Analysis of Packet-Switched Data in a New Basic Rate User-Network Interface of ISDN

Chung-Ju Chang, *Senior Member, IEEE*, and Ming-Dar Dai

***Abstract*— We present here an exact method for analyzing the queueing behavior of packet-switched data in a *new* basic rate user-network interface of ISDN. The new interface of ISDN in this paper adopts statistical multiplexing for providing integrated voice and data services. In order to properly represent the arrival of multirate data service, this interface assumes that the data packet input process is in batch Poisson. Furthermore, it is modeled by a multiple synchronous servers with priority queues and finite waiting room with complete rejection strategy. The analytical method, with an application of the residue theorem in complex variable, can accurately obtain the performance measures of data.**

I. Introduction

THE BASIC rate user-network interface of integrated services digital network (ISDN) has been recommended since 1984 [2]. The channel structure of the interface, namely, $2B + D$, contains two B channels and one D channel. The B channel is of 64 kbps and used for circuit-switched voice or data, while the D channel is of 16 kbps and used for control signaling or packet-switched data.

The initial deployment of current standard ISDN includes both circuit and packet transmission and switching; however, statistical multiplexing and packet switching are better ways for providing integrated voice and data services, considering the adaptability to new services with different bandwidth requirements, the cost savings in transmission/switching systems and in system administration, and the efficiency of network facilities [13]. Thus, it has long been recognized by many researchers that the future evolution of narrowband ISDN would eventually abandon circuit switching for voice and move to an integrated transport based on statistical multiplexing for adapting itself to broadband ISDN [5], [10], [13], [14]. As ISDN matures, the statistical multiplexing facilities and packet-switched fabric will predominate [10]. In the well-developed ISDN, voice would be packetized into packets only in talkspurt periods and be transmitted on B channel, and the remaining B channel can be utilized for transmission of data packets. The above integration no doubt yields higher utilization of the $2B + D$ channels and more flexibilities to provide services. In this paper, such a *new* basic rate user-network interface of ISDN using statistical multiplexing will be studied.

Many researchers [7]–[9], [12] have performed the queueing analyses of the integrated voice and data TDM system where the Poisson process for data packet arrival was assumed and the infinite buffer was considered; they adopted the generating function approach or the matrix analytic method. However, as the *new* basic rate user-network interface will support multirate data services in a mature ISDN, the data packets are better assumed to arrive in *batch Poisson* process. The high bit-rate data service would be a message containing a large number of data packets, while low bit-rate data service would be a message containing a small number of data packets. Notice that data packets are generated by segmentations of message. Also, the buffer capacity is practically assumed finite in this paper. Thus, the model of this interface is improved by describing it as a multiple synchronous servers system with finite waiting room. It contains voice inputs which are characterized by a Markov chain, D channel signaling input which is assumed to be a renewable Bernoulli process, and data input which is assumed to be a batch Poisson process. Since the renewable Bernoulli process for signaling packet is state independent, this model is constituted as a two-dimensional finite-state imbedded Markov chain. The analytical problem is successfully solved by applying the residue theorem in complex variable [3], [4] in deriving the transition probability of data. Finally, the performance measures of data including average message blocking probability, packet delay, and throughput are obtained exactly.

II. System Modeling

The *new* basic rate user-network interface of ISDN which has $2B + D$ output channels is shown in Fig. 1. Here, we assume that the new interface adopts statistical multiplexing function for all services to adapt itself to broadband ISDN. The interface contains three inputs which are independent among themselves. One is the voice packet coming from active voice sources which is carried only on B channel, the other is the D channel signaling packet which is carried only on D channel, and another is the data packet which is carried on the remaining B channel or D channel. The data packet could be in ATM (asynchronous transfer mode) packet format which contains 5 bytes for header and 48 bytes for payload [14]. This $2B + D$ link is timely partitioned into frames of

Paper approved by H. Takagi, the Editor for Queueing and Network Performance of the IEEE Communications Society. Manuscript received May 1, 1992; revised September 25, 1992 and December 17, 1992.

C.-J. Chang is with the Department of Communication Engineering and Center for Telecommunications Research, National Chiao Tung University, Hsinchu 300, Taiwan, China.

M.-D. Dai is with Network Planning Division, Telecommunications Laboratories, Ministry of Communications, Chungli 320, Taiwan, China.

IEEE Log Number 9404742.

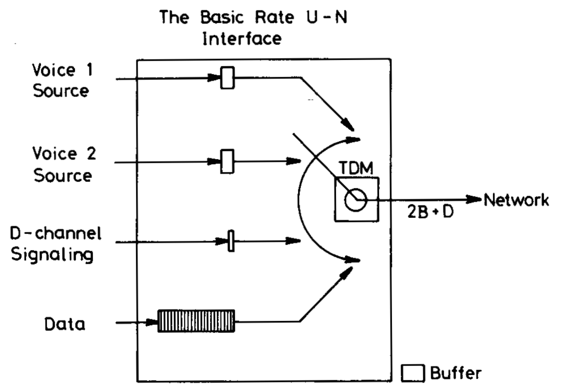


Fig. 1. The configuration of the new basic rate user-network interface of ISDN.

equal duration T which is defined to be the packetization time of a voice packet during talkspurt. In each frame, there are $M = 2N_B + N_D$ time slots in which two N_B slots belong to the two B channels and N_D slots belong to the D channel. The voice packet has the first priority to be served by the B channel and takes N_B slots; the signaling packet has the first priority to be served by the D channel and takes N_D slots. If there are no other voice or signaling packets to be served in a frame, the remaining slots are then allocated to the data packets. Each data packet takes one slot. The seizure and the release of the $2B+D$ channels are assumed to be carried out simultaneously at the beginning and at the end of a frame, respectively. Voice packet, D channel signaling packet, or data packet, will be served immediately at the beginning of the frame if allocated with time slots, and will depart from the interface at the end of that frame. Packets arriving during a frame are buffered but not served even though there are free slots in that frame. In this way, the interface is modeled as a multiple synchronous servers queue in which each server works synchronously with the frame clock.

