Sampling and inference of flow statistics in the Internet

Fabrice Guillemin

Orange Labs, 22300 Lannion
Joint work with INRIA, RAP project
Rennes, 23 October 2009

Introduction

- The Internet is to a large extent "open":
  - there is no signaling between users and the network (unlike the classical PSTN) - the network is not aware of what users are doing
  - there is a wide variety of applications, which can be downloaded by end users without control by the ISP
  - ISP are most of the time not allowed to inspect the payload of IP packets (privacy and confidentiality of personal data)
- Traffic measurements are hence essential in the Internet in order to
  - study the composition of traffic (applications identified through ports or packet patterns, usage in terms of bandwidth, etc.)
  - evaluate the quality perceived by end users
  - supervise the network (anomaly detection, troubleshooting, etc.)

Challenges in traffic measurements

- The speed of transmission links rapidly increases
  - at the access through the deployment of high speed ADSL lines (10 Mbit/s downlink - 1 Mbit/s uplink), fiber access (100 Mbit/s downlink - 10 Mbit/s uplink), and new radio technologies (3G, 3G+, LTE, etc.)
  - in the core with transmission links of 10 Gbit/s links, 100 Gbit/s soon
- The topology of the Internet is very intricate:
  - thousands of autonomous systems intricately interconnected
  - autonomous systems are designed according to ad hoc rules by ISP (absence of best practice guidelines)

$\Rightarrow$ Need for efficient measurement tools for mining traffic, inferring topologies (tomography), computing traffic matrices for network planning, etc.

Measurements from a transmission link

- The exhaustive observation of traffic on a high speed link ($\geq$ 1 Gbit/s) for a few hours leads to the storage of hundreds of Gigabytes
- Exhaustive measurements allow the accurate analysis of the activity of end users in terms of applications, contacted peers, bit rates, perceived quality of service, etc.
- However, exhaustive traffic analysis does not scale, leads to prohibitive storage and computing capacities, and consumes too much bandwidth for retrieving measurement data
- To reduce the amount of data to analyze and the number of measurement devices, sample traffic

$\Rightarrow$ Fundamental question: How to sample traffic?

How to sample traffic?

- Two dimensions: space and time
- In space, in order to reduce the number of measurement devices, choose ad hoc measurement locations:
  - particular points in the network (e.g., peering links)
  - regional areas covering sufficiently large numbers of end users (e.g., close to a BAS),
  - aggregation points (e.g., border routers of backbones)
- In time, select packets to analyze
  - all packets are not equivalent: data packets for characterizing user activity, SYN packets for tracking TCP connection establishments and possibly for detecting DDoS attacks, etc.
  - those packets with the same source and destination addresses, the same source and destination ports and the same protocol type form a **flow**

The concept of flow

Definition (Flow)

A flow is a set of packets with the same source and destination addresses, the same source and destination ports and the same protocol type. A flow is terminated when no packets have been observed for a duration of τ seconds (e.g., a few seconds).

- For instance, a TCP connection, a UDP stream is a flow
- The flow is the closest to user activity and the most relevant to evaluate the quality perceived by the end user
  - estimation of the packet loss rate through TCP segment retransmissions,
  - evaluation of the bit rate of the application
  - ...

Flow sampling vs. packet sampling

- Flow oriented sampling (as well as resource management) would be the **best solution** in the Internet
- However, flow sampling is much more CPU consuming than packet sampling
  - need for identifying flows by maintaining flow tables
  - look up of flow tables upon every packet arrival
  - big flows (elephants) are more relevant for estimating the quality of service for file downloads
  - however short flows (mice) are more relevant for some applications (e.g., web surfing, etc.) and some supervision tasks (e.g., detection of Distributed Denial of Service attacks)

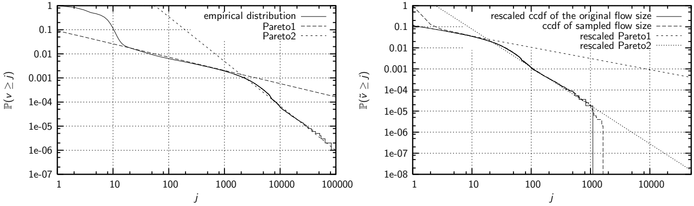
In practice, only packet sampling is implemented (NetFlow by CISCO)

