The HKU Scholars Hub   The University of Hong Kong   香港大學學術庫

Title	A unified performance model for reservation-type multiple-access schemes
Author(s)	Qiu, X; Li, VOK
Citation	Ieee Transactions On Vehicular Technology, 1998, v. 47 n. 1, p. 173-189
Issued Date	1998
URL	http://hdl.handle.net/10722/42771
Rights	©1998 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.

A Unified Performance Model for Reservation-Type Multiple-Access Schemes

Xiaoxin Qiu, *Member, IEEE,* and Victor O. K. Li, *Fellow, IEEE*

Abstract— This paper presents a unified performance model for the integrated voice/data wireless system using reservation-type multiple-access (MA) schemes. It is observed that although these schemes are different in the frame structure and resource assignment procedure, all of them can be described by several common state variables whose evolvement exhibits the Markovian property. Based on this observation, a general Markovian model is developed in this paper. Three performance measures, namely, voice-packet-loss probability, data throughput, and data delay are defined. As a special case, the performance evaluation model for the voice-only system is also presented. Numerical results are given and verified by simulation under both voice-only and integrated scenarios using packet-reservation MA (PRMA), dynamic time-division multiple access (D-TDMA), and resource-auction multiple access (RAMA) as examples. It is found that our analytical model is quite accurate, especially in the region of interest. The impact of system parameters (such as the voice-permission probability, data-retransmission probability, maximum number of voice slots per frame, etc.) on the integrated system performance is also investigated for these three example systems.

Index Terms—Markovian models, multiple access, reservation-based protocols, TDMA.

I. INTRODUCTION

WIRELESS communication has become an important field of activity in telecommunications. The demand for services increases dramatically, and the required service is no longer restricted to the telephone call. The emerging wireless data communication services, such as portable computing, paging, personal e-mail, etc., play an increasingly important role in current wireless systems. It is recognized that these systems will be required to support a wider range of telecommunication applications involving packet data, voice, image, and full-motion video. With the implementation of a truly multimedia and personal communications system, the user will be released from the bondage of the telephone line and enjoy the freedom of telecommunications [7], [11].

One problem associated with the wireless feature is that the system bandwidth is limited. To satisfy the huge service demand and to fulfill the multimedia requirement with limited bandwidth, an efficient multiple-access (MA) scheme is required. The MA scheme will have a significant impact on the user performance, system capacity, and hardware complexity.

A successful MA scheme needs to take full advantage of the characteristics of different kinds of traffic to achieve high multiplexing efficiency. Some of these characteristics are well known. For instance, in packetized voice communication, the user's speech consists of a sequence of talkspurts and silence gaps. In the silence gap, nothing is transmitted, and the channel can be released for other users. Compared to short data communication, which generates only one or two delay-insensitive packets during one connection, the talkspurt, including a sequence of packets generated periodically, can be considered as continuous and delay sensitive. For this kind of continuous and delay-sensitive traffic, the connection-oriented transmission mode is better than the connectionless one, which is preferred in short data transmission. On the other hand, video traffic can also be considered as a sequence of delay-sensitive packets generated continuously. Contrary to conventional voice traffic, it requires much higher data rate and has more stringent performance requirements. When video traffic is integrated in the system, the major problem is how to support the mixed traffic with varying data rates efficiently. In this paper, we only focus on a system with integrated voice and data traffic.

Reservation-based time-division MA (TDMA) schemes have been extensively studied in the literature. Dynamic TDMA (D-TDMA), which was first introduced in satellite communications, is currently proposed for use in the personal communications system (PCS) [6], [9], [17], [19]. Packet-reservation multiple access (PRMA) was proposed by Goodman *et al.* for indoor wireless communications [13], [14], and resource-auction multiple access (RAMA) has been designed for quick resource assignment in the wireless system [1]–[5].[1] All these schemes and their many variations belong to the family of reservation-type MA schemes. In D-TDMA and RAMA, there are two periods in each frame: one period for making reservations using slotted ALOHA and another for information transmission. Part of the system bandwidth is dedicated for reservation. In PRMA, all the system resources are used for information transmission. The reservation is achieved by transmitting information packets directly instead of special reservation packets. The common feature of these reservation-type protocols is that they all have the frame structure and use the reservation mechanism to achieve high multiplexing efficiency in the hybrid transmission mode.

[1] It should be pointed out that the original RAMA dealt with multicarrier systems with one carrier solely dedicated to RAMA while the rest of the carriers were for communications only. In this work, we deal with a single carrier RAMA with some dedicated slots for auctioning in each frame.

There is much work to investigate the performance of these schemes individually, concentrating on the voice-only case. D-TDMA is the same as reservation TDMA when only voice traffic exists in the system. Its performance has been investigated in [17]. The performance of PRMA has been simulated in [8] and analyzed in [14] using equilibrium point analysis (EPA). A two-dimensional (2-D) Markovian model can be found in [15] and [18], and a simplified Markovian model is presented in [16]. The resource assignment capability and voice performance of RAMA are investigated in [1], [2], and [4].

Unfortunately, the performance of the reservation-type integrated voice and data system is still not well understood. The integrated D-TDMA system has been simulated in [19] in the PCS environment, but no analytical model is available to evaluate its performance. In [13], an integrated voice and data PRMA system was modeled by two Markov chains corresponding to voice and data subsystems, respectively. Its performance was evaluated using EPA. However, the deviation between the analytical and simulation results is quite large. A more accurate three-dimensional (3-D) Markovian model can be found in [20]. But it is not easy to solve this 3-D Markov chain. When the system is large, it becomes intractable.

We observe that although the reservation mechanisms and the frame structures of reservation-type protocols are different at a first glance, they can be fully described by several common state variables, and the system transitions exhibit the Markovian property. In this paper, a unified performance evaluation model will be developed for the reservation family of MA schemes based on these observations. Our model can be used to choose system parameters so as to optimize the system performance. Some interesting cases are studied in the section on numerical results. These cases give some insight into the system behavior.

This paper is organized as follows. In Section II, we will briefly describe three representative protocols of the reservation family: PRMA, D-TDMA, and RAMA. The general assumptions for the system are also given. The voice and data terminal models are developed in Section III. Then in Section IV, a unified performance evaluation model is developed for the integrated system. As a special case, the analytical model for the voice-only system is also derived at the end of this section. Numerical examples are given in Section V. The accuracy of the analytical model is verified by simulation. The system behavior will be investigated by varying the system parameters, such as the voice-permission probability, data-retransmission probability, maximum number of voice slots per frame, etc. Finally, we conclude in Section VI.

II. PROTOCOL DESCRIPTIONS

In this section, we will briefly describe three representative protocols of the reservation family: PRMA, D-TDMA, and RAMA. All of them have the frame structure and the contention-based reservation mechanism. For ease of analysis, the following assumptions are made in this paper. First, a single-cell system with one base station is considered. Also, we will focus on the uplink (mobile users to base station)
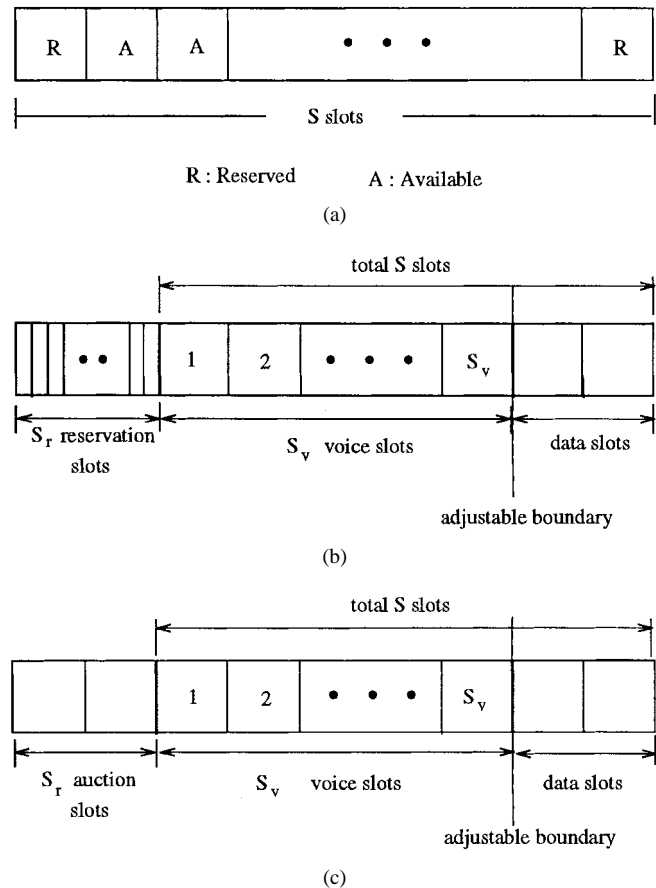

Fig. 1. Frame structures of PRMA, D-TDMA, and RAMA.

performance and assume that the base station is powerful enough to maintain the downlink (base station to mobile users) performance. Immediate feedback is required. This assumption is reasonable for terrestrial wireless systems consisting of small cells with short propagation delay. In addition, the channel is assumed error free for both the uplink and downlink. The failure of transmission is only caused by collisions.

