

# Network Traffic Model: A Case of BIUST Network

<sup>1</sup>Thato Solomon, <sup>2</sup>Adamu Murtala Zungeru,  
<sup>1</sup>Rajalakshmi Selvaraj, <sup>2</sup>Mmoloki Mangwala and <sup>2</sup>Joseph Chuma

<sup>1</sup>Department of Computer Science and Information Systems,

<sup>2</sup>Department of Electrical, Computer and Telecommunication Engineering,

Botswana International University of Science and Technology, Private Bag 16, Palapye, Botswana

## Article history

Received: 26-12-2016

Revised: 23-04-2017

Accepted: 21-06-2017

Corresponding Author:  
Adamu Murtala Zungeru  
Department of Electrical,  
Computer and  
Telecommunication  
Engineering, Botswana  
International University of  
Science and Technology,  
Palapye, Botswana  
Email: zungerum@biust.ac.bw

**Abstract:** Network traffic model and analysis provides the average load, bandwidth requirements and the different application of the available bandwidth for a particular network, in addition to several other details of the network. This paper presents mathematical model used for modelling real world problems using Botswana International University of Science and technology (BIUST) network traffic as a case. Sophisticated analysis of data is done to model the BIUST network with the succor of statistics, as it implies the collection and interpretation of data through mathematical processes called stochastic processes. From the attained results, the model and estimation of packet traffic distribution for BIUST Network based on Pareto distribution, it was perceived that about 20% of the users had about 80% of the bandwidth consumed.

**Keywords:** Mathematical Modelling, Traffic Model, Mathematical Processes, Stochastic Processes, Packet Traffic Distribution, Pareto Distribution

## Introduction

Network traffic modelling is used as the basic for the network applications and for capacity planning of network systems. Given the impact of poor choices in this arena, it is clear that the validity of the underlying models is of critical importance (Wilson, 2016). They are a wide number of mathematical models that could be used to model network traffic depending on the type of network to be modelled. The factors used to evaluate a network are taken directly from the underlying traffic model.

Statistics is concerned with making inferences about the way the world is, based upon things we observe happening. Nature is complex, so the things we see hardly ever conform exactly to simple or elegant mathematical idealizations as the world is full of unpredictability, uncertainty, randomness. Probability is defined as the language of uncertainty and so to understand statistics, we must understand uncertainty, as probability and statistics work hand in hand (Tamhane and Dunlop, 2000). Sophisticated analysis of data will be done to model the BIUST network with the aid of statistics, as it involves the collection and interpretation of data through mathematical processes called stochastic processes.

A stochastic process is simply a probability process that is, any process in nature whose evolution we can analyze successfully in terms of probability. On the empirical side, a discussion of the nature of probability would take us too far afield (and might sidetrack us into philosophy) and on the mathematical side the definitions would require too much high-powered mathematics (Tamhane and Dunlop, 2000).

According to (Tamhane and Dunlop, 2000), random variable associates a unique numerical value with each outcome in the sample space. Usually a random variable it is explained as a real valued function defined over a sample space. Random variable is denoted by a capital letter (e.g.,  $Y$  or  $X$ ) and a particular value taken by a random variable is denoted by corresponding lower case letters ( $y$  or  $x$ ). Random variable can either be discrete or continuous.

The remainder of this paper is organized as follows. Section 2 gives the related work to this paper. Section 3 gives a brief explanation on mathematical modelling. It will also include the modelling process as well as the classification of mathematical models. In section 4, we provided data sets that will be used to model BIUST network and a packet distribution model for BIUST network will also be presented. Finally, the paper is concluded in section 5.

## Related Work

Traffic models reflect our best knowledge of traffic behavior. Latest studies of real telecommunications network traffic data have revealed that teletraffic exposition self-similar (or fractal) properties over a wide range of time scales (Boxma and Cohen, 2000; Radev and Lokshina, 2009). The properties of self-similar telecommunications network traffic are very distinct from properties of traditional models based on Poisson, Markov-modulated Poisson and related processes (Giambene, 2005). Usage of traditional models in networks characterized by self-similar processes can result in to biased conclusions about the performance of analyzed networks (Radev and Lokshina, 2009; Jeong, 2002). Traditional models can result in over-estimation of the network performance (Faraj, 2000), lack of allocation of communication and data processing resources and hence problems in ensuring the QoS. Then, full understanding is that the self-similar nature in teletraffic is a vital issue.

Self-similar teletraffic is seen in LAN and WAN, where superposition of strictly independent alternating ON/OFF traffic models whose ON- or OFF-periods have heavy-tailed distributions with infinite variance can be used to model aggregate network traffic that shows self-similar (or long-range dependent) behavior typical for measured Ethernet LAN traffic over a wide range of time scales (Kushner, 2001).

In ATM network traffic self-similar traffic arriving at an ATM buffer results in a heavy-tailed buffer occupancy distribution buffer cell loss probability reduces with the buffer size, not exponentially as in traditional Markovian models, but hyperbolically.

One more implementation of traffic self-similarity is in Internet traffic, where many characteristics of WWW can be modeled using heavy-tailed distributions, including the distribution of traffic times, the distribution of user requests for documents and the distribution of WWW document sizes (Jeong, 2002).

