大规模高速网络流量测量研究[1]

程 光　龚 俭
（东南大学计算机系　南京 210096）
(email: gcheng, jgong@njnet.edu.cn)

摘要：随着互联网的发展，理解网络行为对于网络管理、规划和发展都有重要意义，网络流量测量是研究网络行为的基础。根据网络测量流量的不同，测量方法分为主动测量和被动测量，每种测量方法有其应用背景和优缺点。为了从不同角度研究网络行为需要定义不同的测度，IETF 的 IPPM 工作组现已经定义了一整套用于流量行为测量的测度。根据测量的环境和应用背景的区别，国外不同的研究机构建立了不同的测量体系结构和测量工具。同时，由于网络带宽越来越大，全流量测量和分析研究越来越困难，为了解决这一问题，近几年，流量抽样测量研究现成为高速网络流量测量的研究重点。
关键词：被动测量、主动测量、测度、测量体系结构、抽样测量

1．引言

Internet 是由上亿台计算机互联而成的全球性网络 [1]，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，因此近年来对大规模互联网络行为的研究成为本领域被关注的热点目标。网络行为测量和分析是网络行为学研究的基础，通过测量分析可以掌握网络的行为的基本特征（base-line），有助于寻找网络行为变化的规律，构造并验证网络行为的数学模型。所以，针对网络行为的测量与分析方法展开系统性的研究将对 Internet 行为学方面的研究取得理论突破具有重要的意义。

目前网络测量根据网络行为研究方向大体可分为拓扑测量、流量性能测量和端至端性能测量。网络的测量方法主要有主动测量和被动测量，两种测量方法各有优缺点，分别用于不同的场合。为了测量需要相应的测量工具，简单的测量工具 ping、traceroute 远远不能满足 Internet 行为测量的需要，CAIDA [2]等国际组织开发了各种专门用于 Internet 网络测量的工具 [3] ，并建立各种网络测量体系结构。由于吉比特以太网和其它高速网络技术的发展，对流量分组进行直接测量几乎成为不可能，同时，大量的流量日志也使流量行为分析相当困难，为了解决这一问题，近几年，流量抽样测量研究现成为高速网络流量测量的研究重点。目前国外网络行为测量学研究发展很快，同时也取得了较大的成绩，本文将分别对网络测量方法、测量测度、测量工具和测量体系结构项目及流量抽样研究等内容进行将专门讨论。

2．网络测量方法

Internet 流量数据有三种形式：被动数据（指定链路数据）、主动数据（端至端数据）和 BGP 路由数据，由此涉及着两种测量方法：被动测量方法和主动测量方法。被动测量方法是从网络中的某一点收集流量信息，如：从交换机、路由器或通过一个单独的设备被动地监听网络链路上的流量来收集数据。被动监测的常用形式是使用类似 RMON 的探测器或 Coral 监测器从交换机或路由器上直接收集流量信息。主动测量方法是为了监测两指定端点之间的性能而向网络中注入流量的方法。主动测量技术通常被网络工程师用来诊断网络问题，然而，近几年来，主动测量技术被网络用户或网络研究人员用来分析指定网络路径的流量行为。CAIDA 的 ISMA [4]每半年举行一次网络测量的专题讨论会，其中测量方法是讨论的重点。下面分别讨论主动测量和被动测量的行为和问题。

[1]本文受国家自然科学基金重点项目 90104031 资助

2.1 主动测量技术

主动意味着测量过程中产生新的网络流量。这些流量也许是为了引起网络部件的特殊响应（如：traceroute），也许是为了查看网络为流量提供服务类型的性能（如：treno）。主动测量给网络增加了潜在的荷载负担，特别是如果没有仔细设计使得该方法产生的流量数最小，那么附加的流量会扰乱网络，歪曲分析结果。如：为了测量在 IP 网络云中瓶颈链路的带宽，定期地向测试路径发送巨大的 TCP 流量，那么由此产生的附加流量可能会产生 Heisenberg 效应而拥塞通过网络云到达这点的路径，并且测量的吞吐量低于瓶颈链路的带宽。

