ELSEVIER

Computer Communications 21 (1998) 243–253

computer communications

Bandwidth assignment paradigms for broadband integrated voice/data networks[1]

Maria C. Yuang*, Yieh R. Haung

Department of Computer Science and Information Engineering, National Chiao Tung University, Hsinchu 30050, Taiwan

Received 5 November 1994; accepted 2 February 1996

Abstract

Broadband Integrated Services Digital Networks (BISDNS)/Asynchronous Transfer Mode (ATM) are expected to support diverse applications demanding different bandwidth and Quality of Services (QoSs). Particularly for integrated voice and data networks, voice traffic results in call losses and data traffic suffers from longer delays should networks have insufficient bandwidth. The main goal of the paper is to analytically determine optimal bandwidth allocated to voice and data by means of a queueing model with heterogeneous arrivals and multiple designated channels. The accuracy of analytical results is confirmed by simulation results. On the basis of the analysis, the paper proposes a bandwidth assignment paradigm for the assignment of network bandwidth to voice and data in an effort to guarantee minimal data delay and voice call blocking probability. The run time complexity of the bandwidth assignment computation and paradigm construction is shown to be polynomial bounded. Consequently, the resulting bandwidth assignment assures QoSs in terms of data delay and voice call blocking probability under various network loads. © 1998 Elsevier Science B.V.

Keywords: Broadband Integrated Services Digital Network (BISDN); Asynchronous Transfer Mode (ATM); Quality of Service (QoS); Bandwidth assignment

1. Introduction

Broadband Integrated Services Digital Networks (BISDNs)/Asynchronous Transfer Mode (ATM) [1–7] are expected to support diverse applications demanding different bandwidth [8–13] and Quality of Services (QoSs) [14–16]. It is thus essential to design networks which furnish the effective and dynamic allocation of the bandwidth to satisfy different service demands. The Virtual Path (VP) concept [1,17–19,24] has been proposed to achieve such goal. A VP can be viewed as a pre-established route through the network with dedicated bandwidth onto which virtual channels (VCs) are multiplexed. The dedicated bandwidth of VPs can be reassigned according to the traffic variation. The network performance can in turn be optimized.

Bandwidth assignment to VPs have been studied in numerous literature [8–13]. Herzberg and Pitsillides [10] proposed a bandwidth assignment scheme to maximize network throughput, while Gerla *et al.* [9] aimed at minimizing mean delay across a network. Hui *et al.* [11] formulated the VP bandwidth assignment problem as a non-linear programming model which minimized the total usage cost. Burgin [8] studied the dynamic behavior of some VP bandwidth assignment algorithms for a range of time-varying offered traffic inputs. In Ref. [12], a distributed VP bandwidth control mechanism was proposed to effectively determine the bandwidth assignment. Logothetis and Shioda [13] presented a centralized VP bandwidth allocation scheme to minimize the worst voice call blocking probability.

In this paper, we focus on the bandwidth assignment to two VPs (e.g. for voice and data) in broadband integrated networks with multiple channels. For multichannel integrated voice/data networks, various bandwidth assignment mechanisms have been proposed. In the fixed-boundary scheme [20,21], a fixed portion of the bandwidth is designated for voice and data. The major drawback of this scheme is the inefficient use of bandwidth. An improved scheme called the flexible-boundary strategy [22,23] was proposed to achieve dynamic sharing of the bandwidth between voice and data. Kekre [23] allowed bandwidth reservation for voice traffic and limited the buffer size for data traffic. The major limitation of this scheme is that the determination of the optimal bandwidth assignment is non-trivial and complicated. The goal of the paper is to

* Corresponding author.

[1] An earlier version of this paper has appeared in IEEE ICC'94.

0140-3664/98/$19.00
PII S0140-3664(97)00040-6

analytically determine optimal bandwidth allocated to voice and data by means of a queueing model with heterogeneous arrivals and multiple designated channels.

The paper first analyzes a multichannel flexible-boundary queueing system which allows bandwidth reservation for both voice and data traffic. This analysis is an extension and rectification of the work of Bhat and Fischer [24]. The accuracy of analytical results is confirmed by simulation results. On the basis of the analysis, the paper proposes a bandwidth assignment paradigm for the assignment of network bandwidth to voice and data in an effort to guarantee minimal data delay and voice call blocking probability. The run time complexity of the bandwidth assignment computation and paradigm construction is shown to be polynomial bounded. Consequently, the resulting bandwidth assignment assures QoSs in terms of data delay and voice call blocking probability under various network loads.

This paper is organized as follows. Section 2 presents the analysis of the multichannel flexible-boundary queueing system. Based on the analysis in Section 2, Section 3 illustrates the formal algorithm for the construction of bandwidth assignment paradigms. Finally, Section 4 concludes the paper.

2. Queueing analysis

2.1. Model

The analysis is based on a continuous-time queueing system (see Fig. 1) with heterogeneous arrivals (voice and data) and *multiple designated channels* (voice, data, and shared channels). Designated voice and data channels are dedicated for voice and data traffic transmissions, respectively, and shared channels can be used by either type of traffic. Voice traffic would balk from the system if legitimate channels are all in use upon its arrival. On the contrary, data traffic would be queued should legitimate channels be busy. In addition, the queue length for data traffic is assumed to be infinite since data traffic is loss sensitive in nature.

The analysis is to determine the blocking probability for voice traffic and the mean delay for data traffic. It is assumed that voice traffic and data traffic are Poisson distributed with arrival rates λ_1 and λ_2, respectively, and have exponential service-time distribution with rates μ_1 and μ_2, respectively. The traffic intensity of voice and data is denoted by ρ_1 (i.e. $\rho_1 = \lambda_1/\mu_1$) and ρ_2 (i.e. $\rho_2 = \lambda_2/\mu_2$), respectively. In addition, as shown in Fig. 1, let s be the total number of channels in the system, and m and n be the number of channels designated for voice and data traffic, respectively. Thus, there are $s-m-n$ shared channels which can be used by either type of traffic. All channels are employed in an FCFS manner.

2.2. Analysis

The system is ergodic [20] if $\lambda_2 < (s-m)\mu_2$. Let P_{ij} ($0 \le i \le s-n$; $j \ge 0$) be the steady state probability of having i voice calls and j data packets in the system. To compute the steady state distribution, several sets of balance equations have to be obtained. First of all, if the number of voice calls in the system is less than m (the number of the designated voice channels), i.e. $i = 0, 1, \ldots, m-1$, a set of balance equations can be attained as

$$\left.\begin{aligned}
&(\lambda_1+\lambda_2+i\mu_1)P_{i0} = (i+1)\mu_1 P_{i+1,0} + \mu_2 P_{i1} + \lambda_1 P_{i-1,0}\\
&(\lambda_1+\lambda_2+i\mu_1+\mu_2)P_{i1} = (i+1)\mu_1 P_{i+1,1} + 2\mu_2 P_{i2} + \lambda_1 P_{i-1,1} + \lambda_2 P_{i0}\\
&(\lambda_1+\lambda_2+i\mu_1+2\mu_2)P_{i2} = (i+1)\mu_1 P_{i+1,2} + 3\mu_2 P_{i3} + \lambda_1 P_{i-1,2} + \lambda_2 P_{i1}\\
&\qquad\vdots\\
&[\lambda_1+\lambda_2+i\mu_1+(s-m-1)\mu_2]P_{i,s-m-1} = (i+1)\mu_1 P_{i+1,s-m-1} + (s-m)\mu_2 P_{i,s-m} + \lambda_1 P_{i-1,s-m-1} + \lambda_2 P_{i,s-m-2}\\
&[\lambda_1+\lambda_2+i\mu_1+(s-m)\mu_2]P_{i,s-m} = (i+1)\mu_1 P_{i+1,s-m} + (s-m)\mu_2 P_{i,s-m+1} + \lambda_1 P_{i-1,s-m} + \lambda_2 P_{i,s-m-1}\\
&[\lambda_1+\lambda_2+i\mu_1+(s-m)\mu_2]P_{i,s-m+1} = (i+1)\mu_1 P_{i+1,s-m+1} + (s-m)\mu_2 P_{i,s-m+2} + \lambda_1 P_{i-1,s-m+1} + \lambda_2 P_{i,s-m}\\
&\qquad\vdots
\end{aligned}\right\} \quad (1)$$

