Numerical example: The effectiveness of the harmonic analysis procedure developed is examined here for a signal given as

$$s(t) = 0.1+0.3\cos(2\pi\times 25t+70^\circ)+0.2\cos(2\pi\times(50/\sqrt{2})t+60^\circ) +0.7\cos(2\pi\times(50/\sqrt{2}+0.5)t+80^\circ)+1.0\cos(2\pi\times 50t) +0.5\cos(2\pi\times\sqrt{3}\times 50t+90^\circ)+0.4\cos(2\pi\times 3\times 50t+40^\circ) \quad (6)$$

This is an exacting test case. The signal $s(t)$ has frequencies that have irrational number relationships to one another (e.g. 50, $50/\sqrt{2}$ and $50/\sqrt{3}$), together with harmonics which have closely similar frequencies (i.e. $50/\sqrt{2}$ = 35.35Hz and $50/\sqrt{2}$ + 0.5 = 35.85Hz).

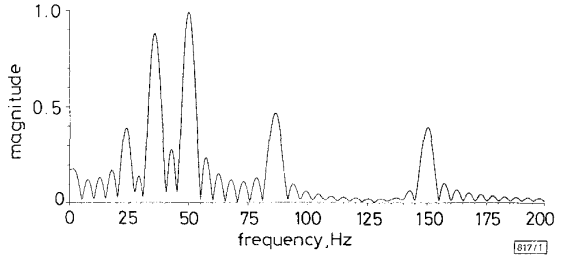

Fig. 1 *Frequency spectrum of test wavefrom as found from DFT analysis*

For our purposes, it is taken that there are 600 equally spaced data samples available corresponding to a time record length of 200ms for a sampling frequency of 3kHz. Initial DFT analysis over the complete record length of 600 samples gives the spectrum of Fig. 1. On the basis of the peaks in the spectrum, DFT analysis identifies the five frequency components summarised in Table 1.

Table 1: Results of harmonic analysis

Frequency			Magnitude			Phase		
Actual	DFT	New method	Actual	DFT	New method	Actual	DFT	New method
Hz	Hz	Hz				deg	deg	deg
0.0	0.0	0.0	0.1	0.175	0.1	0.0	0.0	0.0
25.0	23.9	25.0	0.3	0.389	0.3	70.0	118.9	70.0
35.35	—	35.35	0.2	—	0.2	60.0	—	60.1
35.85	35.9	35.85	0.7	0.878	0.7	80.0	73.79	80.0
50.0	50.0	50.0	1.0	0.992	1.0	0.0	–2.44	1.1×10^{-4}
86.6	86.5	86.6	0.5	0.467	0.5	90.0	91.61	90.0
150.0	150.2	150.0	0.4	0.396	0.4	40.0	30.46	40.0

All the numerous other frequency components of the spectrum of Fig. 1 are spurious and derive from spectral leakage. Of the two frequency components that are separated in frequency by only 0.5Hz, the one that dominates in Fig. 1 is the component of frequency 35.85Hz, whose magnitude is 0.7.

Owing to spectral leakage, the DFT process fails to resolve between the two signal components of frequencies 35.35Hz and 35.85Hz.

The separation of the total data file into training and test partitions is at choice. An initial choice is made here of 500 data samples in the training set, leaving 100 samples for the test set. Therefore, the maximum length for the training set is $M1 = 500$. That for the test set is $M - M1$, or 100.

From the initial DFT processing, $N = 5$. Using the quasi-Newton procedure, error minimisation for the training set leads to convergence at low error. However, the test errors are high, indicating a need to revise the initial value of N. There is no indication from the test error magnitude of whether N should be increased or decreased to lower the error. Choosing first to decrement N to a value of 4, increases the error in both the training set minimisation and the test set. Successively reducing N to 3, then to 2, and then to 1 confirms a trend of increasing error in both the training set and test set. Returning from this sequence to the initial choice of $N = 5$ and then incrementing to $N = 6$ leads to a low error of 10^{-9} in minimisation confined to the training set and 10^{-6} for the test error. For checking purposes, increasing N to 7 and then to 8 confirms levels of test error similar to those for $N = 4$ or 3. The procedure therefore correctly gives the value $N = 6$.

