

基于实测流量的网际分组丢失率分析

米婷^{1,2}, 丁伟^{1,2}

(1. 东南大学 计算机科学与工程学院, 江苏 南京 211189; 2. 江苏省计算机网络技术重点实验室, 江苏 南京 211189)

摘要: 分组丢失率是评价网络性能的一个重要指标。提出了一种利用完整 TCP 流计算实测数据中分组丢失率的方法, 并基于实测数据对 CERNET 江苏省网与外部不同网络分区的分组丢失率进行分析。算法给出了基于实测数据的分组丢失率包括首次分组丢失率 p_n 、整体分组丢失率 p_{sl} 和重传率 p_{rp} 3 个测度。统计分析表明, 在绝大多数情况下, p_n 比 p_{sl} 小, 而 p_{rp} 比 p_n 和 p_{sl} 明显要高, 验证了分组丢失对网络性能的影响以及重传率明显比分组丢失率高。

关键词: 分组丢失率; 重传率; 实测流量; TCP 流; 分区实验

中图分类号: TP393

文献标识码: A

文章编号: 1000-436X(2013)Z2-0023-04

Analysis of packet loss rate based on the measured traffic

MI Ting^{1,2}, DING Wei^{1,2}

(1. School of Computer Science and Engineering, Southeast University, Nanjing 211189, China;

2. The Key Laboratory of Computer Network Technology, Nanjing 211189, China)

Abstract: The packet loss rate is one of key indicators to evaluate the performance of network. A method for calculating packet loss rate in the measured data based on the intact TCP flows was presented. Based on the measured data, the packet loss rate of Jiangsu Province in CERNET network and different partitions outside was analysed. The algorithm gives the packet loss rate including first packet loss rate p_n , overall packet loss rate p_{sl} and retransmission rate p_{rp} . Through statistical analysis, p_n is smaller than p_{sl} in the vast majority of cases, and p_{rp} is significantly higher than p_n and p_{sl} , which proves the impact of packet loss on the performance of network, and retransmission rate is obviously higher than packet loss rate.

Key words: packet loss rate; retransmission rate; measured traffic; TCP flow; partition experiment

1 引言

根据 IPPM 中 RFC2680 的定义, 在 T 时刻源发送分组 P 到宿, 宿若没有接收到该分组即为分组丢失^[1]。端到端分组丢失率是基本的网络指标, 会影响到网络及应用程序的性能^[2]。尽管光纤链路的采用和光纤技术的发展已使分组丢失大大降低, 但网络流量的增长经常超过网络处理能力的提高, 拥塞和缓存溢出仍是 IP 传输需要考虑的因素。

TCP 是互联网中广泛使用的传输层协议, 它提供了一种可靠的面向连接的字节流服务。很多流行的应用程序都使用 TCP, 如 Telnet、Rlogin、FTP

和 SMTP 等, 是网络中使用最多的传输层协议。TCP 报文传输的一个重要特征是传输过程中的序号, 正常到达的报文序列号应该是非递减的, 当到达的报文序号没有按照非递减方式出现时, 网络中即出现乱序。一定程度的乱序会导致接收端认为报文丢失, 引起拥塞窗口的调整和快速重传, 从而可能会导致网络或应用程序的性能下降。当发送端接收到多个重复 ACK 或者出现定时器超时, 发送端会认为报文在传输中丢失, 从而重传该报文, 同样会引起拥塞窗口的调整进而影响网络性能。

目前, 应用程序可以通过主动或被动测量的方法来研究分组丢失率。Barford 和 Sommers 比较了

收稿日期: 2013-09-04

基金项目: 国家重点基础研究发展计划(“973 计划”)基金资助项目(2009CB320505)

Foundation Item: The National Basic Research Program of China (973 Program) (2009CB320505)

主动与被动测量方法,发现目前主动测量分组丢失率的方法存在标准抽样技术固有的高方差性以及终端主机接口易分组丢失的问题^[3],还提出了一种克服标准泊松探测工具测量分组丢失率的低精度缺陷算法^[4]; Benko 和 Veres 提出了一种基于端到端 TCP 分组丢失率估算算法的被动测量方法,通过观测 TCP 序号来定位分组丢失及推算分组丢失率^[5]; Allman 等人通过观测 TCP 发送端的重传报文来推算分组丢失率^[6]。这些研究大多是根据定位分组丢失位置,利用流量模型来估算分组丢失率,难以进行验证。而本文将从实测数据的角度对特定网络的分组丢失情况进行分析,即基于特定网络的实测数据,用统计的方法获取其实际的分组丢失率。

