A New Multifractal Traffic Model and Its Impacts on Buffer Queueing Performances

Jeferson Wilian de Godoy Stênico
School of Electrical and
Computer Engineering
State University of Campinas - UNICAMP
PO Box 6101-13.083-970, Campinas, SP, Brazil
jstenico@decom.fee.unicamp.br

Lee Luan Ling
School of Electrical and
Computer Engineering
State University of Campinas - UNICAMP
PO Box 6101-13.083-970, Campinas, SP, Brazil
lee@decom.fee.unicamp.br

Abstract—**In this paper, we propose a new analytical expression for estimating byte loss probability at a single server queue with multi-scale traffic arrivals using Pareto distribution. In order to make the estimation procedure numerically tractable without losing the accuracy, we assume and demonstrate that an exponential model is adequate for representing the relation between mean square and variance of Pareto distributed traffic processes under different time scale aggregation. Extensive experimental tests validate the efficiency and accuracy of the proposed loss probability estimation approach and its superior performance for applications in network connection with respect to some well-known approaches suggested in the literature.**

Keywords — Multiscale Traffic Modeling; Loss Probability; Network Traffic

I. INTRODUCTION

Today communication networks must deal with a considerable number of different types of traffic, sharing common resources through statistical multiplexing. The efficient sharing of network resources depends on the statistical characteristics of traffic.

Historically the stochastic models widely used to characterize network traffic have been based on Poisson-like approaches, and more generally, Markovian modeling. Later, Leland et al [8] and other subsequent studies have demonstrated that traffic traces of modern high speed data networks exhibit fractal properties, such as self-similarity and long-range dependence (LRD) [13]. These new statistical traffic features, inadequately modeled by classical Poisson and Markov models, can strongly impact the performance of networks [12].

In contrast to the self-similar or monofractal behavior, some recent studies suggest that the measured TCP/IP and WAN ATM traffic flows exhibit a more complex scaling behavior, which is consistent with multifractals [5, 14]. Multifractal based traffic modeling is more general than the monofractal based and provides a more accurate and detailed description of network traffic series in different time scales [16].

Even taking into account the influence of the long-range dependent characteristics, the expected queuing behavior in buffer still cannot be adequately modeled without considering the multifractal nature of traffic [13].

The loss probability and packet delay are key performance measures related to quality of service (QoS) in computer networks, such as TCP-IP and ATM. Several Studies have been conducted in order to characterize the average size of the queue and the distribution of the number of packets in the buffer [4][9][13][15].

In [4] some lower bounds of the loss probability for self-similar processes were derived. In [13] a non-asymptotic multiscale analysis on some queue models based on multifractal cascade concepts were performed. Interesting enough, the analysis is valid for any buffer size, i.e., the approach, named Multiscale Queuing, incorporates the distributions of traffic.

In [9] the authors describe a statistical model for multi-scale traffic, deriving an equation for calculating the loss probability, whereas the input process traffic has lognormal distribution and that only the first two moments are sufficient to characterize the process of traffic. The derived analytical equation for the loss probability estimation is relatively complex and presents convergence when it is used numerically.

Moreover, in [15] the authors used an exponential approximation to model the second-order moment of the traffic process and derived an analytical expression for the loss probability estimation in a single server queue. It was assumed that the input traffic has a lognormal distribution. The derived analytical formula is computationally attractive overcoming the shortcoming of the analytical loss probability estimation equation obtained in [9] in terms of simplicity, accuracy and rapid convergence.

In this paper, we present a new approach for loss probability estimation in a single server link. We consider that the input traffic has a Pareto Distribution. We show how to get the estimates analytically once we assume multi-scale input traffic. Based on this analytical method, we evaluate its potential applications for control admission especially when networks traffic holds multi-scaling characteristics.

The paper is organized as follows: in Section II, we present the definition of the multi-scaling traffic processes, review some their major concepts and analyze the characteristics of the second-order statistical moments. In Section III, we present the derivation of the analytical expression for the loss probability estimation in a single server queue and our proposal for simplifying the analytical expression. In Section IV, we compare the proposed method to some well-known traffic models mentioned in the literature. Finally in Section VI we present our conclusions.