另外，主动测量至少要多个网络部件某种形式的参与。如：ping 用于估计主机 A 到主机 B 的 RTT，需要主机 B 响应 ICMP ECHO 请求信息。有几种形式的合作已经广泛应用在 Internet 上，如：响应 ICMP 请求和匿名 FTP 服务器允许主机 A 和服务器之间进行吞吐量测量，可以将这种合作定义为被动合作。另一种合作方式是主动合作，如果要测量 A 至 B 路由的对称性，从 B 到 A 和从 A 到 B 同样需要进行路由测量，需要 B 也要同样主动参加测量。

跟踪和可视化 Internet 拓扑结构是主动测量最主要的应用，CAIDA 最近开发的 skitter[5]动态测量工具可用于动态发现和绘制全球 Internet 拓扑[6]。同时主动测量技术可以探测网络的特定现象，如发现许多 Internet 端至端的延迟分布具有重尾特征[7]。Internet 的健壮性和可靠性很大程度取决于 ISP 网络有效可靠的路由，Internet 路由行为的分析直接影响下一代网络硬件、软件和操作政策。主动测量应用其它领域还有：评估 IP 地址空间的利用率，路由的不对称性和不稳定性，按网络地址前缀长度的流量分布，BGP 路由表空间使用效率，单播和组播路由不一致的程度等。

2.2 被动测量技术

被动测量是在网络中的一点收集流量信息，如使用路由器或交互机收集数据或者一个独立的设备被动地监测网络链路的流量。被动测量可以完全取消附加流量和 Heisenberg 效应，这些的优点使我们更愿意使用被动测量技术。有些测度使用被动测量获得相当困难：如决定分组所经过的路由。但被动测量的优点使得决定测量之前应该首先考虑被动测量。如果关心的不是完整的 Internet 路由，而是 AS 之间的路由，那么能监测两个对等 BGP 之间的流量，因为流量中包含全部的 AS 之间的路由信息。被动测量技术的遇到的另一个重要问题是目前提出的要求确保隐私和安全问题。

网络流量是采用大小不一的报文传送，收集到的数据可以进行各种流量分析，如：流量中各种应用的成分、报文的长度分布、报文到达时间、性能和路径长度等，这些流量行为的了解能帮助设计下一代互联网设备和体系结构。

网络管理员最感兴趣的被动测量流量是流量的流矩阵，即：有多少流量从一个网络流向另一个网络的表格，这个信息能有助于优化设计决定。不同的流量粒度矩阵有不同的用处，AS 粒度流量矩阵有助于优化拓扑结构；一个公司或大学网络管理者为了了解各部门之间流量交换的情况，可以建立系或工作组粒度的流矩阵；国家粒度的流量矩阵有助于了解各国的开放策略和国际商业前景，[6]美国是世界 Internet 流量主要中转国，71％的其它国家之间的国际流量经过美国。

同时，被动测量还有许多其它的应用，包括：识别、刻画和跟踪网页缓冲和代理的优化配置；网络体系结构的安全危害；拥塞控制算法的有效性；流量增长是由于增加了用户还是每个用户流量的增加；流行协议和应用使用的变化；新的技术和协议（如：组播和 IPv6）的渗透力和影响。以上的被动测量应用是 Internet 流量行为研究的主要内容。

有时为了能够从被动收集的数据中提取某些参数可能需要借助于主动测量。另外，被动测量是应该有尽可能低的丢失率，否则测量的数据将难以进行精确估计。但随着流量速率的增加，保证不丢失数据变得越来越困难，一种可行的解决方法是使用网络流量抽样技术，下面将着重讨论被动测量中抽样技术。

3．网络流量抽样测量技术

随着吉比特以太网的出现和高速网络技术的发展，直接对网络流量进行全分组测量相当困难，另外，大规模流量数据库的维护、数据分析也相当困难。在这种情况下，将统计抽样的方法引入流量测量具有十分重要的意义，使用抽样的方法，不是收集流量的所有报文，而只是选择部分报文。在 1993 年，Claffy [8]进行 NSFNET 主干测量时，首次使用网络流量统计抽样技术，研究使用经典的事件和时间驱动静态抽样方法来减少采集的报文数。Jack Drobisz [9]等人认为这种静态流量抽样方法可能会产生不精确的流量统计资料，考虑了网络流量自相似的特点，对 Claffy 的静态抽样方法加以改进，发展了一种自适应的抽样方法，这种方法能更好地进行网络管理及对测量的流量特性进行评估。