The model for an active voice source at the interface assumes that both the talkspurt and the silence durations are exponentially distributed with means $1/\lambda$ and $1/\mu$, respectively [1], [6]. Since an active voice source generates one packet every packetization time T during talkspurt but none during silence, the number of frames contained in the talkspurt and silence durations are geometrically distributed with means $1/\lambda T$ and $1/\mu T$, respectively, if $\lambda T < 1$ and $\mu T < 1$. And the voice correlation coefficient which describes the correlation behavior between two consecutive voice packets is $w = 1 - \lambda T - \mu T$ [9]. The longer the talkspurt and silence durations are, the higher the voice correlation coefficient is. The D channel signaling considered here is limited to the control signaling and network management information. In general, its activity is very low [11]. We assume that the arrival of signaling packet in a frame is a renewable Bernoulli process with "ι" denoting the probability of no packet arrivals. And the input process of data packets, generated from multirate, multimedia data services, is modeled as a batch Poisson with average message arrival rate λ_d. The message size of data service, in unit of data packet, is assumed to be a positive integer-valued random variable with an arbitrary probability distribution $g(\cdot)$.

For each active voice source and D channel signaling inputs, a buffer with length 1 is provided separately. It is because a voice call in talkspurt generates one voice packet per frame and the voice packet will immediately be transmitted at the beginning of the next frame; and each signaling packet arriving before the end of a frame is transmitted at the beginning of the next frame. But for the data input, a finite buffer with capacity N is assumed. When a data message arrives and the free space in the buffer is not sufficient enough to accommodate the entire message, an overflow of the buffer will occur and the message will be completely rejected to guarantee the integrity of message.

III. Queueing Analysis

Define $[V_x, Q_x]$ as the system state in steady state where V_x and Q_x denote the number of voice packets transmitted and the queue length of data packets at the beginning of the xth frame, respectively. Note that $0 \le V_x \le \zeta$ and $0 \le Q_x \le N$ where ζ is the number of active voice sources in progress and $0 \le \zeta \le 2$. And define S_x as the number of signaling packets transmitted at the beginning of the xth frame, $Sx \in \{0, 1\}$. Also, let π_{ij} be the probability of state $[i, j]$. The state-transition equations can then be expressed as

$$\begin{aligned}\pi_{ij} = &\left[\sum_{m=0}^{\zeta}\sum_{n=0}^{N}\pi_{mn} \times P^{0}_{ij,\,mn}\right] \times \Pr\{S_x = 0\}\\ &+ \left[\sum_{m=0}^{\zeta}\sum_{n=0}^{N}\pi_{mn} \times P^{1}_{ij,\,mn}\right] \times \Pr\{S_x = 1\}\\ &\text{for} \quad i \in \{0, \cdots, \zeta\} \quad \text{and} \quad j \in \{0, 1, \cdots, N\} \qquad (1)\end{aligned}$$

where $P^{k}_{ij,\,mn}$ is the conditional state-transition probability from state $[m, n]$ to state $[i, j]$, given that the channel capacity for voice and data has $M - kN_D$ slots and $k \in \{S_x\}$.

In order to solve the stationary state probability π_{ij} of (1), the conditional state-transition probability $P^{k}_{ij,\,mn}$ must first be obtained. Let $p_{i,\,m}$ be the probability of having "i" voice packets transmitted in a frame while "m" voice packets transmitted in the previous frame, where $0 \le i, m \le \zeta$. And let $q_{c,\,r}$ be the probability of having "c" data packets accepted during a frame, if there are "r" free buffer spaces left for data packets at the beginning of that frame, where $0 \le c \le r \le N$. The queueing behavior of data between two adjacent frames can be characterized by

$$Q_{x+1} = [Q_x - (M - V_x \cdot N_B - S_x \cdot N_D)]^{+} + d_x \qquad (2)$$

where $[\cdot]^{+}$ denotes the maximum of 0 or the argument, $[Q_x - (M - V_x \cdot N_B - S_x \cdot N_D)]^{+}$ is the number of the buffered data packets arriving before frame x but not served yet at the end of the xth frame, and d_x is the number of data packets accepted by the system during the xth frame. Note that the system has a total of M slots at the beginning of a frame in which $V_x \cdot N_B$ and $S_x \cdot N_D$ slots are first allocated to voice and signaling, respectively, and then the rest slots to data. For a given $S_x = k$ and system state transited from $[m, n]$ to $[i, j]$, c and r in $q_{c,r}$ can then be determined by $c = j - [n - (M - m \cdot N_B - k \cdot N_D)]^+$, and $r = N - n$. Since the input processes of voice and data are independent of each other, the $P^k_{ij,mn}$ can be expressed as

$$P^k_{ij,mn} = p_{i,m} \cdot q_{j-[n-(M-m \cdot N_B - k \cdot N_D)]^+,\, N-n}. \quad (3)$$

