面向 IP 流测量的哈希算法研究*

程 光[1,2+], 龚 俭[1,2], 丁 伟[1,2], 徐加羚[1,2]

[1](东南大学 计算机科学与工程系,江苏 南京 210096)

[2](江苏省计算机网络重点实验室,江苏 南京 210096)

A Hash Algorithm for IP Flow Measurement

CHENG Guang[1,2+], GONG Jian[1,2], DING Wei[1,2], XU Jia-Ling[1,2]

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

[2](Jiangsu Provincial Key Laboratory of Computer Network Technology, Nanjing 210096, China)

+ Corresponding author: Phn: +86-25-83794000 ext 213, E-mail: gcheng@njnet.edu.cn, http://www.njnet.edu.cn

Received 2004-04-14; Accepted 2004-11-22

Cheng G, Gong J, Ding W, Xu JL. A hash algorithm for IP flow measurement. *Journal of Software*, 2005,16(5):652−658. DOI: 10.1360/jos160652

Abstract: In order to solve the problems with computing resource and high-speed network traffic, it is necessary to deal with the network traffic by some measuring technologies, such as sampling measurement and load balance, etc, while the hash algorithm is one of the key measuring technologies. In this paper, firstly, a random metric is provided to evaluate the performance of the hash algorithms. Secondly, the randomicity of XOR and shift operations are analyzed, and it is proved that the two operations can improve the bit randomicity. Thirdly, this paper analyzes the four fields of IP packet, such as source IP, destination IP, source port, and destination port, and a hash algorithm named XOR_SHIFT is provided based on the analysis. Finally, using the CERNET backbone traffic and PMA traffic, this paper analyzes the character of the XOR_SHIFT hash algorithm and compares with the performance among XOR_SHIFT, IPSX and CRC32 hash algorithms. This study shows that the XOR_SHIFT hash function provided in this paper has two advantages: algorithm performance and hash randomicity, and it can be applied to measure the high-speed network traffic.

Key words: hash algorithm; network traffic; XOR; shift; traffic measurement

摘 要: 为了解决计算资源和高速网络流量之间的矛盾,需要对 IP 流进行抽样或负载均衡等处理,而哈希算法是资源代价的核心.首先提出评价哈希算法性能的随机测度;其次从理论上证明比特之间异或运算和位移运算能够提高哈希值的随机特性,提出比特流之间哈希算法的原则;然后分析 IP 报文的 4 个字段:源 IP、宿 IP、源端

* Supported by the National Natural Science Foundation of China under Grant No.90104031 (国家自然科学基金); the National Grand Fundamental Research 973 Program of China under Grant No.2003CB314803 (国家重点基础研究发展规划(973)); the Foundation of Southeast University of China under Grant No.9209002157 (东南大学基金)

作者简介: 程光(1973−),男,安徽黄山人,博士,讲师,主要研究领域为网络行为学;龚俭(1957−),男,博士,教授,博士生导师,主要研究领域为网络行为学,网络安全;丁伟(1962−),女,教授,博士生导师,主要研究领域为网络行为学;徐加羚(1979−),男,助教,主要研究领域为网络测量.

口和宿端口的特性,由此提出相关的哈希算法;最后使用 CERNET 主干流量和 PMA 的数据验证算法的性能,并与 IPSX 和 CRC32 算法进行比较.研究表明,基于异或、位移原则的比特流哈希算法的执行效率和哈希值的均匀性两方面具有较好的性质,能够满足高速网络流量测量需求.

