Asymptotic bounds for large buffer overflow probability

Yu. S. Khokhlov

Lomonosov Moscow State University, Moscow, Russia

L. V. Suslov

Lomonosov Moscow State University, Moscow, Russia

O. I. Sidorova

Tver State University, Moscow, Russia

The problem of quality of service estimation is the most important one in the creation and configuration of modern telecommunications systems. For many years, Markov processes have been successfully used in the analysis of voice and text message flows, which in many cases allow us to obtain explicit formulas for calculating the stationary performance characteristics in the models under consideration.

However, information in modern telecommunications systems has a complex structure that is radically different from conventional telephony or mail services. For the first time, this fact was noticed in the early nineties after conducting high-precision measurements of Internet traffic in the Bell laboratory. Analysts noted three important distinguishing features inherent in data flows in computer networks: **self-similarity** with a wide range of traffic aggregation, **slowly decreasing correlation** (**long memory**) of observations, and **heavy tails** of load distributions coming from sources.

Theoretical and empirical studies have shown that ignoring these features leads to serious errors in QoS estimates. Therefore, it is necessary to build new traffic models that would have the necessary properties. Currently, the most popular models of this type are **fractional Brownian motion** and **stable Levy motion**. It turned out that these models are closely related to the heavy tails of the distributions and the connection speed remote sources with the server. Namely, in the case of fast connection we get a fractional Brownian motion, and in the case of slow connection – stable Levi motion. (see, for example, [1]).

In a number of empirical studies (see, for example, [2] and [3]) it has been shown that very often traffic contains both of the components noted above. For such traffic there is no general methods to assess the quality of service. Note only the works ([4] and [**?**]), where the asymptotic lower bounds of the large buffer overflow probability were obtained for incoming streams based on fractional Brownian motion and the sum of independent fractional Brownian motions with different exponents H.

In our report we analyse the nonhomogenous traffic model based on sum of independent Fractional Brownian motion and symmetric α–stable Levy process with different Hurst exponents H_1 and H_2. For such model we find asymptotical bounds for the overflow probability when the size of buffer $b \to \infty$.

Consider a queuing system to which the next input stream is fed:

$$A(t) = mt + \sigma_1 B_{H_1}(t) + \sigma_2 L_\alpha(t), \tag{1}$$

where$m = m_1 + m_2 > 0$ – average input flow rate ; $B_{H_1} = (B_{H_1}(t),\ t \in R^1)$ – fractional Brownian motion with Hurst parameter H_1, $L_\alpha = (L_\alpha(t),\ t \in R^1)$ – symmetric α-stable Levy motion with parameter $\alpha = 1/H_2$. So both components of the input process are self-similar with indexes H_1 and $H_2 = 1/\alpha$.

The process $A(t),\ t \geq 0$ describes the total load received by the communication node in the time interval $[0, t]$.

Assume that $1/2 < H_1, H_2 < 1$ and the processes $B_{H_1}(t),\ L_\alpha(t)$ are independent.

Let the queuing system in question consists of a single device with a constant service rate $C > 0$. Then the traffic intensity is $r = C - m > 0$.

We are interested in estimation of so-called **overflow probability**, i. e. the probability that stationary workload Q exceeds some threshold level b, namely $\varepsilon(b) := P[Q > b]$. Using the method proposed by Norros we get the asymptotic bounds for $\varepsilon(b)$ in the case of large b.

Our main results are the following theorems. Denote $H = \ min(H_1, H_2)$ and $\tilde{H} = \max(H_1, H_2)$.

Theorem 1. *Consider a queuing system in which there is a single service device with a buffer of size b and a constant service rate C. Let the process of the load entering the system in the interval $[0,\ t]$ be described by the model (1). If the traffic intensity is $r = C - m > 0$, then the following asymptotic estimate is valid for the probability $\varepsilon(b)$ that the stationary load will exceed a* *certain level b:*

$$\varepsilon(b) = P[Q > b] \geq C_1 \cdot b^{-(1-H)\alpha}, \quad b \to \infty, \tag{2}$$

where C_1 is some explicitly calculated constant.

Theorem 2. *Consider a queuing system in which there is a single service device with a buffer of size b and a constant service rate C. Let the process of the load entering the system in the interval* $[0, t]$ *be described by the model (1). If the traffic intensity is $r = C - m > 0$, then the following asymptotic estimate is valid for the probability $\varepsilon(b)$ that the stationary load will exceed a certain level b:*

$$\varepsilon(b) = P[Q > b] \leq C_2 \cdot b^{-(1-\tilde{H})\alpha}, \quad b \to \infty, \tag{3}$$

where C_2 is some explicitly calculated constant.

Unfortunately the lower and upper bounds have a different order with respect to b. Similar results can be obtained in the case when the number of components of the incoming stream is greater than 2.

This research was carried out in accordance with the scientific program of the Moscow Center for Fundamental and Applied Mathematics and Faculty of Computational Mathematics and Cybernetics in Moscow University.

References

[1] *Th. Mikosch, S. Resnick, H. Rootzen, A. Stegeman.* (2002) Is network traffic approximated by stable Levy motion or fractional Brownian motion? Ann. Appl. Probab., 12(1), p. 23–68.

[2] *S. Sarvotham, R. Riedi, and R. Baraniuk.* Connection–level analysis and modeling of network traffic, Tech. Rep., ECE Dept., Rice Univ., July 2001.

[3] *S. Sarvotham, R. Riedi, R. Baraniuk.* Connection–level Analysis and Modeling of Network Traffic. In: Proceedings of the 1st ACM SIGCOMM Workshop on Internet Measurement New York, NY, USA: ACM (2001), p. 99–103.

[4] I. Norros. A storage model with self-similar imput. *Queuing Syst.*, **16** (1994), 387-396.