为达到这个目标,需要解决 3 个问题,首先是如何确认实测数据中的分组丢失现象,其次是如何计算分组丢失率,最后是如何合理地将这个计算公式运用到实测数据中。

本文通过流量中完整的 TCP 连接来确定分组丢失,通过首次分组丢失率 p_{fl} 和整体分组丢失率 p_{sl} 作为分组丢失率的计算公式,同时给出了一个重传率 p_{rp} 的公式作为比对。在计算实测数据时,对网络对端(入方向的源地址和出方向的宿地址)进行了分区,并分别计算了所关注的网络到对端的总体和分区分组丢失情况。

2 完整 TCP 流信息的获取

2.1 数据来源与数据选取

本文研究采用被动测量方法,数据来源为 CERNET 江苏省网到国家主干链路,分析数据来源于 <http://iptas.edu.cn>。实验具体选取了 2011 年 3 月、2012 年 3 月以及 2013 年 2 月的一个工作日(即总计 3 个工作日)的下午 14:00~15:00 采集的 trace 进行实验分析。

2.2 完整 TCP 流的获取

实验中,以 16 s 作为超时,对采集的 trace 具体组流散列函数如下:

$$\begin{aligned} a &= sip \wedge dip; & b[0] \wedge &= a[0] \wedge a[2]; \\ b[1] \wedge &= a[1] \wedge a[3]; & b \wedge &= sport \wedge dport; \\ hash &= b \wedge (a \gg 13) \wedge ((a \gg 3) \& 0x7FFF) \end{aligned}$$

其中, sip 为源地址, dip 为宿地址, sport 为源端口, dport 为宿端口。

对分组好的 TCP 流进行报文信息提取,过滤出

以 FIN 或 RST 为结束的 TCP 流作为完整的 TCP 流。为防止 TCP 流的报文数过少而对统计的意义不大,实验中只取报文数大于 4 的完整 TCP 流进行实验统计计算。

3 分组丢失率和重传率测度计算公式

在网络报文传输中,正常顺序到达的网络报文 TCP 序列号应该是非递减的,当到达的序号没有按照非递减方式出现时,网络中即出现乱序。为防止报文乱序接收的现象,在实验中首先对过滤后的完整 TCP 流的报文信息按序列号进行排序。基于 TCP 的超时与重传机制^[7],通过在发送时设置一个定时器,如果当定时器溢出时还没有收到确认,发送端就会重传该数据。所以,对于排序后的报文信息,当检测到相邻报文序列号和 ACK 均相等时,记为一次分组丢失,然后进行统计计算。

3.1 首次分组丢失率 $p_{fl}(p_{first_loss})$

根据文献[8],对于 TCP 的 NewReno 算法,数据传输过程中有多个分组丢失后,NewReno 在快速恢复阶段每隔一个往返时延(RTT)重传一个丢失的分组,直到拥塞窗口的所有丢失分组都被重传。即每次重传或者超时,进入快速恢复状态,一个新的周期就又开始,当所有报文都被应答后才退出此状态。

本文定义首次分组丢失率 $p_{fl}(p_{first_loss})$ 如下

$$p_{fl} = \lim_{n \rightarrow \infty} \frac{n}{\sum_{k=1}^n (A_k + 1)} = \frac{1}{E(A_{\infty}) + 1}$$

其中, A_k 为第 k 周期首次分组丢失时的传输分组数。

本文中每条完整的 TCP 流只取第一次分组丢失之前的数据分组进行统计计算,即对于每条流,此测度分子只记第一次分组丢失,分母是第一次分组丢失之前的数据分组与此次分组丢失的总和。

3.2 整体分组丢失率 $p_{sl}(p_{sum_loss})$

对于同一条流中连续多次报文序列号相等的情况,即连续重传,分子只记丢掉的报文,分母包括所有重传的报文。

根据文献[9]中定义,即

$$p_{sl} = \frac{\text{lost data packets}}{\text{all data packets}}$$

相比 3.1 节,这里是将每条流的所有报文用于计算。对于每次分组丢失,如果 TCP 源不是采取

selective-ack 机制，就可能会重传多个报文，如果均算作分组丢失，就会导致分组丢失率的计算偏大。因此，这里将连续重传的报文只记为一次分组丢失来进行统计计算。

3.3 重传率 p_{rp} ($p_{resent_packets}$)

正如在 3.2 节中所说，为了与分组丢失率进行对比，此处提出重传率的概念。对于每条排序后的 TCP 流，只要检测到相邻的 2 个报文的序列号与 ACK 相等，则记为一次重传。

定义公式如下

$$p_{rp} = \frac{\text{no. of resent packets}}{\text{total no. of sent packets}}$$

本文对所有重传的数据分组进行统计。基于不是 selective-ack 机制的 TCP 源，即当发生分组丢失后，TCP 源会将分组丢失后的所有已经发送过的报文均进行重传，显然，重传的报文数会比分组丢失数大。因而，理论上统计结果会较大。