A. The Voice Transition Probability $p_{i,m}$

As the durations of takspurt and silence are assumed to be exponentially distributed with means $1/\lambda$ and $1/\mu$, respectively, and as the frame time T is assumed to be much smaller than $1/\lambda$ and $1/\mu$, the probability that a call remains in talkspurt (in silence) between two consecutive frames is $1 - \lambda T$ $(1 - \mu T)$, and the probability that a call transits from talkspurt to silence (from silence to talkspurt) is λT (μT). Then the $p_{i,m}$ can easily be obtained by

$$p_{i,m} = \sum_{\kappa=0}^{\min(i,m)} [C^m_\kappa \cdot (1 - \lambda T)^\kappa \cdot (\lambda T)^{m-\kappa}] \times [C^{\zeta-m}_{i-\kappa} \cdot (\mu T)^{i-\kappa} \cdot (1 - \mu T)^{(\zeta-m)-(i-\kappa)}] \quad (4)$$

where $C^m_\kappa = \frac{m!}{\kappa! \cdot (m-\kappa)!}$.

B. The Data Transition Probability $q_{c,r}$

It is expected that an arrival message in the sequence of possible arrival messages during a frame may be either accepted or blocked owing to its inappropriate message size. For examples, if $r = 10$, $c = 6$, and the sequence of possible arrival messages during a frame, expressed in number of data packets and in the order of their appearance, is $\Psi = \{3, 1, 7, 2, 5\}$, then the first, the second, and the fourth messages are acceptable but the third and the fifth (with message size 5) messages are blocked. However, if $r = 10$, $c = 6$, and $\Psi = \{5, 6, 1\}$, then the first (with message size 5) and the third messages are acceptable but the second message is rejected. The derivation of the probability $q_{c,r}$ is a combinatorial problem. There is a difficulty that is arisen in finding out which arrivals able to be accepted by the system among all the possible arrival sequences in order to obtain the exact result. This complicated problem can be successfully solved by applying the residue theorem in complex variable [3], [4]. Nevertheless, in this paper, for computational tractability, some modifications are made so that the multiple-ordered derivations which often occur in the numerical computation of the residue values of poles at $z = 0$ can be avoided. The detailed derivation of $q_{c,r}$ is given in the Appendix.

Once $p_{i,m}$ and $q_{c,r}$ are given, the state probability π_{ij} can be obtained from (1) and the summability-to-one criterion. This analysis is rather elegant and can also be applied to the cases of primary rate user-network interface ($23B + D$ or $30B + D$) or $2B +$ big$D$ in the future.

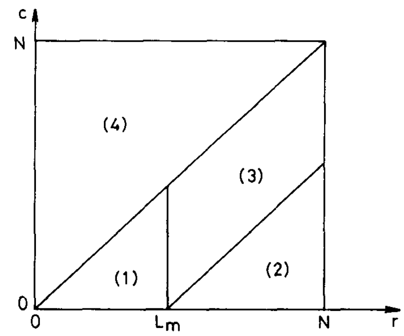


Fig. 2. The four regions for the derivation of data transition probability $q_{c,r}$.

C. The Performance Measures

To derive the average message blocking probability B_M, let $\pi_{ij}(t)$ be the stationary probability of having "i" voice packets and "j" data packets in the system at "t" seconds after the beginning of a frame where $t \in (0, T)$. Obviously, $\pi_{ij}(t = 0^+)$ is just π_{ij}. Also let $\pi^d_j(t)$ be the marginal probability of $\pi_{ij}(t)$, i.e., $\pi^d_j(t) = \sum^\zeta_{i=0} \pi_{ij}(t)$ for $0 \le j \le N$. Similarly, $\pi^d_j(t = 0^+)$ is just π^d_j where π^d_j is the marginal probability of π_{ij}. Because there is no departures during $(0, t)$ for $t \in (0, T)$, $\pi^d_j(t)$ satisfies

$$\pi_u j^d(t) = \sum_{k=0}^{j} q_{j-k,\, N-k}(t) \cdot \pi^d_k(0^+) \quad \text{for} \quad 0 \le j \le N. \quad (5)$$

Define $B_M(t)$ as the probability that a data message will be blocked if it arrives at "t" seconds after the beginning of a frame. Owing to the complete rejection strategy adopted here, this data message will be blocked if its message size is greater than the remaining buffer space at time t. Thus, $B_M(t)$ is given by

$$B_M(t) = \sum_{j=0}^{N} \pi^d_j(t) \cdot G^c(N - j) \quad \text{for} \quad t \in (0, T) \quad (6)$$

where $G^c(\cdot)$ is defined as

$$G^c(\kappa) = \begin{cases} 1, & \kappa = 0, \\ 1 - [g(1) + g(2) + \cdots + g(\kappa)], & \kappa \ge 1. \end{cases} \quad (7)$$

The message arrivals form a Poisson process; any arbitrarily selected message will arrive at any time in a frame with an equal probability of $1/T$. Hence, the average message blocking probability B_M is obtained by

$$B_M = \frac{1}{T}\int_0^T B_M(t)\,dt = \sum_{j=0}^{N}\sum_{k=0}^{j} [\overline{q}_{j-k,\, N-k} \cdot \pi^d_k(0^+) \cdot G^c(N - j)] \quad (8)$$

where

$$\overline{q}_{j-k,\, N-k} \equiv \frac{1}{T}\int_0^T q_{j-k,\, N-k}(t)\,dt. \quad (9)$$

