International Journal for Development of Science and Technology (IJDST)

Vol. 1, Issue 2, September 2017, PP. 1-14

www.ijdst.com

ISSN: 2456-8503

PRODUCTLOG PROBABILITY DISTRIBUTED PACKET SIZES OF VBR VIDEO SEQUENCES FOR LONG AND SHORT RANGE DEPENDENCIES

Mr. Khadija Mkocha, Dept. of Electronics and Telecommunications Engineering, University of Dares Salaam, P. O. Box 33335, Dares Salaam, Tanzania.
Mr. Mussa M. Kissaka, Dept. of Electronics and Telecommunications Engineering, University of Dares Salaam, P. O. Box 33335, Dares Salaam, Tanzania.
Dr. Omar F. Hamad, Dept. of Electronics and Telecommunications Engineering, University of Dares Salaam, P. O. Box 33335, Dares Salaam, Tanzania.

Abstract

Self similar models capture the stochastic fractal nature of Variable Bit Rate (VBR) sequences which exhibit both long memory and short memory. Long Range Dependence (LRD) has been found to be much significant if the system operates in relatively large time scales whereas Short Range Dependence (SRD) seems to be more prominent when the system has short memory with small buffer sizes of till about 1000 cells. Long memory results into a familiar burstiness, even at low resolutions, thus, having consequences to loss probability and to mean queuing delay. The best models for LRD include the Deterministic Chaotic maps, Fractional Gaussian Noise, as well as models of ON/OFF sources. These are either aggregated or with ON and OFF periods distributed according to heavy tailed marginals. On the other hand, the best models for SRD include most classical models such as the Markovian and autoregressive/moving average models and their combinations. Models are now needed that are parsimonious yet capturing the intrinsic properties of multimedia traffic. To the best of our knowledge there are no models that capture both LRD and SRD. This work, therefore, aimed at deriving a marginal probability distribution of packet sizes of VBR video sequences which exhibit both long range and short range dependencies of typical video sequences. An extensive study of the statistical properties of packet sizes of VBR video sequences coupled with the application of the Principle of Maximum Entropy and Stochastic Programming were employed where two important results were obtained. Firstly, the data was found to assume a Product Log (Lambert W) probability distribution. Secondly, the derived probability distribution suggests that the sequences are intrinsically fractal in nature, complex stochastic processes, and possibly, more amenable to circular statistics.

Keywords: *Product log; probability distribution; vbr video sequences; long and short range dependencies; and traffic characterization.*

1. Introduction

Traffic characterization is a crucial part of network planning as it plays a significant role on how best to provision for the required services from the corresponding network (Babic, G. et. al, 1998). Specifically, the behavior of the traffic will directly influence all control functionalities, namely admission control, flow control, error control, as well as congestion control of the traffic within the network.

Before Leland's (1990) work, sources of traffic in any network would simply be modeled as being marginally distributed according to the Poisson distribution or as simple Autoregressive/ Moving Average time series (Lazaris and Koustakis, 2008; Sisodia, et.al., 1998; Adas, 1997). These models eventually fell short of representing the dynamics of traffic sources in the world of packet – switched, packet sources (Tanwir and Perros, 2013; Hlavacs, et.al 2003; Chen, 2007). As a result, towards the end of the twentieth century, a good number of models proliferated including heavy tailed marginal distributions (Zukerman, et.al., 2003; Gomes, et.al., 2009), fractional autoregressive models (Park and Willinger, 2000; Abry, et.al., 2003), chaotic maps (Erramilli, et.al., 1994), as well as the self-similar models (Fras, et al., 2012). It however became apparent that whereas the new models could very well capture the long range dependence (LRD) of network as well as source traffic, short range dependence (SRD) was under-represented and vice versa (Garroppo, et.al., 1997). It is the aim of this work, therefore, to develop a parsimonious stochastic model that at least captures the statistical properties, specifically, the marginal distribution, of the VBR sequence frame sizes.

2. Related Work

Surveys (Tanwir and Perros, 2013; Mohammed and Agamy, 2011; Chandrasekaran, 2006; Adas, 1997), have been done to show the strides taken towards the complete characterization of VBR sources. The general trend is such that the simple and mathematically tractable models are not very accurate whereas the most accurate models are mathematically complex.

On employing parametric methods into modeling the observed VBR sequences as time series some of the authors have tried fitting the statistical properties of the observed sequences into known probability distributions. However, studying the statistical properties of the VBR sequences and comparing to those charted in (Leemis and McQueston, 2008) there is no common probability distribution that fits the profile. To the best of our knowledge, no other author attempted to derive the maximum entropy probability distribution of a time series employing stochastic programming given the constraint that the means are given in form of a probability distribution rather than a constant parameter.

3. Statistical Analysis of VBR Video Sequence Packet Sizes

3.1 Visual Inspection and Given Data

The data utilized in the study involved 25 sequences of different types including newscast, cartoons, as well as action/sports events in 7 to 15 different types of encodings, from the Telecommunications Networks Group (TKN) at the Faculty of Electrical Engineering and Computer Science of the Technische Universitat Berlin. There is an extensive summary of the statistical properties of the sequences by

Fitzek and Reisslein (2000) giving, in addition to the usual statistics concerning mean and peak values of frame rates and frame sizes, some useful statistics of up to the second moments which illustrate further the burst nature of VBR traffic. For this study, only MPEG encoded sequences were used. As illustrated in Figure 1, the traces maintain strong correlations to lags of up to, sometimes, 200. This is strong evidence of long range dependence, where autocorrelation, with increasing lags, decreases slowly, exponentially at best (Fitzek and Reisslein, 200). On the other hand, beyond long range dependence, self – similarity is also evident in the frame size distribution as evidenced in the trace descriptions depicted in Figure 2. In this case, it has been shown that the Hurst parameter is maintained above 0.5 for up to aggregation levels of 800, where in some video sequences the Hurst exponent is maintained above 0.72 over all aggregation levels (Fitzek and Reisslein, 2000).

