网络的行为观测

龚俭　吴桦

（东南大学计算机系　南京 210096）

摘要：本文介绍了网络行为观测的基本概念和有关的研究内容，并对其中的一些重要内容进行了讨论，包括网络行为观测的基本测度，观测方式，主要观测内容，观测数据的存储方法，以及观测数据的处理与显示问题，为这一领域的研究提供了背景信息。

关键词：网络管理，网络测量，性能测度，网络行为。

1. 网络的测量问题

以 Internet 为代表的计算机互联网络已成为人类社会最重要的基础设施之一，对人们的经济与社会活动正产生着日益广泛深入的影响，因此了解的网络行为十分必要。Internet 的规模和管理结构越来越复杂，每个网络管理者只能在自己的管辖范围之内对基础设施进行分析和诊断，因此在全局范围内形成了分布式的运行管理体系。由于缺乏集中的控制，Internet 的行为观测和性能管理因缺乏协调和网络服务提供者 NSP 之间出现利益冲突等因素而成为一个阻碍 Internet 健康发展的重要问题。

首先，NSP 发现保持一个大规模互联网络正常运行已经是一件十分艰巨的任务，他们没有更多的时间、精力和财力来进一步对网络行为进行统计和分析。其次，他们也不知道应该采集和统计些什么数据；由于缺乏公认的模型，因此各自的观测结果未必有可比性。现有的 NSP 基本上都是电信公司，熟悉的是电话网的管理，然而 Internet 的流量有自己的特性，例如连接内报文的达到、流内连接的达到、报文在多个连接中的达到等分布，都不能用传统的电话网数学模型来描述，相关的理论研究远落后于工程技术的发展。第三，早期的 Internet 是免费的，因此用户和 NSP 都没有成本概念，并不强调对资源的合理使用，因此对于网络的性能管理和行为观测没有受到特别的重视。但随着 Internet 服务成为一种商品，NSP 和用户都开始要求能够确认和验证所提供的和所得到的网络服务质量，因此不仅对于 NSP，也对于用户和研究者，都需要有能力预测、验证和实时测量网络的行为。

对有关网络行为观测与分析研究的系统讨论源于 1996 年初美国的应用联网研究国家实验室（NLANR）与 Bellcore 在美国 NSF 的支持下召开的一次有关 Internet 统计与测度分析的研讨会（ISMA），在此之后，依托于美国加州大学圣迭戈分校超级计算中心（SDSC）的 CAIDA（Internet 数据分析联合会）对网络测量的相关理论和方法展开了系统性地研究，以寻求提高 Internet 基础设施可靠性和可扩展性的方法。

从 NSP 角度进行的网络行为观测往往与网络性能管理联系在一起，并使用相应的方法和工具。但是每个 NSP 所作的网络行为观测只能是在自己所管辖的网络范围之内，而且缺

乏验证手段。从用户角度进行的网络行为观测表现为端－端的测量活动，现有的方法基本属于 PING 一类；也有一些研究活动是针对选定路径的性能和可靠性的标准化测量和评估而展开，但总体看来进展缓慢，适用性也较差。

2. 网络行为的基本测度

测度是测量数据采集的依据和分析的基础。网络的行为主要是通过网络的性能来体现的，但由于目的不一样，因此不能简单地套用网络性能的有关测度。由于网络结构和应用的多样性和复杂性，建立同时满足 NSP 和用户评估网络行为需要的测度是一个困难问题。围绕这方面的工作很多，比较著名的有 CAIDA 和 IETF 的 IPPM （The IP Provider Metrics）工作组，通过他们的工作，提出了一些公共的测度，可作为规范观测数据和结果的基础。

- 可用性 (Availability)

对服务是否可用的测量。可用性经常表示为一段时间内资源可供使用的百分比，例如 99.9%，这是一种评定网络是否健康的基本方法。但可用性并不总是与性能相关，例如一个很忙的网络由于太慢而无法使用，但是所有的资源都是可用的。可用性的观测通过对网络工作情况进行统计而获得。

- 带宽 (Bandwidth)

对通信连接容量的测量。这里的容量可以是物理意义的，例如一条 T1 连接的带宽是 1.544Mbps；也可以是逻辑意义的，表明一条连接上分配给某种服务的容量，例如为 T1 上视频点播分配 384Kbps 的带宽。带宽描述了连接的固有属性，对于用户和 NSP 可能需要使用不同的方法来测量。NSP 可通过网络管理系统对带宽进行直接观测；而用户往往只能通过某些测量工具对带宽进行间接测量。