II. MULTI-SCALING TRAFFIC PROCESSES AND THEIR CHARACTERISTICS

Definition 1: Let $X(t)$ be the traffic rate at *t*. Then $W(t) = \int_0^t X(t)dt$ will be the arriving load up to *t*. Denote by $V(t,\Delta t) = W(t+\Delta t) - W(t)$. The average traffic rate is $\lambda = lim_{\Delta t \to \infty} V(\Delta t)/\Delta t$.

Let μ and σ^2 represent the mean and the variance of $V(\Delta t)$.

Given $T > 0$, the accumulative process $W(t)$ is said to be a multi-scaling process at time scale T if all of the following condition are satisfied:

- $W(t)$ has a stationary increment at time scale T, i.e., $V(t,T) = V(t)$.
- $V(t)$ has a Pareto distributed density function with parameter α and k: $f_{V(t)}(v) = \frac{\alpha k^\alpha}{v^{\alpha+1}}$.
- $\mu = \lambda T$.
- There exist an integer $M > 0$, a set $A = \{\beta_i(T): 0 < \beta_i(T) < 1, i \le M\}$, a set $\Phi = \{\phi_i(T): 0 < \phi_i(T) < 1, i \le M, \sum_{i=1;M} \phi_i(T) = 1\}$, and a small constant $\varepsilon > 0$ such that for any $\tau \in \{\tau: T - \varepsilon < \tau < T + \varepsilon, \tau > 0\}$ such that

$$\sigma^2 \sim \sum_{i=1}^{M} \phi_i(T)\tau^{2\beta(T)} \quad (1)$$

The expression (1) means there exists a probability measure for set *A*, and $\beta_i(T)$ occurs with probability $\phi_i(T)$. The continuous version of (1) is

$$\sigma^2 \sim \int_{-\infty}^{+\infty} f_{A(T)}(\beta)\tau^{2\beta} d\beta \quad (2)$$

where $f_{A(T)}$ denotes the probability density function of the scaling exponents $\beta(T)$. Notice that symbol "~" in (2) has the following interpretation: $x(p) \sim y(p)$ implies $lim_{u \to p}(x(u)/y(u)) = c$, where $0 < c < \infty$ is a constant.

A. Second-Order Moments of the Multi-Scaling processes

For simplicity, we assume that the scaling exponents $\beta(T)$ at time scale T of a traffic process follow a normal distribution $N(\tilde{\alpha}, \tilde{\sigma}^2)$ with mean $\tilde{\alpha}$ and variance $\tilde{\sigma}^2$. Here we omit the subscript T for $\tilde{\alpha}$ and $\tilde{\sigma}^2$. Therefore, the variance of the distribution σ^2 of the traffic process at time scale T can be represented as:

$$\sigma^2 \sim \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\tilde{\sigma}} exp\left[-\frac{(\beta-\tilde{\alpha})^2}{2\tilde{\sigma}^2}\right] T^{2\beta} d\beta \quad (3)$$

Let $z = T^{2\beta}$, then $\beta = ln(z)/(2ln(T))$ and $d\beta/dz = dz/(2ln(T)z)$. Then Equations (3) becomes

$$\sigma^2 \sim \int_0^{\infty} z \frac{1}{\sqrt{2\pi}(2ln(T)\tilde{\sigma})z} exp\left[-\frac{(ln(z) - (2ln(T)\tilde{\alpha}))^2}{2(2ln(T)\tilde{\sigma})^2}\right] dz \quad (4)$$

The right hand side of Eq. (4) shows that σ^2 simply has a log-normal distribution $L(\varpi, \theta)$ with parameters $\varpi = 2ln(T)\tilde{\alpha}$ and $\theta = (2ln(T)\tilde{\sigma})^2$. For the log-normal representation given by (4), a simple calculation can show that the mean μ and variance σ^2 of the distribution of the multi-scaling increment traffic process at time scale T are related to ϖ and θ as:

$$\mu = exp(\varpi + \theta^2/2) \quad (5)$$

and

$$\sigma^2 = exp(2\varpi + \theta^2)[exp(\theta^2) - 1] \quad (6)$$

Therefore,

$$\varpi = ln(\mu) - \frac{1}{2} ln\left(\frac{\sigma^2}{\mu^2} + 1\right) \quad (7)$$

and

$$\theta = \sqrt{ln\left(\frac{\sigma^2}{\mu^2} + 1\right)} \quad (8)$$

Under the log-normal distribution of σ^2, it can be shown immediately that

$$\sigma^2 \sim exp[2ln(T)\tilde{\alpha} + 2(ln(T)\tilde{\sigma})^2] = T^{2\tilde{\alpha}} T^{2\tilde{\sigma} ln(T)} \quad (9)$$