Packet sampling

- 1-out-of-k sampling is implemented in CISCO routers (NetFlow)
  - one packet every other k packets is captured by the measurement engine
  - information of captured packets (arrival time, volume, TCP flags, etc.) is inserted into a flow table
  - the flow table is exported when full or upon expiration of a timer in the form of NetFlow records
  - a collector receives NetFlow records and analyzes traffic
- NetFlow has many shortcomings, the most limiting one is that packets are sampled without taking care of the flow level
- Information on big volumes is preserved (law of large numbers) but information on small flows is erased

Challenge: How to infer flow statistics from NetFlow sampled

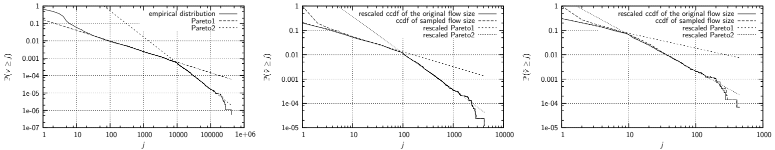
Example 1: Sampling of an exhaustive ADSL traffic trace



(a) Original size (b) Sampled size ($p = 1/100$)

Figure: Distribution of the flow size.

Example 2: Abilene traffic trace (campus traffic)



(a) Original size (b) $p = 1/100$ (c) $p = 1/1000$

Figure: Distribution of the flow size.

Other considerations

- Alternative to deterministic 1-out-of$k$ packet sampling: probabilistic packet sampling
  - Every packet is selected with probability $p = 1/k$
  - Advantages:
    - flows can be considered in isolation
    - derivation of theoretical bounds is easier
  - Probabilistic sampling is not very different from deterministic sampling: if $\tilde{v}^d$ (resp. $\tilde{v}^p$) is the number of sampled packets in a flow under deterministic (resp. probabilistic) sampling, then by Le Cam inequality

$$\|\mathbb{P}(\tilde{v}^d \in \cdot) - \mathbb{P}(\tilde{v}^p \in \cdot)\|_{tv} \leq p\mathbb{E}\left(\frac{v^2}{V}\right) + p^2\mathbb{E}(v).$$

The heterogeneity of Internet traffic

- Internet traffic is not "univariate", its characteristics greatly depend on the dominant applications:
  - Commercial IP traffic has been dominated for the last few years by peer-to-peer applications ($>$ 70% of global traffic in 2002-2007), now streaming traffic (Youtube, Dailymotion, etc.) is of the same order of magnitude in peak hours
  - Campus traffic comprises bigger elephants with higher bit rates
- The distribution of flows in not "unimodular":
  - the flow size of mice is much different from that of elephants
  - Various types of elephants: small and big elephants (piecewise "Pareto" distribution)

$\Rightarrow$ Unmanageable situation for a parametric characterization of flows

Possible solution: adaptation to traffic

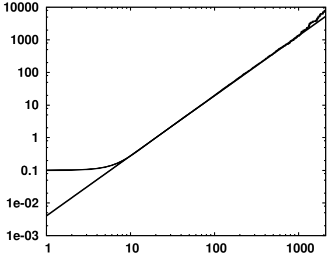
- Choose the observation window and fix thresholds so as to obtain a nice "Pareto" distribution for the number of packets of elephants
- Assumption: there exist Δ, B_{min}, B_{max} and $a > 0$ such that if v is the number of packets transmitted by a flow in Δ time units, then $\mathbb{P}(v \geq x \mid S \geq B_{min}) \sim (B_{min}/x)^a$ for $B_{min} \leq x \leq B_{max}$
- The proportion of elephants with size greater than B_{max} is less than 5%.
- An elephant is a flow with a least B_{min} packets
- The length Δ is chosen so that there is a sufficient number of large flows in time window (warm-up of a few minutes)

Algorithm (exhaustive traffic trace)

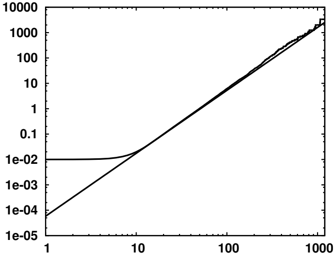
- Δ is fixed so that at least 1000 flows have more than 20 packets.
- B_{max} is defined as the smallest integer such that less than 5% of the flows have a size greater than B_{max}.
- A Least Square Method is performed to get a linear interpolation in a log-log scale of the distribution of sizes between B_{min} and B_{max}. The constant B_{min} is chosen as the smallest integer such that the L_2-distance in the sense of least square method with the approximating straight line is less than 2.10^{-3}. The slope of the line gives the value of the parameter a.

