CERNET 主干信道中 TCP 数据流的一些传输特性

高亚东，丁伟，史冰
(东南大学计算机科学与工程学院,江苏 南京 210096)

摘要：对高速主干网中 IP 数据流的研究，尤其是对 TCP 数据流的研究，对于网络行为学研究具有重要意义。论文利用统计方法，从流、报文、数据量三个角度分析了在 CERNET 江苏省网边界采集的 TCP 报文数据，在此基础上探讨了所观测的高速主干信道中 TCP 数据流的传输特性。基于统计值的相关指标不仅能反映高速主干信道中的 TCP 平均负载效率和重传率，而且在一定程度上反映了网络安全状况。

关键字：TCP 数据流 负载效率 重传率

Some Transport Features of TCP Flows in CERNET Backbone

Gao Yadong, Ding Wei, Shi Bing
(School of Computer Science and Engineering, Southeast University, Nanjing 210096 China)

Abstract: Making researches on IP flows, especially TCP flows in high-speed backbone channels, has significant value for researching network behaviors. We analyzed TCP packets collected from Jiangsu province border of CERNET, using statistical methods. Then, based on the statistical results, we researched some transport features of TCP flows in the high-speed backbone channel which we observed. These statistical values provided a basis for studying some factors, such as TCP load efficiency, retransfer rate, etc. And they also reflected security status of the network to a certain extent.

Key words: TCP, flows, load efficiency, retransfer rate

1 引言

随着互联网迅速普及，网络带宽、用户数量、应用的种类和规模都在迅速增长，网络结构也日益复杂。在网络快速扩张的同时，各种各样的问题不断出现，影响了互联网的正常运行。为了了解网络的运行状况，发现问题的原因所在，从而更好地为用户提供服务，有必要对网络行为进行系统性的测量和研究。这其中，对高速主干网中 IP 数据流的研究，尤其是针对占据绝大部分高速主干带宽的 TCP 数据流的研究，对于网络行为学研究具有重要意义。但是目前全世界只有少数科研机构和互联网运营商具备 1Gbps 以上高速主干信道的数据采集能力，大部分研究人员只能使用仿真的方法研究主干网的行为。由于不是基于真实环境，仿真方法的研究成果可能存在偏差，因而不能正确地揭示网络运行的状况和客观规律。

CERNET 华东（北）地区网络中心承担了国家重大基础研究发展规划（973）课题“网络动态行为和传输控制理论”（2003CB314804）。在该课题的支持下，实现了一个基于大规模高速网络的数据在线采集存储系统 Watcher，采集点设置于 CERNET 江苏省网（JSERNET）边界信道。

本文受国家 973 计划课题“网络动态行为和传输控制理论”（2003CB314804）资助

高亚东（1982—），男，硕士研究生，研究方向：网络行为学。
丁伟（1962—），女，工学博士，教授，博士生导师，主要研究方向：网络体系结构，网络行为学。
史冰（1975—），男，硕士研究生，研究方向：网络行为学。

JSCERNET 只有一个边界出口，其边界路由器通过高速光纤在 CERNET 华东（北）地区网络中心内实现和 CERNET 国家主干的连接，是理想的观测对象。利用 Watcher 系统，采用光纤分光方式，采集了 2005 年 11 月 10 日全天 24 小时流经观测点的数据，截取每一个报文的头部信息，打上时间戳后存储在物理介质中，供研究使用。由于数据量庞大，在本文的研究中，选取了 2005 年 11 月 10 日 19 时至 24 时共 5 小时的数据作为实验数据。下面首先给出四个定义，以此开始对 CERNET 主干信道中 TCP 数据流传输特性的讨论。

2 四个定义

针对 TCP 类型的报文，给出下列四个定义：

定义一 全部数据：某个时段特定信道上传输的全部 TCP 报文的数据总量（含 IP、TCP 报文头部）。

全部报文：某个时段特定信道上传输的全部 TCP 报文。

定义二 协议数据：某个时段特定信道上传输的负载有用户数据的 TCP 报文的数据总量（含 IP、TCP 报文头部）。

协议报文：某个时段特定信道上传输的负载有用户数据的 TCP 报文。

定义三 负载数据：某个时段特定信道上传输的 TCP 报文中用户数据的总量。

负载报文：某个时段特定信道上传输的负载有用户数据的 TCP 报文。

定义四 有效数据：负载数据中去除重传数据，即无重复的用户数据总量。

有效报文：负载报文中去除重传报文，即负载有无重复用户数据的 TCP 报文。

注：协议报文和负载报文的定义相同；下文的讨论基于这四个定义。

3 定义的阐释

TCP 报文中包含大量 SYN、SYN/ACK、ACK 和 FIN 报文，它们是 TCP 协议实现传输控制的手段。这些报文只有 IP 和 TCP 头部，不负载用户数据。另外还有一些报文，例如扫描报文和 DDOS 攻击报文，它们也不负载用户数据。虽然不负载用户数据的报文不仅仅包括攻击报文，但网络正常运行时，其所占比例应在一个相对稳定的合理的范围内。如果其所占比例出现突变，可能显示网络中正发生某个安全事件。而这部分报文的数量正是全部报文和协议报文的数量差值，因此，协议报文与全部报文的数量比值在一定程度上反映了网络的安全状况。

