一种基于包速率自适应的报文抽样算法 *

陈庶樵, 张 果, 朱 柯
(国家数字交换系统工程技术研究中心, 郑州 450002)

摘 要: 针对 NetFlow 抽样概率需手动配置的缺陷,提出了一种基于包速率自适应的分组抽样算法。通过测量包速率,采用预定义测量误差的方法,根据包速率的变化自适应地调整抽样概率,从而在有限资源情况下达到控制测量误差的目的。基于实际互联网数据进行了实验比较,结果显示:与传统的 NetFlow 算法相比,该方法易于实现,测量误差可控,具有高效性和准确性,同时具有资源节约性。

关键词: 流量测量; 包速率; 自适应; 抽样
中图分类号: TP393 **文献标志码**: A **文章编号**: 1001-3695(2010)07-2727-03
doi:10.3969/j.issn.1001-3695.2010.07.092

Algorithm based on packet rate adaptive for packet sampling

CHEN Shu-qiao, ZHANG Guo, ZHU Ke
(*National Digital Switching System Engineering & Technological R&D Center*, *Zhengzhou* 450002, *China*)

Abstract: For the inflexibility of NetFlow's sampling probability, this paper proposed the algorithm based on packet rate adaptive for packet sampling. The algorithm measured the packet rate, employed the predefine measurement error, adaptively adjusted the sampling probability according to the change of packet rate and advanced to control the measurement error. Also conducted experiments based on real network traces. Results demonstrate that the proposed method can implement simplicity, controllability of measurement error with higher efficiency and without sacrificing accuracy, while memory consumption is lower compared with other methods.

Key words: traffic measurement; packet rate; adaptive; sampling

网络流量分析是进行科学的网络建设以及实现高效可控的业务承载的有力支撑与重要保障,但是,链路带宽的快速增长和网络流量的急剧膨胀给流量测量带来了极大的挑战:处理器资源的大量消耗导致路由器整体性能下降;产生的海量数据将占有大部分存储器的资源;输出的流量记录占用大量的传输带宽。IFTF 工作组的 IPFIX(IP flow information export)和 PSAMP(packet sampling)建议使用报文抽样技术进行流量测量。Cisco 在 NetFlow 中引入抽样机制[1]以适应网络的高速化。虽然如此,NetFlow 的抽样方法仍然存在一些缺点[2]:a)使用"1 out of *N*"的静态抽样策略,抽样率需人工配置,造成使用不便;b)所用资源随流量波动而变化,缺乏资源保护功能。本文针对 NetFlow 现存的问题,提出了一种基于报文速率自适应的分组抽样算法(packet rate adaptive sampling algorithm, PRAS)。该算法把单个测量时隙划分为尽可能多的互不重叠的等长子测量时隙,通过实时测量该子时隙的报文速率,在满足预定义误差的情况下,确定对区间固定分组数量的简单随机抽样概率,从而达到控制测量误差的目的。

1 实时报文速率测量

1.1 算法基本结构

由于网络流的高速率特性,致使单位时间内会有大量的数据包,若全部缓存将会需要大量的硬件资源。出于资源限制的考虑,本文采用控制子测量时隙的方法:在一个测量时隙内,将测量时隙化为尽可能多个等长度的子测量时隙;对每个子测量时隙,把在该子测量时隙内到达的所有分组作为抽样样本总体,对该时间段内到达的包进行简单随机抽样。

基于上面的划分测量时隙的思想,设计子测量时隙内的抽样过程。在子测量时隙内,采用基于数据包的抽样策略,通过统计该子测量时隙内的数据包数量计算出满足控制误差条件的抽样概率,对该时间间隔内到达的数据包进行简单随机抽样。示意图如图 1 所示。采用这种抽样设计方法,一方面抽样方法简单、易操作;另一方面,使子测量时隙尽可能地小,便于控制存储器资源和处理器资源,防止被过度占用。同时子测量时隙内的抽样过程是相同和相互独立的。

