一种面向时间粒度的流记录聚合自适应哈希函数

蔡少敏，丁伟，张俊

（东南大学 计算机科学与工程学院，江苏 南京 210000）

摘要：针对 NBOS 系统对 NetFlow 流记录的聚合需求，基于 NetFlow 流记录字段特征和网络流量的季节模型特征，设计并实现了一种面向时间粒度的流记录聚合自适应哈希函数。通过与其他一些常用组流哈希函数实验对比，验证了该哈希函数的均匀性、冲突率以及计算速度等性能测度都具有更好的性质，能够满足高速网络流量测量需求。该函数的自适应性使得其可以适用于 NBOS 的任何安装环境。

关键词：NetFlow; NBOS; 哈希函数; 均匀性; 冲突率; 计算速度

中图分类号：TP393　　文献标识码：A

An adaptive Hash algorithm for NetFlow aggregation in the perspective of time granularity

CAI Shao-min, DING Wei, ZHANG Jun

(Department of Computer Science and Engineering, Southeast University, Nanjing, 210000, China)

Abstract: According to the requirement of NetFlow[1] aggregation in NBOS[1] system, an adaptive Hash algorithm for NetFlow aggregation based on the data character of the NetFlow and season model character of the network traffic was designed and realized. Through comparison with other hash algorithms, it is proved to be the best algorithm in the performance of the uniformity, collision rate and calculating speed, so it can meet the needs of high-speed network traffic measurements. The adaptability of this algorithms make it possible that it can be used in any situation NBOS installs.

Key words: NetFlow; NBOS; Hash function; uniformity; collision rate; calculating speed

随着互联网带宽和规模的快速增长，对其的监控和管理面临越来越大的压力和挑战。为了能更好地了解网络运行的状况，Cisco 首先在其路由器上实现了基于抽样的流记录信息的提供，即 NetFlow 测量 IP 流。基于 NetFlow 信息，CERNET 华东北地区网络中心在国家支撑计划的支持下开发了一个面向 IP 集合的网络流量行为观测系统 NBOS[1]。该系统以子网和节点作为观测对象，以时间粒度为单位，提供观测对象的 QOS、流量合理性和流量热点等有关测度。由于 NetFlow 是在抽样的基础上获得的，同时还采用了 15s 的短流终止间隔，这会导致将一个原本单独的流变成了多个流记录输出。为了更好地反映网络中“流”的原始状态，NBOS 以收到的 NetFlow 数据中的流终止时间为标准，以 5min 为时间粒度，对收到的 NetFlow 流记录按流规范中五元组进行聚合，在这个过程中，需要使用 Hash 函数。

哈希值的随机性是基于流的抽样技术的关键，文献[2-7]分析了几种组流哈希算法，它们也能应用于流聚合的情况。为了给 NBOS 选择最合适的 Hash 函数，经过反复验证，cheng 哈希函数具有最好的效果，但因其需要的计算量超过了系统的承受能力，因此，NBOS 最初选择了一个相对简单的方案。随后我们经过努力发现了一种新的面向时间粒度的流记录聚合

收稿日期：2011-08-25; 修订日期：2011-09-28

基金项目：国家科技支撑计划资助项目(2008BAH37B04)

通讯联系人：丁伟(1962-)，女，江苏南京人，东南大学教授，博士生导师; E-mail: wding@njnet.edu.cn.

自适应的哈希函数 NBOS2_hash，经过试验，证明其在各 Hash 性能测度上可以达到 cheng 哈希的效果，在计算量上可以减少 1/3 达到系统在承受能力方面的要求。

本文将详细介绍这个哈希函数。该函数的主要原理是充分考虑了 NBOS 中哈希函数使用的网络环境、数据环境和硬件环境，利用信息熵理论[8-9]和有关概率理论对现有的哈希函数进行了改进。新的算法的自适应性体现在它是根据网络流量季节模型特征，用以前时间单位的流量特征调整算法的有关参数，以保持哈希函数能够在长时间的运行中具有良好的综合性能。

1 相关工作

NetFlow 流的定义通常采用五元组（源地址、宿地址、源端口、宿端口、协议类型）。由于网络中 98%以上的流为 TCP 流和 UDP 流，这个字段的随机性不够理想，所以不考虑将协议字段作为哈希算法的输入参数。文中参与讨论的哈希函数的输入参数均是四元组的流记录，经过哈希运算输出的是 16~20 位的比特串。实验中比较的哈希函数有 NBOS 现在使用的哈希函数[1]、IPSX[2]、cheng[3]。

