Identifying Elephant Flows Through Periodically Sampled Packets

Tatsuya Mori
NTT Service Integration Labs
Waseda University
mori.tatsuya@lab.ntt.co.jp

Masato Uchida
NTT Service Integration Labs
uchida.masato@lab.ntt.co.jp

Ryoichi Kawahara
NTT Service Integration Labs
kawahara.ryoichi@lab.ntt.co.jp

Jianping Pan
NTT MCL
panjianping@acm.org

Shigeki Goto
Waseda University
goto@goto.info.waseda.ac.jp

ABSTRACT

Identifying elephant flows is very important in developing effective and efficient traffic engineering schemes. In addition, obtaining the statistics of these flows is also very useful for network operation and management. On the other hand, with the rapid growth of link speed in recent years, packet sampling has become a very attractive and scalable means to measure flow statistics; however, it also makes identifying elephant flows become much more difficult. Based on Bayes' theorem, this paper develops techniques and schemes to identify elephant flows in periodically sampled packets. We show that our basic framework is very flexible in making appropriate trade-offs between false positives (misidentified flows) and false negatives (missed elephant flows) with regard to a given sampling frequency. We further validate and evaluate our approach by using some publicly available traces. Our schemes are generic and require *no* per-packet processing; hence, they allow a very cost-effective implementation for being deployed in large-scale high-speed networks.

Categories and Subject Descriptors: C.2.3 [Computer-Communication Networks]: Network Operations—*Network monitoring*

General Terms: Measurement, Theory, Verification

Keywords: measurement, flow statistics, packet sampling, Bayes' theorem, the elephant and mice phenomenon

1. INTRODUCTION

As many measurement-based studies have revealed, flow statistics exhibit strong heavy-tail behaviors in various networks (including the Internet) [5, 8, 12, 15, 16]. This characteristic is often referred to as *the elephant and mice phenomenon* (a.k.a. *the vital few and trivial many rule*); i.e., most flows (mice flows) only have a small number of packets, while a very few flows (elephant flows) have a large number of packets. A noticeable attribute of elephant flows is that they contribute a large portion of the total traffic volume despite being relatively few in the number of flows. For example, for one trace used in this study, about 0.02% of all flows contributed more than 59.3% of the total traffic volume.

Thus, the impact of elephant flows on network performance is significant. This fact makes identifying these flows very important in developing traffic engineering schemes. In addition, knowing the statistics of such flows is also very useful for network operation and management. By quickly identifying elephant flows that are overwhelmingly consuming network resources, network operators can immediately take appropriate actions against individual hosts or networks generating these flows.

To identify elephant flows, traditionally we have to collect all packets in the concerned network, and then extract their flow statistics. As many previous studies have indicated, however, such an approach lacks of scalability [1–4, 7]. For very high-speed links (say, OC-192+), directly measuring all flows is beyond the capability of measurement equipments (i.e., the requirements for CPU power, memory/storage capacity and access speed are overwhelming). As a solution to this problem, recently packet sampling techniques have attracted much attention from both the industry and the research community. For instance, some modern routers have these functions embedded, e.g., NetFlow [11] and sFlow [14]. The Packet Sampling (psamp) Working Group [13] in IETF has been standardizing techniques related to packet sampling.

In this paper, we are particularly interested in the following problem: "*How can we identify elephant flows in sampled packets?*" When answering this question, we adopt the *simplest* form of packet sampling; i.e., the sampling process is completely flow-state independent, and per-packet processing such as flow lookup and packet hashing is totally unnecessary. This form of packet sampling can be easily achieved by using a very simple technique — *periodic sampling*. The purpose of adopting this approach is to reduce the implementation cost and the operation overhead. For ISPs operating large-scale networks with a variety of measurement equipments, a cost-effective implementation is considered crucial in practice.

The main contribution of our work is developing a frame-

IMC'04, October 25–27, 2004, Taormina, Sicily, Italy.

work to find the threshold of sampled packets for a single flow, which can determine whether the flow is an elephant flow in unsampled packets[1]. We show that such a threshold can be calculated by using Bayes' theorem. To do so, we introduce an *a priori* distribution of the number of per-flow packets (i.e., the distribution of the number of per-flow packets in unsampled packets). This approach is very flexible in making appropriate trade-offs between false positives (misidentified flows) and false negatives (missed elephant flows) with regard to a given sampling frequency. Based on our approach, we find, somehow to our surprise, that the thresholds calculated for a variety of *a priori* distributions are quite similar. This observation suggests that a calculated threshold for a network in a certain period can be applicable to the network and other networks for a long run, which also reduces the operation overhead of our schemes.

