用于高速 IPv6 网络流量抽样测量的算法

潘 乔，裴昌幸

(西安电子科技大学 计算机网络与信息安全教育部重点实验室，陕西 西安 710071)

摘要：为了解决高速 IPv6 网络流量测量，提出了一种基于数据包首部内容分析的流量抽样测量算法. 算法将 IPv6 数据包首部内容进行关键字段的掩码匹配，通过 Hash 映射，利用判断 Hash 值是否属于抽样域来决定数据包的采集与否. 其特点是利用信息熵理论，分析 IPv6 数据包首部，选择出熵值较大的字段，将其作为抽样算法掩码匹配的关键字段，这样就避免了对数据包首部内容的全抽样，在保证抽样样本随机性的前提下，有效地减少了运算量. 实验结果表明，总体流量和抽样样本的数据包大小分布函数曲线十分吻合，验证了该算法的正确性.

关键词：IPv6；抽样测量；分布式测量；信息熵

中图分类号：TP393 **文献标识码**：A **文章编号**：1001-2400(2007)03-0377-05

Method for traffic sampling in high-speed IPv6 networks

PAN Qiao, *PEI Chang-xing*

(Ministry of Edu. Key Lab. of Computer Networks and Information Security, Xidian Univ., Xi'an 710071, China)

Abstract: Traffic sampling techniques are widely used for traffic measurements at a high link speed to prevent an exhaustion of resources and to limit the measurement costs. However, the challenge of an effective sampling method for IPv6-based networks is as yet unmet. This paper proposes a traffic sampling measurement method to take the challenge. For ensuring randomness of sample, we use entropy as an evaluation tool to analyze the bit randomness of each byte in IPv6 packet headers, and conclude that the last one byte of the Payload Length field and byte numbers 8, 12, 14, 15 and 16 of the IPv6 source and destination address fields which have both unchangeability during forwarding and high bit entropy values. We estimate whether a packet is sampled based on a hash function computed over the selected bytes. Therefore, the entire packet header content is not taken into account in our sampling method. The advantages of the method are improved randomness of the sample and the runtime efficiency of the sampling algorithm. Finally, through experiments using real IPv6 traffic traces, we prove that the sampled traffic data can correctly reflect the packet size distribution of full packet trace.

Key Words: IPv6; sampling measurement; distributed measurement; information entropy

IPv6 作为下一代互联网的基本网络协议已逐步进入实用阶段，但是在网络测量研究中，直接用于大规模、高速 IPv6 网络流量抽样采集的方法还没有[1]. 目前的抽样方法如果直接用于下一代的 IPv6 网络流量测量在某些方面存在着一定的局限性. 例如，C. Claffy 等人[2]提出的基于事件触发或时间触发的静态抽样方法和 RFC2330[3]推荐采用的基于泊松间隔的抽样方法，以及 Cisco 公司的 Netflow 采用的"1 out of N"抽样方法，这些方法都只适用于单点测量，而仅仅依靠单点的网络流量测量，因为搜集信息不全面，无法完整地反映由 128 bit 巨大地址空间带来的 IPv6 网络规模大幅增加后的网络特征，如数据包的大小分布、协议分布和流量规律等；文[4]研究的轨迹抽样方法，虽然可用于分布式的多点网络流量测量，但存在两方面的问题：它仅

收稿日期：2007-01-16

基金项目：国家自然科学基金资助项目(60132030，60572147)

作者简介：潘 乔(1977-)，男，西安电子科技大学博士研究生.

研究了对 IPv4 数据包内容进行抽样,由于 IPv6 数据包结构与 IPv4 数据包有着很大的改变,所以它不能直接用于抽样采集 IPv6 网络流量;它对数据包首部中所有不变的部分都进行抽样运算,这样使得计算量较大,数据抽样的速度降低,增加了测量上的开销,无法满足网络数据实时采集的要求.