- 基线 (Baseline)

对“正常状态”的测量，较多地为 NSP 使用。这是一种广义的均值概念，用于刻划网络和用户的某种正常行为，需要通过对观测值进行统计处理而获得。通过将当前观测到的行为与相应的基线进行比较，可以将不正常的状态从“正常状态”中区分出来。基线帮助管理者辨认出可能导致故障的突然变化；还可以通过长时间的基线观察为将来的网络规划提供依据。

- 延迟 (Delay)

从网络、连接或设备的一端到另一端的延迟测量，反映网络或系统的负载情况，通过网络的传输延迟和系统的响应时间来表示。做为网络性能的状态参量，延迟总是存在的；而且可以随着网络负载和结构的变化而变化。延迟对于 NSP 和用户都是有意义的，它是辨认出网络潜在问题的关键量度。例如，如果响应时间变大，但是延迟没有改变，可认为问题出在客户机或是服务器。

3. 网络行为的观测方式

网络的行为观测可分为被动测量和主动测量两大类。

被动测量是指在流量通过时，在网络内部某个观测点，从中继系统或某个检测系统中得到的网络流量数据。基于从中继系统采集的方法如 NetFlow，它使用 SNMP 协议从路由器或交换机中得到流量数据，在设备成本上优于使用专门的检测系统的采集方法，使用较为普遍。这种方法的缺点在于当采集频率增加时，会过分占用设备的处理能力，降低网络性能，而频率较小时，由于缓充区的大小有限，会丢失数据，如何在两点间寻求平衡点，是一个比较复杂的问题 。有些文献指出可以用排队论等模型求出最佳点。通过检测系统的方法如 OC3mon，它通过光纤探头得到 155M 链路上的流量，这种测量需要特定的设备，在一定程度上限制了它的应用。被动测量依赖被测网络的流量，对现存的流量无影响，但也有可能无法或很难从观测到的流量中获得所需的信息，例如响应时间。

主动测量表现为为了测量特定两端间性能时，必须引入新的流量；如 ping，traceroue 等网络故障诊断工具都是主动测量工具。这种通过发出测试报文得到观测值的方法常用于问题定位，灵活方便，但要求对网络的原有状态不能产生大的影响，从而需要制定合适的测试方案，开发适用的测试工具。由于主动测量通过产生流量来测试网络的性质，所以可更有效地获得所需的信息，但会对现存流量产生影响，测量结果也会受干扰。

无论对于被动测量还是主动测量，时间的准确性是一个关键问题。虽然很多情况下，并不要求绝对准确，但是在大规模网络中的同步问题仍然是影响测量结果的一个重要因素。

4. 网络行为的观测内容

不同的 NSP 和用户出于各自的兴趣和利益，在网络行为的具体观测内容上既存在差别，也具有共性。总体看来，当前对 Internet 网络行为的观测主要集中在四个方面。

(1) 拓扑测量

通过对网络拓扑的观测可获得网络基础设施结构的直观描述，它分为链路级和路由级两个不同的层次，与传统的网络配置管理所理解的拓扑发现有所不同。拓扑测量通常采用通过在现存的 IPv4 地址空间中从几个源点到一组（可能是上万个）宿点探测通路情况的方法，可获得 Internet 一部分区域的连通度描述和路由结构，在形式上类似于拓扑发现。如果欲观测的网络过于庞大，则往往不可能通过一个拓扑发现动作来进行观测（例如对一个大的 ISP 的拓扑结构发现可能需要几十个小时的操作，几乎已无实时性可言），因此需采用 X－射线成像的方法从多侧面收集数据，然后进行合成显示。拓扑观测的结果可采用拓扑结构表示和地理结构表示两种方法，可以采用三维技术和 VRML 等表示技术。需要开发 IP 地址与地理信息之间的映射方法，以更加直观地向用户表达观测结果。另外通过拓扑观测，还可以检测路由的变化频度和模式以发现替换路由的使用方式；并通过拓扑分析使 NSP 和用户发现全局网络中的关键点等。