A. PRMA

In PRMA, time is divided into slots which are grouped into frames. The frame length is determined by the voice-packet-generation rate. One voice packet is generated in a frame. The frame structure is shown in Fig. 1(a). A slot in a frame is either available or reserved by the voice user. Both voice and data users can contend for the available slots, based on the voice-transmission-permission probability p_t and the data-retransmission probability p_r, respectively. (Here, "voice users" refers to those who have a newly generated talkspurt, but have not obtained reservation yet.) If a voice user succeeds, this slot will be labeled as reserved and this voice user can use the corresponding slot in subsequent frames until the end of the current talkspurt. If a data user succeeds in the contention, this slot is still labeled as available, and the data user can only use this slot in the current frame and no reservation is allowed. Contrary to other reservation protocols, no dedicated reservation packet is needed in PRMA. The information packet is used for channel access [13], [14].
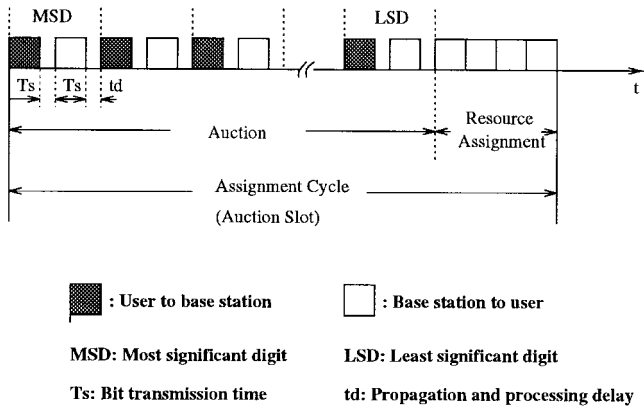
Fig. 2.   Illustration of the resource auction procedure in RAMA.

B. D-TDMA

D-TDMA was first introduced in satellite communications [6], [9], [17] and has recently been proposed as a candidate MA scheme for the third-generation wireless communication systems [19]. Time on the channel is also divided into a contiguous sequence of TDMA frames, which are subdivided into reservation slots, voice slots, and data slots. The frame structure is shown in Fig. 1(b), where there are S_r reservation slots, S_v voice slots, and the remaining slots are exclusively assigned to data. The base station keeps track of the information slots, which are either reserved or available. There are two types of packets being transmitted in the channel: the reservation packet and information packet. The reservation packet is used for the reservation of information slot (either voice or data slots). It often includes the origin and destination addresses for data or the calling and called users' ID's for voice. It is usually much shorter than the information packet.

A terminal generating a new voice talkspurt or a new data packet transmits the appropriate reservation packet in the reservation slots of the next frame, based on the voice-transmission-permission probability p_t or the data-retransmission probability p_r. If there is more than one packet transmitted in the same reservation slot, collision occurs. All the packets are destroyed. At the end of each reservation slot, the successful or unsuccessful reservation will be identified and broadcasted by the base station. The unsuccessful user can retry in the next reservation slot with probability p_t or p_r. This reservation procedure is quite similar to that used in PRMA.

The voice user successfully transmitted in the reservation slot will be assigned one of the available voice slots and will keep using it in subsequent frames until the end of the talkspurt. If there is no voice slot available, the voice user has to recontend in the next frame. After the assignment of voice traffic, data users who transmitted successfully in reservation slots can use all the remaining slots, but cannot reserve them. If there is no slot available, the data user has to recontend in the next frame. No reservation queue is maintained in the base station.

C. RAMA

The frame structure of RAMA is quite similar to that of D-TDMA except that the reservation slots are replaced by

the auction slots, as shown in Fig. 1(c). The assignment of voice slots is subject to a limit S_v. The data user can use the remaining capacity. The base station maintains a list of available and reserved slots. In each auction slot, the available resource will be auctioned to requesting users and assigned to the winner. The auction procedure is based on the user's ID, which is a random number generated by the user when he needs to contend. For instance, the user's ID can be an eight-digit binary number with two additional digits for priority assignment. The number of digits in the ID depends on the number of users in the system. A sufficient number of digits should be used to keep the probability of two independent users generating the same number small. It is assumed here that voice users always have higher priority than data users. The priority digits of voice are "10," and those of data are "00." (Note that a third kind of traffic with priority higher than voice can have priority digits "11.") At the beginning of the auction, the requesting user will transmit its ID in the auction slot, one digit at a time. Following each transmitted digit, the base station will announce the highest value among received digits on the downlink channel. Any user with digit value less than the announced one will drop out from further participation in this auction. Note that multiple users with identical digit value do "collide," but do not cause an erasure or failure [1], [2], [4]. When all the digits have been transmitted, there will be a final winner. The base station will assign this user an available slot, subject to certain constraints (e.g., whether it is voice or data). This auction and assignment procedure is illustrated in Fig. 2. It is clear that if there is a voice user transmitting in an auction slot, data users will always drop out. The users dropping out in the current auction slot can enter the next one. An available information slot will always be assigned to one of the requesting users in the auction, irrespective of the traffic load. This kind of deterministic assignment procedure compares very well to PRMA or D-TDMA, especially when the traffic is high. Under heavy traffic, the contention success probability of PRMA or D-TDMA will degrade very quickly due to frequent collisions.

III. TERMINAL MODELS

A. Voice Terminal Model

In each reservation-type protocol, an active voice user may be in one of three states, namely, the silence state (SS), the reservation state (RS), and the contention state (CS). When the user has nothing to transmit, it is in the silence state. When the user generates a new talkspurt and has not made a reservation yet, it leaves the silence state and enters the contention state. When it contends successfully and obtains a reservation, this user will enter the reservation state. Fig. 3(a) illustrates the transitions among these three states [14]. In this figure, q_s is the probability that a talkspurt ends in a frame, q_a is the probability that a talkspurt is generated in a frame, and P_{CR} is the probability that a terminal gets the reservation in the current frame. We assume the probability that a user returns to the silence state before it obtains a reservation is zero. It is also assumed that the lengths of talkspurts and silence gaps are
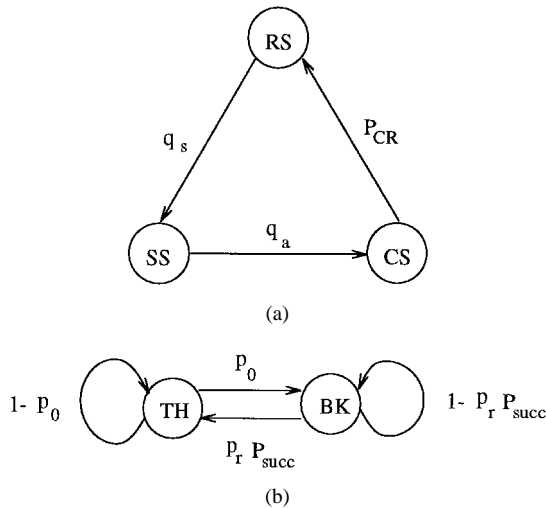
Fig. 3.   Voice and data terminal models.

exponentially distributed with means t_1 and t_2, respectively. (Experimental results show that t_1 is around 1 s, and t_2 is around 1.35 s [14].) If we assume that the frame duration is T s, then

$$q_s = 1 - \exp(-T/t_1) \quad \text{and} \quad q_a = 1 - \exp(-T/t_2). \quad (1)$$

P_{CR} depends on the system parameters such as the permission probability, voice activity factor, channel status, etc. [14].

