DOI：10.7652/xjtuxb201308004

对称性的起因：非业务性网络流量新特性的挖掘与探索

李强[1]，秦涛[2]，管晓宏[1,2]，郑庆华[2]
（1.清华大学自动化系智能与网络化系统研究中心，100084，北京；
2.西安交通大学智能网络与网络安全教育部重点实验室，710049，西安）

摘要：针对当前非业务性网络前后向流量呈现对称性的问题，提出一种基于双向网络流模型的分析方法。与传统的单向网络流模型相比，双向网络流模型以主机为粒度实现流量双向聚合，利用网络流在协议、端口分布和流对称性方面的特征指标刻画主机间的交互行为，从而发现流量对称性产生的根源。基于西安交通大学校园网流量的实验结果表明，1%的主机承载了 90%以上的前向 UDP 流量，是造成流量对称性的根本原因。相比于传统的网络流模型，文中提出的模型和分析方法能快速定位造成流量对称性的主机，更适用于非业务性网络流量的控制与管理。

关键词：流量测量；非业务性网络；流量对称性；UDP 流量
中图分类号：TP393.2 **文献标志码**：A **文章编号**：0253-987X(2013)08-0019-07

Cause of Symmetry for Traffic Volume: Mining and Exploring New Characteristics of Traffic in Non-Business Networks

LI Qiang[1], QIN Tao[2], GUAN Xiaohong[1,2], ZHENG Qinghua[2]
(1. Center for Intelligent and Networked Systems, Tsinghua University, Beijing 100084, China;
2. MOE Key Laboratory for Intelligent Networks and Network Security, Xi'an Jiaotong University, Xi'an 710049, China)

Abstract: Measurement methods based on bi-directional flow model are proposed to investigate the symmetric characteristic of the non-business network traffics. The features of protocol, port distribution and flow size are extracted from host-level aggregated traffic to characterize users' interactive behavior and to find the origins of traffic symmetry. The analysis results using actual traffic traces collected from the campus network of Xi'an Jiaotong University show that 1% of the total hosts bear more than 90% of the forward UDP traffic, which results in the symmetry of total traffic. Compared with the traditional flow model, the proposed methods can quickly locate the hosts that are responsible for traffic symmetry, and are more suitable for traffic control and management in non-business networks.

Keywords: traffic measurement; non-business network; symmetric traffic; UDP traffic

当前，互联网高速发展，网络带宽不断增长，并不断涌现出新兴的网络应用，互联网已经成为人们日常生活中不可或缺的一部分。据中国互联网络信息中心(CNNIC)2012 年 1 月发布的《第 29 次中国

收稿日期：2012-12-23。 作者简介：李强(1983—)，男，博士生；管晓宏(通信作者)，男，教授，博士生导师。 基金项目：国家自然科学基金资助项目(60921003，60110320)；国家科技支撑计划资助项目(2011BAK08B02)；中央高校基本科研业务费专项资金资助项目。
网络出版时间：2013-06-06 网络出版地址：http://www.cnki.net/kcms/detail/61.1069.T.20130606.1650.001.html

互联网络发展状况统计报告》[1]显示,目前搜索引擎是仅次于即时通信的第二大网络应用。与此同时,传统的电子邮件、论坛及 BBS 等的使用率则有所下降。微博已经成为近一半网民使用的重要互联网应用,网络音乐、网络游戏和网络视频所代表的网络娱乐类应用继续稳定增长,特别是网络视频用户增幅明显,已经成为第五大互联网应用。

这些新兴的网络服务应用在丰富人们生活的同时,也使网络流量统计特征呈现出多样性和复杂性。特别是基于 P2P 的文件共享、优酷和土豆等视频分享点播平台以及迅雷下载等多种热门网络应用被大量使用,使得当前非业务性网络(例如企事业单位网络、校园科研网络等)的前向数据量和后向数据量基本相等,网络带宽更加难以管理,并且这些网络服务和应用往往采用 UDP 协议作为文件传输和视频播放的主要传输协议[2-3],极大地改变了传统网络环境下的流量特性。由于 UDP 是面向无连接的传输层协议,不受拥塞控制算法的调节,在近年来 UDP 流量在整体流量中的比例逐步提高的趋势下,会对网络上的核心节点路由器以及网络交换设备带来压力,因此,如何对这些 UDP 流量进行优化管理,提高网络带宽的利用率,从而提高路由能力和效率,就显得很有研究的价值和必要。

