

Research Article

Statistical Analysis of Video Frame Size Distribution Originating from Scalable Video Codec (SVC)

Sima Ahmadpour,¹ Tat-Chee Wan,^{2,3} Zohreh Toghrayee,⁴ and Fariba HematiGazafi⁴

¹Department of Electrical, IT & Computer Sciences, Islamic Azad University, Qazvin Branch, Qazvin, Iran

²National Advanced IPv6 Center, Universiti Sains Malaysia, 11800 Penang, Malaysia

³School of Computer Sciences, Universiti Sains Malaysia (USM), 11800 Penang, Malaysia

⁴Central Bank of the Islamic Republic of Iran, Tehran, Iran

Correspondence should be addressed to Sima Ahmadpour; ahmadpour.sima@gmail.com

Received 30 June 2016; Accepted 15 January 2017; Published 19 March 2017

Academic Editor: Alicia Cordero

Copyright © 2017 Sima Ahmadpour et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Designing an effective and high performance network requires an accurate characterization and modeling of network traffic. The modeling of video frame sizes is normally applied in simulation studies and mathematical analysis and generating streams for testing and compliance purposes. Besides, video traffic assumed as a major source of multimedia traffic in future heterogeneous network. Therefore, the statistical distribution of video data can be used as the inputs for performance modeling of networks. The finding of this paper comprises the theoretical definition of distribution which seems to be relevant to the video trace in terms of its statistical properties and finds the best distribution using both the graphical method and the hypothesis test. The data set used in this article consists of layered video traces generating from Scalable Video Codec (SVC) video compression technique of three different movies.

1. Introduction

Generally, a thorough understanding of the traffic and quality characteristics of encoded video is the basis for traffic modeling and the development of video transport mechanisms [1]. Multimedia transmissions have imposed a huge amount of the today traffic over computer and mobile communication networks. This can be done by simply using a live experiment using real networks and real sources. However, testing real networks is fairly expensive and often it is difficult to come up with realistic results. Another solution to this would be to model the traffic using mathematical analysis or simulation. Trace-driven simulations are thought reliable because they represent an actual traffic load; nevertheless they are usually static and so they provide merely a point representation of the workload space. One more disadvantage of using traces is the difficulty in adjusting parameters and extending the trace if there is a need to continue the simulation beyond the number of packets/frames in the trace file [2]. With this intention, statistical and mathematical traffic models are

assumed as better solutions since they can be used to provide a better understanding of various traffic characteristics. This is because they are stochastic in nature, and hence different realizations that represent the actual data can be obtained by varying model parameters.

Among the various characteristics of video traffic, the following two are of major interest in literature:

- (a) Distribution of frame sizes.
- (b) Autocorrelation Function (ACF) that captures common dependencies between frame sizes in VBR video.

Among all multimedia applications, video services are demonstrated as the most common ones for generating traffic in communication networks. Obviously, the raw video data requires very high transmission bandwidth and large amount of storage space [3]. Therefore, using video compression techniques is highly recommended and there exist different types of network traffic based on their application. The focus of this paper is on video traces generated from a Scalable Video Codec (SVC) as a compression technique.

SVC is an extension of H.264/AVC which is standardized by The Joint Video Team of the ITU-T VCEG and the ISO/IEC MPEG H.264/AVC standard [4]. SVC has been proposed to support bandwidth efficient and loss resilient video streaming. Meanwhile, the encoding structure of SVC includes one base layer and one or more enhancement layers. H.264 SVC supports layer-scalable coding which presents Temporal Scalability, spatial scalability, and quality (SNR) Scalability [5]. SVC provides two types of quality scalability, known as coarse grain scalability (CGS) and medium grain scalability (MGS). In this paper, the statistical analysis of CGS encoding has been pondered.

In this strategy, each layer has an independent prediction procedure (all references have the same quality level) in a similar fashion to the MPEG-2. In fact, the CGS strategy can be regarded as a special case of spatial scalability when consecutive layers have the same resolution [6]. Coarse grain SNR scalable coding is achieved using the concepts for spatial scalability. The same interlayer prediction mechanisms are employed. The only difference is that base and enhancement layers have the same resolution. The CGS only allows a few selected bitrates to be supported in a scalable bitstream. In general, the number of supported rate points is identical to the number of layers. Switching between different CGS layers can only be done at defined points in the bitstream [7].

Communication network measurements have indicated that many quantities which are characterizing the network performance have long-tail probability distributions. The quantities have the tails that decay more slowly than exponential. This long-tail behavior is mostly related to the terms such as file lengths, call holding times, scene lengths in video streams, and intervals between connection requests in Internet traffic. Long-tail distributions can have a significant effect on performance [8]. long-tail service-time distributions lead to long-tail waiting-time distributions in the queues [9]. Since performance models with long-tail distributions are usually difficult to analyze, it is usually difficult to describe them in detail. To address this problem, finding the best distribution among all of the different distributions has come to account. The aim is to derive a statistical distribution to fit the real data accurately. Since they are the most common distributions related to data with long-tails distributions they were chosen. In other words these data have probability distributions with high skewness which are difficult to analysis by usual statistical models. Therefore, nonsymmetric distributions need to be addressed in this study.

The organization of this paper is as follows. The notable frame size distribution related works are presented in Section 2. Section 3 describes the methodology of the study in which different distributions are explained with the statistical properties in detail. Meanwhile, Section 4 performs the result of fitting these distributions to the data based on statistical criteria. Finally, the conclusion of this study is provided in Section 5.

2. Related Works

Several works have been conducted in order to analyze the video frame sizes. The early work performed by Heyman et al.

[10, 11] and Xu and Huang [12] presented the marginal distribution of videoconference encoded by H.261 which were generated by different hardware coders with different coding algorithms is gamma distribution. Aforementioned authors applied this result to design a discrete autoregressive model (DAR) of order one. Krunz and Hughes in [13] modeled the frame sizes which are compressed by MPEG-2 standard. In this study, the best fit for the distribution of frame sizes was lognormal distribution among gamma, Weibull, and lognormal distributions. They used the fitting distribution for three types of frame sizes such as I frames, B frames, and P frames. Fitzek and Reisslein [14] have provided a public available library of frame size sequences including MPEG-4, H. 263, and H.263+ encoded video with a detail of statistical analysis of generated traces. It was shown that the movies as visual content cause a frame generation with gamma-like frame size histogram. Poon and Lo [15] presented that a normal mixture distribution for fitting the sample histogram of video traces encoded by H.261 and H.263. It was proved that this method is better than simple Gamma and Lognormal. Lazaris et al. [16] indicated that Gamma and Lognormal distribution are not always the best fit for MPEG-4 videoconference traffic. Furthermore, they presented the notion that for single videoconference sources the best fit is Pearson type V distribution among all examined distributions. Koumaras et al. [3] discovered that gamma distribution is the best fit for frame sizes compressed by H.264 standard where this fits three types of video frame. Furthermore, Masi et al. in [17] indicated that the Erlang or gamma distributions are fitted appropriately in to the three data sets of actual video frame sizes. In their work the data set consists of two different video compression standards, H.263 and H.264. The best fit of frame size distributions was used to generate packet streams for use in packet level congestion models. Salah et al. in [18] figured out that gamma is a well fitted distribution to the data and Weibull distributions and inverse Gaussian distribution is ranked second after these distributions.

