A Hierarchical Analysis of Access Multiplexers with Multimedia Traffic

Sayeed Ghani* and Mischa Schwartz**
* Dept. of MIS and Computer Science, IBA, Karachi, Pakistan, sghani@iba.edu.pk
** Dept. of Elect. Eng., Columbia University, NY, schwartz@ee.columbia.edu

Abstract

Decomposition approximations have recently been used in the performance evaluation of access networks such as GSM/GPRS and ATM based networks. In this paper we extend the analysis to a network possessing a hierarchy of traffic types, such as voice, data and broadband video. We look at different policies for the transmission of the traffic, and apply a hierarchical Decomposition technique to the performance analysis of the network where necessary. We find that the Decomposition technique is quite useful in predicting the performance of networks where video, voice and data have channel holding times which are each of different orders of magnitude in comparison to each other.[1]

1. Introduction

Various approximation methods have been used in the past for the performance evaluation of Integrated Networks that carry different types of traffic such as voice and data. In particular a *Decomposition approximation*, developed in [1], has been usefully applied in the evaluation of GSM/GPRS (see [2], [3], [4] and [5]) and ATM networks (see [6], [7], [8] and [9]). In this paper we generalize the methodology of [1] which was originally applied to two classes of narrowband traffic – Voice and Data, by including a third category of broadband traffic that of video. This method has also been applied to non-preemptive priority analysis in the case of GSM/GPRS networks [10]. [11] presents an overview of this technique as well as a review of different areas where the approximation has been recently applied.

In order to study the applicability of the decomposition technique to broadband integrated services we consider two alternate broadband integration policies, termed the *Dual Layered Movable Boundary Scheme (DLMBS)*, introduced in section 2, and *Movable Boundary Complete Sharing (MBCS)*, discussed in section 3, and apply the Decomposition technique to their analysis. Although we do briefly compare the two policies in section 4, our goal here is primarily to study the suitability of the Decomposition technique rather than to search for optimal access policies. In section 5 we consider the analysis of data integration with the circuit switched traffic. Finally in section 6 we look at the data performance and compare the approximation results with exact results.

2. DLMBS

In this scheme we extend the voice/data movable boundary scheme that has been earlier analyzed in [1] to incorporate broadband video circuit switched channels. The scheme, which we choose to analyze, is as follows. The frame structure is shown in figure 1. It is similar to the Senet frame structure of figure 1 of [1], with the addition of video channels. As before N_d (narrowband) channels are reserved for data. Data packets are packet switched with a maximum of M packets allowed into the system. N_v channels are shared between voice and data with voice having preemptive priority over data. A single broadband video connection requires b_w channels, defined as a broadband channel. N_w such connections are available for video and are shared with voice and data as follows. The $N_w b_w$ channels are allocated to video, voice and data with video having preemptive priority over voice and data, and voice in turn having preemptive priority over data. Since video is circuit switched it experiences blocking. The total number of channels N, is then $N = N_d + N_v + N_w b_w$. We also define N_c as the number of circuit switched channels. Thus $N_c = N_v + N_w b_w$.

Since we have assumed that video can preempt voice channels, it is logical to also assume a Speech Activity Detection (SAD) model for voice (see e.g. [12], [13]), with voice channels being allocated only for the duration of a talkspurt. This implies a talkspurt model of the voice process, assuming S off-hook voice sources. The SAD model is based on the fact that whereas preemption of a talkspurt by video may be acceptable, the preemption of an entire voice call would clearly not be acceptable.

We describe the scheme outlined above as a *Dual-Layered Movable Boundary Scheme* (DLMBS). It can be thought of as an extension of the Movable Boundary Scheme of [1]. We choose to assign the priorities on the bases of the mean holding times of the services. Since we generally have

$$1/\mu_w >> 1/\mu_v >> 1/\mu_d \qquad (2.1)$$

[1] This work is based on the PhD thesis of Dr. Ghani that was supported by the National Science Foundation under Grant CDR-84-21402 and by the Office of Naval Research under Grant N00014-85-K0371.

with $1/\mu_w$ being the mean holding time of video calls, video is given the highest priority and data the lowest. (2.1) is justified in that typically the holding times of video may range from 30 – 150 minutes, while that of voice may only be a few minutes long.