协议报文和负载报文的概念相同，所以二者的数量必然相同。协议数据和负载数据之间的差别，在于协议数据有传输层和网络层的报头，而负载数据则是纯用户数据。对于每一个 TCP 报文，报文的长度和 IP 报头的长度存储在 IP 报头的特定字段，而 TCP 报头的长度存储在 TCP 报头的特定字段。通过统计这些长度值，可以计算负载数据和协议数据的比值。该比值能反映承载用户数据的 TCP 报文的平均负载效率。

有效数据和负载数据的比值能够反映数据重传情况，客观上也反映了网络的性能。去除重传报文的方法是将数据组成流，在组流过程中对具有相同序列号的报文只计数一次。这其中，采用文献[1]中给出的数据流定义：数据流是符合特定的流规范（specification）和超时（timeout）约束的一系列数据包的集合。对于相同的数据包序列，采用不同的流规范和超时约束可以得到不同的流集合。在网络测量中，广泛使用（源地址，宿地址，源端口，宿端口，协议类型）五元组作为流规范区分不同的流，而超时判定，则经常以 64 秒作为阈值。在下文的实验分析中也以五元组和 64 秒超时阈值进行数据流的区分。对于实验中使用的组流算法，我们已经将它所基于的模型发表在参考文献[4]中，本文不再详细叙述。

4 实验分析

实验中针对四个定义中涉及的数据总量和相应的报文数量进行统计，共 8 个统计数据。另外，针对四种数据类型，分别统计它们在每个时间粒度内对应新产生的流数。因此，共有 12 个统计数据。需要注意的是，对于有效数据对应的流，将重传报文去除后流的报文在时间轴上分布变得更加稀疏，有

可能人为地将一部分流提前判断为超时，此时一个流就可能被截断为数个流，这和我们实验的初衷相悖。因此，在组流过程中，我们一旦判断某个报文是某个流的重传报文，仍需要将其时间戳纳入超时判断，更新它的五元组所对应的流信息结构中维护的最近报文时间戳，而不是简单地将报文丢弃。实验中，我们以 5 秒作为统计粒度，对应 5 个小时的实验数据，共获取了 3600 组统计值。实验程序刚运行时，在内存中新建大量的流信息，有关流数目的统计值偏大。实验程序运行一段时间后，各种统计值将进入相对稳定的状态。我们选取具有代表性的统计结果数据，计算相关比值，绘制成图表，分析其特点（注：图表的时间轴以 2005 年 11 月 10 日 19:00:00 为原点）。

1）从流的数量看：

统计结果显示，协议数据、负载数据和有效数据对应的流数目在每个时间粒度内都相同。这个现象和定义一致，因为协议数据和负载数据的区别仅仅在于是否包括 IP 报头和 TCP 报头部分，负载数据和有效数据的区别也仅仅在于是否包括重传的数据，而判定一个流取决于五元组和超时，是否计算报文的头部长度和重传报文，并不会使标识流的五元组发生变化，加之前文已经说明将重传报文的时间戳纳入超时判定，不会人为将流截断，因此，三者的值必然表现为完全一致。

计算每个时间粒度内协议流和全部流数量的比值，绘制成图 1。图 1 显示，全部流中，只有 3%左右的流是负载有用户数据的协议流。这个结果说明，当前作为互联网最主要传输协议的 TCP，它的数据流中，绝大部分不负载任何用户数据。这其中有一部分流是协议流的反向 ACK 流，它们对于协议流的传输控制是不可缺少的。根据 TCP 协议的传输确认机制可以判定，协议流对应的 ACK 流数目和协议流数目相当，也应该占全部流 3%左右。由此，我们从统计结果可以得出结论，我们所观测的信道中传输的 TCP 流总数的 90%以上和用户数据的传输无关。