The remainder of the paper is organized as follows. Section 2 reviews some related work and compares with our work. Section 3 gives the definition of elephant flows. In Section 4, we describe how to identify elephant flows in periodically sampled packets. We also validate and evaluate our approach by using some public packet traces. Section 5 discusses the way of obtaining such an *a priori* distribution. The effectiveness of our approach in other networks is also discussed. In Section 6, we conclude this paper with a brief summary.

2. RELATED WORK

The problem addressed in this paper has also been discussed by Estan and Varghese [4]. The main idea of their approach is to focus on elephant flows and neglect numerous mice flows, which is quite similar to ours. They proposed two novel techniques, referred to as *sample-and-hold* and *multistage filters*, respectively. Both techniques improve the process of extracting statistics of elephant flows in high-speed networks, while still keeping the memory consumption reasonably low. The main difference of their approach from ours is its requirement for complex per-packet processing, which may increase its implementation cost and operation overhead. On the other hand, the advantage of our approach is due to its simplicity; since it has no requirement for per-packet processing, the implementation cost of our schemes will be much lower.

In another related work, Duffield et al. [3] investigated how to infer unsampled flow statistics (instead of identifying elephant flows) from sampled flow statistics. Their key idea is to use a scaling approach, which is based on the number of sampled SYN packets in TCP flows. Kumar et al. [7] proposed a new technique referred to as space-code Bloom filter (SCBF) for extracting per-flow statistics of traffic in high-speed networks. The key points of their approach are extending the traditional Bloom filter with multiple sets of hash functions and using multi-resolution sampling. Their approach can capture most flow statistics very well, while only requiring a small amount of memory resources. However, identifying elephant flows through their approach requires the identities (e.g., source IP addresses) of potential elephant flows; i.e., it requires *a priori* knowledge of the elephant flows first. Papagiannaki et al. [12] proposed a scheme to classify elephant flows based on both flow volume and time persistence. Their approach successfully isolates

[1]In this paper, unsampled and sampled packets refer to the packet trace before and after the packet sampling process.

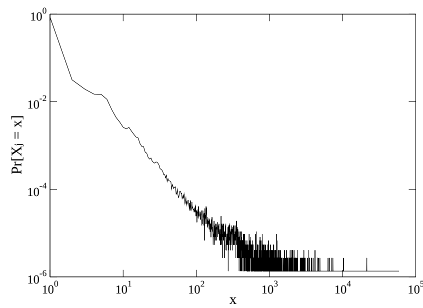

Figure 1: Probability density of X_j for the IPCL trace.

elephant flows that exhibit considerable persistence in time domain. Since their scheme assumes the direct measurement of all flows, some extensions will be required to obtain per-flow statistics in very high-speed networks. Golab et al. [6] proposed a deterministic algorithm to identify frequent items (similar to elephant flows in our context) using a memory-limited sliding window model. Although the algorithm can fulfill its objective with limited memory resources, it still requires per-packet processing, which we want to completely avoid in our context.

3. DEFINITION OF ELEPHANT FLOWS

A qualitative characterization of elephant flows is that they represent the majority of total traffic volume (in the number of packets or bytes) while being relatively few in the number of flows. The quantitative definition of an elephant flow can be arbitrarily determined by network operators according to their own criteria. In this paper, we define elephant flow as a flow that contributes more than 0.1% of all unsampled packets.

For illustration purpose, here we use a packet trace of Internet traffic measured at an OC-48c backbone link by the PMA project of NLANR[2]. In this trace, we considered the first 10^7 packets, referred to as the unsampled IPCL trace or simply trace when the context is clear, which correspond to about 137 seconds of observed traffic. The choice of 10^7 packets is considered reasonable for calculation convenience; the number of packets still allows us to obtain sufficient data for statistical analysis. Moreover, since the trace lasts about 137 seconds (with an average throughput of about 1 Gbps), the identified elephant flows within this time window can give meaningful information for the purpose of traffic engineering and network operation[3]. Throughout this paper, we define a flow by the 5-tuple identity (i.e., source/destination IP addresses, source/destination port numbers, and protocol identifier). Since the session time of some elephant flows can be very long (say, more than 3 hours), we include all partial flows in the time window (e.g., TCP flows with missing SYN or FIN packets) for analysis purpose. Our objective

[2]More precisely, we use the trace IPLS-CLEV-20020814-090000-0 [9].