The voice and data models are the same as in [1]. We use the talkspurt model of [12] for the voice process. Thus we have S off-hook voice sources; each voice source alternating between talkspurt and silence, with talkspurt and silence periods exponentially distributed with means $1/\mu_v$ and $1/\lambda_v$ respectively. We define $\rho_v \equiv \lambda_v/\mu_v$. Data arrivals are Poisson at rate λ_d and service times exponentially distributed with mean $1/\mu_d$. $\rho_d \equiv \lambda_d/\mu_d$ is the data utilization.

For video we assume arrivals are Poisson with rate λ_w, and holding times exponentially distributed with mean $1/\mu_w$. The video utilization is defined as $\rho_w \equiv \lambda_w/\mu_w$. γ_c is defined as the total circuit switched throughput: $\gamma_c \equiv \gamma_v + \gamma_w$, while γ_v, γ_w are the voice and video throughput respectively.

We now define the following: $W(t)$ = number of video calls in the system, $V_T(t)$ = number of off-hook voice sources in talkspurt, $V_C(t)$ = number of channels occupied by voice talkspurts, and $D(t)$ = number of data packets in the system. We refer to $V_T(t)$ as the voice *talkspurts* process, and $V_C(t)$ as the voice *channel* process. It is clear that $V_C(t)$ depends directly on $V_T(t)$ and $W(t)$, and is given by

$$V_C(t) = \min(V_T(t), N_c - b_w W(t)) \quad (2.2)$$

Thus we will be primarily interested in finding the equilibrium distribution of $W(t)$ and $V_T(t)$, after which $V_C(t)$ can easily be determined. At any time the number of talkspurts being frozen out is $V_T(t) - V_C(t)$. If we assume that the holding times of all the services are much greater than the frame length T, then the video, voice and data process can be approximated by a 3-dimensional continuous-time Markov Chain $\{W(t), V_T(t), D(t)\}$. The equilibrium state probability vector $p_{k,i,j}$ is defined as

$$p_{k,i,j} \equiv \text{Probability}\left(W(t) = k, V_T(t) = i, D(t) = j\right),$$
$$k = 0, 1, \ldots, N_w;\quad i = 0, 1, \ldots, S;\quad j = 0, 1, \ldots, M.$$

We define $q(k,i,j;k',i',j')$ to be the transition rates from state (k,i,j) to (k',i',j'). The non-zero transition rates are as follows.

$$\begin{aligned}
&q(k,i,j;k+1,i,j) = \lambda_w && k = 0, 1, \ldots, N_w - 1\\
&q(k,i,j;k-1,i,j) = k\mu_w && k = 1, \ldots, N_w\\
&q(k,i,j;k,i+1,j) = (S-i)\lambda_v && i = 0, 1, \ldots, S-1\\
&q(k,i,j;k,i-1,j) = i\mu_v && i = 1, 2, \ldots, S \quad (2.3)\\
&q(k,i,j;k,i,j+1) = \lambda_d && j = 0, \ldots, M-1\\
&q(k,i,j;k,i,j-1) = \min(j, j_{k,i})\mu_d && j = 0, \ldots, M-1
\end{aligned}$$

For convenience we have defined $j_{k,i}$ to be the number of channels available to data when video and voice talkspurt processes are in state k, i respectively:

$$j_{k,i} \equiv N - \min(kb_w + i, N_c) \quad (2.4)$$

Figure 2 shows an example of the Markov State Transition diagram for the circuit switched processes $\{W(t), V_T(t)\}$, with $N_c = 5$, $N_v = 3$, $N_w = 1$, $b_w = 2$, $S = 6$. The states in which one or more talkspurts are being frozen out, i.e. when $V_T(t) - V_c(t) \geq 0$, are shown by being circled.

