Empirical Performance of Weibull Self-Similar Tele-traffic Model

J. Popoola, R. A. Ipinyomi

***Abstract*— The stringent memoryless assumption in general traffic modeling is often violated due to concurrent or batch arrivals. Batch arrivals often leads to self-similarity or long range dependency problems in Internet traffic. Modeling tele-traffic data in this condition requires the use of heavy tailed distribution. In this paper, we propose a general form of Weibull tele-traffic model for self-similar Internet traffic data. The model empirical performance was observed via Monte-Carlo simulation with the aid of discrete event simulation. Performance analysis for the proposed model alongside standard Poisson/ Exponential model was also achieved. The results from the analysis established the strength of Weibull model over the existing model.**

***Index Terms*— long range dependence , Self-similarity, heavy tailed distribution, burstiness.**

I. INTRODUCTION

Self-similarity in internet traffic occurs when packets of the same burst length arrive at the same time or when packets burst at the same inter-arrival period on the server. Self-similarity is also called infinite variance syndrome. Simply speaking, a process shows self-similarity implies the process is indistinguishable from its scaled versions obtained by averaging the original process within different observation time scales. Missing this property makes some other popular traffic model such as Poisson trace give over-optimized evaluation of network performances. A mathematical description of self-similarity can be concluded as follows; Assume an increment process $X_i\,(i = 1,2,\dots)$ and another process $X_j^m\,(i = 1,2,\dots)$ which is obtained by averaging the values in non-overlapped blocks of size m in X_i, i.e

$$X_j^m = \frac{1}{m}(X_{jm-m+1}, +X_{jm-m+2}+\cdot\ \cdot\ \cdot+X_{jm}) \qquad (1)$$

The process X_i is said to be self-similar if

$$X_j^m \sim m^{H-1}X_i$$

Which implies

$$var(X_j^m) = m^{2H-2}var(X_i) \qquad (2)$$

$m\ (m \geq 1)$ is the scale parameter whereas $H\ (0.5 < H \leq 1)$ is the hurst parameter. For instance when $H = 1$ process X_j^m and process X_i have the same distribution without any decay. The hurst parameter is used to measure the burstiness of a process.

J. Popoola, Department of Statistics, University of Ilorin, P.M.B. 1515, Ilorin, Nigeria.
R. A. Ipinyomi, Department of Statistics, University of Ilorin, P.M.B. 1515, Ilorin, Nigeria

The self-similar process can also contain a property of long-range dependence (Sheluhin, 2007). Long range dependence describes the memory effect, where a current value strongly depends upon the past values, of a stochastic process, and it is characterized by its autocorrelation function. For $0 < H < 1, H \neq 1/2$ the autocorrelation function $r(k)$ for lag k is;

$$r(k) = H(2H-1)k^{-2H-2} \qquad (3)$$

The autocorrelation function decays hyperbolically, as k increases, which means that the autocorrelation function is not summable (Fras et al, 2012). This is opposite to the property of short-range dependence (SRD), where the autocorrelation function decays exponentially and the equation (3) has a finite value. Short and long-range dependence have a common relationship with the value of the Hurst parameter of the self-similar process (Park and Willinger 2000) and (Yolmaz, 2002):

i. $0 < H < 0.5$ →SRD - Short Range Dependence
ii. $0.5 < H < 1$ →LRD - Long Range Dependence

Network performance degrades gradually with increasing self-similarity. The more self-similar the traffic, the longer the queue size. The queue length distribution of self-similar traffic decays more slowly than with poisson sources. However, long-range dependence implies nothing about its short-term correlations which affect performance in small buffers. Additionally, aggregating streams of self-similar traffic typically intensifies the self-similarity (“burstiness”) rather than smoothing it, compounding the problem.