III. LOSS PROBABILITY ESTIMATION WITH MULTI-SCALING INPUT PROCESSES

Let $X(t)$ represent a multi-scaling traffic process with a Pareto distribution as follows:

$$f_{V(t)}(v) = \frac{\alpha k^\alpha}{v^{\alpha+1}} \text{ for } v > k \quad (10)$$

where $\mu = \frac{\alpha k}{\alpha - 1}$ and $\sigma^2 = \left(\frac{k}{\alpha-1}\right)^2 \left(\frac{\alpha}{\alpha-2}\right)$ are mean and variance, respectively.

The distribution parameters α and k can be determined by the knowledge of the μ and σ^2 of the process X(t). In other words, the mean and variance values can be numerically estimated directly from given input network traffic flows. Therefore,

$$\alpha = \frac{\mu^2}{\sigma^2} \quad (11)$$

and

$$k = \mu - \frac{\sigma^2 \mu}{\mu^2} \quad (12)$$

or

$$\alpha = \frac{\mu^2}{\sigma^2} + 2 \quad (13)$$

and

$$k = \frac{\mu^3 + \sigma^2 \mu}{\mu^2 + 2\sigma^2} \quad (14)$$

In this section we derive an analytical expression for loss probability under a single server queue.

We assume that the single queue is stable with buffer capacity big enough to accommodate any eventual transient bursts. Then, the following balance equation can be established:

$$Q(t_0) + V(t - t_0) = Q(t) + O(t - t_0) \quad (15)$$

where $Q(t)$ is the queue length at time *t*, $V(t - t_0) = W(t) - W(t - t_0)$ is the cumulative traffic load in the period $[t, t_0]$, and $O(t - t_0)$ denotes the traffic load leaving in (t_0, t). Here we assume

$$O(t) = C\big(t - I(t)\big) \quad (16)$$

where C is the constant service rate and $I(t)$ denotes the total server idle time of up to t. Let $V(0) = 0$ and $Q(0) = 0$. Therefore $Q(t)$ can be written as:

$$Q(t) = max(V(t) - O(t), 0) \quad (17)$$

Let $Y(t) = V(t) - Ct$ and $\Delta t = CI(t)$, Equation (17) can be expressed as:

$$Q(t) = max(Y(t) + \Delta t, 0) \quad (18)$$

Applying the law of total probability, the loss probability in queue can be calculated as:

$$\begin{aligned} P_{loss}(t) &= P(Q(t) > q) = P(Y(t) + \Delta(t) > q, Y(t) > q) \\ &\quad + P(Y(t) + \Delta(t) > q, Y(t) \le q) \\ &= P(Y(t) > q) + P\big(Y(t) \le q < Y(t) + \Delta(t)\big) \end{aligned} \quad (19)$$

The first term $P(Y(t) > q)$ in (19) is called the absolute loss probability (P_{abs}) and the second term $P\big(Y(t) \le q \le \Delta(t)\big)$ the opportunistic loss probability $\big(P_{opp}\big)$. Assuming $Q(T)$ stationary, letting $\rho = 1 - \eta = 1 - \lambda/C$ and using the result derived by Benes [1], the second term $\big(P_{opp}\big)$ can be written as:

$$\begin{aligned} P_{opp}(t) &= P\big(Y(t) \le q < Y(t) + \Delta(t)\big) \\ &= \rho \int_0^t f_{V(u)}(v)|_{v=Cu+q} du \end{aligned} \quad (20)$$