4 分组丢失率和重传率分区统计

为了更具体清晰地比较 3 个测度之间的关系来体现分组丢失对网络性能的影响，对于这 3 个时间段的 TCP 流，作者根据其方向不同，对其源（宿）地址进行具体的 IP 归属分区。

4.1 分区方案

对 2.2 节中获取的 3 个时间段的单向 TCP 流，基于 trace 在 CERNET 主干江苏省网边界采集，将宿地址是江苏省内的 TCP 流方向记为入方向(in)，

源地址是江苏省内的 TCP 流方向记为出方向(out)。对入方向的源地址和出方向的宿地址的 IP 归属进行分区。国内主要分为 cernet、电信、联通和移动，国外主要分为亚洲、欧洲、北美洲和其他。然后，分别对各分区进行 3 个测度的统计计算，获取各分区的分组丢失率。

4.2 数据分析

对于每个时间段的入方向和出方向，实验中分别选取了 50 000 条报文数大于 4 的完整 TCP 流进行统计计算。

表 1 是 3 个时间段 50 000 条流对于测度 p_n 的相关数据信息的总体统计，即只统计 50 000 条流首次分组丢失之前所包含的报文总数和字节总数。表 2 是 3 个时间段 50 000 条流对于测度 p_{sl} 和 p_{rp} 的相关数据信息的总体统计，即统计每条流所有的数据。

表 3~表 5 分别是 3 个时间段总体与各分区分组丢失率统计的实验结果。

根据表 3，入方向国内移动、国外北美洲分区的分组丢失率较大；出方向国内联通、国外亚洲分区的分组丢失率较大。表 4 入方向国内移动、国外亚洲分区的分组丢失率较大；出方向国内电信和联通、国外亚洲和其他分区的分组丢失率较大。表 5 入方向国内电信、国外亚洲分区的分组丢失率较大；出方向国内联通、国外亚洲分区的分组丢失率较大。根据 3 个表对比，入方向 2013 年总体和各分区 p_n 相比前两年较小，而总体 p_{sl} 相差不大；出方向 2012 年总体 p_n 和 p_{sl} 均较小。

表 1 p_n 相关数据信息总体统计

总报文数和字节数	20130222_1400_1500		20120301_1400_1500		20110311_1400_1500	
	入方向	出方向	入方向	出方向	入方向	出方向
流数	50 000	50 000	50 000	50 000	50 000	50 000
报文数	784 541	439 808	683 612	417 820	584 190	453 951
字节数	813.23 GB	306.09 GB	716.81 GB	294.01 GB	449.42 GB	97.66 GB

表 2 p_{sl} 和 p_{rp} 相关数据信息总体统计

总报文数和字节数	20130222_1400_1500		20120301_1400_1500		20110311_1400_1500	
	入方向	出方向	入方向	出方向	入方向	出方向
流数	50 000	50 000	50 000	50 000	50 000	50 000
报文数	1083 129	580 809	1 040 547	536 905	814 066	670 831
字节数	1 190.22 GB	452.94 GB	1 166.52 GB	434.58 GB	688.60 GB	124.31 GB

表 3 20110311_1400_1500 分组丢失率统计/%

分区 名字	入方向			出方向		
	p_n	p_{sl}	p_{rp}	p_n	p_{sl}	p_{rp}
总体	1.358 98	3.084 15	3.946 12	1.679 26	2.799 21	6.350 48
cernet	1.336 68	3.289 01	4.000 73	1.615 99	2.832 67	6.884 51
电信	1.121 65	2.173 98	3.869 42	1.277 83	2.546 61	4.204 34
联通	1.555 32	3.222 33	4.383 63	2.359 51	3.160 76	6.464 96
移动	2.169 37	3.992 53	10.08 64	1.121 30	1.961 64	2.833 48
亚洲	1.286 11	2.320 06	2.779 32	3.253 76	3.518 23	10.67 03
欧洲	1.604 51	2.349 87	2.573 67	1.289 24	1.700 15	5.203 50
北美	2.392 53	4.327 15	5.009 03	1.775 19	2.906 28	5.219 34
其他	1.167 32	1.540 96	2.838 61	1.838 76	2.549 25	3.708 00

表 4 20120301_1400_1500 分组丢失率统计/%

分区 名字	入方向			出方向		
	p_n	p_{sl}	p_{rp}	p_n	p_{sl}	p_{rp}
总体	1.72 496	3.46 597	3.97 810	1.39 749	2.13 278	2.62 989
cernet	1.89 741	3.87 018	4.44 717	1.12 011	1.83 034	2.14 584
电信	1.19 479	2.72 265	3.11 625	1.16 551	2.18 604	2.51 851
联通	1.33 891	2.51 966	2.92 216	1.85 904	2.55 389	3.27 369
移动	4.78 395	4.90 834	5.38 143	1.43 005	2.00 482	2.33 348
亚洲	2.27 416	2.57 842	4.07 292	2.25 675	2.17 063	4.61 602
欧洲	1.30 698	1.44 168	1.86 107	1.26 450	1.25 069	1.66 759
北美	1.48 380	2.65 094	3.11 755	0.87 549	0.95 848	1.19 094
其他	1.64 993	1.71 647	1.90 718	2.69 495	3.27 198	5.36 810