[3]We confirmed that in many cases, elephant flows did exceed this time window; i.e., an identified elephant flow within the first time window will be recognized as an elephant flow again in the following time windows with high probability. That result is omitted due to space limit.

is to identify elephant flows by using packets sampled from the unsampled trace as quick and accurate as possible.

Figure 1 shows the probability density $\Pr[X_j = i]$ $(i = 1, 2, ...)$ of the IPCL trace, where X_j is the number of packets of the j-th flow. The probability density of X_j clearly decays in an approximate power-law fashion. As many previous measurement-based studies [5, 8, 12, 15, 16] have revealed, this characteristic seems to be intrinsic to Internet traffic. The IPCL trace contains 737,780 flows in the observed time window. Since an elephant flow by our definition is the one that contributes more than 0.1% of the total 10^7 packets, any flow j for which $X_j \geq 10^4$ is considered as an elephant flow in this paper. Under this definition, we have 167 elephant flows in the unsampled IPCL trace; these flows account for more than 59.3% of the total traffic volume (in the number of bytes).

4. IDENTIFYING ELEPHANT FLOWS

In this section, we propose an approach to identify elephant flows by counting the number of sampled packets for individual flows. Our task is to find a threshold determining whether a sampled flow represents an elephant flow in unsampled packets. Our approach is based on Bayes' theorem. Here, we assume that an *a priori* distribution $\Pr[X_j = i]$ is *known* in advance. How to obtain such a $\Pr[X_j = i]$ will be discussed in the next section. We first describe a framework for our approach. Then, we discuss the trade-off between false positives and false negatives with regard to a given sampling frequency. Based on these results, we give a procedure of identifying elephant flows in sampled packets. Finally, we present numerical results to validate and evaluate our approach.

4.1 A framework for our approach

Let n packets be randomly sampled from a population of N packets[4]. The sampling frequency f is defined as $f = n/N$. Let Y_j be the number of sampled packets for a flow j, which has X_j packets in the population (i.e., unsampled packets). Given $X_j = x$, the probability with which Y_j satisfies $Y_j = y$ is

$$\Pr[Y_j = y | X_j = x] = \binom{x}{y}\binom{N-x}{n-y}\bigg/\binom{N}{n}, \qquad (1)$$

which is a hyper-geometric distribution[5].

According to Bayes' theorem, given $Y_j \geq y$, the probability with which X_j satisfies $X_j \geq x$ can be calculated as follows.

$$\begin{aligned} &\Pr[X_j \geq x | Y_j \geq y] \\ = \; &\frac{\sum_{k=x} \Pr[Y_j \geq y | X_j = k] \Pr[X_j = k]}{\sum_{k=1} \Pr[Y_j \geq y | X_j = k] \Pr[X_j = k]}, \end{aligned} \qquad (2)$$

where $\Pr[Y_j \geq y | X_j = x] = 1 - \sum_{i=0}^{y-1} \Pr[Y_j = i | X_j = x]$, and $\Pr[X_j = i]$ is the probability density of X_j shown in Fig. 1.

[4]Precisely speaking, periodic sampling, which will be adopted in our scheme, is different from random sampling. However, since there is a large number of concurrent flows coexisting in a high-speed link, successive packets of a given flow will be interleaved by packets of many other flows, which effectively randomizes the selection of packets of the given flow [3].

[5]To calculate (1), we use its binomial approximation.

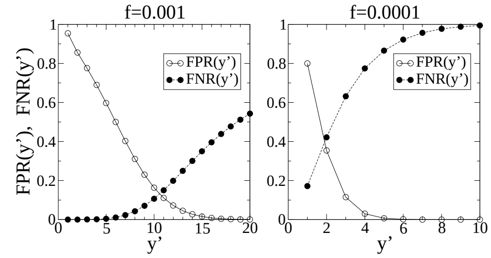

Figure 2: Trade-off between FPR(y') and FNR(y') for the IPCL trace.

Equation (2) means: given *a priori* distribution $\Pr[X_j = i]$, we can use y, the number of sampled packets for a flow, to calculate the probability with which the flow has more than x packets in the population. Let a flow j be an elephant flow if the number of unsampled packets $X_j \geq \hat{x}$; i.e., $\hat{x}$ is the threshold identifying an elephant flow (e.g., $\hat{x} = 10^4$ for the example in the previous section). Then, if $\Pr[X_j \geq \hat{x} | Y_j \geq \hat{y}]$ is close enough to 1 for a given $y = \hat{y}$, the flow j is very likely to have more than $\hat{x}$ packets in the population. In other words, if the number of sampled packets Y_j for a flow j is greater than $\hat{y}$, there is a high probability that the flow represented by the sampled packets is indeed an elephant flow in unsampled packets.