然而,目前国内外针对 UDP 流量的特性监测和度量的研究并不多。张艺濒等人考虑到当前 UDP 流量的重要性,在国内骨干网上针对 TCP 和 UDP 流量,分别在传输层和应用层对流大小、流速率和流持续时间等流量特征指标进行了详细的测量和统计分析工作,并对应用层协议流的大小、快慢和长短进行了分类[4]。张敏等人通过对从美国和瑞典的骨干网上捕获的从 2002 年到 2009 年间的流量数据进行分析,发现 UDP 与 TCP 流量的比值逐年提高,特别是在流数上;他们还对 UDP 流量的端口使用情况进行了统计[5]。张敏和陈常嘉专门对校园网和美国骨干网数据集里的 UDP 流量特性进行了测量分析,主要包括端口号、包长和字节等基本特性[6]。这些工作都是在 UDP 流量逐渐增长的背景下开展的,但都偏重于整体宏观上的分析,只对 UDP 的基本特性进行了测量分析。John 等人对瑞典骨干网上的数据集在 IP、TCP 和 UDP 这些协议层次上比较了前向和后向的流量特性差异[7],但是他们主要针对 TCP 流量做了详细的分析,对 UDP 流量则讨论不深,并且对目前非业务性网络的流量对称性缺乏分析。

针对上述问题,本文以西安交通大学校园网络(XJTU)为研究对象,收集了多个网络流量样本,对非业务性网络流量的对称性进行度量和分析。首先,对流量样本进行宏观的统计分析,发现校园网络流量整体上具有对称性,也即前向和后向的流量基本相当,特别是 TCP 流量占据了主要的后向带宽,而 UDP 流量则占据了主要的前向带宽。随后,提出一种双向网络流模型对造成这种现象的原因进行了深度分析,借此度量了 UDP 流量在前向和后向上的流量分布、网络流大小、连接度变化等特征,发现是一些典型 P2P 应用产生了这些 UDP 流量,并使用独立实验对这一结论进行了验证,对造成非业务性网络流量对称性的原因进行了初步分析和探讨。

1 网络流量数据的整体特性分析

1.1 数据集描述

本文一共使用了 3 个数据集,其中有 2 个是在美国骨干网 OC192 链路上使用被动测量技术采集到的匿名化的数据集,来自 CAIDA(The Cooperative Association for Internet Data Analysis)[8-9],另外一个是利用搭建在 CERNET(中国教育科研网)西北网络中心的高速网络流量分解、存储和监控平台,通过镜像捕获采集到的来自西安交通大学学生区 3 栋学生宿舍楼(约有 1 000 名学生)的网络流量,每个用户具有独立 IP 地址,样本采集自 2010-11-12 从 11:00 到 22:00 的流量。表 1 为数据集的详细信息,从中可以看出,在校园网络中,UDP 流量已经超过了 TCP 流量,成为了当前网络主要的传输协议。

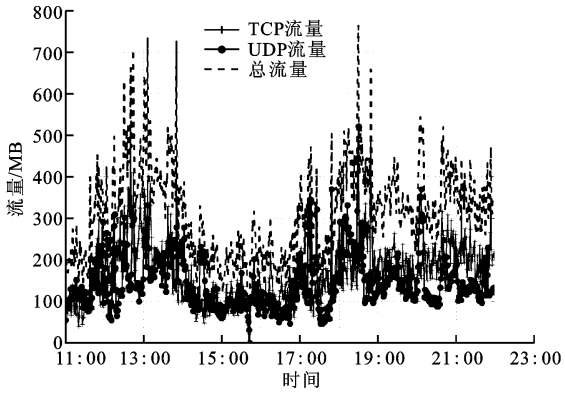
表 1 数据集描述

数据集	日期(时长)	数据包数/10^6	字节量/GB	UDP/TCP 流量参数比	
				包数比	字节量比
CAIDA-2008	2008-05-15 (40 min)	500.40	242.9	0.13	0.08
CAIDA-2011	2011-04-13 (1 h)	511.53	195.7	0.21	0.24
XJTU-2010	2010-11-12 (11 h)	357.01	201.9	1.29	0.88