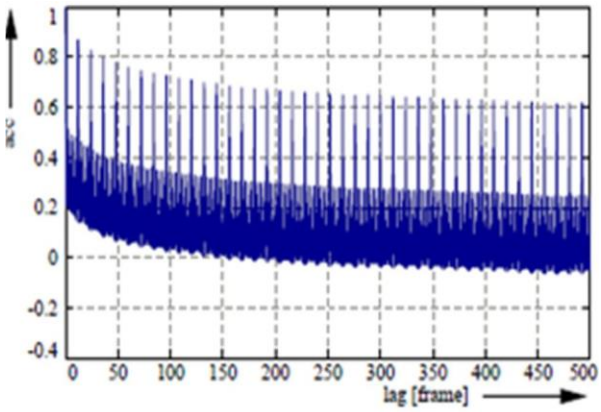


Figure 1: Autocorrelation of frame sizes with increasing lag (fitzek and reisslein, 2000)

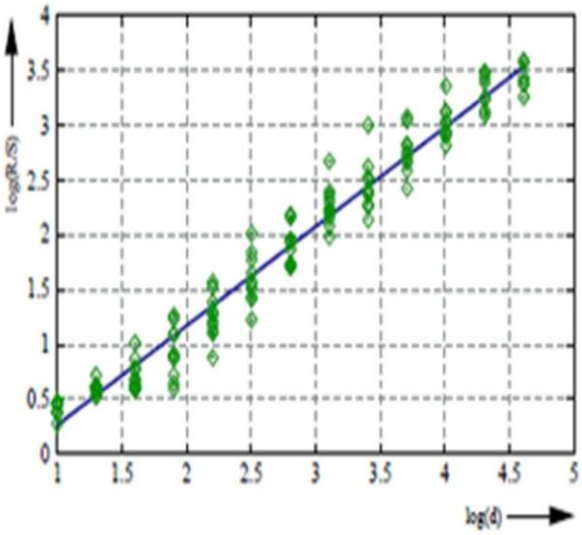
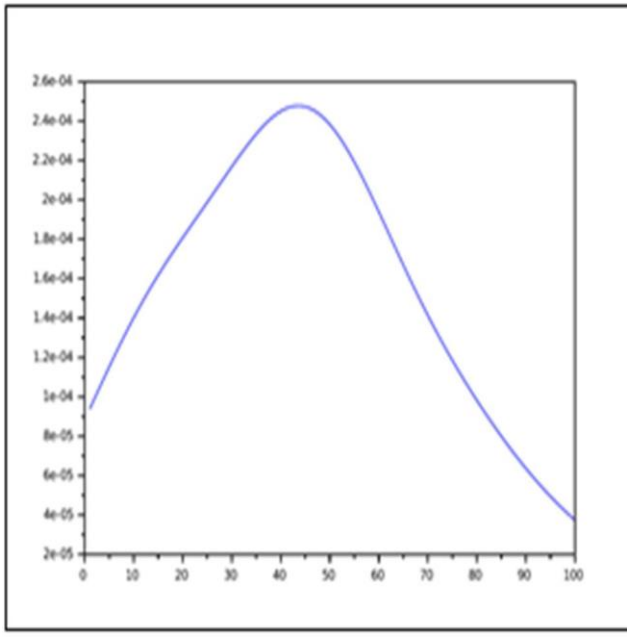


Figure 2: The R/S plot of a typical MPEG VBR sequence at aggregation level α = 1 (fitzek and reisslein, 2000)

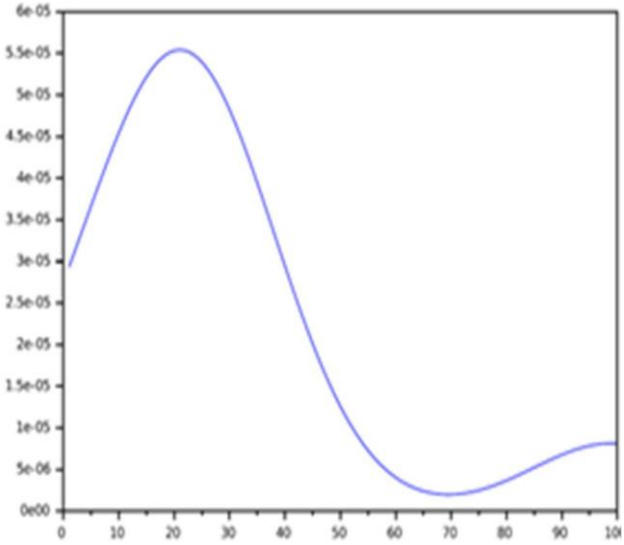
3.2 Profiles estimation of the 1st, 2nd, 3rd and 4th moments of the video sequences

Using the KS density function in Scilab, the densities of the mean frame sizes as well as the mean frame rates across the sequences were simulated. The results obtained were not surprising since means have a reputation of being normally distributed. This is illustrated in Figure 3. In this case, the estimated density had a skewness of -0.243 and a kurtosis of - 1.21, which are both within the acceptable range of normality (George & Mallery, 2010).

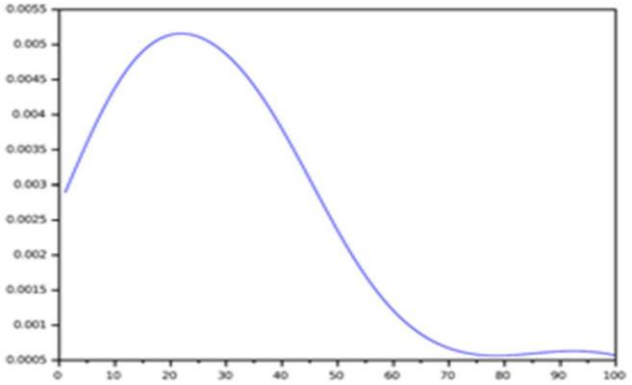


means

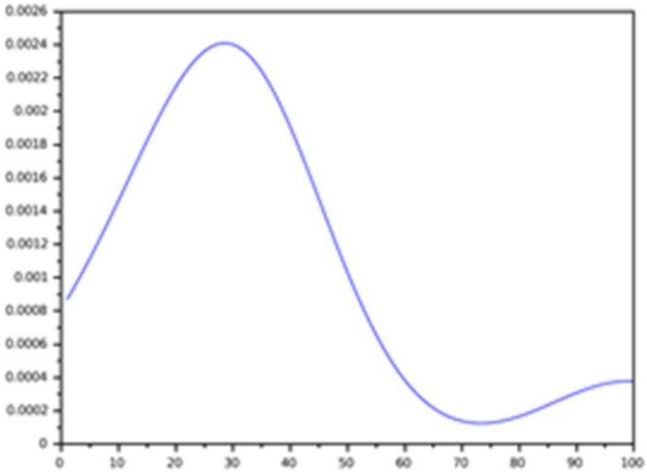
Upon further investigation, some more densities were estimated, corresponding to the maximum frame sizes (maxima), the minimum frame sizes (minima), as well as the second, third and fourth moments. These are all illustrated in Figure 4.



