A CONNECTION ADMISSION CONTROL BASED ON REAL TIME TRAFFIC STATISTICS

Shusaburo Motoyama

Department of Telematics, Campinas State University, Campinas, Brazil

Email: motoyama@dt.fee.unicamp.br

David W. Petr and Victor S. Frost

Telecommunications and Information Sciences Laboratory

The University of Kansas, Lawrence, Kansas, USA

Email: dwp@ tisl.ukans.edu; frost@tisl.ukans.edu

Resumo: Um controle de admissão de conexões baseado em estatística de tráfego em tempo real é apresentado neste artigo. O método consiste em estimar a distribuição das fontes agregadas com base no histograma do número de células chegando num intervalo fixo, τ e é usado em conjunto com a distribuição de novas chamadas, estimada em tempo não-real por medição ou simulação. A nova distribuição de fontes agregadas é obtida usando a técnica da convolução e antes da aceitação de uma nova chamada. A razão de perda de células ("cell loss ratio", ou CLR) é avaliada usando esta nova distribuição, e é verificado se CLR satisfaz a qualidade de serviço ("quality of service", ou QoS) necessária. A viabilidade e a eficácia do método são demonstradas por simulação.

ABSTRACT A connection admission control based on real time traffic statistics is presented in this paper. The method consists of estimating the distribution of the aggregate sources based on the histogram of the number of cells arriving in a fixed interval τ and it is used in conjunction with the new call distribution estimated off-line by measurement or simulation. The new distribution of aggregate sources is obtained using the convolution technique and before the acceptance of a new call. The cell loss ratio (CLR) is evaluated by using this new distribution and it is examined if CLR satisfies required quality of service (QoS). The feasibility and effectiveness of the method are demonstrated by simulation.

1. INTRODUCTION

In Broadband Integrated Service Digital Network (B-ISDN) based on the Asynchronous Transfer Mode (ATM) technique a variety of signals such as video, voice, image, and data are statistically multiplexed and share common switching and transmission resources. Because these signals have different requirements, the bandwidth management and traffic control techniques developed for traditional circuit or packet switching networks are not applicable. For instance, the call admission control in circuit switched network is based on a fixed number of circuits or trunks, so that if the number of calls exceeds the number of trunks, the exceeded calls are blocked. In ATM networks the connection admission control is based on bandwidth allocation. Since each call needs different bandwidth, an elaborated bandwidth management technique is required.

Many connection admission controls have been proposed in the literature. The simplest way is to segregate the total bandwidth into classes of traffic and provide independent control for each class [1]. In this case, statistical gains are obtained only within a class.

Another proposal of the connection admission control is based on a set of precomputed maps as guidelines [2,3,4]. These maps are obtained off-line, through simulation or analysis for a predetermined offered load in the network. This type of admission control needs storage of the resource allocation guidelines in the switching nodes and does not accurately reflect the real-time traffic.

An approach based on equivalent capacity concept was introduced in [5]. This method is based on results of [6] and proposes analytical expressions for computing the equivalent capacity of a source or a number of multiplexed sources. The method conforms well when the source has exponentially distributed packet lengths and silence times but it has some limitations when the source does not resemble exponential distribution and there is correlation between packets (such as video source) [7].

A method based on a dynamic scheme that assigns the bandwidth on the basis of on-line measurement and estimation of an upper bound on the cell loss ratio was proposed in [8,9]. The upper bound on cell loss ratio is derived from the measured number of cells arriving at the network node during a fixed interval, and the source parameters specified by the user.

The connection admission control proposed in this paper is based on the cell loss probability formula presented in [8], but using a different approach to estimate the aggregate sources and new call distributions.

In Section 2 the proposed method is discussed in detail. The measurement considerations to estimate the aggregate sources and new call distributions are presented in Section 3. In Section 4 the feasibility of the method is shown using simulation and examples. Finally, conclusions are presented in Section 5.

Palavras Chaves: ATM, CAC, Real Time Control..

2. THE METHOD

A connection admission control based on distribution of the number of cells arriving during the fixed interval τ was presented in [8]. The method assumes that it is possible to estimate the probability of the number of cells arriving from connected sources. The estimate probability distribution is weighted with the measured frequency distribution in order to find the renewed process distribution, $\hat{p}(k)$. This renewed process distribution is used as the distribution of the connected sources.