By monitoring the test error, the minimum length for the training set required is also identified. For the correct value of $N = 6$, the test error is within the specified tolerance when $M1 = 260$. Increasing $M1$ beyond 260 does not change the test error. The procedure developed therefore discloses a minimum data window length requirement in this example of 260 samples. The final results of harmonic analysis using the error minimisation method of the paper are summarised in Table 1.

Overall, the analysis method guarantees very low errors in finding harmonic frequencies together with the magnitude and phase values of all harmonic components. The accurate identification of the two frequency components that are separated by only 0.5Hz indicates the high resolving power of the method. Spectral leakage errors are avoided completely. The low sensitivity of the method to the length of the training set and of the test set provides a good indication of the robustness of the method in practical use.

Conclusions: The new method of parametric harmonic analysis based on data partitioning achieves a correct identification of the number of frequency components; a good estimate of the minimum record length required; a valid identification of all frequency components; a high resolution between frequency components in the signal waveform for which harmonic analysis is required.

25 March 1996
Electronics Letters Online No: 19960784

T.T. Nguyen (*Energy Systems Centre, The University of Western Australia, Australia*)

References

1 KAY, S.M., and MARPLE, S.L., Jr.: 'Spectrum analysis - A modern perspective', *Proc. IEEE*, 1981, **69**, (11), pp. 1380–1419

Estimates of loss probabilities for delay-sensitive traffic in ATM networks

K.-C. Lai and T.-H. Lee

Indexing terms: Asynchronous transfer mode, Telecommunication traffic

Simple estimates of loss probabilities for heterogeneous delay-sensitive traffic in ATM networks are presented. Cells of different connections can have different loss priorities. Numerical results show that the estimates, which are derived from the bufferless fluid flow model, are close to the actual cell loss probabilities.

Introduction: Loss probability is considered to be an important measure of quality of service (QOS) in ATM networks. To make fast admission control possible, the loss probability must be computed in real-time. Unfortunately, it is often very time consuming to compute the exact loss probability with queueing models.

In this Letter we present simple estimates of loss probabilities for heterogeneous delay sensitive traffic with multiple QOS requirements. Cells of different connections (and within a connection) are allowed to have different loss priorities. The estimates are derived from the bufferless fluid flow model. Numerical results show that the estimates are close to the actual cell loss probabilities.

Estimates of loss probabilities: Consider a multiplexer with M independent connections that generate delay-sensitive traffic. The link capacity is denoted by C. Since all connections generate delay-sensitive traffic, we assume that the buffer size in the multiplexer is small and thus its effect can be neglected.

Assume that the multiplexer supports N loss priorities. Number the priorities 1 to N so that the loss probability of priority i cells is smaller than or equal to that of priority j cells if $i > j$. To cope with multiple loss probability requirements, the link capacity C is divided into N bands denoted by C_1, C_2, ..., C_N which can be

dynamically adjusted to reduce connection blockirig probability. A cell of priority i can share bands C_1, C_2, ..., C_i. Moreover, all cells of priority i or higher share band C_i fairly.

All traffic sources are assumed to be non-negative, bounded, stationary and ergodic. We adopt the fluid flow model [1] to simplify the calculation of loss probability. Let $Z(t)$ denote the aggregate traffic generated by these M mutually independent connections. Under the above assumptions, to compute loss probabilities, it suffices to use a random variable Z to represent the random process $Z(t)$, so long as the density functions of Z and $Z(t)$ are identical for all t. Let Z_j denote the priority j traffic generated by the existing connections. It can be shown that as M approaches infinity

$$\frac{Z_j}{Z} \to \frac{E[Z_j]}{E[Z]} = \frac{\text{MEAN}_j}{\text{MEAN}} := p_j \qquad (1)$$

with probability one [2], where MEAN = $E[Z]$ and $\text{MEAN}_j = E[Z_j]$. We prove a lemma concerning the average loss rate of priority j traffic in the following section.