For modeling single source trace, the best distribution needs to be found. Although there were few studies that analyzed packet size distribution, this article considered frame size analysis as output of SVC layers.

3. Evaluation Setup of Video Sequences

The data set presented in this paper consists of three different video traces with the CIF (352×288) resolution, a frame rate of 30 frames/second, GOP pattern: G16B15, and the quantization parameters (I, P, and B): 48, N/A, 48 taken from [19]. A video trace characterizes an encoded video stream by providing time stamp, frame type (e.g., I, P, or B), frame size (in byte), and PSNR quality for each encoded frame (and layer of a scalable encoding). Video traces can be readily fed into simulation models of video transport systems, thus, facilitating the evaluation of novel transport mechanism. The video traces under this study include the following:

- (i) NBC News sequence (48992 frames) 60 minutes long divided into one base layer and one enhancement layer in which the frame types are intraframes (I frames) and bidirectional frames (B frames).

- (ii) Sony Demo sequence (17,664 frames) 60 minutes long divided into one base layer and one enhancement layer same as the previous video sequence with I and B frame types.
- (iii) Silence of the Lambs sequence (53984 frames) with the similar properties of the former sequences.

Above-mentioned sequences demonstrate video sequences with low or moderate scene changes. The particular encoder tool which is used for encoding purpose is JSVM encoder taken from [19].

4. Methodology

The proper methodology for the paper is as follows:

- (i) Investigate the hypothesized distribution families which are suitable in terms of the overall shape of the data under the study.
- (ii) Estimate the parameters of selected distributions by writing code.
- (iii) Find the best distribution for the data by goodness of fit tests.
- (iv) Investigate the autocorrelation function.

Each of these steps will be described in detail in next subsections. As the distributions studied in this paper are widely used in most literature, they will be addressed in terms of statistical theory and the relevant applications in the video traffic modeling.

4.1. Investigating the Different Distributions. Since there are numerous statistical distributions, it is not common to investigate all of them to find the best one for the data set. Therefore, plotting the density function of the data provides a preliminary point of view to identify what kinds of distributions should be studied.

Figure 1 showed that the shape of density function of NBC News with different layers and frame types is not symmetric and has high skewness. This plot had been performed for other two movies and same results are concluded. However, due to the space limitation, they were not presented for every step of implementation for rest of the article. Hence, to implement the fitting of an appropriate distribution to the data, the more common distributions with skewness include exponential, lognormal, logistic, log-logistic, Weibull, gamma, normal, inverse Gaussian, negative-binomial, and Pearson family distributions. In this article, the Pearson distribution was identified in detail, since it contains four parameters which lead to achieve more appropriate fit.

4.1.1. Pearson Distribution. Since one of the most common used distributions in literature is Pearson distribution, four-parameter system of probability density functions, which is provided by Abate et al. [9], will be explained in more detail here. On the other hand, four parameters compromise location, scale, skewness, and kurtosis of the distribution and describe the data better than the density functions with less parameters. More parameters in a distribution need

more precise description of the data. Depending on some conditions which will be discussed later, there are five Pearson distribution types. In fact, skewness and kurtosis characterize these Pearson types. Let γ_1 and γ_2 denote skewness and kurtosis of the distribution; then some reparameterization is constructed as follows:

$$\begin{aligned}\beta_1 &= \gamma_1^2 \\ \beta_2 &= \gamma_2 + 3.\end{aligned}\tag{1}$$

A valid solution of the following differential equation specifies a Pearson density family:

$$\frac{p'(x)}{p(x)} + \frac{a+x-\lambda}{b_2(x-\lambda)^2 + b_1(x-\lambda) + b_0} = 0;\tag{2}$$

therefore:

$$p'(x) = -p(x) \frac{a+x-\lambda}{b_2(x-\lambda)^2 + b_1(x-\lambda) + b_0}.\tag{3}$$

The different types of Pearson distributions can be obtained by solving (3). It is wise to call equation (3) as a first-order linear differential equation with variable coefficients. For simplicity, λ is assumed zero. Thus, (6) can be rewrite as follows:

$$p'(x) + p(x) \frac{a+x}{b_2x^2 + b_1x + b_0} = 0.\tag{4}$$

To clarify how this equation will be solved, a simple linear order differential equation is described which has the general form of

$$Dy(x) + f(x)y(x) = g(x).\tag{5}$$

At first the proof of (5) is provided in Appendix A in detail and then the assumptions (6) and (7) can be released.

$$f(x) = \frac{a+x}{b_2x^2 + b_1x + b_0} = 0,\tag{6}$$

$$y(x) = p(x)g(x) = 0,$$

$$p(x) \propto e^{-\int ((x-a)/(b_2x^2 + b_1x + b_0))dx}.\tag{7}$$

Now it is time to solve (7) which is the primer beginning of different types of Pearson distribution. According to the discriminate of equation in denominator, there are two general cases which leads to production of different kinds of Pearson distribution which are displayed in next subsection.

(A) Case 1 Nonnegative Discriminate. If the discrimination is positive, the equation contains two roots named a_1 and a_2 , being maybe the same:

$$\begin{aligned}a_1 &= \frac{-b_1 - \sqrt{b_1^2 - 4b_0b_2}}{2b_2} \\ a_2 &= \frac{-b_1 + \sqrt{b_1^2 - 4b_0b_2}}{2b_2}.\end{aligned}\tag{8}$$