Each Internet communication consists of a transfer of information from one computer to another; examples are the downloading of a Web page and the sending of an email message. When a file is transferred, it is not sent across the Internet as a continuous block of bits. Rather the file is broken up into pieces called packets, and each packet is sent individually. Many different protocols collectively carry out the transfer. A protocol is simply a set of rules for communication between computers. Jones (2004) worked on simulation of self-similar processes for modelling self-similar network. He focused more on the development of algorithm that can be used for generating and fitting self-similar process. He also showed that the severity of self-similar process depends on the offspring distribution. Fras et al (2012) developed a simulation tool in the area of measurements, modeling and simulations of the self-similar network traffic. They described a number of facts about self-similarity, long range dependencies and probability, which are used to describe such stochastic processes. They also described the mechanism and models to simulate network traffic in the OPNET Modeler simulation tool. The main goal of their research was to simulate measured network traffic, where they tend to minimize discrepancies between the measured and the simulated network traffic in the sense of packet-rate, bit-rate, bursts intensity, and variances. Alakiri et al (2014) studied the desirability of pareto distribution for modelling modern internet traffic characteristics. They buttressed the inadequacy of poisson /exponential process for modelling modern LAN and WAN network processes. They also talked about burtiness (self-similarity) and correlated (long range dependency) nature of LAN and WAN network processes which violate the standard poisson exponential process. In addition, they suggested future research on comparison of various internet traffic generators (distributions) and use of statistical methods to proof and validate the results of the experiments, which is being examined in this study. This study will further extend the distribution that best describes the self –similar property to the queueing system so as to have better results of network efficiency performance. In this paper, we consider the general distribution of Weibull for modelling self-similar traffic with inherent long range dependency problem.

II. MATERIALS AND METHODS

A. Simulation scheme

The first step is the generation of the sequence of inter-arrival times and transmission times in accordance with the given distributions. The subsequent steps listed below generate queue size N, service start S-st, service end S-end and delay time W; The “queue-size on arrival” N^i is obtained by comparing the arrival time of the current arrivals and the values in the “service ends” column of the previous rows. In particular, the queue size on arrival is equal to the number of customers that arrive before the current customer (previous rows) that their “service ends” time values are greater than the arrival time value of the current arrival. The “service starts” value is the maximum of the “arrival time” value of the current arrival and the “service end” value of the previous arrival. Also notice that if the queue size on arrival of the current arrival is equal to zero, the service start value is equal to the “arrival time” value of the current arrival and if the queue size on arrival of the current arrival is greater than zero the service start value is equal to the “service end” value of the previous arrival. The “service ends” value is simply the sum of the “service starts” and the “service duration” values of the current arrival. The “delay” value is the difference between the “service ends” and the “arrival time” values.

Table I: Simulation parameter

Variables	Parameter
Inter-arrival time	0.01 m/s
Transmission time	0.008 m/s
Bandwidth	150mb
Hurst Index (H)	0.9

B. Weibull Distribution

The Weibull PDF is determined by the parameters k and θ . For $k = 1$ the Weibull distribution is identical to the exponential distribution (fras et al, 2012):

$$f(x;k,\theta) = k\theta(\theta x)^{k-1}e^{-(\theta x)^k};\quad k > 0\,;x > 0 \qquad (4)$$

And the cumulative distribution function (CDF) is:

$$F(x;k,\theta) = 1 - e^{-(\theta x)^k} \qquad (5)$$

The parameters k and θ is the shape and the intensity parameter.

C. Traffic performance of G/G/1/K

Given that $E(T)$ and σ_T^2 are the mean and variance of inter-arrival time into the system, Also if $E(X)$ and σ_X^2 are the mean and variance of packet size transmitted per unit of time. Thus,

$$\rho = \lambda\bar{X}$$

For a stable process when $\rho < 1$,

According to sanjay (2002), there exist a bound for the waiting time on the queue W_q,

$$\frac{\lambda\sigma_X^2 - \bar{X}(2-\rho)}{2(1-\rho)} \le W_q \le \frac{\lambda(\sigma_X^2 + \sigma_T^2)}{2(1-\rho)}$$

As $\rho \to 1$,

$$W_q = \frac{\lambda(\sigma_X^2 + \sigma_T^2)}{2(1-\rho)}$$

By little’s formula, the average waiting time in the system is;

$$W = W_q + \bar{X}$$

Also the average size in the system is;

$$N = \lambda W$$

The expected queue size is;

$$N_q = \lambda W_q$$

III. RESULTS

Table II: Performance measures for the true model and Weibull model.

Traffic Model	Performance Measures	True Model	$We/We/1/K$
Analytical result	N	2.0080	2.5168
	N_q	1.6371	2.0637
	W	0.0200ms	0.0251ms
	W_q	0.0164ms	0.0206ms
Simulation Result	N	1.8952	2.8517
	N_q	1.2500	1.4671
	W	0.0162ms	0.0392ms
	W_q	0.0125ms	0.0347ms