1.1 NBOS 现在使用的哈希函数

NBOS 现在使用的哈希函数最大的特点是计算量最小，但测度性能也相对比较不理想。其算法如下：

a = sip ^ dip; b[0] ^ = a[0] ^ a[2]; b[1] ^ = a[1] ^ a[3]; b^ = sport ^ dport;

hash = (b ^ (a >> 13) ^ ((a >> 3) & 0xFFFF)）& 0xFFFF;

其中 sip 为源地址，dip 为宿地址，sport 为源端口，dport 为宿端口。

1.2 IPSX 哈希

IPSX 哈希函数是 2008 年提出的一种只有位移和异或两种操作的哈希算法，其算法相对简单且计算量相对较小，但是该哈希函数的性能却比较差。算法如下：

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;

其中，f1 = 源地址，f2 = 宿地址，f3 =源和宿端口。中间变量 h1、v1、v2 均是 32 比特串，算法输出的是 h1 的后 16bit 串。

1.3 cheng 哈希函数

cheng 哈希函数是性能最好的组流函数，该算法主要通过对数据源的随机性特征的分析来安排位移和异或，进而达到提高均匀性的目的。其核心思想是使两个比特串异或结果的随机性更平均。即使用源宿地址、源宿端口共 96 位，每 8 位为一个比特串单元，共 12 个比特串。通过使后一个比特串中随机性最好的一位与前串中随机性最差的一位异或，来增加该位的随机性。

这些组流函数并不适合 NBOS 的需求，NBOS 现在使用的哈希函数和 IPSX 哈希均只是采用异或和位移操作，它们的计算量小，计算速度相对较快，但同时均匀性能较差同时冲突率很高；cheng 哈希函数性能很好，但是计算量偏大，不能满足 NBOS 的实时性要求。

为了在运算代价以及性能效果两方面达到平衡，在充分发挥硬件资源的情况下，尽可能提高哈希函数的综合效率，使之更加适用类似 NBOS 系统的需求。本文提出了一种新的面向时间粒度的流记录聚合自适应哈希函数，下面将详细介绍。

2 自适应 Hash 函数

我们将这种新的面向时间粒度的流记录聚合自适应哈希函数称为 NBOS2_hash。该哈希函数能够在运算代价小的情况下实现很好的性能效果，另外，该哈希函数是自适应的，它是根据网络流量季节模型特征，通过上一个周期的流量特征调整算法的有关参数，以保持哈希函数能够在长时间的运行中具有良好的综合性能。

2.1 算法思想

通过增加一个并发的分析模块对上一个周期的数据进行分析，根据网络流量的季节模型特征，上一个周期可以是前一天，也可以是上一周的同一天，或者是两者的组合。

算法思想为：

①统计上一个周期单位时间粒度内 NBOS 流记录的平均数量，结合硬件资源的情况，确定哈希函数的输出空间，充分利用硬件资源；

②利用信息熵理论和概率论中随机分布理论，对哈希函数的输入参数的数字特征进行统计分析，以缩小函数的输入数据量；

③利用异或运算提高运算结果数据的随机程度来设计具体运算操作，提高哈希输出函数的均匀性。

其前两项工作是 NBOS2_hash 的分析模块，功能是经过统计上一个周期 NBOS 系统中的数据，计算出参与哈希函数计算的基本参数，同时传递给 NBOS 控制中心。分析模块使用另外一台主机来完成，不占用 NBOS 系统的任何资源，其网络环境如图 1 所示。

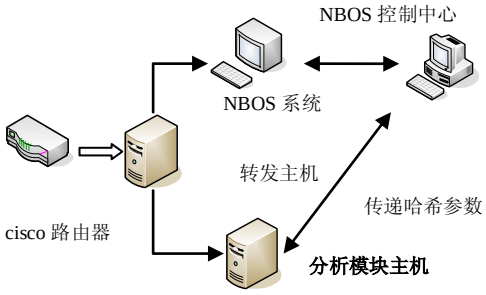

图 1 NBOS 网络环境图

Fig.1 NBOS network environment

2.2 具体算法描述

① 统计上一个周期的NBOS流记录数。取得其单位时间粒度内的平均值。根据其大小，确定哈希函数的输出比特串的长度L，即哈希函数的输出空间为2^L。

② 以时间粒度为单位（比如 5min）统计最近 24h 的 NetFlow 流记录的四元组。统计出 NetFlow 流记录源地址中每位比特对应随机变量的信息熵值，取信息熵值最大的 L 个比特位按先后次序组成一个新的比特串 β，同时记录下这些比特位在比特串中的位置，即相对于最高位的偏移量 γ_n (n=1, 2，…, L)。将剩余的（32− L）个比特位按照原来的先后顺序组成新串，并顺次与源端口、宿 IP 和宿端口顺序拼接成长度为（96− L）的比特串。