关键词: 哈希算法;网络流量;异或;位移;流量测量
中图法分类号: TP393 文献标识码: A

IP 流测量是网络管理和网络流量工程技术研究和发展的重要依托,为此,IETF 组织已建立了 IP 流信息输出工作组 IPFIX[1].但是网络主干流量增长迅速,1997 年,MCI 网络中的并发流数量大约为 25 万条[2],2003 年 CERNET OC48 主干网络流量并发流数达到 300 万条.由此可见,互联网发展“更大、更快”的趋势将更进一步对网络测量产生严重影响.“更大”是指互联网络将采用 IPv6 为基本网络层协议,网络用户的增加和网络服务的多样化;“更快”是指互联网主干带宽达到 OC48(2.5Gbit)甚至 OC192(10Gbit),互联网用户的端到端带宽将达到 100Mbps 以上.为了解决“更大、更快”网络 IP 流测量,Cisco 路由器提供 NetFlow[3]测量 IP 流,并建议网络速度超过 OC-3 时使用抽样测量技术:即在路由器内存中维护抽样报文的流记录.

哈希值的随机性是基于流的抽样技术的关键,文献[4,5]分析了 5 种基于 IP 流字段的哈希算法,可以分为两类方法:一类[6]是直接使用报文中的标识字段,这种方法效率高,但是难以确保足够的哈希值随机性;另一类[7]是使用哈希函数计算哈希值,哈希函数选择是保证哈希值随机性和算法高效性的关键因素,文献[4,5]的实验结果表明,异或运算具有较高的哈希性能,但是没有从理论上证明.本文首先提出评价哈希算法性能的随机测度;其次从理论上证明比特之间异或运算和位移运算能够提高哈希值的随机特性,提出比特流之间哈希算法的原则;然后分析 IP 报文的 4 个字段:源 IP、宿 IP、源端口和宿端口的特性,由此提出相关的哈希算法;最后使用 CERNET 主干流量数据验证算法的性能,并与 IPSX 和 CRC32 算法进行比较.文中使用的数据来自于 CERNET 主干网络流量数据和 NLNAR PMA 的网络流量数据.文章研究表明,基于异或、位移原则的比特流哈希算法的执行效率和哈希值的均匀性两方面具有较好的性质,能够满足高速网络流量测量需求.

1 随机测度定义

传统上采用五元组定义流:协议、源 IP(32bit)(SIP)、宿 IP(32bit)(DIP)、源端口(16bit)(SPORT)、宿端口(16bit)(DPORT),由于网络中 90%以上的流量为 TCP 流量,协议字段中信息量很小,因此本文不考虑将协议字段作为哈希算法的输入参数.两个 IP 字段分成前 16bit 源 IP(bsip)、后 16bit 源 IP(asip)、前 16bit 宿 IP(bdip)和后 16bit 宿 IP(adip),研究一个哈希算法 H,哈希算法以 16bit 串的六元组为输入,一个 16bit 串的哈希值(hashID)为输出,16bit 的伪随机哈希值可以保证抽样精度达到 $1/2^{16}$.我们首先定义比特随机性的测度,以评价流字段各个比特的随机性;其次定义评价比特流的随机测度.

定义 **1**(比特随机测度 ***E***). 定义为位熵 $H(b)=-(p\log_2 p+(1-p)\log_2(1-p))$ 和最大位熵 $H_{\max}(b)=1$ 的比值,表示比特随机程度的测度, $E=H(b)/H_{\max}(b)=H(b)$.由定义可知, $0\le E\le 1$,E 表示随机程度,E 越接近 1,表示随机性越强,即位熵越大;E 接近 0,表示确定性信息越大,位熵越小.

定义 **2**(位流随机测度 ***E***). 定义位流熵 $H(s)=-\sum_{i=0}^{2^s-1}p_i\log_2 p_i$ 和最大位流熵 $H_{\max}(s)$的比值,表示比特流随机性程度, $E=H(s)/H_{\max}(s)=H(s)/s$.

定义 **3**(流标记 **FlowID**). 每个报文具有一个四元组{源 IP、宿 IP、源端口、宿端口},将 IP 字段分成前 16bit 串和后 16bit 串,因此流标记定义为 FlowID={源 IP 前 16bit(bsip)、源 IP 后 16bit(asip)、宿 IP 前 16bit(bdip)、宿 IP 后 16bit(adip)、源端口(sport)、宿端口(dport)}.我们重点探讨和研究这 6 个 16bit 串为输入的哈希算法.