Assuming the most bursty process distribution $\theta(k)$ for a new call with an average bit rate A_V and a peak bit rate R_p, the cell loss probability after the new call acceptance, estimated before the acceptance is given as [8]

$$\hat{B} = \frac{\sum_{k=0}^{\infty} [\,k - C\tau/L\,]^{+}\, \hat{p}(k) * \theta(k)}{\left(\sum_{k=0}^{\infty} k\hat{p}(k) + \tau A_V / L\right)} \tag{1}$$

where

C = Transmission link bandwidth

L = Cell length

$*$ denotes convolution and

$$[x]^{+} = \begin{cases} x, & \text{if } x > 0 \\ 0 & \text{otherwise} \end{cases};$$

If the cell loss probability $\hat{B}$ is less than a specified cell loss probability B, then the call is accepted; otherwise the call is rejected.

In [8] the $\hat{\ }(\ \ )$ is estimated using theoretical consideration and measured estimates.$\theta(k)$ is estimated as the most bursty process distribution. In this paper we use a different approach; we will use measured estimates for both $\hat{\ }(\ \ )$ and $\theta(k)$.

The idea is to get a distribution of the aggregate sources in the network by measurement in periodic time T and the distribution that gives worst cell loss probability is chosen by successive comparisons. The distribution of a new call requesting connection is estimated off-line by measurement or by simulation and it is stored in the switching nodes. The calls are divided into classes and each class has different distribution given by measured (or simulated) values.

By using these approaches we can calculate ˆ( )* θ(k) in Eq.1 by direct convolution or by transform method such as FFT (Fast Fourier Transform). The resultant distribution is used in Eq.1 to calculate the cell loss probability.

In the next section we discuss how the measured estimates of the distributions can be obtained.

3. MEASUREMENT CONSIDERATIONS

Let a time period T divided into N smaller intervals, each of length τ. Let $X_i(t)$ be a random process describing the number of cells arriving in i th interval of length τ. The process $X_i(t)$ is called the rate process.

The mean number of cell arrivals in the interval τ will be

$$M = \frac{\sum_{i=1}^{N} X_i(t)}{N} \tag{2}$$

The mean estimated cell arrival rate is given by

$$\hat{\lambda} = \frac{M}{\tau} = \frac{\sum_{i=1}^{N} X_i(t)}{N\tau} \tag{3}$$

The process X(t) can be used to study characteristics of the aggregate sources in the network. One of the characteristics that can be studied is the correlation of the aggregate sources [10]. For instance, Fig.3.1 shows the autocorrelation of the aggregate sources with 200 voice sources and one video source. Each voice source model used is such that a cell is generated at every 6 ms, representing a telephone traffic at a constant rate of 64 kb/s. The video source model generates fixed length video packets with exponentially distributed times. The number of bytes generated per packet is arranged so that exactly two cells are filled.

The autocorrelation is calculated by counting number of cells at every 50 ms, in a period of 10 seconds, and using standard discrete autocorrelation formula.

