# Characterization and Modeling of IP and protocol specific Internet Traffic using Chaotic Attractors

Eugenio Costamagna, Lorenzo Favalli, and Francesco Tarantola
Department of Electronics, University of Pavia
Pavia, Italy
{author}@ele.unipv.it

***Abstract* -- The behavior of IP based network traffic has been the subject of several works. In this paper we compare two modeling approaches and their ability to capture two basic traffic features: packet size distribution and interarrival times. Data traces have been measured at the border gateway of the campus network of the University of Pavia at different times of the day. The two modeling schemes use respectively well established Hidden Markov Model (HMM) techniques and a stochastic engine derived from chaotic attractors.**

## I. 1. INTRODUCTION

As the usage of data networks is rapidly increasing, there is a growing effort in trying to achieve a better knowledge of both the aggregate as well as single service characteristics. This in turn will provide a better ability to model traffic streams, which is the most basic ability in network operating and planning.

This traditional activity is almost non-existent in today's design of the Internet, where users implicitly accept all the ups and downs of a "best effort" technology. In perspective, Internet is expected to become the main connecting infrastructure, enabling a whole variety of applications, and the notion of "quality of service" (QoS) needs to be included in the Internet paradigm too. So, a large amount of efforts is currently being spent to determine suitable performance measurements, access schemes, and resource assignment procedures, to allow the sharing of this infrastructure among general users and privileged (paying) users, which need not to be trapped in congestions or suffer from sudden bottlenecks. The evaluation of such schemes requires the availability of accurate models of the traffic process.

Traditionally, stochastic models (and specifically traffic models) are often implemented in the form of Markov processes, as they provide a flexible framework that can be customized to fit different system properties. They are also attractive since the statistical behavior of the system can be directly derived from the model. Their drawback is the definition of the state diagram and of its transition probabilities to properly describe the target system. This often leads to large and complex structures with several parameters to be tuned. Moreover, it has been shown that packet traffic exhibits long-range dependence (LRD) that cannot be captured by Markov chains, as they can reproduce the autocorrelation function only in the initial, fast decaying steps [1]. Since LRD is related to some *physical* characteristic of traffic, several studies have been performed to develop models that are intrinsically self-similar.

The discovery of LRD characteristics has raised a strong debate whether or not Markov models, known to be SRD, could be used to engineer network parameters.

Traffic engineering has then be forced in the last ten years to take into account the new features of the traffic generated by interactive services. The first milestones in understanding the nature of traffic in data networks have been the works by Lealand et al. [1] who measured arrival rates on an Ethernet, by Beran et al. [1] characterizing Variable Bit Rate (VBR) video sources and by Park et al. [3] for WWW traffic. The main result derived by all these works is the delineation of a novel phenomenon that is since then referred to as *long range correlation* (LRD, Long Range Dependence) or *self similarity*.

This concept summarizes the high variability of network performance figures at any time scale of observation and is strictly related to the notion of *fractal* or more in general *multi-fractal* behavior, although the two concepts are not completely overlapping [4].

The next step is the effective use of these concepts to generate models of traffic behavior accurate enough for the novel network paradigm. Several approaches have been adopted ranging from use of Fractional Brownian Motion [5] to Heavy Tailed (HT) and Power-Tailed (PT) Distributions [6],[7], via superposition of on-off sources [8], Modulated Markov Processes with HT inter-events distributions and chaotic maps [9][10][11].

Researchers have also pointed out that the self similar behavior is exhibited both as an aggregate property as well as due to single streams, and seems to be mostly related to the interaction between the bursty user behavior with the segmentation and reassembly procedures intrinsic in the TCP-IP protocol stack.

For a thorough description of these and other characteristics, a must is the book by Park and Willinger [12].

It is interesting to notice that despite the overall acceptance of the notion of self similarity as an intrinsic characteristic of network traffic, some authors have pointed out that this may not affect actual performance [13].

Indeed in [14] we too found that a Markov based approach considering hidden Markov models (HMM) [15] of VBR traces was able to provide relatively good long-range characteristics and very good results in a network environment completely simulated with *Network Simulator 2* (NS2). In the same work we compared a chaos based [16] modeling approach that on the contrary provided good LRD features with acceptable but somehow worse result in terms of final delay-loss performances.

These outcomes have led us to start an extensive measurement campaign conducted at the edge router of the campus network at the University of Pavia, to investigate the actual behavior of traffic and to gain insight about the models properties.

In this paper, we compare the ability of a Markov based approach with that of a chaotic one in modeling aggregate traffic. It is only a preliminary work, since only relatively short sequences of aggregate traffic have been taken into account, and it describes some of the first attempts to tune the chaotic procedure. Nonetheless, it will be shown that a low dimensionality Markov model can describe most of the statistical features of the analyzed sequences, and that the chaotic models supply appreciable features in terms of long range statistics.

## II. 2. MODELING APPROACH

Markov models have been applied since the beginning to describe the behavior of traffic events. We tried to use Hidden Markov Models (HMM). In this case, there is no more correspondence between a state and a physical event and the properties are hidden in the model structure, from which the name of hidden Markov models. From a mathematical point of view, an HMM can be described as a 5-ple, $\lambda=(S,V,A,B,\Pi)$, where S is the set of the states of the system, with cardinality N; V is the set of observable values ("the alphabet"), with cardinality M; A is the state transition probabilities matrix; B is the observable probabilities matrix, and $\Pi$ is the initial state vector. The model parametrization is commonly performed using algorithms such as the method of moments or gradient or the Baum Welch (BW) procedure used in this work.

The BW algorithm is very robust in that it always converges, but there is no guarantee that it converges to a global maximum: the optimized parameters may consequently be not the optimal ones in an absolute sense. Furthermore, there is no guidance in the best selection of the number of states. For what the convergence speed is concerned, it is highly influenced by the size of the alphabet of symbols V.

For the chaotic model, the approach is similar to that used in [16], which was derived straightforwardly from the experience gained in simulating error gap series in mobile radio channels and the attenuation process in satellite links. It stems from the observation that a chaotic attractor provides a description of a dynamical system joining a deterministic mechanism to an unpredictable behavior. This translates in a number of properties, the most important being the fact that small differences in sampling times along the trajectory result in completely different evolutions of sampled series.