一种常用的抽样方法是使用固定时间间隔的抽样：周期抽样。周期抽样简单，但是这种方法具有两方面问题：1）如果被测量的统计量本身具有周期行为，那么周期抽样将可能只能测量到周期行为的一部分。2）周期抽样的测量行为可能会干扰被测量的对象。另外一种较合理的抽样方法是“随机附加抽样”：样本之间是相互独立的，抽样间隔是通过一个函数 G(t)随机产生。这种样本抽样的效果取决于分布函数 G(t)。随机附加抽样具有重要的优点，一般来说，它避免了同步的影响，样本的统计量可以得到一个无偏估计。但随机附加抽样也有一些缺点。首先，由于抽样不是按照固定间隔产生，而傅立叶变换技术是假设样本间隔固定，所以抽样样本难以进行频域分析。其次，如果 G(t)不是一个指数分布，样本仍然可能会具有某些可预见性。

可以证明，如果 G(t)是具有比率 λ 的的指数分布：G(t)=1-exp(1-λ*t)，新样本的到达是不可预见的，即，样本是无偏的。泊松抽样不易引起同步，它能精确地进行周期行为的测量，且当新的样本出现时，泊松抽样的方法不易被预先控制。由于泊松抽样的方法具有如此多的优点，RFC2330[10]推荐使用泊松抽样的方法进行 Internet 的流量测量。

4．网络测量测度研究

Internet 取得的巨大成功是标准化的成功，但是 Internet 测量的标准化不满足网络快速增长的需求。甚至象如何测量沿着 Internet 路径的吞吐量和延迟等基本问题都缺乏标准测量框架，诊断问题和判定是否得到合理的性能都越来越困难。在这种情况下，IEFT[11]建立 IP 性能测度工作组（IPPM）[12]来发展一套标准的测度，它们能用来刻画 Internet 数据传送服务的质量、性能和可靠性。设计的这些测度能被网络操作者、终端用户或独立的测试组使用，这些测度并不是好坏判断的标志，而是作为不偏的性能数据测量。

定义网络测度的目的是为了使网络用户和网络服务提供商对网络的性能和可靠性有一个共同精确的认识。为了实现这个目的，IPPM/IETF 通过多次会议讨论认为定义的测度应该满足以下几方面标准，其中主要有：定义的测度需要具体且定义明确；测度的方法应该具有可重复性：在理想情况下，使用同样的方法进行测量，应该得到相同的测量结果。

Paxson[13]等人定义测度为 Internet 组成部分的不同属性。RFC2330 定义测度为“在运作的 Internet 中有一些关于 Internet 性能和可靠性的参量，这些参量的值我们是希望知道的，当这样一个参量被详细说明后，我们称这个参量为一个测度”。下面我们将讨论 IPPM 制定的 RFC 和草案。

IPPM 目前已经制定了 5 个 RFC：IP 性能测度框架（RFC 2330），这篇备忘录的目的是为 IPPM/IETF 发展具体测度定义一个通用的框架结构。测量连接性的 IPPM 测度（RFC 2678），连接性是 Internet 的基本要素，RFC2678 定义 Internet 主机对之间连接性的一组测度：1）瞬间单向连接性：定义在某一时刻一个方向的连接性，2）瞬时双向连接性：定义在某一时刻两个方向的连接性，3）相隔时间连接性：定义两台主机之间一段时间内两个方向的连接性。单向延迟测度（RFC 2679）、单向分组丢失测度（RFC 2680）和往返延迟测度（RFC 2681）均从单个、抽样和统计三个角度来定义测度。

同时，IPPM 还制定了几个测度草案：瞬间分组延迟变化测度定义通过 Internet 路径分组延迟变化的一个测度，是基于连续分组单向延迟差异的统计数据，这个测度在两台主机之间时钟同步和不同步都有效，在不同步时允许评估二者差异。经验定义的块传输能力测度的框架为标准化多种 BTC 测度定了一个框架结构，块传输能力（BTC）是测量网络通过单个 TCP 链路传输大数据量的能力，BTC 的直觉定义是期望