1.2 分组测速缓存模型

从流量测量的需求来看,测速缓存中只需存放分组头部的某些域及其他相关统计信息[3]。由于网络流量随时间不断变化和高速的特性,使得缓存模型采用对数据包循环分批处理的思想:用较多的缓存来保存分组信息,数据包测速和流量统计两个过程同时进行,进而提高测量效率。在本文中给出较为保守的缓存设计:缓存由 1 和 2 两部分组成,每部分可存放 n(n 为子测量时隙内所能达到的分组数的最大值)个分组信息,当子测量时隙结束时,根据分组总数计算出抽样率,再对缓存内分组进行简单随机抽样。缓存存储数据包过程如下:按照到达顺序对连续的 n 个数据包按照子测量时隙($i=1,2,3,\cdots$)进行分批缓存,当 i(即第 i 个时隙)为奇数时,从链路到达且数据包的特

收稿日期: 2009-12-08; **修回日期**: 2010-01-11 **基金项目**: 国家"863"计划资助项目(2008AA01A323)

作者简介:陈庶樵(1973-),男,吉林人,教授,博士,主要研究方向为宽带信息网络;张果(1985-),男,助理工程师,硕士研究生,主要研究方向为网络测量(zhangguo1716@yahoo.com.cn);朱柯(1975-),男,讲师,博士,主要研究方向为网络体系结构.

征信息进入缓存 1,当 i 为偶数时,缓存信息进入缓存 2(图 2)。

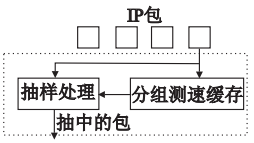

图1 数据包测速抽样示意图

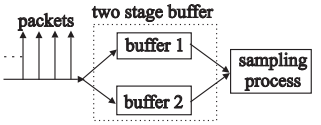

图2 测速缓存设计结构图

2 抽样概率

根据预定义的测量误差,由中心极限定理计算满足误差要求的抽样报文的数量,以确定单个测量时隙的最佳抽样概率。

2.1 符号定义

L 为测量时隙的长度;S_k 为第 k 个测量时隙的数据包个数;S 为整个测量时间段内的数据包个数;P 为子测量时隙报文的抽样概率;α 为子测量时隙的抽样概率进行估计的置信水平;β 为子测量时隙的报文抽样概率的误差范围;Z_α 为标准正态分布下置信水平 α 的分位点;n 为子测量时隙内抽样报文的数量;Y_i 为单个测量时隙内总体报文中的每个报文;n^* 为子测量时隙内满足预定误差的最小抽样报文数;p^* 为子测量时隙的最佳抽样概率;f_i 为整个测量时间段内被创建的每一个流。

2.2 确定最佳抽样概率

在一个测量时隙内,当数据包到达时,若它被作为样本抽取,则先判断流量信息表里有无该流。如果存储器里没有该条目,则创建;否则,该报文所属的流量条目累加计数。抽样概率直接影响了测量的结果,下面介绍如何确定满足预定义测量误差的最佳抽样概率。流量负载 S_k 由前端的流测速缓存得到,则对于此子测量时隙,只要计算出满足误差要求的抽样报文的数量,就能得到该子时隙内的抽样概率。由于网络的突发特性,子测量时隙内数据包的个数是不断变化的,为了不引进额外的流量测量误差,必须把分组抽样的个数严格控制在一定的范围。

在单个测量时隙的抽样为简单随机抽样,由简单随机抽样的性质[4],得到如下引理:

引理 1 从总体 S_k 中抽取样本量为 n 的简单随机样本。若对总体中的每个单元报文 Y_i,定义随机变量 X_i 如下:

$$X_i=\begin{cases}1 & \text{若 } Y_i \text{ 被抽中}\\ 0 & \text{若 } Y_i \text{ 未被抽中}\end{cases}\quad i=1,2,\cdots,S_k \tag{1}$$

则

$$E(X_i)=\frac{n}{S_k};i=1,2,\cdots,S_k \tag{2}$$

$$V(X_i)=\frac{n}{S_k}(1-\frac{n}{S_k});i=1,2,\cdots,S_k \tag{3}$$

证明 由定义可以看出随机事件 X_i 只有两个结果,服从伯努利实验的条件,重复从总体 S_k 中抽取 n 个数据包,那么该试验即为 n 重伯努利实验。所以表示这 S_k 个随机事件的 S_k 个随机变量都服从二项分布,且 $X_i=1$ 的概率为 n/S_k,$X_i=0$ 的概率为 $1-n/S_k$,结论得证。