③ 计算λ=(96-L)/L，将长度为（96− L）的比特串中顺次取出(λ−1)个长度为L的比特串，剩余的比特串为(96−λL)位，分别记为λ_i (i=1, 2, …, λ)。

④ 第λ个比特串的确定：利用③中剩余的(96 −λL)位比特串，与宿端口的最后⌈(1+λ) L −96⌉位顺序拼接而成。

⑤ 计算哈希值：hash=β^ λ_1 ^ λ_2 ^ … ^ λ_λ

算法流程图如图 2 所示，图中虚线框内的工作在分析模块主机上完成。

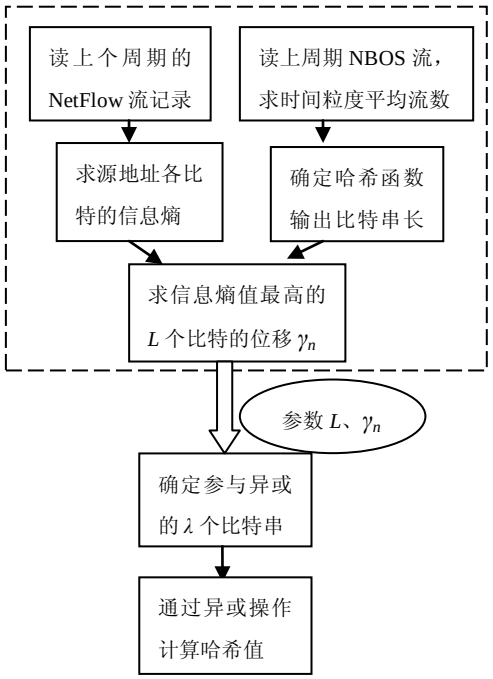

图 2 NBOS2_hash 算法流程图

Fig.2 flow chart of NBOS2_hash algorithm

3 实验及性能比较

3.1 实验

实验数据来源于 NBOS 实际运行的 JSERNET 边界，采集 2011 年 8 月 7 日 10:00 到 11:00 的 NetFlow 数据(表 1)。对四种哈希算法的哈希值，进行了均匀性、冲突率以及计算时间的比对。

表 1 12 个时间粒度内 NetFlow 流记录数据

Tab.1 Netflows in 12 time granularity

时间粒度	1	2	3	4	5	6
流数	605094	624321	620413	651135	663243	645156
时间粒度	7	8	9	10	11	12
流数	671306	664084	673072	665326	653923	652842

对于 IPSX 哈希、cheng 哈希以及 NBOS 系统现在使用哈希算法，都只需要在一段时间范围内一条条读取 NBOS 流四元组，根据相应的哈希算法求得其哈希值；而对于本实验提出的改进的哈希算法，首先需要计算要参与哈希计算的基本参数 L、λ 和 r_n。这些基本参数是通过 NBOS 源数据预分析系统在 JSERNET 某主节点采集到的前一天的 NetFlow 流记录和 NBOS 流记录数，作为哈希函数改进的参考和使用数据。

根据 2011 年 8 月 6 日 24h 的 NBOS 流记录，求得其单位粒度内流数平均值 O_{ave}=55653，L=16；根据这 24h 的 NetFlow 流记录中源 IP 地址各比特位为 1 的概率 p，结合信息熵公式

