软件学报 ISSN 1000-9825, CODEN RUXUEW
Journal of Software,2013,24(2):255−265 [doi: 10.3724/SP.J.1001.2013.04316]

E-mail: jos@iscas.ac.cn
http://www.jos.org.cn
Tel/Fax: +86-10-62562563

基于近似方法的抽样报文流数估计算法*

程 光[1,2], 唐永宁[3]

[1](东南大学 计算机科学与工程学院,江苏 南京 210096)

[2](计算机网络和信息集成教育部重点实验室(东南大学),江苏 南京 210096)

[3](School of Information Technology, Illinois State University, Normal 61790-5150, USA)

通讯作者: 程光, E-mail: gcheng@njnet.edu.cn, http://www.seu.edu.cn

摘 要: 维护每个报文的流记录需要占用大量测量资源.目前已有多种抽样技术估计网络流统计信息,然而精确地估计出流数统计信息是目前的研究难点.提出了 Integral 和 Iteration 两种基于报文抽样样本估计网络流数的算法. Integral 算法只需使用抽样流长为 1 的流数信息就可以近似推导出未抽样的流数.Iteration 算法通过建立迭代函数估计未抽样流数,然后根据未抽样流数和已抽样的流数推断出原始流量的流数.采用 CERNET(China education and research network)骨干网络链路数据将这两种算法与 EM(expectation maximization)算法进行对比,表明 Iteration 算法具有较好的精度和性能.

关键词: 报文抽样;网络测量;流数;二项分布

中图法分类号: TP393 **文献标识码**: A

中文引用格式: 程光,唐永宁.基于近似方法的抽样报文流数估计算法.软件学报,2013,24(2):255−265. http://www.jos.org.cn/1000-9825/4316.htm

英文引用格式: Cheng G, Tang YN. Estimation algorithms of the flow number from sampled packets on approximate approaches. Ruanjian Xuebao/Journal of Software, 2013,24(2):255−265 (in Chinese). http://www.jos.org.cn/1000-9825/4316.htm

Estimation Algorithms of the Flow Number from Sampled Packets on Approximate Approaches

CHENG Guang[1,2], TANG Yong-Ning[3]

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

[2](Key Laboratory of Computer Network and Information Integration (Southeast University), Ministry of Education, Nanjing 210096, China)

[3](School of Information Technology, Illinois State University, Normal 61790-5150, USA)

Corresponding author: CHENG Guang, E-mail: gcheng@njnet.edu.cn, http://www.seu.edu.cn

Abstract: Recording flow statistics for each network packet is resource-intensive. Various sampling techniques are used to estimate flow statistics. However, the estimation accuracy based on the sampling remains a significant challenge. This paper introduces both sampling techniques denoted as Integral and Iteration algorithms, which can accurately infer the number of original flows from the sampled flow records. The Integral algorithm uses only the number of sampled flows with one sampled packet to approximately deduce the number of unsampled flows. The Iteration algorithm can estimate the number of unsampled flows using an iteration method. The number of original flows can be precisely estimated according to both the number of sampled flows and unsampled flows. Both the algorithms are compared to the EM (expectation maximization) algorithm using multiple traffic traces collected from CERNET (China

* 基金项目: 国家自然科学基金(60973123); 国家重点基础研究发展计划(973)(2009CB320505); 江苏省科技计划项目(科技支撑计划——工业部分)(BE2011173)

收稿时间: 2012-03-16; 修改时间: 2012-07-03; 定稿时间: 2012-08-20

education and research network) backbone. The result shows that the Iteration algorithm is superior to the EM algorithm and can provide highly accurate estimation on the number of original flows.

Key words: packet sampling; network measurement; number of flows; binomial distribution

由于处理器能力、缓存容量、网络带宽等硬件资源的限制,测量高速链路中传输的所有报文信息代价过大,现有的网络测量设备大多采用报文抽样的方法进行流量测量,然后对抽样报文采用统计推断方法进行总体统计特性的估计.在 Cisco 路由器中的 Sampled NetFlow[1]功能采用报文抽样功能,IETF(Internet Engineering Task Force)的 PSAMP(packet sampling)和 IPFIX(IP flow information export)工作组[2]提出了几种报文抽样方案,Duffield[3]指出报文抽样可以应用在计费、流量管理、故障检测和 SLA(service level agreement)检测等领域.网络流就是在一个测量时间段内通过一条链路的报文中源 IP、宿 IP、源端口、宿端口和协议这 5 个字段或其中若干个字段的报文集合,流数就是在这些不同报文集合的数量,蠕虫、病毒、扫描等事件都与流数特性相关,流数的测量和监控是网络管理的重要研究领域.准确地测量网络中的流数信息,对网络管理和安全非常重要.

基于随机抽样报文的流数估计方法主要有 3 种.第 1 种方法是直接将抽样报文的流数除以抽样比率.第 2 种方法是简单地使用抽样报文中的流数作为原始流数[4].这两种方法的缺点是没有考虑流量的分布特性,在抽样时造成大量短流信息的丢失,第 1 种方法高估了流数,第 2 种方法低估了流数.第 3 种方法是采用 EM(expectation maximization)算法推断出原始流量的流长分布,然后根据二项分布的方法统计出流数.EM 算法的运行时间很长,且算法大多不能收敛,一般采用定义迭代次数来强行结束算法运行;同时,EM 算法对于大流和短流混合估计造成精度很低.

本文采用抽样报文来推断原始流量中的流数.与传统流数方法相比,其主要区别在于将原始流量的流数估计分成未抽样的流数和已抽样的流数:对抽样流量进行统计,计算出已抽样流量部分;对于未抽样流数的估计,本文提出了 Integral 算法和 Iteration 算法.Integral 算法近似采用积分建立未抽样流数和抽样后流长为 1 的流数之间的关系.Iteration 算法基于原始流量中长流没有被抽样流数的的影响很小,因此采用二项分布估计出短流没有被抽样的概率,然后,基于估计的短流分布推断出未被抽样的流数.本文将已抽样流和未抽样流分开统计,一方面大幅度提高了算法的运行效率,另一方面又能提高估计精度.在短流统计估计中采用了抽样概率补偿方法,对由于未被抽样的流和未在计算范围内的长流采用二项分布进行概率估计,采用概率补偿的机制对短流统计中的未被统计抽样的流进行补偿估计,避免在统计过程中出现的系统误差.

