·博士论坛·

流测量算法综述

刘卫江[1,2,3] 龚 俭[2] 丁 伟[2] 程 光[2]

[1](东南大学计算机科学与技术学科博士后流动站,南京 210096)

[2](东南大学计算机科学与工程系,南京 210096)

[3](渤海大学信息科学与工程学院,辽宁锦州 121003)

E-mail:wjliu@njnet.edu.cn

摘 要 理解网络行为对于网络管理、规划和发展都有重要意义,而流测量是了解网络行为的基础。由于网络的高速与流数量的巨大,使得实时在线的流测量变得很困难。因此各种流测量技术、流测量算法成为研究热点。文章综述了目前利用抽样和哈希技术在流识别和流分布方面取得的成果,并分析了各种算法的优缺点。最后分析了抽样与哈希技术的长处与不足,提出了多种技术相结合的研究方向。

关键词 报文抽样 哈希 流测量 算法

文章编号 1002-8331-(2005)29-0001-03 **文献标识码** A **中图分类号** TP393

A Survey of Flow Measurement Algorithms

Liu Weijiang[1,2,3] Gong Jian[2] Ding Wei[2] Cheng Guang[2]

[1](Post Doctoral Station for Computer Science and Technology,Southeast University,Nanjing 210096)

[2](Department of Computer Science and Engineering,Southeast University,Nanjing 210096)

[3](School of Information Science and Engineering,Bohai University,Jinzhou,Liaoning 121003)

Abstract: Understanding network behavior is very important for network management,planning and development,but it is based upon flow measurement.Due to high speed of network and large number of flows,it is very difficult to measure flow real-timed and online.So a variety of flow measurement techniques and algorithms become a research hot-spot.In this paper,we survey the achievements obtained in the fields of identifying flow and flow distribution by using sampling and hash technique,and evaluate the advantages and disadvantages of some algorithms.Finally,we analyze the advantages and disadvantages of sampling and hash techniques and present the future research trend of various techniques combined.

Keywords: packet sampling,hash,flow measurement,algorithm

1 引言

互联网是上亿台计算机互联成的全球性网络,随着更多的网络服务的提供及用户的不断增加, 网络流量变得越来越大,网络行为也变得越来越复杂。虽然相关的组网与管理技术在不断地完善,但人们对它在局部和整体范围内所体现出的行为特征依然没有一个正确和完整的认识。掌握 Internet 的行为是网络规划、网络管理和网络安全、新网络协议和网络应用设计等诸多研究工作的重要前提。而网络测量是分析、掌握网络行为的基础,通过收集数据或报文踪迹,以定量地分析不同的网络应用在网络中活动规律的技术。

通过定量测量并分析网络,我们可以理解网络流量与网络特征,探讨网络行为和运行规律;通过测量建立起网络性能基线,有效地进行网络监测,合理地分配网络资源,迅速定位网络故障;了解网络端到端、整体甚至局部性能细节,为规划、设计网络提供科学依据。报文(packet)是网络中最小的传输单元,最初网络行为研究主要集中在数据报文层次上,但由于这些研究相对平等地分析每个报文,从而导致对报文间关系及其更高层次信息分析的缺失。针对流(flow)的网络行为研究在很多方面弥补了局限于报文层次研究的不足。对网络"流"的特性进行测量和分析,可以掌握网络某结点或链路的流量细节,如运输协议、应用协议、流量强度情况和用户行为特征等。而且采用流概念记录测量结果,通常可以大大节省存储空间。因此流测量成为网络测量中的一个热点研究方向, 特别是流分布的测量。而流测量的方法的主要问题在于缺乏可测量性。早在 1997 对 MCI 踪迹(traces)的测量显示了超过 250 000 个流存在[1]。而现在的(40Gb/s)高速网络 1min 的流量就存在着上百万个流。由

基金项目:国家 973 重点基础研究发展规划项目(编号:2003CB314803);国家自然科学基金(编号:90104031)资助

作者简介:刘卫江(1969-),男,博士后,副教授,主要研究方向:网络测量、网络行为学。龚俭(1957-),男,博士生导师,教授,主要研究方向为网络安全、网络行为学。丁伟(1962-),女,博士,博士生导师,教授,研究方向为网络管理、网络行为学。程光(1973-),男,博士,副教授,研究方向为网络行为学、网络测量。