Now the analysis can be carried out in a way similar to the method adopted in [1]. By going through the Courtois Decomposition technique section III of [1], we can show that by using diagonal folding we get the following approximation.

$$P_{k,i,j} \approx P^{\bullet}_{k,i,j} = p(j \mid k,i) p_{k,i} \quad (2.5)$$

Here $p^{\bullet}_{k,i,j}$ is the approximation to $p^{\bullet}_{k,i,j}$. $p(j \mid k,i)$ is the conditional probability:

$$p(j \mid k,i) = \Pr(D(t) = j \mid W(t) = k, V_T = i),$$

and $P_{k,i}$ is the marginal probability:

$$p_{k,i} \equiv \Pr(W(t) = k, V_T = i) = \sum_{\forall j} p_{k,i,j}\,.$$

Since the circuit switched services have preemptive priority over data, they are independent of data. Thus in this section and section 3 we will ignore data, and try to get a solution for the circuit switched equilibrium probability distribution $p_{k,i}$. In section 5 we will combine this with the analysis of data.

We first show that the solution to the circuit switched sources can be found exactly. We have noted that video is independent of voice and data, and voice is independent of data. Now although the voice channels process $V_C(t)$ does depend on video, the voice talkspurt process $V_T(t)$ is independent of video. This is clear from the transition rates (2.3). Thus we can write down the solution for the video and voice equilibrium probability distributions $p_{k,i}$ as $p_{k,i} = p_w(k) p_T(i)$. $W(t)$ simply has the Erlang-B distribution:

$$p_w(k) \equiv \Pr(W(t) = k) = \frac{\rho_w^k}{k!} \Big/ \sum_{i=0}^{N_w} \rho_w^i / i! \qquad k = 0, \ldots, N_w \quad (2.6)$$

$V_T(t)$ has the Engset distribution:

$$p_T(i) \equiv \Pr(V_T(t) = i) = \binom{S}{i} \rho_v^i \Big/ (1 + \rho_v)^s \qquad i = 0, 1, \ldots, S \quad (2.7)$$

As stated previously in (2.2) the channel process $V_C(t)$ depends directly on the talkspurt process $V_T(t)$, and $W(t)$ which determines the number of channels available to voice. Thus the conditional equilibrium distribution $p_C(i \mid k) \equiv \Pr(V_C(t) = i \mid W(t) = k)$ is given by:

$$p_C(i \mid k) = \begin{cases} p_T(i) & i = 0, 1, \ldots, i_k - 1 \\ \sum_{i=i_k}^{S} p_T(i) & i = i_k \end{cases} \quad (2.8)$$

Where i_k is the number of channels available for voice given that the video is occupying k broadband channels: $i_k = N_c - b_w k$.

Having obtained $P_{k,i}$, we can now find the voice talkspurt *cutout fraction* ϕ using [14]: $\phi = \frac{\textit{average number of clipped talkspurts}}{\textit{average number of talkspurts}}$. This gives $\phi = \frac{\sum_{k=0}^{N_w}\sum_{i=i_k}^{S}\lceil i - i_k \rceil p_{k,i}}{Sp'}$. p' is the activity factor for a single voice source: $p' = \lambda_v/(\lambda_v + \mu_v)$, and $Sp' = \sum_{i=0}^{S} i p_T(i)$ is the average number of talkspurts. We have used the notation $\lceil x \rceil = \max\{0, x\}$. The video blocking probability is the probability that video is in state N_w: $P_{BW} = p_w(N_w)$. γ_v and γ_w, the voice and video throughput, are by definition given by: $\gamma_v = \sum_{\forall(k,i)\in\Omega_T} i p_{k,i}$, $\gamma_w = \sum_{\forall(k,i)\in\Omega_T} b_w k p_{k,i}$. Here Ω_T is defined as the state space for the video and voice talkspurts process: $\Omega_T \equiv \{(k,i): k = 0,1,\ldots,N_w; i = 0,1,\ldots,S\}$.