本文第 1 节将对流数估计相关方法进行讨论.第 2 节讨论本文提出的 Integral 算法.第 3 节给出 Iteration 算法.第 4 节对 Iteration 算法的相关参数进行讨论.第 5 节通过实验对算法进行分析.第 6 节总结全文.

1 相关研究

相关流数估计的算法主要有 3 个不同方向的研究:抽样报文估计、抽样流估计和数据流算法估计.其中,每个方向的代表作是:(1) Duffield[5]直接采用报文抽样的数据进行流数估计,本文的方法也是基于报文抽样数据的估计方法;(2) Hohn[6]采用流抽样数据进行流数估计;(3) Estan[7]采用数据流算法进行流数估计.

报文抽样的方法是测量器对通过的每个报文采用相同的概率进行独立抽样,报文抽样的优点是抽样方法简单、测量资源需求较少,目前的 Cisco 路由器的抽样 Netflow[1]功能就是采用报文抽样方法.被抽样的报文进行组流以采用如 EM 算法[5,8,9]进行流数估计.EM 算法本质上是一种迭代方法,主要用来求后验分布的极大似然估计,它的每一次迭代由两步组成:E 步(求期望)和 M 步(极大化).EM 算法的主要缺点是具有较高的时间复杂度,同时估计精度较差.目前,用估计流长分布的 EM 算法对原始未抽样流数进行估计,实际操作过程将耗费时间巨大.为了提高抽样报文统计估计精度,Duffield[5]和 Tune[10]分别提出了采用 TCP(transmission control protocol)信息以辅助估计流分布的方法.其中,Duffield[5]方法主要是利用抽样报文中的 TCP SYN 信息以辅助流分布估计.该方法是假设每个 TCP 流正好有一个 SYN 报文.但是这种方法的缺点是部分 TCP 流有多个 SYN 报文,而有些 TCP 流在测量时间段内没有出现 SYN 报文.由于 UDP 报文中没有 SYN 信息,因此 TCP SYN 的方法没有办法估计 UDP 流的分布,同时,网络中 UDP 流的比重越来越大.Tune[10]主要是采用 Fisher 信息理论来分析比较在使

用 TCP 的 SYN,SEQ 等不同 TCP 信息情况下,对 TCP 流估计可能会产生的误差.Tune[10]并没有给出具体的流分布估计算法,同时,Tune[10]在对实际互联网数据进行实验分析时,假设原始流数是已知数.

基于流抽样的方法是测量器对每个通过的流采用相同的概率进行独立抽样,因此对于每个流,要么所有的报文被抽样,要么该流中没有一个报文被抽样,因此,原始流量的流数非常简单地使用抽样流数除以抽样概率[6,11,12].Tune[10]算法由于需要测量出每个 SYN 报文,并且对每个 SYN 报文进行单独抽样,因此本质上该算法实际上是流抽样算法而不是报文抽样算法.基于流抽样方法估计的流数精确,也非常简单,但是测量器对于读取通过每个报文的流标识,需要较高的测量资源,目前还没有一个路由器或测量器采用流抽样方法进行流量抽样.

基于数据流方法[7,13]对每个报文采用哈希函数进行处理,将哈希值记录在一个小哈希空间中,测量结束后,对哈希空间中的值采用统计方法进行估计.数据流方法的优点是需要较少的内存空间,同时具有较高的估计精度;但是由于需要对每个报文进行复杂的操作,需要消耗较多的测量资源.

2 积分推断法(Integral)

流标识定义为源 IP、宿 IP、协议、源端口、宿端口五元组或其子集,如流标识可以定义为上述五元组.一条由 n 个报文构成的流,n 个报文中每个报文按照随机抽样概率 p 独立抽样,则抽样后报文数为 k 的概率服从二项分布(binomial distribution),即经过 n 次独立重复实验中抽取 k 个报文的概率是二项分布公式(1):

$$p(X=k)=\binom{n}{k}p^k(1-p)^{n-k}=b(k;n,p),k=0,1,...,n \tag{1}$$

其中,X 为在 n 次贝努里实验中出现成功的次数;$\binom{n}{k}$ 表示在 n 次抽样实验中抽样到 k 个报文的各种组合情况,称为二项系数(binomial coefficient).假设一个原始流流长 j,按照报文抽样概率 p 抽样后有 i 个报文被抽样,其抽样概率为 $b(i;j,p)$,原始流长 j 的流数可表示为 f_j,其抽样后流长为 i 的流数有 $g_{ji}=f_j\cdot b(i;j,p)$,则抽样流量中流长为 i 的流数有 $g_i=\sum_{j\geqslant i}(f_j\cdot b(i;j,p))$, 抽样流长为 1 的流数为 $g_1=\sum_{i\geqslant 1}i\cdot f_i\cdot p\cdot(1-p)^{i-1}$, 由此,

$$g_1/p=\sum_{i\geqslant 1}i\cdot f_i\cdot(1-p)^{i-1}.$$

我们假设该式是一个对于 p 的连续函数,在等式两边都对 p 进行积分可以得到:

$$g_1\cdot\ln(p)=-\sum_{i\geqslant 1}f_i\cdot(1-p)^i+C.$$

未抽样(即流长为 0)的流数为 $g_0=\sum_{i\geqslant 1}f_i\cdot(1-p)^i$. 由此,我们得到未抽样的流数的一个推断,可以表示为

$$g_0=-g_1\cdot\ln(p)+C,$$

其中,C 为常数.

我们知道,当 p=1 时,所有的报文都被抓取,没有任何流丢失;当 p=1 时,g_0=0,C=0,$g_0=-g_1\cdot\ln(p)$,因此流量的总流数可以估计为公式(2):

$$\hat{m}=-g_1\cdot\ln(p)+\sum_{i=1}g_i \tag{2}$$

当然,公式(2)的推断基于 p 是连续变量,实际上并不一定符合连续变量的假设.另外,g_1 是在抽样概率为 p 的结果,而常数 C=0 是根据 p=1 计算的,因此,公式(2)估计流数会存在较大的误差.公式(2)的优点是,我们仅仅需要抽样概率和 1 个报文抽样流的流数这两个参数就可以很容易地估计出原始流数的大致信息,相对于直接采用抽样流数的方法,估计精度要高很多.在本文第 5 节的实验中我们将发现,该算法对流数的估计精度接近 EM 算法的估计精度,但是与 EM 算法相比,该算法所使用的测量资源基本可以忽略不计.