表 5 20130222_1400_1500 分组丢失率统计/%

分区 名字	入方向			出方向		
	p_n	p_{sl}	p_{rp}	p_n	p_{sl}	p_{rp}
总体	1.13 289	3.34 005	4.17 605	1.71 711	3.39 268	6.30 896
cernet	1.10 144	3.07 733	3.69 162	1.13 723	3.59 318	5.64 242
电信	1.18 077	5.11 173	6.86 263	1.56 720	3.06 407	5.25 777
联通	0.82 656	3.65 498	4.17 398	1.30 874	4.54 685	7.43 027
移动	0.95 179	1.10 466	1.47 912	1.14 820	5.92 798	11.6 186
亚洲	1.39 671	2.06 915	2.90 407	2.18 550	2.39 972	4.09 772
欧洲	0.79 764	1.17 896	1.69 058	1.01 117	1.52 417	2.31 529
北美	0.58 507	1.07 751	1.74 304	0.72 049	0.79 337	1.63 299
其他	0.52 781	0.55 888	0.79 840	0.25 547	0.29 603	0.33 156

从整体上可以看出，除了少数分区的 p_{sl} 比 p_n 小（见表 4 中亚洲和欧洲分区），其他分区和整体的 p_{sl} 均比 p_n 大，且 p_{rp} 明显高于 p_n 和 p_{sl} ，是三者中最大的，这充分说明了分组丢失对网络性能的影响，并且验证了 3.3 节中重传率比分组丢失率高的理论设想。

5 结束语

本文提出了一种利用完整 TCP 流计算实测数据中分组丢失率的方法。首先通过流量中完整的 TCP 连接来确定分组丢失，然后提出了 3 个测度计

算公式，进而将此公式运用到实测数据中，对 CERNET 江苏省网与外部不同网络分区的分组丢失率进行统计。通过比较 p_n 和 p_{sl} ，说明了分组丢失对网络性能的影响。另外，通过定义 p_{rp} ，将其与分组丢失率进行比较分析，体现出重传率明显比分组丢失率高，验证了理论设想。

参考文献:

- [1] ALMES G, KALIDINDI S, ZEKAUSKAS M. A One-Way Packet Loss Metric for IPPM[S]. RFC 2680, 1999.
- [2] WANG Y, HUANG C, LI J. Queen: estimating packet loss rate between arbitrary Internet hosts[A]. Passive and Active Measurements Conference[C]. Seoul, Korea, 2009. 57-66.
- [3] BARFORD P, SOMMERS J. Comparing probe and router-based packet-loss measurement[A]. IEEE Internet Computing[C]. 2004. 50-56.
- [4] SOMMERS J, BARFORD P, DUFFIELD N. Improving accuracy in end-to-end packet loss measurement[A]. ACM SIGCOMM Computer Communication Review[C]. 2005. 157-168.
- [5] BENKO P, VERES A. A passive method for estimating end-to-end TCP packet loss[A]. IEEE Globecom[C]. 2002. 2609-2613.
- [6] ALLMAN M, EDDY W, OSTERMANN S. Estimating loss rates with TCP[A]. ACM Performance Evaluation Review[C]. 2003. 12-24.
- [7] STEVENS W. TCP/IP 详解卷一: 协议[M]. 北京: 机械工业出版社, 2000.226-227.
- [8] STEVENS W. TCP/IP Illustrated Volume 1: the Protocols[M]. Beijing: China Machine Press, 2000.226-227.
- [9] KAJ I, OLSEN J. Throughput modeling and simulation for single connection TCP-tahoe[A]. Proceedings of the 17th International Teletraffic Congress[C]. Salvador, Bahia, Brazil, 2001.705-718.
- [10] WU H, GONG J. Packet loss estimation of TCP flows based on the delayed ACK mechanism[A]. The 12th Asia-Pacific Network Operations and Management Symposium[C]. Jeju, Korea, 2009. 540-543.

作者简介:



米婷 (1990-), 女, 湖北孝感人, 东南大学硕士生, 主要研究方向为网络测量、网络行为学等。



丁伟 (1962-), 女, 江苏南京人, 博士, 东南大学教授, 主要研究方向为计算机集成制造、通用搜索引擎、PKI 证书体系、网络环境下的远程教育和网络行为学等。