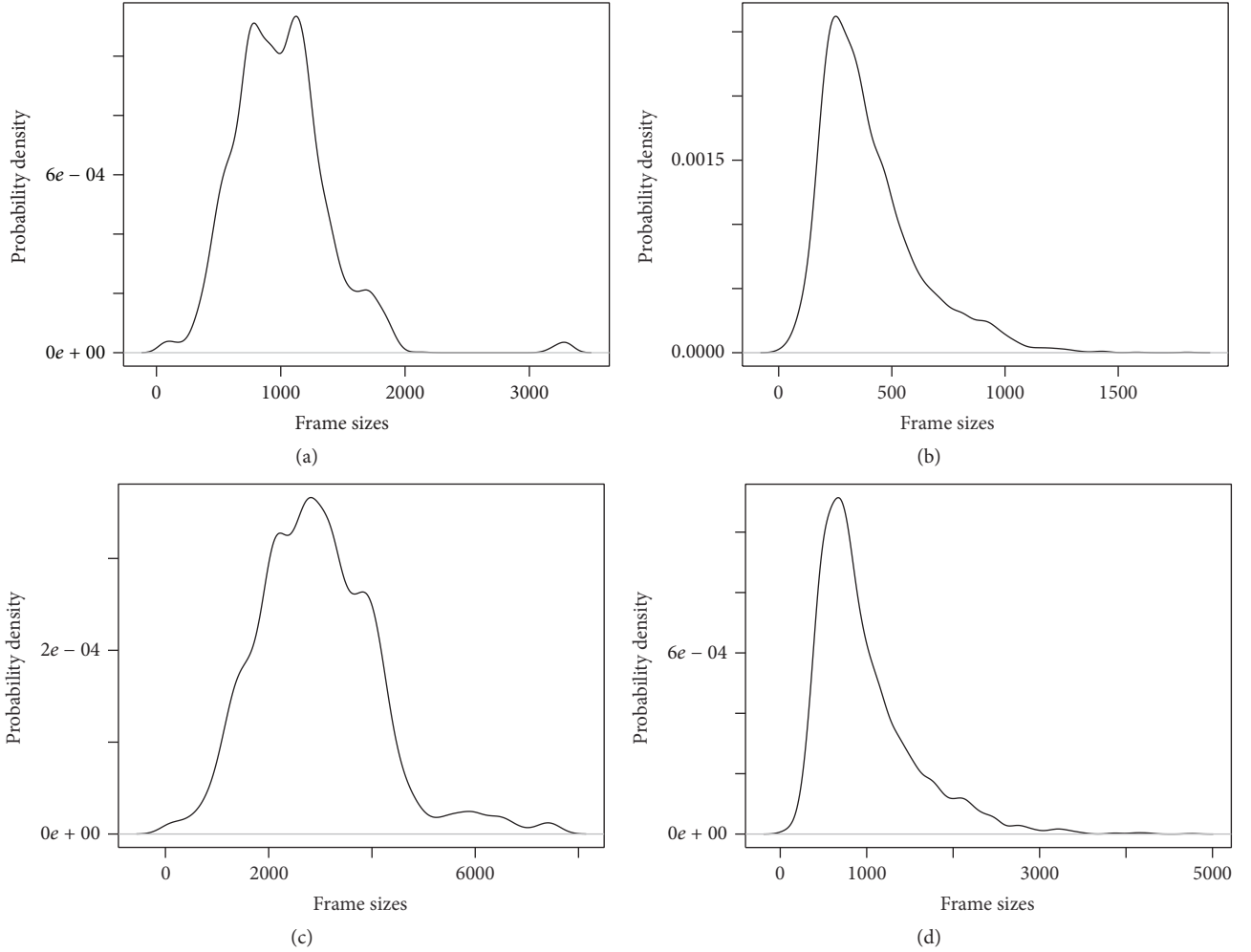


FIGURE 1: (a) The probability density function of I frame, in base layer of NBC News. (b) The probability density functions of B frame, in base layer of NBC News. (c) The probability density function of I frame, in enhancement layer of NBC News. (d) The probability density function of I frame, in enhancement layer of NBC News.

A complete solution of (7) is delivered in Appendix B and the following results have been obtained at the end:

$$k = \frac{1}{b_2(a_1 - a_2)} \Rightarrow p(x) \propto (x - a_1)^{-k(a_1 - a)} (x - a_2)^{k(a_2 - a)}. \quad (9)$$

The density is only proportional to this mathematical phase, 4 types of the Pearson distribution can be obtained as follows.

(I) *Pearson Type 1 Distribution.* The Pearson type 1 distribution occurs when the denominator of (7) includes opposite sign roots, $a_1 < 0 < a_2$. For more detailed proof please refer to the Appendix C.

$$f(x) = \frac{\Gamma(m_1 + m_2 + 2)}{\Gamma(m_1 + 4) \Gamma(m_1 + 1)} \left(\frac{x - \lambda}{a_2 - a_1} \right)^{m_1} \cdot \left(1 - \frac{x - \lambda}{a_2 - a_1} \right)^{m_2}. \quad (10)$$

(II) *The Pearson Type 2 Distribution.* This distribution is a specified case of Beta distribution which is symmetric [20]. The probability density function is as follows:

$$f(x) = \frac{\Gamma(2(m_1 + 1))}{\Gamma(m_1 + 1)^2} \left(\frac{x - \lambda}{a_2 - a_1} \left[1 - \frac{x - \lambda}{a_2 - a_1} \right]^{m_1} \right). \quad (11)$$

(III) *The Pearson Type 3 Distribution.* If scale parameter is allowed to be negative, negative skewness, the Pearson type 3 distribution is obtained as follows [21]:

$$f(x) = \frac{1}{|a_2 - a_1|^{m_1 + 1} \Gamma(m_1 + 1)} |x - \lambda|^{m_1} e^{-(x - \lambda)/(a_2 - a_1)}. \quad (12)$$

(IV) *The Pearson Type 5 Distribution.* Finally, the Pearson type IV distribution is inverse gamma distribution in which

again negative scale parameter is allowed [22]. So, the probability density function is as follows:

$$f(x) = \frac{1}{|a_2 - a_1|^{m_1+1} \Gamma(m_1 + 1)} |x - \lambda|^{-m_1-2} e^{-(a_2-a_1)/(x-\lambda)}. \quad (13)$$

(B) *Case 2: Negative Discriminate-Pearson Type 4 Distribution.* In this case the discriminant of (7) is negative ($b_1^2 - 4b_2b_0 < 0$). The denominator of (7) contains no roots. A detailed proof of Pearson type 4 is given in Appendix D.

$$p(x) = \frac{|\Gamma(m + (v/2)i) / \Gamma(m)|^2}{\alpha \beta (m - 1/2, 1/2)} \left[1 + \left(\frac{x - \lambda}{\alpha} \right)^2 \right]^{-m} \cdot \exp \left[-v \arctan g \left(\frac{x - \lambda}{\alpha} \right) \right]. \quad (14)$$

This is the Pearson type IV distribution [23].

Eventually, there are five types of Pearson distribution based on positive or negative of discriminant cases of quadratic function in the basic $p(x)$ in this distribution. Furthermore, many well-known distributions are special case of these types of the Pearson distributions. Table 1 provides the summary of distributions.

After describing the different distributions, the parameter estimation is the second step of methodology which will be addressed in the next subsection.

4.2. Estimating the Parameters. In statistics, there are three general estimation methods including the following:

- (i) Maximum Likelihood estimation method [24].
- (ii) Least square estimation method.
- (iii) Moment estimation method.

In most statistical analyses, maximum likelihood estimation (MLE) method is used to estimate parameters because it produces unbiased, consistent, and efficient estimates as it is stated by Myung [25]. Based on Kelton and Law [26] the MLE of a distribution are the parameters of that function that maximize the likelihood of the distribution given a set of observational data. Given a set of observational data x and a probability density function (PDF) f the likelihood function is

$$l = \prod_{i=1}^n f(x_i, \text{parameters}). \quad (15)$$

MLE tends to determine the values of the parameters that maximize the function l . Therefore, this paper is focused MLE method as well for parameter estimation.