3 迭代算法(Iteration)

3.1 算法描述

如果一条流中没有一个报文被抽样,则这条流未被抽样.根据公式(1),j 个报文中没有一个报文被抽样的概

率为 $g_{j0}=(1-p)^j$.假设已知原始流长 j 的流数为 f_j,根据公式(1),按照概率 p 抽样后,没有被抽样的流数均值为 $E(g_{j0})=f_j\cdot(1-p)^j$,因此,对于所有流量,没有被抽样的流数均值为公式(3):

$$E(g_0)=\sum_{j\geqslant 1} f_j\cdot(1-p)^j \tag{3}$$

原始流长 j 按照抽样概率抽样到 i 报文的概率服从公式(1),已知原始流长 j 的流数为 f_j,可以计算出原始流长 j 的流被抽样到 i 的流数 $E(g_{ji})=f_j\cdot b(i;j,p)$,因此,对所有原始流量抽样后被抽样到 i 的流数均值为

$$E(g_i)=\sum_{j\geqslant i} E(g_{ji})=\sum_{j\geqslant i} f_j\cdot b(i;j,p).$$

原始流量中的流数 m 等于所有可能被抽样流数的总数之和,即 $\hat{m}=\sum_{i\geqslant 0}E(g_i)=E(g_0)+\sum_{i>0}E(g_i)$. 其中,在抽样流量中,$E(g_0)$是未知的,$\sum_{i>0}E(g_i)$ 等于已知的抽样流量中所有流的流数 $\sum_{i>0}g_i$ 的总和.

由于测量后仅知道报文抽样概率 p 和抽样流量流长 i 的流数 $g_i(i>0)$,那么,如何根据这些已知信息推断出未抽样的流数 $E(g_0)$?由公式(3)可知,$E(g_0)$是由所有 $f_j(j>0)$计算得到的,但原始流量分布 $f_j(j>0)$是一组未知数值.由 $E(g_i)(i>0)$的一次抽样测量的结果 g_i,实际上该问题就变成了通过 g_i 推断出 f_j.由于无法直接从 g_i 推断出 f_j,因此我们假设一个原始流的分布 $\hat{f}_j$,由 $\hat{f}_j$ 生成一个抽样分布 $\hat{g}_i$ 公式(4):

$$\hat{g}_i=\sum_{j\geqslant i}\hat{f}_j\cdot b(i;j,p) \tag{4}$$

如果对于所有的 $|g_i-\hat{g}_i|<\delta, i>0$,则认为假设的原始流分布 $\hat{f}_i$ 等于实际流长分布;否则,将采用 g_i 对 $\hat{f}_j$ 进行修正,反复迭代直到满足条件(5)为止:

$$|g_i-\hat{g}_i|<\delta,\ i>0 \tag{5}$$

该算法的基本特点是,将原始流量的流数估计分成未抽样的流数和已抽样的流数,对抽样流量进行统计,计算出已抽样流量部分;未抽样部分的流数的推断是采用二项分布估计出短流没有被抽样的概率.由于原始流量中长流没有被抽样的概率非常低,直接将长流忽略;而对于短流分布,采用递归的方法估计.

3.2 流长修正方法

如果有 $|g_i-\hat{g}_i|<\delta, i>0$ 条件不满足,则需要对假设的原始流长分布 $\hat{f}_j(j>0)$ 进行修正计算.根据 $\hat{g}_i=\sum_{j\geqslant i}\hat{g}_{ji}=\sum_{j\geqslant i}\hat{f}_j\cdot b(i;j,p)$,将 g_i 和 $\hat{g}_i$ 之间的差距按照 $\hat{f}_j$ 抽样到 $\hat{g}_{ji}$ 等比率分配,即 $\hat{g}'_{ji}=\hat{g}_{ji}\cdot g_i/\hat{g}_i$,计算出所有的修正结果 $\hat{g}'_{ji}, j\geqslant i, i>0$. 根据 $\hat{g}'_{ji}$ 计算出新的原始流估计值 $\hat{f}_j$,假设抽样流长最长为 $i_{\max}$,如果 $j\leqslant i_{\max}$,则原始流长 j 抽样到 $i\in[1,j]$的所有流数修正后的结果为 $\hat{f}'_j=\sum_{i=1}^{j}\hat{g}'_{ji}$,其抽样概率累加和为 $p_j=\sum_{i=1}^{j}b(i;j,p)=1-b(0;j,p)$. 如果 $j>i_{\max}$,则原始流长 j 抽样到 $i\in[1,i_{\max}]$的所有流数修正后的结果为 $\hat{f}'_j=\sum_{i=1}^{i_{\max}}\hat{g}'_{ji}$,其抽样概率累加和为 $p_j=\sum_{i=1}^{i_{\max}}b(i;j,p)$. 修正后的原始流长 j 的流数 $\hat{f}_j=\hat{f}'_j/p_j, j>0$. 将修正后的 $\hat{f}_j$ 代入公式(4)继续计算,直到条件(5)满足为止.

该修正方法具有两个特点:(1) 将二项分布理论抽样 i 后的结果和实际抽样 i 流数之间的差按照各假设原始流长抽样到 i 的比例进行修正;(2) 将原始流长 j 抽样到各个抽样流长后的修正结果按照抽样比例进行概率补偿,实现对抽样流量中不能反映的未抽样流量以及超过最长抽样流长的未知抽样流量进行概率补偿.