3. Complete Sharing (CS)

Before we look at the data analysis for the DLMBS scheme studied in the previous section, we consider another access scheme for the video and voice circuit switched services, called *Complete Sharing.* This has been considered at the length in [15] and [16]. We briefly describe this as follows. As in the previous section, let there be N_c channels available for the circuit switched traffic, i.e. video and voice. The circuit switched services, however, now have equal priority.

We now introduce packet switched data into the above scenario as follows. Let N_d channels be dedicated for data, giving a total of $N = N_d + N_c$ TDM channels. We further allow data to use any of the N_c unused circuit switched channels, with the circuit switched services having preemptive priority over data. The above scheme is thus another movable boundary policy and we call this *Movable Boundary with Complete Sharing* (MBCS).

As in [1], the preemptive priority of circuit switched traffic makes it independent of data. Hence the analysis would first consider the circuit switched traffic and then subsequently consider the combined packet and circuit switched scenario as done in the previous section. For the sake of consistency between the two models, a SAD model for voice is assumed. Without providing a detailed analysis for this model here, we illustrate the results by use of a simple example and summarize the results.

Figure 3 shows the state transition diagram of $\{W(t), V(t)\}$ for $N_c = 5$, $N_{vr} = 3$, $b_w = 2$, $S = 6$ and $N_w = 2$. The states in which one or more talkspurts are being frozen out are shown encircled as in figure 2. $\overline{\Omega}_k$ is defined as the following subset $\overline{\Omega}_k \subset \Omega_T$:

$$\overline{\Omega}_k \equiv \{(k,i): \forall i \in \{0,\ldots,S\} \text{ such that video is not blocked in state } (k,i)\} \quad (3.1)$$

The set of states $\overline{\Omega}_0$, $\overline{\Omega}_1$ are also shown. They correspond to the set of states for $W(t) = 0$ and $W(t) = 1$ respectively, for which video is not blocked. Since video no longer has preemptive priority over voice as was the case in figure 2, we have states between which there is only one directional transition such as between (0, 4) and (1,4). This clearly indicates that the Markov chain does not satisfy the Kolmogorov reversibility criterion given in [18], theorem 1.8, and hence the Markov chain is not reversible. Analysis of such chains is generally much more difficult since a closed form solution does not generally exist. Hence a Decomposition approximation is applied to the circuit switched processes $\{W(t), V_T(t)\}$ by first defining: $\alpha_w \equiv \mu_v/\mu_w$, and making the assumption directly from (2.1) that: $\alpha_w \gg 1$.

This provides us the basis for applying the Decomposition technique to this Markov chain. The result can be shown to be of the form:

$$p_{k,i} \approx p^*_{k,i} = p(i \mid k) p^*_w(k) = p_T(i) p^*_w(k) \quad (3.2)$$

Derivations for the approximations to the voice cutout fraction ϕ^* and blocking probability for video P^*_{BW} can be found, but due to space limitations are not shown here.

Figure 4 and 5 show graphs of ϕ, ϕ^* and P_{BW}, P^*_{BW} respectively, for varying voice and video loads. They compare the Decomposition approximations with exact results for the small system we have considered before with $N_c = 5$, $N_{vr} = 5$, $b_w = 2$, $S = 6$. The figures show that the Decomposition approximation is quite accurate when compared to exact results, which were obtained for a relatively small value of $\alpha_w = 10$.

