A bird-eye view of fluid queues in communication network models: heavy tails and long memory, part I and II

Gennady Samorodnitsky
Cornell University, USA

Abstract

It has been noticed in the last 10 years or so that broadband measurements of teletraffic shows that the data exhibit the following characteristic properties: heavy tails, self-similarity and long-range dependence, starting with an influential sequence of papers including [10] and [14]. A common explanation of the observed self-similarity and long-range dependence is exactly via heavy tails: heavy tailed transmission times, heavy tailed burst lengths, etc. See e.g. [1], [2], [13].

A standard model connecting the empirical facts of heavy tails, self-similarity and long-range dependence is the ON/OFF model. In it, traffic is generated by a large number of independent ON/OFF sources (such as workstations in a big computer space). An ON/OFF source transmits data at a constant rate to a server if it is ON and remains silent if it is OFF. Every individual ON/OFF source generates an ON/OFF process consisting of independent alternating ON- and OFF-periods. The ON-periods are iid and so are the lengths of the OFF periods. Moreover, the ON- and the OFF-periods for each source are independent. Teletraffic is then generated by the superposition of a large number of these iid ON/OFF sources. Support for this model in the form of statistical analysis of Ethernet Local Area Network traffic of individual sources was provided in [14]. One of the conclusions of this study was that the lengths of the ON- and the OFF-periods are heavy tailed and in fact Pareto-like with tail index α between 1 and 2. Further evidence on infinite variance distributions in teletraffic is given in [2], [3], [10], which present evidence of infinite variance Pareto like tails in file lengths, transfer times and idle times in the World Wide Web traffic.

One of the immediate consequences of the assumption of Pareto-like tails with tail index α between 1 and 2 is that a stationary version of the ON/OFF-process of an individual source exhibits LRD in the sense that its covariance function stays positive and is not integrable; see [6] for a mathematical proof. This mathematical fact explains LRD at the individual source level, but not at the level of teletraffic. In the ON/OFF model, teletraffic is considered as the superposition of iid individual ON/OFF processes, and its workload is the integrated superposition of the ON/OFF processes.

An alternative model that generates a traffic with similar statistical properties is the so-called $M/G/\infty$ model, in which sessions arrive according to a Poisson process and session durations are heavy tailed. The key paper [11] showed that for both ON/OFF model and $M/G/\infty$ model, when the number of input "streams" is getting large, and the time scale increases as well, the deviation from the mean of the total work input into the system converges to one of the two well

known processes, either Lévy stable motion, or Fractional Brownian motion, depending on the relative rates at which the number of "streams" and the time scale grow.

From this point of view from the "bird-eye" point of view, a network driven by heavy tailed input looks like performing either stable motion or Fractional Brownian motion around its mean. This behaviout has since been generalized to network of queues in [4], and to random fields by [9].

However, the assumptions underlying this limiting behaviour are very specific, and we will consider the possible limiting behaviour of the queue under a very general scenario of input forming a marked stationary point process, with the points being the moments the jobs arrive to the quere, and the marks being the work requirements for each job. Using the Palm theory, we will see that there are many more different possible types of the limiting behaviour of a queue than just a stable motion or a Fractional Brownian motion. In fact, the Fractional Brownian motion scenario turns out to be much more robust towards chaging the specific assumptions than the stable motion scenario does.

Furthermore, we will see that, in fact, in some cases, the Fractional Brownian motion limit can appear exactly one would, naively, expect a the stable limit instead.

The first lecture will be of an introductory type, while the second lecture will discuss the new results. The latter are based on [12] and [5].

References

[1] Crovella, M. and A. Bestavros: *Explaining world wide web traffic self–similarity*, Preprint available as TR-95-015, 1995.

[2] Crovella, M. and A. Bestavros: *Self–similarity in World Wide Web traffic: evidence and possible causes*, Performance Evaluation Review <u>24</u> (1996) 160–169.

[3] Crovella, M., A. Bestavros and M. Taqqu: *Heavy-tailed probability distributions in the World Wide Web*, In *A PRACTICAL GUIDE TO HEAVY TAILS: Statistical Techniques for Analysing Heavy Tailed Distributions*, M.S.T. Robert Adler, Raisa Feldman, editor. Birkhäuser, Boston (1999) 3–26.

[4] D'Auria, B. and G. Samorodnitsky: *Limit behavior of fluid queues and networks*, Operations Research <u>53</u> (2005) 933–945.

[5] Fasen,V.: *The workload process with a Poisson cluster input can look like a Fractional Brownian motion even in the slow growth regime*, Preprint, 2008.

[6] Heath, D., S. Resnick and G. Samorodnitsky: *Heavy tails and long range dependence in on/off processes and associated fluid models*, Mathematics of Operations Research <u>23</u> (1998) 145–165.

[7] Hernandes-Campos, F., J. Marron, G. Samorodnitsky and F. Smith: *Variable heavy tailed durations in internet traffic*, Performance Evaluation Journal 58 (2004) 261–284.

[8] Hernandes-Campos, F., K. Jeffay, C. Park, J. Marron and S. Resnick: *Extremal dependence: Internet traffic applications*, Stochastic Models 21 (2005) 1–35.

[9] Kaj, I., L. Leskelä, I. Norros and V. Schmidt: *Scaling limits for random fields with long-range dependence*, Annals of Probability 35 (2007) 528–550.

[10] Leland, W., M. Taqqu, W. Willinger and D. Wilson: *On the self-similar nature of Ethernet traffic (extended version)*, IEEE/ACM Transactions on Networking 2 (1994) 1–15.

[11] Mikosch, T. S. Resnick, H. Rootzén and A. Stegeman: *Is network traffic approximated by stable Lévy motion or Fractional Brownian motion?* Annals of Applied Probability 12 (2002) 23–68.

[12] Mikosch,T. and G. Samorodnitsky: *Scaling limits for workload process*, Mathematics of Operations Research 32 (2007) 890–918.

[13] : Paxson, V. and S. Floyd: *Wide area traffic: the failure of Poisson modelling*, IEEE/ACM Transactions on Networking 3 (1994) 226–244.

[14] Willinger, W. M. Taqqu, R. Sherman and D. Wilson: *Self–similarity through high variability: statistical analysis of ethernet LAN traffic at the source level*, Proceedings of the *ACM/SIGCOMM'95*, Cambridge, MA, Computer Communications Review 25 (1995) 100–113.