综合第 3.1 节的公式(4)和公式(5),可以推导出短流估计的 Iteration 迭代算法公式为公式(6):

$$\hat{f}_j^{(k+1)}=\left\{\hat{f}_j^{(k)}\cdot\sum_{i=1}^{\min(j,i_{\max})}\frac{b(i;j,p)\cdot g_i}{\sum_{j\geqslant l\geqslant 1}\hat{f}_j^{(k)}\cdot b(l;j,p)}\right\}\Bigg/\left\{\sum_{i=1}^{\min(j,i_{\max})}b(i;j,p)\right\} \tag{6}$$

由于 Iteration 算法的迭代过程中主要是根据所分配的抽样流量中的分配比重进行迭代,可能会导致迭代后的结果远远偏离实际情况,其中容易出现的情况是,后面中长流的流数迭代值超过了前面的中短流.由于网络流

量的分布基本规律是服从重尾分布,重尾分布的特征是短流的流数大于长流的流数.基于这个特征,我们规定迭代过程中长流的流数只能小于或等于其前面的流数,避免长流的流数大于短流的流数情况的出现.假设流长 i 的第 k 次迭代估计值是 $\hat{g}_i^{(k)}$, 流长 i+1 的第 k 次迭代估计值是 $\hat{g}_{i+1}^{(k)}$,如果 $\hat{g}_{i+1}^{(k)} > \hat{g}_i^{(k)}$, 则设置 $\hat{g}_{i+1}^{(k)} = \hat{g}_i^{(k)}$.

3.3 估计误差

假设知道原始流长 j 流数的估计为 $\hat{f}_j$,以报文抽样概率 p 理论抽样到流长 i 的概率服从二项分布 $b(i;j,p)$, $b(i;j,p)$表示原始流量流长 j 以报文抽样概率 p 抽样到流长 i 的概率,则估计原始流量流长 j 以报文抽样概率 p 抽样到流长 i 的流数为 $g(j,i)=b(i;j,p)\cdot\hat{f}_j$, 其中,$i\in[1,n]$,$j\in[1,m]$.计算所有理论抽样到流长 i 的流数累加和 $\hat{g}_i=\sum_{j=1}^{m}g(j,i), i\in[1,n]$, 如果该假设流数估计分布接近实际分布,则按照二项分布方法抽样后的分布应该接近实际分布,即 $g_i\approx\hat{g}_i$. 已知流长 j 抽样后的二项分布的均值为 $E\xi=j\cdot p$,其方差为 $D\xi=j\cdot p\cdot(1-p)$,累加和 $\hat{g}_i$ 的方差估计为 $D(\hat{g}_i)=D\left(\sum_{j=1}^{m}\hat{f}_j\cdot b(i;j,p)\right)=(p\cdot(1-p))\cdot\sum_{j=1}^{m}\hat{f}_j\cdot j=\Delta\cdot(p\cdot(1-p))$.其中, $\Delta=\sum_{j=1}^{m}\hat{f}_j\cdot j$; 在 $\hat{g}_j\approx g_j$ 情况下,Δ 是常数.

4 算法参数

4.1 流长参数

网络流量具有以下两个特征:(1) 网络流长分布服从重尾分布,其基本特点就是短流的数量远多于长流的数量;(2) 对于一个抽样概率,短流未被抽样的概率大于长流未被抽样的概率.假设一个流长 i,该流未被抽样的概率服从二项分布$(1-p)^i$.基于上述流量的两个特征,如果一个流未抽样的概率小于Δ,考虑流量的重尾分布特性,则流长小于未抽样概率为Δ的流将不再考虑其对未抽样流数的影响.$(1-p)^i<\Delta$,$i>\ln(\Delta)/\ln(1-p)$,i 取值大于等于 $\ln(\Delta)/\ln(1-p)$的最小整数.如果Δ=10%,则当抽样概率 p 取值 10%时,i=22;p=1%,i=230;p=1/256,i=589.

流长为 i 的流被抽样到流长为 j 的流的概率为 $p_{ij}=\binom{i}{j}p^j(1-p)^{i-j}$, 该流被抽样到所有小于等于流长 j 的概率为 $\sum_{k=0}^{j}p_{ik}=\sum_{k=0}^{j}\binom{i}{k}p^k(1-p)^{i-k}$.

如果流长为 i 的流被抽样到小于等于流长 j 的概率需要大于β,即 $\sum_{k=0}^{j}\binom{i}{k}p^k(1-p)^{i-k}>\beta$, 则认为流长为 i 的流以概率 p 抽样后,该流的流长小于等于 j,大于 j 的抽样流可以忽略.如果设β=99%,则按照上面的取值,当 p 取值为 10%时,i=22;p=1%,i=230;p=1/256,i=589 等情况下,j 的取值都等于 6.

下面考虑影响抽样流长j 的原始流长.如果一个原始流长 m 的流,其抽样到流长小于等于 j 的概率累加和小于一个阈值α,即 $\sum_{k=0}^{j}\binom{m}{k}p^k(1-p)^{m-k}<\alpha$, 则认为流长 m 及其以上的流为长流,这些长流对 0~j 的抽样流长的影响可以忽略不计.同样,设置α=0.1,则对于抽样概率 10%,1%,1/256,j=6,按照公式 $\sum_{k=0}^{j}\binom{m}{k}p^k(1-p)^{m-k}<\alpha$ 可以计算出其 m 值分别为 103,1 052 和 2 695.下面针对实际网络流量测量数据,分析其中的流长测量参数.

(1) 影响未抽样流的最大原始流长 i_1

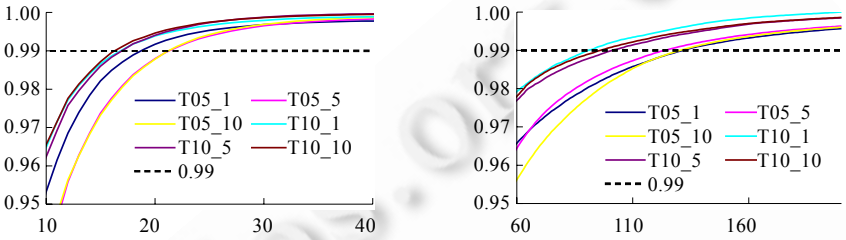
图 1 的 x 轴是原始流长,y 轴是没有被抽样的流数占所有没有被抽样的流数的比重累加和,图 1 分别比较第 5.1 节表 1 中 6 种不同的流量日志的情况.图 1(a)的抽样概率是 10%,由图 1(a)可知,在原始流量中,流长小于 20 的所有流量按照 10%抽样概率占了未抽样流数的 99%以上.因此,为了估计 1/10 抽样概率未被抽样流数,如果参数设置为 99%,则最大原始流长 $i_{\max}$ 仅需设置为 20.图 1(b)的抽样概率为 1%,由图 1(b)可知,原始流量中小于 140 的流长的所有流量按照 0.01 抽样概率占了未抽样流数的 99%以上,则影响未抽样流的最大原始流长 i_1 取值 140.