B. Data Terminal Model

The data terminal is either in the thinking state (TH) waiting for the generation of a new data packet or in the backlogged state (BK) trying to transmit the generated packet. When a packet is generated, the user is forbidden from generating a new one until the current one is successfully transmitted. We consider the delayed first transmission (DFT) operation mode in this paper, i.e., when a new data packet is generated, the user enters the backlogged state immediately and the packet is transmitted according to the retransmission probability p_r in each available slot. It is shown in [10] that the DFT mode has almost the same performance as the immediate first transmission (IFT) mode. We use DFT here just for the analytical simplicity. We assume that the data packet is generated independently in each data user following the Bernoulli distribution with parameter p_0 packets/frame. The state transition of a data terminal is illustrated in Fig. 3(b), where P_{succ} is the transmission success probability.

IV. UNIFIED PERFORMANCE EVALUATION MODEL FOR INTEGRATED VOICE AND DATA SYSTEM

It is observed that although the reservation mechanisms and frame structures of different reservation-type protocols are different, each can be fully described by three common state variables whose evolvement is a Markov process. Based on this observation, a unified performance evaluation model for the integrated voice and data system using reservation-type MA schemes will be developed in this section. The stationary distribution obtained from this model is used to derive statistics

of interest such as the voice-packet-loss probability, data throughput, and delay. As a special case, the performance evaluation model for the voice-only system will be presented at the end.

A. Unified Performance Evaluation Model

Denote the total number of information slots in a frame as S (including both voice and data slots), the number of reservation (auction) slots in D-TDMA (RAMA) as S_r, and the maximum number of voice slots per frame in D-TDMA or RAMA as S_v. Note that S_v may equal S if no slots are dedicated to data transmission.

Given N_v voice users and N_d data users, the system can be fully described by three state variables $\{N_r, N_s, N_b\}$, namely, the number of voice users in the reservation state, the number of voice users in the silence state, and the number of data users in the backlogged state. The number of voice users in the contention state will be $N_c = N_v - N_r - N_s$, and the number of data users in the thinking state will be $N_t = N_d - N_b$.

We model the evolvement of the system as an embedded Markov process with the embedded point at the beginning of each frame. Since the system state in the current frame only depends on that in the immediately preceding one, this embedded process can be modeled by a 3-D Markov chain. This Markov chain is ergodic since it is obviously irreducible and has a finite number of states. Therefore, the stationary distribution of system states exists. We denote it as $\underline{\Pi}$

$$\underline{\Pi} = \{\pi(n_r, n_s, n_b)\} = \{P(N_r = n_r, N_s = n_s, N_b = n_b)\}.$$

To solve a 3-D Markov chain is computationally expensive. Hence, we employ a simplification.

1) Simplification: We observe that the voice user leaving or entering the silence state only depends on whether a new talkspurt is generated or the current one is ended. It does not depend on how many voice users are contending or how many data users are backlogged. Therefore, we can split the 3-D Markov chain into two subprocesses [12]: the speaking-silence one and the reservation one. The speaking-silence process, whose state variable is N_s, describes the above transition of voice users' entering and leaving the silence state. We denote its stationary distribution as $\underline{\Pi}^{(s)}$

$$\underline{\Pi}^{(s)} = \{\pi(n_s)\} = \{P(N_s = n_s)\}.$$

On the other hand, the reservation process models the reservation or transmission status of the system. It is a 2-D process whose state variables are $\{N_r, N_b\}$. Its evolvement depends on the value of N_s. Although not completely true with varying N_s, the stationary distribution of this process is assumed to exist, denoted as $\underline{\Pi}^{(r,b|s)}$ [12]

$$\underline{\Pi}^{(r,b|s)} = \{\pi(n_r, n_b \,|\, n_s)\}$$
$$= \{P(N_r = n_r, N_b = n_b \,|\, N_s = n_s)\}.$$

Mathematically, this approach can be illustrated as

$$\pi(n_r, n_s, n_b) = \pi(n_r, n_b \,|\, n_s)\pi(n_s). \quad (2)$$

$\underline{\Pi}^{(s)}$ and $\underline{\Pi}^{(r,b|s)}$ will be evaluated instead of $\underline{\Pi}$. With this approach, the computational complexity can be reduced.

a) Find $\underline{\Pi}^{(s)}$: To simplify the notation, we will use $B(n, m, p)$ to denote the Binomial distribution with parameters (n, p)

$$B(n, m, p) = \binom{n}{m} p^m (1-p)^{n-m}. \tag{3}$$

Since the embedded point is at the beginning of each frame, there may be more than one voice user having state transitions. The speaking-silence process is a multibirth/multideath process. We need to construct the one-step transition probability matrix and solve the flow balance equations to find $\underline{\Pi}^{(s)}$. Denote the one-step transition probability matrix of the speaking-silence process as

$$\mathbf{P}^{(s)} = \{P(N_s(x+1) = j \mid N_s(x) = i)\}$$

where $i, j = 0, \ldots, N_v$ and $N_s(x)$ is the number of voice users in silence state at the beginning of the xth frame. Given q_a and q_s

$$P(N_s(x+1) = j \mid N_s(x) = i)$$
$$= \sum_{l=\max(0,i-j)}^{\min(i,N_v-j)} B(i, l, q_a) B(N_v - i, j - (i-l), q_s) \tag{4}$$

where l is the number of voice users leaving the silence state (becoming active) in the current frame. With the one-step transition probabilities, we can solve the following flow balance equations

$$\underline{\Pi}^{(s)} = \underline{\Pi}^{(s)} \mathbf{P}^{(s)} \quad \text{and} \quad \sum_{n_s} \pi(n_s) = 1 \tag{5}$$

and find the stationary distribution of the speaking-silence process, $\underline{\Pi}^{(s)}$.

b) Find $\underline{\Pi}^{(r,b|s)}$: Given the number of voice users in the silence state, $N_s = n_s$, the reservation process is a 2-D multibirth/multideath process. We also need to construct the one-step transition probability matrix to evaluate the stationary distribution. First, we define three probability functions which will be frequently used.

1) $\xi_v(s,t)$ is the probability that with s contending voice users and t backlogged data users, there is a voice success in the current **reservation** slot

$$\xi_v(s,t) = B(s, 1, p_t) B(t, 0, p_r). \tag{6}$$

2) $\xi_d(s,t)$ is the probability that with s contending voice users and t backlogged data users, there is a data success in the current **reservation** slot

$$\xi_d(s,t) = B(s, 0, p_t) B(t, 1, p_r). \tag{7}$$

3) $\Theta(s,t;c;l_1,k_1)$ is the probability that among s contending voice users and t backlogged data users, there are l_1 voice successes and k_1 data successes in c **reservation** slots

$$\Theta(s,t;c;l_1,k_1)$$
$$= [1 - \xi_v(s,t) - \xi_d(s,t)]\Theta(s,t;c-1;l_1,k_1)$$
$$+ \xi_v(s,t)\Theta(s-1,t;c-1;l_1-1,k_1)$$
$$+ \xi_d(s,t)\Theta(s,t-1;c-1;l_1,k_1-1). \tag{8}$$

The ending condition of this recursive evaluation is given in (9) at the bottom of the page.