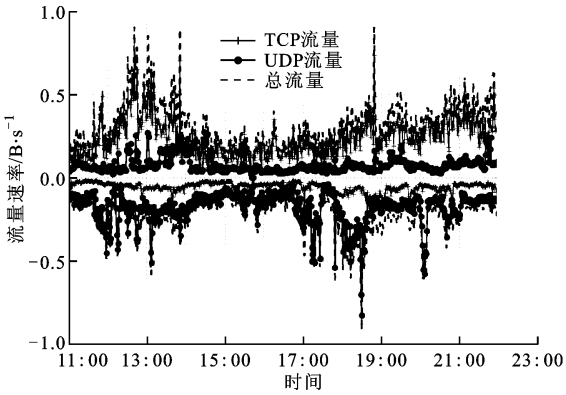
1.2 整体特性分析

从表 1 中显示的 2 个 CAIDA 数据集的 UDP 与 TCP 流量在数据包数和字节量上的比值可以看

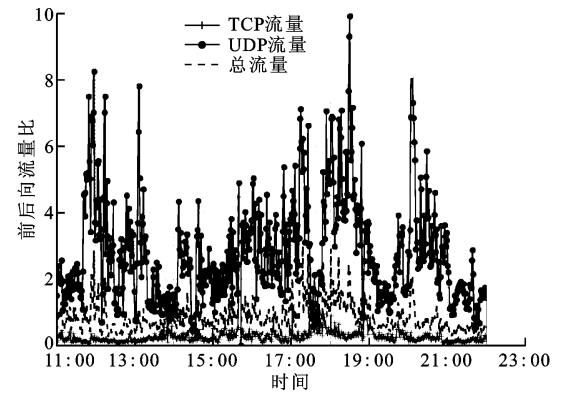
到，目前 UDP 流量占的比例高于 2008 年时的比例，在 XJTU-2010 数据集上可以看到，UDP 和 TCP 流量参数的比值也是较大的，后面将只对 XJTU-2010 数据集的 UDP 流量进行分析。图 1 显示了西安交大校园网络流量的变化情况，时间窗口大小取为 1 min。从图 1a 显示的整体流量上可以看出，校园网流量随着时间发生变化，但短时间内其带宽较稳定，同时 UDP 流量也几乎与 TCP 流量相当。图 1b 显示了 TCP、UDP 及总流量速率在前向和后向随时间的变化情况，从图中可以明显看到，在后向流量上 TCP 速率较大，这也符合非业务网络的基本特征：网络的主要应用目的是获取资源。但是，在前向流量上 UDP 速率较大，这是造成非业务网络流量对称性的基本原因。为进一步探究网络流对称性产生的原因，我们分析了前、后向网络流流速的比值，分析结果如图 1c 所示。由图示结果可以看出，前、后向流量的总体流量速率基本相等，其比值围绕 1 有少许波动，但是 UDP 的前向速率远远大于其后向速率，有时差别在 10 倍以上。经统计发现，在前向流量中 UDP 流量占据了 76.5%，这是造成对称性的基本原因。要理解这种新的特性，需要对这些 UDP 流量的特性进行测量并深入分析。如何获取 UDP 流量的特性，对网络管理具有重要的意义。



(a)整体流量



(b)前后向流量速率的变化



(c)前后向流量速率比值的变化

前向流量速率为负值，后向流量速率为正值

图 1　西安交通大学校园网络流量示意图

2　双向网络流模型及其特征指标

2.1　双向网络流模型的建立

双向网络流定义为在一段时间内的 2 个源宿节点间具有相同五元组属性信息的所有数据包的集合。实际中在处理数据包时，考虑到硬件性能以及内存等方面的限制，通常需要设置超时机制。

(1)设置最大流闲置时间 τ_{out}。一条流如果在一段时间内没有新的数据包到达，则认为超时，这条流视为结束。后面如果再有新的数据包到达，则认为是一条新流的开始。

(2)设置最大流持续时间 T_{out}。一条流的持续时间如果大于 T_{out} 即认为超时，也会被认为结束。新达到的数据包被视为新流的第一个数据包。

因此，对一条双向流，其中任意 2 个相邻数据包的时间间隔都小于最大流闲置时间 τ_{out}，并且流的持续时长也小于 T_{out}。用集合 $A=\{a_1,a_2,\cdots,a_N\}$ 表示在发送方 S 与接收方 R 之间交互通信的所有 N 个网络数据包。用 $a_i=(t_i,G_i,c_i)$ 表示交互双方之间的第 i 个数据包，其中 t_i 是数据包的达到时间，G_i 是表示数据包的五元组 $(s_{\text{ip},i},s_{\text{port},i},d_{\text{ip},i},d_{\text{port},i},\text{protocol}_i)$，$c_i$ 表示数据包中包含的一些属性特征，如包大小等；用 $\tau_i=t_{i+1}-t_i(1\leqslant i\leqslant N-1)$ 表示相邻数据包的时间间隔。对于集合 A 中一条从 a_i 开始的含有 n 个数据包的双向流，可以表示为 $P=\{a_k\mid i\leqslant k\leqslant i+n-1,\tau_{\text{m}}\leqslant\tau_{\text{out}},i\leqslant m<n,t_n-t_1\leqslant T_{\text{out}}\}$。这种双向流建立方法是考虑交互双方之间数据包的到达顺序，只要相邻数据包的间隔没有超过 τ_{out}，就属于同一条双向流。