4.3. Find the Best Distribution. The final step of the methodology is finding the best distribution for the data set. In statistics, there are two common approaches to fit the best distribution to the data: graphical methods and hypothesis testing which will be explained in next subsections.

4.3.1. Graphical Methods. Graphical method is a way to represent the results from the fitting process. It visually, evaluates how well a distribution matches with the input data. The most widely used and powerful goodness of fit test is quantile-quantile (Q-Q). Q-Q plot is used to calculate the quantile values of two probability distributions and plot them against each other. A point (x, y) on the plot is related to one of the quantiles of the data (x -axes) plotted against the quantiles of the considered distribution (y -axes). Therefore, if the fit is good, the points of the plot will lie approximately along the line $y = x$ (45° reference line).

4.3.2. Hypothesis Test. Generally, in statistic, there are three different statistical hypotheses [27] known as follows:

- (a) Comparing two models which are placed in H_0 and H_1 .
- (b) Testing parameters of a hypothesized model which are put in H_0 and H_1 .
- (c) Comparing two different distributions which are stated in H_0 and H_1 .

The last one is the goal of this study. To test all of these hypotheses, the appropriate statistic, a variable with a specified distribution, needs to be established. One popular statistic to test the considered hypothesis in this study is the Kolmogorov-Smirnov (K-S) statistic which is constructed based on the cumulative distribution functions. K-S test is a nonparametric and distribution free test [28]. The K-S test applies the maximum vertical deviation among the two plots and can be identified in the manner similar to the ones described in [16].

$$D = \sup_x \{|F(x) - G(x)|\}, \quad (16)$$

where $F(x)$ is an empirical distribution function of the real video traffic and $G(x)$ is the cumulative distribution function of the model.

The small values of D lead to accepting the null hypothesis in this study:

$$H_0: F(x) = G(x) \quad \text{versus} \quad H_1: F(x) \neq G(x). \quad (17)$$

The main result needs to be obtained based on p value. If p value is less than 0.05, usual significant level, the null hypothesis, H_0 , is rejected. The distribution $F(x)$ is a suitable probability function for the data if the p value is larger than 0.05.

4.4. Autocorrelation Function. Capturing the ACF structure of VBR video traffic is actually more challenging because of the fact that VBR traces exhibit both Long-Range Dependent (LRD) and short-range dependent (SRD) properties. If the autocorrelation function decays exponentially fast it can be referred to as an SRD process, but if it decays slowly, then the source is referred to as an LRD process.

As it can be seen from the Figure 2 the autocorrelation remains high even for large numbers of lags and it decay very slowly; both these facts are a clear indication of the

TABLE 1: The special cases of Pearson types.

Distribution	Form of distribution	Type
Beta distribution	$f(x; \alpha, \beta) = \frac{1}{B(\alpha, \beta)} x^{\alpha-1} (1-x)^{\beta-1}$	1
Cauchy distribution	$f(x; x_0, \gamma) = \frac{1}{n} \left\{ \frac{\gamma}{(x-x_0)^2 + \gamma} \right\}$	4
Chi-square distribution	$f(x; k) = \frac{x^{(k/2)-1} e^{-x/2}}{2^{k/2} \Gamma(k/2)}$	3
Continuous uniform distribution	$f(x) = \frac{1}{b-a} \quad a \leq x \leq b$	1
Gamma distribution	$f(x; k, \theta) = \frac{1}{\theta^k} \frac{1}{\Gamma(k)} x^{k-1} e^{-x/\theta}$	3
Inverse-chi-squared distribution	$f(x; k) = \frac{2^{-k/2}}{\Gamma(k/2)} x^{-k/2-1} e^{-1/2x}$	5
Inverse Gamma distribution	$f(x; \alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{-\alpha-1} e^{-\beta/x}$	5

importance of LRD. So, strong autocorrelation coefficients are found due to the periodic repetition of I, B, and P frames, and the autocorrelation function has a very slow exponential decrease.

5. Result and Discussion

As mentioned, in this paper, SVC video trace includes two layers called base layer and enhancement layer; each layer consists of I frames and B frames. The video traces that are considered here belong to NBC News, Silence of the Lambs, and Sony Demo with low or moderate scene changes. To fit the different distributions and choose the best one, the data are analyzed in R software [29], statistical software, which provides the K-S statistics, its p value, and the Q-Q plots. For the first step the graphical evaluation of the distributions using Q-Q plot and K-S test was performed.

5.1. Q-Q Plot. All distributions being studied should be fitted to the data to determine the closest distribution to the actual data. To do so, the Q-Q plot was employed to detect the closer graph to the reference line, which is the best fit.

As it can be seen from Figure 3, the exponential distribution is further from the reference line compared to the other distributions in both figures. The Pearson curve is quite nearer to the reference line in comparison with other curves graphically. Although the tails are quite far in bigger size of frames, Pearson distribution can be considered as the best fit. In other words, the Pearson distribution type IV is the best distribution for NBC News in all cases which is one of the four-parameter distributions describing the data with high accuracy degree.

5.2. K-S Test. In order to validate the previous results, another test, known as K-S, is performed on all of the examined data.

This test is able to determine if two datasets are significantly different. One of the advantages of K-S test is that it makes no assumption about the distribution of the data [16]. The closer the graph is to the reference line, the less it was subjected to the K-S test.

Figure 4 showed that the K-S graph belonging to the Pearson distribution was the closest to the reference line. This was confirmed by the value of K-S test, which was 0.0502 for the NBC News of base layer and 0.0294 for the NBC News of the enhancement layer.

5.3. Statistical Evaluation. In order to provide the statistical confirmation to the visual finding from the earlier section, a statistical test has been performed. The lower K-S value causes the better fit. This can be verified by p value in a way that the highest p value belongs to the better fit.

Table 2 indicates that the best distribution, which is fitted to the data, is the Pearson distribution type IV for the NBC News base layer and the enhancement layer (I and B frames). These findings were derived from the K-S tests and the p values. As for base layer of the I frames in the NBC News, the K-S test and p value related to the Pearson distribution type IV are 0.0502, which is the lowest, and 0.1105, which is the highest among the other distributions. In the same vein, the corresponding values for the B frames are 0.0223 and 0.2214, respectively. Although the p value exceeds 0.05 (significant level for testing), this value is better than the others, making it suitable for the best distribution in the context of this sequence of frame sizes.