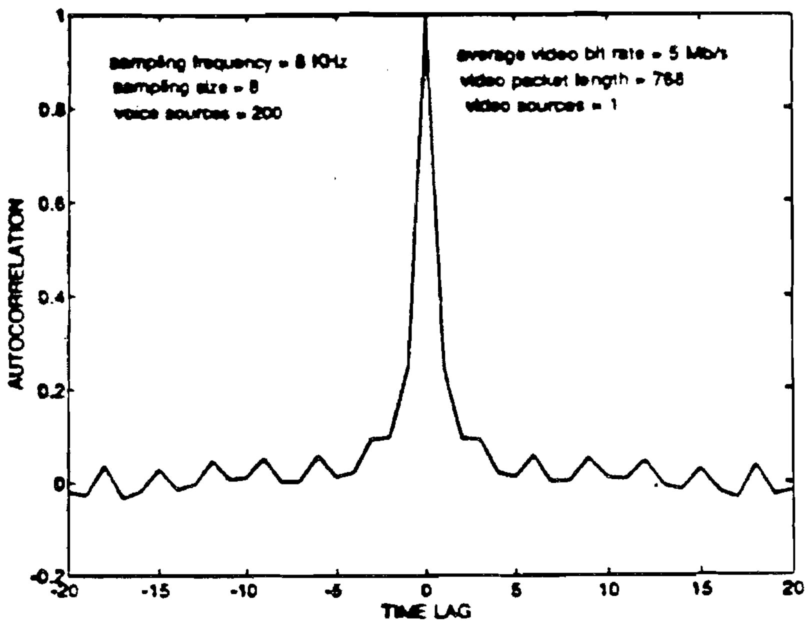


Figure 3.1 Autocorrelation function

The correlation shows that the aggregate sources have periodic components (voice sources) and a reasonable correlation at time origin due to the video source.

This example shows that rate process can point out very well some characteristics of the aggregate sources.

Another characteristic that can be studied by rate process is the probability distribution of the aggregate sources. The rate process can be easily translated to the probability distribution by using the histogram of the number of cells in τ interval. For example, Fig.3.2 shows the probability distribution obtained through rate process for Poisson source.

An advantage of using the histogram as probability distribution is that a relatively small number of discrete values can represent the distribution. Thus, the operations with the distribution can be quickly done by using the direct convolution or the transform method such as FFT (Fast Fourier Transform). This characteristic is very important for proposed connection admission control explained in the previous section.

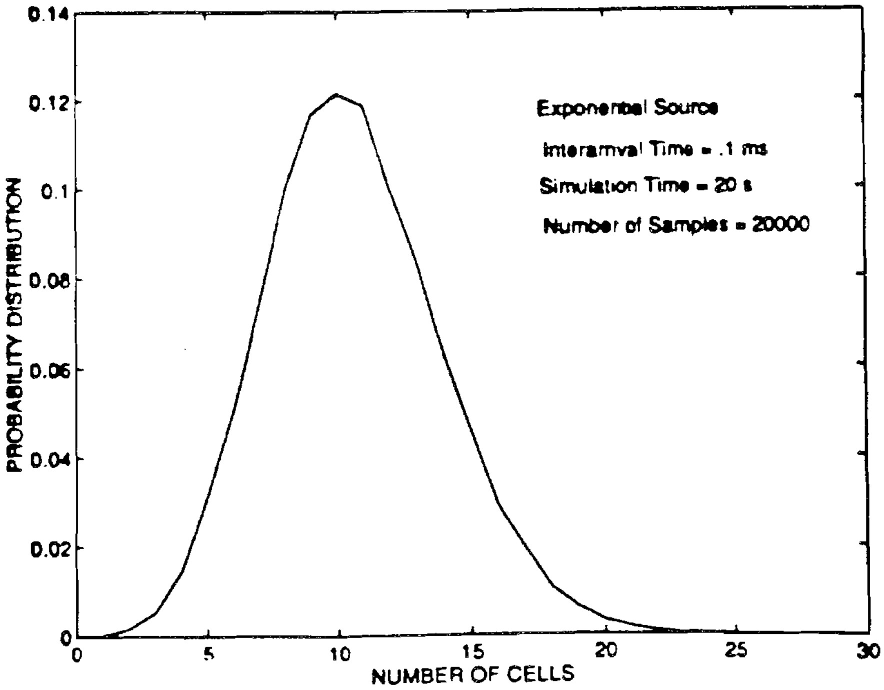


Figure 3.2 Histogram-based Poisson distribution obtained through rate process.

4. SIMULATION AND EXAMPLES

Fig.4.1 shows the simulation scheme implemented to verify the feasibility of the proposed admission control. The simulation tool used is Block Oriented Network Simulator (BONES)[13]. A simple FIFO is used to queue the cells. The block named Channel is used to model the processing delay to transmit a cell. The Data Collector blocks make the counting of the cells in an interval τ and store this information. The block Sink performs the network node function and removes the cells. The feedback at the channel output is used to indicate the end of a cell transmission. The parameters indicated by small arrows are specified before the simulation and can be easily modified.

4.1 Source Models

The On-Off sources have been successfully used in the literature to model the aggregate sources of the broadband integrated services networks. Thus, in this paper, video, data and voice sources are all modeled as on-off sources with different parameters.