In TCP/IP network traffic the transfer of files or messages shows that the reliable transmission and flow control mechanisms of TCP serves to maintain long range dependent structure which include heavy-tailed file size distributions (Bobbio *et al.*, 2013). The relationship between self-similar traffic and network performance is defined, as captured by performance measures such as packet loss rate, retransmission rate and queueing delay, where increased self-similarity results in degradation of performance and queueing delay exhibits a dramatic increase as self-similarity increases.

The self-similarity observed in video traffic provides possibility for developing models for Variable Bit Rate (VBR) video traffic using heavy-tailed distributions (Radev, 2005). The autocorrelation of the VBR video sequence decay hyperbolically and can be model educing Fractional Autoregressive Integrated Moving-Average

(F-ARIMA) and Fractional Gaussian Noise (FGN) self-similar processes (Radev and Lokshina, 2009; Ravid and Lokshina, 2007).

The impact of self-similar models on queuing performance is important and the main trends in such findings are connected with (a) permission traffic modeling for high speed networks, (b) efficient simulation of actual network traffic and (c) analyzing queuing models and protocols under realistic traffic scenarios (Mehdi, 2003). The traditional models of teletraffic that assume independent arrivals, based on Poisson processes, Markov-modulated Poisson processes and other related processes are not able to capture the self-similar nature of teletraffic (Hayes and Ganesh, 2004).

In another work presented by (Radev and Lokshina, 2009), the time series of self-similar processes exhibit burstiness over a wide range of timescales. Self-similarity can statistically detail wireless IP network traffic that is bursty on numerous time scales. Modeling and simulation of self-similar telecommunications network traffic can be done with the generators of synthetic self-similar sequences, which are separated into two practical classes: The sequential generators and the fixed-length sequence generators. The fixed-length sequence generators for simulation of self-similar wireless IP network traffic are contemplated in this study (Radev and Lokshina, 2009).

## Mathematical Modelling

We propose mathematical modeling of BIUST network traffic using an adopted model called Pareto Distribution model which is used for modelling real world problems. The basis of the distribution is that a high proportion of a population has low income while only a few people have very high incomes (Erlina, 2011).

### Overview of Mathematical Modelling

Mathematical modelling is by and large comprehended as the way toward applying mathematics to a real world problem with a perspective of comprehension the last mentioned (Osterbo, 2003; Frost and Melamed, 1994). One can contend that mathematical modelling is the same as applying mathematics where we additionally begin with a real world problem; we apply the essential mathematics, yet after having found the result we no more consider the underlying issue aside from maybe to check if our answer makes sense. This is not the situation with mathematical modelling where the utilization of mathematics is more to understand this real world problem. The modeling process might possibly result to solving the problem altogether however it will reveal insight to the circumstance being studied (Osterbo, 2003). Figure 1 shows the key phases in modelling process. The adopted model that is used to model the network is based in BIUST Network.

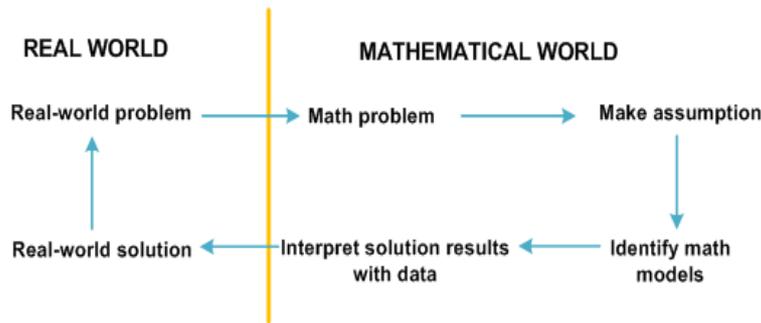


Fig. 1. Modelling process (Osterbo, 2003)

### Categories of Mathematical Models

#### Continuous-Time Source Models

Continuous time is mostly interested in stochastic process to represent the time-varying source rate  $X(t)$  or the set of packet arrival times  $\{t_1, t_2, t_3, t_4, \dots\}$ . This model includes Uniform distribution, Gamma distribution, Exponential distribution, Beta distribution and Pareto distribution (Wilson, 2016; Tamhane and Dunlop, 2000; Frost and Melamed, 1994; Yang and Petropalu, 2001; Chandrasekaran, 1994).