For the enhancement layer of I frames in the NBC News, the lowest K-S test, which was 0.0332, and the highest p -value, which was 0.0474, belong to Pearson type IV. Similarly, for the B frames, those values are 0.0204 and 0.2824, respectively, which proves that the Pearson distribution type IV is the best. A similar form of analysis was done for both Silence of the

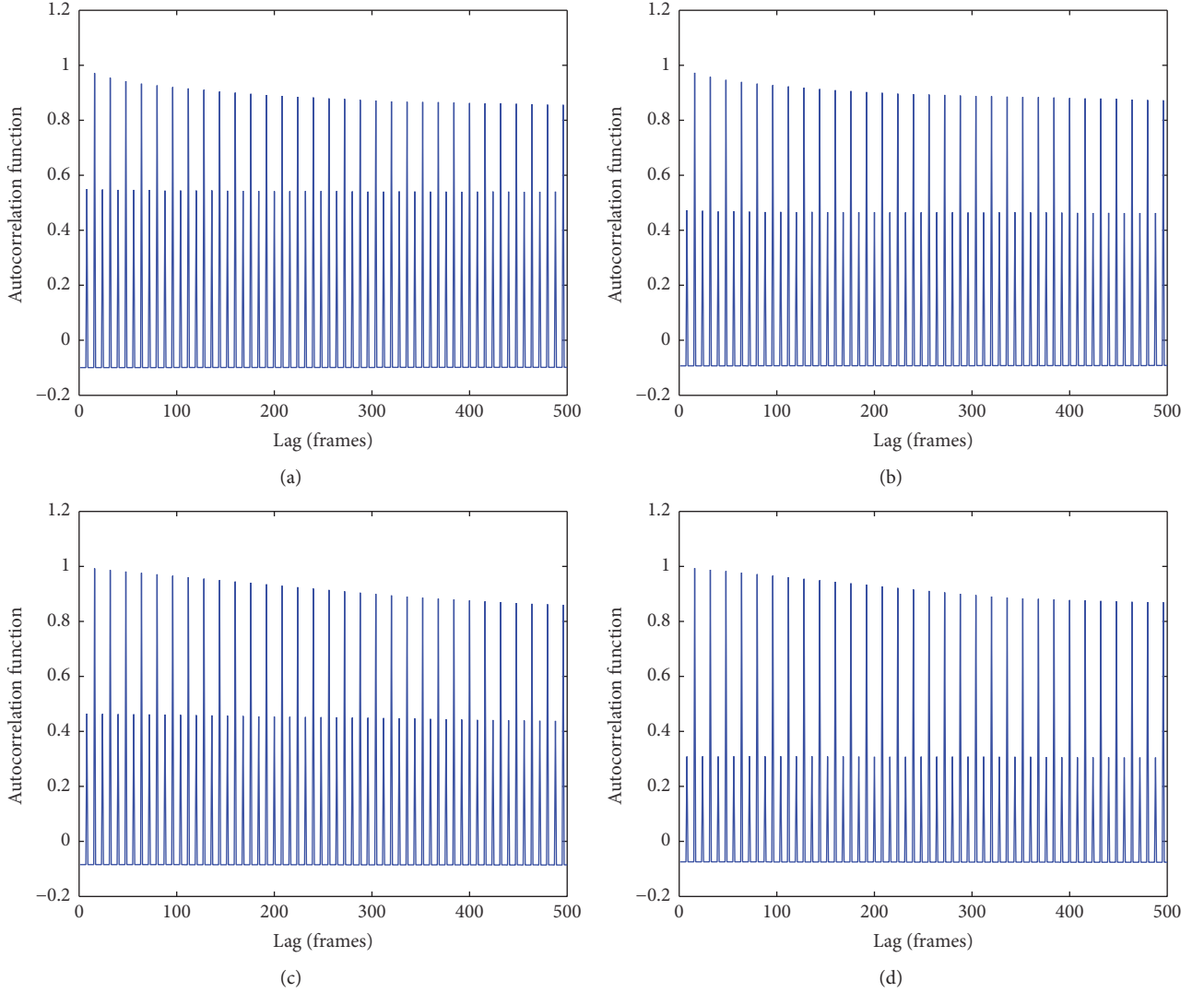


FIGURE 2: (a) and (b) ACF of NBC NEWS base layer and enhancement layer and (c) and (d) ACF of Sony Demo base layer and enhancement layer.

Lambs and Sony Demo in both I and B frames of the base layer and enhancement layer and again, due to the presence of space limitation, the results are not shown. The results show that for all of the p values, which are effectively less than 0.05, none of the distributions are well fitted to these data, and, however, the lowest K-S test value or the maximum p value can be chosen for an appropriate distribution for the data. Among those distributions with the skewness mentioned above, the Pearson distribution resulted in the best fitting for all video traces in the study (for both I frames and B frames).

6. Conclusion

In this study, frame types and the corresponding layers of SVC video traces are considered as data set for which the different distributions are investigated. Eventually, the density function of frame sizes has right skewness tail. It is also true for other video traces under the study, although the limitation

of space does not allow showing all of them here. As it was discussed, the Pearson distributions are appropriate to fit this data set. To achieve this finding, the analysis was started with plotting the density function of data to get some information on the data. Then by applying the Q-Q plot and K-S test, each of distribution was fitted to the data. Finally, based on the k-s statistic and the p value, the best distribution was chosen which was shown by the Q-Q plot too. Although there have been some studies in which the Pearson distributions were chosen as the best distribution, this result was for H.264 and MPEG encoded video data. Except the B frame trace of base layer, it is worth noting that the findings of this study result in the Pearson distributions as the best ones related to SVC video data for the first time and this is the first contribution of this paper. This result has been used to provide a traffic model in [30] by same author.

To the best of our knowledge, none of the existing researches explains all of the relevant distributions details in

TABLE 2: The statistical evaluation for the distributions of NBC News.

Movie	Distributions	Parameters	K-S statistic	p value	Best fit
I frame in base layer	<i>Pearson IV</i>	$m = 53.5727, \mu = -186.5132, location = -2052.722, scale = 1714.155$	0.0292	0.1105	Pearson IV
	Weibull	Shape = 3.1016, scale = 1100.322	0.0335	0.0014	
	Gamma	Shape = 7.1376, Rate = 0.0072	0.0433	2.20E - 16	
	Exponential	Rate = 1.0157e - 03	0.3341	2.20E - 16	
	Lognormal	Location = 6.8204, shape = 0.4116	0.0628	7.71e - 11	
	Normal	Location = 984.4909, scale = 342.2462	0.5146	0.0049	
B frame in base layer	Logistic	Location = 979.0349, scale = 195.6807	0.0547	0.0013	
	<i>Pearson IV</i>	$m = 3.4146, \mu = -11.9247, location = -20.0412, scale = 171.6336$	0.0252	0.2214	
	Weibull	Shape = 1.9873, scale = 456.1013	0.0327	1.47E - 14	
	Gamma	Shape = 3.9226, Rate = 0.0097	0.0679	2.20E - 16	
	Exponential	Rate = 2.4851E - 03	0.3176	2.20E - 16	
	Lognormal	Location = 5.8645, shape = 0.5268	0.0946	0.0668	
	Normal	Location = 984.4909, scale = 342.2462	0.5027	2.20E - 16	
	Logistic	Location = 374.6126, scale = 114.4575	0.0865	2.20E - 16	