Intuitively, $\Pr[X_j \geq \hat{x} | Y_j \geq \hat{y}]$ will increase with $\hat{y}$. However, there is an intrinsic trade-off between false positives and false negatives, which prevents us from choosing an arbitrary $\hat{y}$. Our framework allows us to quantify this trade-off and make proper choices, as we shall see shortly.

4.2 Trade-off between false probabilities

If we assume that a flow for which Y_j satisfies $Y_j \geq y'$ is an elephant flow, then the false positive ratio (FPR) and the false negative rate (FNR) can be defined as follows.

$$\begin{aligned} \mathrm{FPR}(y') &\stackrel{def}{=} \Pr[\text{detected flows are not elephant flows}] \\ &= 1 - \Pr\left[X_j \geq \hat{x} | Y_j \geq y'\right] \end{aligned} \qquad (3)$$

$$\begin{aligned} \mathrm{FNR}(y') &\stackrel{def}{=} \Pr[\text{elephant flows are not detected}] \\ &= 1 - \Pr\left[Y_j \geq y' | X_j \geq \hat{x}\right] \\ &= 1 - \frac{\sum_{k=\hat{x}} \Pr[Y_j \geq y' | X_j = k] \Pr[X_j = k]}{\sum_{k=\hat{x}} \Pr[X_j = k]} \end{aligned} \qquad (4)$$

These equations can be calculated by using (2). Ideally, we should find $y' = \hat{y}$ such that both FPR($\hat{y}$) and FNR($\hat{y}$) are minimized. However, as shown below, there is a trade-off between the two metrics. The trade-off becomes particularly critical when the sampling frequency f is very low. Again, we illustrate this trade-off by using the IPCL trace.

Figure 2 shows FPR(y') and FNR(y') $(y' = 1, 2, ...)$ calculated by using (3) and (4). Here, the number of unsampled packets is $N = 10^7$, and the threshold to identify an elephant flow is $\hat{x} = 10^4$. Sampling frequency f is 10^{-3} and 10^{-4}, respectively. For $\Pr[X_j = i]$, we used the one from the IPCL trace, i.e., the distribution shown in Fig. 1. In both cases, an increase in y' leads to a decrease in FPR(y') and an increase in FNR(y'); i.e., there is an unavoidable trade-off between the two metrics. We can also see that the trade-off is more critical at the lower sampling frequency (i.e., $f = 10^{-4}$).

Table 1: $\hat{y}$, FPR($\hat{y}$), and FNR($\hat{y}$) for the IPCL trace by calculation

f	$\hat{y}$	FPR($\hat{y}$)	FNR($\hat{y}$)
10^{-3}	13	0.045	0.250
10^{-4}	4	0.030	0.774

4.3 Flow identification procedure

According to the result shown in the previous subsection, we have to make appropriate trade-offs between false positives and false negatives when identifying elephant flows. This subsection presents such a flow identification procedure designed with these observations kept in mind.

We consider the following policy as a guideline, i.e., *the false positive ratio should be reasonably low while reducing the false negative rate for a given sampling frequency.* According to this policy, we should keep the false positive ratio low enough without sacrificing the false negative rate too much, which is the consequence of the trade-off. Nevertheless, identifying elephant flows with a sufficiently low false positive ratio still provides very useful information for traffic engineering and network operation. First, we can avoid mistreating *non-elephant* flows (e.g., shaping their packet rate) when the false positive ratio is sufficiently low. Second, although keeping a low false positive ratio may cause a higher false negative rate, the amount of traffic generated by the identified flows is already significant. For example, in the IPCL trace, the 10 *heaviest* elephant flows account for about 10% of the total traffic volume.

According to the policy, our goal is to find $y' = \hat{y}$ such that FPR($\hat{y}$) becomes reasonably low. However, merely increasing $\hat{y}$ will lead to an increase in FNR($\hat{y}$), due to the revealed trade-off. So, we obtain the threshold $\hat{y}$ with the following constraint

$$\hat{y} = \min_{y'} \left\{ y' \mid \mathrm{FPR}(y') \leq \epsilon \right\}, \tag{5}$$

where ϵ specifies a tolerable false positive ratio. Equation (5) guarantees an FPR is lower than ϵ (say, 0.05), while keeping the corresponding FNR as low as possible.