由于大量随机变量 X_i 服从同一分布且相互独立,由中心极限定理得到引理 2。

引理 2 由于随机变量 X_1, X_2,$\cdots$,X_{S_k} 相互独立,服从同一分布,并且具有数学期望和方差:$\mu=E(X_i)=n/S_k$,$\sigma^2=V(X_i)=n/S_k(1-n/S_k)$。令 $\overline{X}=\frac{1}{S_k}\sum_{i=1}^{S_k}X_i$,则对于大规模的总体 S_k,得到

$$\overline{X}=\frac{1}{S_k}\sum_{i=1}^{S_k}X_i\sim N[\mu,\sigma^2/S_k] \tag{4}$$

其中:$N[\mu,\sigma^2/S_k]$ 是以 μ 为均值、σ^2/S_k 为方差的正态分布。

证明 结合已知条件,由中心极限定理可知:

$$\frac{\frac{1}{S_k}\sum_{i=1}^{S_k}X_i-\mu}{\sigma/\sqrt{S_k}}\overset{\text{近似地}}{\sim}N[0,1] \tag{5}$$

又因为 $\overline{X}=\frac{1}{S_k}\sum_{i=1}^{S_k}X_i$,所以有 $\frac{\overline{X}-\mu}{\sigma/\sqrt{S_k}}\overset{\text{近似地}}{\sim}N[0,1]$,即 $\overline{X}=\frac{1}{S_k}\times\sum_{i=1}^{S_k}X_i\sim N[\mu,\sigma^2/S_k]$,故得证。

在单个测量时隙的抽样概率满足等式 $P=n/S_k$,又因为抽样概率 $P=\frac{1}{S_k}\sum_{i=1}^{S_k}X_i$,将 P 代入引理 2,可得 $\sqrt{S_k}[P-\mu]/\sigma\sim N[0,1]$。要想控制抽样误差,只需预先定义单个测量时隙抽样概率的估计精度即可。单个测量时隙内,以置信水平 α,使得报文抽样概率的估计值 P 满足:$P\in(-\beta/2,\beta/2)$,即单个测量时隙,报文的抽样概率以大于等于 α 的概率,存在 $\pm\beta/2$ 的误差。

由引理 1、2 的特性得到,以置信水平 α 进行估计抽样概率的置信区间为

$$P\pm Z_\alpha\sqrt{\frac{\frac{n}{S_k}(1-\frac{n}{S_k})}{S_k}}\Rightarrow P\pm Z_\alpha\sqrt{\frac{P(1-P)}{S_k}}$$

当 $P=\frac{1}{2}$ 时,$P(1-P)$ 有最大值 1/4,此时,置信区间有最大范围 $\pm Z_\alpha\sqrt{\frac{1}{4S_k}}$。欲满足预定的误差范围,只需置信区间的最大范围小于预定义的误差范围,即

$$Z_\alpha\sqrt{\frac{1}{4S_k}}\leqslant\frac{\beta}{2}\Rightarrow Z_\alpha\sqrt{\frac{1}{8n}}\leqslant\frac{\beta}{2}\Rightarrow n\geqslant n^*=\frac{Z_\alpha^2}{2\beta^2} \tag{6}$$

以上推出的抽样概率,由于其一方面减少了抽样样本的数量,降低了抽样过程对系统所造成的额外负荷,另一方面,又可以满足预定义的误差水平,故称之为最优抽样概率。所以子测量时隙内的最佳抽样概率为 $p^*=n^*/S_k$。要得到 p^*,需要定义 (α,β),再由测速缓存得到第 k 个子测量时隙内的分组个数 S_k。

3 流量测量算法流程

该测量算法主要包括以下五个部分:首先预定义测量误差;分隔定义测量时隙为 K 个子时隙,在子测量时隙内,根据得到的数据包速率计算最佳抽样概率;然后利用该抽样概率对数据包进行简单随机抽样;创建、更新或者输出流记录,测量时隙结束时输出结果。该算法的流程如图 3 所示。