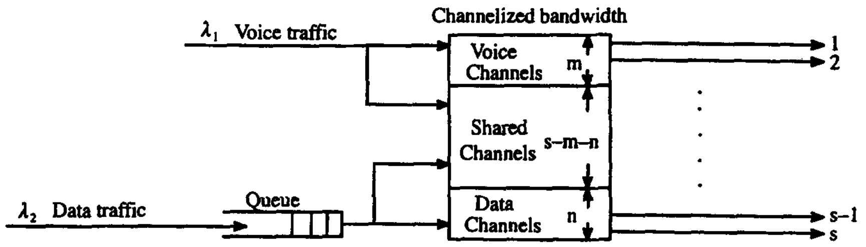


Fig. 1. The queueing model.

Secondly, if the number of voice calls in the system is equal to m, i.e. $i = m$, the set of balance equations can be given as

$$\left.\begin{aligned}
&(\lambda_1+\lambda_2+i\mu_1)P_{i0}=(i+1)\mu_1P_{i+1,0}+\mu_2P_{i1}+\lambda_1P_{i-1,0}\\
&(\lambda_1+\lambda_2+i\mu_1+\mu_2)P_{i1}=(i+1)\mu_1P_{i+1,1}+2\mu_2P_{i2}+\lambda_1P_{i-1,1}+\lambda_2P_{i0}\\
&(\lambda_1+\lambda_2+i\mu_1+2\mu_2)P_{i2}=(i+1)\mu_1P_{i+1,2}+3\mu_2P_{i3}+\lambda_1P_{i-1,2}+\lambda_2P_{i1}\\
&\vdots\\
&[\lambda_1+\lambda_2+i\mu_1+(s-i-1)\mu_2]P_{i,s-i-1}=(i+1)\mu_1P_{i+1,s-i-1}+(s-i)\mu_2P_{i,s-i}+\lambda_1P_{i-1,s-i-1}+\lambda_2P_{i,s-i-2}\\
&[\lambda_2+i\mu_1+(s-i)\mu_2]P_{i,s-i}=(i+1)\mu_1P_{i+1,s-i}+(s-i)\mu_2P_{i,s-i+1}+\lambda_1P_{i-1,s-i}+\lambda_2P_{i,s-i-1}\\
&[\lambda_2+i\mu_1+(s-i)\mu_2]P_{i,s-i+1}=(i+1)\mu_1P_{i+1,s-i+1}+(s-i)\mu_2P_{i,s-i+2}+\lambda_1P_{i-1,s-i+1}+\lambda_2P_{i,s-i}\\
&\vdots
\end{aligned}\right\}\quad(2)$$

Thirdly, if the number of voice calls in the system exceeds m but is less than the total number of channels eligible for voice traffic, i.e. $i = m + 1, m + 2, \ldots s - n - 1$, the set of balance equations can be expressed as

$$\left.\begin{aligned}
&(\lambda_1+\lambda_2+i\mu_1)P_{i0}=(i+1)\mu_1P_{i+1,0}+\mu_2P_{i1}+\lambda_1P_{i-1,0}\\
&(\lambda_1+\lambda_2+i\mu_1+\mu_2)P_{i1}=(i+1)\mu_1P_{i+1,1}+2\mu_2P_{i2}+\lambda_1P_{i-1,1}+\lambda_2P_{i0}\\
&(\lambda_1+\lambda_2+i\mu_1+2\mu_2)P_{i2}=(i+1)\mu_1P_{i+1,2}+3\mu_2P_{i3}+\lambda_1P_{i-1,2}+\lambda_2P_{i1}\\
&\vdots\\
&[\lambda_1+\lambda_2+i\mu_1+(s-i-1)\mu_2]P_{i,s-i-1}=(i+1)\mu_1P_{i+1,s-i-1}+(s-i)\mu_2P_{i,s-i}+\lambda_1P_{i-1,s-i-1}+\lambda_2P_{i,s-i-2}\\
&[\lambda_2+i\mu_1+(s-i)\mu_2]P_{i,s-i}=(i+1)\mu_1P_{i+1,s-i}+(s-i)\mu_2P_{i,s-i+1}+\lambda_1P_{i-1,s-i}+\lambda_2P_{i,s-i-1}\\
&[\lambda_2+i\mu_1+(s-i)\mu_2]P_{i,s-i+1}=(i+1)\mu_1P_{i+1,s-i+1}+(s-i)\mu_2P_{i,s-i+2}+\lambda_2P_{i,s-i}\\
&[\lambda_2+i\mu_1+(s-i)\mu_2]P_{i,s-i+2}=(i+1)\mu_1P_{i+1,s-i+2}+(s-i)\mu_2P_{i,s-i+3}+\lambda_2P_{i,s-i+1}\\
&\vdots
\end{aligned}\right\}\quad(3)$$

Finally, if the number of voice calls in the system is equal to $s-n$ (the maximum number of channels eligible for voice traffic), i.e. $i = s-n$, the set of balance equation can be obtained as

$$\left.\begin{aligned}
&(\lambda_2+i\mu_1)P_{i0}=\mu_2P_{i1}+\lambda_1P_{i-1,0}\\
&(\lambda_2+i\mu_1+\mu_2)P_{i1}=2\mu_2P_{i2}+\lambda_1P_{i-1,1}+\lambda_2P_{i0}\\
&(\lambda_2+i\mu_1+2\mu_2)P_{i2}=3\mu_2P_{i3}+\lambda_1P_{i-1,2}+\lambda_2P_{i1}\\
&\vdots\\
&[\lambda_2+i\mu_1+(s-i-1)\mu_2]P_{i,s-i-1}=(s-i)\mu_2P_{i,s-i}+\lambda_1P_{i-1,s-i-1}+\lambda_2P_{i,s-i-2}\\
&[\lambda_2+i\mu_1+(s-i)\mu_2]P_{i,s-i}=(s-i)\mu_2P_{i,s-i+1}+\lambda_1P_{i-1,s-i}+\lambda_2P_{i,s-i-1}\\
&[\lambda_2+i\mu_1+(s-i)\mu_2]P_{i,s-i+1}=(s-i)\mu_2P_{i,s-i+2}+\lambda_2P_{i,s-i}\\
&[\lambda_2+i\mu_1+(s-i)\mu_2]P_{i,s-i+2}=(s-i)\mu_2P_{i,s-i+3}+\lambda_2P_{i,s-i+1}\\
&\vdots
\end{aligned}\right\}\quad(4)$$