2 比特随机运算的分析

哈希算法需满足的两点要求:(1) 生成的哈希序列随机性尽可能大;(2) 算法简单、高效.本节将首先分析两

个比特随机运算的随机性质,然后分析异或运算的性质并证明相关的定理.

2.1 二元比特运算分析

首先分析所有可能的二元比特运算,两个比特之间可能的运算有 $2^2\times 2^2=16$ 种,见表 1.对于比特运算而言,运算结果受制于真值表,实际上运算的种类因此是有限的.可以通过分析两个比特之间的二元运算,归纳获得对多比特运算有参考意义的结论.

Table 1 Truth table of binary bit-wise operation

表 1 二元比特运算真值表

0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1
0	1	0	0	0	0	1	1	1	1	0	0	0	0	1	1	1	1
1	0	0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1
1	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1
A	B	0	&		A		B	$\oplus$	|		$\overline{\oplus}$	$\overline{B}$		$\overline{A}$			1

显然,对于恒 0 运算和恒 1 运算生成比特的位熵为 0,无法保持良好的随机性,因此不能使用.对于产生结果中 0 和 1 的比例为 3:1 或 1:3 的 8 种运算($C_4^1+C_4^3$),如果 A 和 B 的位熵均较高,那么在 A 和 B 的相关性低(或无相关性)的情况下,生成出的结果中 0 和 1 的比例必然差异很大,所以位熵不会很高.当然,可以通过选择两个位熵很低的比特 A 和 B 以期获得好的结果位熵.例如 A 和 B 都是非常可能出现 1 的,那么与(&)操作后的位熵反而可能得到很大改善.但由于选择这样位熵特性的比特并不直观,所以难以确定这两个比特所应该具备的特性.因此,一般可以选择的操作为 0 和 1 的比例为 2:2 的 6 种(C_4^2)运算,对于 A, $\overline{A}$ 操作,其位熵等于 A 的位熵.同理,B 和 $\overline{B}$ 操作的位熵等于 B 的位熵.因此仅需分析异或运算和同或运算,且发现这两种运算的位熵相同.

2.2 异或运算分析

2.2.1 异或运算的随机性分析

定理 1. 两个独立不相关的离散比特随机变量ξ和η之间的异或运算所得的比特随机变量$\xi,\zeta=\xi\oplus\eta$,其位熵 $H(\zeta)\geq\max(H(\xi),H(\eta))$.

证明:设随机变量ξ出现 0 的概率为 p,出现 1 的概率为 1–p;随机变量ξ出现 0 的概率为 q,出现 1 的概率为 1–q.因此随机变量ξ出现 0 的概率为 $p_\xi^0=p\cdot q+(1-p)\cdot(1-q)$,出现 1 的概率为 $p_\xi^1=1-p_\xi^0=p+q-2p\cdot q$.位熵的大小可以由 0,1 概率和最大位熵概率 0.5 之间的距离决定,如果距离 0.5 越近,则随机变量ξ的位熵越大,反之越小.因此在证明定理 1 时,使用最大位熵概率 0.5 之间的距离为比较位熵大小测度.随机变量ξ距离概率中心的距离为 $D(\xi)=2\cdot(0.5-p)^2$,随机变量ζ距离概率中心的距离为 $D(\zeta)=2(0.5+2pq-p-q)^2$,因此证明 $H(\zeta)\geq H(\xi)$ 就等同于证明 $D(\zeta)\leq D(\xi)$,即证明 $2(0.5-p)^2-2(0.5+2pq-p-q)^2\geq 0$ (0<=p,q<=1)即 $D(\zeta)\leq D(\xi)$,因此 $H(\zeta)\geq H(\xi)$,同理可以证得 $H(\zeta)\geq H(\eta)$.故定理 1 成立. □