Synthetic traces (10^6 flows with a Pareto distribution)

1. Pareto $a = 1.85$. Estimation: $\hat{a} = 1.84$, $B_{min} = 9$, $B_{max} = 100$
2. Pareto $a = 2.5$. Estimation: $\hat{a} = 2.48$, $B_{min} = 11$, $B_{max} = 65$



(1)



(2)

Experimental results

Characteristics of traffic traces considered in experiments

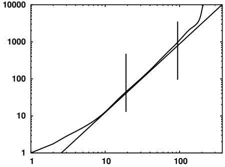
Name	Nb. IP packets	Nb. TCP Flows	Duration
ADSL Trace A	271 455 718	20 949 331	2 hours
ADSL Trace B Upstream	54 396 226	2 648 193	2 hours
ADSL Trace B Downstream	53 391 874	2 107 379	2 hours
Abilene III Trace A	62 875 146	1 654 410	8 minutes
Abilene III Trace B	47 706 252	1 826 380	8 minutes

Statistics of the elephants for the different traffic traces.

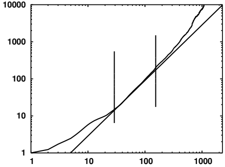
	ADSL A	ADSL B Up	ADSL B Down	Abilene A	Abilene B
Δ (sec)	5	15	15	2	2
B_{min}	20	29	39	89	79
B_{max}	94	154	128	324	312
a	1.85	1.97	1.50	1.30	1.28

The Abilene traces 20040601-193121-1.gz (trace A) and 20040601-194000-0.gz (trace B) can be found at the url http://pma.nlanr.net/Traces/Traces/long/ipls/3/.

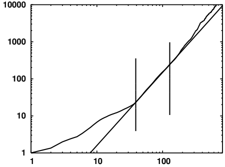
Results with the ADSL traffic trace



$\Delta = 5$ s - Trace A - downlink

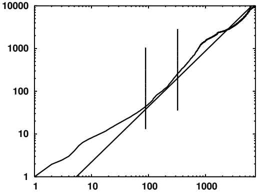


$\Delta = 15$ s - Trace B -downlink

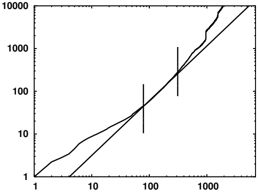


$\Delta = 15$ s - Trace B - uplink

Experimental results (Abilene)



$\Delta = 2$ s



$\Delta = 2$ s

In the Abilene traces, elephants are bigger, the Pareto approximation covers a larger range of flow sizes than in the ADSL trace but the observation window is much smaller to preserve the validity of the Pareto approximation

Problem statement

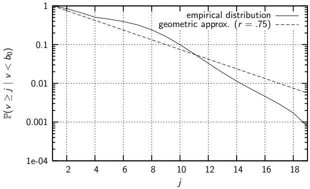
- Problem statement: infer the characteristics of flows from sampled data
- Methods in the technical literature
  - Duffield *et al*
    - The quantity $\hat{V} = k\hat{v}$ is an unbiased estimator of the initial flow size v and the error $\sqrt{\hat{V}} \leq \sqrt{kv}$ (these results rely on the use of the central limit theorem)
    - The initial number of flows is inferred by using the number of SYN messages ($\hat{N} = k.n_{SYN}$)
    - The method cannot be applied to small or moderate flow sizes (the original flow size should be $O(k)$)
  - Hohn and Veitch:
    - Assume probabilistic sampling
    - Use generating functions to recover the original flow size distribution
    - The inversion procedure is however unstable

Parametric approach

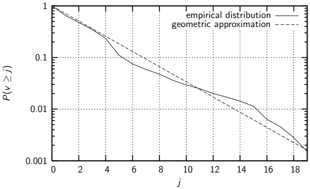
- Assumptions on the form of traffic: two types of flows (mice and elephants)
  - Mice: a mouse is a flow with less than 20 packets
  - Elephants: an elephant is a flow comprising at least 20 packets
- Assumptions on packet sampling
  - Mixing condition: If K TCP flows are active during a time interval of duration Δ, at each sampling instant a packet of the ith flow is chosen with probability v_i/V where v_i is the number of packets of the ith flow and $V = v_1 + \cdots + v_K$.
  - Negligibility assumption: In any window of length Δ, the number of packets of every flow is negligible when compared to the total number of packets V in the observation window. There specifically exists some $0 < \varepsilon \ll 1$ such that for all $i = 1, \ldots, K$, $v_i/V < \varepsilon$.