得到一条理想 TCP 链路的平均数据速率，然而 IETF 标准允许多种拥塞控制算法，传输算法的多样性造成了标准化测度的困难。单向丢失模式采用测度，丢失模式或丢失分布是决定用户观察性能的关键参数，对于相同的丢失率，两个不同丢失的分布可能产生对性能的感觉极其不同，丢失模式的影响对使用一个适应协议（如：TCP）的非实时应用也是相当重要，为了获得分组丢失模式，这篇草案中提出了两个衍生测度，丢失距离和丢失周期：丢失周期测度俘捉丢失的频率和长度（突发性），丢失长度测度俘捉丢失周期之间的间距。周期流量网络性能测量是关于一个 IP 网络应用层性能测量的概念，最初的动机是交互式的周期流（如，基于 IP 的多媒体会议）的 QoS，但应用层测量的思想可以有更广泛的应用。

同时 IPPM 制定了单向延迟测量协议。通过标准化收集 IPPM 单向延迟测量技术建立一个环境，在这个环境中可以通过在比目前更广阔的 Internet 路径网络上收集 IPPM 测度数据。一个特别引人注目的设想是 OWDP 服务器的广泛部署，OWDP 服务器测量单向延迟就象使用 ping 测量往返时间一样普通，另外 OWDP 包括加密、安全、符合逻辑的控制和测试功能独立、支持小测试分组等。

5．网络测量体系结构研究

历史上 Internet 一直不能很好地测量，随着网络地规模的增加，这种问题变得越来越严重。CAIDA 等组织通过发展 Internet 测量体系结构来解决网络测量问题，这些体系结构由分布在网络不同位置的专有工作站的测量平台构成，这些工作站相互协作，相互交换测试流量来测量网络路径和网络云的属性。

测量体系结构有以下几个用途：1、通过从不同点探测网络的特性，诊断网络内部的性能问题；2、测量广大范围网络路径的属性，来研究网络行为和进化[13]；3、测量从 ISP 网络的一个端点到网络另一端点的探测流量，评价不同 ISP 的性能；4、综合 ISP 性能的评估，为 ISP 提出合理的建议，促使他们优化自己的网络。

测量 Internet 端至端性能的特性需要分析一系列不同 ISP 的行为，没有一个 ISP 能完全依靠自己的网络完整的监控这一行为，因此，体系结构的各部分需要相互合作共同管理。Internet 的异质性决定测量体系结构不能只属于一个单位或组织，如果希望得到广泛地部署，ISP 必须能自己决定是否配置测量平台，有以下动机决定他们这样的行为：1、体系结构有利于在他们的网络或对等地网络中监控和调试性能问题，使他们能更好地操纵他们的网络；2、将来，ISP 和用户之间的“服务水平约定”将提供在线测量，使用户能检测他们是否真正获得协议中规定的服务；3、ISP 可以限制使用他们的平台来测量他们自己的网络，而且他们可以利用自己的平台有效地测量竞争者的网络性能。ISP 可以使用“测量政策”来限制公众利用他们的平台测量流量的类型。测量政策指谁能用哪个平台进行什么样的测量。

CAIDA 归纳了目前国际上重要的监测 Internet 流量的测量体系结构项目。根据测量方法的不同，测量体系结构也可以分为主动测量体系结构和被动测量体系结构。下面我们来讨论具有有代表性的主动测量体系结构项目和被动测量体系结构项目。