Each video source has a peak bit rate of 33 Mb/s and average bit rate equal to 13.5 Mb/s [8]. Each data source has a peak bit rate of 10 Mb/s and average bit rate equal to 1 Mb/s [11]. Each voice source has ON periods, during which cells are generated at 32 kb/s. The average duration of ON and OFF periods are 352 ms and 650 ms, respectively [12].

4.2 Simulation Parameters

The transmission link bandwidth C is assumed to be 150 Mb/s and the buffer size K is limited to 100. The cell length L is 53 bytes. It is assumed that the maximum admissible delay Tmax satisfies the relationship [8]

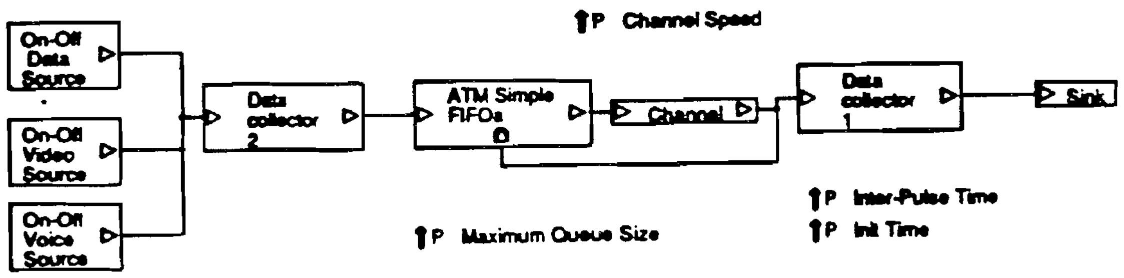


Figure 4.1 Simulation scheme

Tmax = KL/C

We assume that this maximum delay is also the fixed time interval τ in which the number of the cells is counted. Thus, the fixed interval τ for given simulation parameters is .28267 ms. The total number of intervals N is assumed to be 20000.

4.3 Examples

Let us consider an example with video and voice sources. There are 4 video sources and 1065 voice sources connected to the network. Another video source wants to connect to the network. We will consider that to satisfy the QoS (Quality of Service) the cell loss probability must be less than 10^{-2}. This high cell loss probability is chosen to save computing time because it is very difficult to get low cell loss probability in simulation.

In Fig.4.2.a the probability distribution of the aggregate sources with 4 video sources and 1065 voice sources is shown. This distribution was obtained through the histogram of the number of cells counted in τ interval.