Table 1 lists $\hat{y}$, FPR($\hat{y}$), and FNR($\hat{y}$) as calculated by following (5) for the examples shown in Fig. 2. Here, we have $\epsilon = 0.05$. When $f = 10^{-3}$, if $\hat{y}$, the number of sampled packets for a flow, is greater than 13, the flow is very likely to be an elephant flow, since FPR($\hat{y}$) $\leq \epsilon = 0.05$. Due to the fact that FNR($\hat{y}$) ≈ 0.25, about 25% of elephant flows in unsampled packets will be missed on average.

When $f = 10^{-4}$, FNR($\hat{y}$) becomes much higher if $\hat{y} = 4$; i.e., more elephant flows in unsampled packets will be missed due to an ultra-low sampling frequency and a bounded false positive ratio. However, the identified elephant flows are more likely to be elephant flows, which is also meaningful for traffic engineering and network operation. The result also suggests that for such a low f, it is advisable to allow a higher tolerable FPR, or to have a higher f when affordable, if false negatives have greater impact on network operation than false positives. Since the number of elephant flows in the IPCL trace is 167, we can expect that on average about 38 (i.e., $167 \times (1 - 0.774)$) *heavy* elephant flows will be identified according to our approach.

In a summary, the procedure for identifying elephant flows in sampled packets is enumerated as follows.

Step 1: Determine (i) the number of packets N in the population, (ii) a sampling frequency f, (iii) a threshold $\hat{x}$ determining an elephant flow in unsampled packets, (iv) an *a priori* distribution of X_j of unsampled packets (i.e., $\Pr[X_j = i]$), and (v) a threshold ϵ that FPR should satisfy.

Step 2: Calculate $\hat{y}$ by using the above (i) – (v) and (5).

Step 3: Conduct periodic packet sampling, and meanwhile count the number of per-flow packets.

Step 4: If the number of sampled packets for a flow is greater than $\hat{y}$, the flow is identified as an elephant flow.

Table 2: FPR and FNR through packet sampling of the IPCL trace

f	$\hat{n_e}$	n_e	FPR	FNR
10^{-3}	134	127	0.053	0.240
10^{-4}	38	38	0.000	0.772

4.4 Performance evaluation

In this subsection, we evaluate our approach through an actual packet sampling process for the IPCL trace. Following the procedure described in the previous subsection, we first determine the following parameters (ref. **Step 1**): $N = 10^7$, $f \in \{10^{-3}, 10^{-4}\}$, and $\hat{x} = 10^4$. We again use the distribution of the IPCL trace shown in Fig. 1 as the *a priori* distribution of X_j. Second, we use the calculated $\hat{y}$ listed in Table 1 (**Step 2**). We then conduct the actual packet sampling for the IPCL trace and count the number of per-flow packets (**Step 3**). We use periodic sampling, which is the simplest form to implement; i.e., we periodically sample every (N/n)-th packet. Finally, we investigate the flows for which the sampled packets satisfy $Y_j \geq \hat{y}$ (**Step 4**). The number of identified elephant flows in sampled packets is denoted as $\hat{n}_e$. Among these $\hat{n}_e$ flows, there are n_e actual elephant flows by our definition. We use N_e to denote the number of total elephant flows in unsampled packets (recall that $N_e = 167$ for the IPCL trace, as stated in Section 4).

Here, the false positive ratio FPR and false negative rate FNR can be approximated by FPR $= 1 - n_e/\hat{n}_e$, FNR $= 1 - n_e/N_e$.

FPR and FNR through packet sampling of the IPCL trace are listed in Table 2. As we can see, they are in good agreement with those listed in Table 1, which were calculated from (3) and (4), respectively. This result confirms that our approach can effectively identify elephant flows in sampled packets. With a periodic sampling, such an approach is also very efficient.

5. A PRIORI DISTRIBUTION OF X_j

Identifying elephant flows accurately requires an appropriately chosen $\hat{y}$. We have shown that such a $\hat{y}$ can be obtained by following (5). To do so, we need know N, f, and $\hat{x}$, and $\Pr[X_j = i]$; i.e., we have to obtain the distribution of X_j in unsampled packets. There are three possible approaches: A) use the distribution measured previously at the same link; B) infer the unsampled flow statistics (e.g., per-flow packet distribution) from sampled flow statistics; C) utilize the power-law characteristic of Internet traffic.