定理 1 表明,相互独立的两个比特随机事件之间异或可以增加结果的随机性.如果网络流量各比特之间存在联系,对于两个比特随机变量ξ和η之间在某种程度上也是相互联系的,即它们之间存在统计依赖关系.因此得知——随机变量ξ取值条件下的条件位熵 $H(\eta|\xi)$总是不大于另一个随机变量η的无条件位熵 $H(\eta)$.因此,$H(\eta|\xi)$表示了已知ξ后η残留的不确定度,如果已知ξ,η的不确定度的减少量为 $H(\eta)$–$H(\eta|\xi)$.同样,如果已知η,ξ的随机程度的减少量为 $H(\xi)$–$H(\xi|\eta)$.如果两个比特随机变量ξ和η之间相关性很大,则获得的位熵特性将无法保证.若随机变量ξ={0,1,0,1,0,1},而随机变量η={0,1,0,1,0,1},虽然随机变量ξ和η的位熵 $H(\xi)$=1,$H(\eta)$=1,但是两者之间的异或 $\zeta=\xi\oplus\eta$ ={0,0,0,0,0,0,0,0},其位熵 $H(\zeta)$=0.因此两个随机变量由于相关性为 1,其异或运算以后位熵小于计算之前的位熵.下面需要研究两个随机变量相关性达到多少以后,异或运算的位熵开始减少.

定义 4. 随机变量ξ和η之间的相关系数定义为

$$\rho=\mathrm{cov}(\xi,\eta)/\sqrt{\mathrm{cov}(\xi,\xi)\mathrm{cov}(\eta,\eta)}=(E(\xi\eta)-pq)/\sqrt{pq(1-p)(1-q)} \tag{1}$$

相关系数ρ是刻画随机变量ξ和η之间相依性的数字特征,相关系数等于 0 时表示不相关,而接近 1 时表示线性相关.

定理 2. 对于服从 0,1 分布的比特随机变量ξ和η,随机变量ξ中 1 事件出现的概率为 p,η中 1 事件出现的概

率为 $q,E(\xi)=p,E(\eta)=q$.异或后的随机变量 $\zeta=\xi\oplus\eta$ 位熵 $H(\zeta)>=\max(H(\xi),H(\eta))$的条件是随机变量$\xi,\eta$之间的相关系数 $\rho=\mathrm{cov}(\xi,\eta)/\sqrt{\mathrm{cov}(\xi,\xi)\mathrm{cov}(\eta,\eta)}=(E(\xi\eta)-pq)/\sqrt{pq(1-p)(1-q)}$.

$$|\rho|\le\min\left(\left|\frac{2p+q-1-2pq}{2\sqrt{pq(1-p)(1-q)}}\right|,\left|\frac{q-2pq}{2\sqrt{pq(1-p)(1-q)}}\right|\right) \tag{2}$$

证明:对于服从 0,1 分布的比特随机变量ξ和η,随机变量ξ中 1 事件出现的概率为 p,η中 1 事件出现的概率为 $q,E(\xi)=p,E(\eta)=q$.其协方差: $\mathrm{cov}(\xi,\eta)=E[(\xi-E\xi)(\eta-E\eta)]=E(\xi\eta)-pq$, $\mathrm{cov}(\xi,\xi)=E(\xi\xi)-E(\xi)E(\xi)=p(1-p)$,相应地, $\mathrm{cov}(\eta,\eta)=q(1-q)$.ξ和 η的相关系数为 $\rho=cov(\xi,\eta)/\sqrt{cov(\xi,\xi)cov(\eta,\eta)}=(E(\xi\eta)-pq)/\sqrt{pq(1-p)(1-q)}$, $E(\xi\eta)=\rho\sqrt{pq(1-p)(1-q)}+pq$.随机变量(ξ,η)出现的事件有 4 种(00,01,10,11),设每种事件出现的概率分别为 $p_{00},p_{01},p_{10},p_{11}$,这 4 个概率之间的关系是: $p_{01}=q-p_{11}$, $p_{10}=p-p_{11}$, $p_{00}=1-p-q+p_{11}$,由于 0×0=0×1=1×0=0, 1×1=1,因此 $E(\xi\eta)=p_{11}$. $\zeta=\xi\oplus\eta$ 随机变量的 0 事件出现的概率 $p_0=p_{00}+p_{11}=1-p-q+2p_{11}$, $p_1=p_{01}+p_{10}=p+q-2pq-2\rho\sqrt{pq(1-p)(1-q)}$.设随机变量ξ,η对应的位熵分别为 $H(\xi),H(\eta)$,设 $H(\xi)>=H(\eta)$,比特事件 0,1 出现的概率距 0.5 中心的距离可以认为是位熵的大小,同时可知 0,1 出现的概率以 0.5 为中心左右对称.因此 0,1 概率之间的距离反映位熵的大小,两者之间的距离越近,则位熵越大,否则位熵越小.随机变量ξ的 0,1 概率之间的间距为$|2p-1|$,随机变量η的 0,1 概率之间的距离为$|2q-1|$,如果 $H(\xi)>=H(\eta)$,则$|2p-1|<=|2q-1|$.