在形成双向流之后，将双向流 P 中的第一个数据包 a_1 的方向定义为“前向”，方向从 $(s_{ip,1}, s_{port,1})$ 到 $(d_{ip,1}, d_{port,1})$。一条双向流中所有前向数据包组成的集合称为前向流。类似地，把从 $(d_{ip,1}, d_{port,1})$ 到 $(s_{ip,1}, s_{port,1})$ 的方向定义为“后向”，将所有后向数据包组成的集合称为后向流。如果一条双向流中只包含了一个方向的数据包（即缺少前向数据包或后向数据包），称为伪双向流。

2.2 双向网络流特征指标的建立

双向网络流描述了用户间的交互式行为，由于它包含了 2 个方向的信息，因此可以给流量特性提供更丰富的描述和刻画。

特征指标 1：前后向包数比，定义为前向流中的数据包数与后向流中的数据包数之比。对于伪双向流，该特征指标无意义。

特征指标 2：前后向字节量比，定义为前向流的字节量与后向流的字节量之比。对于伪双向流，该特征指标无意义。

特征指标 1 和 2 其实反映了网络流的对称度，该指标的值越接近 1，则表明网络流的对称性越明显。按照流量采集监控范围，将校园内的学生主机划分为内部主机，其他主机为外部主机。用矩阵 $\boldsymbol{M}$ 来描述内部受监控主机与外部主机之间的网络交互行为，用 $\boldsymbol{M}(t)$ 表示在 t 时刻的采样值。假设在 t 时刻观测到 m 台内部主机和 n 台外部主机，则矩阵 $\boldsymbol{M}$ 可以用下式描述

$$\boldsymbol{M}(t)=\begin{bmatrix}(I_{11},O_{11}) & (I_{12},O_{12}) & \cdots & (I_{1n},O_{1n})\\ (I_{21},O_{21}) & (I_{22},O_{22}) & \cdots & (I_{2n},O_{2n})\\ \vdots & \vdots & & \vdots\\ (I_{m1},O_{m1}) & (I_{m2},O_{m2}) & \cdots & (I_{mn},O_{mn})\end{bmatrix} \tag{1}$$

式中：I_{ij} 表示从外部主机 j 到内部主机 i 的字节量；O_{ij} 表示从内部主机 i 到外部主机 j 的字节量。

在此基础上，我们给出描述内部主机流量特性的 2 个特征指标。

特征指标 3：内部主机的出连接度 $D_{CO,i}$，定义为在时间窗口 T 内，内部主机所连接的不同目的主机的数量，可以用下式描述

$$D_{CO,i}=\|\{O_{ij}\,|O_{ij}\neq 0,\quad 1\leqslant j\leqslant n\}\|,\quad 1\leqslant i\leqslant m \tag{2}$$

式中：i 表示第 i 台内部主机。$D_{CO,i}(t)$ 为 t 时刻的出连接度。

特征指标 4：内部主机的入连接度 $D_{CI,i}$，定义为在时间窗口 T 内，与内部主机连接的外部主机的数量，可以用下式描述

$$D_{CI,i}=\|\{I_{ij}\,|I_{ij}\neq 0,\quad 1\leqslant j\leqslant n\}\|,\quad 1\leqslant i\leqslant m \tag{3}$$

式中：i 表示第 i 台内部主机。$D_{CI,i}(t)$ 表示 t 时刻的入连接度。

3 UDP 流量特性度量分析

按照 1.1 小节中双向流的定义及其组成方法，对数据集 XJTU-2010 的 UDP 流量进行处理，组成双向流。参考 NetFlow 的文档规范[10]，这里取参数 τ_{out} 为 5 min，T_{out} 为 30 min。后面涉及到的流量特征指标分析，时间窗大小 T 的取值设为 1 min。