Approach A assumes that the distribution is similar if measured at the same link in similar periods (say, in daily busy hours). However, this approach lacks scalability, since it requires the direct measurement of flow characteristics. Approach B uses the result given by Duffield et al. [3]; i.e., with the estimated number of original flows, the unsam-

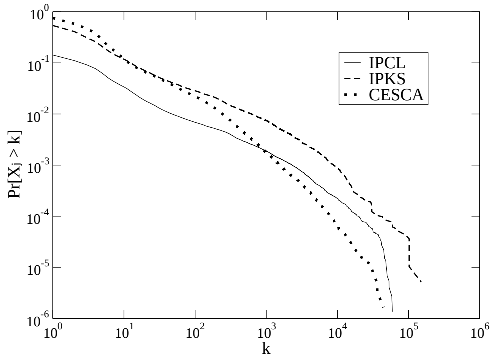

Figure 3: Complementary cumulative distributions of X_j for the three traces.

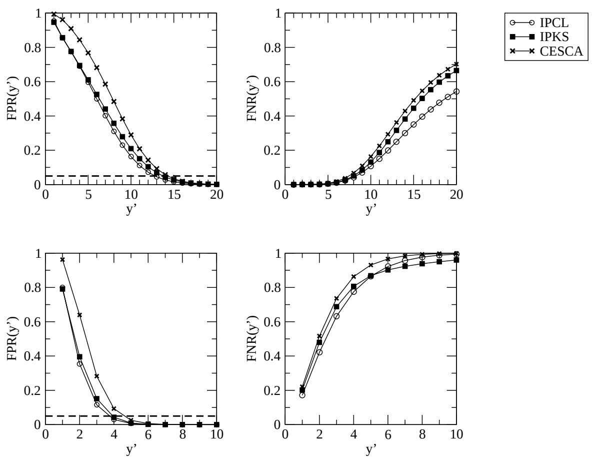

Figure 4: FPR(y') and FNR(y') for the three traces: $f = 10^{-3}$ (top) and $f = 10^{-4}$ (bottom).

pled per-flow packet distribution $\Pr[X_j = i]$ can be inferred from the observation of sampled per-flow packet distribution $\Pr[Y_j = i]$. Approach C utilizes the observation that the distribution of the number of per-flow packets for Internet traffic seems to decay in an approximately power-law fashion.

In the following, we focus on Approach C. We investigate how $\Pr[X_j = i]$ affects $\hat{y}$. First, we look into several empirical distributions obtained from packet traces measured in different networks. We then augment our investigation with the case of a family of theoretical Pareto distributions. We find, somehow to our surprise, that the calculated thresholds $\hat{y}$ are similar for these empirical and theoretical distributions. This result suggests that a threshold obtained for one network can be used as an approximation for other networks as long as their flow statistics also exhibit heavy-tail characteristics.

5.1 Empirical distributions

To supplement the IPCL trace, we used two more unidirectional packet traces of Internet traffic. The first one was measured at the same location as the IPCL trace, but at other OC-48c backbone links, i.e., those between Indianapolis and Kansas City [9]. We refer to this trace as the IPKS trace[6]. The second one was measured at a Gigabit Ethernet link connecting the Anella Cientifica to the global Internet via RedIRIS [10]. We refer to this trace as CESCA[7]. Each link had an adequate volume of traffic during the measurement period.

Table 3: $\hat{y}$ for the three traces by calculation

trace	$\hat{y}$ $(f = 10^{-3})$	$\hat{y}$ $(f = 10^{-4})$
IPCL	13	4
IPKS	14	4
CESCA	15	5

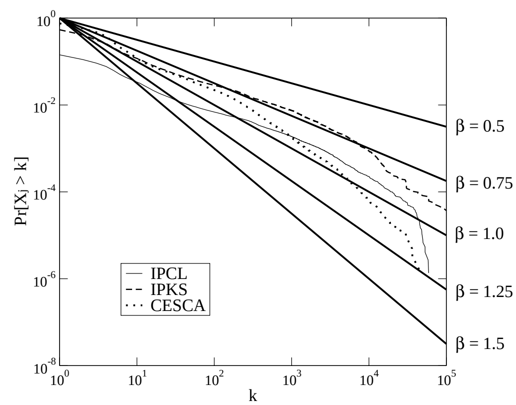

Figure 5: Complementary cumulative distributions of X_j for the Pareto distributions.