目前国外主要的主动测量体系结构项目有：CAIDA 开发了 skitter 工具[5]测量从一个源点到数千个目标点的 IP 路径和 RTT，使用主动探测来测量获取和跟踪全球 Internet 拓扑。Paxson 等人[13]在 Internet 的 50 个站点上部署了网络探测幽灵，详细研究了路由、延迟、丢包和 TCP 的动态性。IPMA 项目[14]的路由协议收集器是一个广域部署的探测幽灵，为了理解 Internet 路由的动态行为，这些幽灵跟踪路由协议的更新。Pinger 项目[15、16]监测通过高能物理研究所（HEP）的站点，pinger 根据一个固定的时刻表发送一系列的 ICMP 请求信息。WAND 项目[17、18]进行一些单向延迟和丢包测量，使用 GPS 来同步时钟，WAND 能被动获取 ATM 位元分组，另外 WAND 还开发了以太网卡驱动，可以在 sonet 接口上获取 IP 分组。RIPE[19]是非常类似于 Surveyor 的项目，也是 IPPM/IETF 的单向延迟和丢包测度的一个应用，它的使用范围是同 RIPE 联系的欧洲网络。NIMI[20]是建立一个网络测量通用平台的测量体系结构，NIMI 原型测量延迟、丢包和单向路由，NIMI 为满足 IPPM 单向延迟测度注意同步时钟。

同时 CAIDA 等组织也发展 Internet 被动测量体系结构测量网络。其中重要的被动测量体系结构有：CoralReef[21]是 CAIDA 开发的一套完整的软件包，提供用于进行被动测量的网卡驱动、库函数、及流量

分析软件，用于刻画 UCSD 和商业 Internet 之间一条链路的流量特性。NLANR 的测量和运行分析组(MOAT)正建立一个网络分析体系结构(NAI)[22]，该体系结构使 Internet 的系统服务模型和测度得到一个更好的理解，体系结构包括对分组头记录分析的被动测量、主动测量、特定服务器上的 SNMP 信息和基于 BGP 数据的 Internet 路由相关信息等。NPACI[23]网络气象服务，周期地监测分布式系统和动态地预测各种网络的性能。WAND 项目[24]为统计分析和仿真模型构造建立 Internet 流量模型。

6．结语

网络流量行为测量是理解网络行为的首要条件，根据流量行为的客观数据判断网络行为的状况。测量网络行为之前，需要定义描绘网络行为的刻度，如：响应时间、丢包率、利用率和可用性等，IPPM/IETF 为进行网络测量定义了测量体系结构和各种测度。根据研究对象和定义测度的不同有不同的测量方法：主动测量和被动测量，它们有各自的优缺点，针对不同的应用。由于目前高速网络的出现给被动测量带来了困难，近几年抽样的被动测量技术的成为流量测量的重点内容。由于网络测量是网络安全、网络管理和规划的基础，出于不同的目的，以美国为首的很多国家都建立了测量体系结构项目，同时开发了各种流量测量工具。本文从测量方法、抽样测量、测度定义及测量体系结构项目等角度对当今测量技术的发展状况进行了讨论。

根据以上讨论的几方面问题，进行网络流量测量需要考虑以下几方面问题：首先根据研究问题选择测量方法，采用主动测量方法还是被动测量方法。其次选择网络中测量点的位置，测量点位置的选择需要考虑以下几个方面：要求知道测量网络的拓扑结构；测量点能发现在使用的 IP 地址范围；测量点最好能监控每条链路，较为简单的方法是只监控最忙的链路或者监控点设在网络的边界。第三：需要考虑测量的测度，可以从 IPPM/IETF 定义的测度中选择一种或几种，当然根据实际问题的不同，也可以定义自己的测度，但定义的测度最好能符合 RFC2330 中的测度定义要求。最后涉及的问题是数据的收集和归档，需要测量和存储怎样的数据？数据怎样存储和访问？为了用户易于访问数据，提供什么样的接口，如每天或每星期的 Web 页面。

CERNET 华东（北）地区网络中心为网络安全、网络计费管理和网络行为研究等，综合国外测量体系结构项目及 IPPM/IETF 相关网络测量文档，建立了用于监控 OC48 的 CERNET 主干光纤的被动测量体系结构。目前正在进一步研究提高测量体系结构的高速网络流量采样技术和抽样流量测量技术。根据网络测量流量，已建立了相应的网络管理系统、网络计费系统和网络安全监测系统。当然，测量只是让我们对网络行为有了感性认识，为了网络行为有理性认识，还要在此基础上借用分析、建模的方法来研究。

参 考 文 献

[1] Number of Hosts advertised in the DNS, Internet Domain Survey, July 2000 http://www.isc.org/ds/WWW-200007/index.html.