To simplify Eq. (1)–(4), let us define the moment-generating function of the occupancy distribution of data traffic as

$$\Pi_i(z)=\sum_{j=0}^{\infty}P_{ij}z^j, \quad (5)$$

where $|z| < 1$ and $0 \le i \le s-n$. First, for i being less than m, multiplying each term in Eq. (1) by z^j summing overall $j \ge 0$, and replacing the right term of Eq. (5) by $\Pi_i(z)$, Eq. (1) can be expressed as

$$D_i(z)\Pi_i(z) = (i+1)\mu_1 z\Pi_{i+1}(z) + \lambda_1 z\Pi_{i-1}(z) + (z-1)\mu_2 \times \sum_{j=0}^{s-m-1}(s-m-j)P_{ij}z^j, \quad (6)$$

where $D_i(z) = -\lambda_2 z^2 + [\lambda_1 + \lambda_2 + i\mu_1 + (s-m)\mu_2]z - (s-m)\mu_2$ and $\Pi_{-1}(z) = 0$. Similarly, for $i = m$, Eq. (2) can be expressed as

$$D_i(z)\Pi_i(z) = (i+1)\mu_1 z\Pi_{i+1}(z) + \lambda_1 z\Pi_{i-1}(z) + (z-1)\mu_2 \times \sum_{j=0}^{s-m-1}(s-m-j)P_{ij}z^j - \lambda_1 \sum_{j=0}^{s-m-1} P_{ij}z^{j+1}, \quad (7)$$

where $D_i(z) = -\lambda_2 z^2 + [\lambda_2 + i\mu_1 + (s-m)\mu_2]z - (s-m)\mu_2$ and $\Pi_{-1}(z) = 0$. For $m+1 \le i \le s-n-1$, Eq. (3) becomes

$$D_i(z)\Pi_i(z) = (i+1)\mu_1 z\Pi_{i+1}(z) + (z-1)\mu_2 \times \sum_{j=0}^{s-i-1}(s-i-j)P_{ij}z^j - \lambda_1 \sum_{j=0}^{s-i-1} P_{ij}z^{j+1} + \lambda_1 \times \sum_{j=0}^{s-i} P_{i-1,j}z^{j+1}, \quad (8)$$

where $D_i(z) = -\lambda_2 z^2 + [\lambda_2 + i\mu_1 + (s-i)\mu_2]z - (s-i)\mu_2$. Finally, for $i = s-n$, Eq. (4) can be similarly derived as

$$D_i(z)\Pi_i(z) = (z-1)\mu_2 \sum_{j=0}^{s-i-1}(s-i-j)P_{ij}z^j + \lambda_1 \times \sum_{j=0}^{s-i} P_{i-1,j}z^{j+1}, \quad (9)$$

where $D_i(z) = -\lambda_2 z^2 + [\lambda_2 + i\mu_1 + (s-i)\mu_2]z - (s-i)\mu_2$. Now, summing Eq. (6)–(9), we obtain

$$\sum_{i=0}^{s-n} S_i(z)\Pi_i(z) = [(s-m)\mu_2 - \lambda_2 z]\sum_{i=0}^{m}\Pi_i(z) + \sum_{i=m+1}^{s-n}[(s-i)\mu_2 - \lambda_2 z]\Pi_i(z) = \sum_{i=0}^{m}\sum_{j=0}^{s-m-1}(s-m-j)\mu_2 P_{ij}z^j + \sum_{i=m+1}^{s-n}\sum_{j=0}^{s-i-1}(s-i-j)\mu_2 P_{ij}z^j. \quad (10)$$

Differentiating Eq. (10) with respect to z and setting $z = 1$, one gets

$$-\lambda_2 + [(s-m)\mu_2 - \lambda_2]\sum_{i=0}^{s-n}\Pi_i'(1) - \sum_{i=1}^{s-m-n}(i\mu_2)\Pi_{i+m}'(1) = \sum_{i=0}^{m}\sum_{j=1}^{s-m-1} j(s-m-j)\mu_2 P_{ij} + \sum_{i=m+1}^{s-n}\sum_{j=1}^{s-i-1} j(s-i-j)\mu_2 P_{ij}. \quad (11)$$

Replacing λ_2/μ_2 by ρ_2, the mean number of data packets in the system can thus be derived as

$$\sum_{i=0}^{s-n}\Pi_i'(1) = \frac{\sum_{i=0}^{m}\sum_{j=1}^{s-m-1} j(s-m-j)P_{ij} + \sum_{i=m+1}^{s-n}\sum_{j=1}^{s-i-1} j(s-i-j)P_{ij} + \sum_{i=m+1}^{s-n}(i-m)\Pi_i'(1) + \rho_2}{s-m-\rho_2}. \quad (12)$$

Using Eq. (12) and applying Little's Formula, we can obtain the mean data delay as

$$T = \text{mean data delay} = \frac{\sum_{i=0}^{s-n}\Pi_i'(1)}{\lambda_2}. \quad (13)$$

Consequently, as shown in Eq. (12) and (13), to compute T, $\Pi_i'(1)$ ($i = m+1$ to $s-n$) and P_{ij} have to be solved. Now, let us temporarily turn our attention to the blocking probability (P_B) for voice traffic. Notice that P_B be expressed as

$$P_B = \text{voice blocking probability } = 1 - \left[\Pi_0(1) + \Pi_1(1) + \ldots + \Pi_{m-1}(1) + \sum_{i=m}^{s-n-1}\sum_{j=0}^{s-i-1} P_{ij}\right]. \quad (14)$$

Setting $z = 1$ in Eq. (6), (7), and (8), one can derive

$$\Pi_i(1) = \left[\frac{i+1}{\rho_1}\right]\Pi_{i+1}(1), \text{ for } i = 0, 1, \ldots, m-1; \quad (15)$$

and

$$\sum_{j=0}^{s-i-1} P_{ij} = \left[\frac{i+1}{\rho_1}\right]\Pi_{i+1}(1), \text{ for } i = m, m+1, \ldots, s-n-1. \quad (16)$$

Since $\sum_{i=0}^{s-n}\Pi_i(1) = 1$, from Eq. (14), (15), and (16), P_B can be expressed as

$$P_B = [\Pi_m(1) + \Pi_{m+1}(1) + \ldots + \Pi_{s-n}(1)] - \frac{1}{\rho_1}[(m+1)\Pi_{m+1}(1) + (m+2)\Pi_{m+2}(1) + \ldots + (s-n)\Pi_{s-n}(1)] = \Pi_m(1) + \sum_{i=m+1}^{s-n}\left(1 - \frac{i}{\rho_1}\right)\Pi_i(1). \quad (17)$$

$\Pi_m(1)$ in Eq. (17) can be computed recursively from Eq. (15) as

$$\Pi_m(1)=\frac{1-\sum_{i=m+1}^{s-n}\Pi_i(1)}{\left[\left(\frac{m!}{0!}\right)\left(\frac{1}{\rho_1}\right)^m+\left(\frac{m!}{1!}\right)\left(\frac{1}{\rho_1}\right)^{m-1}+\ldots+\frac{m!}{(m-1)!}\left(\frac{1}{\rho_1}\right)+1\right]}=\frac{1-\sum_{i=m+1}^{s-n}\Pi_i(1)}{\sum_{j=0}^{m}\left(\frac{m!}{j!}\right)\left(\frac{1}{\rho_1}\right)^{m-j}}. \quad (18)$$

Now in order to compute T and P_B, we must calculate $\Pi_i(1)$ and $\Pi_i'(1)$, where $m+1\le i\le s-n$, and the steady state probabilities, P_{ij}. Differentiating Eq. (8) and (9) with respect to z and setting $z=1$, we obtain $\Pi_i'(1)$ $(m+1\le i\le s-n)$ recursively. Similarly, setting $z=1$ in Eq. (8) and (9), one can derive $\Pi_i(1)$, where $m+1\le i\le s-n$. Consequently, by Eq. (12), the terms left undetermined are the following steady state probabilities:

(1) P_{ij}, where $i\le m$ and $j\le s-m-1$;

and

(2) P_{ij}, where $m+1\le i\le s-n$ and $j\le s-i-1$.