With the help of these functions, the one-step transition probability of the reservation process, conditioned on the number of silent voice users $N_s = n_s$, can be evaluated as

$$P(N_r(x+1) = m, N_b(x+1) = n \mid N_r(x) = i, N_b(x) = j)$$
$$= \sum_{l=\max(0,m-i)}^{U_l} \sum_{k=\max(0,j-n)}^{U_k} \Phi(N_v - n_s - i, j; S_c, S - i; l, k)$$
$$\cdot B(i+l, m, 1 - q_s) B(N_d - j, n - (j-k), p_0) \tag{10}$$

where l is the number of voice users who obtain reservations successfully in the current frame and k is the number of data users successfully transmitted in the same frame. U_l and U_k are the upper limits of l and k, respectively. $\Phi(N_v - n_s - i, j; S_c, S - i; l, k)$ is the probability that among $(N_v - n_s - i)$ contending voice users and j backlogged data users, l voice users obtain reservations and k data users transmit information packets successfully in a frame with S_c reservation slots and $(S - i)$ available information slots. U_l, U_k, S_c, and $\Phi(\cdot)$ are protocol dependent, which are given as follows for PRMA, D-TDMA, and RAMA, respectively.

1) For PRMA, there are no dedicated reservation slots. The users try to access the channel by transmitting information packets directly. In addition, there is no limit on the number of voice slots which can be used in a frame. Therefore

$$S_c^{\text{PRMA}} = S - i$$
$$U_l^{\text{PRMA}} = \min(N_v - n_s - i, S - i)$$
$$U_k^{\text{PRMA}} = \min(j, N_d - n, S - (i+l))$$
$$\Phi(N_v - n_s - i, j; S_c, S - i; l, k)$$
$$= \Theta(N_v - n_s - i, j; \ S - i; \ l, k). \tag{11}$$

2) For D-TDMA, in each frame, the number of reservation slots is always S_r. The maximum number of voice slots per frame is S_v. In D-TDMA, successfully transmitting a reservation packet is not equivalent to making a successful reservation or transmitting an information

$$\Theta(s,t;c;l_1,k_1) = \begin{cases} [1 - \xi_v(s,t) - \xi_d(s,t)]^c & l_1 = k_1 = 0 \\ 0 & c < (l_1 + k_1) \text{ or } l_1 < 0 \text{ or} \\ & k_1 < 0. \end{cases}$$
$$\tag{9}$$

packet. When the traffic is heavy, some users may successfully transmit reservation packets, but cannot make reservations due to the shortage of available information slots. Therefore, we have (12), given at the bottom of the page.

3) For RAMA, in each frame, the number of auction slots is always S_r. The maximum number of voice slots per frame is also S_v. The probability $\Phi(N_v - n_s - i, j; S_r, S - i; l, k)$ can be greatly simplified due to the deterministic assignment feature of RAMA. In addition, from the protocol description, we know that the data user is always dropped if there is a voice user transmitting in the auction slot. Therefore, if the number of requesting voice users is larger than the number of auction slots, no data packet can succeed in this frame. Hence, we have (13), given at the bottom of the page.

With the one-step transition probabilities, we can construct the one-step transition matrix $\mathbf{P}^{(r,b|s)}$ and then solve the following flow balance equations to get $\underline{\Pi}^{(r,b|s)}$:

$$\underline{\Pi}^{(r,b|s)} = \underline{\Pi}^{(r,b|s)}\mathbf{P}^{(r,b|s)} \quad \text{and} \quad \sum_{n_r}\sum_{n_b}\pi(n_r, n_b \mid n_s) = 1 \tag{14}$$

B. Performance Measures

1) Voice Performance: Since the voice packet needs immediate delivery, it has to be discarded if its delay exceeds a certain limit. Because this packet dropping causes performance degradation, it should be limited. The packet-loss probability is used as the voice performance measure for this purpose. The packet-loss probability is defined as the average fraction of packets in a talkspurt dropped due to delayed transmission. In this paper, we assume that the packet will be dropped if it cannot be transmitted in a frame, i.e., the delay limit of voice packets is the frame duration T. Usually, a packet-loss probability less than 1% is acceptable to the user. Let $\mathbf{E}(x)$ be the expected value of a random variable x. With the stationary distribution of system states, $\pi(n_r, n_b, n_s)$, we can define the packet-loss probability P_{loss} as [18]

$$P_{\text{loss}} = \frac{\mathbf{E}(N_c) - \mathbf{E}(R)}{\mathbf{E}(N_c) + \mathbf{E}(N_r)}. \tag{15}$$

Here

$$\mathbf{E}(N_c) = \mathbf{E}(\text{num. of voice users in the contention state})$$
$$= \sum_{n_r=0}^{U_r}\sum_{n_s=0}^{N_v-n_r}\sum_{n_b=0}^{N_d}(N_v - n_r - n_s)\pi(n_r, n_s, n_b) \tag{16}$$

$$\mathbf{E}(N_r) = \mathbf{E}(\text{num. of voice users in the reservation state})$$
$$= \sum_{n_r=0}^{U_r}\sum_{n_s=0}^{N_v-n_r}\sum_{n_b=0}^{N_d}n_r\pi(n_r, n_s, n_b) \tag{17}$$

$$S_c^{\text{D-TDMA}} = S_r$$
$$U_l^{\text{D-TDMA}} = \min(N_v - n_s - i, S_r, S_v - i)$$
$$U_k^{\text{D-TDMA}} = \min(j, N_d - n, S_r - l, S - (i+l))$$
$$\Phi(N_v - n_s - i, j; S_c, S - i; l, k) = \begin{cases} \Theta(N_v - n_s - i, j; \ S_r; \ l, k) \\ \quad l < S_v - i, l + k < S - i \\ \sum_{l_1=l}^{\min(S_r-k, N_v-n_s-i)}\Theta(N_v - n_s - i, j; \ S_r; \ l_1, k) \\ \quad l = S_v - i, l + k < S - i \\ \sum_{k_1=k}^{\min(S_r-l, j)}\Theta(N_v - n_s - i, j; \ S_r; \ l, k_1) \\ \quad l < S_v - i, l + k = S - i \\ \sum_{l_1=l}^{\min(S_r-k, N_v-n_s-i)}\sum_{k_1=k}^{\min(S_r-l_1, j)}\Theta(N_v - n_s - i, j; \ S_r; \ l_1, k_1) \\ \quad l = S_v - i, l + k = S - i. \end{cases} \tag{12}$$

$$S_c^{\text{RAMA}} = S_r$$
$$U_l^{\text{RAMA}} = \min(N_v - n_s - i, S_r, S_v - i)$$
$$U_k^{\text{RAMA}} = \min(j, N_d - n, S_r - l, S - (i+l))$$
$$\Phi(N_v - n_s - i, j; \ S_r, S - i; \ l, k) = \begin{cases} 1 & S_r \leq N_v - n_s - i, \ l = \min(N_v - n_s - i, S_r, S_v - i) \\ & \text{and } k = 0 \\ 1 & S_r > N_v - n_s - i, \ l = \min(N_v - n_s - i, S_r, S_v - i) \\ & \text{and } k = \min(j, N_d - n, S_r - (N_v - n_s - i), S - (i+l)) \\ 0 & \text{otherwise.} \end{cases} \tag{13}$$
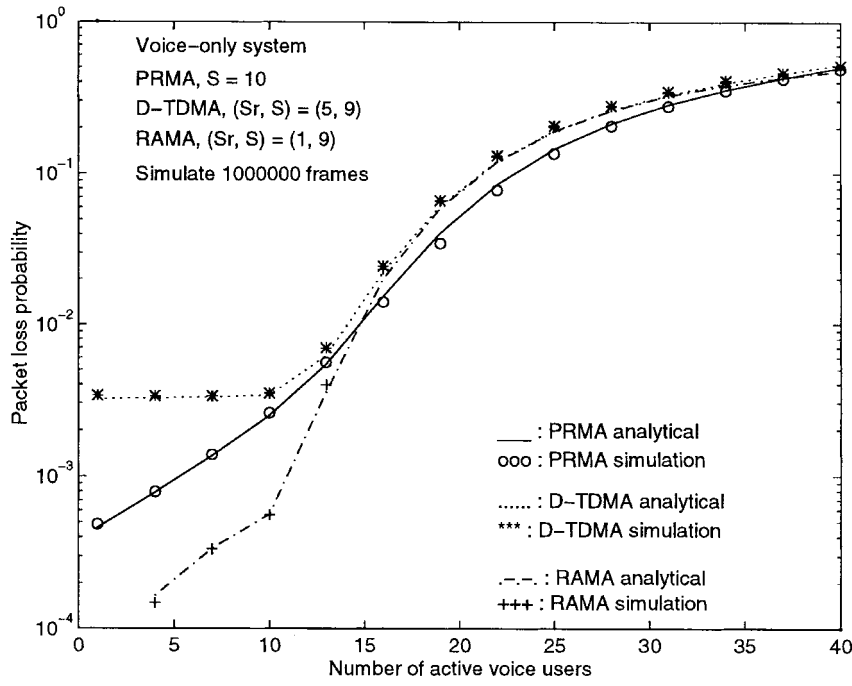
Fig. 4. Voice-packet-loss probabilities of D-TDMA, PRMA, and RAMA in voice-only scenario.