4. Comparison of DLMBS vs. MBCS

Figures 6 to 8 compare the two schemes considered in sections 2 (DLMBS) and 3 (MBCS). The examples are the same as considered before in figure 2 and 3 respectively. The comparisons use exact solutions for both the cases. Figure 6 looks at the total throughput γ_c, defined earlier, for varying voice and video loads. The pairs of curves correspond to ρ_w = 0, 0.5, 1, 2 and 5 for continuously varying voice load ρ_v, from 0 to 5. We see that for a given video load, γ_c of MBCS is higher at low voice loads, whereas γ_c of DLMBS is higher at high voice loads. Similarly, for a fixed voice load, γ_c of

DLMBS is higher for low video loads and γ_c of MBCS is higher at high video loads.

This suggests that if γ_c is considered to be the performance criterion, then for our example, at low voice loads and high video loads MBCS is better. At high voice loads and low video loads DLMBS performs better.

5. Data Analysis

Having found the equilibrium probabilities for the two forms of circuit switched policy, we now look at the data analysis. The Decomposition approximation can now be simply written down directly from (2.9) of [1] as:

$$p_{j,k,i} \approx p^*_{j,k,i} \equiv \Pr(D(t)=j \mid W(t)=k, V_T(t)=i) \times \Pr(W(t)=k, V_T(t)=i) = p(j \mid k,i)\, p_{k,i}. \quad (5.1)$$

In the case of MBCS, $p_{k,i}$ is estimated by a Decomposition approximation $p^*_{k,i}$ as given by (3.2), and we get $p^*_{k,i,j} = p(j \mid k,i)p(i \mid k)p^*_w(k)$. This is an example of hierarchical Decomposition.

6. Results and Conclusions

We first look at the systems we have already considered for the circuit switched services given in figures 2 and 3. We add N_d=3 dedicated data channels, which gives N = 8. Thus we have $N=8$, $N_c=5$, $N_d=3$, $b_w=2$, $S=6$. For DLMBS $N_v=3$, and for MBCS $N_{vr}=5$. For DLMBS we assume $\rho_v=1$ and $\rho_w=1$. This gives ϕ = 8.07%, $P_{BW}=50\%$ and $\gamma_c=3.76$. For MBCS we set the voice load such that γ_c is the same, i.e. 3.76. This occurs when $\rho_v=1.17$, resulting in $\phi=7.76\%$, $P_{BW}=61.3\%$. Thus we note that as far as the circuit switched performance is concerned, DLMBS is superior in terms of γ_c.

The above are exact results. If we use the approximation for the MBCS scheme we get $\rho_v=1.18$, $\phi=7.1\%$, $P_{BW}=61.3\%$ for $\gamma_c=3.76$. These are quite close to the above exact results. Figure 9 shows the mean data queue length vs. data load for the two schemes under the above voice/video loading condition, using the approximation for MBCS (i.e. $\rho_v=1.18$). For each scheme we plot three curves, the Decomposition approximation, a *Zero approximation* defined in [19] which serves as a useful bound for the case when α approaches zero, and simulation results. The simulation uses $\alpha=10$, $\alpha_w=10$. As expected the simulation results lie between the two approximations.

Next we consider a larger system $N=36$, $N_c=24$, $N_d=12$, $N_w=2$, $b_w=6$. For DLMBS we set $N_v=12$, and for MBCS we choose $N_{vr}=24$. For DLMBS we let $\rho_v=1$ and $\rho_w=1$. This gives $\phi=4.73\%$, $P_{BW}=20\%$, and $\gamma_c=19.1$. For MBCS as was done earlier we increase the voice load until γ_c is approximately the same. Thus for $\gamma_c=19.1$ we set $\rho_v=1.26$, which results in $\phi=1.75\%$, $P_{BW}=55.3\%$. Thus we note that as far as the circuit switched performance is concerned, DLMBS is again superior in terms of γ_c.

Figure 10 shows the mean data queue length vs. data load for the two schemes under similar loading conditions, using the Decomposition approximation for MBCS as before. The simulation again uses $\alpha=10$, $\alpha_w=10$. We see an important result that the Decomposition and Zero approximations are drawn closer together as compared to the smaller system. This result was also seen in the voice/data systems considered in [1].