Note that Eq. (1)–(4) do not provide enough independent equations to solve these steady state probabilities. This is because $(s-n)$ additional unknowns, that is, P_{is-m} (for $i=1$ to $m-1$) and P_{s-i} (for $i=m$ to $s-n$), have to be resolved. Therefore, we must discover another $(s-n)$ additional independent equations. First of all, setting $z=1$ in Eq. (10), we obtain one additional equation. Moreover, let ϕ_i be the root of $D_i(z)$ inside (0,1), where

$$\phi_i=\frac{\lambda_2+i\mu_1+(s-i)\mu_2-\sqrt{[\lambda_2+i\mu_1+(s-i)\mu_2]^2-4(s-i)\lambda_2\mu_2}}{2\lambda_2},$$

where $m+1\le i\le s-n-1$. Equating the zeros on both sides of Eq. (8) provides $(s-m-n-1)$ additional equations to be used for finding the steady state probabilities. Now, we have to discover the last m equations. For $i\le m$, let

$$A_i(z)=D_i(z)-\frac{i\lambda_1\mu_1z^2}{A_{i-1}(z)}, \quad (19)$$

where $A_{-1}(z)=1$. From Eq. (6), (7), and (19), we have

$$A_m(z)\Pi_m(z)=(m+1)\mu_1z\Pi_{m+1}(z)-\lambda_1\sum_{j=0}^{s-m-1}P_{mj}z^{j+1}$$
$$+(z-1)\mu_2\left[\sum_{i=0}^{m}\frac{(\lambda_1z)^{m-i}\sum_{j=0}^{s-m-1}(s-m-j)P_{ij}z^j}{\prod_{k=i}^{m-1}A_k(z)}\right], \quad (20)$$

where $\Pi_{k=m}^{m-1}A_k(z)=1$. Notice that $A_m(Z)$ has m distinct roots between (0,1). The proof is presented in Appendix. Equating the zeros on both sides of Eq. (20) provides the last m equations needed. Accordingly, the mean delay of data traffic and the blocking probability of voice calls can thus be obtained directly from Eq. (13) and (17).

2.3. Analytical and simulation results

To verify the accuracy of the analysis, we undertook an event-driven simulation. In the simulation, we assumed that there were two types of arrivals (voice and data) and six channels in the system. Voice traffic would balk from the system at the lack of free channels at the time of its arrival. In contrast, data traffic would be buffered should all channels be busy. Both types of traffic are handled in an FIFO manner.

Fig. 2 shows the mean data delay as a function of the offered data-traffic intensity. The figure demonstrates that our analytic results agree with simulation results with negligible discrepancy. Moreover, the figure reveals the high sensitivity of the mean data delay to a ($=\mu_2/\mu_1$, i.e. the ratio of voice service time to data service time). We have observed a salient phenomenon which is that the data delay oscillates as the data-traffic intensity increases. The larger a is, the more the data delay oscillates. This phenomenon can be inferred from the following fact. Since voice calls often last on the order of minutes, while the duration of data transmissions may elapse only on the order of milliseconds, the periodic termination of voice calls can decrease data delays profoundly. Fig. 3 shows the same effect on the oscillations of the mean data delay with a constant a and different voice-traffic intensity.

On the contrary, we have discovered that the voice call blocking probability is irrelevant to a. As shown in Fig. 4, the blocking probability increases with the voice-traffic intensity. The effect on the mean data delay under various bandwidth assignments (i.e. m and n) is depicted in Figs. 5 and 6. In particular, Fig. 5 shows that the greater number of channels reserved for data, the later the delay oscillation takes place. Fig. 6 further illustrates the enormous increase in the mean data delay with m. Finally, Figs. 7 and 8 are plots of the voice call blocking probability as a function of ρ_2 under different bandwidth assignments. Not surprisingly, the blocking probability increases with n and decreases with m.

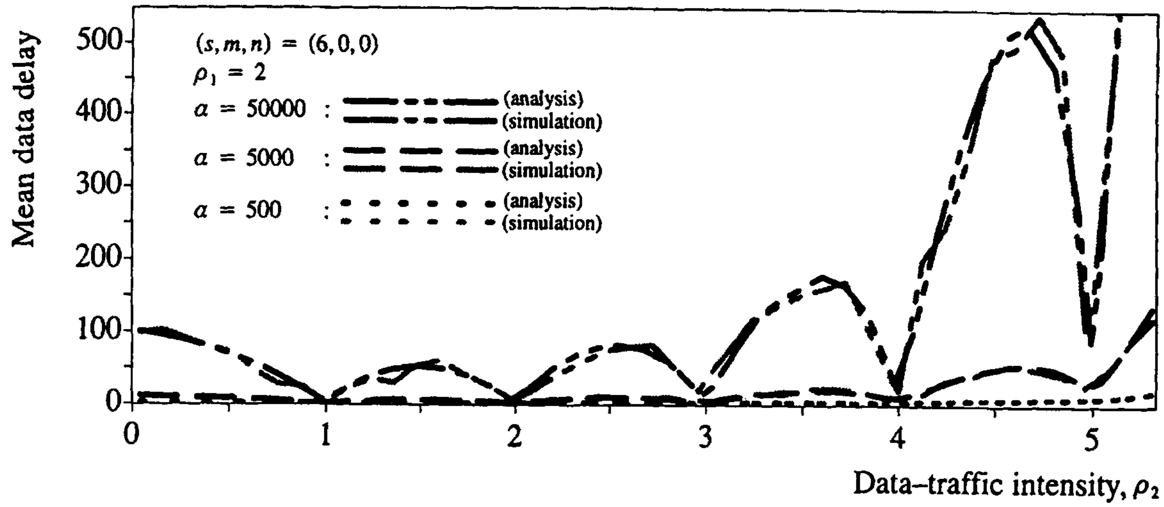


Fig. 2. Mean data delay under various a.