To derive the mean data packet delay D_P, let N_Q be the average number of data packets in the system observed at arbitrary time t, and let $\lambda_{d,\,\mathrm{eff}}$ be the average effective entering rate of data packets at time t, where $t \in (0, T)$. They can be obtained by

$$N_Q = \sum_{j=1}^{N} j \cdot \overline{\pi}_j^d, \tag{10}$$

and

$$\lambda_{d,\,\mathrm{eff}} = \lambda_d \cdot \sum_{j=0}^{N-1} \overline{\pi}_j^d \cdot \left[\sum_{k=1}^{N-j} k \cdot g(k)\right] \tag{11}$$

where $\overline{\pi}_j^d = \frac{1}{T}\int_0^T \pi_j^d(t)\,dt$. From Little's formula D_P in units of frames, can be obtained by

$$D_P = N_Q / \lambda_{d,\,\mathrm{eff}}. \tag{12}$$

Throughput S is defined as the average number of data packets completing service per frame. As mentioned previously, voice and D channel signaling have higher priorities to transmit via B channel and D channel, respectively; data, signaling and voice packets take 1, N_D, and N_B slots, respectively. Thus, throughput S is given by

$$S = \sum_{i=0}^{\zeta}\sum_{j=1}^{N}\sum_{k=0}^{1} \mathrm{Min}\,(j,\, M - i \cdot N_B - k \cdot N_D) \times \pi_{ij} \times P_r\{S_x = k\}. \tag{13}$$

Incidentally, the mean message delay D_M, defined as the average time from the instant an acceptable message arrives to the instant the last packet of the message leaves the system, is given by

$$D_M = \sum_{j=0}^{N-1} \overline{\pi}_j^d \cdot \sum_{i=1}^{N-j} \frac{g(i)}{[1 - G^c(N-j)]} \cdot \left[0.5 + (i+j) \cdot \frac{1}{S}\right] \tag{14}$$

where $[1 - G^c(N-j)]$ is a normalization factor since the message delay is considered for unblocked message arrivals; the item of 0.5 denotes the residual period measured from the time of message arrival to the beginning of the next frame; and S is the throughput of data.

IV. Numerical Examples and Discussions

In all the numerical examples shown in this section, $T = 30$ ms, $N_B = 4$, $N_D = 1$, $N = 50$, $\iota = 0.9$, and $\zeta = 2$ are assumed. Two probability distributions of message size are investigated. The first one, called truncated geometric distribution (TG) with parameter p, denotes that low bit-rate data traffic is more likely to happen than high bit-rate data traffic. Its $g(\cdot)$ is given by

$$g(k) = \begin{cases} C \cdot p \cdot (1-p)^k, & 1 \le k \le L_m, \\ 0, & \text{otherwise} \end{cases} \tag{15}$$

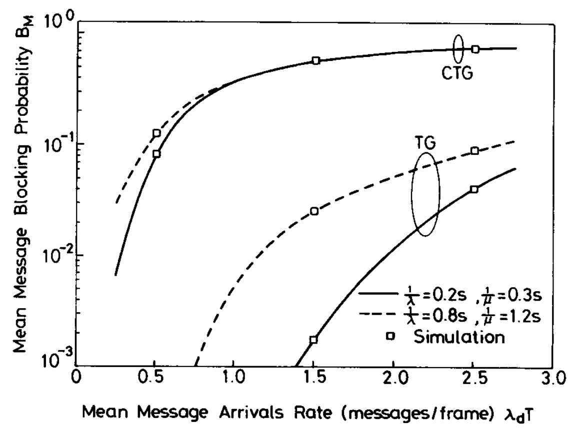


Fig. 3. The mean message blocking probability B_M versus the mean message arrival rate $\lambda_d T$.

where L_m is the maximum number of packets for data message, and C is a normalization constant such that $g(1) + \cdots + g(L_m) = 1$. Clearly, C is given by

$$C = \left[\sum_{k=1}^{L_m} p \cdot (1-p)^k\right]^{-1} = [(1-p)\{1-[1-p]^{L_m}\}]^{-1}. \tag{16}$$

In contrast with TG, the second one, called the contruncated geometric distribution (CTG), denotes that low bit-rate traffic is less likely to happen than high bit-rate data traffic. We also assume that the probability of having an arrival with a size of L_m data packets in CTG is equal to that of having an arrival with a size of 1 data packet in TG; the probability of having an arrival with size of $L_m - 1$ data packets in CTG is equal to that of having an arrival with size of two data packets in TG and so on. In the following numerical examples, $L_m = 10$ and $p = 0.5$ for both TG and CTG distributions are assumed. Hence, the mean packet length per message arrival is two packets for TG but nine packets for CTG. Also, two sets of talkspurt and silence durations, $1/\lambda = 0.2$ s, $1/\mu = 0.3$ s, with the voice correlation coefficient $w = 0.75$ and $1/\lambda = 1.2$ s, $1/\mu = 1.8$ s with $w = 0.96$, are considered to investigate the voice correlation effect on performance measures of data.

Figs. 3–5 show average message blocking probability B_M, average packet delay D_P, and throughput the data S versus message arrival rate $\lambda_d T$, respectively. First, we observe that, for both TG and CTG, all of these three performance measures increase with the increment of the message arrival rate. These are intuitively reasonable. We also observe that CTG has larger blocking probability, packet delay and throughput than TG. It is because CTG has a heavier traffic load than TG for any given $\lambda_d T$.