于网络的高速与流数量的巨大,使得实时在线的流测量变得非常困难。因此各种流测量技术、流测量算法成为研究热点,并且取得一系列的成果。目前在流测量方法上主要是使用了抽样和哈希技术,研究的重点是流识别和流分布的算法。本文综述了利用抽样和哈希技术在流识别和流分布方面取得的主要成果。第二节介绍了流的基本概念,第三节介绍了利用抽样技术在测量流方面的成果。第四节介绍了利用哈希技术在测量流方面的成果。在第五节分析了这两种技术的优缺点并提出了未来工作的一个研究方向并结束全文。

2 流的概念

IP 网中的"流"概念可定义为对一个呼叫或连接的人为逻辑对应[2],流是流量的一部分,由起始时间和停止时间定界。与流相关的属性值(源/目的地址、分组计数、字节计数等)具有聚合性质,反映了在起始和停止范围发生的事件。流的起始时间对给定的流是固定的,其停止时间可能随该流的持续时间增加。当流在某时间间隔内无新分组到来时,则称该流已终止,否则称该流为活跃流。由于研究的背景不同,对于流采用了不同的定义。在文中我们采用文献[3]中的定义。

定义 1 流是指在某时间段内通过一个观测点的具有共同性质的报文集合。在文中 TCP 流是指具有相同的源 IP、宿 IP、源端口、宿端口的 TCP 报文集合,普通流是指具有相同的源 IP、宿 IP、源端口、宿端口的 TCP 或 UDP 报文集合(不考虑协议)。

定义 2 抽样流是指在上述定义的流中以概率 $p=\frac{1}{N}$ 进行概率抽样而得到的报文集合。

在本文中为叙述方便称定义 1 中的流为原始流。

定义 3 流长度是流中包含报文的数量。

3 抽样测量

3.1 报文抽样技术

早在 1993 年 K Claffy[4]就系统研究了基于时间和基于报文到达次序为抽样的激发机制,分析系统抽样、随机抽样分层的测量技术。在 1998 年,Cozzani[5]研究基于报文内容抽样的测量技术。进入 21 世纪以后,由于 2.5Gbps 和 10Gbps 高速主干网络的普遍使用和网络测量的广泛应用,抽样测量技术有了较快发展。2003 年 IETF 成立了报文抽样测量工作组(PSAMP)[6]专门用于研究报文抽样测量技术。美国 AT&T 实验室的 N.G. Duffield,2001 年发表的论文[7]提出了 Trajectory 基于报文内容的抽样技术,后来他又在论文[8]中研究关于 Trajectory 抽样应用。同时抽样测量技术也开始应用于网络产品中,如 Cisco 的 Netflow[9]和 NetranMet 测量器[10]。抽样会造成一些内在信息的损失,因此希望能够使用推断的方法来减少这种损失。假若平均每 N 个报文抽出一个,报文的总数可以能被估计为抽到的报文数的 N 倍,若抽样报文选择与报文大小无关,报文的总字节数也可以按同样的方式估计。但是原始流量的更详细的特征不是这么容易被估计的,包括一个流中的报文数量及此流中全部报文的字节总量等。直观地看,在路由器中使用报文抽样时,长流比短流更容易被抽到。因此简单地把流长度乘上 N 倍,不是对原始流长度的一个好的估计。