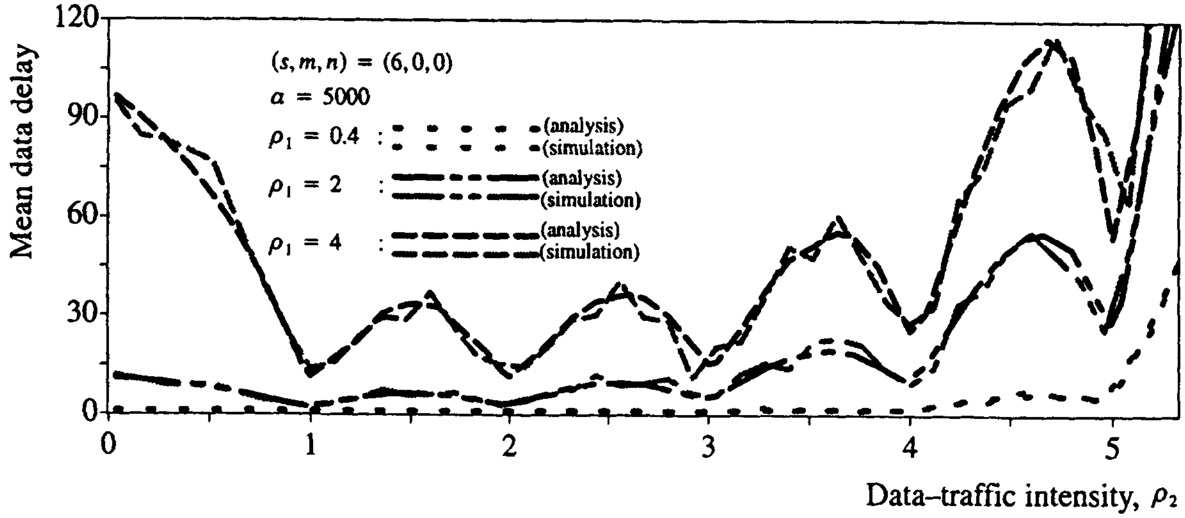


Fig. 3. Mean data delay under various ρ_1.

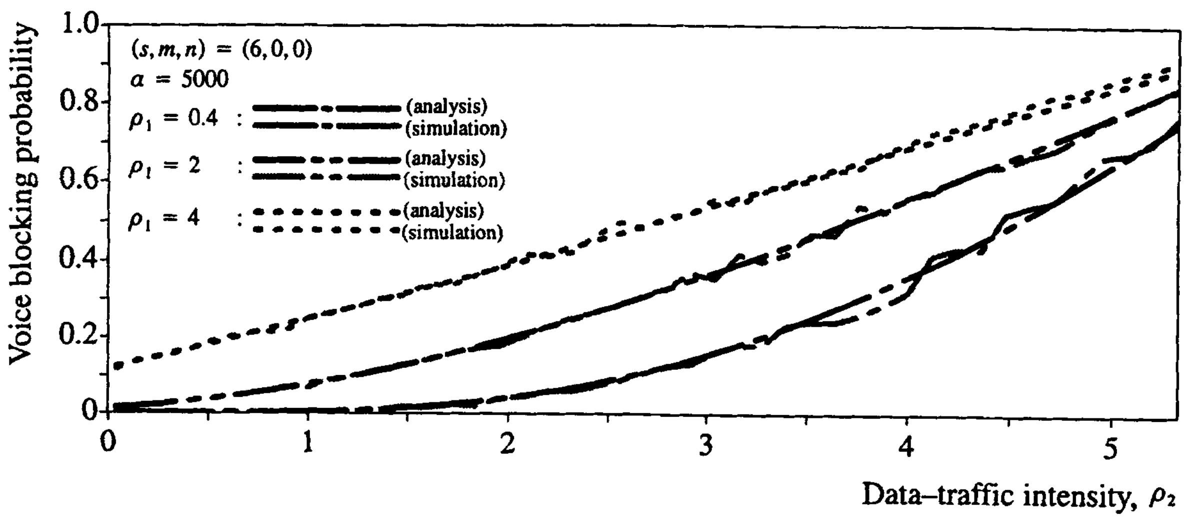


Fig. 4. Voice-call blocking probability.

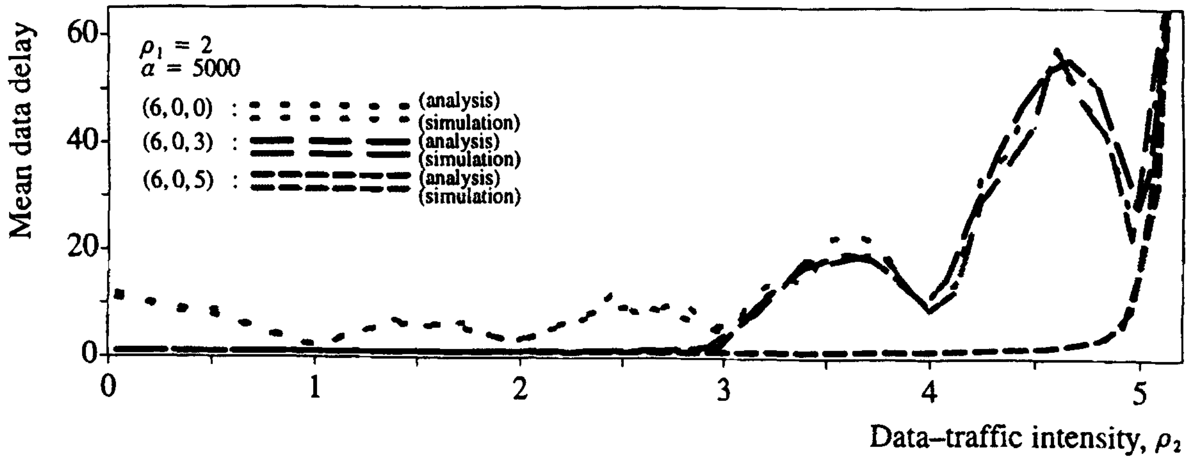


Fig. 5. Mean data delay under various n.

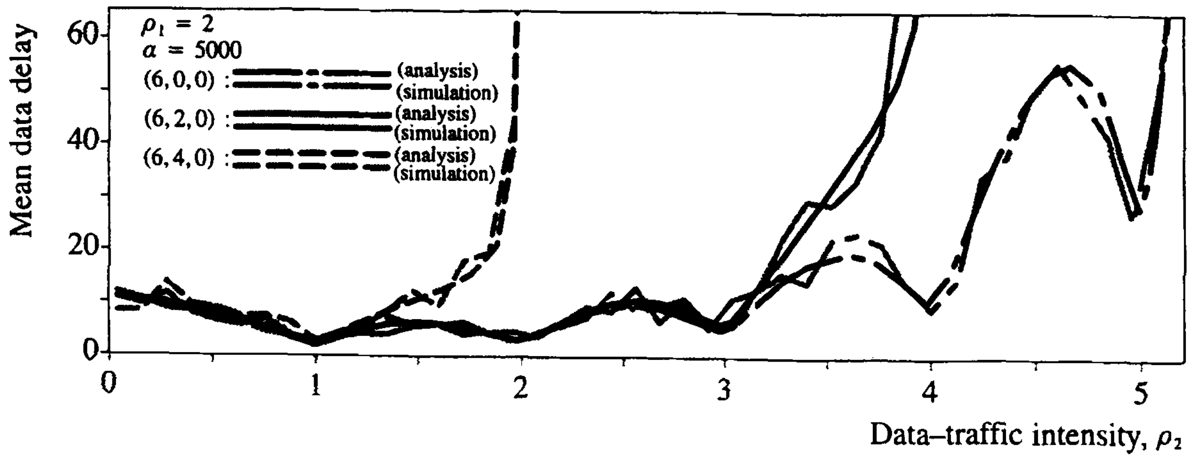


Fig. 6. Mean data delay under various m.