In both figures 9 and 10 we make the following observation. The difference in data performance between the two circuit switched policies is minimal when γ_c is kept constant over the two policies. We also note the existence of ρ_d^0 defined as the data utilization where all the curves seem to intersect, and that for $\rho_d < \rho_d^0$, the DLMBS scheme results in larger mean queue lengths than MBCS and for $\rho_d > \rho_d^0$, MBCS gives larger queue length. We also note that the Decomposition approximation reflects this fact, i.e. the Decomposition approximation of the mean queue length for DLMBS is greater than that of MBCS for $\rho_d < \rho_d^0$, and vice versa for $\rho_d > \rho_d^0$.

In conclusion, what we have seen in this analysis is that the Decomposition approximation can be usefully extended to the performance analysis of networks with multi-layered traffic, each having their own distinct characteristics. We saw that in certain cases, as in MBCS, the approximation can be applied in a hierarchical manner as well, while still giving reasonably accurate results both for the circuit-switched as well as for the packet-switched traffic. In addition we also compared two policy schemes for combining circuit-switched traffic with broadband video and narrowband voice.

7. References

[1] S. Ghani and M. Schwartz, "A Decomposition Approximation for the Analysis of Voice/Data Integration," in *IEEE Trans. Commun.*, vol. 42, pp. 2441-2452, July 1994.

[2] S. Ni, S. Häggman, "GPRS performance estimation in GSM circuit switched services and GPRS shared resource systems", *IEEE WCNC 1999*.

[3] Y.-D. Lin et al., "Two-stage dynamic uplink channel and slot assignment for GPRS", *IEEE ICC 2001*.

[4] H.-Y. Shin, J.-L. C. Wu, "The study of quick download mechanism in an infrastructural wireless environment", *IEEE GLOBECOM 2002*.

[5] H.-H. Liu et al., "Delay analysis of integrated voice and data service for GPRS", *IEEE Commun. Letters*, no. 8, Aug 2002.

[6] R. Bolla et al. "An adaptive neural network admission controller for dynamic bandwidth allocation", *Systems, Man and Cybernetics, Part B, IEEE Trans. on*, Vol: 28, Issue: 4, Aug 1998, page(s): 592-601.

[7] R. Bolla et al., "A hierarchical control structure for multimedia access networks", *IEEE ICC 1999.*

[8] F. Davoli, P. Maryni, "A two-level stochastic approximation for admission control and bandwidth allocation", *IEEE J. Selected Areas Commun.*, no. 2, Feb. 2000, pp. 222-233.

[9] R. Bolla et al., "Adaptive bandwidth allocation methods in the satellite environment", *IEEE ICC 2001.*

[10] S. Ghani and M. Schwartz, "A decomposition approximation for the Performance Evaluation of Non-Preemptive Priority in GSM/GPRS," submitted to *Broadband Wireless Networking Symposium 2004*, to be held in San Jose on 25-29 October 2004.

[11] S. Ghani, "Use of Approximation Methods in the Analysis of Communication Networks with Heterogeneous Traffic," *Proc. International Workshop on Frontiers of Information Technology*, Islamabad, 23-24 Dec. 2003.

[12] P. O'Reilly, "Performance Analysis of Data in Burst Switching", *IEEE Trans. on Commun.*, vol. Com-34, No. 12, December 1986, pp. 1259-1263.

[13] P. O'Reilly and S. Ghani, "Data Performance in Burst Switching when the Voice Silence Periods have a Hyperexponential Distribution", *IEEE Trans. Commun.*, vol. Com-35, Oct. 1987, pp. 1109 – 1112.

[14] C.J. Weinstein. Fractional speech loss and taker activity model for tasi and for packet-switched speech. *IEEE Trans. Commun.*, COM-26(8):1253-1257, August 1978.