The modeling procedure is thoroughly described in [3] and is based on a weighted sum of components derived from seven Lorenz strange attractors [9] as in the following formula:

$$F\ [\ f\ \_{i}(k_i\ u_i\ )] \qquad (1)$$

where:

- f (•) is a polynomial or exponential function: we present here results obtained with exponentials;
- the $u_i$ variables are selected coordinates or geometric distances sampled on the attractor trajectories (the x coordinate in the current case), and the $k_i$ are suitable weights;
- index i is ranging from 1 to 7;
- Weights and sampling distances are optimized to match the characteristics of the target sequence;
- F(•) is a probability shaping function, used to match the distribution of the sample amplitudes derived from (1) to the target time series. This function is obtained, after some preliminary sample generation runs of the model, by means of an inversion procedure for the cumulative probability function described in [16]. The whole procedure has been found to result in a very good correspondence with the target in all cases.

Note that we are not attempting to describe any *physical* characteristic of the underlying system. To optimize the parameters in (1), various cost functions have been implemented, tailored to the actual problem: in the current case, emphasis was done to medium term moving correlations, short term moving covariances and histograms, and to mean differences, averaged over the sequence length, among samples taken at medium term distances, i.e., 20, 100 and 1000 samples.

## III. 3. MEASUREMENTS SET UP AND RESULTS

The campus network at the University of Pavia comprises several fast Ethernet branches connected by switches to a single gateway that routes the outgoing traffic to the backbone national infrastructure. For a complete network plan see [10]. Our monitoring station running `tcpdump` was inserted just prior the gateway so that it could monitor all traffic over the network. The analysis has been performed at different times of the day and the large amount of data has been filtered to extract data subsets relative to different parameters at the aggregate level. Work is in progress to classify the data according to protocol and service, and to identify specific performance measures for a connection carrying H323 service between the campus in Pavia and a remote campus in Mantova: the two premises are about 100 Km apart and are connected via a 2Mb HDSL link. Among all the possible traffic characteristics, our work has been focused on two parameters: packet size and packet interarrival time. The two modeling approaches presented in the previous section have been customized as follows:

- HMM: a number of states from 2 to 10 has been used. Since the results were not greatly improved for larger numbers of states this has been considered to be the optimal number. The dynamic of the data sets needs also to be discretized for the Baum Welch algorithm to converge in reasonable time. The presented results are for a number of symbols in the alphabet equal to 50. With these numbers, a sequence of about $10^6$ events is processed in a few hours on a 1.5 GHz PIV computer.

- CHAOS: the optimization procedure of the generator is currently based on the experience gained working on radio channels, mainly with the Nelder and Mead simplex procedure [19]. Consequently the optimization procedure is not effective in avoiding to be trapped in local minima, and the quality of the results of any optimization run has to be carefully evaluated.

Apparently, for traffic sequences Lorenz attractors lead to results better than the Duffing one, and five attractors instead of seven in the sum (1) seem to provide results somewhat similar: to determine the minimum sufficient number of attractor remains an open issue. Probably, when suitable or near suitable numbers of attractors are imposed, the optimization procedure is more effective with the lower numbers than with the higher. We program to undertake experiments with tabu search procedures to compare the results, as for error gap series in [16].

The lengths of the target and HMM sequences are of the order of 500,000 samples; the lengths of the chaotic sequences are of the order of 50,000 samples. The optimization procedure of these requires CPU times comparable to HMM sequences, but some care in manipulating optimization weights. The sample dimensions are normalized for all sequences, to have maximum values of about 1,500 units.

*1) Packet sizes, aggregate level.*

In Figure 1. we compare to the target the results obtained for moving autocorrelation functions computed over intervals of 2000 samples and averaged for two chaotic and one HMM sequences. As seen, the results are very good, as are for the cumulative distributions (not reported). As expected, notable differences are found in Figure 2. where the variation coefficients are shown: the HMM is unable to produce a curve similar to the target, and only the exponential function version of the chaotic model is capable of mimicking it

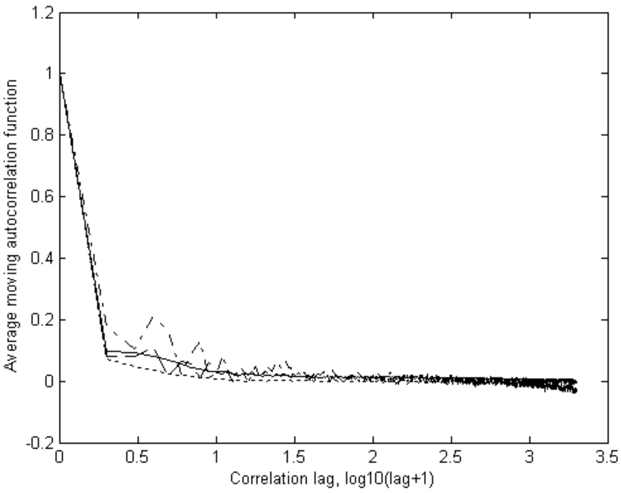


Figure 1. .Moving autocorrelation function for packet size sequences averaged over the whole sequence. Solid line: target sequence; dashed line: chaotic sequence, exponential functions (see text); dash/dotted line: chaotic sequence, square functions; dotted line: HMM

These behaviors are confirmed by the strongly related Variance-Time plots in Figure 3. The asymptotic behaviors of these are related in turn to the Hurst parameters of the process, and the curves lying in the upper part of the right hand figure are more likely to exhibit LRD properties.

In Figure 4. we show the appreciable result obtained in terms of averaged moving covariances (six windows) for the chaotic sequence, exponential functions. The corresponding result obtained using square functions is shown in Figure 5. with more objectional features. The HMM sequence leads to curves very similar to the target.