笔者提出了基于 IPv6 数据包首部分析的抽样算法,该算法首先用掩码匹配选择出 IPv6 数据包首部的关键字段,将其进行 Hash 映射,判断映射出的值是否落入抽样域以决定数据包的采集与否,从而完成 IPv6 流量的抽样测量. 该算法的特点是从 IPv6 数据包首部信息入手,利用信息熵理论对大量的实际 IPv6 网络流量数据进行统计分析,比较 IPv6 数据包首部中各个字段的比特位随机性,选择出随机性好的字段,将其作为掩码匹配选择的关键字段,这样就可以在保证抽样样本的随机性的前提下,不用进行整个数据包首部内容的全抽样,减少了运算量,提高了数据抽样采集的速度. 另外,又由于该算法是基于在传输中数据包首部中内容不会被改变的字段的抽样,所以在网络中的各个测量节点上,只要配置同一抽样算法和同一抽样域,对于相同流量数据在不同的节点测量得到的抽样样本是一样的,利用这个特点,可将该算法用于分布式的网络流量多点测量,为更全面地分析网络性能提供数据.

1 算法描述

算法流程如图 1. 为了便于说明,定义 IPv6 数据包首部 P:$\{0,1\}^l$,关键字段的掩码 Mask_key:$\{0,1\}^m$,其中 l 和 m 分别表示 IPv6 数据包首部和掩码的大小. IPv6 数据包首部的大小是固定的 40 字节,所以掩码的大小也为 40 字节. 算法工作步骤说明如下:

(1) 对输入的 IPv6 流量数据的数据包首部进行掩码匹配选择出关键字段,即 P 与 Mask_key 按位进行逻辑"与"运算,用"&"表示,则有 P_c 为 P & Mask_key. 这样,掩码匹配的结果 P_c:$\{0,1\}^{l_C}$ 中就只保留下了 IPv6 数据包首部中的关键字段,它的大小为 l_c,在下节算法分析中具体讨论关键字段的定义和选择它的原因.

(2) 将 P_c 作为 Hash 映射的输入,得到一个 n 字节的二进制 0,1 序列的集合,即 $\{0,1\}^{l_C} \rightarrow \{0,1\}^n$,映射的结果决定是否抽样采集一个数据包.

(3) 判断 Hash 映射后值 $\{0,1\}^n$,如果落入抽样域 D,即 $\{0,1\}^n \in D$,则抽样该数据包;否则,不抽样.

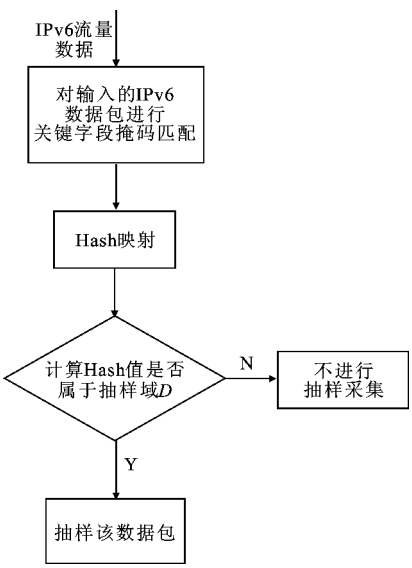

图 1 抽样算法流程图

选择使用 Hash 映射的原因是由于其映射出的值是均匀分布于整个值域上的,例如 Hash 值长度为 n,则 Hash 映射的值域是 $[0,2^n-1]$,对于 IPv6 流量数据中的任一数据包,通过 Hash 映射以后能以相同的概率取 0 到 2^n-1 中的每一个值. 这样抽样域不管落在值域的任何一段上,只要长度相等,就不会影响抽样样本的随机性.

2 算法分析

算法中定义的 IPv6 数据包首部的关键字段需同时满足以下两个条件:

(1) 在数据包传输过程中,包首部中内容不会被改变的字段. 如 IPv4 和 IPv6 数据包首部中的源和目的 IP 地址等,由于 IPv4 和 IPv6 协议的不同,它们的包首部中这样的字段也不尽相同.

(2) 数据包首部中随机程度较大的字段.

条件(1)保证了抽样算法在网络中的各个测量节点上所抽样采集到的数据的一致性,由于该字段是在传输中不会被改变的,所以对于同一网络中的同一流量数据,在节点 A 和节点 B 用掩码匹配选出的关键字段是一样的,Hash 映射后的结果是相同的. 如果采样相同的抽样域,那么抽样样本也将是相同,满足了分布式网络流量测量的要求.