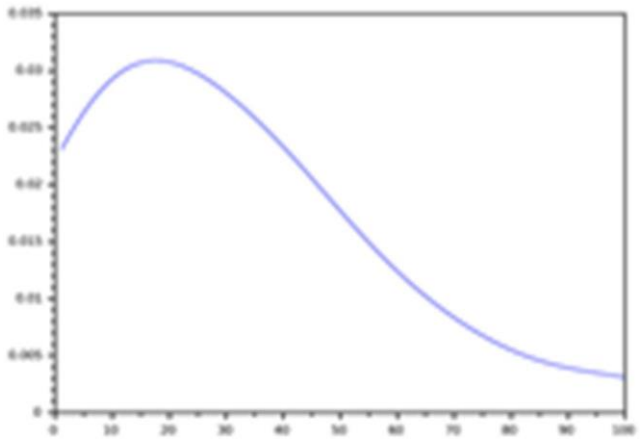
(a) Frame sizes



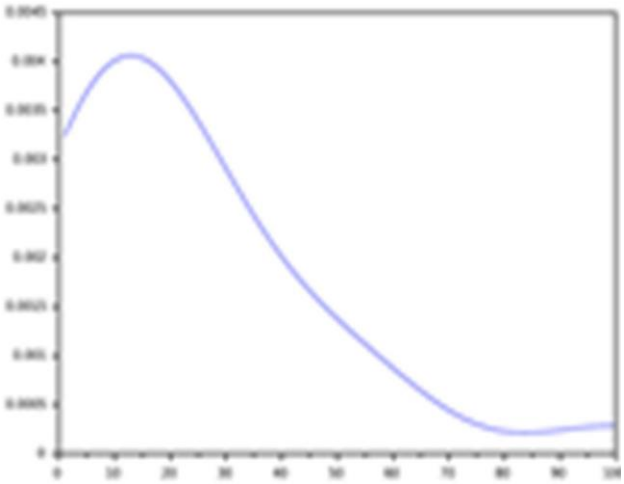
(b) Maxima



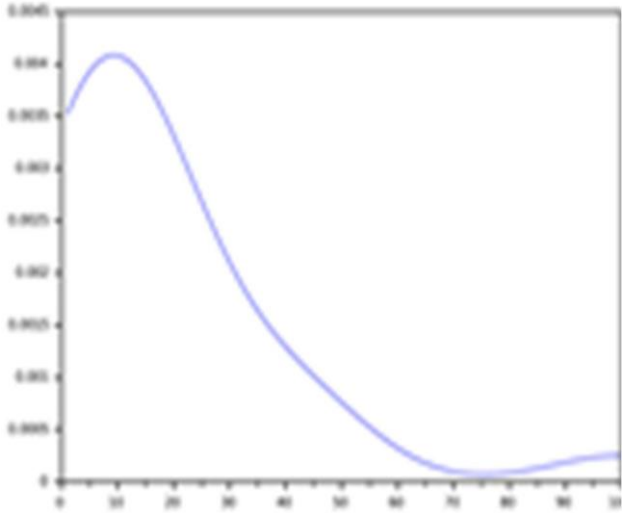
(c) Minima



(d) 2nd Order moments



(e) 3rd Order moments



(f) 4th Order moments

Figure 4: Kernel density estimations of frame sizes and their higher order moments

The motivation behind this analysis is a publication (Leemis and McQueston, 2008) which shows how different distributions behave, and are related to one another, as shown in Figure 5. Interestingly, there is no any distribution documented whose density matches the VBR traces, namely; showing scaling, being equally distributed to its minima and maxima, as well as its higher moments. It is only logical to assume at this point that there is yet an unknown distribution that fits MPEG VBR sequences.

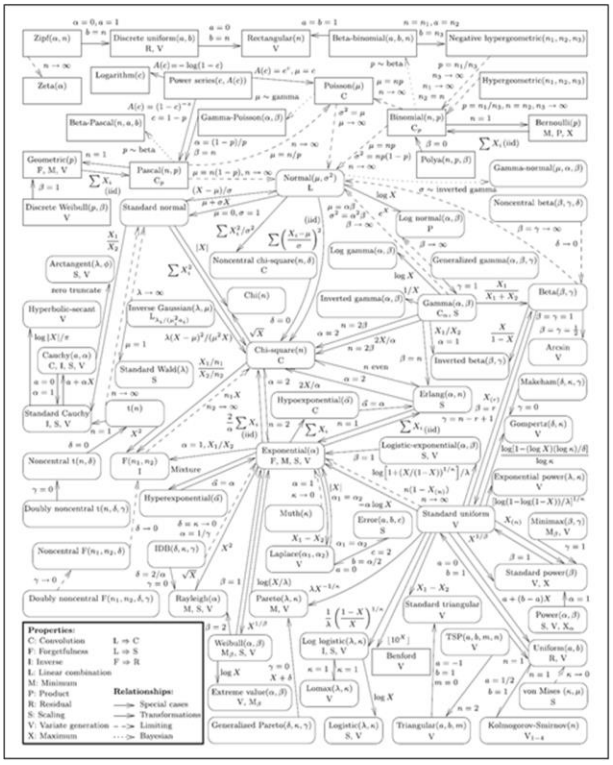


Figure 5: Relationships among common probability distributions (leemis and mcqueston, 2008)

4. Maximum Entropy Distribution for Packet Sizes of VBR Video Sequences

4.1 Formulation of the Optimization Problem

Let the unknown probability be p(x). The objective is to find the distribution p(x) which maximizes the entropy of the packet sizes of any given video sequence. It follows that, the objective function is: -

$$\text{Maximize} \quad -\int_x p(x)\log p(x)dx \qquad (1)$$

Previous work utilizing the principle of Maximum Entropy set up constant mean and variance as constraints. However, analysis of the traces done above shows that the means are normally distributed. The constraint is therefore set up to be probabilistic in this case. It follows that equation (1) will be maximized subject to the following constraint: -

$$\int_x xp(x)\,dx \sim N(\mu,\sigma^2) \qquad (2)$$

However, an equation (2) is not correctly written for optimization. To express it in terms of the distribution functions of the Gaussian density, one possible scenario from equation (2), we have: -

$$\int_{x_i} x_i\, p(x_i)\,dx \le m_i \qquad (3)$$

Where mi is the mean frame size of the it sequence, which is normally distributed. Using the probabilistic approach of chance constrained stochastic programming (Shapiro et.al. 2009); equation (3) can be