$$\mathbf{E}(R) = \mathbf{E}(\text{num. of successful voice reservations}$$

$$\text{per frame}) = \sum_{n_r=0}^{U_r} \sum_{n_s=0}^{N_v-n_r} \sum_{n_b=0}^{N_d}$$

$$\left[\sum_l \sum_k l\Phi(N_v - n_r - n_s, n_b;\right.$$

$$\left. S_c, S - n_r; l, k) \right] \cdot \pi(n_r, n_s, n_b)$$

$$(18)$$

where $\Phi(N_v - n_r - n_s, n_b;\ S_c, S - n_r;\ l, k)$ is defined in (11) for PRMA, in (12) for D-TDMA, in (13) for RAMA, and as follows.

1) For PRMA

$$S_c^{\text{PRMA}} = S - n_r$$
$$U_r^{\text{PRMA}} = \min(N_v, S)$$
$$0 \le l \le \min(N_v - n_r - n_s, S - n_r)$$
$$0 \le k \le \min(n_b, S - (n_r + l)). \tag{19}$$

2) For D-TDMA and RAMA

$$S_c^{\text{D-TDMA(RAMA)}} = S_r$$
$$U_r^{\text{D-TDMA(RAMA)}} = \min(N_v, S_v)$$
$$0 \le l \le \min(N_v - n_r - n_s, S_r, S_v - n_r)$$
$$0 \le k \le \min(n_b, S_r - l, S - (n_r + l)). \tag{20}$$

2) Data Performance:

We use the average throughput and delay as the data performance measures. As usual, the average data throughput β is defined as the average number of data packets successfully transmitted in a frame

$$\beta = \sum_{n_r=0}^{U_r} \sum_{n_s=0}^{N_v-n_r} \sum_{n_b=0}^{N_d} \left[\sum_l \sum_k k\Phi(N_v - n_r\right.$$

$$\left. - n_s, n_b; S_c, S - n_r; l, k) \right]$$

$$\pi(n_r, n_s, n_b) \tag{21}$$

where S_c and U_r and the ranges of l and k are determined by (19) for PRMA and by (20) for D-TDMA and RAMA. The evaluation of function $\Phi(\cdot)$ is the same as before for each protocol.

The average delay D_d is the average time that a data packet spends in the buffer until the beginning of the successful transmission. By Little's Formula

$$D_d = \frac{\mathbf{E}(N_b)}{\beta} - 1 \tag{22}$$

where

$$\mathbf{E}(N_b) = \mathbf{E}(\text{num. of backlogged data users in a frame})$$

$$= \sum_{n_r=0}^{U_r} \sum_{n_s=0}^{N_v-n_r} \sum_{n_b=0}^{N_d} n_b \pi(n_r, n_s, n_b). \tag{23}$$
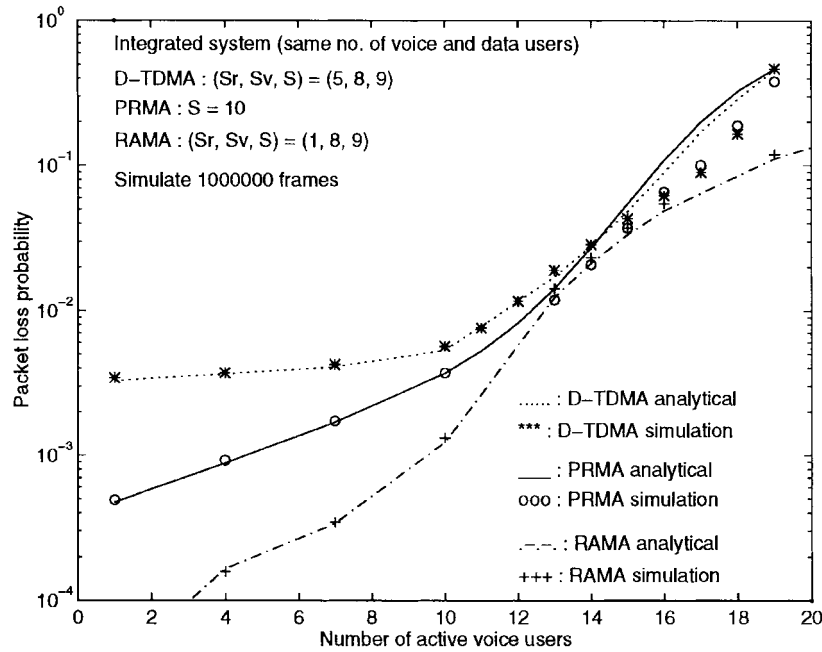
Fig. 5.   Voice-packet-loss probability of integrated D-TDMA, PRMA, and RAMA systems.
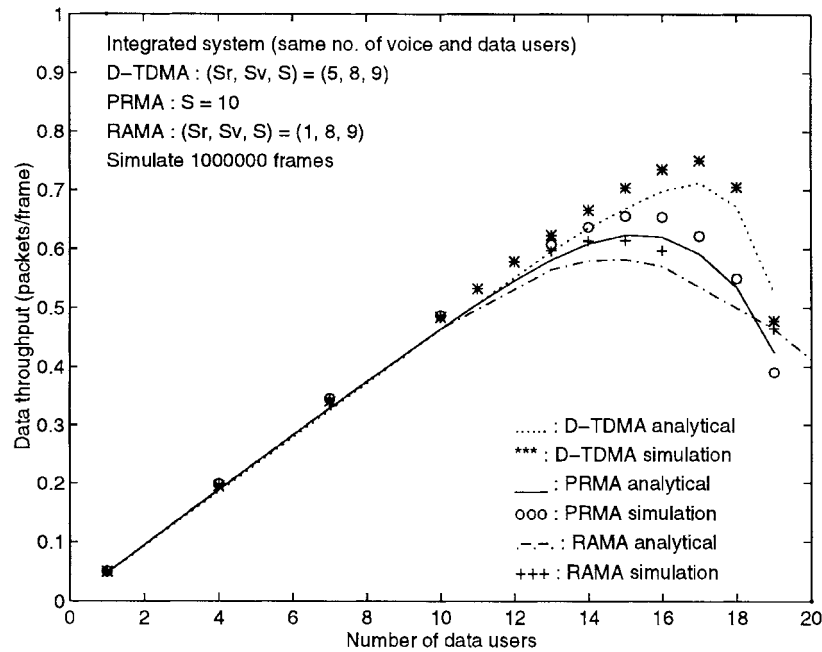

Fig. 6.   Data throughput performance of integrated D-TDMA, PRMA, and RAMA systems.

Again, U_r is determined by (19) for PRMA and (20) for D-TDMA (RAMA).