需要证明在相关系数ρ到达多少时,其位熵 $H(\zeta)=H(\xi\oplus\eta)=\max(H(\xi),H(\zeta))=H(\xi)$,即证明

$$|2p-1|=|p_0-p_1|=|1+4pq+4\rho\sqrt{pq(1-p)(1-q)}-2p-2q|.$$

$$\rho_1=\frac{2p+q-1-2pq}{2\sqrt{pq(1-p)(1-q)}},$$

$$\rho_2=\frac{q-2pq}{2\sqrt{pq(1-p)(1-q)}}.$$

因此相关系数满足下列条件

$$|\rho|\le\min\left(\left|\frac{2p+q-1-2pq}{2\sqrt{pq(1-p)(1-q)}}\right|,\left|\frac{q-2pq}{2\sqrt{pq(1-p)(1-q)}}\right|\right).$$

异或后的位熵会增加,证明完毕. □

对 CERNET 数据分析表明,FlowID 6 个字段的比特之间均满足定理 2 的条件.如图 1 所示,是源宿 IP 第 32bit 之间的相关系数ρ,从图 1 可以看出,$|\rho|$曲线在$|\rho_1|,|\rho_2|$下方,满足定理 2 的条件,所以两个比特之间进行异或计算会增加随机性.6 个流字段之间进行异或计算可以增加伪随机序列的随机测度值,因此在基于异或的哈希函数不变的情况下,将全部的 6 个流字段作为哈希函数的输入,所得到的随机序列的随机测度值最大.

根据异或运算的结合率和交换率两个性质可以容易得出,流字段之间以不同的顺序进行异或运算不会改变伪随机序列的随机测度值.因此异或哈希算法可以不用考虑不同字段之间的运算次序.

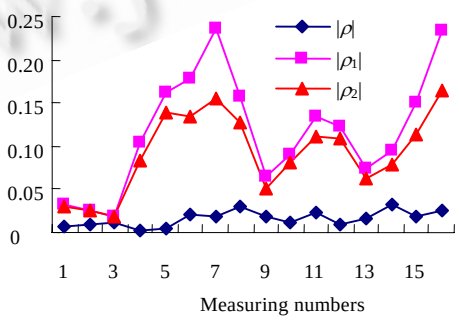

Fig.1 Correlation coefficient of the 32th bit in source IP and destination IP

图 1 源宿 IP 第 32 比特的相关系数

3 位移运算分析

本文第 2 节从理论上分析的异或运算可以增加伪随机序列的随机测度值,下面将研究比特位移对伪随机序列随机测度值的影响.继续使用前文介绍的 CERNET 主干流量数据 asip∧adip,bdip∧dport 和 bsip∧sport 这 3 种组合的随机测度值进行分析.图 2 是这 3 种组合位移和随机测度之间的关系图,从图中可以明显看出,流字段位移 2~6bit 异或生成的伪随机序列的随机测度值大于不位移或位移其他的值.图 3 是 asip∧adip,bdip∧dport 和 bsip∧sport 三者之间异或后生成伪随机序列的随机测度值,其中 shifti 表示 asip,bdip 和 bsip 进行循环位移 *i* 个比特,其中 shift0 表示不进行循环位移,分不同的时间总共测量 20 次,每次连续测量 200 000 条流.从图中可以看出,循环位移 1~6bit 的随机测度值大于循环位移为 0 的随机测度值,这 6 组平均值的差别最大不超过 0.000 4,其中平均随机测度值最大的是 shift3.