Assumptions on mice

Distribution of the size of mice



ADSL trace

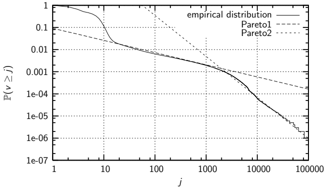


Abilene B trace

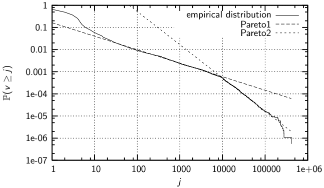
Geometric assumption: There exists some $b_0 > 0$ such that for $1 \leq j < b_0$, $\mathbb{P}(v = j) = (1 - r)r^{j-1}/(1 - r^{b_0})$ for some $r > 0$. In practice, we take $b_0 = 20$.

Assumptions on elephants (complete trace)

Distribution of the size of elephants



ADSL trace



Abilene B trace

Piecewise Pareto assumption: There exist some $m > 0$ and some integers $j_0 < j_1 < \ldots < j_m = \infty$ such that for $\ell = 1, \ldots, m$ and $j \in [j_{\ell-1}, j_\ell]$, v has a Pareto distribution of the form

$$\mathbb{P}(v \geq j) = \mathbb{P}(v \geq j_{\ell-1}) \, (j_{\ell-1}/j)^{a_\ell}$$

for some shape parameters $a_\ell > 0$.

Theoretical results

Proposition (Via Le Cam's inequality and Chen-Stein method)

Let W_j (resp. W_j^+) be the number of sampled flows with (resp. at least) j packets and K the original number of flows. Then,

$$\left| \frac{\mathbb{E}(W_j)}{K} - \mathbb{Q}_j \right| \leq \mathbb{E}\left(\min(pv, 1) \frac{v}{V} \right) < \varepsilon,$$

$$\left| \frac{\mathbb{E}(W_j^+)}{K} - \sum_{\ell \geq j} \mathbb{Q}_\ell \right| \leq \mathbb{E}\left(\min(pv, 1) \frac{v}{V} \right) < \varepsilon,$$

where $\mathbb{Q}$ is the probability distribution defined for $j \geq 0$ by $\mathbb{Q}_j = \mathbb{E}\left(\frac{(pv)^j}{j!} e^{-pv} \right)$, $p = 1/k$ is the sampling rate, and v is the number of packets in a flow.

Theoretical results (cont'd)

Corollary

If v has a Pareto distribution, then for all $j > a$

$$\lim_{K \to +\infty} \frac{\mathbb{E}(W_{j+1})}{\mathbb{E}(W_j)} = 1 - \frac{a+1}{j+1} + O((pb)^{j-a}), \qquad (1)$$

$$\lim_{K \to +\infty} \frac{\mathbb{E}(W_j)}{K} = a(pb)^a \frac{\Gamma(j-a)}{j!} + O((pb)^j), \qquad (2)$$

$$\lim_{K \to +\infty} \frac{\mathbb{E}(W_j^+)}{K} = (pb)^a \frac{\Gamma(j-a)}{(j-1)!} + O\left(\frac{(pb)^j}{1-pb}\right) \qquad (3)$$

The first equation gives a means of estimating the shape parameter a.

Theoretical results (cont'd)

Proposition (Obtained via Laplace method)

If v has a Weibull or Pareto distribution,

$$\lim_{j\to\infty}\lim_{K\to\infty}\frac{\mathbb{E}(W_j^+)}{K\mathbb{P}(v\geq j/p)}=1.$$

This proposition is valid even if the distribution is piecewise Pareto. If for $j_{m-1}\leq j\leq j_m$

$$\mathbb{P}(\tilde{v}\geq j)=\mathbb{P}(\tilde{v}\geq j_{m-1})\,(j_{m-1}/j)^{a_m}$$

then for $\frac{j_{m-1}}{p}\leq j\leq\frac{j_m}{p}$

$$\mathbb{P}(v\geq j)\sim\nu\mathbb{P}\left(\tilde{v}\geq j_{m-1}\right)\left(j_{m-1}/(pj)\right)^{a_m},$$

Inference procedure

- Geometric assumption for the size of mice and piecewise Pareto for the size of elephants
- Assume that $\mathbb{P}(v \geq j) = \mathbb{P}(v \geq b_0)(b_0/j)^{a_1}$ for $b_0 \leq j \leq j_1/p$ (reasonable if p is not too small)
- Estimate the shape parameters from the sampled flow size
- Use the Poisson approximation

$$\mathbb{P}(\tilde{v} = j) \sim \frac{\mathbb{P}(v < b_0)}{\nu} \sum_{\ell=1}^{\infty} (1-r) r^{\ell} \frac{(p\ell)^j}{j!} e^{-p\ell} + \frac{1}{\nu} \sum_{\ell=b_0}^{\infty} \frac{(p\ell)^j}{j!} e^{-p\ell} \mathbb{P}(v = \ell). \quad (4)$$