#### Uniform Distribution Model

A uniform distribution arises in situations where all values are “equally likely” over an interval. Specifically, the Probability Density Function (PDF) of a uniform distribution is constant over an interval. A random variable  $X$  has a uniform distribution over the interval  $[a, b]$  (denoted by  $X \sim U[a, b]$ ). The PDF is given by:

$$f(x) = \begin{cases} \frac{1}{b-a}, & a \leq x \leq b \\ 0, & \text{otherwise.} \end{cases} \quad (1)$$

The Cumulative Distribution Function (CDF) of a uniform random variable is given by:

$$F(x) = \begin{cases} 0, & x < a \\ \frac{x-a}{b-a}, & a \leq x \leq b \\ 1, & x > b \end{cases} \quad (2)$$

#### Gamma Distribution Model

A random variable is said to be gamma distribution with parameters  $\gamma > 0$  and  $r$  if its PDF is given by:

$$f(x) = \frac{\gamma^r x^{r-1} e^{-\gamma x}}{\tau(r)} \text{ for } x \geq 0 \quad (3)$$

where the gamma function,  $\tau(r)$ , is defined by:

$$\tau(r) = \int_0^{\infty} u^{r-1} e^{-u} du \text{ for } r > 0 \quad (4)$$

For positive integer values of  $r$ , it can be shown that  $\tau(r) = (r-1)!$  this special case of gamma distribution, known as the *Erlang distribution*, is used in queuing theory to model waiting times. A shorthand notation  $X \sim \text{Gamma}(\gamma, r)$  denotes that  $X$  has a gamma distribution with parameters  $\gamma$  and  $r$ .

#### Exponential Distribution Model

The exponential distribution is a continuous analog of the geometric distribution; as such it is an example of continuous waiting time distribution. The PDF of an exponential random variable  $X$  with parameter  $\gamma > 0$  is:

$$f(x) = \gamma e^{-\gamma x} \text{ for } x \geq 0 \quad (5)$$

The CDF of an exponential random variable is given by:

$$F(x) = \int_0^x \gamma e^{-\gamma y} dy = 1 - e^{-\gamma x} \text{ for } x \geq 0 \quad (6)$$

#### Beta Distribution Model

The beta distribution provides a flexible way to model many types of measurements that have finite ranges. A random variable has a beta distribution on the interval  $[0, 1]$  with parameters  $a$  and  $b$  (denoted by  $X \sim \text{Beta}(a, b)$ ) if its PDF is given by:

$$f(x) = \frac{1}{B(a, b)} x^{a-1} (1-x)^{b-1} \text{ for } 0 \leq x \leq 1 \quad (7)$$

where,  $B(a, b)$  is the beta function defined by:

$$B(a, b) = \frac{\tau(a)\tau(b)}{\tau(a+b)} \quad (8)$$

Note that a random variable having a finite range other than  $[0, 1]$  can always be transformed to the  $[0, 1]$  range.

The  $U[0,1]$  distribution is a special case of the beta distribution when  $a = b = 1$ , in which case above PDF reduces to  $f(x) = \frac{1}{B(1,1)} = 1$ .

### Pareto Distribution Model

Pareto distribution is a skewed, heavy-tailed distribution that is sometimes used to model the distribution of incomes. The law was developed by Vilfredo Pareto in 1897 and he firstly included it in one of his works in which he attempted to prove that the distribution of incomes and wealth in society is not random that a consistent pattern appears throughout history, in all parts of the world and in all societies (Alzaatreh and Famoye, 2012). Pareto distribution was one of the most famous but much-criticized law of income distribution.

The PDF of a Pareto distribution is given by:

$$f(x) = \frac{ab^a}{x^{a+1}} \quad (9)$$

For  $b = 1$ :

$$f(x) = \frac{a}{x^{a+1}} \quad (10)$$

The CDF of a Pareto random variable is given by:

$$f(x) = 1 - \left(\frac{1}{x}\right)^a \text{ for } b = 1 \quad (11)$$

where, “ $a$ ” is a shape parameter and “ $b$ ” represents location parameter of the network.

### Discrete-Time Source Models

This model includes Bernoulli distribution, Binomial distribution, Poisson distribution, hypergeometric distribution, multinomial distribution and geometric distribution (Wilson, 2016; Tamhane and Dunlop, 2000; Frost and Melamed, 1994; Yang and Petropalu, 2001; Chandrasekaran, 1994).

### Bernoulli Distribution Model

Bernoulli random variable is referred to as a random variable that can take only two values, say 0 and 1. The Bernoulli distribution is a useful model for dichotomous outcomes. Some examples are the sex of a baby (male or female), the outcome of an experiment (success or failure) and the toss of a coin (head or tail). An experiment with a dichotomous outcome is called a Bernoulli trial.

Suppose that an item drawn at random from a production process can be either defective or non-

defective. Let  $p$  denote the fraction of the defective items produced by the process. Then the probabilities of the possible outcomes for an item randomly drawn from this process are  $P(\text{Defective}) = p$  and  $P(\text{Non-defective}) = 1 - p$ . A Bernoulli random variable can be defined as  $X = 1$  if the item is defective and 0 if item is non-defective with the following distribution:

$$f(x) = P(X = x) = \begin{cases} p & \text{if } x = 1 \\ 1 - p & \text{if } x = 0 \end{cases} \quad (12)$$

### Binomial Distribution Model

In statistics and probability theory, Binomial distribution is a common distribution role of discrete processes in which a stationary probability is dominant for each independently generated value (Chandrasekaran, 1994; Alzaatreh and Famoye, 2012; Ascombe, 1949; Bevrani and Shariffar, 2014; Smith, 2015). Simply a binomial distribution can be explained as the sum of autonomous and identically distributed Bernoulli random variables.

