

在图 3 根据自己的实验数据画出的对比坐标图中,可以看出 PRE 算法和 RTTM 算法得到的 RTT 的值在某些时间点上还是有误差。实验网络环境是校园局域网,在这种网络环境下网络传输时延相对较大,测试测量得到的 PAI 和 PSI 就会有数值误差,如果再按照第一节讨论 PRE 算法时假定 PAI≈PSI 就会产生实验误差[12]。

从图 3 的总体上来看,在不考虑某些时间点的网络传输时延过大的问题,该算法可以较好地测量 RTT,从实验可以证明 PRE 算法是可以有效地测量往返时延。

参考文献:

[1] Jiang Hao,Dovrolis C. Passive estimation of TCP round trip times[J]. ACM Computer Communication Review,2002,32(3):75-88.

[2] Veal B,Li Kang,Lowenthal D. New methods for passive estimation of TCP round-trip times[J]. Passive and Active Measurement,2005,12(4):12-33.

[3] 张轶博,雷振明. 一种被动式 RTT 测量算法[J]. 北京邮电大学学报,2004(5):85-89.

[4] 刘秋让,倪红波. TCP 拥塞方法控制解决方法分析及评价[J]. 计算机工程,2001,27(6):65-66.

[5] 庞 胜,宋光龙. 一种在 IP 网络中间节点被动测量 RTT 的测量新方法[J]. 中国数据通信,2005,11(4):312-314.

[6] 杨 乐. TCP 往返时延被动测量方法综述[J]. 大众商务,2009(14):312-314.

[7] 黄鹂声,周明天,汪文勇. 骨干网络中的被动模式 RTT 测量方法[J]. 计算机应用研究,2009,26(8):3086-3089.

[8] 谢希仁. 计算机网络[M]. 第 5 版. 北京:电子工业出版社,2009:187-220.

[9] Jacobson V. TCP extensions for high performance[S]. RFC 1323,1992.

[10] Postel. Transmission control protocol introduction[S]. RFC 793,1981.

[11] Veal B. Framework for IP performance metrics[S]. RFC 2330,1998.

[12] 徐海东,常 傅,宋俊德. 基于业务管理的网络监控策略[J]. 北京邮电大学学报,2003,26(3):46-50.

(上接第 82 页)

[2] 王 辉. 主成分分析及支持向量机在人脸识别中的应用[J]. 计算机技术与发展,2006,16(8):24-27.

[3] 徐 勇,张 海,周森鑫,等. 基于统计学习理论的人脸识别方法研究[J]. 计算机技术与发展,2007,17(11):124-126.

[4] Lerdsud-wichai C,Addel-mpttaleb M. Algorithm for multiple faces tracking[C]//IEEE International Conference on Multimedia and Expo. [s. l.]:[s. n.],2003:777-780.

[5] Zhao W,Chellappa R,Rosenfeld A,et al. Face recognition:a literature survey[J]. ACM Computation Survey,2003,35(4):399-458.

[6] Neri A,Colonnese S,Russo G. Automatic moving object and background separation[J]. Signal Processing,1998,66(2):219-232.

[7] 刘晓宁,周明全,耿国华. 基于单张二维照片的三维姿态计算[J]. 计算机工程,2006,32(6):232-233.

[8] 李盛文,鲍苏苏. 基于 PCA+Adaboot 算法的人脸识别技术[J]. 计算机工程与应用,2010,46(4):170-173.

[9] 李春明. 视频图像中的运动人体检测和人脸识别[D]. 西安:西安电子科技大学,2005.

[10] 李晓莉,达飞鹏. 基于排除算法的快速三维人脸识别方法[J]. 自动化学报,2010,36(1):153-158.

[11] 孙晓丽,宋国乡,冯象初,等. 基于噪声-纹理检测算子的图像去噪方法[J]. 电子学报,2007,35(7):1372-1376.

[12] 于世华. 基于人脸识别的上机辅助身份验证系统[D]. 长春:东北师范大学,2008.