CCITT SGXV
Working Party XV/1
Experts Group for ATM Video Coding

Document AVC - 6
May 23, 1991

**Subject:** Statistical Analysis of Video Teleconference Traffic in ATM Networks

**Source:** Bellcore

**Purpose:** Discussion

## 1. MODELING ISSUES STUDIED AND CONCLUSIONS

For this contribution[1] [2] we analyzed a 30 minute sequence of video teleconference data in order to answer the following question: What statistical models characterize the data accurately and what models of video sources are accurate enough to be used in traffic studies? The traffic data that we analyzed consists of a sequence indicating the number of cells per frame for 48,500 frames of a teleconference and was obtained by recording the output of a VBR video coder during a 30 minute teleconference. We tested the accuracy of our synthetic traffic models by simulation and comparison with results obtained using the recorded traffic data. A key feature of our study is the use of a long sequence of "real" data. Most previous studies have used very short sequences (for instance, a few seconds) of data to address these issues. We believe that the use of such short sequences necessarily sacrifices accuracy.

Our major conclusions are: (1) Unlike some previous studies, the number of cells per frame for video teleconferences is not normally distributed. Instead, it follows a gamma (or negative binomial) distribution. Also, in the absence of scene cuts and scene changes the number of cells per frame is a stationary process. (2) For traffic studies, neither an autoregressive model of order 2 nor a two-state Markov chain model is good because they do not model correctly (either underestimate or overestimate) the occurrence of large values (of number of cells per frame) and these large values are a primary factor in determining cell-loss rates. The order 2 autoregressive model, however, fits the data well in a statistical sense. (3) A detailed Markov chain model (with about 60 states) is sufficiently accurate for use in traffic studies. We have been able to synthesize this model from just the peak rate, mean, variance, and first-order autrocorrelation coefficient of the traffic.

## 2. SUMMARY OF STATISTICAL ANALYSIS AND SIMULATION

The histogram of the data is shown in Fig. 1(a). In this figure, the dashed curve with the asterisks is a smoothed version of the empirical density function.

This histogram has the general shape of a negative binomial distribution, which is shown as the solid curve in Fig. 2(a). The probability function for the negative binomial distribution is given by