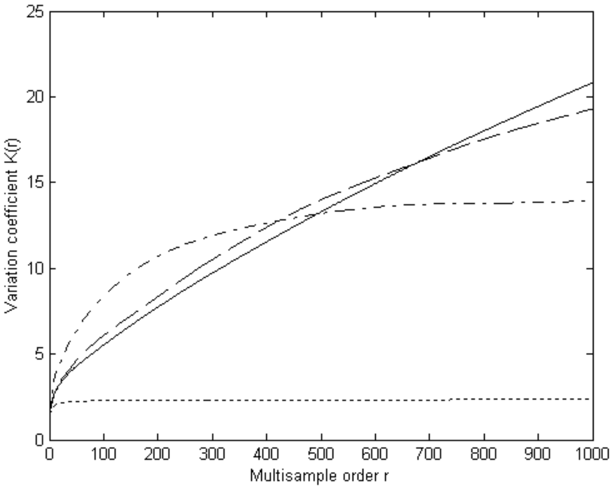


Figure 2. Variation coefficients for the same sequences in figure 1. Solid line: target sequence; dashed line: chaotic sequence, exponential functions (see text); dash/dotted line: chaotic sequence, square functions; dotted line: HMM

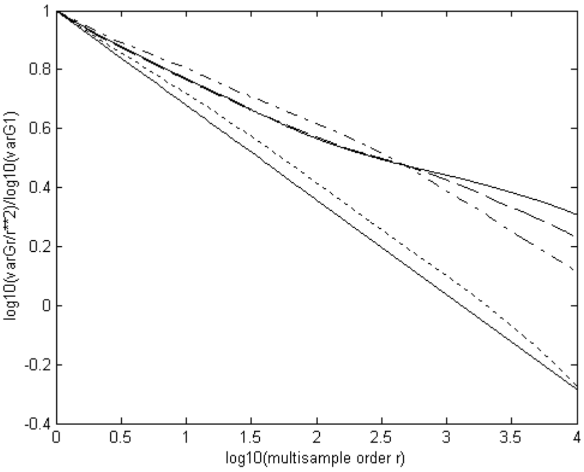


Figure 3. Variance time plots. Solid line: target sequence; dashed line: chaotic sequence, exponential functions (see text); dash/dotted line: chaotic sequence, square functions; dotted line: HMM

*2) Interarrival times, aggregate level.*

From Figure 6. to Figure 10. we present for interarrival time sequences relative to the above packet size sequence the same comparison as in the previous figures. The same comments remain appropriated.

*3) Packet sizes, WWW traffic.*

Figure 11. and Figure 12. present the comparison of the Variation coefficients and of the Variance-Time plots, using only the exponential version of the chaotic model.

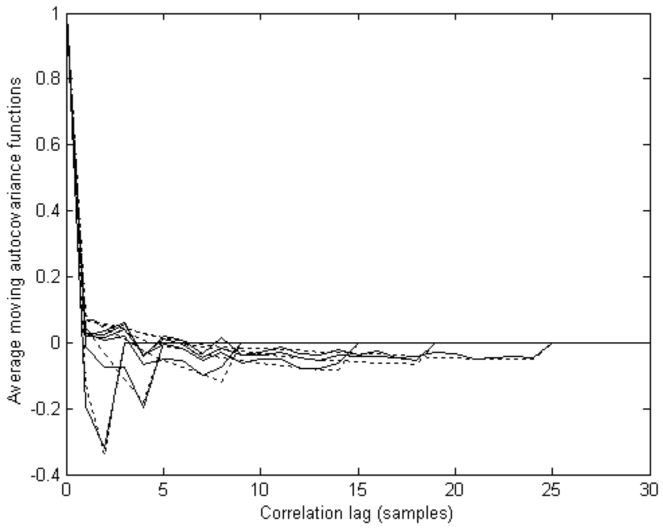


Figure 4. Moving Covariances. Solid line: chaotic sequence; dashed line: target sequence

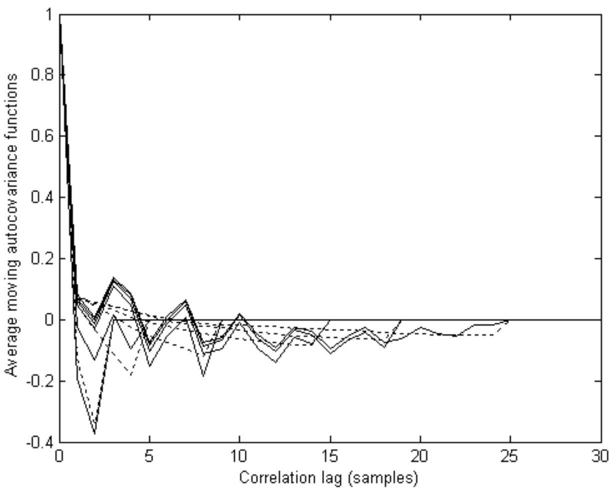


Figure 5. Moving Covariances. Solid line: chaotic sequence; dashed line: target sequence.

The behaviors looks similar to the corresponding curves for aggregate traffic, and similar are the performances of the models.

In Figure 13. the cumulative distributions are shown, corroborating the past experience on good performances from all models in mimicking this feature.

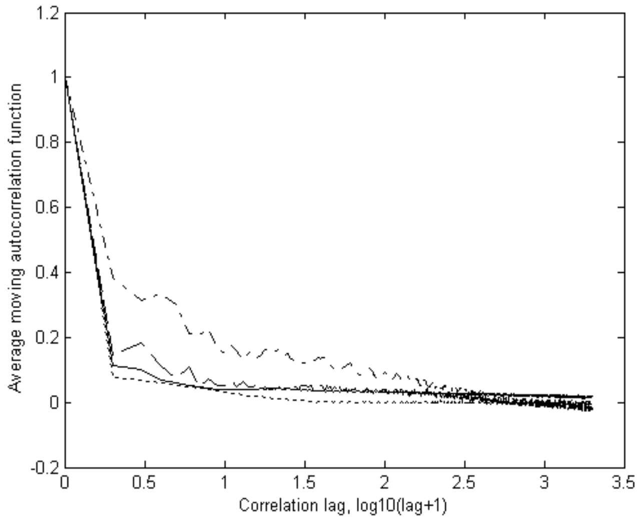


Figure 6. Moving autocorrelation function for interarrival time sequences averaged over the whole sequence. Solid line: target sequence; dashed line: chaotic sequence, exponential functions (see text); dash/dotted line: chaotic sequence, square functions; dotted line: HMM