Note that longer average talkspurt and silence periods will result in a higher voice correlation. From Figs. 3 and 4, we also find that both B_M and D_P increase as the voice correlation increases. Intuitively, a longer talkspurt duration makes more data packets to queue due to insufficient time slots assigned, and thus leads to a larger queue length. On the other hand,

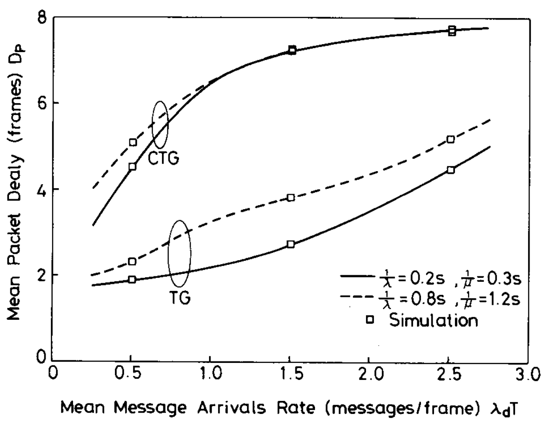


Fig. 4. The mean packet delay D_P versus the mean message arrival rate $\lambda_d T$.

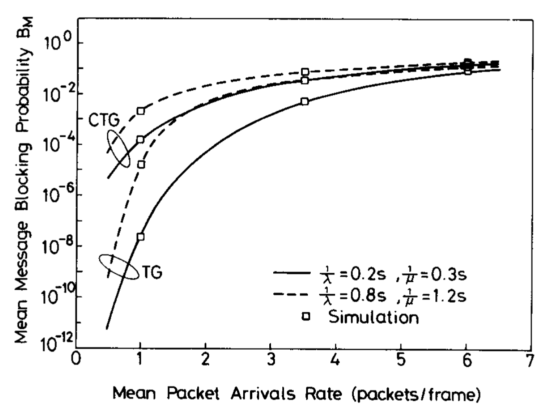


Fig. 6. The mean message blocking probability B_M versus the mean packet arrival rate.

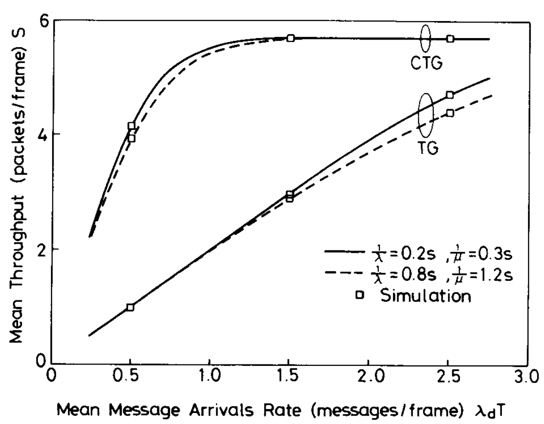


Fig. 5. The mean throughput S versus the mean message arrival rate $\lambda_d T$.

a longer silence duration can only decrease a little in queue length because the data queue length is inherently small in the silence period. Consequently, the *overall* average queue length increases as the voice correlation increases, although voice activity $\mu/(\lambda+\mu)$ is kept constant. This was also observed by Li and Mark [9] and by Sriram, Varshney, Shanthikumar [12] in infinite-buffer cases. Hence, we have the results.

From Fig. 5, however, we also observe that throughput S decreasesas the voice correlation increases. The behavior is, in fact, consistent with those shown in Figs. 3 and 4 and is explained as follows. Since the voice activity is kept constant, the average number of time slots remaining for data packets does not change. However, it is believable that the number of time slots remaining for data packets during a talkspurt period is *few* and then almost all of the accepted data packets are queued. This explains that the longer data queue resulted from longer talkspurt periods does not result in its throughput higher. On the other hand, the number of time slots remaining for data packets in a frame during a silence period is *many* and the data buffer will thus become empty most of time near the end of the silence period. It means that, for a finite waiting queue, a longer silence period wastes more time slots. Thus, the data throughput for higher voice correlation is smaller.

Also, the performance measures versus the mean packet arrival rate are illustrated here. For a given mean packet arrival rate, obviously, the message arrival rate λ_d of CTG would be smaller than that of TG. Fig. 6 shows the average message blocking probability B_M versus the mean packet arrival rate. The blocking probability also tends to increase as the packet arrival rate increases. In addition, we observe that CTG still has larger B_M than TG although their traffic intensities are the same for a given mean packet arrival rate. It is due to the fact that messages of CTG have larger probabilities with more packets than those of TG.

As far as packet delay D_P and throughput S shown in Figs. 7 and 8 are concerned, we see that for a given *low* mean packet arrival rate, say below 3 packets/frame, both TG and CTG have almost the same throughput. It is because most arrivals can be accepted to receive service by the system in the low traffic situation. The accepted messages of TG, however, have lower probabilities with more packets than those of CTG. This implies that the delay for CTG will be larger than that for TG. On the other hand, for a given *high* mean packet arrival rate, the mean system occupancy is almost full most of the time. Only shorter messages have the probabilities to be accepted; many longer messages are blocked, especially for CTG. As a result, the effective packet entering rate of CTG being less than that of TG. Hence, D_P and S for CTG are smaller than those of TG. Furthermore, we observe that the high voice correlation coefficient similarly damages the queueing behaviors of data in these figures.