3.2 长流的识别

对于很多应用,了解长流可能就足够了[11,12]。2002 年 Cristian Estan 在文[13]提出了 sample and hold 算法用于发现长流有效地解决了如何在报文抽样情况下获取和维护流信息的问题。它的基本思想是对报文进行抽样,并且对每一个报文都进行处理。当一个流的一个报文被抽到之后,在一个内存的哈希表中建立了这个流长度的计数器,这个流后序的每一个报文都更新这个计数器,一直到测量结束。这实质上是通过抽样报文来抽流的方法,由于长流的报文数量大,因此长流被抽到的概率也相当大。这种方法可以精确地识别长流,用的内存也很小。它的缺点是对每一个报文都要进行处理和访问内存,因此要求内存的速度达到线速,给测量系统很大的压力。2004 年 Tatsuya Mori 在文[14]提出了一个新的识别长流的机制。这一机制的关健是由抽样流的报文数来确定这个原始流是否为长流的阀值。这一阀值是通过由 Bayes 定理计算在平衡错误肯定和错误否定的基础上得到的。而且各种先验分布对阀值的影响不大,这意味着在一个网络上得到的阀值已在另一个网上应用。这种识别机制的好处是非常简单,易于实现,而且不需每个报文都要进行处理,适用于高速网络。它的缺点是精度不高,错误的肯定率与错误的否定率的平衡不易实现。

3.3 流分布的估计

流分布信息在很多的网络测量和监控的应用中是有用的。首先,流分布信息可以允许服务提供者去推断它们网络的使用形式。第二,流分布信息可以帮助部分地检测到一个导致全局网络的从一种模式到另一种模式变化的事件的存在。进一步,流分布也可帮我们检测到各种类型的网络安全攻击,如 Ddos 和网络蠕虫。在 Ddos 攻击中,如果攻击者利用了伪造 IP 地址,我们将看到一个长度为 1 的流的大量增长。还有,被保留各种网的历史流分布信息也帮助我们研究它的随着时间的进化。对于由抽样数据获得原始流信息的研究是 Duffield 在文[15]首先开始的,提出对由抽样流数据推断出原始流数据的思想,特别是对平均流长度的推断,又在文[16]提出了两种推断方法:比例法(scaling method)和 EM 算法。比例法算法简单,由抽样数据很容易就可以计算出原始的未抽样的统计数据,但这种方法只能用于对 TCP 流的抽样估计,抽样时要对 SYN 报文进行统计,对于没有 SYN 报文的 UDP 流不适用。EM 算法求流长度分布的极大似估计的迭代求解方法。它不要求协议信息,尽管它也可以利用这一信息,因此它既适用于 TCP 流也适用于普通流。缺点是计算复杂性及被推断流的总数的缺少控制。对于它还有一个挑战是好的循环结束标准的选择,不当的结束标准可能造成尾部的震荡。

4 哈希测量

4.1 Bloom filter

Bloom filter 是用来表示一个集合的数据结构,它支持成员查询,随机存储。自从它在 1970 年[17]由 Bloom 提出后,就在

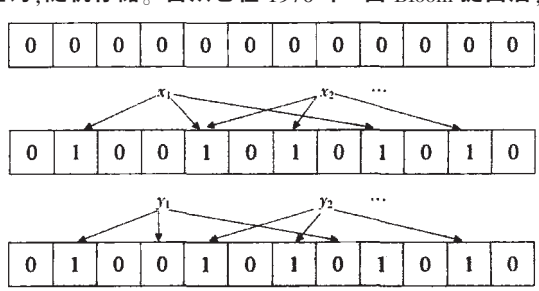

图 1 一个 Bloom filter 例子

数据库应用中被使用，近年在网络也受到了广泛关注[18]。

原理 假设所申请的内存大小为 m 比特位，该方法创建 k 个相互独立的哈希函数，能将数据集均匀映射到$[1\cdots m]$中去。对任何元素，利用哈希函数进行计算，得到 k 个$[1\cdots m]$之间的数，并将内存空间中这 k 个对应比特位都置为 1。

图 1 提供了一个例子。开始时数组的初始化为 0，集合中每个元素 x_i 被哈希 k 次，每次产生一个比特位置数，并把相应的比特位置 1。为了检查 y 是不是在这个集合中，哈希它 k 次并检查相应的比特位。y_1 不在这个集合中，由于在它的一个比特位是 0。而 y_2 或是在集合或是产生了一个错误的肯定。该算法的最大特点是，仅使用一小块远小于数据集数据范围的内存空间表示数据集，并且各个数据仍然能被区分开来。尽管存在错误的肯定，但当错误的概率足够小的时候，大量空间的节省使得这一缺点微不足道了。