Lemma 1: Let $L_j(Z)$ denote the average loss rate of priority j traffic. It holds that

$$L_j(Z) = E\left[Z_j\left(1 - \sum_{i=1}^{j}\frac{C_i}{\hat{Z}_i}\right)^+\right]$$

where

$$\hat{Z}_i = \sum_{k=1}^{N} Z_k$$

Proof: We prove this lemma by induction. Let $loss_j(Z)$ denote the loss rate of priority j traffic so that $L_j(Z) = E[loss_j(Z)]$. Also, let $res_j(Z)$ represent the traffic that will compete for band C_j. Then we have

$$loss_j(Z) = (res_j(Z) - C_j)^+ \frac{Z_j}{Z_j + \hat{Z}_{j+1}} = (res_j(Z) - C_j)^+ \frac{Z_j}{\hat{Z}_j} \qquad (2)$$

and

$$res_{j+1}(Z) = (res_j(Z) - C_j)^+ \frac{\hat{Z}_{j+1}}{\hat{Z}_j} \qquad (3)$$

Hence, by eqns. 2 and 3 we have

$$res_{j+1}(Z) = loss_j(Z)\frac{\hat{Z}_{j+1}}{Z_j} \qquad (4)$$

For $j = 1$, we have $res_1(Z) = Z$ and

$$loss_1(Z) = Z_1\left(1 - \frac{C_1}{\hat{Z}_1}\right)^+$$

Suppose that

$$loss_n(Z) = Z_n\left(1 - \sum_{i=1}^{n}\frac{C_i}{\hat{Z}_i}\right)^+$$

For $j = n + 1 \le N$, we have, by eqn. 4

$$loss_{n+1}(Z) = [res_{n+1} - C_{n+1}]^+ \frac{Z_{n+1}}{\hat{Z}_{n+1}} = \left[loss_n(Z)\frac{\hat{Z}_{n+1}}{Z_n} - C_{n+1}\right]^+ \frac{Z_{n+1}}{\hat{Z}_{n+1}} = \left[Z_{n+1}\left(1 - \sum_{i=1}^{n+1}\frac{C_i}{\hat{Z}_i}\right)\right]^+$$

Therefore, it holds that

$$L_j(Z) = E\left[Z_j\left(1 - \sum_{i=1}^{j}\frac{C_i}{\hat{Z}_i}\right)^+\right]$$

This completes the proof of Lemma 1.

Let P_j denote the loss probability of priority j traffic. It is clear that

$$P_j = \frac{L_j(Z)}{E[Z_j]}$$

which is difficult to compute in real-time. Fortunately, when the number of connections is large we have

$$P_j = \frac{L_j(Z)}{E[Z_j]} = \frac{E\left[Z_j\left(1 - \sum_{i=1}^{j}\frac{C_i}{\hat{Z}_i}\right)^+\right]}{p_j E[Z]} \simeq \frac{E\left[p_j Z\left(1 - \sum_{i=1}^{j}\frac{C_i}{\hat{p}_i Z}\right)^+\right]}{p_j E[Z]} = \frac{E[Z - \Omega_j]^+}{E[Z]} := Pv_j \qquad (5)$$

where

$$\Omega_j = \sum_{i=1}^{j}\frac{C_i}{\hat{p}_i} \quad \text{and} \quad \hat{p}_i = \sum_{k=1}^{N} p_k$$

We suggest using Pv_j to approximate P_j. Notice that p_i can be easily obtained if the variables MEAN_i and MEAN are stored. Hence, Ω_j can be computed in real time. In a real system the bands C_1, C_2, ..., C_N may be dynamically adjusted to reduce connection blocking probability. In this case Ω_j needs to be recomputed, which involves only a few divisions and thus can be accomplished in real time.

The difficulty to obtain Pv_j is the computation of $E[Z-\Omega_j]^+$. A real-time computation algorithm for $E[Z-\Omega_j]^+$ can be found in [3].

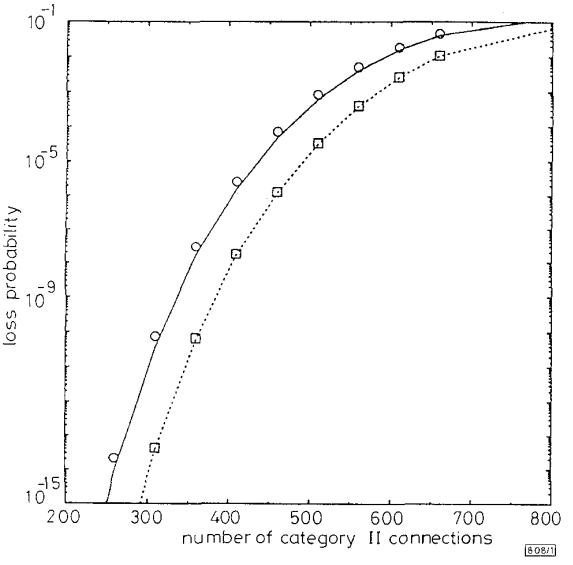

Fig. 1 *Loss probability against number of category II connections*

Number of category I connections is 100
$d_2^{(1)} = 0$, $d_2^{(2)} = 1$
——— P_1
····· P_2
○ Pv_1
□ Pv_2