C. Special Case: Voice-Only System

We will derive the performance model for the voice-only system in this section as a special case of the above integrated model. Without data, $\Theta(\cdot)$ in (8) can be simplified to

$$\Theta(s; c; l_1)$$
$$= P(l_1 \text{ out of } s \text{ contending voice users successfully}$$
$$\text{transmitting in } c \text{ \textbf{reservation} slots})$$

$$= [1 - \xi_v(s)]\Theta(s; c - 1; l_1)$$
$$+ \xi_v(s)\Theta(s - 1; c - 1; l_1 - 1) \qquad (24)$$

where $\xi_v(s)$ is equal to $\xi_v(s, 0)$ in (6). The ending condition of this recursive evaluation is

$$\Theta(s; c; l_1) = \begin{cases} [1 - \xi_v(s)]^c, & l_1 = 0 \\ 0, & c < l_1. \end{cases} \qquad (25)$$

In the voice-only scenario, the system can be described by two state variables $\{N_r, N_s\}$, and its evolvement can be modeled as a 2-D Markov Chain. We can again split this 2-D Markov Chain into the speaking-silence process and the reservation process. Each of them is a one-dimensional (1-D)
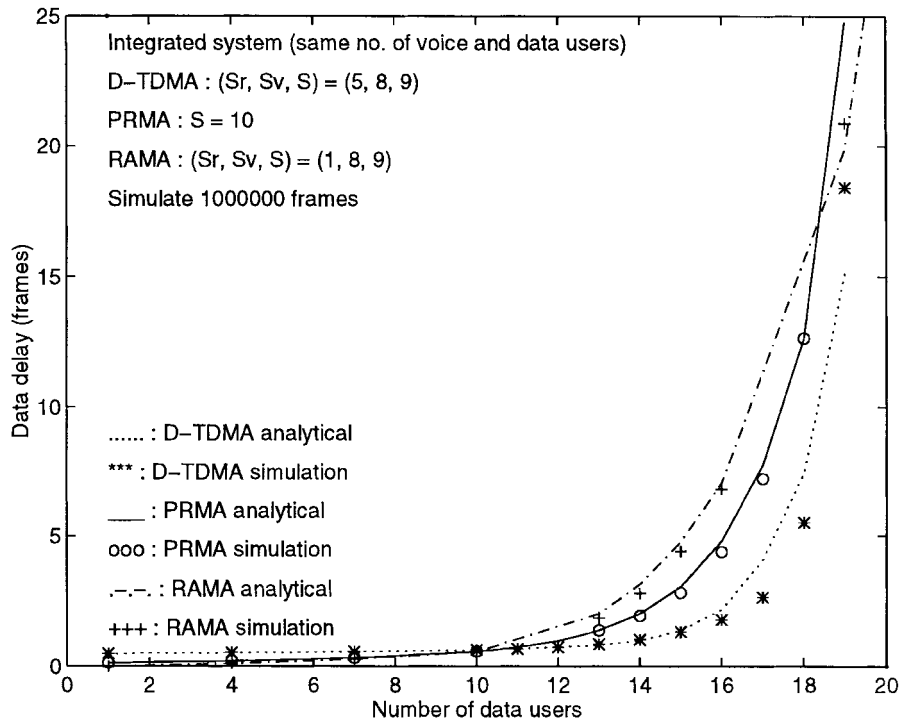
Fig. 7. Data-delay performance of integrated D-TDMA, PRMA, and RAMA systems.

process which can be solved easily. The stationary distribution $\underline{\Pi}^{(s)}$ of the speaking-silence process can be evaluated by using exactly the same approach as that in the integrated model.

Now, the reservation process is only a 1-D multi-birth/multideath process, with N_r being its state variable. It is a conditional process, depending on the value of N_s. The one-step transition probability, conditioned on $N_s = n_s$, can be evaluated as

$$
\begin{aligned}
&P(N_r(x+1) = m \mid N_r(x) = i) \\
&= \sum_{l=\max(0,m-i)}^{U_l} \Phi(N_v - n_s - i; S_c, S - i; l) \\
&\quad \times B(i+l, m, 1-q_s)
\end{aligned}
\tag{26}
$$

where l is the number of successful voice reservations in the current frame. Still, U_l, S_c, and $\Phi(\cdot)$ are protocol dependent.

1) For PRMA

$$
\begin{aligned}
S_c^{\text{PRMA}} &= S - i \\
U_l^{\text{PRMA}} &= \min(N_v - n_s - i, S - i) \\
\Phi(N_v &- n_s - i; S_c, S - i; l) \\
&= \Theta(N_v - n_s - i; S - i; l).
\end{aligned}
\tag{27}
$$

2) For D-TDMA, see (28), given at the bottom of the page.
3) For RAMA, see (29), given at the bottom of the page.

With the one-step transition probabilities, we can construct the one-step transition matrix $\mathbf{P}^{(r|s)}$ and then solve the flow balance equations in (14) to get the stationary distribution of reservation process, $\underline{\Pi}^{(r|s)}$.

With $\underline{\Pi}^{(s)}$ and $\underline{\Pi}^{(r|s)}$, $\underline{\Pi}$ can be evaluated. Again, we use the voice-packet-loss probability as the performance measure,

$$
\begin{aligned}
S_c^{\text{D-TDMA}} &= S_r \\
U_l^{\text{D-TDMA}} &= \min(N_v - n_s - i, S_r, S - i) \\
\Phi(N_v - n_s - i; S_c, S - i; l) &= \begin{cases} \Theta(N_v - n_s - i; S_r; l), & l < S - i \\ \sum_{l_1=l}^{\min(N_v - n_s - i, S_r)} \Theta(N_v - n_s - i; S_r; l_1), & l = S - i. \end{cases}
\end{aligned}
\tag{28}
$$

$$
\begin{aligned}
S_c^{\text{RAMA}} &= S_r \\
U_l^{\text{RAMA}} &= \min(N_v - n_s - i, S_r, S - i) \\
\Phi(N_v - n_s - i; S_c, S - i; l) &= \begin{cases} 1 & l = \min(N_v - n_s - i, S_r, S - i) \\ 0 & \text{otherwise.} \end{cases}
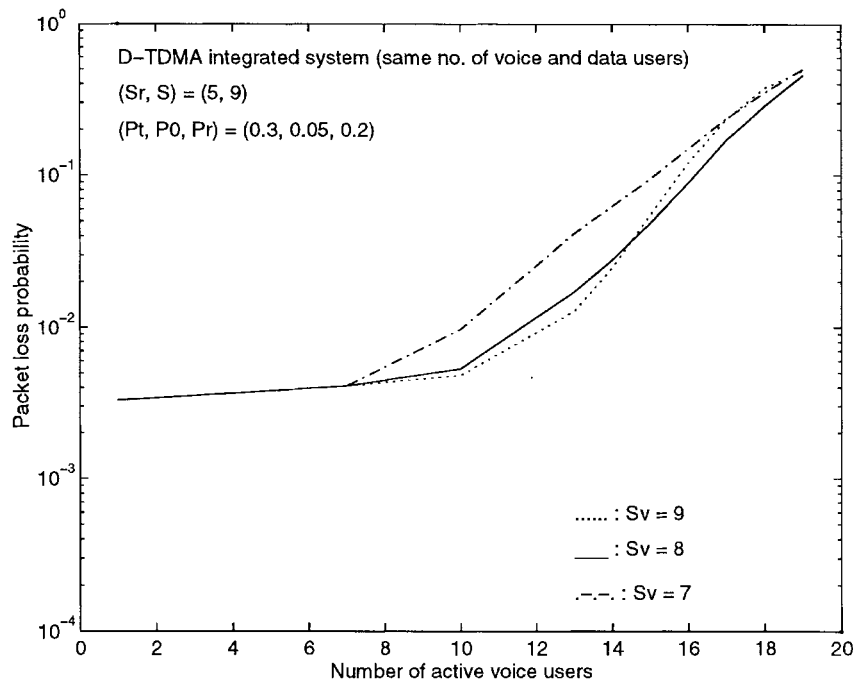\end{aligned}
\tag{29}
$$

Fig. 8.   Impact of S_v on the packet-loss probability of integrated D-TDMA.

