

A new regenerative estimator for effective bandwidth prediction

I.Vorobieva (Petrozavodsk University), G.Procissi(University of Pisa), E.Morozov (Institute of Applied Mathematical Research), M.Pagano (University of Pisa)

Motivation: Lately computer networks become more and more widespread. Sometimes a problem of guaranteeing of reliable and effective data transmission is critically important. In this connection it is essential to focus attention on such events, that can entail faulty operation or become a reason of loss of a message. One of such events is a great load of network components (for example, big queues of packets in units of a network). One can solve this problem using superfluous resources. But it is clear, that this method is economically unprofitable. At the same time insufficient resources can't provide demanded parameters of work of a network.

Problem: What minimal volume of resources is able to provide given quality of service? There is also inverse problem: what quality of service can be guaranteed using given volume of resources?

Theoretical base: Typically the load of network units is moderate. So it's drastic increase is a large deviation from typical values. So our research is based on the Large Deviation Theory (LDT).

LDT: LDT is concerned with the rate of a convergence of some probabilities to zero. Let's describe it. A large deviation from prospective value (for example, a mean value) is a rare event, and in limit probability of such deviation is zero. LDT shows, that for wide class of rare events such probabilities turn out to have exponential form.

$$\mathbb{P}(X > a) \approx e^{-\delta a}$$

LDT and queues: In the view of a mathematical model a unit of a network is a queuing system. The main results of LDT, that allow to calculate probabilities of great load of queuing system, are:

$$\lim_{b \rightarrow \infty} (Q > b) = -\delta,$$

where

$$\delta = \sup\{\theta > 0 : \Lambda(\theta) < \theta C\},$$

where

$$\Lambda(\theta) = \log \mathbb{E}e^{\theta X}.$$

Here X is a number of arrivals during a time unit and Q is a queue size. Function $\Lambda(\theta)$ is called rate-function, because it determines the rate of convergence of probability of rare event to zero.

Aim of the research: LDT in connection with queuing systems can be applied for solution of many practical problems,

concerning guaranteeing of quality of service in computer networks. One of such problems is estimating of minimal bandwidth of the network unit, that can provide given quality of service. Our research is devoted to solution of this problem.

Original experiment: In the article [1] is described experiment, the aim of which was to predict minimal rate, with which server should process udp-packets to provide given quality of service. Under quality of service we mean here maximal loss-ratio of up-d-packets (the ratio of the number of refused udp-packets to the total number of arrivals).

Experiment was run over Ethernet, at that incoming traffic was monitored on the Sun workstation.

Estimation technic: All period of observation was divided into several time interval with fixed length. Then the number of udp-packets, arrived at the server during each time interval, was calculated.

Denote the number of packets, incoming during i -th time interval, as X_i . Then assume block sums

$$\tilde{X}_1 := \sum_{k=1}^B X_k,$$

$$\tilde{X}_2 := \sum_{k=B+1}^{2B} X_k, \dots,$$

where B is chosen block size. Let's note, that value of B is fixed

and chosen arbitrarily.

Then let's find parameter θ^* , that will be useful for further calculations.

$$\theta^* = -\frac{\log \Gamma_0}{b},$$

where Γ_0 is the maximal loss-ratio, and b is a buffer-size of the node.

Now using values X_i и θ^* let's estimate a value of the rate-function

$$\hat{\lambda}_B^T(\theta) := \frac{1}{B} \log \frac{B}{T} \sum_{i=1}^{T/B} e^{\theta \tilde{X}_i},$$

where T is the total time of observation.

Then effective bandwidth estimator is:

$$\hat{s}(b, \Gamma_0) := \frac{\hat{\lambda}(\theta^*)}{\theta^*}$$

Problem: What block size B should be used? The report on the experiment contains estimators of effective bandwidth, that was calculated using different values of B . At that these estimators differ for 5-7%.

Refined estimator: We supposed, that in the case of regenerative structure of the input stream, the estimator could be refined, if we would group values X_i not in blocks with some length B , but using cycles of regeneration. So, if β_k are moments

of regeneration of the input stream, then

$$\tilde{X}_k := \sum_{i=\beta_k}^{\beta_{k+1}-1} X_i.$$

All other calculations are the same. Besides, such way of grouping seems to be more motivated, because in this case random variables \tilde{X}_k would be really independent in view of independence of regeneration cycles. It allows to prove an optimality of the estimator.

To check our assumption about higher efficiency of the estimator, based on the regeneration cycles, an imitating experiment was run. Within the limits of the experiment we developed a program, that models tandem network with two single server stations.

In such system arrivals are serviced consecutively: at first, on the first station, then on the second one. The input stream to the second station is regenerative, and moments, when the buffer of the first station is empty, are moments of regeneration.

An example of results of the program are shown on the fig. 1.

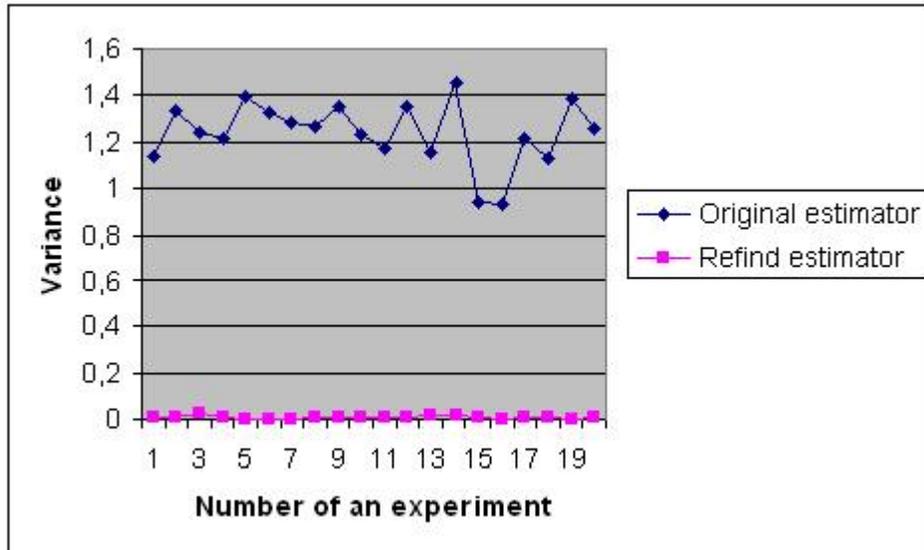


Fig. 1

Results: Results of the imitating experiment showed, that in overwhelming majority of cases a variance of the estimator, calculated using regeneration cycles is lower, then a variance of the estimator, calculated using blocks with arbitrary length. It is known, that an estimator with lower variance is considered to be more effective. So we can say, that imitating modeling confirms our assumption: our modification of the technic of effective bandwidth prediction allows to calculate not only more motivated, but also more effective estimator in case of a regenerative input stream.

References

- [1] S. Crosby, I. Leslie, M. Huggard, J. T. Lewis, B. McGurk and R. Russel. Predicting bandwidth requirements of ATM and Ethernet traffic, 1996.
- [2] A. Ganesh, N. O'Connell and D. Wischik. Big Queues. Springer-Verlag, Berlin, 2004.
- [3] E. Morozov. Communication systems: rare event simulation and effective bandwidths. Pamplona, University of Navarre, 2004.