Converted into a goal constraint by assigning an arbitrary reliability on how well the constraint in equation (3) should be satisfied.

$$\text{Let } P_m = \text{Prob}\left\{\int_{x_i} x_i\, p(x_i)dx \le m_i\right\} \qquad (4)$$

Setting the reliability criterion to be 99%, it follows that

$$\text{Prob}\left\{\int_{x_i} x_i\, p(x_i)dx \le m_i\right\} \ge 0.99 \qquad (5)$$

And by the definition of the Cumulative Distribution Function F, equation (5) can be written as: -

$$F_x\left(\int_x xp(x)dx\right) \ge 0.99 \qquad (6)$$

In other words,

$$\int_x xp(x)dx \ge F^{-1}(0.99) \qquad (7)$$

It follows that, since the mean is normally distributed, F-1(0.99) can be estimated, using the CDF of the standard Gaussian distribution as

$$F^{-1}(0.99) = \mu + \varphi^{-1}(0.99)\sigma \approx \mu + 2.33\sigma \qquad (8)$$

4.1. Solution to the Maximization Problem

It follows that; equation (2) can be rewritten as:

$$\int_x xp(x)dx >= \mu + 2.33\,\sigma \qquad (9)$$

and the resulting Lagrangian is

$$L = -\int p(x)\log p(x) - \lambda\left[\int xp(x)dx - \mu - 2.33\sigma\right] \qquad (10)$$

To find the extreme of this function with respect to p(x), Leibniz Integral rule was applied to get

$$\frac{dL}{dp(x)} = \int_x p(x) \cdot \left(\frac{1}{p(x)}\right) + \log p(x) dx + \int_x x\, dx$$

$$= -\left(x + \frac{1}{p(x)} \log p(x) - x \right) \frac{x^2}{2} + c = 0$$

(11)

Which implies that

$$\frac{-1}{p(x)} \log p(x) + \frac{x^2}{2} + c = 0$$

(12)

From the equation, it can be seen that p(x) can be solved using the Lambert W equation. Using the Wolfram Alpha online solver, we get

$$p(x) = e^{-\left\{ W_{-1}\left[\left(\frac{-x^2}{2} - c \right) \right] \right\}},$$

where,

$$2c + x2 \neq 0 \qquad (13)$$

which is illustrated in Figure 6 for a few values of x.

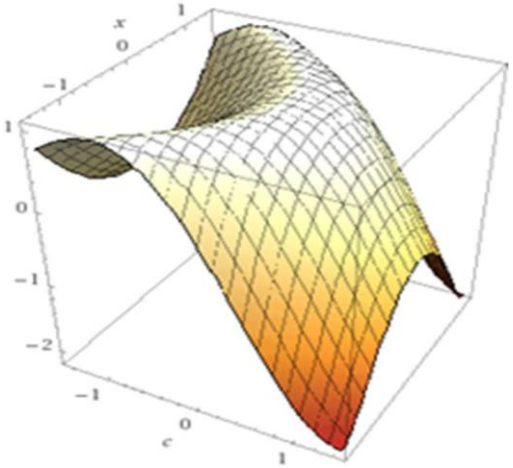


Figure 6: A three dimensional plot of the real part of the derived pdf

There are two things to take note of at this point. To begin with, p(x) is a general solution (not particular). Secondly, Figure 4.8 shows only the real part of p(x), implying that p(x) is a complex process. The resulting CDF is similarly calculated using the online solver and the resulting expression is given by equation 14. Figure 7 is the plot showing the (real part of the) CDF.

$$\int e^{W(-1)\left(-\frac{x^2}{2}-c\right)} dx = \frac{e^{-c\,W(-1)} \operatorname{erf}\left(x\sqrt{\frac{W(-1)}{2}}\right)}{\sqrt{\frac{2\,W(-1)}{x}}} + \text{constant}$$

(14)

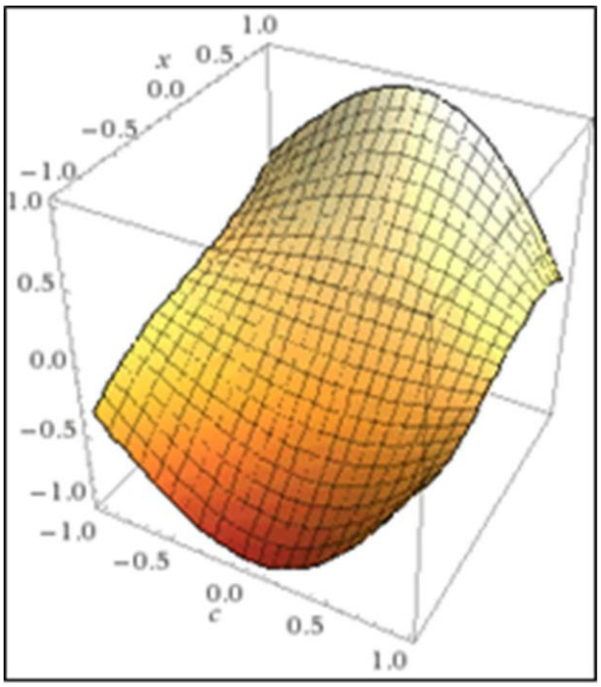


Figure 7: The cumulative distribution function

5. Analysis of the Method

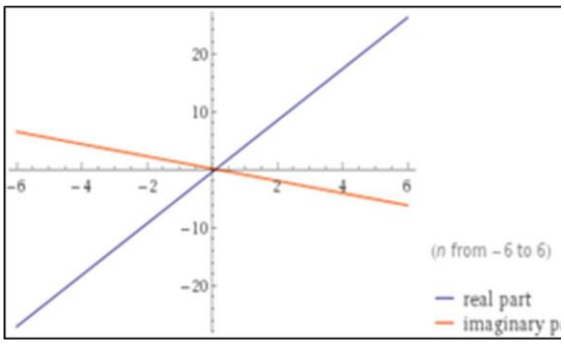
5.1 Complex Nature of the Probability Distribution

The fact that the derived probability distribution is complex in nature, qualifies the observed stochastic process to be expressed as an observation in the context of either the Riemann manifolds or real multidimensional spaces $\Re n$ (Bender, et.al. 2010). In this sense, the stochastic process is therefore more amenable to analysis under circular/directional statistics, which is coincidentally a typical characteristic of temporal data (Mardia and Jupp, 2000). It is only logical at this point to assume that this observation is the result of the intricate temporal relations of the data series under analysis, manifested as self similarity. In addition, the circular behavior of the series could be the result of repeated I-P-B frame structures within a GOP, as well as the inter-GOP similarities.