TABLE I
SYSTEM PARAMETERS FOR VOICE-ONLY D-TDMA, PRMA, AND RAMA

Protocol	S_r	S	p_t
D-TDMA	5	9	0.3
PRMA	–	10	0.3
RAMA	1	9	1

which is defined in (15), but with

$$\mathbf{E}(N_c) = \sum_{n_r=0}^{U_r} \sum_{n_s=0}^{N_v-n_r} (N_v - n_r - n_s)\pi(n_r, n_s) \qquad (30)$$

$$\mathbf{E}(N_r) = \sum_{n_r=0}^{U_r} \sum_{n_s=0}^{N_v-n_r} n_r\pi(n_r, n_s) \qquad (31)$$

$$\mathbf{E}(R) = \sum_{n_r=0}^{U_r} \sum_{n_s=0}^{N_v-n_r} \sum_{l} l$$
$$\times \Phi(N_v - n_r - n_s; S_c, S - n_r; l)\pi(n_r, n_s) \quad (32)$$

where $S_c = S - n_r$ for PRMA and $S_c = S_r$ for D-TDMA or RAMA and

$$U_r = \min(S, N_v)$$

and

$$0 \le l \le \min(N_v - n_r - n_s, S_c, S - n_r).$$

V. NUMERICAL EXAMPLES

We present numerical results for PRMA, D-TDMA, and RAMA, using the performance evaluation model in Section IV. The analytical results are compared to those from simulation. Here, the event-driven simulation model is used with the source traffic assumptions described earlier. One million frames are simulated for each case, and appropriate statistics (e.g., voice-packet-loss probability, average data throughput, and average data delay) are obtained. Then, the impact of system parameters, such as the voice-permission probability, the data-retransmission probability, and the maximum number of voice slots per frame, etc., on the system behavior is also investigated in this section. Note that in this paper, to simplify the analysis, we assume the same number of data and voice users in the system. One information slot in D-TDMA is assumed divided into five reservation slots and one information slot in RAMA is equivalent to one auction slot. In addition, the frame length T is set to 16 ms.

A. Voice-Only System

The performance of voice-only D-TDMA, PRMA, and RAMA systems will be evaluated using our general model under specified system parameters, which are summarized in Table I. Here, the total bandwidth used in these three systems is the same. With these parameters, the analytical results are evaluated and compared to those from simulations in Fig. 4. It is observed that our analytical results match the simulation results very well. It is also found that when the traffic is light, RAMA provides the best performance since it can efficiently resolve the collision problem in the contention-based reservation-type protocols. When the traffic is heavy, PRMA provides slightly better performance than D-TDMA and RAMA with specified parameters. This is because in D-TDMA and RAMA, reservation or auction slots are fixed overhead, which consumes system bandwidth. On the contrary, all the slots can be used for the information transmission in PRMA.
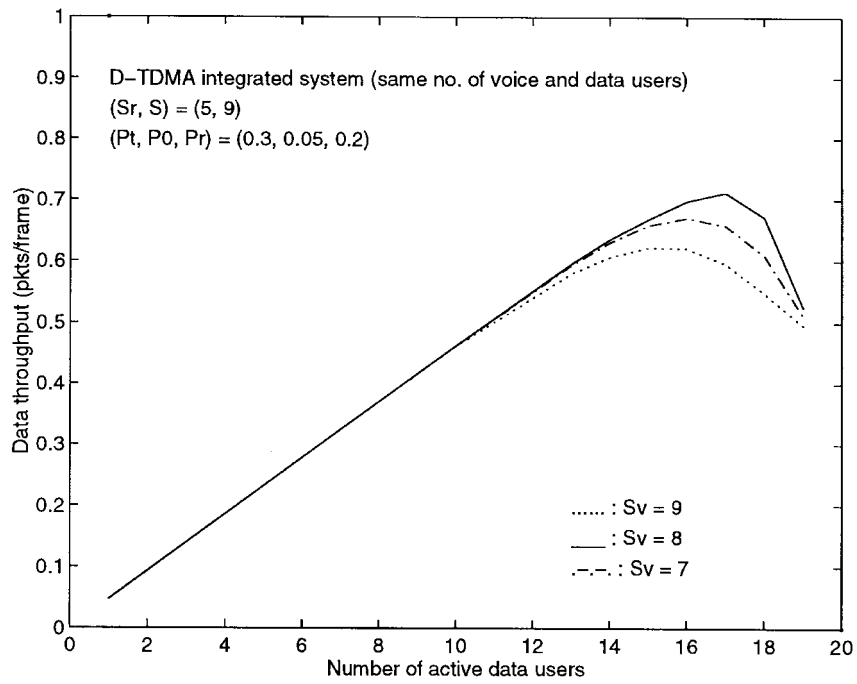
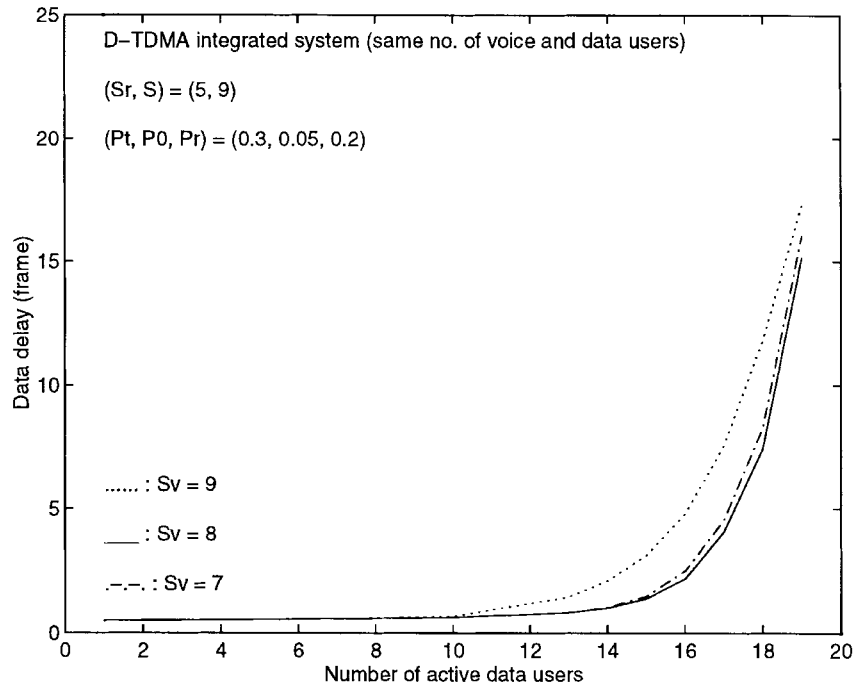Fig. 9. Impact of S_v on the data throughput of integrated D-TDMA.

Fig. 10. Impact of S_v on the data delay of integrated D-TDMA.

TABLE II
SYSTEM PARAMETERS FOR INTEGRATED D-TDMA, PRMA, AND RAMA

Protocol	S_r	S_v	S	p_t	p_0	p_r
D-TDMA	5	8	9	0.3	0.05	0.2
PRMA	–	10	10	0.3	0.05	0.2
RAMA	1	8	9	1	0.05	1