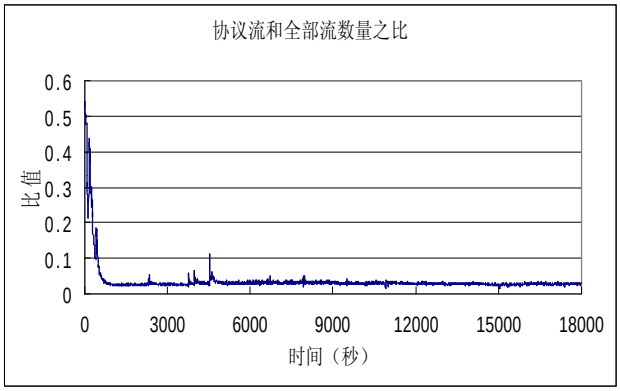

图 1 协议流与全部流数量比值分布图

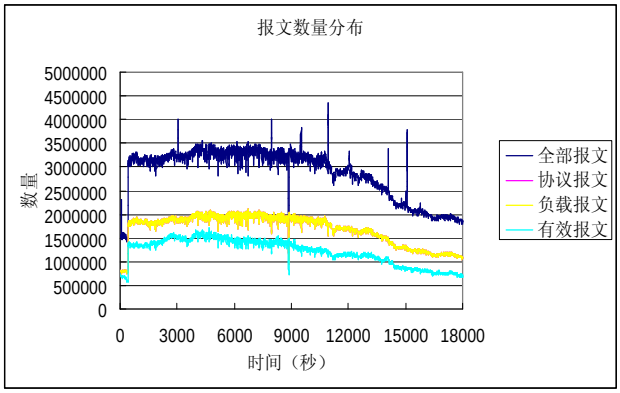

图 2 报文数量分布图

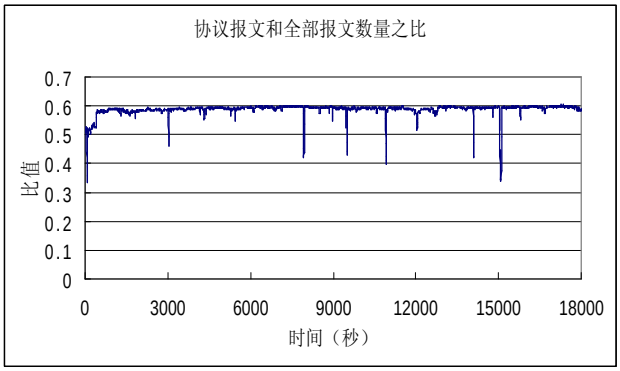

图 3 协议报文和全部报文数量比值

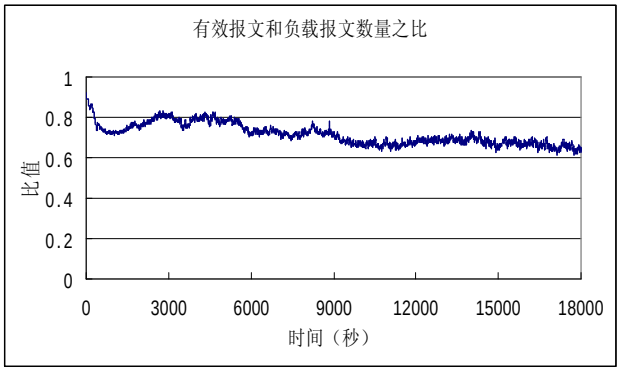

图 4 有效报文和负载报文数量比值分布图

2) 从报文数量看：

协议报文和负载报文的定义相同，因此二者的数量在统计数据上也表现为严格相等，图 2 中的中间一条线是两条分布曲线的重合。从图 2 可以看出，四种类型报文数量分布曲线的变化趋势基本一致。图 2 中全部报文数量的分布曲线比另三条曲线毛刺更多，这说明网络中的突发情况更多地表现为不负载用户数据的报文数量发生突变，这很可能由扫描攻击、DOS 攻击，以及蠕虫爆发等安全事件引起。

图 3 显示，协议报文占全部报文的比例接近 60%。结合图 1，可以发现，所观测的高速主干信道中，占总数 3%左右的 TCP 流中包含的报文数占TCP 总报文数接近60%，即大部分 TCP 报文只属于很少的 TCP 流。对应图2中全部报文数量分布曲线出现毛刺的地方，图 3 中出现了比值向下的突变，由此我们认为协议报文和全部报文数量比值能够反映网络中的突发情况，在一定程度上反映了网络可能发生安全事件。

图 4 显示，有效报文和负载报文的比值在 0.6 至 0.8 附近波动，可以推算报文重传率在 20%左右至 40%之间波动，由此我们可以发现，TCP 协议在重传报文数量方面要以较高的重传代价来保证用户数据可靠传输。

3) 从数据量看

将全部数据的统计值绘制成分布曲线，与图 3 中全部报文的数量分布曲线相比，总体变化趋势相同，但很少有大幅度的变化，没有突发的毛刺。这说明，网络中突发事件对报文数量波动影响较大，而对数据总量波动影响相对较小。限于篇幅，本文没有列出数据量分布图。