In addition to the previous numerical computation of the theoretical results, some simulations are also carried out to support the analysis. There is a very good agreement between analysis and simulation.

V. Concluding Remarks

An exact method for analyzing the queueing behavior of packet-switched data in a new basic rate user-network

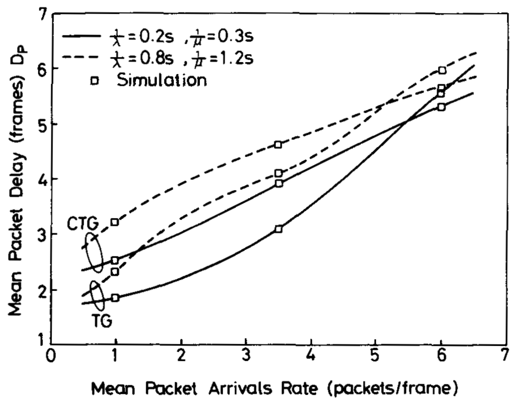


Fig. 7. The mean packet delay D_P versus the mean packet arrival rate.

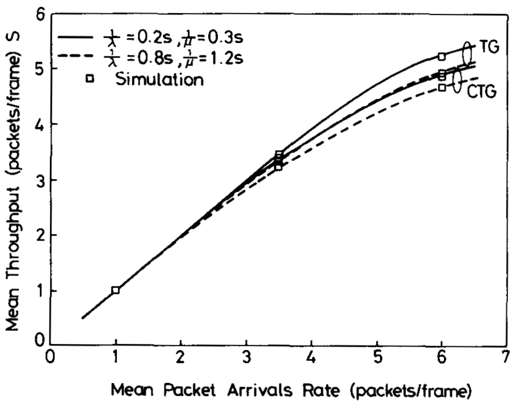


Fig. 8. The mean throughput S versus the mean packet arrival rate.

interface of ISDN is presented in this paper. This ISDN interface assumes statistical multiplexing for both voice and data services and is sophisticatedly modeled as a multiple synchronous servers with priority queues as well as finite waiting room with complete rejection strategy. The voice input is modeled as a discrete-time Markov chain which can characterize both the steady and the correlation behavior of voice. The data input is assumed to be a batch Poisson process which can characterize the multirate data services. And the signaling input is approximately modeled as a Bernoulli process. The analytical problem is successfully solved by an application of the residue theorem in complex variable. The performance measures of data including message blocking probability, packet delay and throughput are obtained exactly.

Appendix

A. Derivation of Transition Probability $q_{c,r}$

We here derive the transition probability $q_{c,r}(t)$which is defined as the probability of having "c" data packets accepted during $(0, t)$ if there are "r" free buffer spaces left at the beginning of the frame. Obviously, $q_{c,r} = \lim_{t\to T} q_{c,r}(t)$. In the derivation of $q_{c,r}(t)$, four regions shown in Fig. 2 are considered. They are i) Region 1, $0 \le c \le r < L_m$, ii) Region 2, $r - c \ge L_m$, iii) Region 3, $r \ge L_m$ and $0 \le r - c < L_m$, and iv) Region 4, $c > r$ where L_m is the maximum size of an arriving message. Owing to the additional considerations of Regions 2 and 3, the computational problem of multiple-ordered derivations can be avoided.

i) Region 1: $0 \le c \le r < L_m$.

In this region, the typical arrival sequence is

$$\beta_0\alpha_1\beta_1\alpha_2\beta_2\cdots\alpha_{\kappa-1}\beta_{\kappa-1}\alpha_\kappa\beta_\kappa$$

where α_j's are the acceptable arrivals which may contain more than one, and β_j's are the blocked arrivals which may contain more than one or empty, $0 \le j \le \kappa$. For given c and r, $q_{c,r}(t)$ is obtained by [3], [4]

$$\begin{aligned} q_{c,r}(t) &= \frac{1}{2\pi i} \oint \frac{A_{c,r}(z)\cdot \exp(-\lambda_d t(1-z))}{B_{c,r}(z)}\, dz \\ &= \Sigma \operatorname{Res}\left[\frac{A_{c,r}(z)\cdot \exp(-\lambda_d t(1-z))}{B_{c,r}(z)}\right], \end{aligned} \quad \text{(A.1)}$$

where $i \equiv \sqrt{-1}$

$$A_{c,r}(z) = \begin{cases} 1; & \text{if } c = 0, \\ g(1); & \text{if } c = 1, \\ \sum_{j=2}^{c} g(j)\cdot A_{c-j,r}(z) & \\ \quad \cdot\left[\prod_{\kappa=c-j+1}^{c-1}(z - G^c(r-\kappa))\right] & \\ \quad + g(1)\cdot A_{c-1,r}(z); & \text{if } c \ge 2. \end{cases} \quad \text{(A.2)}$$

and

$$B_{c,r}(z) = \prod_{\kappa=0}^{c}(z - G^c(r-\kappa)) \quad \text{(A.3)}$$

$g(\cdot)$ is the probability distribution of message size and $G^c(\cdot)$ is defined in (7). The closed integration contour on complex z in (A.1) is carried out along a circle whose radius is greater than any root of $B_{c,r}(z)$.

ii) Region 2: $r - c \ge L_m$.