4.2 流识别

2002 年 Cristian Estan 在文[13]提出了 Multistage filters 算法用于发现长流，并维护了长流的信息。这是一个与 Bloom filter 相似，但不同的技术。它的基本思想如图 2 所示。构建的模块是并行的多个哈希站。每一个哈希站是一个独立的哈希函数，当一个流标识为 F 的报文到来时，每个站把它哈希到本站表的相应位置，这一位置是一个计数器，初值为零，每哈希一次增加这个报文的大小。由于同一个流的所有报文都要被哈希到同一位置，因此如果流 F 发出的报文大小和超过了阀值 T，这时每一个哈希站表的相应的计数器值也一定超过了这个阀值 T，这时可以把这个长流识别出来，在内存中建立这个流的一标识项来记录它的信息。但也存在错误识别的问题，这个概率值可以计算出来。

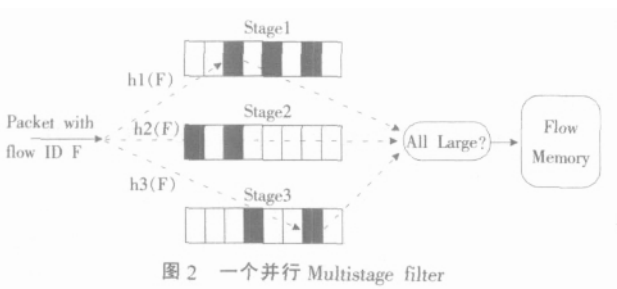

图 2 一个并行 Multistage filter

2004 年文[19]建立了空间 bloom filter 结构，使用最大似然技术估计和均值估计每个 IP 流的大小。它不需维护流的状态信息，需查询时把流信息输入进行查询计算。

4.3 流总数和流分布的估计

2003 年 Cristian Estan[20]使用哈希技术给出了一族 Bitmap 算法来统计活动 IP 流的数量，这些算法实现简单，有些可以用硬件实现，速度快可实现在线处理。Abhishek Kumar 等 2004 年[21]建立 IP 流哈希映射，使用贝叶斯统计估计流的大小分布。这一数据收集模式是以简单的计数数组为基础，结构简单，利用很少的 SRAM 能够在高速链路运行。但由于没有足够的时间和空间来解决哈希冲突，从计数数组得到的流分布是高度变形了的流分布数据。为了得到真实的分布，构建了 EM 算法来恢复这种变形。为提供足够空间的计数数组，算法依赖于对于流总数的估计。为克服这一困难，又构造了一个多分辨率数组，可以有效地解决估计精度问题。这一机制对于单报文流给出了一个估计公式 $n_1=y_1\times e^r$（n_1 是单报文流的数量，y_1 是数组单元值为 1 的单元个数，$r=\frac{n}{m}$，这里 n 是流总数估计，m 是计数器的个数），在精确地估计了流总数的条件下这是一个相当精确地估计。对于多分辨率数组来说，实际上是一个对流抽样的过程，理论上讲对于流分布的估计由流抽样要比报文抽样精确一些[22]。这个对于流分布的估计算法的缺点是计算复杂性太高，尤其是对于大的数据量（收集的时间长，流数量多的时候）计算是一个问题，而且对大流估计值偏高一些。

5 结论

由于高速网络流量给测量带来的负担，为了控制测量操作的资源消耗，选择抽样成为必然。它的缺点是：首先，由于 IP 流存在重尾特性，抽样报文仍需要大量的资源维护流信息；其次，抽样测量丢失了部分详细信息，在总体估计上精度难以保证。再者，由于存储大量的抽样信息，使得测量数据不能进行长期分析和处理。采用哈希技术的特点是存储空间小，处理速度快，可以对每一个报文都进行处理，可以实现在线处理，且流量近似估计精度高。缺点是哈希方法主要是保留 IP 流大小信息，但是丢失 IP 流的五元组地址信息；传统的 bloom filter 数据结构能够记录一个流是否已经存在，这种结构只能标记一个流的存在，我们不能知道这条流的具体大小。而且由于对每一个报文都进行处理，要求内存的速度也要快，用 SRAM。综上所述，流测量技术是近年来的国际上网络研究热点之一，但还只是限定在某些单元技术上，如：抽样技术和哈希技术。抽样技术和哈希技术在测量研究中各有优缺点，需要将不同近似方法相结合使用。因此以基于 IP 流的哈希函数为核心，综合抽样、哈希等多项近似技术的互联网测量、分析研究将是流测量的一个研究方向。（收稿日期：2005 年 7 月）