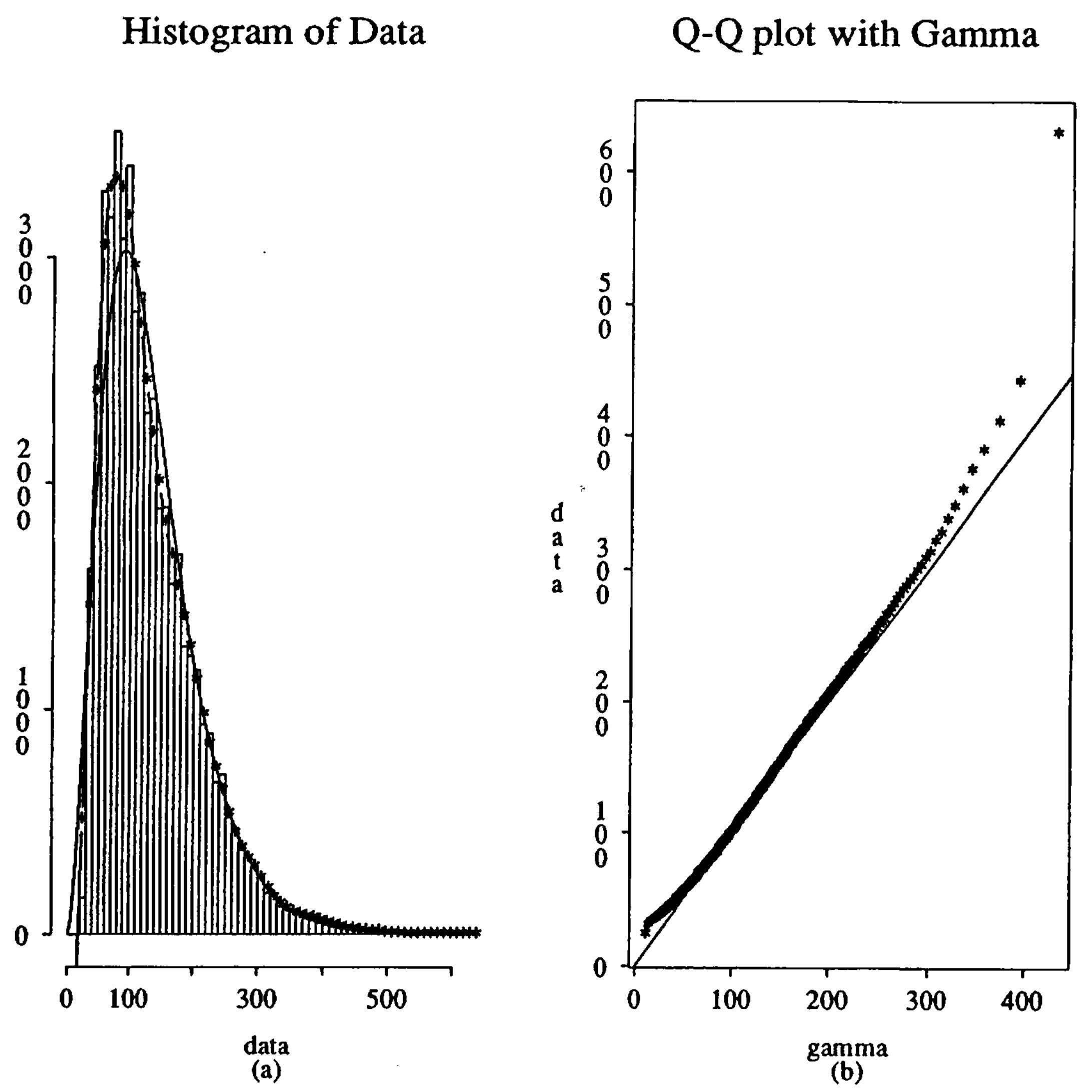


**Figure 1.** Histogram and Q-Q Plot of Data vs. Gamma Distribution

$$f_k = \binom{k+r-1}{k} p^r q^k = \binom{-r}{k} p^r (-q)^k ,$$

for $k=0,1,\ldots$ . Here, $0<p<1$, $q=1-p$, and $r>0$. The parameter $r$ need not be an integer. The mean and variance of this distribution are

$$m = \frac{r(1-p)}{p} \quad \text{and} \quad v = \frac{r(1-p)}{p^2} .$$

For our data, the mean and variance are $m=130.2967$ and $v=5536.873$; the method of moments produces the estimates

$$\hat{p} = 0.02353 \quad \text{and} \quad \hat{r} = 3.140 \tag{1}$$

A powerful goodness-of-fit test is the Q-Q plot, which plots the quantiles of the data vs. the quantiles of the fitted distribution. The Q-Q plot for our data and the gamma distribution with parameters given in (2) is shown in Fig. 1(b). The Q-Q plot shows that the fit is very good except for the right-hand tail, where the gamma distribution has too little probability. The * in the Northeast corner represents the 4 data points that are larger than 600. The data points larger than 400 constitute 0.60 per cent of the data and they account for the other noticeable departures from the straight line. These points can be classified as outliers from a statistical model, but they are important for traffic models, because cell losses are strongly influenced by the presence of frames with many cells.

A goodness-of-fit test based on moments also gives more evidence that the gamma distribution is a good fit to the data.

The empirical density function for the bit rate for 30 minutes of VBR coding data for a video conference shown in Fig. (4a) in Verbiest, Pinnoo, and Vosten[3] also appears to follow a gamma distribution. Those authors postulated a Normal distribution. To examine the hypothesis that the gamma distribution is a better fit to the data, we computed the critical bit rate $b^*$ say, with the property that the proportion of bit rates that are above $b^*$ is $10^{-7}$ from a fitted gamma distribution, and compared our results to those given in[3] Table 1.

**TABLE 1.** Comparisons among distributions

| number of multiplexes | 16 | 32 | 64 | 128 |
|---|---|---|---|---|
| | Peak load per source [mbit/s] | | | |
| measured distr. | 6.50 | 5.65 | 5.15 | 4.81 |
| Normal distr. | 5.93 | 5.42 | 5.04 | 4.77 |
| Gamma distr. | 6.30 | 5.63 | 5.16 | 4.84 |
| | per cent differences | | | |
| Normal distr. | 8.73 | 4.00 | 2.10 | 0.79 |
| Gamma distr. | 3.08 | 0.36 | -0.24 | -0.62 |

In Table 1 it is clear that the gamma distribution does better for this goodness of fit measure than does the Normal distribution.