条件(2)可以满足抽样过程的样本随机性. 为了能够正确估计总体信息,抽样算法原理要求样本具有良

好的随机性,样本的随机程度越大,对总体的估计就越准确. 由此把关键字段定义为数据包首部中那些随机性高的字段,这样就提高了抽样样本的随机性.

在信息论中,"熵"是随机变量不确定度的度量[4],所以可以用熵来评测数据包的随机程度,用这一概念来研究 IPv6 数据包首部中各个比特位的随机性. 由于 IPv6 数据包首部是由 $\{0,1\}^{l_H}$ 组成的,其中 l_H 是包首部的大小,所以其比特位 b 只有 0 或 1 两种取值的可能性,令 $b=\begin{cases}1, & \text{概率为 } p\\ 0, & \text{概率为 } 1-p\end{cases}$,则比特位 b 的熵 $H(b)$ 为

$$H(b) = -p\log_2 p-(1-p)\log_2(1-p) \quad . \tag{1}$$

熵 $H(b)$ 的值越大,其变量的随机性就越大,由式(1)可见当 $p=0$ 或 1 时,$H(b)=0$ 表示变量恒为 1 或 0,从而不再是随机的;而当 $p=1/2$ 时,$H(b)=1$ 熵值最大,表示变量的随机度也达到最大. 这样,由条件(1)和条件(2)定义的关键字段,对应于 IPv6 数据包首部中的比特位,就可描述为在传输过程中不被改变且熵值最大的比特位.

下面基于对 IPv6 数据包首部的分析来讨论如何选择关键字段. 参考 RFC2460[6] IPv6 协议规范中 IPv6 包首部的定义,如图 2,利用来自于广域集成分布环境项目(WIDE)骨干网络 6bone 的实际 IPv6 网络流量数据[7],对从 2005 年 9 月 1 日至 9 月 30 日共 30 天,总计 60,000,000 个 IPv6 数据包首部计算了它们比特位的熵,分析如下:

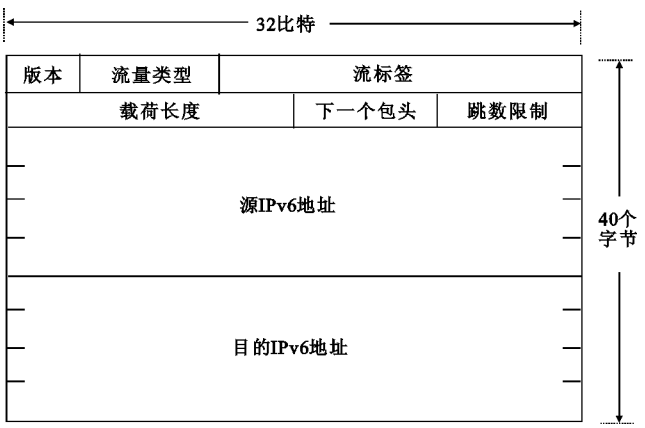

图 2　IPv6 数据包首部格式

(1) 版本字段(4 比特),记录 IP 版本号. 因为这里的数据都是 IPv6,所以该字段的值总是为 6,是确定值,所以其熵为 0.

(2) 流量类型字段(8 比特),该字段其功能类似于 IPv4 包首部的服务类型字段,在 IPv6 中主要是区分业务编码点(DSCP),标记一个 IPv6 数据包,以此指明数据包应该如何处理. 由于目前在请求注释文档(RFC)中并没有对该字段明确定义,在实际 IPv6 流量数据中,也仅有 0.62%的数据包设置了该字段,其他大部分都默认为 0,所以它的各个比特位熵值近似可以看作 0.