$$H(p)=-p\log_2 p-(1-p)\log_2(1-p),$$

求得各比特位的比特熵如图 3 所示。

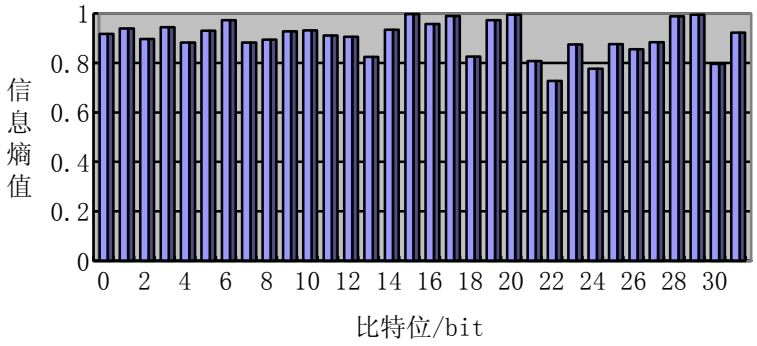

图 3 源 IP 地址各比特位的信息熵

Fig.3 entropy of each bit in source IP address

由图 3 可得，γ_n (n=1,2，…, L）取源 IP 的第 15、29、20、17、28、6、16、3、1、14、10、5、9、31、0、11 位，依次拼接成 β 比特串，λ=5，λ_1 =源 IP 的第 12、2、8、27、7、4、25、23、26、18、19、13、21、30、24、22 位组成的比特串，λ_2=源端口，λ_3=宿 IP 前 16 位，λ_4=宿 IP 后 16 位，λ_5=宿端口，最后计算其哈希值。

3.2 性能比较

3.2.1 均匀性

卡方检验其均匀性，依次取时间粒度（5min）内的 NetFlow 流记录，共进行 12 次(1h)实验。根据统计量的卡方拟合优度[10]检验，分别计算 NBOS 现在使用的哈希函数、改进后的哈希函数、IPSX 哈希函数和 Cheng 哈希函数在各个时间粒度内的均匀性。

哈希函数均匀性对比图如下图所示。

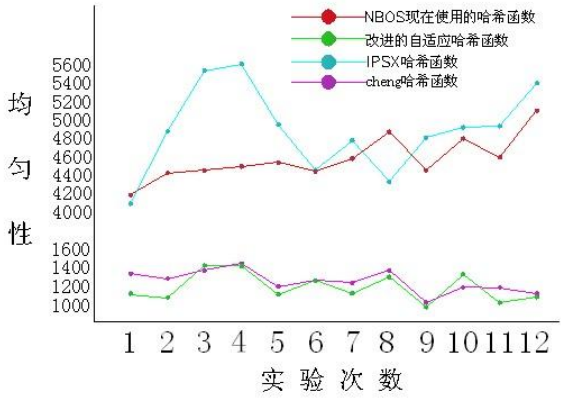

图 4 哈希函数的均匀性对比图

Fig.4 uniformity comparison chart of four hash functions

从这一个小时数据的测试结果可以看出，IPSX 哈希函数的均匀性能最差，且很不稳定，NBOS 现在使用的哈希函数的均匀性也很不理想；而改进后的哈希函数和 cheng 哈希函数均匀性较其他两个哈希函数好很多，其映射结果可以看作是均匀分布的，而且改进后的自适应哈希函数整体呈现最好的均匀性。

3.2.2 冲突率

本文选择参考文献[10]中定义的哈希函数冲突率公式来评价本文提出的哈希函数。

$$f=\frac{1}{\frac{N(N+1)}{2}}\sum_{i=1}^{M}\frac{N_i(N_i+1)}{2}$$ 。

对不同的哈希函数来说，该公式的数值越大，说明哈希函数的冲突率越高。经过 12 次的实验，4 种哈希函数的冲突率如图 5 所示。

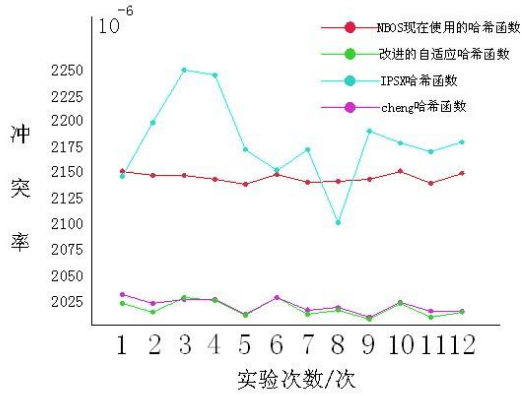

图 5 哈希函数的冲突率对比图

Fig.5 collision rate comparison chart of four hash functions

从实验结果可以看出，改进后的哈希函数和 Cheng 哈希算法的冲突率非常相近，并且非常稳定。

由图 4 和图 5 可以看出，改进后的哈希函数在每个时间粒度内的均匀性和冲突率比其他几种哈希算法要优越，并且每次实验结果都比较稳定，这与初始串的选择是紧密相关的。初始串是由比特位随机变量信息熵最高的多个比特组成，在一定程度上保证了运算结果随性能的稳定程度。