## 2.1 A Markov Chain Model

A Markov chain model was chosen because it had the potential of producing clumps of large frames. One of the Markov chain models which has been proposed before is the two-state model. In the two-state model, the states are a low rate and the peak rate. These rates are measured in cells per frame; for our data they are 25 and 625 respectively. The transition probability from the low rate to the peak rate is p, and the transition probability from the peak rate to the low rate is q. One equation relating p and q is obtained by fixing the mean number of cells in a frame. A second equation is obtained by matching the correlation function. The autocorrelation function for a two-state Markov chain is an exponentially decreasing function, so there is only one parameter. In Fig. 2, the autocorrelation function for the real data appears to be an exponential function, and this was confirmed by more detailed plots. Our data give the estimates $\hat{p}$=0.0026 and $\hat{q}$=0.0123. The two-state model does not work well; it overestimates the loss probability by several orders of magnitude. This is because fitting the mean of the model to the mean of the data forces the variance of the model to be too large. This, in turn, leads to an excess of large frames.

Our Markov chain model is created as follows. Let $X_n$ be the number of cells in frame n, and $Y_n$ be the integer part of $X_n/10$. We propose to model $\{Y_n;\ n=1,2,\ldots,N\}$ as a Markov chain with transition matrix $P=(p_{ij})$. We estimate $(p_{ij})$ in the usual way[4]:

$$\hat{p}_{ij} = \frac{\text{number of transitions i to j}}{\text{number of transitions out of i}}$$

when the denominator is greater than zero. Since the smallest frame in our data has 25 cells, we set $\hat{p}_{i1}=1$ when there are no transitions out of state i; this does not affect the stationary distribution.

The autocorrelation function of the Markov chain is shown in Fig. 3, and it gives the best fit of all the models we tried. The steady-state distribution of the Markov chain is a good match to the distribution of the data, as shown in Fig. 2(a). The Q-Q plot of the distribution of a sample path of the Markov chain and the data is almost a straight line, as shown in Fig. 2(b). A simulated sample path of this model produced 401 frames with more than 400 cells, 51 of which had more than 500 cells, so there are enough (perhaps too many) large cells.

For each statistical model we generated ten realizations of 45,000 frames. These were used in the same way as the real data, as described above. The ten realizations were treated as independent and identically distributed samples, and approximate 95 per cent confidence intervals were computed by standard methods. The results are shown in the table below. The DAR(1) results shown in the table are for the case where the transition matrix for the Markov chain were computed from the four traffic parameter mentioned earlier.

**TABLE 2.** Confidence intervals for P{loss} for various source models

| Source | Probability of Loss × $10^6$ for various delays | | | | |
|---|---|---|---|---|---|
| | Maximum delay allowed (ms) | | | | |
| | 1 | 2 | 3 | 4 | 5 |
| Real Data | 2070.0 | 527.0 | 141.0 | 33.3 | 2.88 |
| Markov chain | 2280 ± 577 | 628. ± 215. | 185. ± 87.5 | 58.3 ± 39.1 | 21.2 ± 19.1 |
| DAR(1) | 2250 ± 512 | 604. ± 171. | 160. ± 52.6 | 35.0 ± 19.9 | 5.80 ±3.54 |
| AR(2) | 1501 ± 250 | 233 ± 71.0 | 39.3 ± 23.1 | 6.01 ± 5.64 | .134 ± .297 |
| Markov chain (two state) | 18404 ± 5681 | 18030 ± 5615 | 17926 ± 5593 | 17825 ± 5572 | 17727 ± 5551 |

In Table 2 we see that the confidence intervals from the Markov chain model surrounds the "true" values obtained from the real data, and those from the other model are far away from the "true" values. With the commonly chosen 5 per cent probability of rejecting a true hypothesis, the classical statistical tests of hypotheses would reject the hypotheses that the AR(2) or the two-state Markov chain models fit the real data, and not reject the hypothesis that the Markov chain fits the real data.