Figure 3 shows complementary cumulative distributions $\Pr[X_j > k] = 1 - \sum_{i=1}^{k} \Pr[X_j = i]$ for these three traces. Although their distributions are different, they all exhibit heavy-tail behaviors and decay in an approximately power-law fashion. Then, we use the distribution $\Pr[X_j = i]$ for each trace to calculate FPR(y') and FNR(y') (see Fig. 4), with $N = 10^7$, $f \in \{10^{-3}, 10^{-4}\}$, and $\hat{x} = 10^4$; four subfigures show similar tendencies of FPR(y') and FNR(y') for all these three traces. Next, we use (5) to calculate $\hat{y}$, with $\epsilon = 0.05$. Table 3 lists the calculated thresholds of identifying elephant flows in sampled packets. As we can see, these thresholds ($\hat{y}$) are very similar for traces with different per-flow packet distributions.

5.2 Theoretical distributions

We use the Pareto distribution as a theoretical distribution, since it is appropriate for evaluating behaviors for which the complementary cumulative distribution decays in a power-law fashion. This distribution is often referred to as the basis of *the elephant and mice phenomenon*. The Pareto distribution is defined by $\Pr[X_j \le x] = 1 - (\alpha/x)^{\beta}$ $(x \ge \alpha)$, where $\alpha > 0$ is a location parameter, and $\beta > 0$ is a shape parameter. With the above notation, the probability density function is defined as $\Pr[X_j = x] = \beta\alpha^{\beta}/x^{\beta+1}$. Figure 5 illustrates the complementary cumulative distribution of the Pareto distributions ($\alpha \le x \le 10^5$). Here, $\alpha = 1.0$, and $\beta \in \{0.5, 0.75, 1.0, 1.25, 1.5\}$. The empirical distributions are also plotted for comparison purpose.

[6]We use the first 10^7 packets (corresponding to about 124 seconds) in the trace IPLS-KSCY-20020814-105000-1 [9].

[7]We use the first 10^7 packets (corresponding to about 85 seconds) in the trace 20040219-120000-a [10].

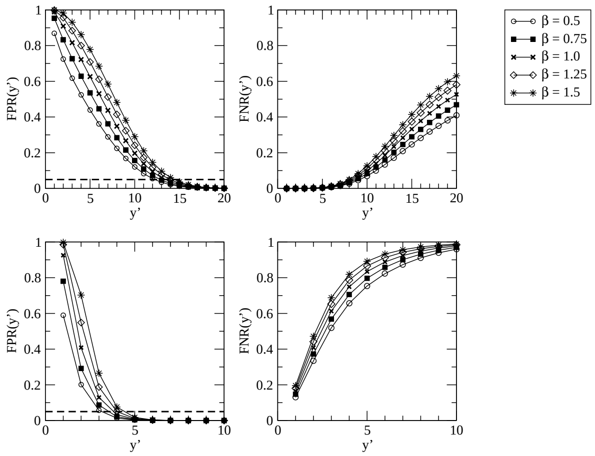

Figure 6: FPR(y') and FNR(y') for the Pareto distributions: $f = 10^{-3}$ (top) and $f = 10^{-4}$ (bottom).

Table 4: $\hat{y}$ for the Pareto distributions by calculation

β	$\hat{y}$ $(f = 10^{-3})$	$\hat{y}$ $(f = 10^{-4})$
0.5	13	4
0.75	13	4
1.0	14	4
1.25	14	4
1.5	15	5

For each of the above Pareto distributions, we calculate FPR and FNR by using (3) and (4) (see Fig. 6), and $\hat{y}$ by using (5) (see Table 4). Here, $N = 10^7$, $f \in \{10^{-3}, 10^{-4}\}$, $\hat{x} = 10^4$, and $\epsilon = 0.05$. Both FPR(y') and FNR(y') show very similar tendencies for all distributions. The calculated thresholds ($\hat{y}$) are similar for the Pareto distributions over a relatively wide range of β, and are similar to those for the empirical distributions.

As shown in Fig. 5, the empirical distributions of X_j for the three traces roughly follow $\Pr[X_j > x] \sim x^{-\beta}$, where β is close to 1. Other measurement-based studies have reported that flow statistics exhibit such characteristics [5, 8, 12, 15, 16] in many different networks (including the Internet). Therefore, we may conclude that the value of $\hat{y}$ calculated by using the Pareto distribution (say, with $\beta = 1.0$) can be used as approximations for a wide range of other networks as long as their per-flow packet statistics also follow heavy-tail distributions.