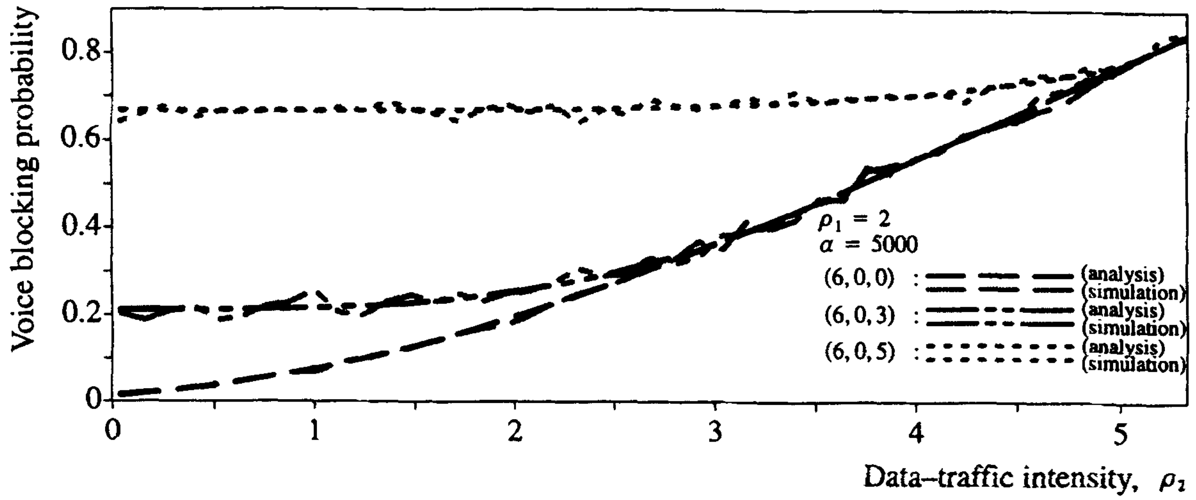


Fig. 7. Blacking probability under various n.

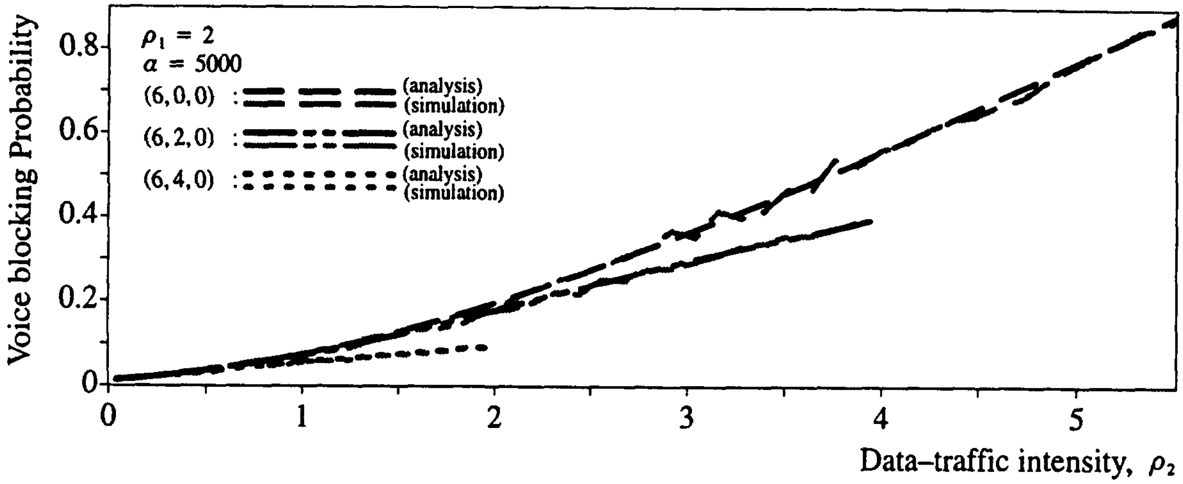


Fig. 8. Blocking probability under various m.

3. Paradigm construction algorithm

Based on the analytic results, the bandwidth assignment paradigm can be constructed as follows.

3.1. Algorithm for paradigm construction

Given λ_1, λ_2, μ_1, μ_2, T^+ (the QoS of data traffic), and P_B^+ (the QoS of voice traffic)

for designated voice channel $m = 0$ **to** $s - 1$
{**for** designated data channel $n = 0$ **to** $s - m - 1$
{compute mean data delay (T) under assigned channel usage (s, m, n) (by Eq. (13)); compute voice blocking probability (P_B) under assigned channel usage (s, m, n) (by Eq. (14));
if ($T \le T^+$) add (s, m, n) to the data-QoS paradigm;
if ($P_B \le P_B^+$) add (s, m, n) to the voice-QoS paradigm;
if ($T \le T^+$ **and** $P_B \le P_B^+$) add (s, m, n) to the system-QoS paradigm}
(End of Algorithm)

3.2. Complexity

The run time complexity of the paradigm construction is computed as follows. First of all, notice that the steady state probabilities are computed by means of Gaussian elimination. The complexity of Gaussian elimination [25] for N unknowns and N simultaneous equations is $O(N^3)$. Thus, in the case of the paradigm construction, N becomes $(m + 1)(s - m) + (s - m - n)^2$. In addition, a network with s channels has $s(s + 1)/2$ combinations of (s,m,n). Consequently, the run time complexity of the paradigm construction is $(s(s + 1)/2) \times O(N^3)$. It is worth noting that, since the construction of the paradigm is polynomial bounded, the recomputation of the paradigm due to the rearrangement of VPs can be performed in semi-real time. Consequently, the channelized bandwidth can be managed accurately and adaptively.

Using the paradigm construction algorithm, we construct the bandwidth assignment paradigms for a system with six channels. The constructed paradigms realize the sets of

Table 1
Paradigm satisfying voice QoS

ρ_2	ρ_1		
	0.5	1	2
0.5	(6,0,0), (6,1,0), (6,2,0), (6,3,0) (6,0,1), (6,1,1), (6,2,1), (6,3,1) (6,0,2), (6,1,2), (6,2,2), (6,3,2) (6,0,3), (6,1,3), (6,2,3), (6,4,0) (6,0,4), (6,1,4), (6,5,0), (6,4,1)	(6,0,0), (6,1,0), (6,2,0), (6,3,0) (6,0.1), (6,1,1), (6,2,1), (6,3,1) (6,0,2), (6,1,2), (6,2,2), (6,3,2) (6,0,3), (6,1,3), (6,2,3), (6,4,0) (6,5,0), (6,4,1)	(6,0,0), (6,1,0), (6,2,0), (6,3,0) (6,0,1), (6,1,1), (6,2,1), (6,3,1) (6,4,0), (6,4,1), (6,5,0)
1	(6,0,0), (6.1,0), (6,2,0), (6,3,0) (6,0,1), (6,1,1), (6,2,1), (6,3,1) (6,0,2), (6,1,2), (6,2,2), (6,3,2) (6,0,3), (6,1,3), (6,2,3), (6,4,0) (6,0,4), (6,1,4), (6,5,0), (6,4,1)	(6,0,0), (6,1,0), (6,2,0), (6,3,0) (6,0,1), (6,1,1), (6,2,1), (6,3,1) (6,0,2), (6,1,2), (6,2,2), (6,3,2) (6,0,3), (6,1,3), (6,2,3), (6,4,0) (6,5,0), (6,4,1)	(6,0,0), (6,1,0), (6,2,0), (6,3,0) (6,0,1), (6,1,1), (6,2,1), (6,3,1) (6,4,0), (6,4,1), (6,5,0)
2	(6,0,0), (6,1,0), (6,2,0), (6,3,0) (6,0,1), (6,1,1), (6,2,1), (6,3,1) (6,0,2), (6,1,2), (6,2,2), (6,3,2) (6,0,3), (6,1,3), (6,2,3), (6,4,0) (6,0,4), (6.1,4), (6,5,0), (6,4,1)	(6,0,0), (6,1,0), (6,2,0), (6,3,0) (6,0,1), (6,1,1), (6,2,1), (6,3,1) (6,0,2), (6,1,2), (6,2,2), (6,3,2) (6,0,3), (6,1,3). (6,2,3), (6,4,0) (6,5,0), (6,4,1)	(6,3,0), (6,3,1), (6,4,0), (6,4,1) (6,5,0)