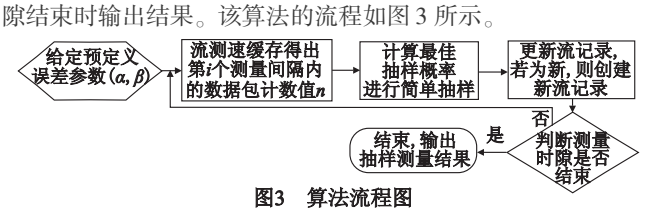

图3 算法流程图

4 实验结果与分析

本章利用来自 NLANR[5] 互联网 IPLS→CLEV 链路上的实际流量数据对抽样算法进行验证。在实验中,在流量的探测点,采用〈源 IP,目的 IP,协议类型〉来区分各个流。

4.1 抽样算法的分组抽样估计无偏性

时间分层分组抽样具有以下性质[6]:时间分层分组抽样

得到的任何一个流,它的分组数的估计值具有无偏性;任何一个流,它的字节数的估计值具有无偏性。本文提出的基于包速率自适应的分组抽样,属于时间分层分组抽样的范畴,因此该算法也具有以上性质。图 4 是利用 PRAS 算法在固定的 100 个测量时隙内, 20 次测量结果的分组误差散点分布图,其中 x 轴是流的分组数,y 轴是相对误差。可以发现,估计误差逐渐随着分组数的增加而减小;误差点随 $y=0$ 水平线的对称性说明了分组数估计的无偏性。

4.2　抽样算法的缓存大小和包抽样误差关系

图 5 中的曲线是使用 PRAS 算法对连续 300 个测量时隙的流量得到的结果。可以看出,本文算法当测速缓存变化时,对应的测量估计误差随之变化,因为在单个子时隙内,缓存越大,能够被统计到的分组越多,能够更好地控制测量误差,估计就越准确。同时,聚合业务流的流量越大,其相对标准差越小;流记录的条目数量越多,其相对标准差越小。因为同样条件下,聚合业务流的流量越大,属于该流的报文被抽中的概率越高,单位时间被抽中的报文数量就越多,测量误差就会越小。正是由于可利用的缓存资源和输出的流记录数量成正比关系,随着记录的流的条目数量越多,其相对标准差也会越小。

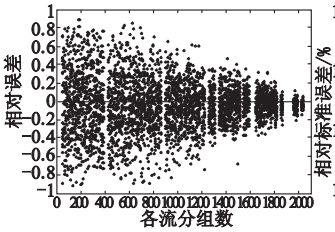

图4　分组数相对误差分布图

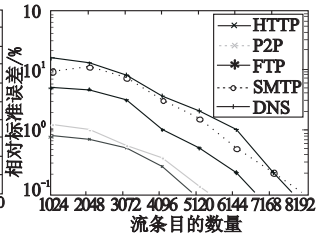

图5　流分组数估计值的相对标准差曲线

4.3　时间子区间划分与抽样估计误差关系

由于链路的流量速率比较大,不可能对所有的测量时隙进行仿真对比。为了更好地仿真,现只需要对其中某些测量时隙进行验证即可。测量时隙固定,缓存大小固定,通过调整子测量时隙来验证子时隙和相关误差的关系,随着测量子时隙的变大,聚合流的测量估计误差也随之变大(图 6)。

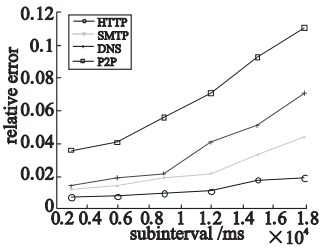

图6　缓存大小固定,测量子时隙大小和误差的关系

4.4　与传统方法的比较

本文利用实际的网络流量数据,从抽样概率和存储器资源两个方面,分别对文献[1]中 NetFlow 的静态抽样机制、文献[2]中的 Adaptive NetFlow、本文的 PRAS 算法进行仿真对比。在确定 NetFlow 的静态抽样概率时,首先运行 PRAS 算法,根据最后的分组总数和抽样分组数计算一个等效抽样率;然后把这个抽样率赋给静态抽样概率。取定 $\alpha=0.998$, $\beta=0.01$,得到单个测量时隙的抽样样本的数量 $n^*=Z_\alpha^2/2\beta=15\ 000$。为便于比较,仿照文献[2],设置"time bin"为 60 s,由于测量时隙的长度等于 6 s,在 Adaptive NetFlow 的一个"time bin"中包含 10 个 PRAS 算法的测量区间。PRAS 算法在一个"time bin"里,抽样率不需要人工配置,可以根据数据包速率自适应地调整:流量较大时,降低抽样率,减轻测量过程对系统造成的负荷;流量较小时,增加抽样率,提高流量测量的准确性。