Also, the absolute loss probability (P_{abs}) can be written as an integral:

$$\begin{aligned} P_{abs}(t) &= P(Y(t) > q) = P(V(t) > Ct + q) \\ &= \int_{Ct+q}^{\infty} f_{V(t)}(v) dv \end{aligned} \quad (21)$$

Thus, the fully characterized queuing behavior of eventually any traffic type in term of information loss is given by:

$$P_{loss}(t) = \int_{Ct+q}^{\infty} f_{V(t)}(v) dPv + \rho \int_0^t f_{V(u)}(v)|_{v=Cu+q} du \quad (22)$$

The first term on the right side of Eq. (21) can be further detailedly expressed as:

$$P_{abs}(t) = \int_{Ct+q}^{\infty} f_{V(t)}(v) dv = \left(\frac{k}{x}\right)^{\alpha} \text{ for } x \ge k \quad (23)$$

Thus, the loss probability under the stationary state assumption is:

$$P_{steady}(t) = \lim_{t\to\infty} P_{loss}(t) = \rho \sup_{t>0} \left\{ \int_0^t f_{V(u)}(v)|_{v=Cu+q} du \right\} \quad (24)$$

or

$$P_{steady}(t) = \left(1 - \frac{\lambda}{C}\right) \int_0^{\infty} \frac{\alpha k^{\alpha}}{v^{\alpha+1}} |_{v=Cu+q} du \quad (25)$$

Note that for multi-scaling traffic series the variables α and k can be calculated using equations (11) and (12) or (13) and (14), respectively. Substituting the relations given by the equations (11) and (12) into (25), the loss probability can be estimated by:

$$P_{steady}(t) = \left(1 - \frac{\lambda}{C}\right) \int_0^{\infty} \frac{\left(\frac{\mu^2}{\sigma^2}\right)\left(\mu - \frac{\sigma^2 \mu}{\mu^2}\right)^{\frac{\mu^2}{\sigma^2}}}{(Ct+q)^{\frac{\mu^2}{\sigma^2}+1}} dt \quad (26)$$

Again, now substituting the relations given by the equations (13) and (14) into (26), the loss probability can be estimated by:

$$P_{steady}(t) = \left(1 - \frac{\lambda}{C}\right) \int_0^{\infty} \frac{\left(\frac{\mu^2}{\sigma^2}+2\right)\left(\frac{\mu^3+\sigma^2\mu}{\mu^2+\sigma^2}\right)^{\frac{\mu^2}{\sigma^2}+2}}{(Ct+q)^{\frac{\mu^2}{\sigma^2}+3}} dt \quad (27)$$

where $\mu = \lambda T$ and $\sigma^2 = T^{2\tilde{\alpha}} T^{2\tilde{\sigma} ln(T)}$.

A. Our Approach for Loss Probability Estimation

In this work, we propose our approach for loss probability estimation. The major motivation of the proposed approach is to reduce the complexity of the numerical integration to be carried out in expressions (26) and (27). In other words, we propose the exponential approximation given in (28) to describe the relation between the square mean and the variance under time scale T in order to make the analytical expression for loss probability estimation simpler, more efficient without losing the accuracy of the estimates.

$$\frac{\mu^2}{\sigma^2} \cong aexp(bx) \quad (28)$$

where a and b are two parameters of the exponential functional model used for the desired fitting. In general, parameters a and b of the exponential fitting function is determined from applying the minimum mean square error approximation.

For illustration purpose, the green colored curve in Figure 1 is the best exponential function fitting for real network traffic lbl_pkt_5 [7].

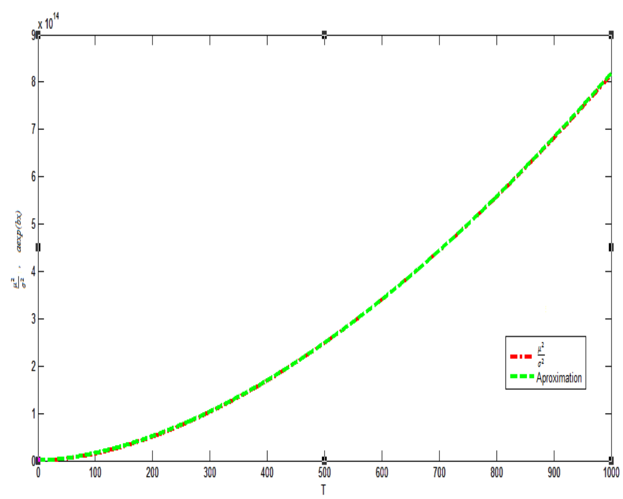

Fig 1: Approximation

Under the exponential fitting function modeling given by (28), we have two expression for the steady state loss probability:

$$P_{steady}(t) = \left(1 - \frac{\lambda}{C}\right) \int_0^{\infty} \frac{(aexp(bx))\left((\lambda x) - (aexp(bx))^{-1}(\lambda x)\right)^{aexp(bx)}}{(Cx+q)^{aexp(bx)+1}} dx \tag{29}$$

or

$$P_{steady}(t) = \left(1 - \frac{\lambda}{C}\right) \int_0^{\infty} \frac{(aexp(bx)+2)\left((\lambda x)\left(\frac{aexp(bx)+1}{aexp(bx)+2}\right)\right)^{aexp(bx)+2}}{(Cx+q)^{aexp(bx)+3}} dx \tag{30}$$

Figure 2 shows how the loss probability changes in terms of buffer size for two loss probability estimation equation given by (29) and (30) for traffic lbl_pkt_5. Clearly, two loss probability curves are very close.

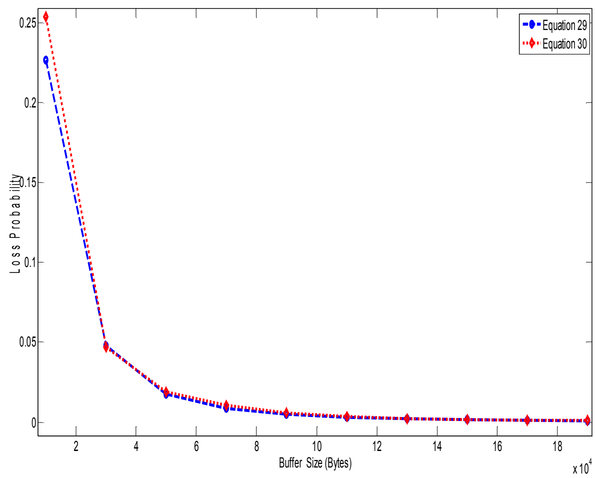

Fig 2: Differences between Equations 29 and 30

IV. Experimental Investigations

In this section we evaluate our approach for loss probability estimation and present our method for traffic admission control and dynamic resource allocation.

A. Loss Probability Estimation:

Table I summarizes the queuing system configuration (server capacity and buffer size) of the single server queue used in the simulation.

Table I: Queuing System Configuration

Traffic Trace	*Server Capacity (Bytes/s)*	*Buffer Size (Bytes)*
lbl_pkt_5	1.4×10^4	3×10^4
dec_pkt_1	12×10^5	3×10^5

Table II compares the loss probability estimates (in number of bytes) for these traffic traces feeding a single server queue scheme defined in Table I, under the following methodologies, namely:

- Simulation: by simulations;
- the Duffield: Duffield's method [4];
- Lognormal: the proposed exponential approach for variance with normal distribution and traffic having lognormal distribution;
- MSQ: Multiscale Queue (MSQ) [13];
- CDTSQ: Critical Dyadic Time-Scale Queue (CDTSQ) [13];
- Proposed: our approach proposed in this paper.

Notice that the Duffield's method provides a lower bound of loss probability $P(Q > b)$ for self-similar processes. "Lognormal", "MSQ" and "CDTSQ" are three multi-scale analyses for network traffic with long-range-dependence [13]. Our proposed approach in this work can be viewed as an alternative and improved version for the Lognormal method proposed in [15].