[2] CAIDA Homepage, http://www.caida.org.

[3] CAIDA Tools site, http://www.caida.org/tools/.

[4] ISMA web page, http://www.caida.org/outreach/isma/.

[5] skitter Web Site, http://www.caida.org/tools/measurement/skitter/

[6] B. Huffaker, Marina Fomenkov, David Moore, Evi Nemeth, K. Claffy, Measurements of the Internet topology in Asia- pacific Region, 2000, http://www.caida.org/outreach/papers/asia_paper/

[7] Kc Claffy, Sean McCreary, Internet measurement and data analysis: passive and active measurement, ASA99 paper, http://www.caida.org/outreach/papers/Nae/4hansen.html

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

[9] Jack Drobisz, Kenneth J. Christensen, Adaptive Sampling Methods to Determine Network Traffic Statistics including the Hurst

Parameter, 23rd. Annual Conference on Local Computer Networks, October 11-14, 1998.
[10] V. Paxson, G. Almes, J. Mahdavi, M. Mathis, Framework for IP Performance Metrics, IETF RFC 2330, May 1998.
[11] The Internet Engineering Task Force, http://www.ietf.org/
[12] IP Performance Metrics (ippm), http://www.ietf.org/html.charters/ippm-charter.html
[13] V. Paxson, Towards a Framework for Defining Internet Performance Metrics, Proceedings of INET 96.
[14] C. Labovitz, et al., "The Internet Performance and Analysis Project,", http://www.merit.edu/ipma/.
[15] W. A. Matthews, R. L. A. Cottrell, and D. E. Martin, "Internet Monitoring in the HEP Community," Proceedings of Computing in High Energy Physics 1998 (CHEP98), Aug. 1998.
[16] L. Cottrell, "PingER Tools," http://www.slac.stanford.edu/xorg/icfa/ntf/tool.html, May 1998.
[17] I. D. Graham, S. F. Donnelly, S. Martin, J. Martens, and J. G. Cleary, "Nonintrusive and Accurate Measurement of Unidirectional Delay and Delay Variation on the Internet," Proc. INET '98, Jul. 1998.
[18] I. D. Graham, et al., Waikato Applied Network Dynamics (WAND) Project homepage, http://atm.cs.waikato.ac.nz/wand/
[19] RIPE Network Coordination Centre, Test Traffic Project Homepage, http://www.ripe.net/test-traffic/index.html
[20] V. Paxson, Jamshid Mahdavi, Andrew Adams, Matt Mathis, An Architecture for Large-Scale Internet Measurement, IEEE Communications, 1998.
[21] CoralReef, http://www.caida.org/tools/measurement/coralreef/
[22] NLANR(MOAT)-PMA Passive Measurement and Analysis, http://moat.nlanr.net/PMA/
[23] NPACI's Network Weather Service, http://nws.cs.utk.edu/
[24] WAND (Waikato Applied Network Dynamics) Project, http://wand.cs.waikato.ac.nz/

A research on traffic measurement in a large-scale high-speed network

Guang Cheng（程光），Jian Gong（龚俭）
(Computer Department of Southeast University　　Nanjing 210096, P.R.China)

Abstract: It is very important to understand network behavior while the development of networks due to network management, network planning and network development. And the measurement of network traffic is the base of the study of network behavior. According to the difference of measured traffic, measurement method is consisted of passive measurement and active measurement that have their advantage and disadvantage. It must define a number of metrics to study the different network traffic behavior. The IPPM of IETF has defined a number of metrics to study traffic behavior. Due to the difference of measured environment and application, many organizations have built various measure architectures and measure tools. In addition, it is more difficult to measure and analysis full traffic trace while the bandwidth is increasing larger, so the traffic sampling measure becomes the focus of the study of high-speed network traffic measurement in recent years.
Keywords: Passive Measurement, Active Measurement, Metrics, Measure Architecture, Sampling Measurement

作者简介：程光，男，28 岁，博士研究生，主要研究方向为网络行为学、网络管理。龚俭，男，43 岁，教授、博士生导师，主要研究方向为网络安全、网络管理、网络行为学