(2) 最大抽样流长 j_{max}

在图 2 中,x 轴是抽样流长,y 轴是原始流抽样到抽样流长的累加概率和,sample_10 曲线表示原始流长为 20 的流按照 0.1 的概率抽样后的累加概率和,sample_100 曲线标识原始流长为 140 的流按照 0.01 的概率抽样后的累加概率和.由图 2 可知,对于 0.1 和 0.01 抽样概率的最大流长,99%流数抽样后小于等于 6,也就是说,99%以上的信息落在抽样流长 6 以内.因此,我们选择的最大抽样流长 j_{max} 取值为 6,以保证 99%流数的覆盖率.

(3) 影响短流的最大原始流长 i_{max}

根据上面的分析可知,对未抽样流数影响累计和超过 99%的原始流长对于 10%抽样概率的原始流长最长在 20,而 1%抽样概率的原始流长在 140.对于未抽样影响的流,99%的流数落在抽样后流长 6 以内.因此,为了考虑未抽样的流数,我们只需测量分析抽样流长在 6 以内的流量分布情况.图 3 的 x 轴表示原始流长,y 轴表示以 1%的抽样概率后,抽样流长为 6 的流来自原始流长的累加百分率.图中的水平虚线表示累加百分率 0.95.由图 3 可知,1%抽样概率的抽样流长为 6 的 95%以上的流数来源于小于 1 000 的原始流长.



(a) 抽样概率 10% (b) 抽样概率 1%

Fig.1 Original flow length and the unsampled probability of the flow

图 1 原始流长和未被抽样流的概率

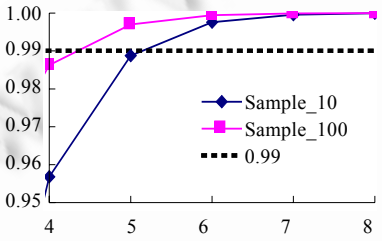


Fig.2 The maximal flow length of the sampled short flows

图 2 抽样短流最大流长

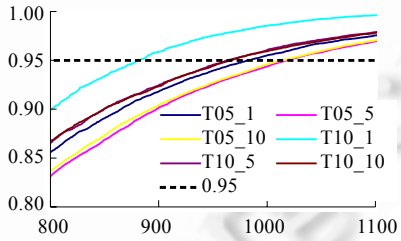


Fig.3 Short flows affected by the maximal original flow length

图 3 影响短流的最大原始流长

4.2 初始值参数

原始流流长分布初始值的选取对流数估计精度的影响较大,选择不合适的初始值可能导致最终的结果不收敛.初始值的选取越接近真实结果,最终估计的流长分布的结果也就越好.一个简单初始值的方案是设置抽样流长分布等于初始原始流长分布,但是在抽样流量中有些抽样流长没有流数为 0,如果直接使初始原始流量流长的流数也等于 0,则根据公式(6),无论经过多少次迭代,该流长的估计流数都等于 0.为了避免初始流长出现 0 的情况,对于某个初始流长为 0,将该初始流长赋值为 1.

下面我们采用 T10_5 的数据进行 1/100 抽样,对不同的初始值进行对比.我们选择的初始值分布分别是:(1) 所有流长均有相同的流长分布(average);(2) 流长分布服从λ=1 的泊松分布(Poisson);(3) 流长分布等于抽样流长分布(sample);(4) 流长分布等于抽样流长分布,但是流长为 1 的流数减少 50%.所有的算法迭代次数设定为 50 次,分别比较估计的原始流数的误差.

图 4 的 x 轴表示 4 个不同的初始值分布,y 轴表示采用 Iteration 算法对于每种初始值估计所产生的流数估计误差.由图 4 可知,选择不同的初始值,对流数的估计精度误差从超过 50%到小于 5%以内.同时,由图 4 可知,直接采用抽样流量作为初始分布估计,效果好于 Poisson 和均匀分布的初始分布,但是抽样分布效果差于对流长为 1 的流数按照抽样流数的 50%作为流长为 1 的流数初始值.

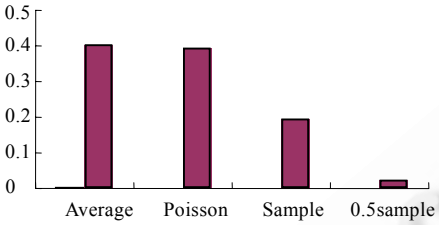


Fig.4 Comparison of the initial parameters

图 4 初始参数比较

抽样流长分布接近原始流长分布,因此可以使用抽样流长分布作为原始流流长初始值,可以得到较为理想的结果.图 5 是流长在 1~100 之间的抽样流数和原始流数的各流长所占比例的关系,其中,x 轴是流长,y 轴表示对应流长的流数在所有流数中所占的比重.从图 5 中可以看出,抽样流中流长为 1 的流数高于原始流长中的比例,其他流长分布接近于原始流长的流量分布.在该实例中,抽样流量中流长为 1 的占总抽样流数的 40%,而原始流量中流长为 1 的占总流数的大约 20%.如果我们将抽样流长为 1 的初始流数减少 50%,则是图 5(b).可以看出,抽样流长分布与原始流长分布非常接近.由图 4 同时可知,0.5sample 的结果远好于 sample 的结果.

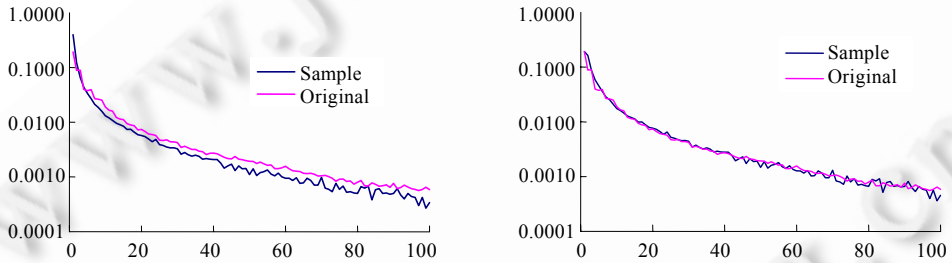


Fig.5 Proportion relationship between the sampled flow number and the original flow number of the flow length from 1 to 100

图 5 流长在 1~100 之间的抽样流数和原始流数的各流长所占比例的关系

从上面的分析来看,初始值的选择对最终结果和循环迭代次数有直接影响.由于对流量抽样以后,流长分布向仅 1 个报文的流集中,也就是说,抽样比率越大,则其中仅 1 个报文的流所占的比重也越大,同时,仅 1 个报文的流对非抽样的流数贡献最大.如果直接采用抽样流长分布作为原始流长分布的初始值假设,则最终估计的结果将使得流长为 1 的流数的比重远高于实际值,确定初始流长为 1 的流数将对最终估计精度存在直接影响.但是在实际测量过程中没有在线的流长分布,为了简单处理,仅仅根据抽样流量分布,采用抽样流长为 1 的流数的 50%作为仅 1 个报文的流的初始值.

