A Comment on Simulating LRD Traffic with Pareto ON/OFF Sources

Thomas Bohnert, Edmundo Monteiro

Abstract— **It is common practice to evaluate the performance of Measurement Based Admission Control under traffic, which exhibits Long-Range-Dependency (LRD) and Self-Similarity (SS) effects. These phenomena are generated by multiplexing Markov ON/OFF sources with Pareto distributed ON and OFF times. Due to dependencies on a myriad of parameters, however, it can be questioned if resulting traffic actually exhibits SS and LRD effects. Further fostered by the lack of evaluation tools, the proof for is commonly missing. For that reason, in this paper we evaluate the robustness of SS and LRD effects in artificially generated traffic to parameter settings. Using a typical simulation for performance evaluation, we measure the degree of SS and LRD.**

Keywords—**Quality of Service, Long-Range-Dependency, Self-Similarity,**

I. INTRODUCTION

INDEPENDENT from any Quality of Service (QoS) supporting architecture, limits on quality measures can only be provided as long as offered load does not exceed available capacity. Henceforth, an essential component for any QoS architecture is Admission Control (AC) which limits competition for resources in times of high demand in advance.

In contrast to many congestion control methods, AC is preventive. Based on various concepts, at flow arrival an AC algorithm determines current and future resource demand accounting for QoS requirements of already admitted flows and the admission requester. Only if QoS bounds for both can be guaranteed, strictly or statistically, a requester is admitted; else rejected.

Clearly, the major challenge for AC is precise estimation of future resource availability as admission argument. Two different approaches exist, Parameter Based AC (PBAC)[1] or Measurement Based AC (MBAC). The former, PBAC, is based on *a priori* source descriptors and theoretical source models. This approach however, is doomed to fail in multi-service and packet oriented architectures like the Internet cause of its myriad of applications and where source behaviour is subject to distortion due to multiple buffering in routers.

Not being able to accurately model the network, measurement based approaches appeal attractive. Disposable resources are estimated by sampling the arrival process and applying probabilistic models. This complies to shift from static source models based on *a priori* knowledge to a dynamic and adaptive approach based on actual measured quantities.

The performance of a measurement based estimator, and thus of an MBAC algorithm, inherently depends on features of the sampled time series, i.e. the packet arrival process, due to limited sample size.

One of the striking features of Internet traffic is *Asymptotic Second-Order Self-Similarity*. Henceforth, the performance of MBAC algorithms is commonly evaluated under SS traffic conditions using a aggregate of multiplexed Markov ON/OFF sources - in general *without* evaluating if this condition actually is met. Based on this finding and with the aim to evaluate the impact of SS on MBAC performance, in consequent work, the aim of this paper is to empirically verify the robustness of SS traffic generation with this source model.

The remaining sections are Sec. II where we briefly introduce the concept of self-similarity, Sec. III with a presentation of results and finally Sec. IV with conclusions and a vista for future tasks.

II. BACKGROUND

Self-Similarity is a scaling phenomena - the shape of a feature repeats itself independent on the scale of observation. In a network context, a time series X of packet arrivals (or bytes) counted in intervals of fixed length τ, this means independence on τ - the resulting time series resembles itself for each intercall length. More formally, let $X = \{X(t), t \geq 1\}$ be a stationary process and let

$$X^{(m)}(k) = 1/m \sum_{t=(k-1)m+1}^{km} X(t), \quad k = 1, 2, 3 \ldots \tag{1}$$

be the level m aggregated sequence by dividing X in non-overlapping blocks of size k. The time series exhibit strict SS if it matches condition

$$X \doteq m^{(1-H)} X^{(m)} \tag{2}$$

However, this proved to be too strict - actual network traffic is *asymptotic* self-similar, i.e. with $m \to \infty$. The degree of SS is expressed by the Hurst parameter, $0.5 \leq H \leq 1$, and increases towards one. Clearly, the Auto-Correlation-Function (ACF) of SS processes resembles itself for each scale of observation. Consequently, network traffic is called *asymptotic second-order self-similar*.