(3) 流标签字段(20 比特),用来标识 IPv6 数据包的一个流. 由于该项是 IPv6 协议中新增的. 目前,大部分数据包不属于特定的流,如 SMTP,FTP 以及 WWW 这些原本为 IPv4 而设计的,本身就没有流标签字段,不能处理,当前互联网工程任务组(IETF)的标准也没有详细说明怎样管理和处理这个标签,所以多数数据包把它默认为 0. 通过实际的数据分析,发现只有 5.09%的数据设置了该位,它的各个比特位熵值也可以近似看作 0.

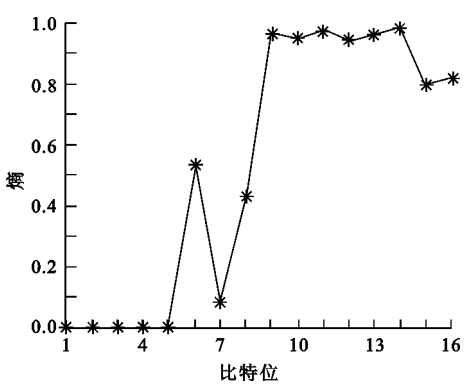

图 3　载荷长度字段的各比特熵

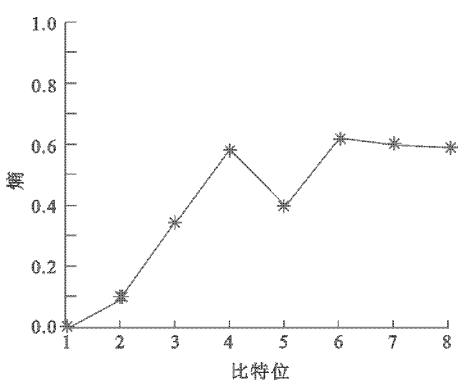

图 4　下一个包首部字段的各比特熵

(4) 载荷长度字段(16 比特),表示 IPv6 包首部后数据包其余部分的长度. 因为 IPv6 中分段只能在源节点进行,该数据包所经过的中间路由器不能再进行任何分段,所以与 IPv4 中长度字段在传输中可能会被改变不同的是:IPv6 的载荷长度字段在传输过程中是不变的. 计算它的各个比特熵,如图 3. 可以看出,载荷长度的后 8 比特位的熵较大,其比特位熵的平均为0. 922 5比特,随机性较强,考虑选作关键字段.

(5) 下一个包首部字段(8 比特),指在 IPv6 包首部后所跟的字段的协议类型. 根据数据计算它的各比特位熵值,如图 4. 由于协议类型有限,对 IPv6 数据统计表明 TCP 数据占总流量的 82.16%,UDP 数据占总流量 12.15%,ICMPv6 占 4.71%,其他的仅占 0.98%,所以该字段的随机性不强,计算出的其各个比特位熵的平均也仅为 0.3970 比特,虽然它在传输过程中不被改变,但不作为关键字段考虑.

(6) 跳数限制字段(8 比特),每当一个节点对包进行一次转发之后,该字段就会被减 1. 因为它在数据包传输的过程中会改变,所以不能作为关键字段.

(7) 源地址(128 比特)和目的地址字段(128 比特),IPv6 与 IPv4 地址最大的差别在于地址空间的长度:IPv6 地址 128 比特的长度使它拥有了更大的地址空间. 由于地址空间很大,实际中大部分 IPv6 的地址空间还没有被分配. 目前用的最多的就是可聚合全球单播地址,它占了 12.5%的 IPv6 的地址空间,互联网地址授权委员会(IANA)分配的公用地址的前缀仅有 2001::/16,2002::/16 和 3FFE::/16. 对目前 IPv6 实际网络地址数据统计,分析的数据中也以这 3 个地址前缀为主. 分别计算源地址和目的地址的各个位的熵,如图 5. 可以看出源地址、目的地址的第 8 字节和后半部分第 12,14,15,16 字节的熵值较大,且变化幅度不大,通过计算它们熵值的均值,如图 6,这些字节的熵的均值也是最大的,所以选其作为抽样的掩码匹配关键字段.