3.1 UDP 伪双向流特性度量分析

在处理得到 UDP 双向流后，发现有一些双向流其实只在一个方向上有数据包传输，也即伪双向流。经过统计，这些伪双向流在字节量上仅占整体 UDP 流量的 2.3%。选取一个 5 min 的时间段，对校内主机在这 5 min 内产生的伪双向流的出、入连接度的累积分布情况进行测量，结果见图 2，从中可以看到，大部分产生这些伪双向流的主机都具有很小的入连接度和出连接度。

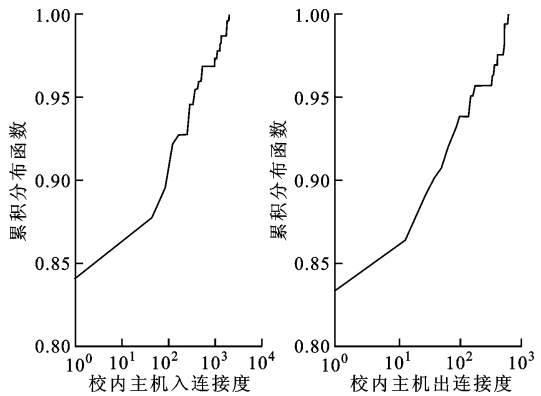


图 2 校内主机出入连接度的累积分布

对于那些出连接度很大的伪双向流，分析发现都是向校外的很多主机发送数据包，但是并没有收到数据包，而且这些主机的通信端口都不同。选取 5 台主机并给出其伪双向流的特征统计结果，见表 2。结合网络监控的先验知识，网络扫描行为往往由 1 台主机对很多台主机的相同端口发送数据包，或者是对 1 台主机的多个不同端口发送数据包。根据表 2 的统计结果可以断定，这些伪双向流是由类似端口扫描等异常行为产生的流量数据。

表 2　部分伪 UDP 双向流特征统计(前向)

IP 地址	伪双向流数	端口数	出连接度	平均包数
IP_1	1 077	256	564	8
IP_2	891	502	536	1
IP_3	936	520	644	2
IP_4	867	252	435	7
IP_5	732	364	386	1

3.2　基于双向流模型的 UDP 流量特性分析

在过滤掉 UDP 伪双向流之后，对剩余的 UDP 双向流进行分析。图 3 显示了在整个采集时间段内，所有受监控在线主机在其端口上的流量统计情况。图中的每个点表示对应主机在对应端口上的流量，点的色度对应着该流量与整体流量的比值，点的大小与流量的大小成正比，即点越大，UDP 流量所占的比例也越大。从图 3 中可以看到，各主机在其端口上产生的 UDP 流量分布并不均匀，绝大部分主机在端口上产生的 UDP 流量都很小，并且这些应用较少使用端口号位于 31000 至 48000 之间的端口。经统计，UDP 流量最大的前 10 台主机占据了 93.2%的总 UDP 流量，并且集中分布在少数端口上，这种特性便于对前向 UDP 流量进行管理。