4.3 循环结束判断

设一个流有 j 个报文,其报文以概率 p 抽样 i 个报文的概率服从二项分布 $b(i;j,p)$,对于有 f_j 个流长为 j 的流抽样到流长为 i 的流服从以概率 $b(i;j,p)$的二项分布,其抽样到流长为 i 的流的均值为 $\overline{g}_{ji} = f_j \cdot b(i;j,p)$, 其抽样的

方差为 $\sigma_{ji}^2 = f_j \cdot b(i;j,p) \cdot (1-b(i;j,p))$. 因此,所有原始流抽样到流长为 i 的流数均值为 $\bar{g}_i = \sum_{j=i}^{\max} f_j \cdot b(i;j,p)$, 其抽样方差为 $\sigma_i^2 = \sum_{j=i}^{\max} f_j \cdot b(i;j,p) \cdot (1-b(i;j,p))$. 实际抽样流长 i 的流数 g_i 服从均值为 $\bar{g}_i$ 方差为 σ_i 的正态分布 $N(\bar{g}_i,\sigma_i)$, 对于给定的置信度 $100(1-\alpha)\%$有 $P\{-z_{\alpha/2} < (f_i - \bar{f}_i)/\sigma_i\} = 1-\alpha$. 由不等式 $-z_{\alpha/2} < (g_i - \bar{g}_i)/\sigma_i < z_{\alpha/2}$ 推导出 $\bar{g}_i - z_{\alpha/2}\sigma_i < g_i < \bar{g}_i + z_{\alpha/2}\sigma_i$, 故所求的置信度 $100(1-\alpha)\%$的置信区间为 $(\bar{g}_i - z_{\alpha/2}\sigma_i, \bar{g}_i + z_{\alpha/2}\sigma_i)$.

如果取α=0.05,则 $z_{\alpha/2}=z_{0.025}$=1.96.

如果实际抽样的流长 i 的流数 g_i 落在抽样区间 $(\bar{g}_i - 1.96\sigma_i, \bar{g}_i + 1.96\sigma_i)$, 则表示假设的原始流长分布对于抽样流长 i 的分布服从假设检验,不需要对抽样流长 i 的流分布进行调整;如果 $g_i \notin (\bar{g}_i - 1.96\sigma_i, \bar{g}_i + 1.96\sigma_i)$, 则需要对落在抽样流长 i 的各原始流数进行重新分配.如果所有抽样流长的流数均落在置信区间中,则推断该假设原始流长分布以 95%的置信度服从原始流长分布.

5 实验分析

5.1 实验数据

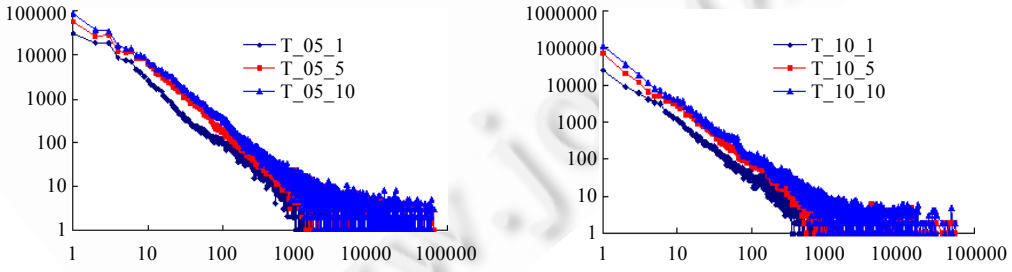
表 1 中的实验数据来自 CERNET(China education and research network)华东北主干网上 2005 年 10 月 11 日和 2010 年 6 月 6 日的 Trace 全报文数据,其中,2010 年的数据是按照源宿 IP 地址异或计算后取其中的 1/4 流量,实际上是按照流抽样方式抽样 1/4 流量.这两组数据分别取 1 分钟、5 分钟、10 分钟对应为 2.4GB,12GB,24GB 的数据并以 1/10,1/100 分别进行抽样,利用本文所提出的 Integral 和 Iteration 的流数估计算法进行估计计算.其中,表 1 中的流数是指在 1 个时间粒度内,具有不同源宿 IP 地址集合的数量.

Table 1 Packet trace datasets in the experiment

表 1 实验报文数据

序号	采集时间	数据名称	时间粒度(分钟)	报文数	流数
1	2005.10.11	T05_1	1	37 936 657	642 213
2	2005.10.11	T05_5	5	188 921 329	1 840 592
3	2005.10.11	T05_10	10	383 533 834	3 132 344
4	2010.6.6	T10_1	1	9 899 685	315 019
5	2010.6.6	T10_5	5	48 487 970	991 283
6	2010.6.6	T10_10	10	93 874 607	1 787 016

图 6 给出了两个不同时间采集的数据的原始分布情况,其中,x 轴表示原始流长,y 轴表示所对应原始流长的流数在所有流数中所占的比重.图 6 表明,表 1 中的 6 组数据的流长分布均体现出明显的重尾分布特性.