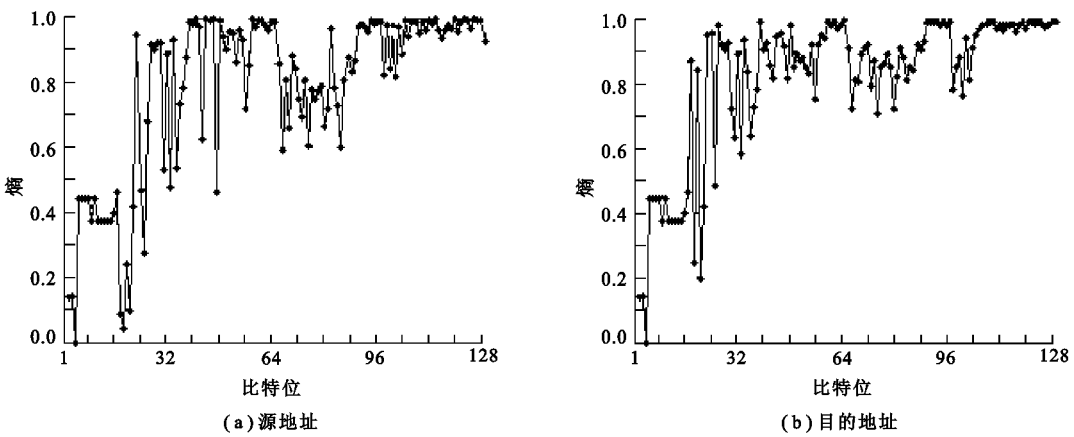

图 5 源地址和目的地址的各比特熵

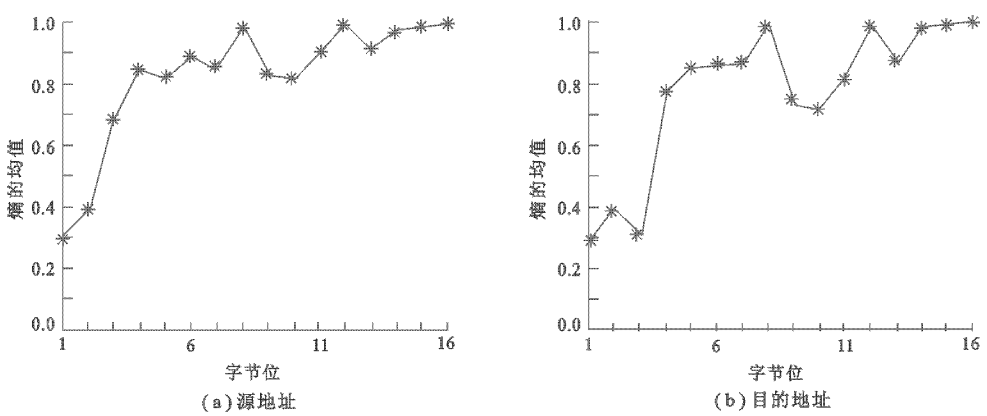

图 6 源地址和目的地址各比特熵的均值分布

综合上述分析,选定 IPv6 数据包首部中比特位熵较大的载荷长度的后 1 字节,源地址和目的地址的第 8,12,14,15 和 16 字节这 6 个字节作为抽样算法中的关键字段.

3 实验结果与分析

为了验证文中提出的抽样算法的正确性,软件编程实现了该算法,采用上文所提的 WIDE 骨干网络

6bone 的 IPv6 数据,用文中的算法做该流量的抽样采集实验. 在实验中,将 IPv6 数据包首部中的载荷长度后 1 字节,源地址和目的地址的第 8,12,14,15 和 16 字节,共 6 个字节(48 比特)作为关键字段进行掩码匹配. 哈希函数采用文[8]的 Bob 算法,它是 CRC32 算法的一种改进算法,其输入以字节为单位且可变长度,输出为 32 比特的值,则哈希函数的值域为$[0, 2^{32}-1]$,所以抽样比率就在$[1/2^{32}, 1]$之间.