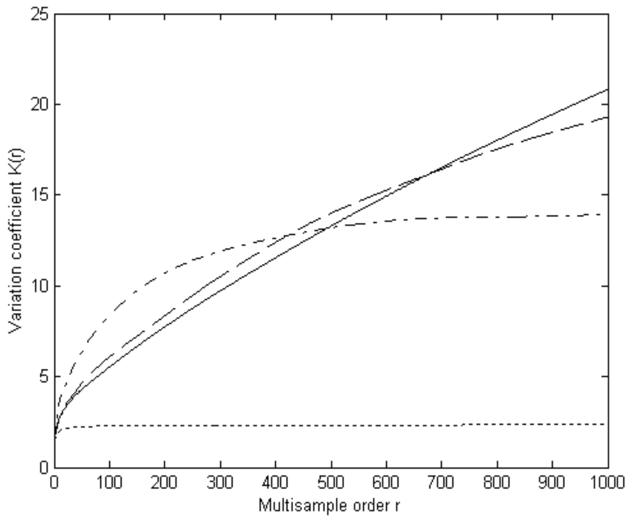


Figure 7. Variation coefficients for the same sequences in figure 6.

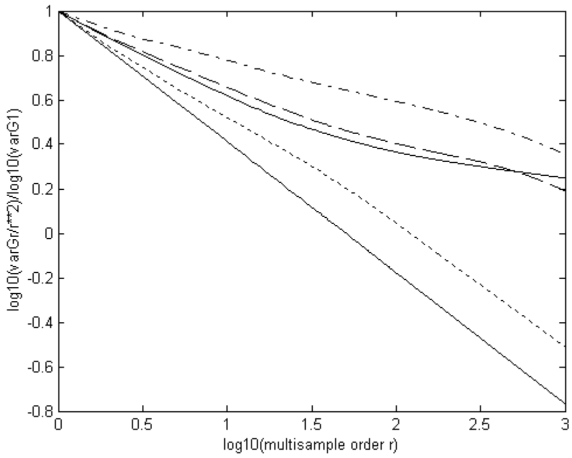


Figure 8. Variance time plots. Solid line: target sequence; dashed line: chaotic sequence; dotted line: HMM

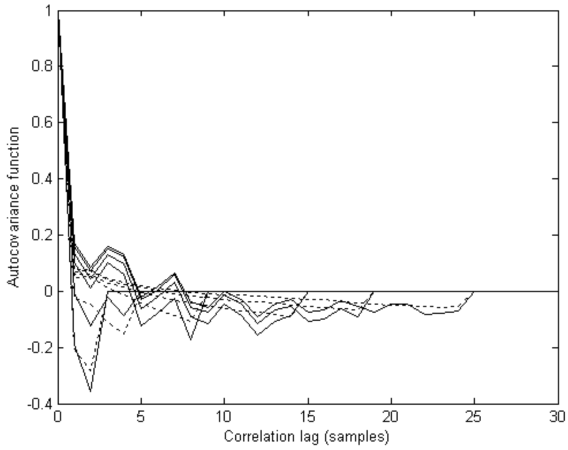


Figure 10. Moving covariances. Solid lines: chaotic sequence; dotted lines: target sequence

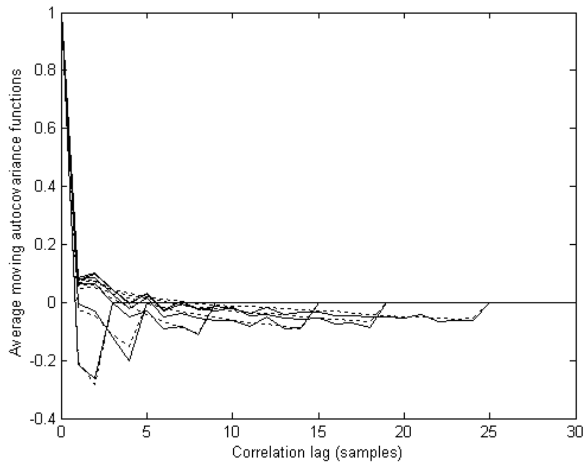


Figure 9. Moving covariances. Solid lines: chaotic sequence; dotted lines: target sequence

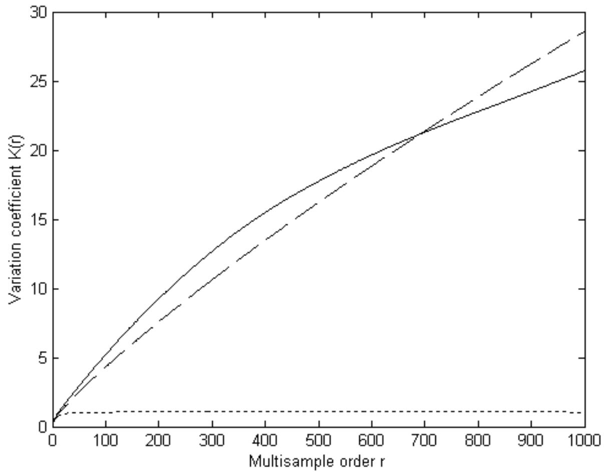


Figure 11. Variation coefficients for the WWW traffic sequence. Solid line: target sequence; dashed line: chaotic sequence; dotted line: HMM

## IV. NETWORK SIMULATION

To verify the ability of the models to predict the behavior of the queue buffer, a very simple network has been implemented in Network Simulator 2 (NS2) with our traces mixing in a node with other traffic having a Pareto distribution. We look here to the queue behavior in infinite buffers to enhance the effects of the trace characteristics, avoiding any queue size reset due to limited buffer capacities.

We introduce in Figure 14. the buffer queue behavior, as derived by the Network Simulator, when supplied with measured traffic, i.e., a measured packet size trace and a companion interarrival time trace, which will be recalled in the following as *target*. Trace length is of the order of 500,000 samples, and simulations was performed to 1000 seconds of traffic.