(2) 负载测量

通过在观测点收集网络负载数据并进行分析，可以获得网络的流量特性，这是网络最重要的行为特征之一。负载测量的主要研究内容包括不同链路速率（OC3，OC12，OC48，DS3）、不同接口类型的监测能力，封装和成帧的有效处理方式，不能实现对流量的完全采集时的性能评估方法，测试粒度的可配置性，负载测量的安全性和可管理性等等。负载测量需要从网络的关键点被动检测流量信息，计量方法可按报文数或字节数统计，内容包括流量中的应用分类（看看支持流量控制的友好协议，如 TCP，和不支持流量控制的不友好协议，如 UDP，的比例），报文长度分布，报文间隔时间，性能，通路长度等，这些指标对设计下一代交换设备很有用。对于运行管理员而言，不同粒度的流量分类矩阵很重要，它表明了特定网络之间的流量情况，是规划这些网络之间的路由交换政策和互联点的依据。

(3) 性能测量

性能测量是传统网络性能管理的主要内容，是网络可用性和使用率评估的依据。对于网络行为观测而言，性能测量不仅是 NSP 的管理员所能进行的活动，也是用户评估和验证网络服务的重要手段。因此用于网络行为观测的性能测量除了传统的主动测试端－端之间的传输延迟外，还要研究单向延迟的测量，非对称路由的测量，以及测量活动连接的容量而不影响性能的方法等。此外，网络服务提供者还可以根据性能测量的结果规划设计 cache 结构，以优化网络的传输负载。

(4) 路由测量

路由测量主要针对 BGP 的路由表进行，分析在指定的时间和地点 AS 的互联关系。通过路由测量可以发现向周边 ISP 输出的效果，拓扑变化对网络性能的影响，新路由政策的结果，单个网络对阻塞或拓扑变化的反应，基础设施由于依赖特定关键通路而产生的脆弱性等等。另外路由测量可以检测实际的信息传输行为与路由政策的差别，发现确定性能下的优化路由。

5. 观测数据的存储方法

为了长期保存和使用方便，必须采取有效的方法压缩和存储得到的海量观测数据，可以分为集中式和分布式。

集中式方法是将数据存放在一个特大型数据库里，由观测分析系统根据需要提炼摘要。由于数据量巨大，原始数据一般无法存放较长时间，需要定时删除或存储在其它媒介如磁带上。这种方法对观测分析系统的宿主机压力很大，而且大量的原始数据从代理机发至处理系统会占用大量的网络带宽。

分布式方法将观测数据存放在分布式数据库中，同时使用某种关联方法来帮助分析系统获得所需的数据。例如哥伦比亚大学计算机系设计了一种较为有效的分布式存储方法：将观察到的大量数据简化为表示网络状态的索引值，这种索引值类似于证券市场的道琼斯指数，

称为健康函数。常用的健康函数是一些 MIB 值组合起来构成状态参量，如某系统利用率可表示为：$U(t)=\frac{(ifInOctets+ifOutOctets)*8}{(ifSpeed*sysUptime*100)}$，链路相对于传输总量的输入瞬时错误率为：$E(t)=\frac{ifInErrors}{(ifInUcastPackets+ifInNUcastPkts)}$，然后将这些状态参量进行线性组合，得到健康函数。如 $H\left(\vec{e},\vec{u}\right)=\vec{A}\vec{e}+\vec{B}\vec{u}$，其中 $\vec{A},\vec{B}$ 为权重因子。这些健康函数与系统的安装、配置和使用有关，并有可能随时间变化，不能地定义成标准形式的 MIB，必须有针对地定义。代理机根据定义在本地算出索引值后，发送给处理系统。这种方法节约了传输带宽和处理时间，减轻了处理系统的负担。

6. 观测数据的分析和显示

通过对网络行为观测数据的分析，可以发现网络和用户的行为是否恰当，网络协议和网络应用的设计是否合适，网络的管理是否合理。具体的分析指标包括流量中各种协议和应用的比例；流长度的分布；报文长度分布；IP 分段的统计；地址前缀长度统计；IP 地址空间的使用效率；ISP 的路由不稳定性和非对称路由的范围变化特性；不同网络地址前缀的流量分布；BGP 路由表空间的使用效率（例如聚合程度）；传输流和路由对某一小部分地址或实体的偏好；常规和组播路由之间的不平衡程度；删除某个特定的 AS 之后对网络连通性的定量影响等等。