Numerical examples: For the numerical examples, we assume C = 150Mbps and $N = 2$. The link capacity is partitioned into two bands C_1 and C_2, where C_1 = 9/10 C and C_2 = 1/10 C. All traffic sources are assumed to be on–off sources. Two categories of connection are considered. A connection of category I has peak bit rate $\text{MAX}^{(1)}$ = 64kbit/s and average bit rate $\text{AVG}^{(1)}$ = 32kbit/s. The peak bit rate and average bit rate of a category II connection are $\text{MAX}^{(2)}$ = 2Mbit/s and $\text{AVG}^{(2)}$ = 0.2Mbit/s, respectively. The percentage of high priority cells of category I and category II connections are denoted by $d_2^{(1)}$ and $d_2^{(2)}$, respectively. In our study, $d_2^{(1)}$ is chosen to be 0 (i.e. all cells are of low priority) and $d_2^{(2)} = 1$ (i.e. all cells are of high priority). Fig. 1 illustrates the curves for Pv_j and the actual cell loss probability P_j. It can be seen that Pv_j is an excellent approximation of the actual cell loss probability.

Conclusion: We have presented asymptotic estimates of loss probabilities of all priority classes for delay-sensitive traffic with multiple quality of service requirements. The asymptotic estimates can be computed in real time and are close to actual loss probabilities. The estimates can be used to design real-time resource allocation schemes. We have also performed numerical examples for other cases (which are not presented here owing to space limitation), and have found that Pv_j seems to be an upper bound of P_j.

25 March 1996
Electronics Letters Online No: 19960783

K.-C. Lai (*Institute of Electronics, National Chiao Tung University, Hsinchu, Taiwan 300, Republic of China*)

T.-H. Lee (*Institute of Communication Engineering, National Chiao Tung University, Hsinchu, Taiwan 300, Republic of China*)

References

1 SAITO, H.: 'Teletraffic technologies in ATM networks' (Artech House, 1994)
2 NEVEU, J.: 'Mathematical foundations of the calculus of probability' (Holden-Day, San Francisco, 1965)
3 LEE, T.H., LAI, K.C., and DUANN, S.I.: 'Real-time call admission control for ATM networks with heterogeneous bursty traffic'. IEEE ICC, 1994, **1**, pp. 80–85
4 ESAKI, H.: 'Call admission control method in ATM networks'. IEEE ICC, 1992, **3**, pp. 1628–1633

Frame decorrelation for noise-robust speech recognition

H.Y. Jung, D.Y. Kim and C.K. Un

Indexing terms: Speech recognition, Signal processing

The authors propose a frame decorrelation method to cope with background noise in speech recognition. Since noise is modelled as a stationary perturbation in most cases, it is effective to reduce slow-varying components. One example ofusing this principle is the highpass scheme. The proposed method has the same property as the highpass scheme. It transforms feature vector sequences into decorrelated sequences and enhances transition regions. Simulation results show that this method is effective for speech with significant noise, and works better than other highpass methods.

Introduction: It is well known that the performance of a speech recognition system becomes degraded severely when training and testing are carried out at different noise levels. To overcome this problem, many researchers used the spectral subtraction method developed in the context of speech enhancement, but in this method noise in corrupted speech must be estimated, which is a very difficult and unsolved problem in real environments. This necessitates a technique that requires no noise estimation. The proposed decorrelation method is a kind of highpass scheme [1, 2]. With this method slowly varying components can slowly be reduced to deal with background noise, thus enabling the stationary regions to be more attenuated than the transition regions between speech segments. This effect can be considered as decorrelation since the stationary regions are more correlated than the transition regions. The frame decorrelation method removes correlations between feature vectors. It transforms feature vector sequences into decorrelated sequences and enhances the transition regions as in the highpass scheme. The enhancement of transition regions will provide good discrimination for speech recognition. According to speaker-independent isolated word recognition experiments, the proposed method is effective for significantly noisy speech and yields better performance than other highpass schemes.