Table 2
Paradigm satisfying data QoS

ρ_2	ρ_1		
	0.5	1	2
0.5	(6,0,0), (6,1,0), (6,2,0), (6,3,0) (6,0,1), (6,1,1), (6,2,1), (6,3,1) (6,0,2), (6,1,2), (6,2,2), (6,3,2) (6,0,3), (6,1,3), (6,2,3), (6,4,0) (6,0,4), (6,1,4), (6,5,0), (6,4,1) (6,0,5)	(6,0,0), (6,1,0), (6,2,0), (6,3,0) (6,0,1), (6,1,1), (6,2,1), (6,3,1) (6,0,2), (6,1,2), (6,2,2), (6,3,2) (6,0,3), (6,1,3), (6,2,3), (6,4,0) (6,0,4), (6,1,4), (6,5,0), (6,4,1) (6,0,5)	(6,0,1), (6,1,1), (6,2,1), (6,3,1) (6,0,2), (6,1,2), (6,2,2), (6,3,2) (6,0,3), (6,1,3), (6,2,3) (6,0,4), (6,1,4), (6,4,1) (6,0,5)
1	(6,0,0), (6,1,0), (6,2,0), (6,3,0) (6,0,1), (6,1,1), (6,2,1), (6,3,1) (6,0,2), (6,1,2), (6,2,2), (6,3,2) (6,0,3), (6,1,3), (6,2,3) (6,0,4), (6,1,4) (6,0,5)	(6,0,0), (6,1,0), (6,2,0), (6,3,0) (6,0,1), (6,1,1), (6,2,1), (6,3,1) (6,0,2), (6,1,2), (6,2,2), (6,3,2) (6,0,3), (6,1,3), (6,2,3) (6,04), (6,1,4) (6,0,5)	(6,0,1), (6,1,1), (6,2,1), (6,3,1) (6,0,2), (6,1,2), (6,2,2), (6,3,2) (6,0,3), (6,1,3), (6,2,3) (6,0,4), (6,1,4) (6,0,5)
2	(6,0,0), (6,1,0), (6,2,0), (6,3,0) (6,0,1), (6,1,1), (6,2,1), (6,3,1) (6,0,2) (6,1,2), (6,2,2), (6,3,2) (6,0,3), (6,1,3), (6,2,3) (6,0,4), (6,1,4) (6,0,5)	(6,0,0), (6,1,0), (6,2,0), (6,3,0) (6,0,1), (6,1,1), (6,2,1), (6,3,1) (6,0,2), (6,1,2), (6,2,2), (6,3,2) (6,0,3), (6,1,3), (6,2,3) (6,0,4), (6,1,4) (6,0,5)	(6,0,1), (6,1,1), (6,2,1), (6,3,1) (6,0,2), (6,1,2), (6,2,2), (6,3,2) (6,0,3), (6,1,3), (6,2,3) (6,0,4), (6,1,4) (6,0,5)

optimal bandwidth assignments satisfying QoS of no more than 10 units of data delay and 0.15 of voice call blocking probability. Table 1 shows a bandwidth assignment paradigm satisfying only the QoS of voice traffic. For instance, the set of bandwidth assignments {(6,3,0), (6,3,1), (6,4,0), (6,4,1), (6,5,0)} guarantees that the blocking probability is less than 0.15 under load $\rho_1 \leq 2$ and $\rho_2 \leq 2$. Table 2 shows a bandwidth assignment paradigm assuring the QoS of data traffic. Finally, Table 3 presents a bandwidth assignment paradigm assuring the QoS of both types of traffic.

The maximum values of ρ_1 and ρ_2 presented in the tables are carefully selected so that there exists at least one optimal bandwidth assignment under a heavy load. If the network load increases greatly, an optimal bandwidth assignment may not exist. If so, an approximate assignment can be used instead of an optimal one. This is also true if a more stringent QoS is required. Under light loads, there may exist more than one optimal assignment selection satisfying QoS. The ultimate bandwidth assignment can thus be selected based on more significant requirements.

4. Conclusions

This paper first analyzed a multichannel flexible-boundary queueing system allowing reservations for both voice and data traffic, The system was based on a continuous-time model. The derivation of the mean data delay and the voice-call blocking probability was presented. One salient result was that the data delay oscillates as the data-traffic intensity increases. Another finding was that the mean data delay is highly sensitive to a, the ratio of voice

Table 3
Paradigm satisfying system QoS

ρ_2	ρ_1		
	0.5	1	2
0.5	(6,0,0), (6,1,0), (6,2,0), (6,3,0) (6,0.1), (6,1,1), (6,2,1), (6,3,1) (6,0,2), (6,1,2), (6,2,2), (6,3,2) (6,0,4), (6,1,4), (6,5,0), (6,4,1)	(6,0,0), (6,1,0), (6,2,0), (6,3,0) (6,0,1), (6,1,1), (6,2,1), (6,3,1) (6,0,2), (6,1,2), (6,2,2), (6,3,2) (6,5,0), (6,4,1)	(6,0,1), (6,1,1), (6,2,1), (6,3,1) (6,4,1)
1	(6,0,0), (6,1,0), (6,2,0), (6,3,0) (6,0,1), (6,1,1), (6,2,1), (6,3;1) (6,0,2), (6,1,2), (6,2,2), (6,3,2) (6,0,3), (6,1,3), (6,2,3) (6,0,4), (6,1,4)	(6,0,0), (6,1,0), (6,2,0), (6,3,0) (6,0,1), (6,1,1), (6,2,1), (6,3,1) (6,0,2), (6,1,2), (6,2,2), (6,3,2) (6,0,3), (6,1,3), (6,2,3)	(6,0,1), (6,1,1), (6,2,1), (6,3,1)
2	(6,0,0), (6,1,0), (6,2,0), (6,3,0) (6,0,1), (6,1,1), (6,2,1), (6,3,1) (6,0,2), (6,1,2), (6,2,2), (6,3,2) (6,0,3), (6,1,3), (6,2,3) (6,0,4), (6,1,4)	(6,0,0), (6,1,0), (6,2,0), (6,3,0) (6,0,1), (6,1,1), (6,2,1), (6,3,1) (6,0,2), (6,1,2), (6,2,2), (6,3,2) (6,0,3), (6,1,3), (6,2,3)	(6,3,1)

service time to data service time. The paper then proposed a paradigm construction algorithm for the bandwidth assignment to voice and data. In addition, the bandwidth assignment paradigms for a six-channel system was constructed to realize the sets of optimal bandwidth assignments satisfying QoSs under three types of network loads. Finally, the paper indicated that the construction of the paradigm requires polynomial-bounded run time complexity ($(s(s+1)/2) \times O(N^3)$). This fact allows the recomputation of the paradigm due to the rearrangement of VPs to be performed in semi-real time.