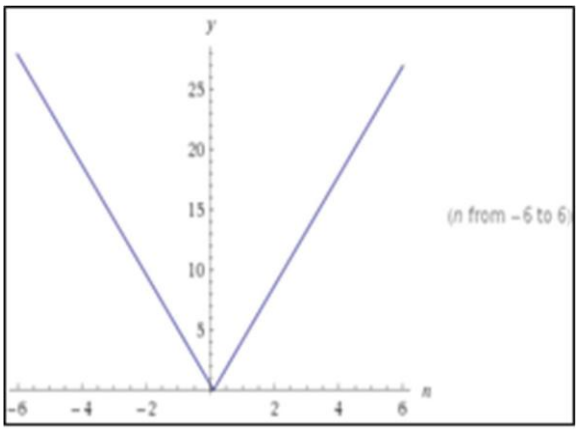
5.2 The Significance of the Constant 'c'

The integration constant *c* was solved once more using the online tool and it was evaluated to the expression given by equation 15 as illustrated in Figure 8 (a). The magnitude of *c* is as shown in the plot in Figure 8(b), where c is zero when n = 0.1.

$$c = \frac{\log\left(-\frac{2\,W(-1)}{\pi}\right) + 4\,i\,\pi\,n}{2\,W(-1)} \quad \text{and } n \in \mathbb{Z} \qquad (15)$$



5.2.1 In value



5.2.2 In magnitude

Figure 8: The constant 'c' as evaluated

The representation of the constant c is clearly that of a directional parameter of circular data. It has been noted that linear statistics can be applied to circular data for small standard deviations. Assuming in this case that some values of c satisfy this condition, in this study, we will restrict the values of c to those whose real values visually resemble the kernel density estimates of the traces and continue with methodology from linear statistics for validation purposes.

5.3 Visual inspection of 'c'

Generally, it ranges from -1 to 1 in the CDF plot with the physical behavior similar to any typical CDF in linear statistics, where the maximum value is 1. The investigation, was therefore carried out to, numerically, determine the value of c. It should be pointed out here that the value of c sought after here is that which results into the PDF being limited to only real values.

The analysis of the complex PDF has been left for future work. With negative values of c, the real part is dominant but negative, whereas when c is positive, the imaginary part is dominant with a positively valued real part. At c = 0, the PDF is completely real. By iteratively substituting different values of c into the PDF equation, taking values less than c = 0.4, the PDF becomes completely real with these values of x (0, 2/e). The values: 0.3, 0.2, 0.15, 0.13, 0.12, 0.125, 0.123, 0.124, 0.1235, 0.12375 etc. The iterated values are shown in Figure 9 whereas the corresponding. PDFs are tabulated in Table 1 and Table 2.

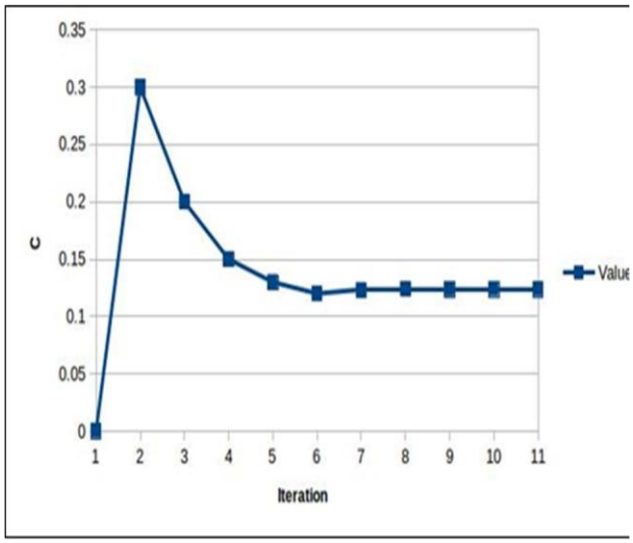


Figure 9: Iteration of different values of c to find the range within which the pdf is completely real