Moreover, the ACF of network traffic exhibits a significant feature, it decays only very slowly to zero. Mathematically, this means strong *dependence* between samples (random variables) over large times - so-called LRD (Joseph Effect). Formally, the ACF reads as $C_X(k) \sim C_\gamma k^{(H-1)/2}$ with a finite constant C_γ and the Hurst H parameter like above. In the rest of the paper we use SS to refer to network traffic with SS and LRD effects.

Willinger et al. showed in [1] that SS network traffic can be generated using two state Markov ON/OFF source models. In ON state, a source emits packets on peak rate and no packets in OFF state. State sojourn times are drawn from two individually parametrised Pareto distributions with a CDF $F(x) = 1 - (\frac{b}{x})^\alpha$. Likely the major finding off their work is the relation $H = \frac{3-\alpha}{2}$.

Informally, SS and LRD makes traffic bursty over different time scales. However, mainly due to its asymptotic character and because of limited time scales of interest - real buffer sizes are inherently limited - the actual impact of SS and LRD is still an open issue [2]. Henceforth, the significance of H as a traffic parameter still remains open.

III. EVALUATION

As we are only interested in the arrival process, a simple scenario of two source nodes, one router and one destination suffices. Pareto ON/OFF (POO) sources are uniformly distributed on source nodes. The arrival process, number of *bytes*, is sampled at the router over intervals τ of $100ms$.

To simulate a realistic scenario, flow arrivals follow a Poisson process with mean μ_a. Holding times are exponentially distributed with mean μ_h. Characteristics of POO sources are configured by the shape parameter α, average ON and OFF time, t_{on} and t_{off} and peak rate ρ. Default settings are listed in Tab. I

Simulation duration is $140000s$ where the first $20000s$ are discarded to evaluate the system in equilibrium. As experimental framework we use the NS-2[2] network simulator.

The authors are with the Laboratory of Communications and Telematics, Centre of Informatics and Systems (CISUC), University of Coimbra, Portugal. E-mail: {tbohnert,edmundo}@dei.uc.pt.

[1] Not to confuse with Probe Based Admission Control. The latter is a special kind of MBAC

[2] http://www.isi.edu/nsnam/ns/

TABLE I

SIMULATION STANDARD SETTINGS

μ_a	μ_h	α	t_{on}	t_{off}	ρ
7.5	$300s$	1.2	$500ms$	$500ms$	$256Kbit/s$

As mentioned in Sec. II, the degree of SS is expressed in H. Thus, to evaluate SS in the arrival process we have to determine H. Using SELFIS[3], we choose four Hurst estimators: Abry-Veicth based on wavelets H_{av}, Periodogram H_p and the related Whittle H_w estimator, both computing in the frequency domain and the R/S statistic H, the initial method developed by the engineer Hurst.

A. *Robustness to Flow Arrival Patterns*

In the first simulations we evaluate the robustness of the SS phenomena using multiplexed POO sources for different flow arrival patterns. We choose the default setup, see Tab. I, and varied μ_a. Results are presented in Tab. II.

TABLE II

ROBUSTNESS TO DIFFERENT FLOW ARRIVAL PATTERNS

μ_a	p	H_{av}	H_p	H_w	H
15	20	0.928	1.005	0.879	0.754
7.5	40	0.930	1.016	0.884	0.719
3.75	80	0.925	1.003	0.879	0.677

This setup examines the dependency of H on mean traffic intensity. The mean number of flows in the system is $p = \frac{\mu_h}{\mu_a}$. A source sends with mean rate $\nu = \rho\frac{t_{on}}{t_{on}+t_{off}}$ and according default settings for t_{on} and t_{off}, in this case $\nu = \frac{\rho}{2}$. Intuitively, one would expect that with a growing number of active sources, the Hurst parameter increases too. However, except for H_w, for all other estimators the value of H and thus the degree of SS decreases with an increasing mean number of sources.

Regarding the estimated value of H and the analytical relation $H = \frac{3-\alpha}{2}$, it seems that only the AV-Estimator and the Whittle-Estimator approach the theoretical value of $H = \frac{3-1.2}{2} = 0.9$, while the R/S-Estimator seems to underestimate the degree with a large error.