Supplying the simulator with a packet size and an interarrival time series both derived from HMMs, we obtained the buffer queue curve shown in Figure 15. We see that some features of the curve in Figure 14. are mimicked, although the whole result is not very exciting. However, when the HMM packet size trace is supplied together with the measured interarrival times, the result, reported in Figure 16. becomes very good.

Supplying the simulator with chaotic packet sizes and interarrival times, the result is very modest, as shown in Figure 17. and no better behavior is shown when measured interarrival times are used. This in spite of the fact that the time behavior of the chaotic trace looks better than the HMM and better statistics for either packet lengths or interarrival times were obtained in [16]: probably, using simulated traces for both the input sequences, any correlation between them is lost, while this correlation is important when assessing the effects of packet arrivals on queue sizes.

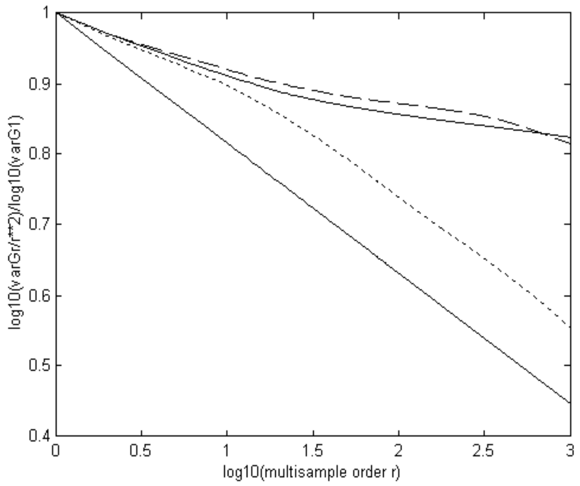


Figure 12. Variance time plot. Solid line: target sequence; dashed line: chaotic sequence; dotted line: HMM

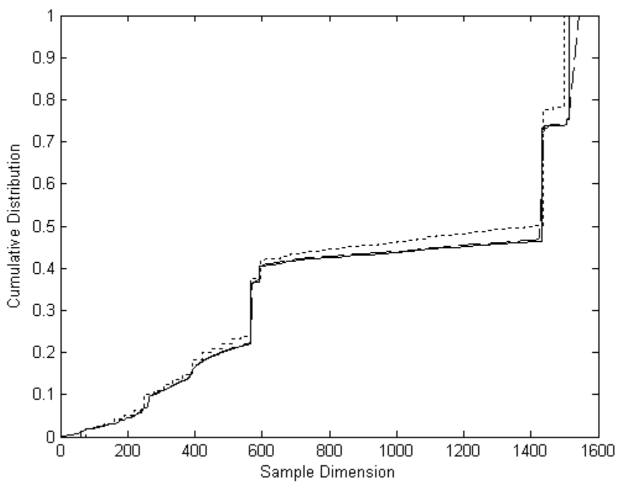


Figure 13. Cumulative distribution function. Solid line: target sequence; dashed line: chaotic sequence; dotted line: HMM

This can contribute to explain the failure of the chaotic trace, as no connection with local properties of the current target trace is pursued in simulation, and only a likely realization of the stochastic process is generated.

This is supported by the results obtained with a different set of traffic measurements, in which the only interesting buffer size trace was obtained supplying to the network HMM simulated interarrival times and measured packet sizes, and again the traces from chaotic sequences were even worse than from HMMs.

A better insight can be gained reducing the process to the observation of a single time series, for instance, generating from the measured packet sizes and interarrival times volume sequences sampled at fixed time. In the following we use the same target sequences of above to derive traffic samples (i.e., cumulated number of bytes) at any 3 ms intervals and the HMM and chaotic sequences generated for the new target show the time behavior as in Figure 18. The variation coefficient curves for the HMM and the two chaotic sequences are shown in Figure 19. and Figure 20. (the scale is determined by the amplitude normalization used in the computation). As for the previous simulated traces, the HMM curve is significantly lower than the target. A better result was obtained from the chaotic model (upper curve in Figure 20. ) but, to investigate the effects of this parameter, a chaotic trace with very low variation coefficient has been generated too. Building on the flexibility of the chaotic model, this was obtained without significant changes in the other statistics. With reference to [19], we optimize weighted sums of exponential of vector moduli sampled on the trajectories of seven attractors.

The queue sizes after supplying the Network Simulator (in the same simulation conditions) are shown in from Figure 21. to Figure 24. We note that the chaotic trace is very good when variation coefficients are well matched to those of the target: also the peak values and the distribution of the time intervals between them are well represented. With the variation coefficients of the lower trace in Figure 20. the overall appearance of the queue trace is somewhat similar, but the peaks are lost. The HMM curve captures significant features of the target, but the ratio between peaks and mean value is completely different, and the local appearance of the draw is objectionably dense thus accounting for the non-bursty behavior of the model tied to its poor LRD characteristics.

Notice that various experiments show that the variation coefficient behavior is by far more effective in causing these effects than accuracy in matching other details of the target statistical features, as ringing windowed correlation or covariance functions.

## V. CONCLUSIONS

A first assessment of the capabilities of HMM and of chaotic models to mimic relevant features of aggregate and WWW traffic sequences has been performed. We tried to mimic statistical features defined according to previous experiences in modeling error gap processes and attenuation time series. At first glance, the models show behavior similar to those already experienced, very good for short term statistics in the case of HMMs, and good for long term features in the case of chaotic models. The ability of different modeling approaches to predict the queueing behavior in an IP network has also been investigated feeding the traffic traces to the input of a FIFO queue of infinite length have been presented. The more important findings can be related to the fact that the Markov based approach by itself is unable to predict the largest peaks of buffer occupancy unless used in some hybrid way with real data. On the contrary, the chaos based traffic generator has proven itself to be very reliable and robust provided that it possesses a complete information of the traffic process: under these conditions, the simulations performed apparently show that a relevant part in the optimization of the model relies on the so called variation coefficient. Further studies are needed to achieve a better insight of this behavior and to compare the proposed chaotic model with other traffic models in the literature.