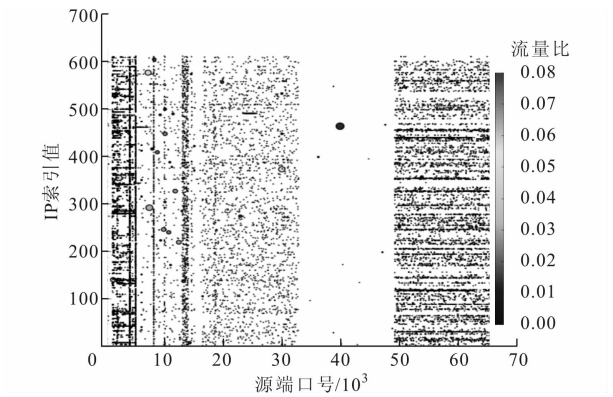


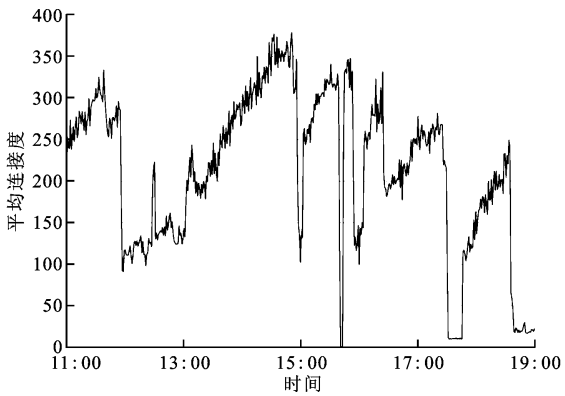
图 3　校内主机的 UDP 流量在端口上的分布

在图 3 中选取 4 台在源端口上流量最大的主机，按流量大小降序排列显示其统计结果，列于表 3。经过分析发现，它们都具有类似的流量特性，如主机连接度较大等。

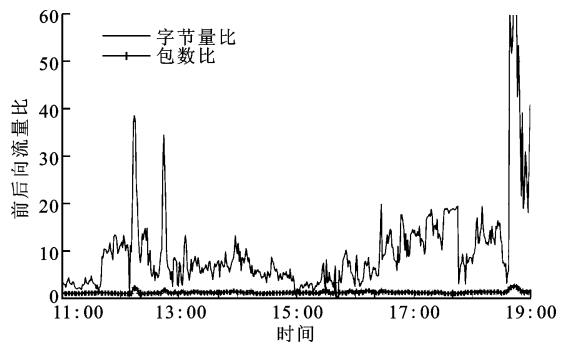
表 3　主机 UDP 源端口号及其流量统计

主机	UDP 源端口号	UDP 流量/GB	UDP 流量与总流量之比/%
H_1	2146	13.17	99.9
H_2	39830	8.75	99.9
H_3	1874	5.88	99.3
H_4	7129	5.39	99.2

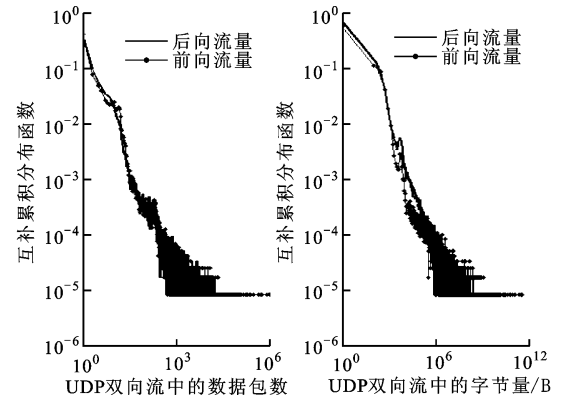
针对这 4 台主要网络主机，首先分析它们的连接度随着时间的变化关系，结果如图 4a 所示，时间窗口取为 1 min。由图中的分析结果可以看出，这些主机具有很大的连接度，表明它们在很短的时间内和数百台主机存在交互行为。进一步分析发现，这部分连接度和字节量在不同目的主机上的分布还是比较稳定的，说明在高连接度的情况下，每个网络连接所占据的流量分布基本稳定，进一步说明这种流量数据是由一种网络应用产生的。图 4b 显示的是这 4 台主机产生的 UDP 双向流的前后向包数比



(a)平均连接度变化



(b)前后向流量比的变化



(c)互补累积分布函数的变化

图 4　4 台校园网络主机的 UDP 流量特征分析结果

以及前后向字节量比的变化情况，时间窗口为 1 min。可以看到，前向 UDP 流量大于后向 UDP 流量，特别是字节量的比值很大，而前、后向包数的比值很小，而且比较稳定，由此也反映出这些 UDP 流量的前向数据包大于后向数据包。

图 4c 给出了这 4 台主机的 UDP 网络流在前、后 2 个方向上的包数和字节量的互补累积分布情况。对图 4c 中的 2 个图进行比较可以看到，在这 2 个方向上 UDP 双向流的包数的互补累积分布函数(CCDF)曲线比较近似，而且大量的 UDP 流的包数都很少，但是在字节量上，这 2 个方向上的 UDP 流大小的分布主要在 1 MB 左右有区别，而且前向 UDP 流的字节量分布在尾部更重些，即前向的网络流大小偏大，这也说明了产生这种网络流的应用的特点——向外部主机提供数据服务。

至此，已经对校园网络 4 台主机所产生的 UDP 流量特征进行了分析，亦即对 90%以上的前向 UDP 流量特性进行了度量。对于单一主机，其流量特征是比较类似的，例如，主机 H_4 在时间段 11:00 到 16:30 内的连接度在 300 到 600 之间变化，前、后向包数的比值稳定在 1 附近，比值较小，而前、后向字节量的比值较大，在 5 到 15 之间变化。根据上述分析，发现产生这些 UDP 流量的网络应用具有如下特征：