6. CONCLUSION

In this paper, we have described how to identify elephant flows through counting periodically sampled packets. The key is to find the threshold of per-flow packets in sampled packets which can reliably indicate whether or not a flow is actually an elephant flow in unsampled packets. We have shown that such a threshold can be obtained based on Bayes' theorem, with a proper trade-off of false positives and false negatives. Moreover, we have found that for various *a priori* distributions, the calculated thresholds are quite similar. This observation suggests that a threshold obtained for one network will be applicable to other networks exhibiting similar per-flow packet distributions. Although our current scheme only focus on identifying elephant flows instead of exacting detailed flow statistics, we believe that identifying elephant flows is an important step toward the goal of the latter during network operation and traffic engineering.

The advantage of our approach is due to its simplicity. Periodic sampling and stateless flow identification without per-packet processing can be easily implemented in any contemporary high-end PC, which is of course a very cost-effective solution. In addition, our analytical framework quantifies the intrinsic trade-off in flow identification and can provide insights on how to choose appropriate parameters. For ISPs operating large-scale networks with a variety of measurement equipments, these features are considered very crucial in practice.

Acknowledgments

The authors would like to acknowledge the people of the PMA project for making their packet traces publicly available to the networking research community. We also wish to thank anonymous reviewers, who have provided valuable comments on an early version of this paper.

7. REFERENCES

[1] N. Duffield, C. Lund, and M. Thorup, "Charging from Sampled Network Usage," ACM SIGCOMM Internet Measurement Workshop, California, November, 2001.

[2] N. Duffield, C. Lund, and M. Thorup, "Properties and Prediction of Flow Statistics from Sampled Packet Streams," ACM SIGCOMM Internet Measurement Workshop, Marseille, France, November, 2002.

[3] N. Duffield, C. Lund, and M. Thorup, "Estimating Flow Distributions from Sampled Flow Statistics," In Proceedings of ACM SIGCOMM, pp. 325–336, August 2003.

[4] C. Estan and G. Varghese, "New Directions in Traffic Measurement and Accounting," In Proceedings of ACM SIGCOMM, pp. 323–336, August 2002.

[5] S. Ben Fredj, T. Bonald, A. Proutiere, G. Regnie, and J. Roberts, "Statistical bandwidth sharing: a study of congestion at flow level," In Proceedings of ACM SIGCOMM, pp. 111–122, August 2001.

[6] L. Golab, D. DeHaan, E. Demaine, and A. Lopez-Ortiz, "Identifying Frequent Items in Sliding Windows over On-Line Packet Streams," ACM SIGCOMM Internet Measurement Conference, Florida, October, 2003.

[7] A. Kumar, J. Xu, J. Wang, O. Spatschek, and L. Li, "Space-Code Bloom Filter for Efficient Per-Flow Traffic Measurement," In proceedings of IEEE INFOCOM, Hong Kong, China, March 2004.

[8] T. Mori, R. Kawahara, S. Naito, and S. Goto, "On the characteristics of Internet Traffic variability: Spikes and Elephants," In Proceedings of IEEE/IPSJ SAINT, pp. 99–106, Tokyo, Japan, Jan 2004

[9] NLANR: Abilene-I data set, http://pma.nlanr.net/Traces/long/ipls1.html

[10] NLANR: CESCA-I data set, http://pma.nlanr.net/Special/cesc1.html

[11] Cisco NetFlow, http://www.cisco.com/warp/public/732/netflow/index.html

[12] K. Papagiannaki, N. Taft, S. Bhattacharya, P. Thiran, K. Salamatian, and C. Diot, "On the feasibility of identifying elephants in internet backbone traffic. Sprint ATL Technical Report TR01-ATL-110918," Sprint Labs, November 2001.

[13] IETF Packet Sampling (psamp) Working Group, http://www.ietf.org/html.charters/psamp-charter.html

[14] InMon sFlow Probe, http://www.inmon.com/products/probes.php

[15] K. Thompson, G. J. Miller, and R. Wilder, "Wide-area internet traffic patterns and characteristics," IEEE Network, vol. 11, no. 6, pp. 10–23, November/December 1997.

[16] Y. Zhang, L. Breslau, V. Paxson, and S. Shenker, "On the Characteristics and Origins of Internet Flow Rates," In Proceedings of ACM SIGCOMM, pp. 309–322, August 2002.