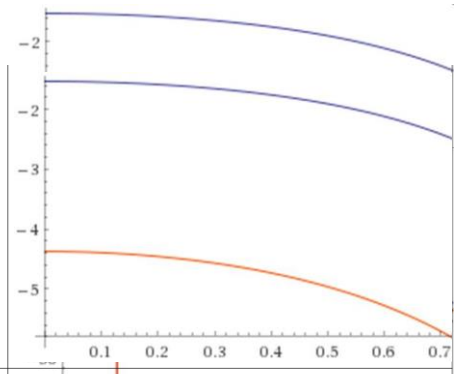
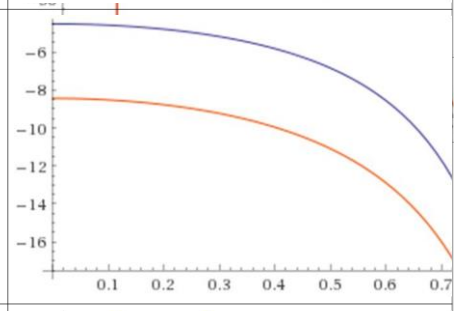
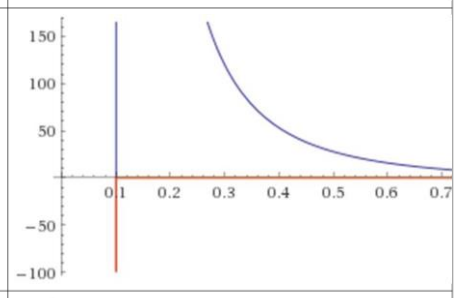
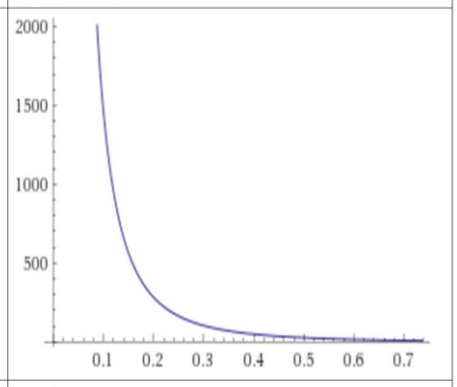
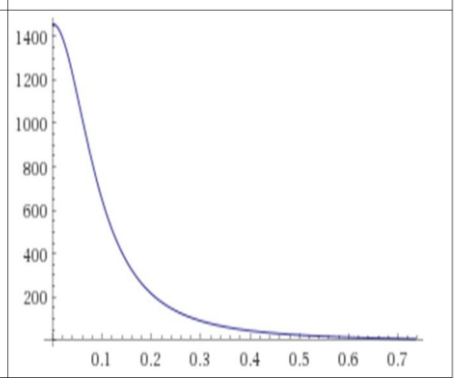
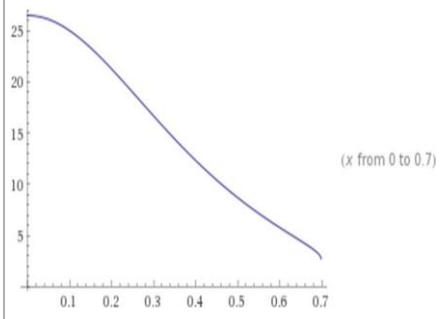
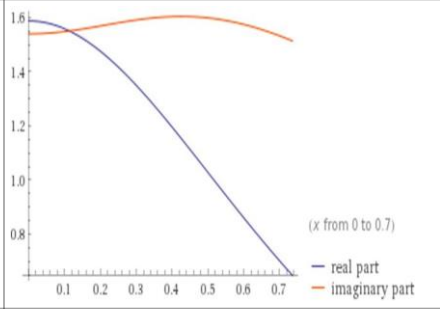
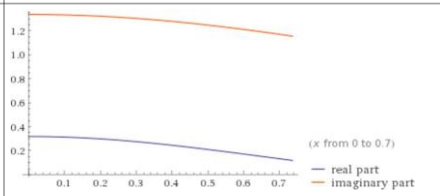
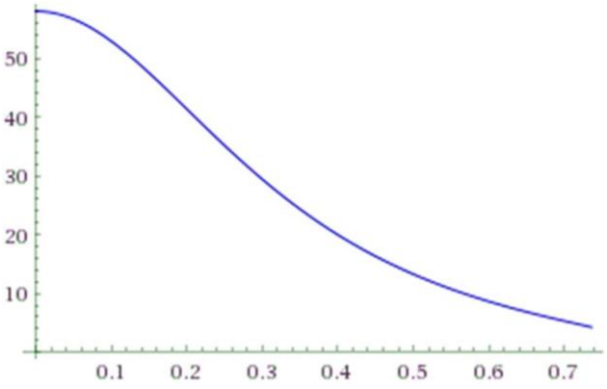
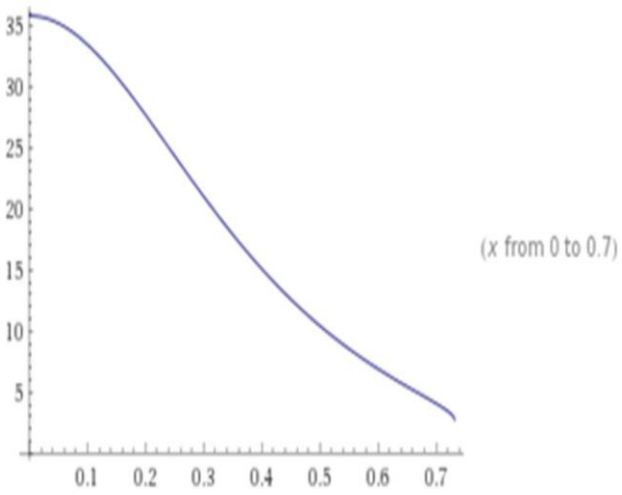
-1	
-0.5	
-0.005	
0	
0.005	

Table 1: PDF Plots with C ranging from -1 to +1

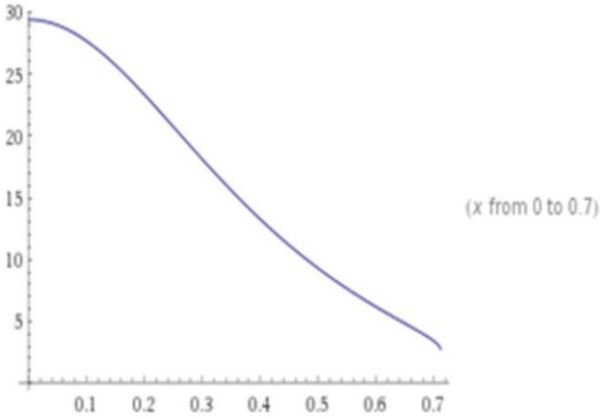
c	plot
0.12359693	
0.5	
1	



5.3.1 For c = 0.0700



5.3.2 For c = 0.1000



5.3.3 For c = 0.1150

Figure 10: Some PDFs with different values of c in the real range

An interesting observation was made from the magnitude of the PDF when plotted in Wolfram Alfa as depicted in Figure 10. An excerpt of the figure, for values of x ranging from about -3 to + 1.5 is magnified and compared to Figure 4(a), in Figure 11 where there is a significant resemblance between the two plots. Further comparison is however left for future work.