Binomial distribution is now generally used in analysis of data in almost every field of human interrogation since it was elucidated in connection with games of chance. According to (Yang and Petropalu, 2001) Binomial distribution is normally used to model the number of resource accessible, number of packets that reach the destination without misfortune and the number of bits in error in a packet. It applies to any dominant number ( $n$ ) of repetitions of an independent process that yields a certain outcome with the identical probability ( $p$ ) on each repetition. For instance, it provides a formula for the probability of acquiring 10 sixes in 50 rolls of a die. In a proof distributed after the death of Swiss mathematician Jakob Bernoulli in 1713, it was discovered that the probability of  $k$  such outcomes in  $n$  repetitions is equivalent to the  $k$ th term (where  $k$  begins with 0) in the expansion of the binomial expression  $(p + q)^n$ , where  $q = 1 - p$  thus the name binomial distribution. In the example of throwing a die, the probability of turning up any number on every roll is 1 out of 6. The probability of turning up 10 sixes in 50 moves, then, is equal to the 10th term beginning with the 0th term in the expansion of  $\left(\frac{5}{6} + \frac{1}{6}\right)^{50}$ , or 0.115586.

The British statistician Ronald Fisher used the binomial distribution to disclose evidence of possible scientific deception in 1936, in the renowned experiments on pea genetics which was reported by Gregor Mendel the recounted Austrian botanist in 1866. With the use of binomial distribution, Fisher observed that one of Mandel's laws of inheritance would dictate that the number of yellow peas in one of Mandel's experiments would have a binomial distribution with  $n =$

8,023 and  $p = \frac{3}{4}$ , for a normal of  $np \cong 6,017$  yellow peas. Fisher discovered exceptional understanding between this number and Mendel's information, which demonstrated 6,022 yellow peas out of 8,023. One would expect the figure to be close, but a figure that close should occur only once in 10 times. In addition, Fisher discovered that all seven results in Mendel's pea experiments were extremely close to the anticipated values-even in an instance where Mendel's calculations contained one defect. The analysis of Fisher sparked a lengthy strife that remains unresolved until today (Ascombe, 1949; Bevrani and Sharififar, 2014).

The general form of the Probability Mass Function (PMF) of a binomial random variable  $X$  with parameters  $n$  and  $p$  (denoted by  $X \sim \text{Bin}(n, p)$ ) is derived as follows: The probability of obtained  $x$  successes and  $n-x$  failures in a particular way (e.g., the first  $x$ -trials resulting in successes and the last  $n-x$  trials resulting in failures) is  $\binom{n}{x} p^x (1-p)^{n-x}$ , because the trial are independent. There are a total of  $\binom{n}{x}$  ways of distributing  $x$  successes and  $n-x$  failures among  $n$  trials. Therefore PMF is given by:

$$f(x) = P(X = x) = \binom{n}{x} p^x (1-p)^{n-x} \text{ for } x = 0, 1, \dots, n \quad (13)$$

#### Poisson Distribution Model

The Poisson distribution is a standout amongst the most vital and generally utilized statistical distributions due to its memory less capabilities. The Poisson distribution is a limited form of the binomial distribution. It is regularly used to portray the pattern of random point-like events in 1-, 2- and 3-dimensions or, all the more commonly, to give the model to uncertainty against which an observed event pattern in time or space might be matched. On the off chance that events happen arbitrarily and autonomously, at a continuous rate (in time) or with a continuous density (in space), then the count of these events per unit time or per unit region will fit in with a Poisson distribution and the example of event is depicted as a Poisson process. The distribution of the length of intervals between events (or waiting times) in a one-Dimensional (1D) Poisson procedure is an Exponential distribution (Smith, 2015; Zou, 2016).

Siméon-Denis Poisson firstly applied Poisson distribution in 1830 to describe the number of times a gambler would win a rarely won game of chance in a large number of tries (Erlina, 2011). Poisson distribution is mostly used to model call arrivals, number of tasks in the system, number of requests to a server, number of failed components and message length (Yang and Petropalu, 2001; Chandrasekaran, 1994). Poisson as it was shown in (13) that the binomial PMF is given by

$P(X = x) = \binom{n}{x} p^x (1-p)^{n-x}$ . When  $n \rightarrow \infty$  and  $p \rightarrow 0$  in such a way that  $np$  approaches a positive constant  $\gamma$  is the limiting binomial PMF can be shown to be:

$$f(x) = P(X = x) = \frac{e^{-\gamma} \gamma^x}{x!}, x = 0, 1, 2, \dots \quad (14)$$

which is the PMF.

#### Hypergeometric Distribution Model

The hypergeometric distribution is, fundamentally, an extraordinary type of the Binomial. Hypergeometric distribution is applied when testing is performed from a predictable population without substitution hence making trials dependent on each other. While the Binomial expect that there are  $n$  independent trials of an experiment, with a fixed probability,  $p$ , which is the same for each event, the hypergeometric deals with the circumstance in which the populace size,  $N$ , from which events are tested, is generally small ( $<100$ ) and inspecting happens without substitution thus the probabilities are not generally the same. A straightforward similarity is selecting balls from a pack or urn containing a blend of red and black. The Binomial applies if, after every random selection the ball is supplanted, whilst the hypergeometric deals with the situation where the balls are not supplanted, so the following ball selection is drawn from a slightly different overall blend of red and dark. Clearly if there are countless in the urn and we are not selecting too much, there is viably no contrast between the two distributions.