(1)主机连接度高，即主机在同一时间连接多个其他主机进行通信，并发连接数很大，并且高连接度分布相对稳定；

(2)扮演了服务器角色，因为主机的前向流量远远高于后向流量，这通常表明主机提供了一些网络服务；

(3)主机的通信端口随机分配，并且不是常见的网络应用端口，符合 P2P 应用端口选取策略。

根据这些 UDP 流量的特点，初步推断它们属于 P2P 流量。为进一步验证这一推断，借助文献[11]中提出的基于传输层行为特征的 P2P 流量识别算法 PTP 来进行流量识别。该算法主要使用了 2 条启发式来检查一条数据流的源和目的之间的流量行为特性。①TCP/UDP 启发式：如果源 IP 和目的 IP 之间同时使用了 TCP 和 UDP 来进行通信，那么它们之间所有除去使用知名端口的数据流都认为是 P2P 流量；②{IP，Port}启发式：如果与某对{IP，Port}进行通信的不同 IP 和端口数量很接近，则认为这对{IP，Port}是属于 P2P 的，所有包含这对{IP，Port}的数据流都被认为是 P2P 流量。同时，为了减小误报率，还使用了一些其他启发式用于过滤如邮件、DNS、Web、网络游戏、端口扫描等流量。

使用该 PTP 算法对整个流量数据集的 TCP 和 UDP 流进行 P2P 流量识别，结果显示有 44.2%的数据流是 P2P 流，它们占据了 29.2%的总流量。此外，还对表 3 列出的 4 台主机在其对应端口上产生的流量统计了其中 P2P 流量所占的比例，结果如表 4 所示，从中可以看到，这几台主机在对应的 UDP 端口上至少有 70%的流量被识别为是 P2P 流量。

我们将在后续工作中对这些 UDP 流量进行深入分析，以识别出这些流量是什么应用产生的，并建立识别模板，定量分析这些业务流量与对称性的关系，从而实现对出口带宽流量的管理和控制，提高带宽利用率。

表 4 主机 UDP 源端口号及其 P2P 流量统计

主机	UDP 源端口号	P2P 流量/GB	P2P 流量与 UDP 流量之比/%
H_1	2146	12.95	98.2
H_2	39830	6.19	70.7
H_3	1874	5.88	99.9
H_4	7129	4.57	84.0

4 总结与展望

本文对真实网络环境下收集到的流量样本进行了度量和分析，发现了非业务性网络流量的对称性特点，并深入分析了对称性产生的原因；提出了双向网络流模型，并以西安交通大学校园网络为例，对校园网络流量进行了分析，发现整体流量在前向和后向上基本相当，具备一定的对称性，同时前、后向的数据传输速率也具有对称性。通过对前、后向流量的统计分析，发现前向 UDP 流量速率远远大于其后向速率，是对称性产生的主要原因。通过细粒度的分析，发现当前 UDP 流量主要由少数校内主机的个别端口所产生，占据了 90%以上的前向 UDP 流量。进一步的实验分析与验证确定了它们主要是由一些典型 P2P 应用产生的。本文的分析结果全面描述了当前由于 P2P 广泛应用所造成的非业务网络流量的新特性及其起因，为合理优化和管理占据大量出口带宽的 UDP 流量奠定了基础。在下一步工作中，将对产生 UDP 流量的应用业务流量进行深入分析和识别，并对 UDP 流量的拥塞控制等进行研究，同时还将针对伪双向流进一步分析其行为特性。

参考文献：

[1] 中国互联网络信息中心. 第 29 次中国互联网络发展状况统计报告 [R/OL]. [2012-07-15]. http://www.cnnic.net.cn/dtygg/dtgg/201201/t20120116_23667.html.

[2] SILVERSTON T, FOURMAUX O, BOTTA A, et al. Traffic analysis of peer-to-peer IPTV communities [J]. Computer Networks, 2009, 53(4): 470-484.

[3] ZHANG Min, JOHN W, CHEN Changjia. Architecture and download behavior of Xunlei: a measurement-based study [C]// Proceedings of International Conference on Education Technology and Computer. Piscataway, NJ, USA: IEEE, 2010: 494-498.

[4] 张艺濒，张志斌，赵咏，等. TCP 与 UDP 网络流量对比分析研究 [J]. 计算机应用研究，2010(6)：2192-2197.
ZHANG Yi-bin, ZHANG Zhi-bin, ZHAO Yong, et al. Comparative analysis on TCP and UDP network traffic [J]. Application Research of Computers, 2010(6): 2192-2197.