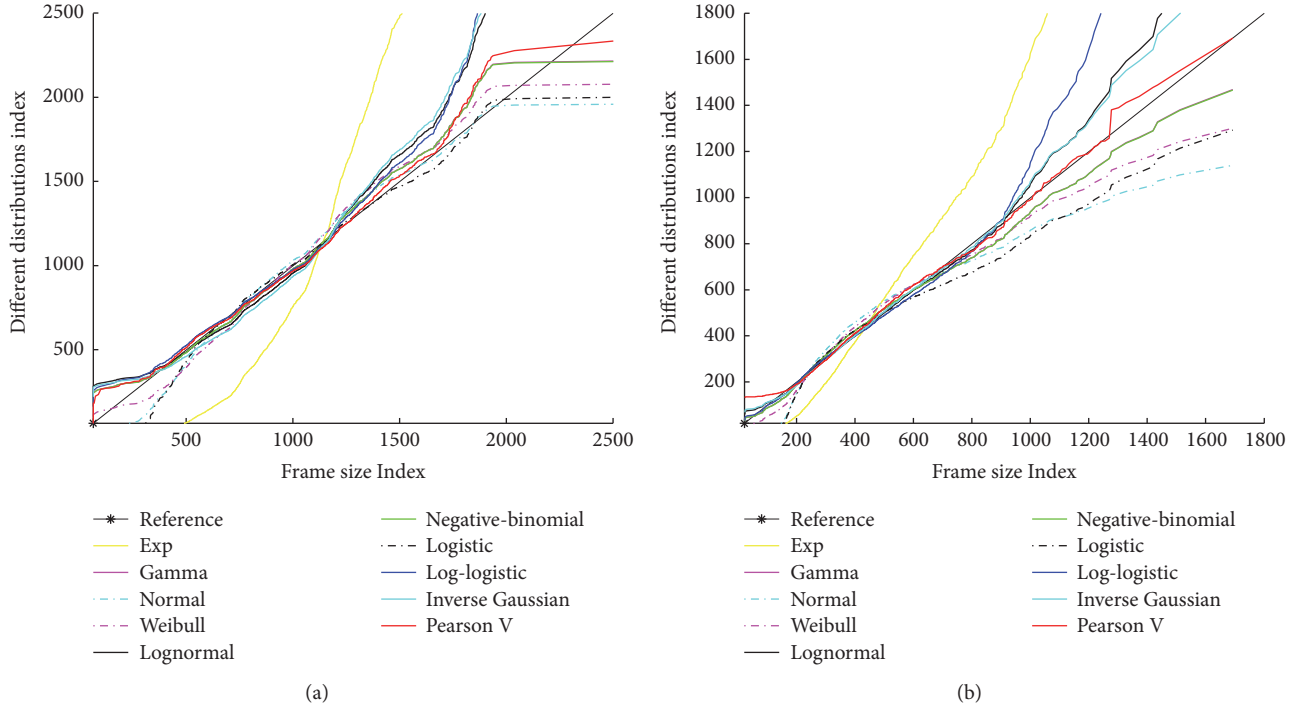


FIGURE 3: (a) The Q-Q plot for I frame, base layer of NBC News. (b) The Q-Q plot for B frame, base layer of NBC News.

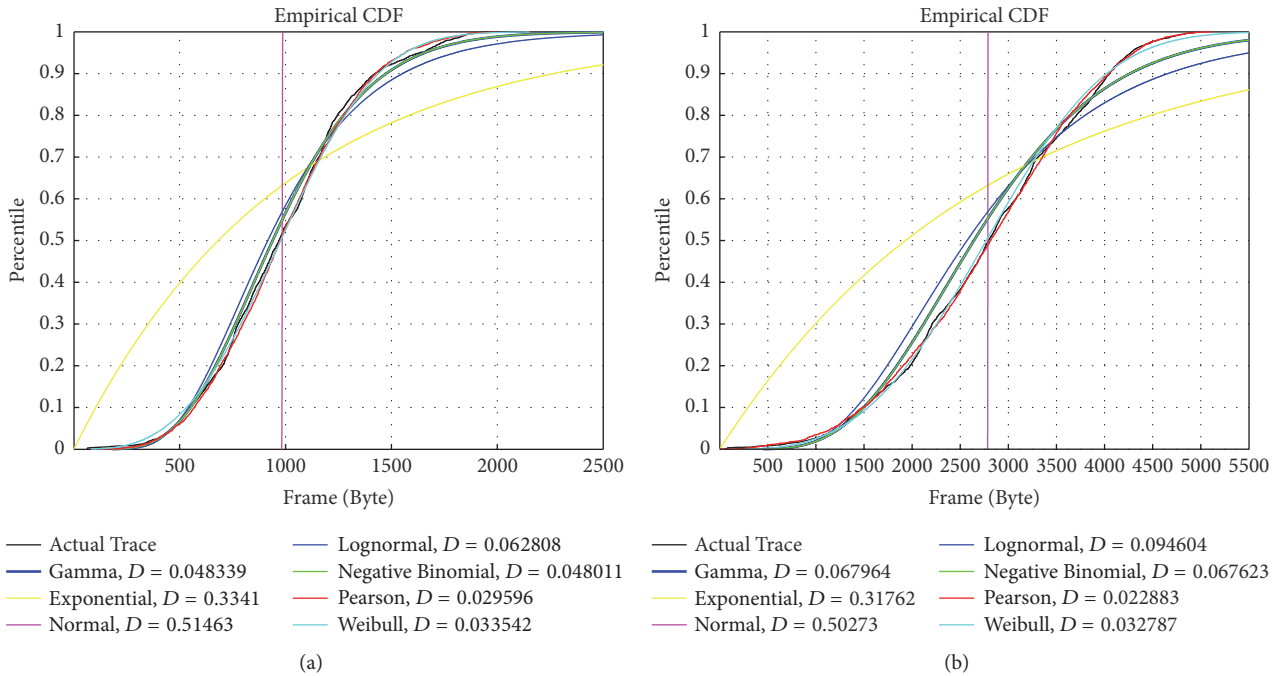


FIGURE 4: (a) The K-S test of the frame size of I frames in the NBC News base layer and (b) the K-S test of the frame size of I frames in the NBC News enhancement layer.

terms of concepts of parameters and theory. Furthermore, the work presented in this paper can be considered as an applicable collection of statistical distributions which are applied in the computer science and it can be regarded as the second contribution of this paper. Nevertheless, to

find the best distribution for some data set under specified situations in which none of these distributions cannot be chosen regarding all of statistical criteria, applying some conversion or the mixture distribution for the data will be done as future works.