[15] Belkacem Kraimeche. *Traffic Access Control Strategies in Integrated Services Digital Networks.* PhD thesis, Columbia University, 1984.

[16] B. Kraimeche, M. Schwartz, "Analysis of Traffic Access Control Strategies in Integrated Service Networks", *IEEE Trans. Commun.*, no. 10, October 1985, pp. 1085-1093.

[17] Burt M. Silverman. *Implementation, Measurement, and Analysis of Access Control Strategies for ISDN.* PhD thesis, Columbia University, 1989.

[18] Frank P. Kelly. *Reversibility and Stochastic Networks.* John Wiley and Sons, Chichester, 1979.

[19] S. Ghani, "A Decomposition Approximation for Integrated Networks", PhD Dissertation, Columbia University, 1990.

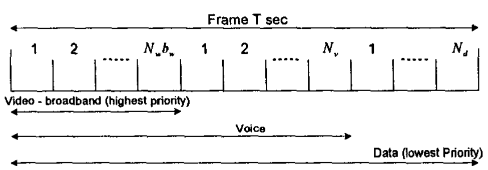
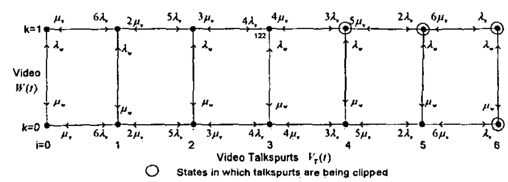
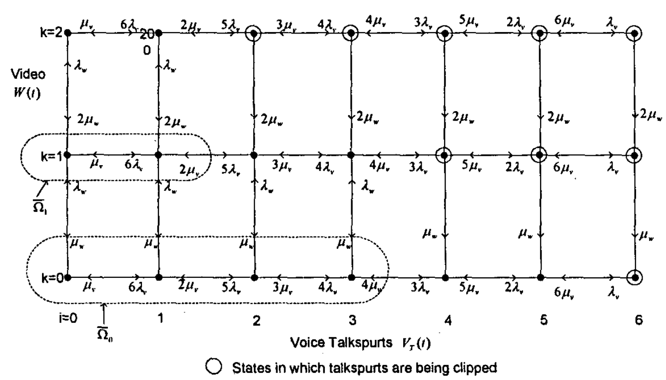


Figure 1. DLMBS frame structure



$N_c = 5$, $N_v = 3$, $N_w = 1$, $b_w = 2$, $S = 6$.

Figure 2. DLMBS State Transition Diagram for Video & Voice



$N_c = 5$, $N_{vr} = 5$, $N_w = 2$, $b_w = 2$, $S = 6$.