Appendix

It can be shown that, from Eq. (6) and (19),

$$A_i(0)A_i(1) < 0, \quad \text{for } i = 0, 1, \ldots, m-1.$$

Therefore, $A_i(z)$ must have one root between (0,1), for $i = 0, 1, \ldots, m-1$. Furthermore, referring to Eq. (19), for $i = 0, 1, \ldots, m$, we get

$$A_i(z)A_{i-1}(z) = D_i(z)A_{i-1}(z) - i\lambda_1\mu_1 z^2, \tag{21}$$

where $A_{-1}(z) = 1$. Since $A_i(z)$ has one root between (0,1), expression $D_i(z)A_{i-1}(z) - i\lambda_1 u_1 z^2$ also has one root between (0,1), for $i = 0, 1, \ldots, m-1$. Using Eq. (21) and after algebraic simplification, we can express $A_m(z)$ as

$$A_m(z) = \frac{D_0(z)\left[D_1(z)A_0(z) - \lambda_1 u_1 z^2\right]\left[D_2(z)A_1(z) - 2\lambda_1 u_1 z^2\right]\ldots\left[D_m(z)A_{m-1}(z) - m\lambda_1 u_1 z^2\right]}{[A_0(z)]^2[A_1(z)]^2\ldots[A_{m-1}(z)]^2}. \tag{22}$$

Eq. (22) states the fact that $A_m(z)$ has m roots between (0,1). Finally, it can be easily proven that if z_i is the root of $A_i(z)$, z_j is the root of $A_j(z)$, and $i \neq j$, then $z_i \neq z_j$. This concludes our proof that $A_m(z)$ has m distinct roots between (0,1).

To determine m roots of $A_m(z)$, we developed a computer program which used the Horner's rule [26] and the Newton–Raphson method with synthetic division [27] supposing that Newton–Raphson method converges after K iterations in the worst case. In addition, using the Horner's rule, $A_m(z)$ can be evaluated in $O(m)$ steps. Hence, the complexity of finding m roots of $A_m(z)$ is $K \times O(m^2)$.

References

[1] D.E. Mcdysan, D.L. Spohn, ATM: Theory and Application, McGraw–Hill, 1995.

[2] U. Black, ATM: Foundation for Broadband Networks, Prentice Hall Series, 1995.

[3] R. Handel, M.N. Huber, Integrated Broadband Networks: An Introduction to ATM-Based Networks, Addison–Wesley, 1993.

[4] R. Handel *et al.*, ATM Networks: Concepts, Protocols, Applications, Addison–Wesley, 1994.

[5] M. de Prycker, Asynchronous Transfer Mode: Solution for Broadband ISDN, Ellis Horwood, 1993.

[6] D. Minoli, Broadband Network Analysis and Design, Artech House, 1993.

[7] S. Minzer, Broadband ISDN and asynchronous transfer mode (ATM), IEEE Comm. Mag., September, 1989.

[8] J. Burgin, Dynamic capacity management in the BISDN, Int. J. Digital Analog Comm. System 3, 1990.

[9] M. Gerla *et al.*, Topology design and bandwidth allocation in ATM nets, IEEE J. Selected Areas Comm., October, 1989.

[10] M. Herzberg, A. Pitsillides, A Hierarchical Approach for the Bandwidth Allocation, Management, and Control in BISDN, Proc. ICC'93, 1993.

[11] J.Y. Hui *et al.*, A layered broadband switching architecture with physical or virtual path configurations, IEEE J. Selected Areas Comm., December, 1991.

[12] R. Kawamura *et al.*, Fast VP-bandwidth management with distributed control in ATM networks, IEICE Trans. Comm. E77-B (1) 1994.

[13] M. Logothetis, S. Shioda, Centralized virtual path bandwidth allocation scheme for ATM networks, IEICE Trans. Comm. E75-B (10) 1992.

[14] J.F. Kurose, Open issues and challenges in providing quality of service guarantees in high speed networks, Proc. Computer Comm. Rev., January, 1993.

[15] R. Nagarajan *et al.*, On defining, computing and guaranteeing quality of service in high speed networks, IEEE Proc. INFOCOM'92, 1992.

[16] A.A. Lazar *et al.*, Control of resources in broadband networks with quality of service guarantees, IEEE Comm. Mag., October, 1991.

[17] J. Burgin, D. Dorman, Broadband ISDN resource management: The role of virtual paths, IEEE Comm. Mag., September, 1991.

[18] S. Ohta, K. Sato, Dynamic bandwidth control of the virtual path in an asynchronous transfer mode network, IEEE Trans. Comm. 40 (7) 1992.

[19] K. Sato *et al.*, Broad-band ATM network architecture based on virtual paths, IEEE Trans. Comm. 38 (8) 1990.

[20] L. Kleinrock, Queueing Systems, Vol. I: Theory, John Wiley, 1975.

[21] L. Kleinrock, Queueing Systems, Vol. II: Computer Applications, John Wiley, 1975.

[22] M. Fischer, T. Harris, A model for evaluating the performance of an integrated circuit- and packet-switched multiplex structure, IEEE Trans. Comm. COM-24, 1976.

[23] H. Kekre, C. Saxena, A finite waiting room queueing model with multiple servers having markovian interruptions and its applications, Comput. Elec. Eng. 5, 1978.

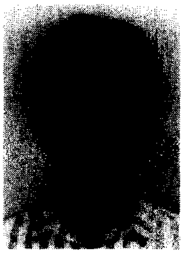
[24] U.N. Bhat, M.J. Fischer, Multichannel queueing systems with heterogeneous class of arrivals, Naval Res. Logist. Quarterly 23 (2) 1976.

[25] R. Sedgewick, Algorithms, Addison Wesley, 1990.

[26] U. Manber, Introduction to Algorithm, A Creative Approach, Addison Wesley, 1989.

[27] A. Constantinides, Applied Numerical Methods with Personal Computers, McGraw–Hill, 1987.

Maria C. Yuang received a B.S. degree in applied mathematics from the National Chiao Tung University, Taiwan, in 1978; an M.S. degree in computer science from the University of Maryland, College Park, Maryland, in 1981; and a Ph.D. degree in electrical engineering and computer science from the Polytechnic University, Brooklyn, New York, in 1989. From 1981 to 1990, she was with AT&T Bell Laboratories and Bell Communications Research (Bellcore), where she was a member of technical staff working on high speed networking and protocol engineering. She has been an associate professor in computer science and information engineering at the National Chiao Tung University, Taiwan, since 1990. Her current research interests include high speed networking, multimedia communications, performance modelling and analysis, and ATM network management.



Yieh R. Haung was born in Taiwan, 1968. He received a B.S. degree in Computer Science from SooChow University, Taiwan, in 1991, and an M.S. degree in Computer Science and Information Engineering from National Chiao Tung University, Taiwan, in 1993. He is currently a Ph.D. candidate in the same department. His current research interests include performance modelling and analysis, high speed networking, ATM multicast switch design, and personal communication services network.