Estimation of the power spectrum: Here we need to estimate the power spectrum of feature vectors to apply this to the decorrelation procedure. This power spectrum is an important factor for decorrelation between feature vectors. It is estimated from statistics of feature vectors, and filter bank outputs are used as feature vectors. We assume that the filter bank vector $O(f, t)$ (where f represents frequency and t represents frame index) is independent of frequency, and estimate individually the power spectrum in each frequency band. The power spectrum of frequency band f_o is represented by

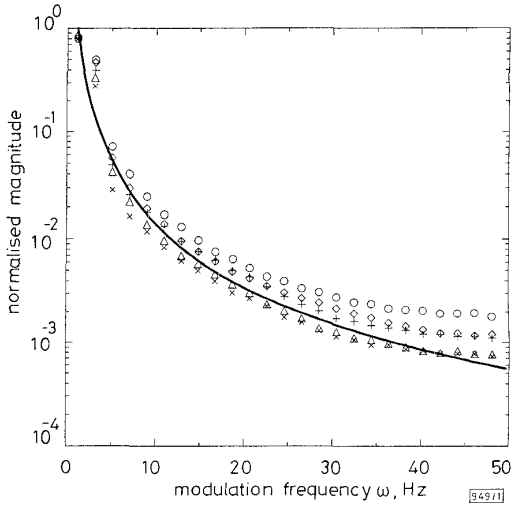

Fig. 1 *Power spectrum of observation vectors*
——— w^2, ○ 215Hz
◇ 585 Hz, △ 1259 Hz
+ 2492Hz, × 4745Hz

$$|S(f_o, w)|^2 = \left| \frac{1}{T} \int_0^T o(f_o, t) e^{j2\pi wt} dt \right|^2 \quad (1)$$

where T is the frame length, $o(f_o, t)$ is a feature coefficient at frequency f_o, and w is a modulation frequency that describes the temporal variation of subband energy. Fig. 1 shows the average power spectrum obtained using 1500 words uttered by 20 male speakers. It has a similar envelope in each subband, and the spectral envelopes are approximately a function of $1/w^2$. Therefore, the power spectrum in each subband is approximated by

$$|S(f_o, w)|^2 \simeq \frac{1}{w^2} \quad (2)$$

Frame decorrelation procedure: We assume that each feature vector $O(f, t_t)$ is independent of frequency, and apply the decorrelation procedure to remove correlations between feature coefficient sequences in each subband, i.e. the decorrelation procedure transforms feature sequences into decorrelated sequences. Assuming that the sequence of feature coefficients at subband f_o, $O(o(f_o, t_1), o(f_o, t_2), ..., o(f_o, t_T))$ is transformed into $Y = (y(f_o, t_1), y(f_o, t_2), ..., y(f_o, f_T))$ by a decorrelation filter D, $y(f_o, t)$ is represented as a temporal convolution of D and O. Here, in order for Y to be decorrelated, the correlation function of Y must satisfy the following [3]:

$$E(y(f_o, t) y^*(f_o, t')) = E((D \cdot O)(D \cdot O')^*) = \delta(f_o; t - t') \quad (3)$$

where $\delta(f_o; t - t')$ is the Kronecker delta function which is one when $t = t'$. Thus, the Fourier transform of eqn. 3 is given by

$$D(f_o, w) R(f_o, w) D^*(f_o, w) = 1 \quad (4)$$

where $R(f_o, w)$ is the power spectrum of feature sequences in subband f_o, and is represented by the sum of signal and noise power spectra i.e. $|S(f_o, w)|^2 + |N(w)|^2$. Note that, since the decorrelation filter removes slowly varying components, a lowpass filter is needed to remove fast varying noise. We use the least-squares optimal filtering method to obtain the lowpass filter. This filter transforms the corrupted signal $(S(f_o, w) + N(w))$ into the signal that is close to the uncorrupted signal $S(f_o, w)$. The optimal filter obtained is

$$\Phi(f_o, w) = \frac{|S(f_o, w)|^2}{|S(f_o, w)|^2 + |N(w)|^2} \quad (5)$$