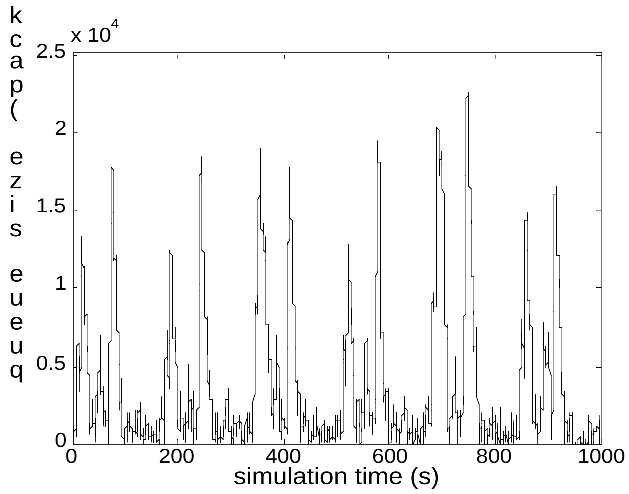


Figure 14. Buffer queue size over 1000 s of simulation, using target sequence 1.

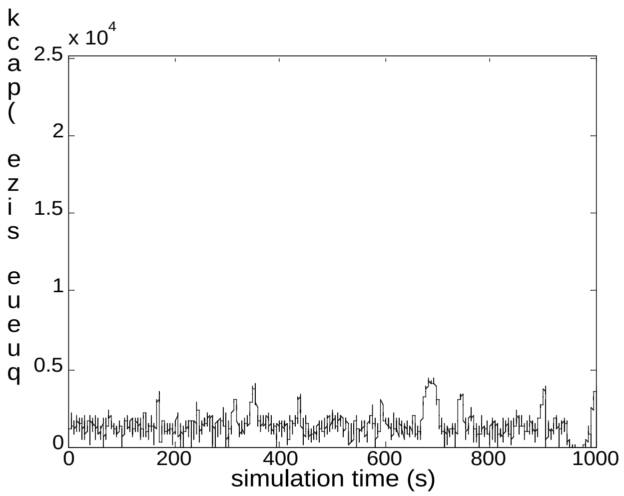


Figure 15. Buffer queue size, HMM sequence

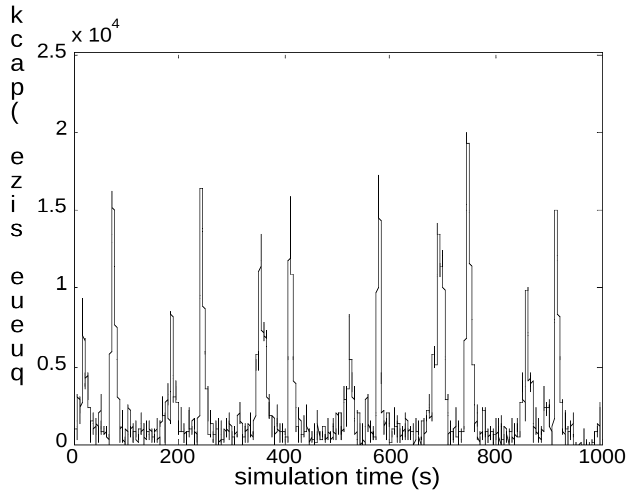


Figure 16. Buffer queue size, using target sequence for interarrival time, HMM for packet size.

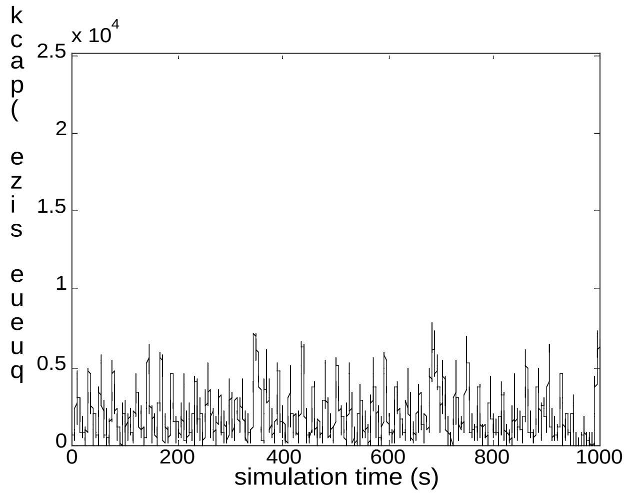


Figure 17. Buffer queue size, chaotic sequence.

## ACKNOWLEDGEMENTS

The authors wish to acknowledge the kindness and competence of the Director of the 'Centro di Calcolo', Dr. Flavio Ferlini and of Dr. Marisa Alicanti. Thanks are owed also to Mr. Massimiliano Oldani and Mr. Federico Zingale, who performed the measurements in partial fulfillment of their graduation.

## .REFERENCES

[1] W.E. Leland, M.S. Taqqu, W. Willinger, D.V. Wilson, "On the self similar nature of Ethernet traffic (Extended version)," IEEE/ACM Tr. In Networking, Vol. 2, No. 1 , Feb 1994, pp. 1 –15.

[2] J. Beran, R. Sherman, M. S. Taqqu, W. Willinger: "Long-range dependence in variable bit rate video traffic," *IEEE Trans. Comm.*, Vol. 43, No. 2/3/4, pp. 1566-1578, Feb.-Apr. 1995.

[3] K. Park, G. Kim, and M.E. Crovella, "On the Relationship Between File Sizes, Transport Protocols, and Self-Similar Network Traffic," Boston University Technical Report, BU-CS-96-016. available at http://www.cs.bu.edu/faculty/crovella/paper-archive/files-protocols/TR-96-016.ps

[4] P. Abry, R. Baraniuk, P. Flandrin, R. Riedi, D. Veitch, "The Multiscale Nature of Network Traffic: Discovery, Analysis, and Modelling," *IEEE Signal Processing Magazine* vol 19, no 3, pp 28-46, May 2002.