Two key assumptions underlying the binomial distribution are that:

- the Bernoulli trials are independent
- Each Bernoulli trial has the same probability of success

These assumptions are valid when a random sample is drawn from an infinite or a very large population<sup>2</sup> of items of which a fraction has a specific attribute. When the population is finite, the assumptions are valid if each randomly sampled item is returned to the population before the next draw. This is called sampling with replacement. But in practice we generally use sampling without replacement. When the population is finite, sampling without replacement creates dependence among the successive Bernoulli trials and the probability of success changes as successive items are drawn. We now derive hypergeometric distribution.

Let  $N$  be the size of the population in which  $M$  items have a specific attribute. We randomly sample  $n$  items from the population without replacement. First we find

the number of ways to draw  $x$  items from the  $M$  with the attribute and  $(n-x)$  items from the remaining  $(N-M)$  items without the attribute. This is given by  $\binom{M}{x} \binom{N-M}{n-x}$ . We then divide it by the number of ways to sample  $n$  items from the population of size  $N$  without any restriction, which is  $\binom{N}{n}$ . Thus:

$$f(x) = P(X = x) = \frac{\binom{M}{x} \binom{N-M}{n-x}}{\binom{N}{n}} \quad (15)$$

This is referred to as the hypergeometric distribution with parameter  $N$ ,  $M$  and  $n$ . If the sampling fraction  $n/N$  is small ( $\leq 0.10$ ), then this distribution is well approximated by the binomial distribution (13) with parameters  $n$  and  $p = M/N$ .

#### Multinomial Distribution Model

The binomial distribution applies when we have a fixed number of independent Bernoulli trials with constant probabilities  $p$  and  $1-p$  for the two outcomes. It is very useful in large number of applications in ecology. In some situations there are more than two possible outcomes in each trial; e.g., a respondent ethnic may be classified as Batawana, Bangwato, Babirwa, Batswapong, Bakgatla, Bakgalagadi, or other. For such trials we need a generalization of the binomial distribution to model the frequencies of different outcomes. Consider a fixed number  $n$  of trials where each trial can result in one of  $k \geq 2$  outcomes and the probabilities of the outcomes,  $p_1, p_2, \dots, p_k$ , are the same from trial to trial, with  $p_1 + p_2 + \dots + p_k = 1$ . Let  $X_1 + X_2 + \dots + X_k = n$ . The joint multivariate distribution of  $X_1 + X_2 + \dots + X_k$  is called the multinomial distribution and is given by:

$$f(x_1, x_2, \dots, x_k) = P(X_1 = x_1, X_2 = x_2, \dots, X_k = x_k) = \frac{n!}{x_1! x_2! \dots x_k!} p_1^{x_1} p_2^{x_2} \dots p_k^{x_k} \quad (16)$$

where,  $x_i \geq 0$  for all  $i$  and  $x_1 + x_2 + \dots + x_k = n$ . This formula can be derived by using the same argument that was used in deriving the formula for the binomial distribution. Specifically,  $p_1^{x_1} + p_2^{x_2} + \dots + p_k^{x_k}$  gives the probability that outcome 1 occurs  $x_1$  times, outcome 2 occurs  $x_2$  times, etc. in a specified order.

#### Geometric Distribution Model

The geometric distribution models the number of independent and identical distribution. Bernoulli trials

needed to obtain the first success. It is an example of a discrete waiting-time distribution, i.e., the distribution of discrete time to an event. Here the number of required trials is the random variable of interest. As an example, consider playing on a slot machine until hitting the jackpot. Let the probability of hitting the jackpot on any attempt represented as  $p$ . The sample space is:

$$\{S, FS, FFS, FFFS, FFFF, \dots\}$$

where,  $S$  denotes a “success” and  $F$  denotes “failure.” Let  $X$  be the number of trials required to hit the jackpot. Assuming that the attempts are independent and  $p$  remains constant, the probability of hitting the jackpot on the  $x$ th attempt is:

$$f(x) = P(X = x) = \underbrace{P(F) \times P(F) \times \dots \times P(F) \times P(S)}_{x-1 \text{ unsuccessful attempts}} \quad (17)$$

$$= (1-p)^{x-1} p, x = 1, 2, \dots$$

This is a geometric distribution with parameter  $p$ . Its CDF is given by:

$$F(x) = P(X \leq x) = \sum_{k=1}^x (1-p)^{k-1} p = 1 - (1-p)^x \quad (18)$$

## Experimental Results and Discussions

### BIUST Network Packet Intervals

#### Live Packet Caption in Morning Time

The data collection was from the live BIUST Network in the morning between 10:00:23 to 10:15:23 on the 06-07-2015. The data is presented in Table 1.

#### Day Live Packet Caption for a Working Hours

The data collection was from the live BIUST Network for the whole day between for a period of 8 hours on the 10-09-2015. The data is presented in Table 2. The data monitoring and capture was reported in (Solomon *et al.*, 2016).