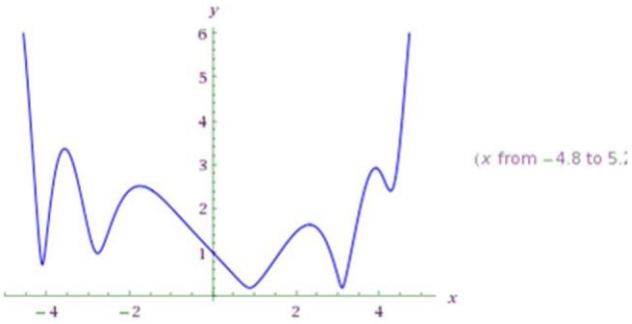
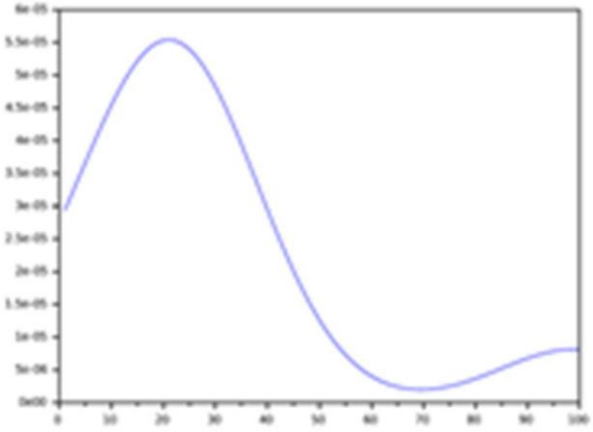
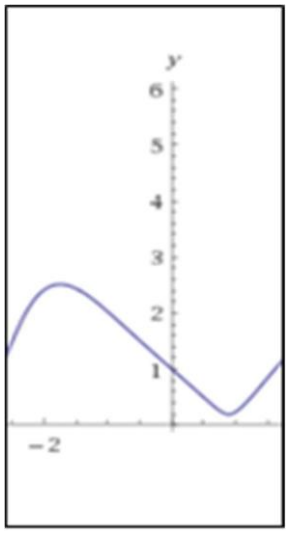


Figure 10: The magnitude plot of the complex PDF



(a)



(b)

Figure 11: A Comparison of the estimated PDF of a video sequence (a) to an excerpt from the magnitude plot of the derived PDF

6. Model Validation

In an arbitrary manner, the value of c = 0.1 Was chosen for the purpose of validating the model. The chosen value was inserted into equation 13 to get the *truncated PDF*. The resulting plot is shown in Figure 12. The PDF was used to generate samples that would be compared to the VBR sequences using Q-Q plots.

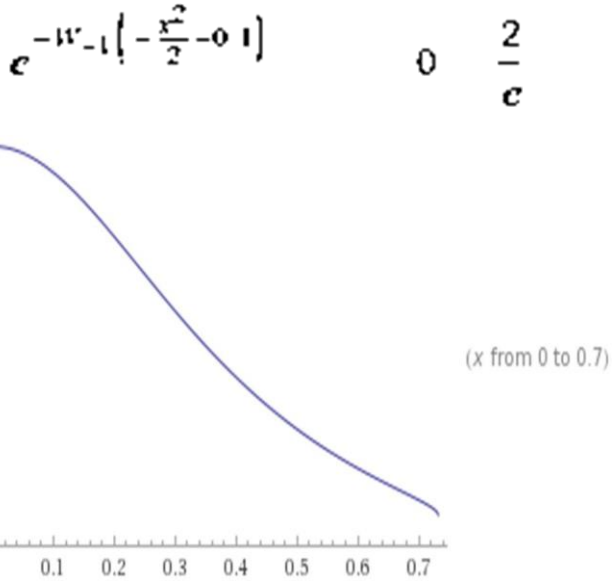


Figure 12: Truncated PDF

4.2 Generating the Traces

The traces were generated using the Acceptance–Rejection Method in the R Studio. The corresponding code for the generation of the samples is given in the Appendix. Figure 13 shows a sample histogram from the generated samples.

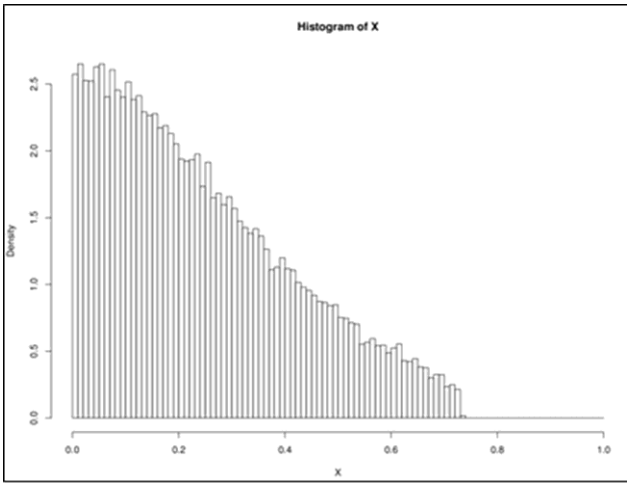
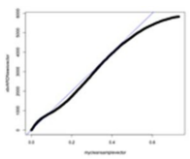
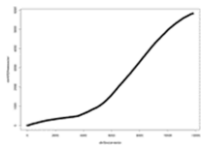
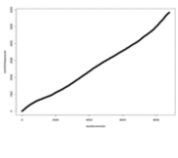
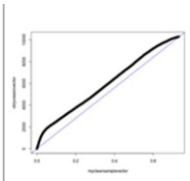
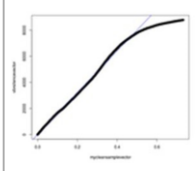
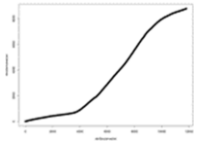
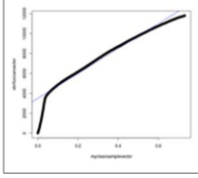
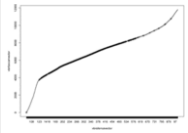


Figure 13: Histogram of samples from the derived PDF

The Q-Q Plots from the Derived Samples, Q-Q plots were drawn to compare the distributions of the frame sizes of the video sequences from the derived PDF as illustrated in Table 3. VBR sequences chosen included an animated sequence, an action movie, a news commentary and a thriller/ action movie. Table three below shows the Q-Q plots between the derived samples and the above mentioned sequences, as well as a few plots between different sequences in either the same or different categories.