3.2.3 计算速度

在网络数据流量很大时，流记录到达的时间间隔非常短，这就要求哈希函数的运算速度尽可能快。本文主要通过哈希函数的运算操作次数来判断其计算速度。

哈希函数的操作次数比较如表 2 所示。

表 2 四种哈希函数操作运算次数比较表

Tab.2 comparison of four hash functions' computing times

哈希函数	异或操作	与操作	或操作	移位操作	总操作次数
NBOS2_hash	5	16	0	1	22
NBOS	9	2	0	2	13
IPSX	8	1	0	9	18
cheng	20	3	0	12	35

本文所提出的哈希函数与 Cheng 哈希和 IPSX 哈希相比，大大减少了运算量，因为它用到的移位操作很少，在时间粒度内 NetFlow 流量非常大的情况下，改进后的哈希算法具有更高的操作速度，参数自适应调整机制也更能满足 NBOS 实时处理的需求。综合均匀性、冲突率以及计算速度这三个测度考虑，本文提出的哈希函数在随机性上比 IPSX 和 cheng 哈希更好，在速度上也比 cheng 哈希快很多，是最适合面向时间粒度的流记录聚合的哈希函数。

4 结语

本文主要从改进 NBOS 中使用哈希函数的目的入手，对哈希函数的需求进行了更深入的分析。相对于前人的工作，贡献主要体现在以下几个方面：①利用单位时间内的流记录数确定哈希函数的输出空间，保持哈希桶的平均长度低于 1.5，有效防止冲突率过高；②将信息熵理论应用于输入比特串的选择，便于优化哈希函数的输入参数，从而提高输出值的随机程度。输出值的随机程度越高，就意味着哈希桶被均匀占用的可能性越大，从而有效地改善哈希函数的性能；③两个不相关的随机变量进行异或操作，可以使结果的随机性能不低于参与运算的两个比特位随机变量的随机性能，为哈希函数的运算操作的选取提供了理论依据。通过对 IP 地址结构的分析，可以结合信息熵理论更好的把握 NetFlow 流记录四元组的信息特征，使得改进后的自适应哈希函数适用范围更广，性能更好，同时对信息熵理论应用于哈希函数进行理论分析和实验验证。在理论的支撑下，结合实验，验证了本文提出的哈希函数具有相对较好的综合性能，更适用与 NBOS 类似的面向时间粒度流记录聚合的特殊环境。

参考文献：

[1] 顾文杰.网络行为观测系统 NBOS 的并行化设计和实现[D].南京：东南大学计算机科学与工程学院, 2010.

[2] 夏青, 丁伟. 一种基于数据相关性特征的网络数据流 HASH 函数[C] //张凤祥. 全国计算机新科技与计算机教育论文集: 第 1 卷. 四川成都: 西南交通大学出版社. 2009: 150-153.

[3] Cheng Guang, Gong Jian, A Method to Device Flow Hash Algorithm Based Traffic Character[C] // IEEE International Conference on Communication Technology Proceedings . Hangzhou: Hangzhou Dianzi University. 2008.

[4] Cheng Guang, Zhao Wei, Gong Jian. XOR Hashing Algorithms to Measured Flows at the High-Speed Link [C]// Second International Conference on Future Generation Communication and Networking . Computer Science and Engineering. 2008.

[5] Cheng Guang, Gong Jian, Ding Wei. Distributed sampling measurement model in a high speed network based on statistical analysis[J]. Chinese Journal of Computers, 2003, 26 (10): 1266-1273.

[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[J]. 计算机学报, 2003, 26(10): 1266-1272.

[7] 程光, 龚俭, 丁伟. 面向 IP 流测量的哈希算法研究[J]. 软件学报, 2005, 16(05): 652-658.

[8] 潘乔. 网络测量中的抽样技术研究[D]. 西安：西安电子科技大学计算机工程, 2007.

[9] 赖宏图. 基于熵理论的真实 IPv6 地址随机测度分析[J]. 福建工程学院学报, 2010, 8(1): 75-76.

[10] 强士卿, 程光. 基于流的哈希函数比较分析研究[J]. 南京师范大学学报, 2008, 8(4):25-28.