#### Sizes and Frequency of Occurrence of the Packets

The first two columns of Table 1 show the values in the log file obtained from Wireshark and the remaining columns are derived from it. The first line shows 0 packets with sizes ranging from 0-19 (bytes). Using the concept of intervals, class limits and the midpoint of a class from statistical theory, the fourth column of Table 1, Average Packet Length ( $APL_i$ ), is obtained.  $APL$  for the interval  $i$  is given by:

$$APL_i = \frac{x_{mi} + x_{Mi}}{2}, 1 \leq i \leq 10 \quad (19)$$

In which,  $x_{mi}$  and  $x_{Mi}$  are the lower and higher values of the  $i$ -th interval, respectively. Column "PL" in Table 1. PL shows the packet size (in bytes) and FO is the frequency of occurrence of the packets.

APL Standard value APLs shown in Table 1 is obtained from Equation 20 dividing the value of APLi by MTU (i.e., Maximum Transmission Unit) (1500), standard for Ethernet networks. In the last column of the same table, the values FD or (FDs) are obtained dividing

the number of occurrence of the packet size by the total number of packets, given by Equation 21:

$$APL_s = \frac{APL_i}{1500} \quad 1 \leq s \leq 10 \quad (20)$$

$$FD_s = \frac{FO_i}{\sum_{i=1}^{10} FO_i} \quad 1 \leq s \leq 10 \quad (21)$$

Table 1. Live packet caption in morning time

$i$	PL	FO	APLi	APLs	FDs
1	0-19	0	9.5	0.00633	0.000000000
2	20-39	0	29.5	0.19670	0.000000000
3	40-79	27801	59.5	0.03967	0.451065970
4	80-159	16063	119.5	0.07967	0.260619139
5	160-319	2034	239.5	0.15967	0.033001266
6	320-639	2072	49.5	0.03300	0.033617808
7	640-1279	927	959.5	0.63967	0.015040400
8	1280-2559	11711	1919.5	1.27967	0.190008761
9	2560-5119	990	3839.5	2.55967	0.016062563
10	5120-10239	36	7679.5	5.11967	0.000584093
Total		61634	14905.0		

Table 2. Live packet caption in working hours

$i$	PL	FO	APLi	APLs	FDs
1	0-19	0	9.5	0.00633	0.000000000
2	20-39	0	29.5	0.1967	0.000000000
3	40-79	46783	59.5	0.03967	0.026132732
4	80-159	1563211	119.5	0.07967	0.873201256
5	160-319	115409	239.5	0.15967	0.064466847
6	320-639	46708	49.5	0.03300	0.026090838
7	640-1279	2654	959.5	0.63967	0.001482510
8	1280-2559	11976	1919.5	1.27967	0.003389729
9	2560-5119	3264	3839.5	2.55967	0.001823253
10	5120-10239	202	7679.5	5.11967	0.000112836
Total		1790207	14905.0		

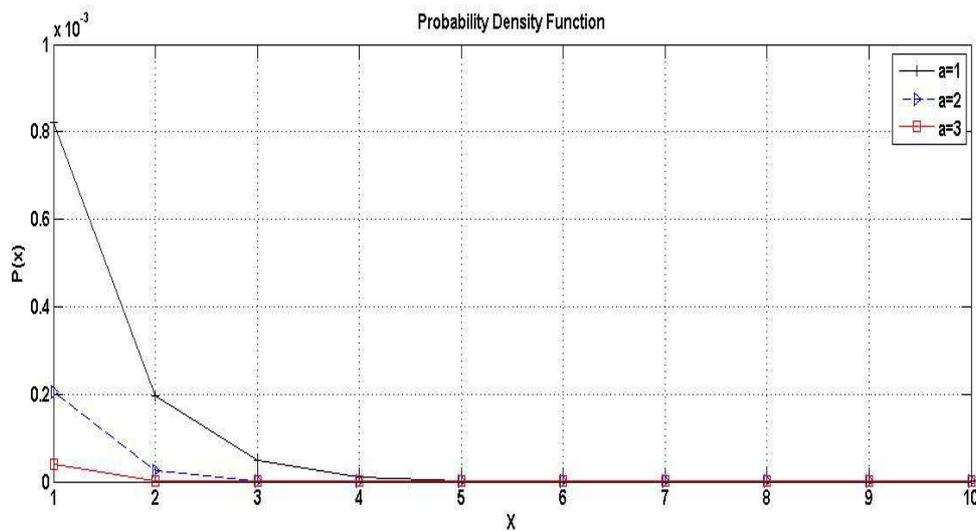


Fig. 2. Probability Density Function (PDF)