Appendix

A. Proof of Equation (5)

A simple linear order differential equation is described which has the following general form:

$$Dy(x) + f(x)y(x) = g(x), \quad (\text{Equation (5)}). \quad (\text{A.1})$$

Both sides of this equation are multiplied by $e^{\int f(x)dx}$ so we have

$$\begin{aligned} Dy(x)e^{\int f(x)dx} + f(x)y(x)e^{\int f(x)dx} \\ = g(x)e^{\int f(x)dx}. \end{aligned} \quad (\text{A.2})$$

Based on the product rule in differential analysis, it equals to

$$D\left(y(x)e^{\int f(x)dx}\right) = g(x)e^{\int f(x)dx}; \quad (\text{A.3})$$

then

$$\begin{aligned} y(x)e^{\int f(x)dx} &= \int g(x)e^{\int f(x)dx}dx + c \\ y(x) &= \frac{\int g(x)e^{\int f(x)dx}dx + c}{e^{\int f(x)dx}}. \end{aligned} \quad (\text{A.4})$$

The assumption (7) comes from the comparison between (6) and

$$y(x) = e^{-\int f(x)dx} \left(\int g(x)e^{\int f(x)dx}dx + k \right). \quad (\text{A.5})$$

B. The Solution of Equation (7)

This appendix covers the solution of (7) in more detail. In case of positive discrimination, there are two real roots like a_1 and a_2 . Therefore, the quadratic function can be rewritten as follows:

$$f(x) = b_2(x - a_1)(x - a_2). \quad (\text{B.1})$$

Rewriting (7) as

$$p(x) \propto \exp \left\{ -\frac{1}{b_2} \int \frac{x - a}{(x - a_1)(x - a_2)} dx \right\} \quad (\text{B.2})$$

then the solution result of the integral is as follows:

$$\begin{aligned} &\int \frac{x - a}{(x - a_1)(x - a_2)} dx \\ &= \int \frac{a_1 - a}{(x - a_1)(a_1 - a_2)} dx \\ &\quad + \int \frac{a - a_2}{(x - a_2)(a_1 - a_2)} dx \\ &= \frac{(a_1 - a) \ln(x - a_1)}{(a_1 - a_2)} + \frac{(a - a_2) \ln(x - a_2)}{(a_1 - a_2)} + c. \end{aligned} \quad (\text{B.3})$$

Considering $k = 1/b_2(a_1 - a_2)$, (7) can be rewritten as $p(x) \propto (x - a_1)^{-k(a_1 - a)}(x - a_2)^{k(a_2 - a)}$.

C. The Proof of Pearson Distribution Type 1

Appendix C presents the proof of Pearson type 1 for better understanding to the readers. To find the density function of X , it can be written as a linear function of Y and its density function is in [31].

$$X = a_1 + y(a_2 - a_1) \quad \text{where } 0 < y < 1$$

$$p(y) \propto \left(\frac{a_1 - a_2}{a_1} y \right)^{v(a_1 - a)} \left(\frac{a_2 - a_1}{a_1} (1 - y) \right)^{v(a_2 - a)} \quad (\text{C.1})$$

$$p(y) \propto y^{m_1} (1 - y)^{m_2},$$

where $m_1 = (a - a_1)/b_2(a_1 - a_2)$ and $m_2 = (a - a_2)/b_2(a_1 - a_2)$ indicate the shape parameter

$$\frac{X - \lambda}{a_2 - a_1} \sim \beta(m_1 + 1, m_2 + 1), \quad (\text{C.2})$$

where $\lambda = \mu_1 - (a_2 - a_1)((m_1 + 1)/(m_1 + m_2 + 2))a_1$ is a location parameter and $(a_2 - a_1)$ is scale parameter.

Pearson type 1: $f(x)$

$$\begin{aligned} &= \frac{\Gamma(m_1 + m_2 + 2)}{\Gamma(m_1 + 4)\Gamma(m_1 + 1)} \left(\frac{x - \lambda}{a_2 - a_1} \right)^{m_1} \\ &\cdot \left(1 - \frac{x - \lambda}{a_2 - a_1} \right)^{m_2}. \end{aligned} \quad (\text{C.3})$$

D. The Proof of Pearson Distribution Type 4

In this case, there are no roots but by defining a new variable as below, the Pearson distribution type 4 is produced and $p(x)$ can be obtained.

$$\begin{aligned} y &= x + \frac{b_1}{2b_2}, \\ \alpha &= \frac{\sqrt{4b_2b_0 - b_1^2}}{2b_2}. \end{aligned} \quad (\text{D.1})$$

So if

$$h(x) = b_2x^2 + b_1x + b_0 \quad (\text{D.2})$$

$$\text{then } h(x) = b_2(y^2 + \alpha^2),$$

$$p(y) \propto \exp \left(-\frac{1}{b_2} \int \frac{y - b_1/2b_2 - a}{y^2 + \alpha^2} dy \right) \quad (\text{D.3})$$

on the other hand, the integral can be solved as follows:

$$\begin{aligned} \int \frac{y - b_1/2b_2 - a}{y^2 + \alpha^2} dy &= \frac{1}{2} \ln(y^2 + \alpha^2) \\ &\quad - \frac{2b_2a + b_1}{2b_2\alpha} \arctan g\left(\frac{y}{\alpha}\right) \\ &\quad + c_0; \end{aligned} \quad (\text{D.4})$$

therefore

$$p(y) \propto \left[1 + \frac{y^2}{\alpha^2}\right]^{-m} \exp\left[-v \arctan g\left(\frac{y}{\alpha}\right)\right], \quad (\text{D.5})$$

where

$$\begin{aligned} m &= \frac{1}{2b_2}, \\ v &= -\frac{2b_2a + b_1}{2b_2\alpha}, \end{aligned} \quad (\text{D.6})$$

As it was mentioned, λ was ignored in the first equation for simplicity but here it is included to the equation

$$\begin{aligned} &\text{Pearson type 4 (Equation (14)): } p(x) \\ &= \frac{|\Gamma(m + (v/2)i) / \Gamma(m)|^2}{\alpha\beta(m - 1/2, 1/2)} \left[1 + \left(\frac{x - \lambda}{\alpha}\right)^2\right]^{-m} \\ &\cdot \exp\left[-v \arctan g\left(\frac{x - \lambda}{\alpha}\right)\right]. \end{aligned} \quad (\text{D.7})$$

Competing Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

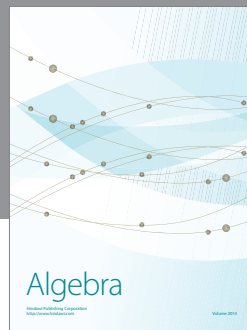
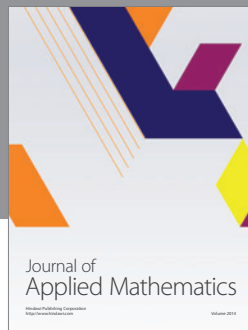
Acknowledgments

This work was supported in part by Universiti Sains Malaysia through Grant no. 304/PNAV/650817/C127.