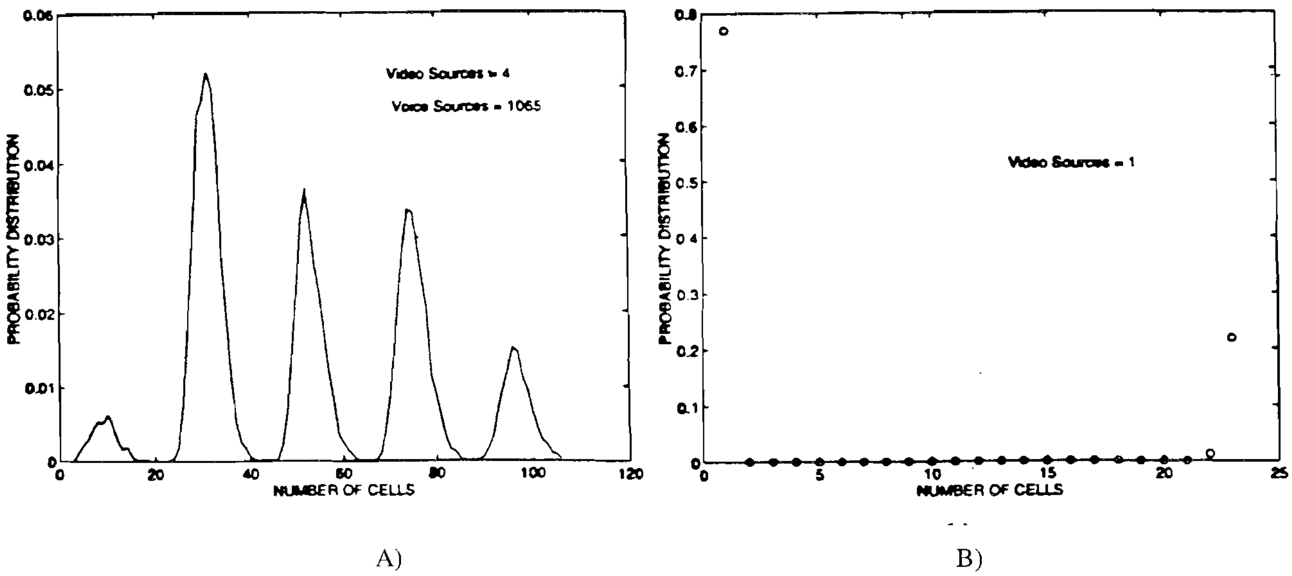
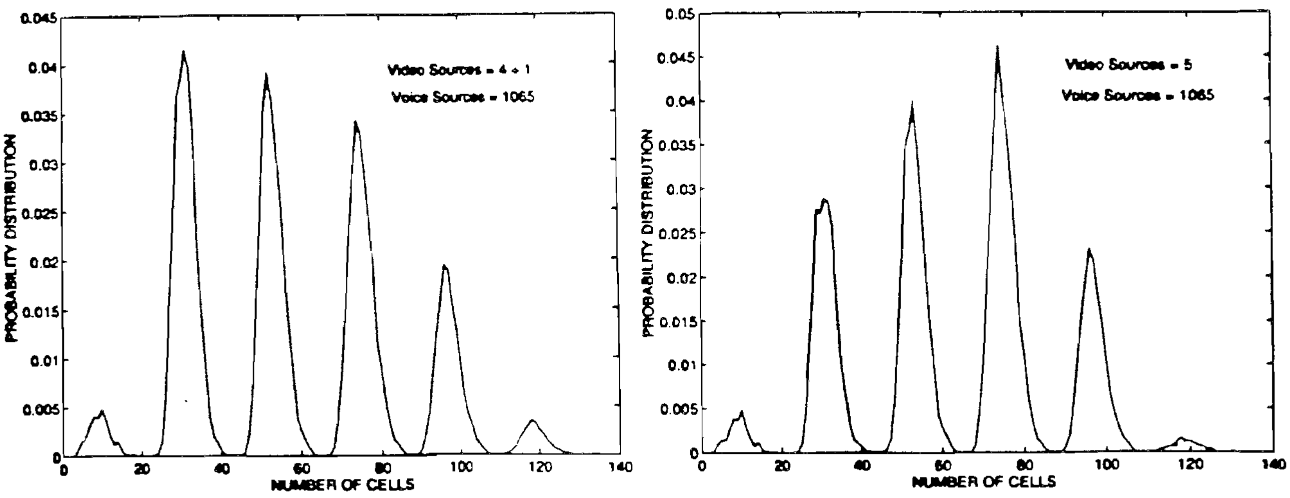


Figure 4.2 Histogram-based probability distributions for aggregate sources and one video source. Two classes of traffic.

Fig.4.2.b shows the probability distribution of only one video source, that is the distribution of the new call. In this case the distribution was obtained considering only one video source in the network.

Having these two distributions we can estimate, before the acceptance of the new call, the aggregate distribution of 5 video sources and 1065 voice sources by calculating the convolution of these two distributions. Fig.4.3.a shows the new estimated distribution of these aggregate sources before the call is accepted and Fig.4.3.b shows the distribution after the call is accepted.

Table I shows the cell loss ratio for the estimated using the Eq.1 and the simulated cases. It can be seen that the new video source can be accepted under specified QoS. Table I also shows that after the acceptance of this new video source, if another video source (total 6 video sources) attempts to enter in the network, that video source will be rejected.



a) b)

Figure 4.3 Probability distribution for two aggregate classes of traffic. a) Estimated before acceptance b) simulatedafter

TABLE I CLR for estimated and simulated cases. Two classes of traffic.