[5] K.R. Krishnan, "A new class of performance result for fractional brownian traffic model," *Queueing System* , No. 22, 1996.

[6] Z. Xiaoyun, Y. Jie, J. Doyle, “Heavy tails, generalized coding, and optimal Web layout,” *IEEE INFOCOM 2001*, vol. 3 , pp. 1617 –1626.

[7] M. Greiner, M. Jobmann, L. Lipsky, “The Importance of Power-Tail Distributions for Modeling Queueing Systems,” *Operations Research*, Vol 47, No. 2, March-April1999. available from http://www.eng2.uconn.edu/~lester/papers/

[8] P. Pruthi, A. Erramilli, “Heavy-Tailed ON/OFF Source Behavior and Self-Similar Traffic,” *Proceedings of the IEEE ICC'95*, Seattle, June 18-22, 1995, pp. 445 –450.

[9] A. Veres and M. Boda, “The chaotic nature of TCP congestion control,” in *Proc. IEEE Infocom 2000*, Tel Aviv, Israel, 26-30 Mar. 2000, pp. 1725-1723.

[10] M. Conti, E. Gregori, A. Larsson: “Study of the impact of MPEG-1 correlations on video-sources statistical multiplexing,” *IEEE Journal on Sel. Areas in Comm.* , Vol. 14, No. 7, pp. 1455-1471, Sept. 1996.

[11] E. Costamagna, L. Favalli, P. Gamba, and G. Iacovoni, “A simple model for VBR video traffic based on chaotic maps: validation through evaluation of ATM multiplexer QoS parameters,” *IEEE - ICC’98*, Atlanta (GA), June 7-11, 1998, pp. 568-572.

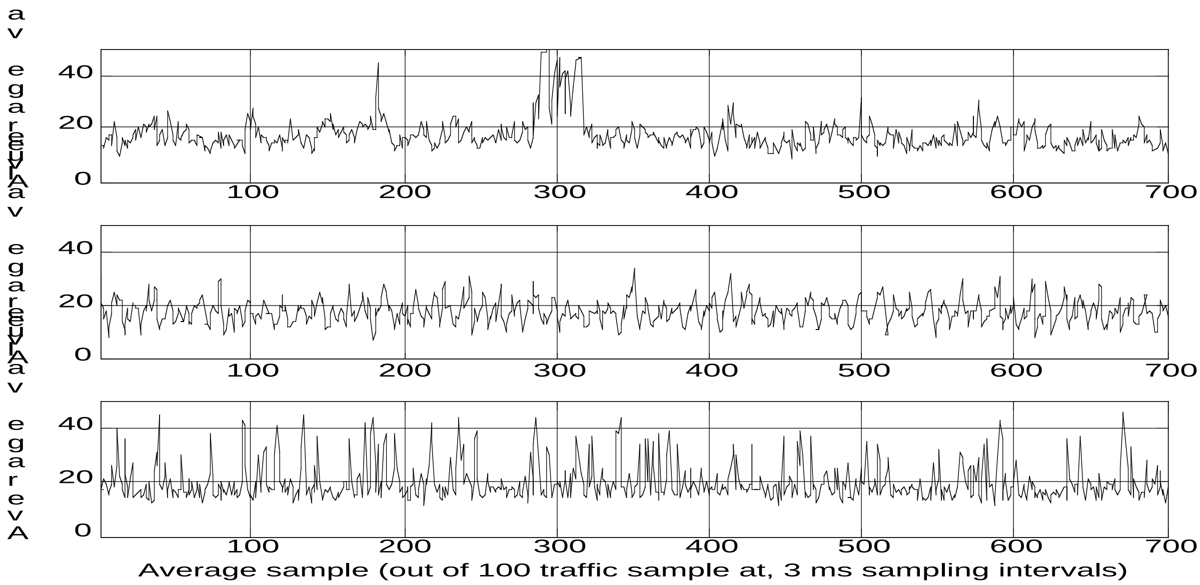


Figure 18. Average value of target, chaotic and HMM sequences from top to bottom.

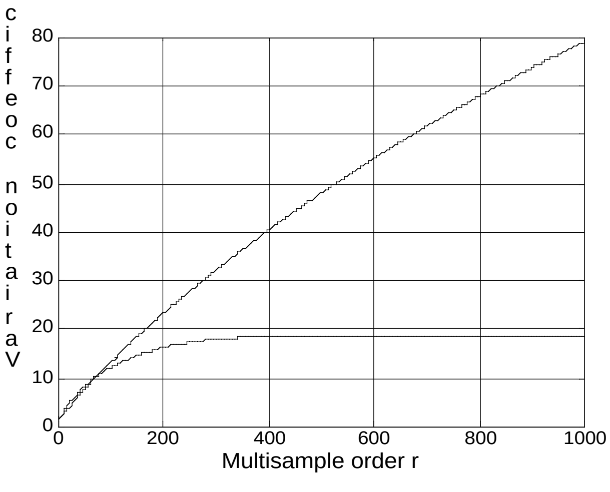


Figure 19. Variation coefficients of target (solid line) and HMM (dashed line) sequences

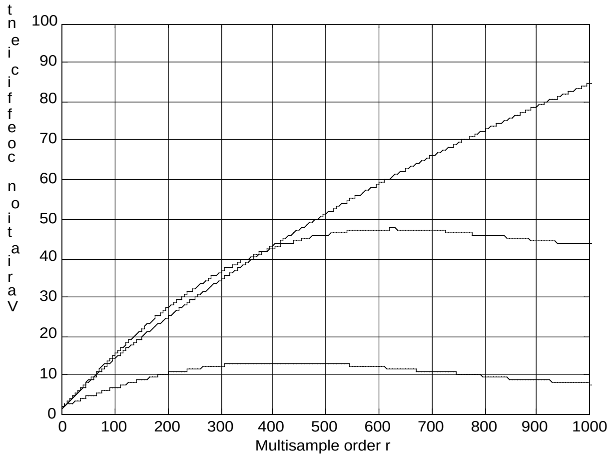


Figure 20. Variation coefficients of target (solid line) and two chaotic sequences (dashed and dashed-dotted lines)

[12] K. Park, and W. Willinger, eds., *Self-similar network traffic and performance evaluation.* J. Wiley and Sons, 2000.