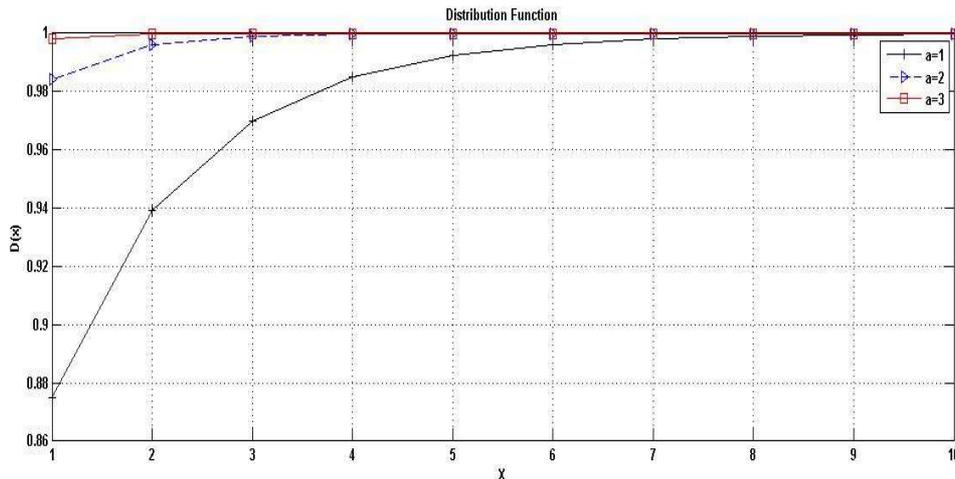


Fig. 3. Cumulative Distribution Function (CDF)

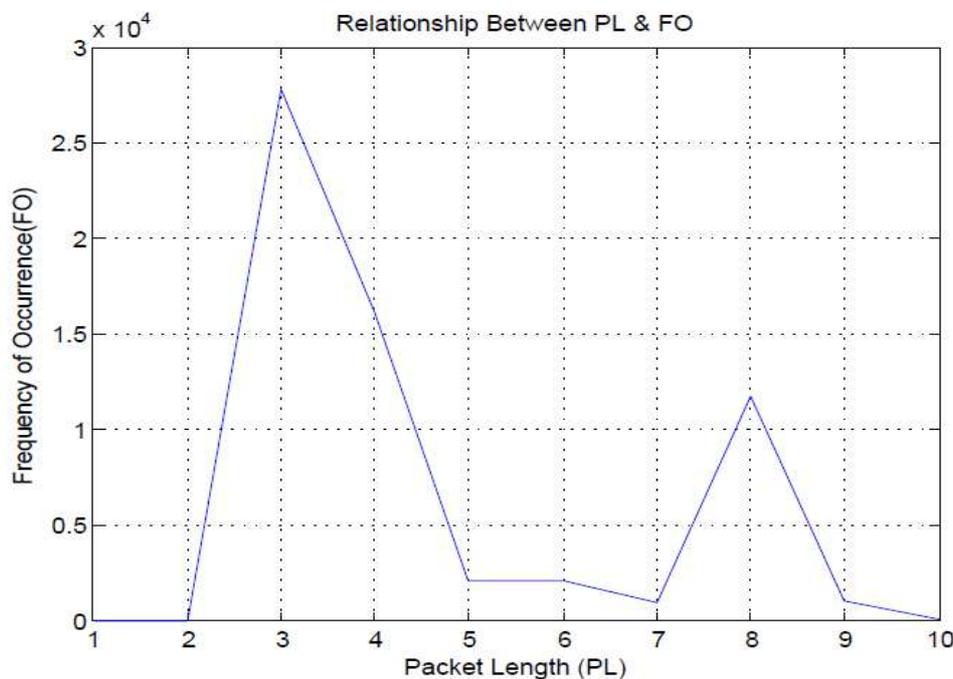


Fig. 4. Relationship between PL and FO

### Mathematical Model

The mathematical model is based on the analysis of Table 1, data input and data output of the system and being modelled through the use of Matlab software. The mathematical model used is Pareto distribution which its probability density function is given by Equation 10. The parameter  $a$  was varied such that  $1 \leq a \leq 3$ . This is presented in Fig. 2 and 3. In Fig. 2, it was observed that the PDF curve of BIUST network with a parameter of 1 reaches 0 for  $x$  value of 5, with a parameter of 2 it reaches 0 at for  $x$  value of 4 and a parameter of 3 at for  $x$  value of 3. In

this Figures, the  $a$  value was adjusted to show variation in the shape parameter and to also note the shape of the PDF. For selected values of the parameter, the simulation was run to compare the empirical density function to the PDF.

The cumulative distribution function curve of BIUST network is then presented in Fig. 3. In this Figure,  $D(x)$  scale axis was adjusted to show that the differences between the curves depend on the value of parameter  $a$ . This confirms the results of Vilfredo Pareto (Alzaatreh and Famoye, 2012) as the Figures show skewed distribution with heavy or slowly decaying tails as much of the data is in the tails.

As Pareto distribution uses the 80-20 rule corresponds to a particular value of parameter ( $\alpha$ ), it was observed that about 20% of the users had about 80% of the bandwidth consumed. The PDF graphs shows that the probability or fraction is rather high at a small value of  $PL$  and then decreases steadily as value of  $PL$  increases. Rather the CDF curve is high at higher values of  $PL$  then decreases steadily as value of  $PL$  decreases. Figure 4 shows the relationship between Packet Length (PL) and Frequency of Occurrence (FO). The result shows that, the Frequency of Occurrence (FO) is much higher at the packet length of 3 which corresponds to (40-79 bytes).