Table II Loss Probability Estimates

Traffic Trace	*lbl_pkt_5*	*dec_pkt_1*
Simulation	8.14×10^{-4}	1.30×10^{-3}
Duffield	8.02×10^{-18}	5.61×10^{-19}
Lognormal	2.31×10^{-4}	4.39×10^{-5}
MSQ	2.05×10^{-6}	3.13×10^{-6}
CDTSQ	9.86×10^{-7}	1.45×10^{-6}
Proposed	4.92×10^{-4}	3.847×10^{-3}

Figure 3 to 6 compare how loss probability estimates vary in function of buffer size and different serve capacities, respectively, for the lbl_pkt_5 and dec_pkt_1 traces. Again, the proposed approach provides considerably better performances.

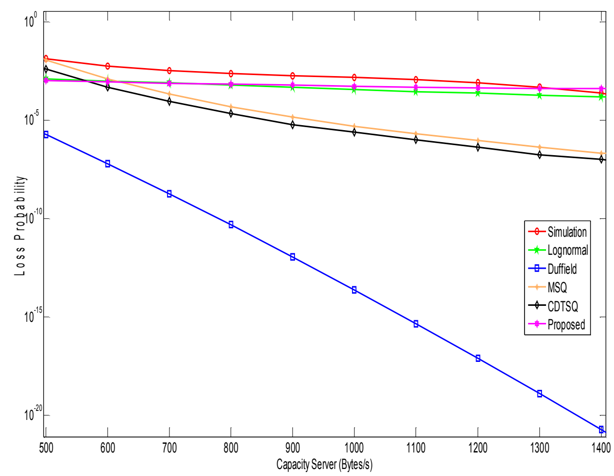

Fig.3. Loss Probability versus Server Capacity for the traffic trace lbl_pkt_5

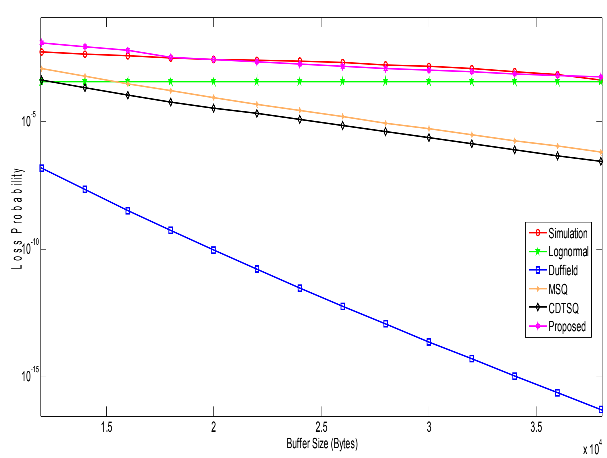

Fig.4. Loss Probability versus Size of Buffer for the traffic trace lbl_pkt_5

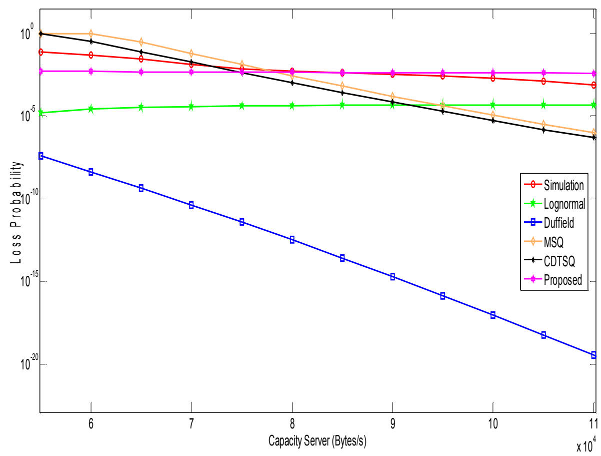

Fig. 5. . Loss Probability versus Server Capacity for the traffic trace dec_pkt_1