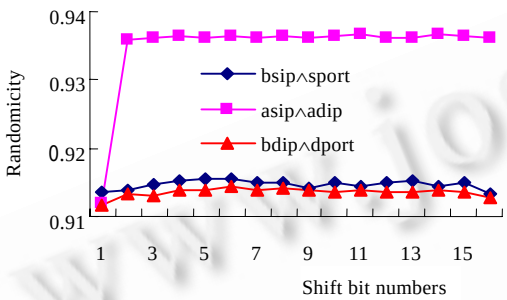

Fig.2 Shift and randomicity
图 2 位移和随机测度之间的关系

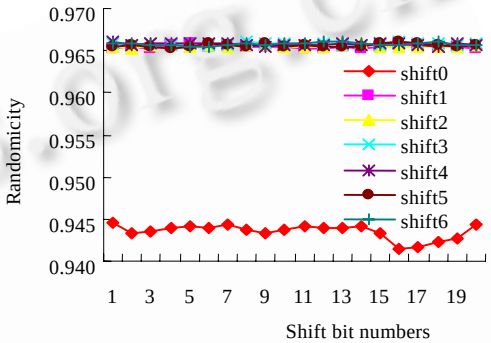

Fig.3 Shift and randomicity of FlowID
图 3 FlowID 的位移数和随机测度值

4 哈希算法及其性能分析

根据前文对 IP 流字段异或、位移分析和设计原则,我们提出了一种基于异或、位移算法,同时给出 IETF 工作组 PSAMP 提出的 IPSX 和 CRC32 算法,然后对这 3 种算法的随机特性进行比较.设 bsip 为源 IP 前 16bit、asip 为源 IP 后 16bit、bdip 为宿 IP 前 16bit、adip 为宿 IP 后 16bit、sport 为源端口、dport 为宿端口;算法使用 XOR(∧),右移(>>)和左移(<<)操作;hash1,hash 为 16bit 无符号整型.

4.1 异或位移哈希算法

从上文分析可知,循环位移 3bit 的平均随机测度值最大,因此下面给出循环位移 3bit 的异或、位移哈希算法(XOR_SHIFT).

```
hash1=asip<<3|asip>>(16-3);   hash1=hash1∧adip;   hash=hash1;
hash1=bsip<<3|bsip>>(16-3);   hash1=hash1∧sport;  hash∧=hash1;
hash1=bdip<<3|bdip>>(16-3);   hash1=hash1∧dport;  hash∧=hash1.
```

其中 hash 是异或、位移算法返回的哈希序列.

4.2 IPSX哈希算法

*f*1=IP 源地址,*f*2=IP 宿地址,*f*3=TCP 或 UDP 报文的源端口和宿端口.中间变量 *h*1,*v*1,*v*2 均是 32 比特串.算法的输出是 *h*1 的后 16bit 串.

```
v1=f1∧f2;     v2=f3;          h1=v1<<8;     h1∧=v1>>4;     h1∧=v1>>12;
h1∧=v1>>16;   h1∧=v2<<6;      h1∧=v2<<10;   h1∧=v2<<14;    h1∧=v2>>7.
```

4.3 CRC32哈希算法

CRC32 哈希算法具有较小的冲突性,它的输出是一个具有足够随机化的 32bit 串,但是该算法的执行性能很低,对于高速网络主干 IP 流选择,这种算法不是很合适.