因为时间分层分组抽样具有如下性质[6]:流 f_j 的分组数估计值的相对标准差具有上界 $1/\sqrt{M\cdot S^j/S}$,其中 M 为可利用的缓存资源。本文中的算法也具有以上性质,通过链路实际流量数据的处理,表 1 中七种业务流的流量百分比的数据可以仿真对比分组数的相对误差所占百分比的关系。

表 1　等效抽样概率聚合业务流的流量百分比误差比较

aggregate flows	HTTP	P2P	FTP	SMTP	RTSP	NETBOIS	DNS
percent/%	51.21	20.35	8.34	3.45	0.42	0.06	1.50
static sample	0.008	0.012	0.055	0.097	0.050	0.041	0.149
adaptiveNetFlow	0.012	0.012	0.051	0.132	0.040	0.043	0.158
PRAS	0.008	0.013	0.052	0.099	0.049	0.033	0.156

表 2、3 是对三种算法的准确性(利用相对误差来衡量)的对比。为了比较的公平性,需要保证三种算法所使用的缓存资源相同,表 2、3 分别是在 4 k 和 8 k 的缓存资源的情况下得到的。可以看到,本文算法并没有引入额外误差。同时随着缓存的增大,各种聚合业务流估计误差均有所降低。

表 2　4 k 情况下聚合业务流的流量百分比误差比较

aggregate flows	HTTP	P2P	FTP	SMTP	RTSP	NETBOIS	DNS
percent/%	51.21	20.35	8.34	3.45	0.42	0.06	1.50
static sample	0.009	0.013	0.058	0.098	0.060	0.045	0.160
adaptive NetFlow	0.013	0.013	0.053	0.135	0.050	0.044	0.163
PRAS	0.009	0.014	0.054	0.096	0.051	0.036	0.158

表 3　8 k 概率情况下聚合业务流的流量百分比误差比较

aggregate flows	HTTP	P2P	FTP	SMTP	RTSP	NETBOIS	DNS
percent/%	51.21	20.35	8.34	3.45	0.42	0.06	1.50
static sample	0.006	0.010	0.050	0.085	0.045	0.038	0.135
adaptive NetFlow	0.009	0.010	0.049	0.099	0.039	0.040	0.143
PRAS	0.006	0.011	0.050	0.085	0.045	0.030	0.149

5　结束语

本文方法具有抽样估计的无偏性、抽样概率自适应等优点,并且易于实现、操控简单。仿真实验结果显示,本文的方法在保证测量准确性的同时,很好地控制了内存资源的使用。另外,如何实现抽样数据包的线速保存和统计将是本文下一步的研究方向。

参考文献:

[1] CISCO SYSTEM, Random sampled NetFlow[EB/OL]. (2005). http://www. cisco. com/en/US/products/sw/iosswrel/ps5207/products_feature_guide09186a00801a7618. html.

[2] ESTAN C, KEYS K, MOORE D, *et al*. Building a better NetFlow [C]//Proc of Conference on Applications, Technologies, Architectures, and Protocols for Computer. New York: ACM Press, 2004: 245-256.

[3] 张震,汪斌强,朱珂. 流量测量的关键技术分析与研究[J]. 计算机应用研究, 2009, 26(9):3442-3447.

[4] 杜子芳. 抽样技术及其应用[M]. 北京: 清华大学出版社, 2005: 124-132.

[5] NLANR [EB/OL]. (2006). http://pma. nlanr. net.

[6] 王洪波, 韦安明, 林宇, 等. 流测量中基于测量缓冲区的时间分层分组抽样[J]. 软件学报, 2006, 17(8):1775-1784.