Figure 3. MBCS State Transition Diagram for Video & Voice

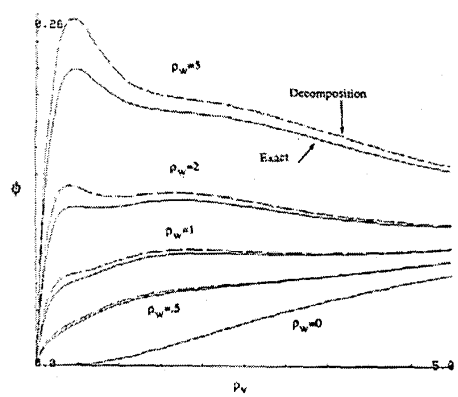


Figure 4. Decomposition and Exact ϕ for MBCS

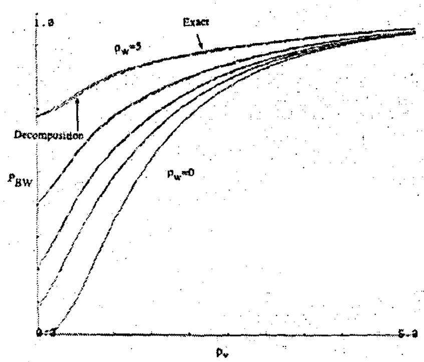


Figure 5. Decomposition and Exact P_{BW} for MBCS

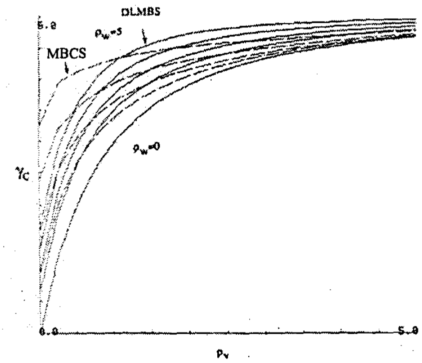


Figure 6. DLMBS vs. MBCS: Total throughput (γ_c)

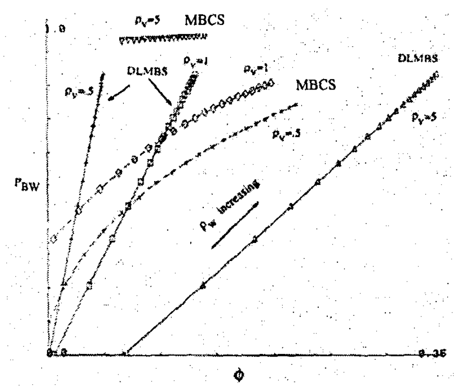


Figure 7. DLMBS vs. MBCS: P_{BW} vs. ϕ, varying ρ_w

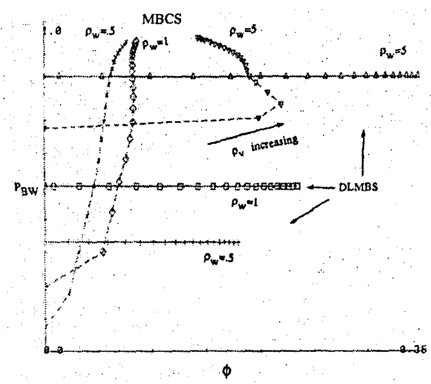


Figure 8. DLMBS vs. MBCS: P_{BW} vs. ϕ, varying ρ_v

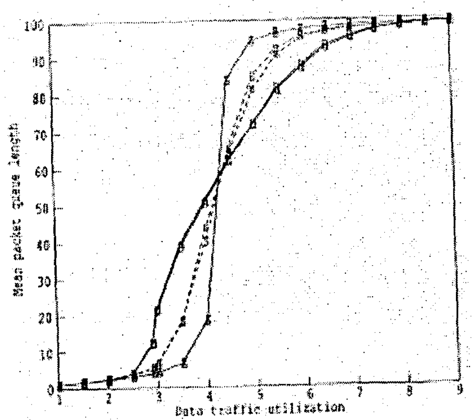
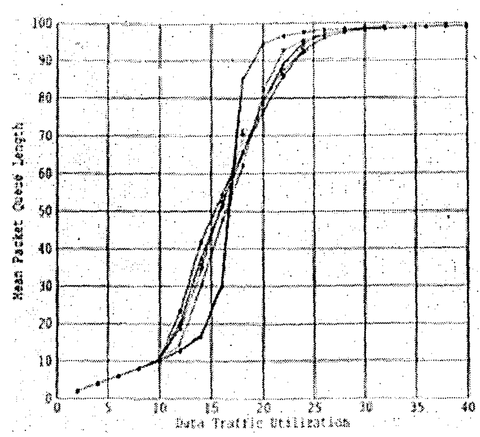


Figure 9. DLMBS vs. MBCS: EQ_D
Solid = Decomp. & Zero Approx.; Dotted=Simulation; D,Z,S = DLMBS; d,z,s=MBCS



N_c=24, N_d=12, b_w=6, S=30

Figure 10. DLMBS vs. MBCS: EQ_D for 24 channel system