C = 150 Mb/s K = 100			
Video Sources	Voice Sources	Estimated	Simulated
5	1065	0.0078	0.0030
6	1065	0.0140	0.0123

Let us consider now the case with three different classes of traffics. In this case we assume that 3 video sources, 2500 voice sources and 19 data sources are already connected to the network and the number of data sources will be increased. Fig.4.4.a, b and c show the distributions of the aggregate sources obtained in different measurement times and Fig.4.4.d shows one data source distribution.

These different measurement times were obtained using different seeds of simulation. Since the measurement time ($T = \tau N = 5.652s$) is short, different characteristics of the distributions are observed and they represent only samples of the whole distribution. The idea of this method is to use these distributions, but choosing only one distribution that gives the worst cell loss probability. The criterion to choose the worst distribution is based on the comparison of the tails of the distributions. The summation of the tail that exceeds the size of the buffer (buffer size = 100, in our example) is calculated and it is compared to the tail of the next distribution obtained by measurement. If the first summation is greater then the second, the first is kept as the aggregate distribution. Otherwise, the second distribution is now used as new distribution. Thus, the worst measured distribution is used for CLR calculation.

Table II Tail summations for different data sources.

Video Sources = 3	Voice Sources = 2500		
Data Sources	Σ Tails, measured in different times		
19	.0438	.0750 *	.0409
20	.0440	.0786 *	.0409
21	.0452	.0790 *	.0529
22	.0484	.0965 *	.0426
23	.0539	.0909	.0539

The tail summations for each of the distributions in Fig.4.4.a, b and c are .0438, .0750 and .0409, respectively. Thus, the distribution that gives worst CLR in this example is the distribution shown in Fig.4.4.b. Table II shows the tail summations for different load of data sources. As it is pointed out by (*) in the table, the greatest value is chosen in each load.

Table III shows the results of simulation and estimated values using the proposed method.

The notation (x+1) in the Table III means that x data sources are connected and one more is requesting connection. The estimated probability is calculated according to the proposed method before the data source is accepted and the simulated probability is obtained assuming that the new call is accepted.

mea It can be seen in the Table III that in some situation an increased load gives less CLR. It means that in particular time the surement captured a distribution that it is advantageous in CLR terms. But since in this method a periodic measurement is done, a less advantageous distribution pattern is always captured as can be observed in the case of (21+1) data sources. In time t3 the CLR found in simulation is less than (20+1) data sources, but in time t2 the measurement

TABLE III CLR for estimated and simulated cases. Three classes of traffic.

C = 150 Mb/s	K = 100		Video Sources = 3		Voice Sources = 2500		
Data	t1		t2		t3		Simul.
19 + 1	.0050	.0040	.0087	.0078	.0047	.0036	.0055
20 + 1	.0049	.0043	.0090	.0077	.0046	.0051	.0057
21 + 1	.0053	.0044	.0090	.0086	.0063	.0038	.0056
22 + 1	.0053	.0062	.0104	.0084	.0050	.0058	.0068

captured a distribution that gives a higher CLR.

Sometimes the estimated value underestimates the CLR, as can be observed in the case of (22+1) data sources. By using the tail comparison criterion we can keep the CLR overestimated. By the criterion the load with (22+1) data sources is not acceptable and the new data source is rejected, despite the average CLR obtained in the simulation (.0068) satisfies the specified QoS.

In the dynamic situation in which the connections are released, the following procedure could be adopted to choose the distribution. The distribution obtained by measurement soon after the releasing of a connection is adopted as the new