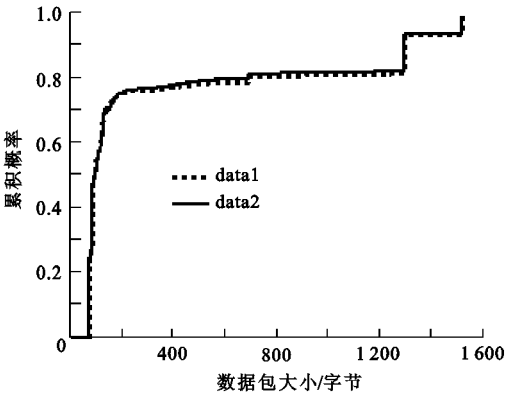

图 7　数据包大小累积分布函数

数据包长度的分布特性是网络流量测量研究的重要对象,通过分析实验结果中总体和抽样样本数据包长度的分布情况来验证抽样方法的正确性. 设$[t,t+T]$时段内的总体包长度序列为$L_{\text{full}}=\{l_1,l_2,\cdots,l_m\}$,抽样样本包长度序列为$L_{\text{smp}}=\{l'_1,l'_2,\cdots,l'_n\}$,$m\gg n\gg 1$. 将数据包长度序列 L_{full} 和 L_{smp} 看作是随机变量,以数据包的长度字节为间隔单位进行频数统计,分别计算出其累积分布函数值,进行拟合后如图 7. 图中 data1 和 data2 分别为总体包序列和抽样样本包序列的累积分布函数值,抽样比率为1/1 000. 总体包序列平均长度为 355.37 字节,标准方差为 497.20 字节;抽样样本包序列平均长度为 358.55 字节,标准方差为 500.23 字节,由图 7 可以看出 data1 和 data2 的数据包长度累积分布特性曲线非常吻合,所以用抽样样本可以反映总体数据包大小的分布情况.

4　结 束 语

提出了一种新的基于 IPv6 数据包首部分析的抽样算法. 利用信息熵理论,对 IPv6 数据包首部中的各个字段比特位的熵进行统计比较,计算表明包首部中的载荷长度后 1 字节,源地址和目的地址的第 8,12,14,15 和 16 字节的随机性最强. 将这 6 个字节作为关键字段掩码匹配后进行 Hash 映射,通过判断 Hash 映射后的值是否属于抽样域,来完成 IPv6 网络流量的抽样采集. 该算法是基于在传输中数据包首部内容不会被改变的字段的抽样,所以在网络中的各个测量节点上,只要配置同一抽样算法和同一抽样域,就可以用于分布式的大规模、高速 IPv6 网络流量的抽样采集,它对于相同流量数据在不同的节点测量得到的抽样样本是一样的,满足了分布式网络性能分析的要求. 最后对实际 IPv6 网络流量做了抽样采集实验,分析了实验结果中总体流量和样本的数据包大小的累积分布函数曲线,它具有非常好的吻合度,从而证明了该算法的正确性.

参考文献:

[1] Zhu Changhua, Pei Changxing, Li Jiandong, et al. Network Measurement and Its Key Technologies[J]. Journal of Xidian University, 2002, 29(6): 813-819.

[2] Claffy K C, Polyzos G C, Braun H W. Application of Sampling Methodologies to Network Traffic Characterization[J]. ACM SIGCOMM Computer Communication Review, 1993, 23 (4): 194-203.

[3] Paxson V, Almes G, Mahdavi J, et al. RFC2330[EB/OL]. [2005-01-18]. http://www.faqs.org/rfrs/rfc2330.html.

[4] Duffield N G. Sampling for Passive Internet Measurement: a Review[J]. Statistical Science, 2004, 19(3): 472-498.

[5] 程光, 龚俭, 丁伟. 基于分组标识的网络流量抽样测量模型[J]. 电子学报, 2002, 30(12A): 1986-1990.

[6] Deering S, Hinden R. RFC2460[EB/OL]. [2005-03-10]. http://www.faqs.org/rfrs/rfc2460.html.

[7] WIDE 6bone. Daily Trace at an WIDE 6bone Line[EB/OL]. [2006-06-06]. http://tracer.csl.sony.co.jp/mawi.

[8] Zseby T, Fokus F, Molina M, et al. Sampling and Filtering Techniques for IP Packet Selection[EB/OL]. [2005-07-19]. http://www.ietf.org/html.charters/psamp-charter.html.

(编辑: 齐淑娟)