(a) 2005 年数据的原始长度分布 (b) 2010 年数据的原始长度分布

Fig.6 Original length distribution of the traffic traces

图 6 流量数据的原始长度分布

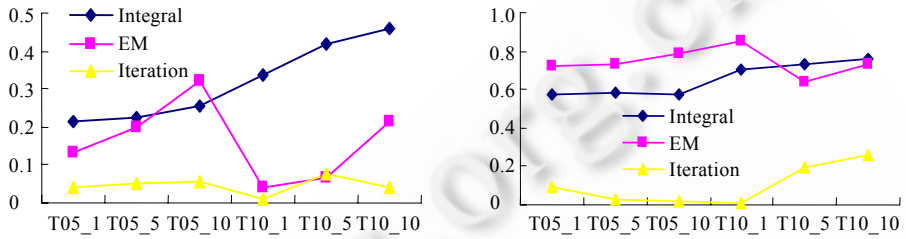
下面给出评价实验性能的测度指标:流数相对误差测度 FNE, $FNE = |\hat{M} - M| / M$, 其中, $\hat{M}$ 是采用抽样估计

算法所估计的流数,M 是实际测量的流数.定义的 FNE 反映估计的流数与实际流数之间的相对误差.

5.2 流数估计算法之间的比较

本文的 Iteration 算法抽样流长取最大值为 6,原始流长 1/10,1/100 抽样比率的最大原始流流长分别取值为 103 和 1 052.由于 Duffield[5]的 EM 算法没有定义具体的收敛条件,仅仅是事先规定循环次数作为 EM 算法循环结束的标志,在本节的比较中,EM 算法的循环次数采用本文的 Iteration 算法的循环次数.两种算法的初始流数分布采用抽样流长.

图 7 是采用表 1 的数据对 3 种算法抽样精度与原始流量大小之间关系的比较,其中,图 7(a)采用 1/10 抽样比率参数,图 7(b)采用 1/100 抽样比率参数,分别采用 EM 算法、Integral 算法、Iteration 算法流数估计的相对误差.在图 7 中,x 轴对应表 1 的 6 种数据集,y 轴表示流数估计相对误差.



(a) 1/10 抽样数据估计 (b) 1/100 抽样数据估计

Fig.7 Comparison of the relationship between the sampling accuracy of three algorithms and the original traffic size

图 7 3 种算法抽样精度与原始流量大小之间关系的比较

由图 7(a)可知,除了 T10_5 数据采用 EM 算法估计误差 6.6%,Iteration 算法的估计误差是 7.9%以外,其他数据的 EM 估计误差超过 Iteration 算法的估计 3 倍以上.总体上,Integral 算法估计精度较差.图 7(b)中所有数据的 Iteration 算法估计误差比 EM 估计误差精确 3 倍~10 倍左右,同时,Integral 算法精度也略高于 EM 算法,这表明 EM 算法受抽样比例影响较大.分析其原因是,由于从抽样概率理论上看,一般最小抽样样本阈值在 30,抽样样本数量小于一定阈值以后,概率理论推断方法将失效.本文将对最小抽样阈值将不做讨论,具体问题可以参考概率论相关的教材.EM 算法由于需要对所有抽样流长进行处理,而大多数流长的抽样样本流数都小于抽样估计所需要的最小阈值,因此抽样概率越小,EM 估计算法将越差.而 Iteration 算法仅考虑抽样流长小于等于 6 个报文的流数,其样本的数量远远超过样本最小阈值.

图 7 也表明,在相同的抽样概率参数控制下,不同流数大小的数据日志对 Iteration 抽样估计算法的精度影响不大,如图 7(a)中的不同流量大小下的抽样估计精度基本上都在 5%,而图 7(b)的抽样精度在不同流量下的精度基本上在 10%左右.

在图 8 中,x 轴表示原始流量的流数,y 轴表示估计的流数.图 8 表明在不同流数大小下,Iteration 算法和 EM 算法在 1/10 抽样参数下估计流数和实际流数之间的对比关系,其中,Original 曲线表示原始流数曲线,EM 曲线表示采用 EM 算法估计的流数曲线,Iteration 曲线表示采用 Iteration 算法估计的曲线.从图 8 中可以看出:Iteration 算法估计的流数比较稳定,受原始流数影响不大;而 EM 算法随着流数影响波动较大,原始流数越大,其估计算法越不精确.

图 9 对 T10_10 数据在 1/10 抽样下分别采用 Iteration 算法和 EM 算法估计流长小于 103 的流数分布情况.流长为 103 是本文第 4.1 节所研究的 1/10 抽样时 Iteration 算法的最大原始流流长.由图 9 可知,Iteration 估计曲线更接近原始流长分布曲线.

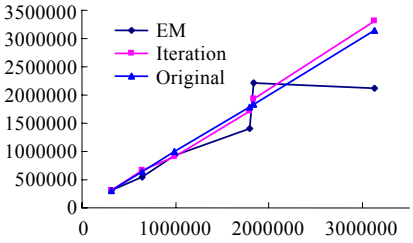


Fig.8 Relationship between the size of the number of flows and the estimation algorithms

图 8 流数大小和估计算法之间的关系

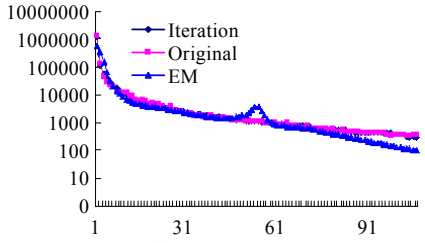


Fig.9 Comparison of the estimated length of the first 100 flows

图 9 前 100 个流长估计的比较

5.3 性能对比

Iteration 算法完成一次全部迭代时间复杂度为 $O(i_{\max}^2 j_{\max})$, EM 算法是 $O(I_{\max}^2 J_{\max})$. 以 T10_1 的 1 分钟的 Trace 全报文数据为例,按 1/10 抽样比例进行计算,实验中取ε=0.01,则 $j_{\max}$=5,$i_{\max}$=103,那么 Iteration 算法的 $i_{\max}^2 j_{\max}=103^2\cdot 5$, 而 EM 算法中则 $I_{\max}^2 J_{\max}=12320^2\cdot 1232$. Integral 算法不需要任何迭代计算,只需计算 1 次即可完成,需要的系统计算时间可以忽略不计.表 2 给出了 T10_1,T10_5,T10_10 这 3 组数据分别按 1/10 抽样,Integral 算法、EM 算法和 Iteration 这 3 种算法进行 3 次迭代,程序的耗时对比情况.从实验结果来看,Iteration 算法所需要的时间远远小于 EM 算法,同时我们还注意到,不同数量的流量日志对 Iteration 算法的计算时间几乎没有影响.其原因是,影响 Iteration 算法性能的 $j_{\max}$ 和 $i_{\max}$ 参数只受抽样参数影响,而与所需估计的流量大小无关.