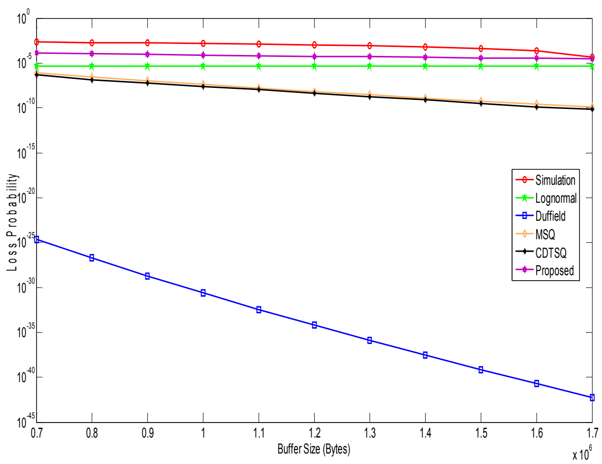

Fig. 6 Loss Probability versus Size of Buffer for the traffic trace dec_pkt_1

V. CONCLUSION

In this paper, we present an analytical expression for estimating the byte loss probability at a single server queue with multifractal traffic arrivals. Initially, we address the theory concerning multifractal processes, especially the Hölder exponents of the multifractal traffic traces. Next, we focus our attention on the second order statistics for multifractal traffic processes. More specifically, we assume that an exponential model is adequate for representing the variance of the traffic process under different time scale aggregation. Then, we compare the performance of the proposed approach with some other relevant approaches (e.g., monofractal based methods, MSQ (multiscale queue) and CDTSQ (Critical dyadic time-scale queue)) using real traffic traces. The simulation results shows that the proposed estimation of loss probability is simple, accurate.

REFERENCES

[1] Benes V., "General Stochastic Processes in theory of Queues", Reading, MA: Addison Wesley,1963.

[2] Beran, J. et al.(1995) Long range dependence in VBR video traffic. IEEE Transactions on Communications, v.43, p.1566-1579.

[3] Crovella, M. E.; Bestavros,(1996) A. Self-similarity in world wide web traffic - evidence and possible causes. Proceedings of ACM Sigmetrics, p.160-169.

[4] Duffield,N.G. & O'Connell,N. "Large Deviations and Overflow Probabilities for the Generalsingles-Server queue, with Applications",Dublin Institute dor Advanced Studies-applied Probability.

[5] Feldmann, A.; Gilbert, A.; Willinger, W. and Kurtz, T. G (1998). The changing nature of network traffic: Scaling phenomena. ACM Computer Communication Review, v.28, p.5-29. Group, Tech.1 Rep.Disas-STP-93-30. 1993.

[6] Garret M. W.; Willinger, W.(1994) Analysis, modeling and generation of self-similar VBR vide traffic. Proceedings of ACM Sigcomm, p.269-280.

[7] http://ita.ee.lbl.gov/html/traces.html

[8] Leland W., Taqqu M., Willinger W., & Wilson D., "On the self-similar nature of Ethernet traffic (extended version)", IEEE/ACM Transactions on Networking, v.2,n.1, pp 1-15, Feb 1994.

[9] Liu N. X. & Baras J. S., "Statistical Modeling and Performance Analysis of Multi-Scale Traffic", Proceedings of Twenty – Second Annual Joint Conference of the IEEE Computer and Communications Societies, INFOCOM 2003, pp.1837-1847, San Francisco, CA, USA, March/April 2003.

[10] Norros,I(1994). A storage model with self-similar input, queueing Systems,v.16,pp.38

[11] Park, K.; Willinger, W. (2000) (Eds.) Self-similar network traffic and performance evaluation. New York: Wiley.

[12] Paxson, V.; Floyd S.,(1995) Wide area traffic: The failure of poisson modeling. IEEE/ACM Transactions on Networking, v.3, n.3, p.226-244.

[13] Ribeiro V. J., Riedi R. H., Crouse M. S., & Baraniuk R. G. "Multiscale Queueing Analysis of Long-Range-Dependent Network Traffic", IEEE INFOCOM 2000, pp. 1026-1035.,Tel Aviv, Israel.

[14] Riedi R. H. & Véhel J. L., "Multifractal Properties of TCP Traffic: A Numerical Study", Technical Report 3129, INRIA Rocquencourt, March 1997.

[15] Stenico, J.W.G, Lee.L.L. A Multifractal Based Dynamic Bandwidth Allocation Approach for Network Traffic Flows, IEEE ICC2010 – International Conference on Communications 23-27 may 2010 – Cape Town, South Africa.

[16] Taqqu M. S., Teverovsky V., & Willinger W., "Is Network Traffic Self-Similar or Multifractal?", Fractals, vol. 5, pp. 63-74, 1997.