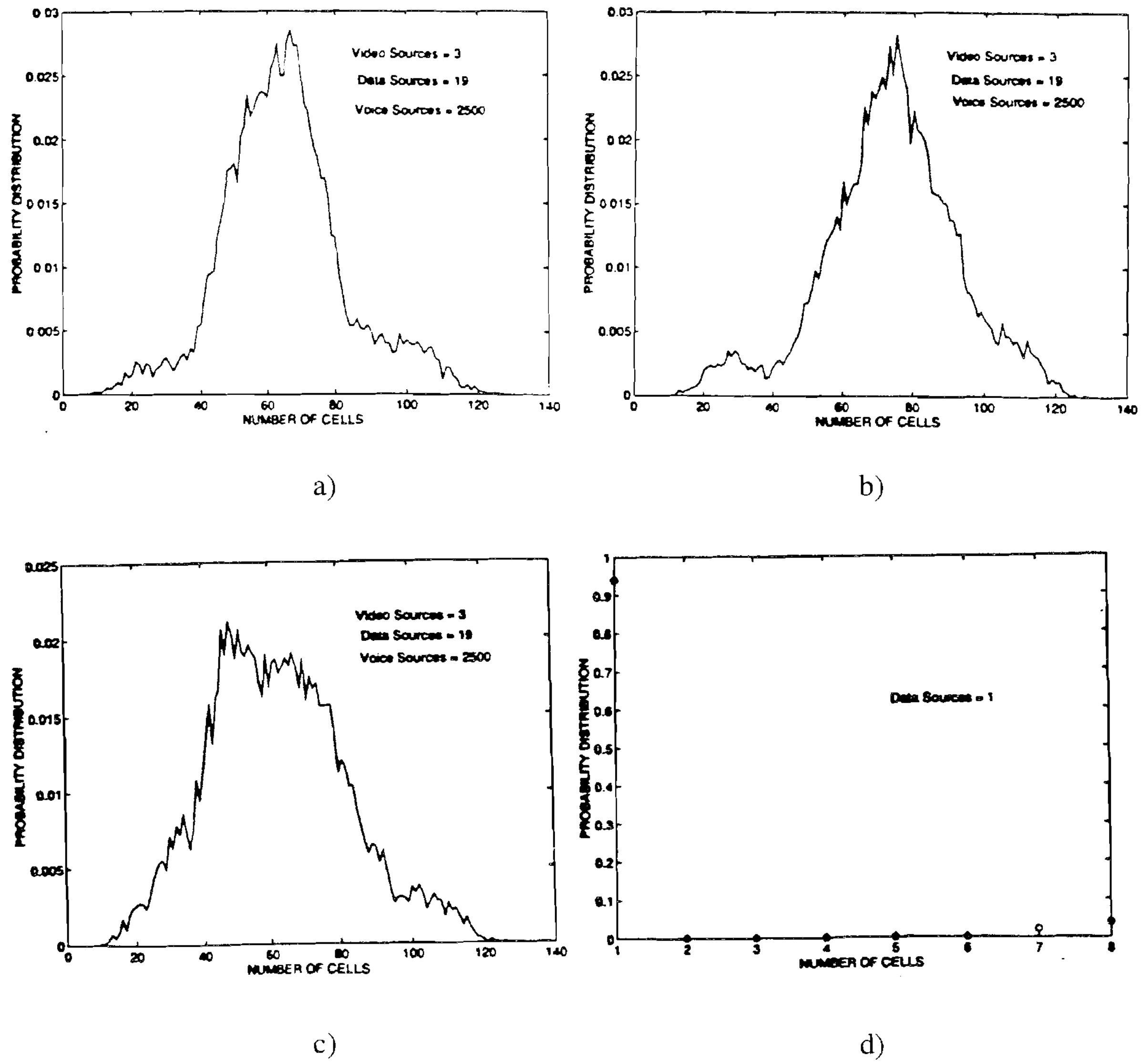


Figure 4.4 Probability Distributions for three mixed classes of traffic. a), b) and c) Distributions obtained in different measurement times. d) One data source distribution.

distribution; after that the same comparison criterion based on the summation of tails explained previously can be used to get the worst distribution.

5. CONCLUSIONS

A connection admission control based on real time traffic statistics was presented in this paper. The method consisted of estimating the distribution of the aggregate sources based on the histogram of the number of cells arriving in a fixed interval τ, and it was used in conjunction with the new call distribution estimated off-line by measurement or simulation. By using convolution technique the new distribution of the aggregate sources was determined and then CLR was evaluated. Since distributions obtained by measurement represented only partial distributions a tail comparison technique was proposed. In this proposed connection admission control the estimated CLR is overestimated because a worst distribution pattern is chosen among distributions obtained in different measurement times.

6. REFERENCES

[1] F. Vakil and H. Saito, "On Congestion Control in ATM Networks," IEEE Magazine on Lightwave Telecommunications Systems, August 1991.