Table 2 Computational comparison among the three algorithms

表 2 3 种算法运行时间对比

Trace	Integral 算法(ms)	Iteration 算法(ms)	EM 算法(s)
T10_1	<1	132	250
T10_5	<1	144	1 206
T10_10	<1	143	2 033

下面分析 3 种算法的内存占用情况.Integral 算法只需要流长为 1 的流数,其空间复杂度为 $O(1)$.由于 Iteration 算法和 EM 算法中主要内存占用是二项分布矩阵,因此 Iteration 算法的空间复杂度为 $O(i_{\max}j_{\max})$,EM 算法的空间复杂度是 $O(I_{\max}J_{\max})$.以前面的 T10_1 数据为例,$j_{\max}$=5,$i_{\max}$=103,那么 Iteration 算法的 $i_{\max}j_{\max}$=103·5=515,而 EM 算法中则 $I_{\max}J_{\max}$=12320·1232.分别采用 3 种算法对 T10_1 数据的 1/10 抽样样本处理下的使用内存情况进行比较,实验结果表明,Integral 算法内存使用小于 1KB,Iteration 算法内存使用小于 10KB,而 EM 算法的内存占用在 60MB.

6 结束语

本文的贡献主要表现在以下几个方面:(1) 将流数估计问题转换成未抽样流数的估计问题,对于已抽样的流数,直接从抽样数据中获得,流数估计误差仅仅来自于未抽样的流数估计.(2) 对未抽样流数的估计方法是,首先估计出对未抽样流有影响的中短流,而把几乎不影响未抽样流数的长流直接剔除.这样一方面可以提高估计算法的效率,减少因大流样本不足而产生的估计误差,另一方面,也可以降低长流对中短流流长估计所带来的误差.(3) 通过统计推断确定对未抽样流量影响最大的抽样短流和最大的原始流流长两个重要算法估计参数.

本文提出的 Iteration 算法与现有的 EM 算法相比,具有较高的执行效率和流数估计精度.同时给出了各类初始值及迭代循环结束的判断方法,使得该算法具有可操作性.该算法可以应用于路由器中对目前抽样 NetFlow 数据中的流数和主机数等估计,以推断网络规模的扩展程度和异常流量的扩展范围.Integral 算法虽然在算法估计精度上比 Iteration 算法和 EM 算法都差,但在时空复杂度上却是其他方法无法比拟的.

对于 Integral 算法,下一步工作将进一步考虑增加对抽样流长大于 1 的流数对算法公式的影响,以提高 Integral 算法的估计精度.对于 Iteration 算法,下一步工作将考虑将该算法应用到基于抽样 NetFlow 流的网络行为观测系统(network behavior observation system,简称 NBOS)中[4],同时,采用 Iteration 算法对网络主机数量等其他参数进行估计的研究方法.

References:

[1] Duffield N. Sampling for passive Internet measurement: A review. Statistical Science, 2004,19(3):472−498. [doi: 10.1214/088342304000000206]

[2] Packet sampling (psamp)—Charter. 2009. http://www.ietf.org/html.charters/psamp-charter.html

[3] Using NetFlow filtering or sampling to select the network traffic to track. 2009. http://www.cisco.com/en/US/docs/ios/ios_xe/netflow/configuration/guide/nflow_filt_samp_traff_xe.html

[4] IP trace distribution system. 2011. http://iptas.edu.cn/src/system.php

[5] Duffield N, Lund C, Thorup M. Estimating flow distributions from sampled flow statistics. IEEE-ACM Trans. on Networking, 2005,13(5):933−946. [doi: 10.1109/TNET.2005.852874]

[6] Hohn N, Veitch D. Inverting sampled traffic. IEEE/ACM Trans. on Networking, 2006,14(1):68−80. [doi: 10.1109/TNET.2005.863456]

[7] Estan C, Varghese G, Fisk M. Bitmap algorithms for counting active flows on high speed links. IEEE-ACM Trans. on Networking, 2006,14(5):925−937. [doi: 10.1109/TNET.2006.882836]

[8] Dempster AP, Laird NM, Rubin DB. Maximum likelihood from incomplete data via the EM algorithm. Journal of the Royal Statistical Society, Series B (Statistical Methodology), 1977,39(1):1−38.

[9] Liu WJ, Gong J, Ding W, Cheng G. Method for estimation of flow length distributions based on least square method. Journal of Southeast University (Natural Science Edition), 2006,34(5):467−471 (in Chinese with English abstract).

[10] Tune P, Veitch D. Towards optimal sampling for flow size estimation. In: Proc. of the Internet Measurement Conf. (IMC 2008). New York: Association for Computing Machinery, 2008. 243−255 [doi: 10.1145/1452520.1452550]

[11] Yang LL, George M. Sampled based estimation of network traffic flow characteristics. In: Proc. of the INFOCOM 2007. New York: Institute of Electrical and Electronics Engineers Inc., 2007. 1775−1783. [doi: 10.1109/INFCOM.2007.207]

[12] Kawahara1 R, Ishibashi K, Mori1 T, Kamiyama N, Harada S, Hasegawa H, Asano S. Detection accuracy of network anomalies using sampled flow statistics. Int'l Journal of Network Management, 2011,21(6):513−535. [doi: 10.1002/nem.777]

[13] Kumar A, Sung M, Xu J, Wang J. Data streaming algorithms for efficient and accurate estimation of flow size distribution. In: Proc. of the SIGMETRICS 2004/Performance 2004. New York: Association for Computing Machinery, 2004. 177−188. [doi: 10.1145/1005686.1005709]

附中文参考文献:

[9] 刘卫江,龚俭,丁伟,程光.一种基于最小二乘法的流长度分布估计方法.东南大学学报(自然科学版),2006,34(5):467−471.

程光(1973－),男,安徽黄山人,博士,教授,博士生导师,CCF 高级会员,主要研究领域为网络测量,网络管理,网络安全.
E-mail: gcheng@njnet.edu.cn

唐永宁(1968－),男,博士,助理教授,主要研究领域为大规模分布式系统,网络安全,流量分类,数据挖掘.
E-mail: ytang@ilstu.edu