4.4 3种算法比较

IP 流四元组随机性分析采用两组数据:CERNET 主干和美国 NLANR 的 PMA 数据[8].CERNET 数据是 2004 年 4 月 17 日从 0:00 至 24:00,在我国东北地区网络中心对 CERNET 主干网络主干测量流量,连续测量 24 小时数据,每个报文截取报文头 44 字节加上 8 字节的时戳,总共测量到 18G 报文,本文中使用的 CERNET 数据全部来自于这组数据源.第 2 组数据来源是美国应用网络研究国家实验室(NLANR)的被动测量和分析工作组(PMA),PMA 在 HPC 网络中设置多个测量点被动测量 Internet 数据,其中本文使用从节点 AIX 和 TXS 于 2002 年 9 月 30 日测量的报文,其数据格式为 TSH,每个报文 44 字节长度,TSH 数据格式包含整个 IP 报头和 TCP 报头的内容.

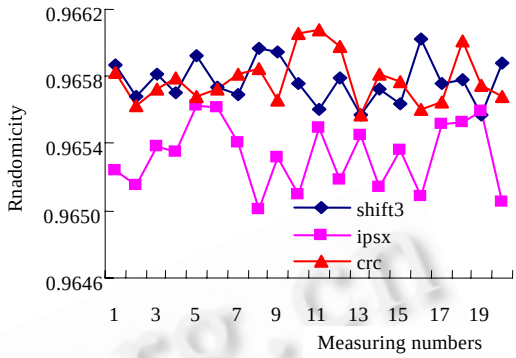

Fig. 4 Comparison among three algorithms
图 4 3 种算法随机性比较

CERNET 数据源中抽取 20 组,每个小时的数据中抽取一组,每组中的数据有 200 000 条流,分别使用异或、位移哈希算法、IPSX 和 CRC32 这 3 种算法哈希得到伪随机序列的测度值如图 4 所示.另一组数据来源于 PMA,表 3 为 AIX 和 TXS 节点测量数据表,其中 AIX 有 4 组数据,AIX1,AIX2,AIX3,AIX4 分别表示其中的每组数据,AIX 表示这 4 组数据的总和;TXS 共有 5 组数据,TXS1,TXS2,TXS3,TXS4,TXS5 分别表示其中的每组数据,TXS 表示这 5 组数据的总和.

Table 2 Comparison among three algorithms using AIX and TXS dataset
表 2 AIX 和 TXS 数据 3 种哈希算法随机测度值比较

	AIX	AIX1	AIX2	AIX3	AIX4	TXS
Number of packets	1 931 949	621 103	605 741	393 969	311 136	2 049 940
XOR_SHIFT	0.995 18	0.990 9	0.989 6	0.988 7	0.985 6	0.993 6
IPSX	0.787 51	0.761 4	0.770 3	0.760 7	0.777 4	0.792 6
CRC	0.996 05	0.991 7	0.990 9	0.988 6	0.986 2	0.995 3
	TXS1	TXS2	TXS3	TXS4	TXS5	
Number of packets	456 544	363 157	493 418	374 960	361 861	
XOR_SHIFT	0.974 3	0.986 0	0.990 0	0.987 4	0.979 3	
IPSX	0.780 3	0.745 3	0.746 4	0.748 1	0.781 4	
CRC	0.978 0	0.987 9	0.991 0	0.987 6	0.981 2	

异或循环哈希算法(XOR_SHIFT),IPSX,CRC 这 3 种算法本质上都是位移和异或操作的组合,都具有时间复杂度为 $O(1)$的运算.对于每个报文,XOR_SHIFT 哈希算法要进行 3 次与操作、6 次位移操作、5 次异或操作;IPSX 对于每个报文需要进行 8 次异或操作、8 次位移操作.假设异或操作和与操作消耗硬件资源相等,则 XOR_SHIFT 哈希算法比 IPSX 算法对于每个报文少 2 次位移操作.而每个报文 CRC32 运算需要异或运算至少 1 536 次操作,与操作 3 072 次操作,因此 XOR_SHIFT 运算效率远高于 CRC32 运算.虽然提出 CRC32 简化的 BOB[9]哈希函数,但 BOB 哈希函数的异或操作和与操作仍有 300 多次.