图 5 显示，一般情况下，协议数据占全部数据 97%以上，而前文的统计结果显示，它对应的协议流只占全部流数量的 3%左右，对应的协议报文数占全部报文数则接近 60%。图 1、3、5 共同表明，很少的 TCP 流包含了大部分 TCP 报文，承载了绝大部分数据量。

图 6 显示，一般情况下负载数据和协议数据之比在 96%到 96.5%之间。据我们对文中使用的 5 小时实验数据的统计显示，TCP 报文的网络层和传输层报文头长度之和平均值是 43.57 字节，由此可推算协议报文的平均长度在 1089 到 1245 之间，这和我们对协议报文长度的统计结果相符。该平均长度和 MTU 的差值说明协议报文的负载区域并没有得到充分利用，当然这是和网络中各种各样的具体应用相关联的。

将有效数据和负载数据的比值绘制成分布曲线后，发现与图 4 所显示的有效报文和负载报文的数量比值分布曲线具有相似的变化特征，用户数据的重传率也是在 20%左右至 40%之间波动，这说明 TCP 协议在实际使用的带宽方面要以较高的重传代价来保证用户数据可靠传输。限于篇幅，加之和图 4 类似，这里没有列出有效数据和负载数据的比值分布图。

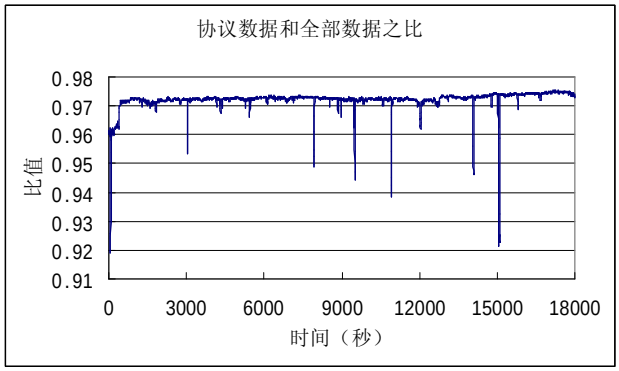

图 5 协议数据和全部数据比值分布图

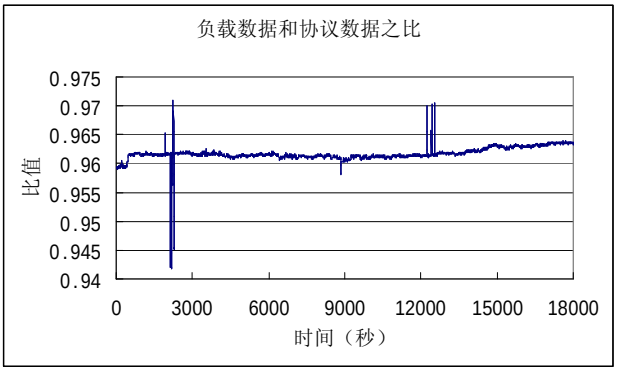

图 6 负载数据和协议数据比值分布图

5 总结

文章通过实验分析表明，在所观测的高速主干信道中，当前互联网中最主要的传输协议TCP的数据流中只有3%左右实际承载了用户数据，与之相当，有3%左右的TCP数据流用于用户数据的传输控制，其余90%以上的TCP数据流和用户数据的传输无关。这承载了用户数据的占总数3%的TCP数据流，包含了大部分TCP报文和绝大部分数据量。此外，我们发现在我们所观测的信道中，TCP协议在保证用户数据可靠传输的同时，需要重传大量报文和用户数据，在占用带宽等方面需要付出较高的重传代价。文章中绘制成分布曲线的有关比值，不仅反映了TCP负载效率和重传率，能够为相关研究提供直观依据，而且在一定程度上反映了网络安全状况，有助于对安全事件的预警发现。

参考文献：

[1] Ryu B, Cheney D, Braun H.W. Internet Flow Characterization: Adaptive Timeout Strategy and Statistical Modeling[J]. In Workshop on Passive and Active Measurement(PAM), Apr, 2001.

[2] Jun Li, Minho Sung, Jun Xu, Large-scale IP traceback in high-speed Internet: practical techniques and theoretical foundation, Security and Privacy, 2004. Proceedings.2004 IEEE Symposium on 9-12 May 2004, P115-129.

[3] 张宏莉，方滨兴，胡铭曾等，Internet测量和分析综述，2003 Journal of Software, Vol.14, No.1, P110-116.

[4] 高亚东，周明中，丁伟，高速网络中的数据流信息提取模型，计算机时代，2004年12月，Vol.22,No.12,P31-33。

附：

作者联系方式：

姓名：高亚东

地址：东南大学计算机学院 CERNET 华东（北）地区网络中心

邮编：210096

电话：025－83792000－304　　13601587453

E-mail：ydgao@njnet.edu.cn