参考文献

1.Thomsom k Miller，G J，Wilder R.Wide-area traffic patterns and characteristics[J].IEEE Network Magazine，1997；11(6)：10~23

2.N Brownlee，C Mills，G Ruth.Traffic Flow Measurement：Architecture[S].RFC 2722，1999-10

3.Requirements for IP Flow Information Export (IPFIX).http://www.ietf.org/rfc/rfc3917.txt

4.Claffy K，Polyzos G，Braun H.Application of Sampling Methodologies to Network Traffic Characterization[C].In：Proceedings of ACM SIGCOMM '93，1993

5.Cozzani I，Giordano S.A passive test and measurement system：traffic sampling for QoS evaluation[C].In：Global Telecommunications Conference，GLOBECOM 1998，The Bridge to Global Integration，IEEE，1998：1236~1241

6.Packet Sampling(psamp).http://www.ietf.org/html.charters/psamp-charter.html，2005-02-02

7.Duffield N G，Grossglauser M.Trajectory Sampling for Direct Traffic Observation[J].IEEE/ACM Trans on Networking，2001；9(3)：280~292

8.Duffield N G，Grossglauser M.Trajectory Sampling with Unreliable Reporting[C].In：IEEE Infocom 2004，Hongkong，2004-03

9.Sampled Cisco.http://www.cisco.com/en/US/products/sw/iosswrel/ps1829/products_feature_guide09186a 0080081201.html，2002-12

10.NeTraMet Version 4.4 Now Available.http://www2.auckland.ac.nz/net/Accounting/ntm.Release.note.html，2002-12

（下转 7 页）

变化对预测结果的影响，本文实验结果证明使用 31 个隐层节点时效果最佳。经过多次实验，调整隐层结点数与权重都不能再提高预测精度，由此证明需要通过其它方法来改善网络。出现这种情况的原因在于，输入信息不足以利用神经网络的能力来找到更好的相关性。表 1 为径向基网的预测结果，其中只列出了隐层结点数为 10、25、31 和 37 个的情况（这几种情况较有代表性）。隐层节点数从 10 增加至 31 的过程中，预测准确度基本上是逐渐增加的。

表 1　简单径向基网的预测结果

隐层节点数	精度
10	61.1%
25	62.7%
31	63.6%
37	63.1%

第二步将径向基函数网级连。级连以后，对窗口大小也需要通过实验来确定，经试验认为第一层子网的窗口宽度 w=15 时效果最佳。将第一层的输出作为输入送入第二层，经实验第二层父网的窗口宽度选取 w=11。预测准确率与窗口大小有关，窗口过小的时候，对测试集预测的表现受到影响。表 2 为使用级连径向基网络的预测结果。

表 2　级连径向基网络的预测结果

精度	C_α	C_β	C_{coil}
68.8%	0.57	0.45	0.46

同时本项研究还进行了序列比对的尝试，对每一个位置根据多序列比对的结果计算相应氨基酸出现的频率，得到特征序列作为第一层的输入。然后，同第二步一样，将第一层的输出作为输入送入第二层。使用进化信息使预测非同源蛋白质二级结构的精度上升了一个多百分点，获得这种成功正是由于从输入模式中抽取了重要的信息。

5　结论

本文在分析了神经网络蛋白质二级结构预测方法的基础上，提出了基于径向基函数网络的预测方法。同时研究了蛋白质二级结构预测算法研究中的数据选取、网络结构和参数确定对网络性能的影响，实验预测准确率平均可以达到 69%左右，研究结果表明基于径向基函数网络预测的可行性和有效性。本文研究的蛋白质二级结构预测方法中的隐层神经元数目选取、窗口宽度设计，网络参数选择均需要进一步的研究工作。