[2] G. Gallassi, G. Rigolio, and L. Fratta, "ATM Bandwidth Assignment and Bandwidth Enforcement Policies," Proceedings of Globecom'89, 1989.

[3] G. Gallassi, G. Rigolio, and L. Verri, "Resource Management and Dimensioning in ATM Networks" IEEE Network Magazine, May 1990.

[4] M. Decina, T. Toniatti, P. Vaccari, and L. Verri, "Bandwidth Assignment and Virtual Call Blocking in ATM Networks," Proceeding of Infocom'90, 1990.

[5] R. Guerin, H. Ahmadi, and M. Naghshineh, "Equivalent Capacity and Its Application to Bandwidth Allocation in High-Speed Networks," IEEE Journal on Selected Areas on Comm., Vol.9, No 7, September 1991.

[6] D. Anick, D. Mitra, and M. M. Sondhi, "Stochastic Theory of a Data Handling System with Multiple Sources," Bell Sys. Tech. J. Vol. 61, No 8, October 1982.

[7] C. Braun, D. W. Petr, and T. G. Kelly, "Performance Evaluation of Equivalent Capacity for Admission Control," Proceedings of IEEE Wichita Conference on Comm., Networking and Signal Processing, April 1994.

[8] H. Saito and K. Shiomoto, "Dynamic Call Admission Control in ATM Networks", IEEE Journal on Selected Areas in Comm, Vol. 9, No 7, September 1991.

[9] H. Saito, "New Dimensioning Concept for ATM Networks," Proceedings of 7th International Teletraffic Congress (ITC) Seminar, Morristown, NJ, October 1990.

[10] S. Motoyama and V. S. Frost, "Mixed Traffic Characteristics in ATM Networks," Proceedings of IEEE Wichita Conference on Comm., Networking and Signal Processing, April 1994.

[11] A. Baiocchi, N. B. Melazzi, M. Listanti, A. Roveri, and R. Winkler, "Loss Performance Analysis of an ATM Multiplexer Loaded with High-Speed On-Off Sources," IEEE Journal on Selected Areas in Comm, Vol. 9, No 3, April 1991

[12] H. Heffes and D. M. Lucantoni, "A Markov Modulated Characterization of Packetized Voice and Data Traffic and Related Statistical Multiplexer Performance," IEEE J. Select. Areas Comm., Vol. 4, No 6, 1986..

[13] K. S. Shanmugan, V. S. Frost, W. W. LaRue, "A Block Oriented Network Simulator (BONeS)", Simulation, Vol. 58, No. 2, 1992 .

Shusaburo Motoyama Formou-se em Engenharia Elétrica pela EESC-USP em 1972, obteve os títulos de Mestre em Engenharia Elétrica pela FEC-UNICAMP em 1975 e Doutor em EngenhariaElétrica pela Universidade de Tóquio, Japão em 1982. Desde agosto de 1975 é professor da FEE-UNICAMP. Foi professor colaborador do Departamento de Computação da UFSCar no período de novembro 1987 a dezembro de 1988. Foi pesquisador visitante no TISL-Telecommunication and Information Sciences Laboratory, da Universidade de Kansas, USA no período de junho de 1993 a junho de 1995. Suas áreas de interesse atual incluem Redes de Alta velocidade, Comutação Digital e Processamento Distribuído.

Victor S. Frost Formou-se em Electrical Engineering pela University of Kansas em 1977, obteve os títulos de MSEE pela University of Kansas em 1978 e PhDEE pela University of Kansas em 1982. Foi Visiting Scientist no German Aerospace Research Establishment Institute no período de julho de 1981 a dezembro de 1981. É atualmente Professor of University of Kansas e Diretor do TISL Telecommunications and Information Sciences Laboratory. Suas areas de interesse incluem Integrated Telecommunication Networks, Traffic/Service Integration e Congestion Control.

David W. Petr Formou-se em Electrical Engineering pela Southern Methodist University em 1976,obteve os títulos de MSEE pela Stanford University em 1978 e de PhDEE pela University of Kansas em 1989. É atualmente Associate Professor of University of Kansas. Suas areas de interesse incluem Integrated Communications Networks, Performance Analysis and Simulation e Adaptive Systems.