## Conclusion

This paper presents a Model used in traffic engineering to predict network performance and to evaluate congestion control schemes. From the result, it shows that, traffic models must have a manageable number of parameters and the estimation of these parameters must be simple. BIUST network was used as a case for the packet distribution model and Pareto distribution approach was adopted and it was observed that the PDF curve changes when there is a change in parameter. As the research was conducted within the boundaries of BIUST network and the attained results might be difficult to generalize to other campus networks due to the difference in network setups and policies implemented. As such, future studies may focus on applying the research on different campus networks so that the results can be generated across different network setups.

## Acknowledgement

I am very much grateful to the management of Botswana International University of Science and Technology (BIUST) for the sponsorship and their support throughout the research.

## Author's Contributions

All authors contributed equally to the implementation of the present paper.

## Conflict of Interest Declaration

The author(s) declare(s) that there is no conflict of interest regarding the publication of this paper.

## References

Alzaatreh, A. and F. Famoye, 2012. Gamma-pareto distribution and its applications. *J. Modern Applied Stat. Meth.*, 11: 78-94.

- Ascombe, F.J., 1949. The statistical analysis of insect counts based on the negative binomial distribution. *Biometrics*, 5: 165-173. DOI: 10.2307/3001918
- Bevrani, H. and S. Sharififar, 2014. An approximation to binomial distribution. *J. Stat. Theory Pract.*, 14: 1-8.
- Bobbio, A., A. Horvath, M. Scarpa and M. Telek, 2003. Acyclic discrete phase type distributions: Properties and a parameter estimation algorithm. *Performance Evaluat.*, 54: 1-32. DOI: 10.1016/S0166-5316(03)00044-0
- Boxma, O. and I. Cohen, 2000. *Self-Similar Network Traffic and Performance*. 1st Edn., John Wiley and Sons, New York.
- Castro, E., A. Kumar, M.S. Alencar and I.E. Fonseca, 2016. A packet distribution model for computer networks.
- Chandrasekaran, B., 1994. Survey of network traffic models. *IEEE Commun. Magaz.*,
- Erlina, 2011. What is mathematical modeling?
- Faraj, R., 2000. Modeling and analysis of self-similar traffic in ATM networks. PhD Thesis.
- Frost, V.S. and B. Melamed, 1994. Traffic modeling for telecommunications networks. *IEEE Commun. Magaz.*, 32: 70-81. DOI: 10.1109/35.267444
- Giambene, G., V. Kueh and C. Párraga Niebla, 2005. Packet scheduling techniques for HSDPA and MBMS transmissions in satellite UMTS. *Proceedings of the 23rd AIAA International Communications Satellite Systems Conference*, Sept. 25-28, Rome, Italy.
- Hayes, B.J. and T.V.J. Ganesh, 2004. *Modeling and Analysis of Telecommunications Networks*. 1st Edn., John Wiley and Sons, New York.
- Jeong, H.D., 2002. Modelling of self-similar teletraffic for simulation. PhD Thesis, University of Canterbury, USA.
- Kushner, H.J., 2001. *Heavy Traffic Analysis of Controlled Queueing and Communication Networks*. 1st Edn., Springer, New York, ISBN-10: 0387952640, pp: 515.
- Mehdi, J., 2003. *Stochastic Models in Queueing Theory*. 1st Edn., Academic Press, Amsterdam, ISBN-10: 0124874622, pp: 482.
- Osterbo, O., 2003. Models for calculating end-to-end delay in packet networks. *Teletraffic Sci. Eng.*, 5: 1231-1240. DOI: 10.1016/S1388-3437(03)80272-X
- Radev, D. and I. Lokshina, 2009. Modeling and simulation of self-similar wireless IP network traffic. *Proceedings of the Wireless Telecommunications Symposium*, Apr. 22-24, IEEE Xplore Press, pp: 1-6. DOI: 10.1109/WTS.2009.5068931
- Radev, D., 2005. Fluid flow analysis for long-range dependent traffic. *Proceedings of the Computer Science, (CS' 05)*, Chalkidiki, Greece.

- Ravid, D. and I. Lokshina, 2007. Modeling and simulation of self-similar teletraffic. Proceedings of the Industrial Simulation Conference, (ISC' 07), Delft, The Netherlands.
- Smith, M.J.D., 2015. Statistical analysis handbook.
- Solomon, T., A.M. Zungeru and R. Sevaraj, 2016. Network traffic monitoring in an industrial environment. Proceedings of the 3rd International Conference on Electrical, Electronics, Computer Engineering and their Applications, Apr. 21-23, IEEE Xplore Press, pp: 133-139.  
DOI: 10.1109/EECEA.2016.7470779
- Tamhane, A.C. and D.D. Dunlop, 2000. Statistics and Data Analysis "for Elementary to Intermediate. 1st Edn., Prentice-Hall, USA.
- Wilson, M., 2016. A historical view of network traffic models.
- Yang, X. and A.P. Petropalu, 2001. The extended alternating fractal renewal process for modeling traffic in high-speed communication networks. IEEE Trans. Signal Process., 49: 1349-1363.  
DOI: 10.1109/78.928689
- Zou, X., 2016. Source traffic modeling and generation. Computer Science Department, KSU.