Table 3: Q-Q Plots

Derived PDF	Sports Action	Action Thriller
News Commentary		
		
Action Movie 1		
		
Action Thriller		
		
Sports Action		
		

7. Results Discussion

Compared to the derived PDF, all the sequences are consistently linear within the mid-range of values suggesting that there is a resemblance between the distributions of the VBR sequence frame sizes and that which was derived in this study. It is however, also noted that significant differences are at the extreme values. The highest values in the plots consistently bend downwards from the linear trend suggesting fatter tails in the actual distribution as compared to the derived distribution. This observation could be explained by the fact that the tails of the derived distribution were truncated so as to limit the PDF within real values. Similarly, lower values in the Q-Q plots have consistently higher slopes when compared to the linear trend lines suggesting more skewness than that captured by the derived distribution. The truncated support of x used in generating the samples from the derived distribution where x was taken to be greater than 0.

8. Conclusion

A very interesting result in this case is the fact that VBR traffic appears to be a complex stochastic process. At a glance, by visually comparing the empirically estimated probability density function and the real part of the derived general PDF, they appear to agree with each other. In addition there is an imaginary component to the PDF which might explain why all the existing models capture only certain features of the VBR traffic at a time rather than comprehensively.

9. Future Work

To extend this work further, it will entail the modification and validation of the derived model in the context of directional statistics. It is intended to find out the regular distribution which could analytically be wrapped to result into the Product Log distribution. It is further expected that the new model will be used to measure the manifestation of SRD as well as LRD.

References

1. Abry, P., Flandrin, P., Taqqu, M. S. and Veitch, D. (2003), Self-Similarity and Long-Range Dependence through the Wavelet Lens, Theory and Applications of Long Range dependence, pp. 527- 556
2. Adas, A. (1997), Traffic Models in Broadband Networks, IEEE Communications Magazine, July 1997. Retrieved from http://www.land.ufrj.br/~verissimo/mscthesis/bibref/biblio52.pdf
3. Babic, G., Vandalore, B. and Jain, R.(1998), Analysis and Modeling of Traffic in Modern Data Communication Networks, Department of Computer and Information Science, Ohio State University, Retrieved April, 2013 from http://arvix.org/pdf/cs/9809056
4. Bender, C. M., Hook, D. W., Meisinger, P. N. and Wang, Q. (2010), Probability Density in the Complex Plane, Annals of Physics, Volume 325, Issue 11, November 2010, pp 2332 – 2362, Retrieved February 2017 from https://arxiv.org/pdf/0912.4659.pdf
5. Chandrasekaran, B. (2006), Survey of Network Traffic Models, Retrieved from http://www.cse.wustl.edu/~jain/cse567-06/ftp/traffic_models3.pdf
6. Chen, T. M. (2007), Network Traffic Modeling, Chapter in the Handbook of Computer Networks, Bidgoli, H. (ed), Wiley. Retrieved from https://pdfs.Semanticscholar.org/5091/e0fv30f8ff50ec47f43affc2bf08fac5dff0.pdf
7. Erramilli A., Singh, R.P. and Pruthi, P. (1994), Chaotic Maps as Models of packet Traffic, ITC 14, retrieved April 2017 from http://citeseerx.ist.psu.edu/viewoc/download? doi:10.1=1=46.7039 & rap=rap1&type=pdf
8. Fras, M., Mohorko, J. and Cucej, Z. (2012), Modeling of Statistical Data Sources Based on Measured Network Traffic, Simulation, Volume 88, issue 10, October 2012

9. Garroppo, R.G., Giordano, S., Pagano, M. and Russo, F. (1997), Modeling and Queueing performance Evaluation of VBR Packet – Video Traffic, Semantic Scholar, Retrieved fromhttps://pdfs.semanticscholar.org/8cb4/93614fe3a0d6c558be9b75ffed2d33689c0.pdf?_ga=1.149015670
10. George, D., and Mallery, M. (2010), SPSS for Windows Step by Step: A Simple Guide and Reference, 17.0 update (10a ed.) Boston 1747936083.491756229
11. Gomes, J. V. P., Inacio, P. R. M., Lakic, B., Freire, M. M, Silva, H. and Monteiro, P. P. (2009), Studies for Modeling Basic Aspects of Source Traffic, Proceedings of the Seventh Conference on Telecommunications (ConfTele 2009) pp. 327 – 330, Santa Maria da Feira, Portugal, May 2009
12. Hlavacs, H., Kotsis, G. and Steinkellner, C. (2003), Traffic Source Modeling, Technical report no. TR – 99101, Institute of Applied Computer Science and Information Systems, University of Vienna
13. Hossein Bidgoli (ed.), WileyFitzek, F.H. P and Reisslein, M. (2000), MPEG-4 and H.263 Video Traces for Network Performance Evaluation, TKN Technical Report, TKN group
14. TU Berlin, retrieved August 2016 from trace.eas.asu.edu/conferences/Techrep/TKN0006.pdf
15. Lazaris, A. and Koustakis, P. (2008), Modeling Video Traffic from Multiplexed H.264 Videoconference Streams, IEEE GLOBECOM 2008 proceedings Leemis L. M. and McQueston, J. (2008), Univariate Distribution Relationships, Teacher's Corner, the American Statistician, February 2008, Vol 62, No 1
16. Mardia, K. V. and Jupp, P.E. (2000), Directional Statistics, Wiley Series in Probability and Statistics, John Wiley & Sons, Ltd., Sussex, UK
17. Mohammed, A.M. and Agamy, A. F. (2011), A Survey on the common network traffic source 136.pdf models, international journal of computer networks (IJCN), Volume-3 Issue May/June 2011. Retrieved fromhttp://www.cscjournals.org/manuscript/Journals/IJCN/Volume3/Issue2/IJCN
18. Park, K. and Willinger, W. (Ed.)(2000), Wavelets for the Analysis, Estimation and Synthesis of Scaling Data, in Self Similar Network Traffic and performance Evaluation, Wiley G, pp. 39-68
19. Shapiro, A., Dentcheva, D. and Ruszczyński, A. (2009), Lectures on Stochastic Programming, Modeling and Theory, Society for Industrial and Applied Mathematics and the Mathematical Programming Society, Philadelphia, USA
20. Sisodia, G., De, S., Hedley, M. and Guan, L. (1998), New Statistical model for VBR Video Traffic in ATM Networks, IEEE
21. Tanwir, S. And Perros, H. (20130, A Survey of VBR Video Traffic Models, IEEE Communications Surveys & Tutorials
22. Zukerman, M. Neame, T. and Addie, R. G. (2003), Internet Traffic Modeling and Future Technology Implications, IEEE INFOCOM 2003