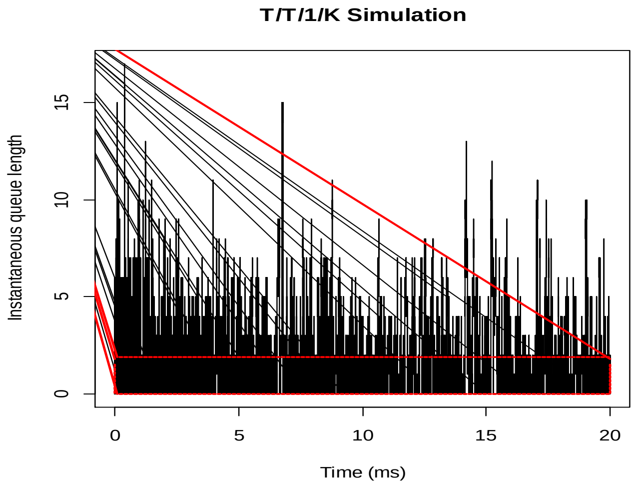

Fig 1: Traffic behavior of the true self-similar traffic over time

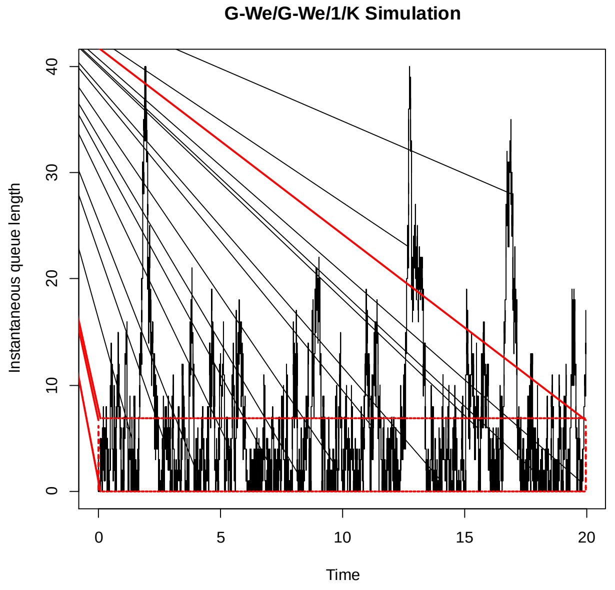

Fig 2: Traffic behavior of Weibull self-similar traffic over time

Table 2 presents the results of the simulation for the tele-traffic model as specified in table 1. The analytical result implies the result using the standard equation for G/G/1/K while the empirical is the simulation result. The two results affirm that the Weibull general model is adequate for modelling the combined problems of self-similarity and long-range dependency. The results also established possibility of queue in the presence self-similarity and long range dependency.

IV. CONCLUSION AND RECOMMENDATION

In this paper, simulation of network traffic with inherent self-similar and long range dependency was presented. In adequacy of standard poisson/exponential process and the corresponding traffic model was also presented. The supremacy of Weibull distribution was also observed for both arrival and transmission processes. Thus it is strongly recommended that to use Weibull distribution when monitoring traffics in internet networks.

REFERENCES

[1] Alakiri, H. O., Oladeji, F. O., Benjamin, B.C., Okolie, C. C., Okikiola, F. M., (2014); The desirability of pareto distribution for modelling modern internet characteristics; International Journal of novel research and applied sciences (IJNREAS) 1 (1).

[2] Akaike, H. (1973): “ Information Theory and an Extension of the Maximum Likelihood Principle.”, In B. petov and F.csake, eds., second international symposium on Information Theory. Budapest: Akademiai Kiado.

[3] Erlang, A.K., Brockmeyer, E., H.L. Halstrom, H.L., & Arns, J.(1948). The Copenhagen Telephone Company. Proceedings of the C.C.I.F. ("Le comité consultatif international des communications téléphoniques à grande distance"), Montreux.

[4] Yõlmaz, H. (2002); “IP over DVB: Managment of self-similarity, Master of Science, Boğaziçi University.

[5] Park, K. and Willinger, W. (2000), “Self-Similar Network Traffic and Performance Evaluation, John Wiley & Sons,.

[6] Kostia, A. and Uri, Y. (2008); Retrial networks with finite buffers and application to internet data traffic; Journal of probability in informational sciences; vol 22 (4).

[7] Sheluhin, O. Smolskiy, S. and Osin, A. (2007), “Self-Similar Processes in Telecommunications, John Wiley & Sons.

[8] Schwarz, G E. (1978): "Estimating the dimension of a model". Annals of Statistics 6 (2): 461–464.

[9] Tran-Gia, P. Staehle, D. and Leibnitz, K. (2001). Source Traffic Modeling of Wireless Applications. International Journal of Electronics and Communications (AEÜ). Vol. 55, No. 1.