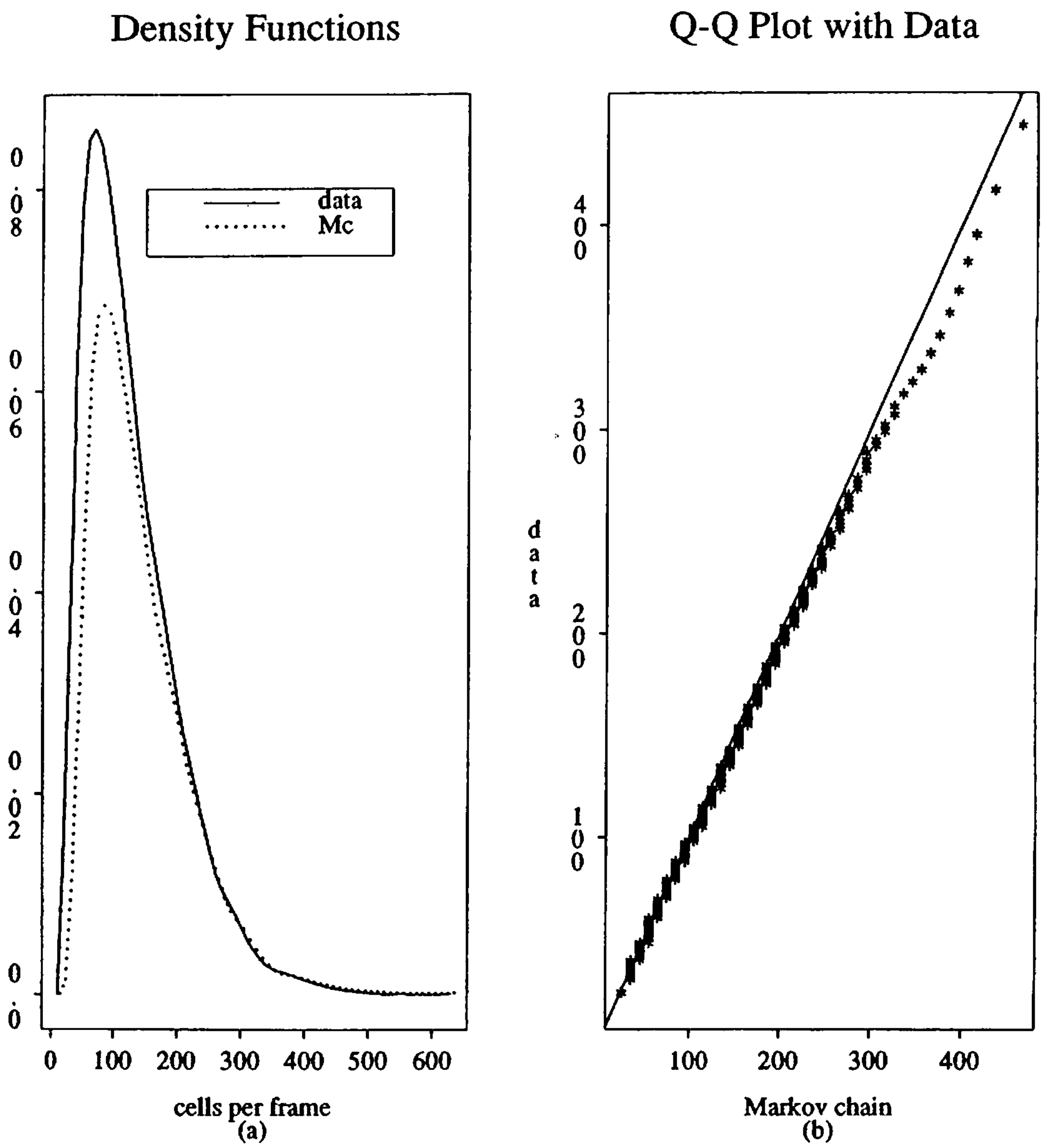


**Figure 2.** Density function and Q-Q plot of the Markov chain

*REFERENCES*


1. D. Heyman, A. Tabatabai and T. V. Lakshman, "Statistical Analysis and Simulation Study of VideoTeleconference Traffic in ATM Networks," submitted to IEEE Journal of Selected Areas in Communications.

2. D. Heyman, A. Tabatabai and T. V. Lakshman, "Analysis of VideoTeleconference Traffic in ATM Networks," submitted to GLOBECOM 91.

3. W. Verbiest, L. Pinnoo, and B. Voeten, "The Impact of the ATM Concept on Video Coding," *IEEE J. on Select. Areas Commun.*, Vol. SAC-6, No. 9, December 1988

4. P. Billingsley, *Statistical Inference for Markov Processes*, U. of Chicago Press, Chicago, 1961.