[5] ZHANG Min, DUSI M, JOHN W, et al. Analysis of UDP traffic usage on Internet backbone links [C]// Proceedings of the 2009 Ninth Annual International Symposium on Applications and the Internet. Piscataway, NJ, USA: IEEE, 2009: 280-281.

[6] 张敏，陈常嘉. 基于测量的 UDP 流特性分析 [J]. 北京交通大学学报，2010，34(5)：131-134.
ZHANG Min, CHEN Changjia. Measurement-based characterization of UDP flowing [J]. Journal of Beijing Jiaotong University, 2010, 34(5): 131-134.

[7] JOHN W, TAFVELIN S. Differences between in-and outbound Internet backbone traffic [C/OL]// Proceedings of TERENA Networking Conference 2007. [2008-11-19]. http://tnc2007.terena.org/programme/presentations/show4fa2.html.

[8] SHANNON C, ABEN E, CLAFFY K, et al. The CAIDA anonymized 2008 Internet traces [EB/OL]. (2008-05-15) [2011-09-08]. http://www.caida.org/data/passive/passive_2008_dataset.xml.

[9] CLAFFY K, ANDERSEN D, HICK P. The CAIDA anonymized 2011 Internet traces [EB/OL]. (2011-04-13) [2011-06-12]. http://www.caida.org/data/passive/passive_2011_dataset.xml.

[10] Cisco Corp. NetFlow services solutions guide [EB/OL]. [2011-12-01]. http://www.cisco.com/en/US/docs/ios/soluti ons_docs/netflow/nfwhite.html.

[11] KARAGIANNIS T, BROIDO A, FALOUTSOS M, et al. Transport layer identification of P2P traffic [C]// Proceedings of the 4th ACM SIGCOMM Conference on Internet Measurement. New York, USA: ACM, 2004: 121-134.

[本刊相关文献链接]

汪志伟，曹建福，郑辑光. 一种面向分簇无线传感器网络的多信道跨层协议. 2013，47(6)：61-67. [doi:10.7652/xjtuxb201306011]

脱立恒，倪宏，刘学. 一种网络冗余流量消除算法. 2013，47(4)：22-27. [doi:10.7652/xjtuxb201304005]

夏秦，王志文，卢柯. 入侵检测系统利用信息熵检测网络攻击的方法. 2013，47(2)：14-19. [doi:10.7652/xjtuxb201302003]

温彦，刘晨，韩燕波. 一种用户主导的跨组织数据按需集成方法. 2013，47(2)：116-123. [doi:10.7652/xjtuxb201302020]

张赛，徐恪，李海涛. 微博类社交网络中信息传播的测量与分析. 2013，47(2)：124-130. [doi:10.7652/xjtuxb201302021]

陈国强，王宇平. 采用离散粒子群算法的复杂网络重叠社团检测. 2013，47(1)：107-113. [doi:10.7652/xjtuxb201301021]

田丰，桂小林，杨攀，等. 采用类别相似度聚合的关联文本分类方法. 2012，46(12)：6-11. [doi:10.7652/xjtuxb201212002]

伍文，孟相如，马志强，等. 模块化动态博弈的网络可生存性态势跟踪方法. 2012，46(12)：18-23. [doi:10.7652/xjtuxb201212004]

翟治年，奚建清，卢亚辉，等. 任务状态敏感的访问控制模型及其有色网仿真. 2012，46(12)：85-91. [doi:10.7652/xjtuxb201212015]

邵必林，边根庆，张维琪，等. 采用 k-均值聚类算法的资源搜索模型研究. 2012，46(10)：55-59. [doi:10.7652/xjtuxb201210010]

杜文超，陈庶樵，胡宇翔. 面向网络流的自适应正则表达式分组匹配算法. 2012，46(8)：49-53. [doi:10.7652/xjtuxb201208009]

杨柳静，秦涛，王晨旭. 应用交互式网络流模型的高速网络异常行为检测与控制. 2012，46(6)：58-65. [doi:10.7652/xjtuxb201206011]

丁要军，蔡皖东. 采用两阶段策略模型(KTSVM)的 P2P 流量识别方法. 2012，46(2)：45-50. [doi:10.7652/xjtuxb201202008]

侯重远，江汉红，芮万智，等. 工业网络流量异常检测的概率主成分分析法. 2012，46(2)：70-75. [doi:10.7652/xjtuxb201202012]

褚伟波，蔡忠闽，管晓宏，等. 采用子网流量组合优化的网络流量分割方法. 2011，45(12)：22-27. [doi:10.7652/xjtuxb201112005]

(编辑　葛赵青　武红江)