[13] D.P. Heyman, and T. V. Lakshman, “What are the implications of long range dependance for VBR-video traffic engineering?,” *IEEE/ACM Trans. on Networking*, vol. 4, N. 3, pp. 301-317, June 1996.

[14] E. Costamagna, L. Favalli, F. Tarantola , “Modeling Of Multiplexed Video Streams In Ip Networks,” *IEEE Packet Video 2002*, Pittsburg, PA, 24-27 aprile 2002.

[15] L.R Rabiner and B.H. Juang, “An introduction to hidden Markov models,” *IEEE ASSP Mag.*, vol. 3, pp. 4- 16, Jan. 1986.

[16] E. Costamagna, L. Favalli, and P. Gamba, “Multipath Channel Modeling With Chaotic Attractors,” *Proc. IEEE*, Vol. 90 pp 842-859, May 2002.

[17] H. O. Peitgen, H. Jurgens, and D. Saupe, *Chaos and Fractals: New Frontiers of Science*, Springer-Verlag, 1992.

[18] http://www.unipv.it/CDC/ccframe.htm

[19] W.H. Press, S.A. Teukolsky, W.T. Wetterling and B.P. Flannery, *Numerical Recipes in FORTRAN, the Art of Scientific Computing*. Cambridge University Press, 1992.

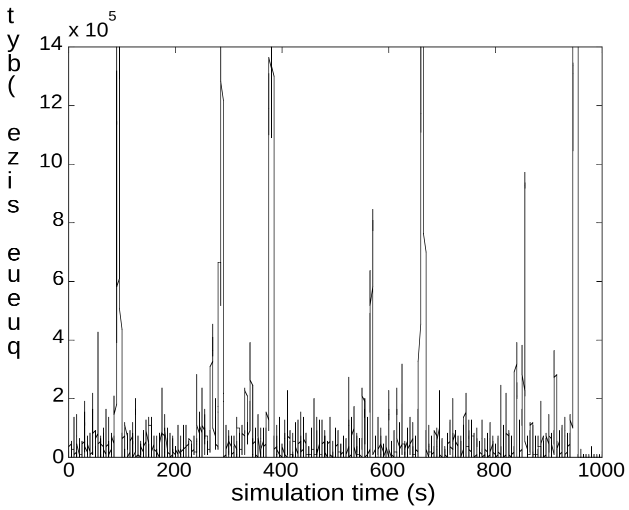


Figure 21. Buffer queue size over 1000 s of simulation at fixed interarrival time of 3 ms with the target sequence.

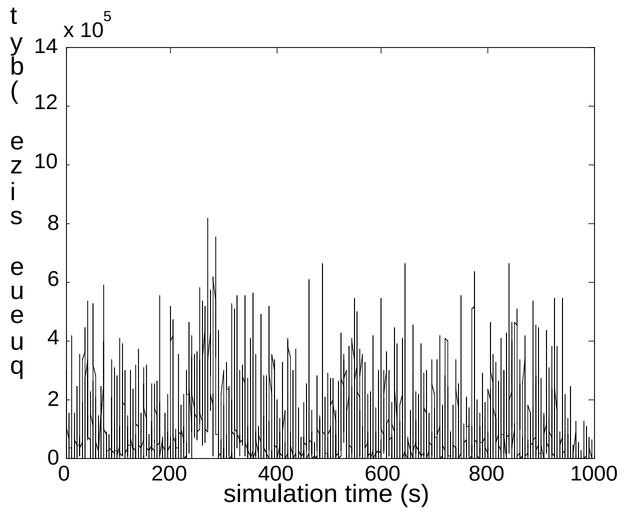


Figure 22. As in Fig. 21, but HMM sequence.

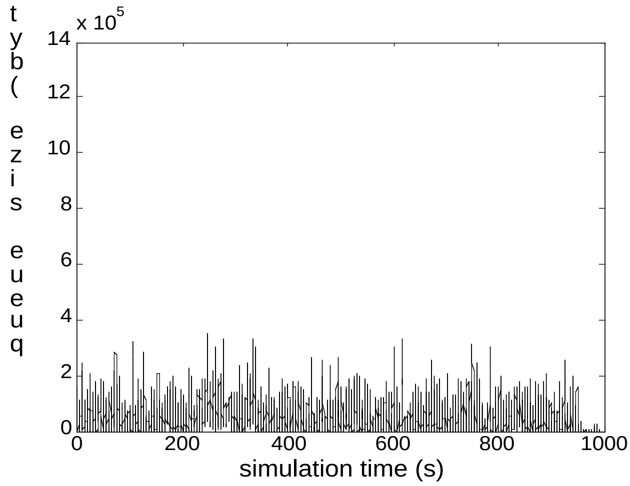


Figure 23. As in Fig. 21 and 22 but chaotic trace with low variation coefficients.

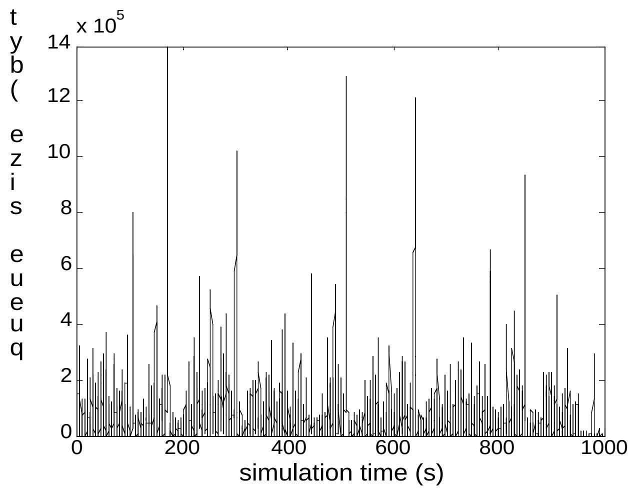


Figure 24. As in Fig 21 and 22 but with chaotic trace with variation coefficients matched to the target.