Since the remaining space of the buffer at the end of a frame, $r - c$, is still larger than or equal to the maximum size of arrivals in this region, there is no blocked arrivals in this frame. Thus the typical arrival sequence is

$$\alpha_1 \quad \alpha_2\cdots\alpha_{\kappa-1} \quad \alpha_\kappa.$$

Then $q_{c,r}(t)$ can be obtained by

$$q_{c,r}(t) = \sum_{\kappa=\lfloor c/L_m\rfloor}^{c} g^{(\kappa)}(c)\cdot\frac{(\lambda_d t)^\kappa}{\kappa!}\cdot\exp(-\lambda_d t) \quad \text{(A.4)}$$

where $g^{(\kappa)}(c)$ is the κ-fold discrete convolution of $\{g(\cdot)\}$ with itself and $\lfloor\cdot\rfloor$ denotes the smallest integer greater than or equal to the argument. The upper bound of κ in the summation of (A.4) denotes the possible greatest number of arrivals accepted. It is just the case of $\kappa = c$ and $\alpha_1 = \alpha_2 = \cdots = \alpha_c = 1$. The lowest bound of κ denotes the possible least number of arrivals accepted. It is just the case that all arrivals or all arrivals except the last one are with the size of L_m packets.

iii) Region 3: $r \geq L_m$ and $0 \leq r - c < L_m$.

This is the case that the remaining buffer space is larger than or equal to the maximum size of arrivals at the beginning of a frame but it is insufficient to accommodate the largest arrival at some time in the middle of the frame. The typical arrival sequence is

$$\alpha_1\alpha_2\cdots\alpha_1\alpha_{i+1}\beta_{i+1}\alpha_{i+2}\beta_{i+2}\cdots\alpha_\kappa\beta_\kappa$$

where $\alpha_1\cdots\alpha_{i+1}$ is an acceptable subsequence as the typical arrival sequence in Region 2 and $\beta_{i+1}\alpha_{i+2}\beta_{i+2}\cdots\alpha_\kappa\beta_\kappa$ is a subsequence as the typical arrival sequence in Region 1. Consequently, the arrival sequence in Region 3 can be regarded as a combination of an arrival sequence in Region 2 followed by an arrival sequence in Region 1. The derivation of $q_{c,r}(t)$ in this region is then given in the following.

Let t_ξ be the time between the $(\xi - 1)$st and the ξ th arrival in a frame, $\xi \geq 2$, and t_1 be the time between the beginning of the frame and the first arrival. The acceptable subsequence $\alpha_1\cdots\alpha_i\alpha_{i+1}$ which is the former part of the typical arrival sequence in this region is subject to $\sum_{j=1}^{i}\alpha_j = x$, $\alpha_{i+1} = y$, $x \leq r - L_m$, $x + y > r - L_m$, and $0 < \tau = \sum_{\xi=1}^{i+1} t_\xi < t$. The conditions of $x \leq r - L_m$ and $x + y > r - L_m$ state that the arrival sequence $\alpha_1\cdots\alpha_i\alpha_{i+1}\beta_{i+1}\alpha_{i+2}\beta_{i+2}\cdots\alpha_\kappa\beta_\kappa$ first has an acceptable arrival subsequence as in Region 2 before the arrival of α_{i+1} and then has an arrival subsequence as in Region 1 after the arrival of α_{i+1}. Owing to the memoryless property of Poisson arrivals, we define a probability of having "c" data packets accepted during $(0, t)$ if there are "r" free buffer spaces left at the beginning of the frame and the first acceptable subsequence of the typical arrival sequence totally has $x + y$ packets in $[0, \tau)$ and $x \leq r - L_m$, $x + y > r - L_m$. We denote this probability by $q_{c,r}(x, y, \tau)$ and obtain it by

$$q_{c,r}(x, y, \tau) = \sum_{i=\lfloor x/L_m\rfloor}^{x} g^{(i)}(x)\cdot g(y)\cdot\left[\lambda_d\cdot\frac{(\lambda_d\tau)^i}{i!}\cdot e^{-\lambda_d\tau}\right] \times\left\{\Sigma\,\mathrm{Res}\left[\frac{A_{c-x-y,\,r-x-y}(z)}{B_{c-x-y,\,r-x-y}(z)}\cdot\exp\left[-\lambda_d\cdot(t-\tau)(1-z)\right]\right]\right\} \qquad \text{(A.5)}$$

where the $\left[\lambda_d\cdot\frac{(\lambda_d\tau)^i}{i!}\cdot e^{-\lambda_d\tau}\right]$ is the Erlang type-i distribution. Finally, we can obtain the $q_{c,r}(t)$ in this region by summing up all possible values of x, y and then time-averaging of τ over $(0, t)$ of $q_{c,r}(x, y, \tau)$. Thus, we have

$$q_{c,r}(t) = \sum_{x=\mathrm{Max}(0,r-2L_m+1)}^{r-L_m}\ \sum_{y=\mathrm{Max}(1,r-L_m-x+1)}^{\mathrm{Min}(L_m,c-x)}\ \sum_{i=\lfloor x/L_m\rfloor}^{x} g^{(i)}(x)\cdot g(y)\cdot\left\{\Sigma\,\mathrm{Res}\left\{\frac{A_{c-x-y,\,r-x-y}(z)}{B_{c-x-y,\,r-x-y}(z)}\cdot\exp\left[-\lambda_d t(1-z)\right]\times\frac{1}{z^{i+1}}\cdot\left[1-\sum_{k=0}^{i}\frac{(\lambda_d tz)^k}{k!}\cdot\exp\left(-\lambda_d tz\right)\right]\right\}\right\} \qquad \text{(A.6)}$$