Inference procedure(cont'd)

- Estimate the quantity $\eta \stackrel{def}{=} \mathbb{P}(\tilde{v} \geq j)/(b_0 p/j)^{a_1}$ for $j \in \{j_0, \ldots, j_1\}$ (by assumption independent of j)
- The number of elephants is $K_e = \eta \tilde{K}$ ($\tilde{K}$ being the number of sampled flows)
- Use Equation (4) for $j = 1, , \ldots, j_0 - 1$ for estimating the parameter r and the number of mice K_m
- The initial number of flows is $K = K_e + K_m$ and the probability of sampling a flow is $\nu = K_s/K$

Experimental results with an ADSL traffic trace ($p = 1/100$, $K_s = 1, 120, 546$)

	a_1	a_2	r	K_e	K_m	ν
experimental	.52	1.81	.75	343,004	19.8e6	.057
estimated	.54	1.81	.84	336,163	20.1e6	.054

Adaptation of the observation window

- By reducing the observation window, the elephant size is Pareto
- The shape parameter is estimated by using Equation (1)

$$a \sim a(j) \stackrel{\text{def.}}{=} (j+1)\left(1 - \frac{\mathbb{E}(W_{j+1})}{\mathbb{E}(W_j)}\right) - 1, \qquad (5)$$

- The number of elephants is given by

$$K \sim K(j) \stackrel{\text{def.}}{=} \frac{j!\,\mathbb{E}(W_j)}{a(j)(pB_{min})^{a(j)}\Gamma(j - a(j))}.$$

Algorithm

Algorithm used to identify Δ and the Pareto parameter from sampled traffic.

— Choose Δ so that $80 \leq \mathbb{E}[W_2] \leq 100$;

— Choose j so that $|a(j) - a(j+1)|$ computed with Equation (5) is minimized with for all j such that $\mathbb{E}[W_j] \geq 5$.

— B_{min} is the smallest integer so that the probability that a flow of size greater than B_{min} is sampled more than j times is greater than $p/10$;

Experimental results

Elephants for the France Telecom ADSL and the Abilene traffic traces.

	ADSL A	ADSL B Up	ADSL B Down	Abilene A	Abilene B
B_{min}	20	29	39	89	79
estimated B_{min}	21	45	45	77	77

Estimations of the Number of Elephants from Sampled traffic

Trace	Δ	j	$\mathbb{E}(W_j)$	$\mathbb{E}(W_{j+1})$	a_{exp}	$a(j)$	K_{exp}	$K(j)$
ADSL A	5s	3	12.89	3.33	1.85	1.95	943.71	1031.04
ADSL B Do	15s	4	9.7	4.75	1.49	1.55	414.90	404.13
ADSL B Up	15s	4	7.46	2.97	1.97	2.00	453.01	462.68
ABILENE A	1s	5	6.04	3.21	1.38	1.81	217.44	270.79
ABILENE B	1s	5	6.1	3.7	1.36	1.51	209.12	197.12

Conclusion

- The heterogeneity of Internet traffic leads to complex parametric models
- Flow sampling would be better but requires more CPU than packet sampling
- Information of small flows is lost through packet sampling
- Inference of flow statistics can be done via a parametric approach but is fragile
- A possible solution: adapt the observation window to simplify the characterization but the global view of flows is lost

Bibliographical support

1. C. Fricker, F. Guillemin and P. Robert. An identification problem in an urn and ball model with heavy tailed distributions. To appear in Probability in Engineering and Informational Science, 2009.
2. Y. Chabchoub, C. Fricker, F. Guillemin, and P. Robert. On the statistical characterization of flows in the Internet with application to sampling. To appear in Computer Communications, 2009.
3. Y. Chabchoub, C. Fricker, F. Guillemin, and P. Robert. Inference of flow statistics via packet sampling in the Internet. IEEE Communication Letters 12(12), December 2008.
4. Y. Chabchoub, C. Fricker, F. Guillemin, and P. Robert. Deterministic versus probabilistic sampling in the Internet. Proc. International Teletraffic Congress (ITC) 20, Ottawa, Canada, 2007.
5. N. Ben Azzouna, F. Guillemin, S. Poisson, P. Robert, C. Fricker, and N. Antunes. Inverting sampled ADSL traffic. Proc. ICC 2005, May 16-20 2005, Seoul, Korea.