（收稿日期：2005 年 7 月）

参考文献

1.King R D，Sternberg M J E.Identification and application of the concepts important for accurate and reliable protein secondary structure prediction[J].Prot Sci，1996；(5)：2298~2231

2.杨国慧，周春光，胡成全等.基于改进贝叶斯网络模型的蛋白质二级结构预测算法[J].自然科学进展，2003；13(6)：667~670

3.Rost B.Rising accuracy of protein secondary structure prediction.In：Chasman D ed.Protein structure determination，analysis and modeling for drug discovery，Dekker，New York，2002：207~249

4.Rost B，Sander C.Improved prediction of protein secondary structure by use of sequence profiles and neural networks[J].Proc Natl Acad Sci USA，1993；90：7558~7562

5.Altschul S F，Madden T L，Schaffer，A A Zhang et al.Gapped BLAST and PSI-BLAST：A new generation of protein database search programs[J].Nucl Acids Res，1997；(25)：3389~3402

6.Jones D T.Protein secondary structure prediction based on position-specific scoring matrices[J].J Mol Biol，1999；292(2)：195~202

7.McGuffin L J，Bryson K，Jones D T.The PSIPRED protein structure prediction server[J].Bioinformatics，2000；(16)：404~405

8.Rost B，Sander C.Combining evolutionary information and neural networks to predict protein secondary structure[J].Proteins，1994；(19)：55~72

9.Przybylski，B Rost.Alignments grow，secondary structure prediction improves[J].Proteins，2002；46：197~205

10.L S Wang，Y Xu.SEGID：Identifying Interesting Segments in (Multiple)Sequence Alignments[J].Bioinformatics，2003；(19)：297~298

11.Chandonia JM，Karplus M.The importance of larger data sets for protein secondary structure prediction with neural networks[J].Protein Science，1996；(5)：768~774

（上接 3 页）

11.Feldmann A，GreenbergA，Lund C Reingold et al.Driving traffic demands for operational IP networks：Methodology and experience[J].IEEE/ACM Transaction Networking，2001-06：265~279

12.Duffield N G，Lund C，Thorup M.Charging from sampled network usage[C].In：ACM SIGCOMM Internet Measurement Workshop 2001，San Francisco，CA，2001-11

13.Estan C，Varghese G.New Directions in Traffic Measurement and Accounting[C].In：SIGCOMM2002

14.Tatsuya Mori，Masato Uchida Ryoichi Kawahara.Identifying Elephant Flows Through Periodically Sampled Packets[C].In：IMC 2004

15.Duffield N G，Lund C，Thorup M.Properties and Prediction of Flow Statistics from Sampled Packet Streams[C].In：ACM SIGCOMM Internet Measurement Workshop 2002，Marseille，France，2002-11

16.Duffield N G，Lund C，Thorup M.Estimating Flow Distributions from Sampled Flow Statistics[C].In：ACM SIGCOMM，Karlsruhe，Germany，2003-08：325~336

17.B Bloom.Space/time tradeoffs in in hash coding with allowable errors[J].CACM，1970；13(7)：422~426

18.Andrei Broder，Michael Mitzenmacher.Network Applications of Bloom Filters：A Survey[J].Internet Math，2003；(4)：485~509

19.Abhishek Kumar，Jun Xu，Li Li et al.Space Code Bloom Filter for Efficient Traffic Flow Measurement[C].In：Proceedings of ACM/USENIX Internet Measurement Conference，Miami，FL，2003-10

20.Estan C，Varghese G.Bitmap algorithms for counting active flows on high speed links[C].In：Proc ACM SIGCOMM Internet Measurement Conference，2003-10

21.Abhishek Kumar，Minho Sung，Jun (Jim)Xu et al.Data streaming algorithms for efficient and accurate estimation of flow size distribution[J].ACM SIGMETRICS，2004

22.Nicolas Hohn，Darryl Veitch.Inverting Sampled Traffic[C].In:Internet Measurement Conference 2003