where the upper bound of x in the summation of (A.6) comes from the fact that the buffer at least has to have L_m left after the acceptance of $\alpha_1\cdots\alpha_i$. Note that $\sum_{j=1}^{i}\alpha_j = x$. The lower bound comes from the facts that $x \geq 0$ and the buffer enters into Region 1 after the arrival of α_{i+1}. The latter fact results in $x > r - L_m - y \geq r - 2L_m$. The upper bound in the summation of y results from the facts that the maximum length of α_{i+1} cannot be greater than L_m and $x + y$ cannot be greater than c, while the lower bound results from the similar reasons that the lower bound of x does.

iv) Region 4: $c > r$.

Since the remaining buffer space "r" is insufficient toaccept a message of which length "c" is greater than r. Thus, $q_{c,r}(t) = 0$.

Acknowledgment

The authors wish to thank the anonymous reviewers for their constructive comments in improving the manuscript.

References

[1] P. T. Brandy, "A statistical analysis of on-off patterns in 16 conversations," *Bell Syst. Tech. J.*, vol. 47, pp. 73–81, Jan. 1968.

[2] CCITT Study Group XVIII I Series Recommendations, 1984.

[3] J. F. Chang and R. F. Chang, "The application of the residue theorem to the study of a finite queue with batch Poisson arrivals and synchronous servers," *SIAM J. Appl. Math.*, vol. 44, no. 3, pp. 646–656, June 1984.

[4] C. J. Chang and J. F. Chang, "The effect of routing and buffer sharing on the behavior of a finite queue with batch Poisson inputs and synchronous servers," *SIAM J. Appl. Math.*, vol. 47, no. 2, pp. 352–366, Apr. 1987.

[5] T. M. Chen and D. G. Messerschmitt, "Integrated voice/data switching," *IEEE Commun. Mag.*, vol. 26, pp. 16–26, June 1988.

[6] J. N. Daingle and J. S. Langford, "Models of analysis of packet voice communications systems," *IEEE J. Select. Areas Commun.*, vol. SAC-4, pp. 847–855, 1986.

[7] 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 Commun.*, vol. SAC-4, pp. 856–867, Sept. 1986.

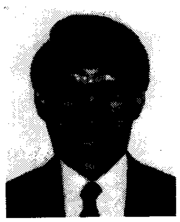
[8] S.-Q. Li, "Traffic characterization for integrated services network," *IEEE Trans. Commun.*, vol. 38, pp. 1231–1243, Aug. 1990.

[9] S.-Q. Li and J. W. Mark, "Performance analysis of an integrated subscriber loop: A basic access to ISDN," in *Proc. IEEE GLOBECOM'86*, Dec. 1986, pp. 27.5.1–27.5.5.

[10] M. Malek, "Integrated voice and data communications overview," *IEEE Commun. Mag.*, vol. 26, pp. 5–15, June 1988.

[11] H. Murata, "Traffic characteristics of control signals in an ISDN switching system," *Comput. Networks ISDN Syst.*, no. 10, pp. 203–210, 1985.

[12] K. Sriram, P. K. Varshney, and J. G. Shanthikumar, "Discrete-time analysis of integrated voice/data multiplexers with and without speech activity detectors," *IEEE J. Select. Areas Commun.*, vol. SAC-1, pp. 1124–1132, Dec. 1983.
[13] J. S. Turner, "Design of an integrated services packet network," *IEEE J. Select. Areas Commun.*, vol. SAC-4, pp. 1373–1380, Nov. 1986.
[14] P. E. White, "The role of the broadband integrated services digital network," *IEEE Commun. Mag.*, vol. 29, pp. 116–119, Mar. 1991.



Chung-Ju Chang (S'81–M'85–SM'94) received the B.E. and M.E. degrees in electronics engineering from National Chiao Tung University, Hsinchu, Taiwan, in 1972 and 1976, respectively, and Ph.D. degree in electrical engineering from National Taiwan University in 1985.

From 1976 to 1988 he was with Telecommunication Laboratories, Directorate General of Telecommunications, Ministry of Communications, Republic of China, as a Design Engineer, Supervisor, Project Manager, and then Division Director. There, he was involved in designing digital switching system, ISDN user-network interface, and ISDN service and technology trials. In the meantime, he also acted as a Science and Technical Advisor for Minister of Ministry of Communications from 1987 to 1989. In August 1988, he joined the faculty of the Department of Communication Engineering and Center for Telecommunications Research, National Chiao Tung University, where he is currently Professor and Director for the Institute of Communication Engineering.. His research interests include performance analysis, broadband networks, and mobile radio networks.

Dr. Chang is a member of the Chinese Institute of Engineers (CIE).

Ming-Dar Dai received the B.E. degree in electrical engineering from National Sun Yet San University, Taiwan, in 1986, and the M.E. degree in communication engineering from National Chiao Tung University, Taiwan, in 1988.

He is now working on switching software development in SIEMENS systems limited, Taipei, Taiwan. Before this, he joined developments of data communications and telecommunications products or prototypes using protocols such as X.25, IN/STP, CCS7, and ISDN in Sertek International Inc. and in Telecommunication Laboratories, Taiwan. His research and development interests include broadband network, protocol design, and performance analysis.