从图 4 和表 2 可以看出,XOR_SHIFT 算法的随机测度值和 CRC32 算法接近,但是高于 IPSX 算法,CERNET 主干流量的 XOR_SHIFT 和 CRC32 算法的随机测度值与 TXS 和 AIX 流量的随机测度值接近,而 CERNET 流量的 IPSX 随机测度值大于 TXS 和 AIX 流量的随机测度值.其原因是 CERNET 主干比 TXS 和 AIX 覆盖网络范围要大,因而 CERNET 的 FlowID 随机测度值比 TXS,AIX 要大.同时,由于 IPSX 算法没有考虑 IP 流的本身特性,因此不能获得较高的随机测度值.

5 结 论

文献[4,5]通过大量的实验发现,XOR 和 CRC16 算法具有较好的随机性,但是并没有说明随机性较好的原因,且给出的随机测度不能统一描述不同的比特流序列,同时也没有给出面向 IP 流的哈希算法.PSAMP 提出的

IPSX没有考虑IP流的本身特性,没有讨论设计算法的理论依据,其算法的随机性不理想且难以满足网络测量的实际使用;而其建议的CRC32算法时间复杂度太大,不利于高速网络流量测量.

本文相对于前人的工作,贡献主要体现在3个方面:(1) 提出了一个评价哈希算法随机特性的测度,随机测度为评价哈希算法的优劣提供了测度指标.(2) 从理论上分析了异或运算和位移运算的随机特性,并提出了设计基于IP流哈希算法的基本原则.(3) 提出了基于IP流特性的哈希算法,循环位移和异或相结合的哈希算法(XOR_SHIFT),通过实验表明,循环位移3,4,6个比特的哈希函数生成的伪随机序列随机测度值较大,并与PSAMP提出的CRC32和IPSX两种哈希算法进行比较.研究表明,XOR_SHIFT哈希算法生成伪随机序列的随机测度值在统计上和CRC32哈希函数生成的伪随机序列随机测度值差别很小,但执行效率高于CRC32算法.XOR_SHIFT算法执行效率和IPSX接近,但是算法生成的伪随机序列的均匀性高于IPSX.由于本文的XOR_SHIFT算法设计方法是基于IP流的特性,传统的CRC32和IPSX算法是从通用性角度研究,因此本文的算法在执行效率和随机特性两个方面优于其他算法.

References:

[1] IP Flow information export (ipfix). 2004. http://www.ietf.org/html.charters/ipfix-charter.html

[2] Thompson K, Miller G, Wilder R. Wide area Internet traffic patterns and characteristics. IEEE Network, 1997,11(6):10−23.

[3] Cisco Netflow. 2004. http://www.cisco.com/warp/public/732/Tech/nmp/netflow/index.shtml

[4] Jain R. A comparison of hashing schemes for address lookup in computer networks. IEEE Trans. on Communications, 1992,40(3): 1570−1573.

[5] Cao Z, Wang Z, Zegura E. Performance of hashing-based schemes for Internet load balancing. In: Nokia FB, ed. Proc. of the IEEE INFOCOM 2000. Piscataway: IEEE Computer and Communications Societies, 2000. 332−341.

[6] Cheng G, Gong J, Ding W. Distributed sampling measurement model in a high speed network based on statistical analysis. Chinese Journal of Computers, 2003,26(10):1266−1273 (in Chinese with English abstract).

[7] Duffield NG, Grossglauser M. Trajectory sampling for direct traffic observation. IEEE/ACM Trans. on Networking, 2001,9(3): 280−292.

[8] NLANR network traffic packet header traces. 2004. http://pma.nlanr.net/Traces/

[9] Niccolini S, Molina M, Duffield N. Hash functions description for packet selection. 2003. http://www.watersprings.org/pub/id/draft-niccolini-hash-descr-00.txt

附中文参考文献:

[6] 程光,龚俭,丁伟.基于统计分析的高速网络分布式抽样测量模型.计算机学报,2003,26(10):1266−1273.