对观测数据最简单的处理就是依据一定的要求对数据进行摘要，形成实时或历史的图表提供给管理员和用户。较高级的数据处理是依据历史数据来及时发现当前网络行为的异常，并能够对网络行为的变化趋势进行一定的预测或展望。由于网络各部分的状态以及一个部分各时间的状态是互相关联的，因此有可能从历史状态大致推测网络的变化趋势。但是影响网络状态的因素复杂多样，一些网络状态的突变是由一些非常的外部事件如信道升级、关键设备发生故障等引起的，因此状态的变化又带有随机性，用一个确定的模型来完全拟合网络的过去和现在及未来的状况是非常困难的。有些研究者尝试了借助于统计数学模型方法或其它一些数学手段，并利用历史状况来推导能够大致地描述网络行为变化规律的模型，例如使用季节模型、最大熵谱方法和奇异谱方法等等。

为了给使用者提供方便，网络行为的报告基本上都是基于 WEB 的图表，它们具有直观、数据容量大、清晰等优点。较为常见的图表形式有曲线图、饼图和直方图，表达的内容包括实时图表、历史图表和趋势图表。这些图表不仅有传统的以时间轴为主线的二维表现形式，也有三维的和使用 VRML 等技术的高级表达形式以更清晰地描述网络的行为。例如用二维时间轴表示的三维统计图可表达一段时间内同一时刻的网络行为比较；使用 VRML 技术可

以获得可动态旋转和放大缩小的三维拓扑结构表示。

此外，位映象显示也是网络行为的一种图形化表示方法。例如，可以对显示进行设置，使得当某个连接的使用率升高时，它在网络地图中的图像就变粗或改变颜色；网络连接出现错误时其图形可以闪烁；网络连接发生高出错率时可用虚线表示；当设备使用率上升时，它在网络图上的图像可以变大；出错率正在上升的网络设备可以显示出一条裂纹等等。利用位映象显示发的好处是可以对网络设备和连接的一般性能一目了然。

7. 结论

网络行为观测是网络行为学的重要内容，它从全局的角度观察网络的性能变化，为网络的服务质量控制和网络的运行管理提供了新的思路，已经成为网络管理领域中一个活跃的研究分支。

随着计算机互联网络应用的日益广泛，了解网络行为的需求越来越迫切。然而现实的情况是 ISP 们将主要精力放在了自己网络的管理和用户需求的满足，缺乏彼此之间的合作来协调对整个网络行为的控制，例如处理性能问题和安全事件；而且缺乏作为判别依据的历史数据 (baseline) 来发现系统行为的变化。另外网络行为观测的政策与法规也不能适应当前的需要。例如用户或 ISP 信息的私有性与进行网络行为观测的关系并不清晰，这就给确定观测粒度和观测结果的使用范围带来了困难。网络行为观测的概念主要来自于网络的性能管理，目前面临的问题分为两个方面：高速网络的行为观测和性能管理的可用性。前者主要涉及网络测量的问题：现有的监测工具不适应高速信道；OC－3 和 OC－12 信道用得越来越多；传统的网管性能数据采集不能满足对于流量模式和发展趋势分析管理的需要；后者涉及用户对网络行为的验证和了解能力：用户缺乏适当的工具和方法来确信或证明自己得到了（或没有得到）应得到的服务质量。

参考文献

1. ISMA Report - 9 /98 Internet Statistics and Metrics Analysis: Engineering Data and Analysis Workshop Report
2. http://www.ietf.org/html.charters/ippm-charter.html
3. http://www.caida.org
4. Claffy, k & Monk, T. What's next for internet data analysis? in IEEE Special Issue on Communications in the 21st Century 85, 1563-1571 (1997).
5. Claffy, k, Miller, G. & Thompson K. The nature of the beast: recent traffic measurements from an Internet backbone. in Proceedings of INET'98(ISOC, Washington, DC, 1998).
6. J. Apisdorf, k claffy, K. Thompson, and R. Wilder.OC3mon: Flexible, Affordable, High-Performance Statistics Collection, http://www.isoc.org/isoc/whatis/conferences/inet/97/

proceedings /F1/F1_2.HTM

7. K. Thompson, G. Miller, and R. Wilder. Wide Area Internet Traffic Patterns and Characteristics, IEEE Network, Nov 1997.

* 本文得到国家 863－317－01－03－99 重大课题资助。

** 龚俭 教授、博士生导师，主要研究兴趣包括网络安全、网络管理、网络体系结构；吴桦 硕士、助教，目前主要研究网络性能管理和网络行为预测。

通信地址：江苏南京 东南大学计算机系 邮编 210096 联系电话：025－3614341

电子邮件：jgong@njnet.edu.cn