However, despite of variation in estimated values, for this limited set of simulations, SS and LRD effects seem to be insensible to different values of μ_a, the issue under investigation in this paper.

B. *Robustness to Burst and Idle Times*

Statistical features of the traffic aggregate are cumulative outcomes of individual sources properties. For POO sources, these properties are determined by the triple t_{on}, t_{off} and α. With the next set of experiments we evaluate SS effects for different source individuals. As before, we apply the default settings, see Tab. I, but diversify t_{on} and t_{off}. Results are presented in Tab. III.

TABLE III

ROBUSTNESS TO BURST AND IDLE TIMES

t_{on}	t_{off}	p_{on}	H_{av}	H_p	H_w	H
200	100	$8.6 * 10^{-8}$	1.033	1.020	0.972	0.693
800	200	$1.32 * 4^{-4}$	0.955	1.042	0.916	0.675
200	800	$1.09 * 10^{-28}$	0.932	0.981	0.876	0.751

As a flow is a set of packet bursts (comprising one to n packets) with a maximum inter-burst time, these experiments can be considered as an increase of investigation granularity. A source can be in ON state with probability p_{on} and in OFF state with probability $1 - p_{on}$ where $p_{on} = \frac{t_{on}}{t_{on}+t_{off}}$.

For the first experiment with a very short burst period t_{on} and a still shorter idle time t_{off} the degree of SS clearly departs from the theoretical level. This suggests some sensibility of the SS degree to high frequent state switching.

The two remaining experiments are carried out to show the sensibility to somewhat extreme burst to idle time ratios. Clearly, the first source with and burst time t_{on} is much burstier as the last one. Interestingly, although the great difference in p_{on}, the values for H are relatively close and in contrary to the first set of experiments, the value of H increases with traffic intensity.

Finally, it is worth to notice that the first three estimators compute rather consistent estimates and only the R/S-Estimator departs significantly.

C. *Diversifying α*

As introduced in [1], the degree of SS is purely a function of α formulated by the functional $H = \frac{3-\alpha}{2}$. The results presented corroborate this result, especially as we have no measure of the accuracy of the applied estimators. To further foster this assumption, we finally examine the simulation default setup for different α levels. Results are presented in Tab. IV.

TABLE IV

DIVERSIFYING α

α	H_t	H_{av}	H_p	H_w	H
1.8	0.6	1.032	0.819	0.999	0.731
1.6	0.7	1.001	0.893	0.999	0.732
1.4	0.8	0.950	0.949	0.992	0.739

The results are somewhat surprising. As α is the crucial parameter, for non of the settings the theoretical values is approached. The reason can be either inaccuracy of the estimators, simulator implementation of POO sources or finally the experimental setup itself as each sampling setup is inherently subject to statistical noise.

IV. CONCLUSION

In conclusion we can say, that using POO traffic with Poisson arrival flow arrival processes and exponential distributed holding times generates SS traffic in a robust manner regarding parameter settings. Stepping ahead, another conclusion is, that this limited number of simulations does not allow any statement about the gain of Hurst parameter monitoring for congestion control mechanisms, as for example MBAC. No clear pattern was identifiable and some results were even contradictory to theoretic foundations. Henceforth, the only conclusion in this context is: Further rigorous investigation is necessary..

ACKNOWLEDGEMENTS

This work was supported by the Portuguese Ministry of Science and High Education (MCES), and by European Union FEDER under program POSI (project QoSMap).

REFERENCES

[1] Walter Willinger and Murad S. Taqqu and Robert Sherman and Daniel V. Wilson, Self-similarity through high-variability: statistical analysis of Ethernet LAN traffic at the source level, IEEE/ACM Transactions on Networking, vol. 5, num. 1, pgs 71-86, 1997.

[2] Finite Timescale Range of Interest for Self-Similar Traffic Measurements, Modelling and Performance Analysis, IEEE International Conference on Networks, ICON'03, September 2003, pp.7-12, Sydney, 2003

---

[3] http://www.cs.ucr.edu/ tkarag/Selfis/Selfis.html