References

- [1] R. Gupta, A. Pulipaka, P. Seeling, L. J. Karam, and M. Reisslein, "H.264 coarse grain scalable (CGS) and medium grain scalable (MGS) encoded video: a trace based traffic and quality evaluation," *IEEE Transactions on Broadcasting*, vol. 58, no. 3, pp. 428–439, 2012.
- [2] S. Tanwir and H. Perros, "A survey of VBR video traffic models," *IEEE Communications Surveys and Tutorials*, vol. 15, no. 4, pp. 1778–1802, 2013.
- [3] H. Koumaras, C. Skianis, G. Gardikis, and A. Kourtis, "Analysis of H.264 video encoded traffic," in *Proceedings of the 5th International Network Conference (INC '05)*, pp. 441–448, Samos, Greece, July 2005.
- [4] C.-M. Lin, J. K. Zao, W.-H. Peng, C.-C. Hu, H.-M. Chen, and C.-K. Yang, "Bandwidth efficient video streaming based upon multipath SVC multicasting," in *Proceedings of the International Wireless Communications and Mobile Computing Conference (IWCMC '08)*, pp. 406–412, Crete Island, Greece, August 2008.
- [5] H. Schwarz, D. Marpe, and T. Wiegand, "Overview of the scalable video coding extension of the H.264/AVC standard," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 17, no. 9, pp. 1103–1120, 2007.
- [6] I. Unanue, I. Urteaga, R. Husemann et al., "A tutorial on H.264/SVC scalable video coding and its tradeoff between quality, coding efficiency and performance," in *Recent Advances on Video Coding*, J. Del Ser Lorente, Ed., chapter 1, InTech, Rijeka, Croatia, 2011.
- [7] Z. Shahid, M. Chaumont, and W. Puech, "Scalable video coding," in *Effective Video Coding for Multimedia Applications*, InTech, 2011.
- [8] A. Feldmann and W. Whitt, "Fitting mixtures of exponentials to long tail distributions to analyze network performance molds," in *Proceedings of the 16th Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM '97)*, Kobe, Japan, April 1997.
- [9] J. Abate, G. L. Choudhury, and W. Whitt, "Waiting-time tail probabilities in queues with long-tail service-time distributions," *Queueing Systems*, vol. 16, no. 3, pp. 311–338, 1994.
- [10] D. P. Heyman, A. Tabatabai, and T. V. Lakshman, "Statistical analysis and simulation study of video teleconference traffic in ATM networks," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 2, no. 1, pp. 49–59, 1992.
- [11] D. Heyman, T. V. Lakshman, A. Tabatabai, and H. Heeke, "Modeling teleconference traffic from VBR video coders," in *Proceedings of the IEEE International Conference on Communications (ICC '94), Conference Record, 'Serving Humanity through Communications' (SUPERCOMM/ICC '94)*, vol. 3, pp. 1744–1748, May 1994.
- [12] S. Xu and Z. Huang, "A gamma autoregressive video model on ATM networks," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 8, no. 2, pp. 138–142, 1998.
- [13] M. Krunz and H. Hughes, "A traffic for MPEG-coded VBR streams," *ACM SIGMETRICS Performance Evaluation Review*, vol. 23, no. 1, pp. 47–55, 1995.
- [14] F. H. P. Fitzek and M. Reisslein, "MPEG-4 and H.263 video traces for network performance evaluation," *IEEE Network*, vol. 15, no. 6, pp. 40–54, 2001.
- [15] W.-C. Poon and K.-T. Lo, "A refined version of M/G/ ∞ processes for modelling VBR video traffic," *Computer Communications*, vol. 24, no. 11, pp. 1105–1114, 2001.
- [16] A. Lazaris, P. Koutsakis, and M. Paterakis, "A new model for video traffic originating from multiplexed MPEG-4 videoconference streams," *Performance Evaluation*, vol. 65, no. 1, pp. 51–70, 2008.
- [17] D. M. B. Masi, M. J. Fischer, and D. A. Garbin, "Video frame size distribution analysis," *The Telecommunications Review*, vol. 19, pp. 74–86, 2008.
- [18] K. Salah, F. Al-Haidari, M. H. Omar, and A. Chaudhry, "Statistical analysis of H.264 video frame size distribution," *IET Communications*, vol. 5, no. 14, pp. 1978–1986, 2011.
- [19] Video Trace Library, <http://trace.eas.asu.edu/>.
- [20] J. A. D. Garcia and R. Gutierrez-Jaimez, "Matricvariate and matrix multivariate Pearson type II distributions," <https://arxiv.org/abs/1011.5083v1>.
- [21] A. C. Olshen, "Transformations of the pearson type III distribution," *The Annals of Mathematical Statistics*, vol. 9, no. 3, pp. 176–200, 1938.
- [22] W. P. Elderton and N. L. Johnson, *Systems of frequency curves*, Cambridge University Press, London, UK, 1969.
- [23] J. Heinrich, *A Guide to the Pearson Type IV Distribution*, 2004, <http://www-cdf.fnal.gov/physics/statistics/notes/cdf6820-pearson4.pdf>.
- [24] Y. Nagahara, "The PDF and CF of Pearson type IV distributions and the ML estimation of the parameters," *Statistics and Probability Letters*, vol. 43, no. 3, pp. 251–264, 1999.

- [25] I. J. Myung, "Tutorial on maximum likelihood estimation," *Journal of Mathematical Psychology*, vol. 47, no. 1, pp. 90–100, 2003.
- [26] W. D. Kelton and A. M. Law, *Simulation Modeling and Analysis*, McGraw-Hill, Boston, Mass, USA, 2000.
- [27] G. Casella and R. L. Berger, "Statistical inference," *Journal of Time Series Analysis*, vol. 9, 2001.
- [28] M. Asai and U. Dashzeveg, "A distribution-free test for symmetry with an application to S&P index returns," *Applied Economics Letters*, vol. 15, no. 6, pp. 461–464, 2008.
- [29] D. Stockwell, *Niche Modeling Predictions from Statistical Distributions*, Chapman & Hall/CRC Mathematical and Computational Biology Series, Chapman & Hall/CRC Taylor & Francis, Boca Raton, Fla, USA, 2007.
- [30] S. Ahmadpour, Z. Toghrayee, and T.-C. Wan, "An extended discrete autoregressive model for variable bit rate video traffic encoded by scalable video codec," *International Journal of Communication Systems*, vol. 28, no. 18, pp. 2239–2254, 2015.
- [31] K. Pearson, "Mathematical contributions to the theory of evolution.—X. Supplement to a memoir on skew variation," *Proceedings of the Royal Society of London*, vol. 68, no. 442–450, pp. 372–373, 1901.



Hindawi

Submit your manuscripts at
<https://www.hindawi.com>