B. Integrated Voice and Data System

Now, we will investigate the integrated performance of reservation-type systems, also using D-TDMA, PRMA, and RAMA as examples. First, we justify the correctness of the analytical model by comparing the analytical results to those from simulation. Then, the impact of various system
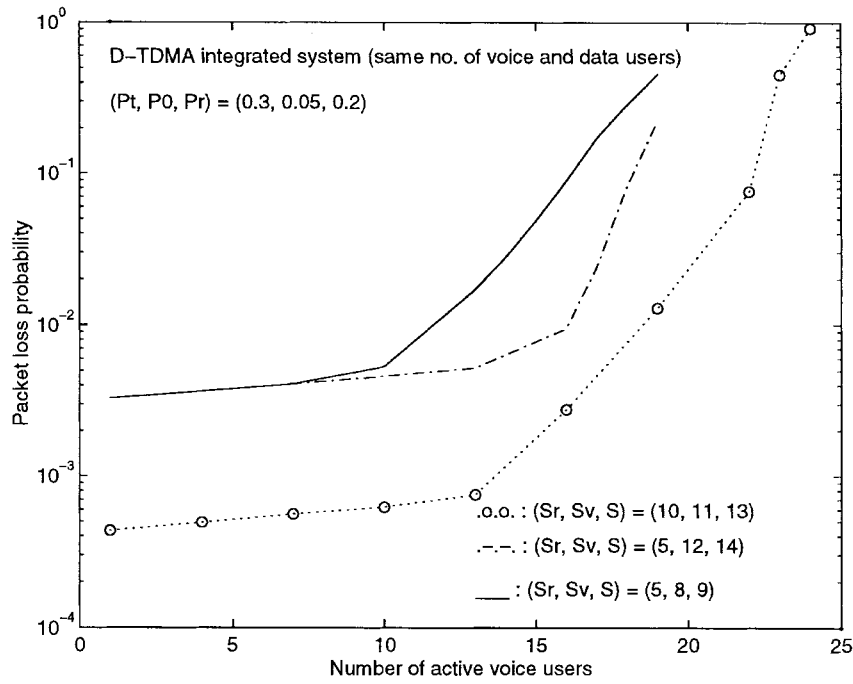
Fig. 11. Impact of (S_r, S_v, S) on the voice-packet-loss probability of integrated D-TDMA.
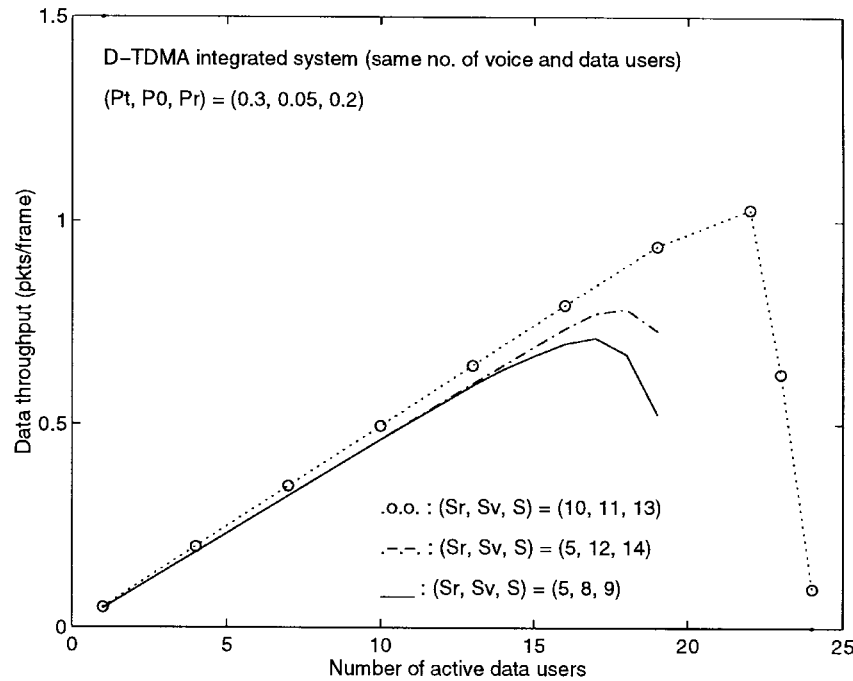

Fig. 12. Impact of (S_r, S_v, S) on the data throughput performance of integrated D-TDMA.

parameters on the system performance will be studied for D-TDMA, PRMA, and RAMA.

Under the specific system parameters summarized in Table II, the voice-packet-loss probability, the data throughput and delay performance of these three systems are shown in Figs. 5–7, respectively. The analytical results are compared to the simulation. It is found that our unified model is accurate, especially in the region of interest (i.e., the region with the voice-packet-loss probability less than 1% and with reasonable

data delay). But when the system becomes too congested, the analytical results will be pessimistic. This is because we assume the probability that a voice user returns to the silence state before it obtains a reservation is zero. However, in a heavily congested system, all the packets in a talkspurt may be dropped in which case this probability is nonzero.

Comparing these three systems, it is observed that RAMA always provides the best voice performance and almost always the worst data performance (except under very heavy traffic
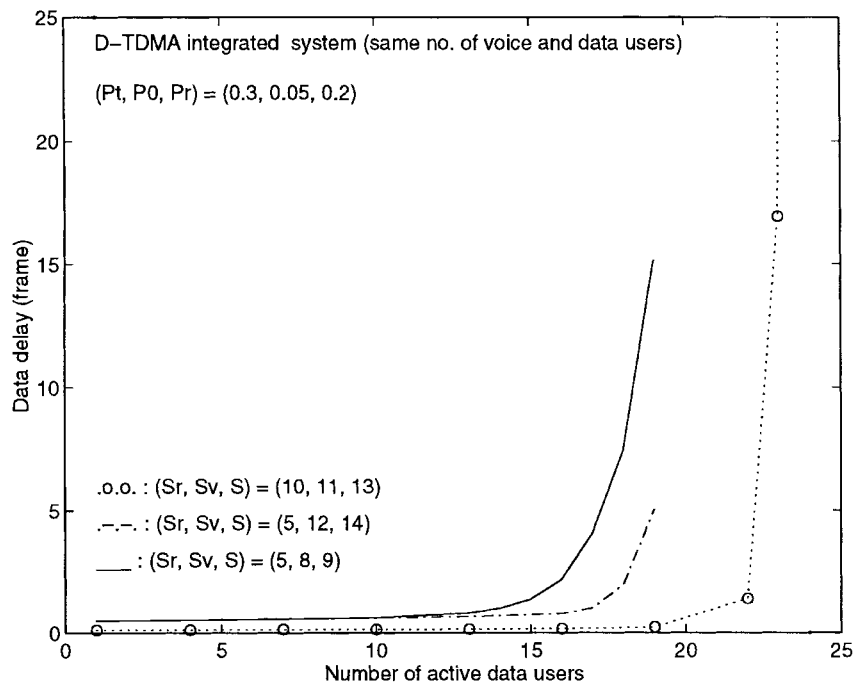
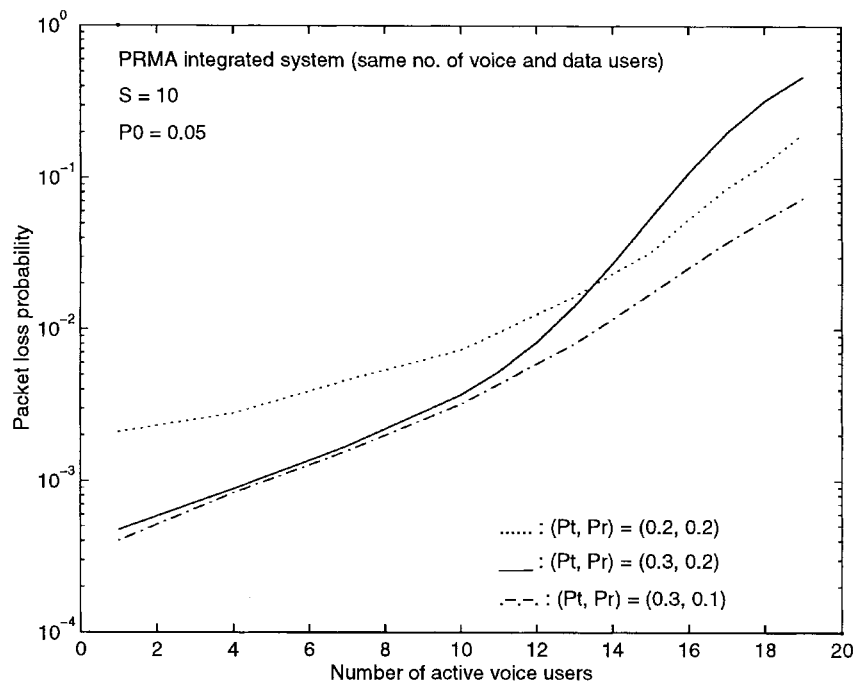Fig. 13.   Impact of (S_r, S_v, S) on the data-delay performance of integrated D-TDMA.

Fig. 14.   Impact of (p_t, p_r) on the voice loss probability of integrated PRMA.

scenario). This is because in RAMA, the voice traffic is given much higher priority than that in D-TDMA and PRMA. If there is a voice user transmitting in the auction slot, data users always drop out. This guarantees the voice performance at the expense of the data performance. Of course, the tradeoff in voice and data performance in RAMA can also be controlled by giving voice user a permission probability. When the traffic is light, RAMA is superior to D-TDMA and PRMA when both the voice and data performance are considered. It is

also observed that when the number of users increases, the performance of RAMA degrades much more slowly than that of D-TDMA or PRMA. This is because in RAMA, an available information slot will always be assigned to a requesting user, independent of the traffic load. On the other hand, the unsuccessful contentions in D-TDMA and PRMA under heavy traffic cause the degradation in system performance.

Now, the impact of various system parameters (such as the voice-permission probability, data-retransmission probability,
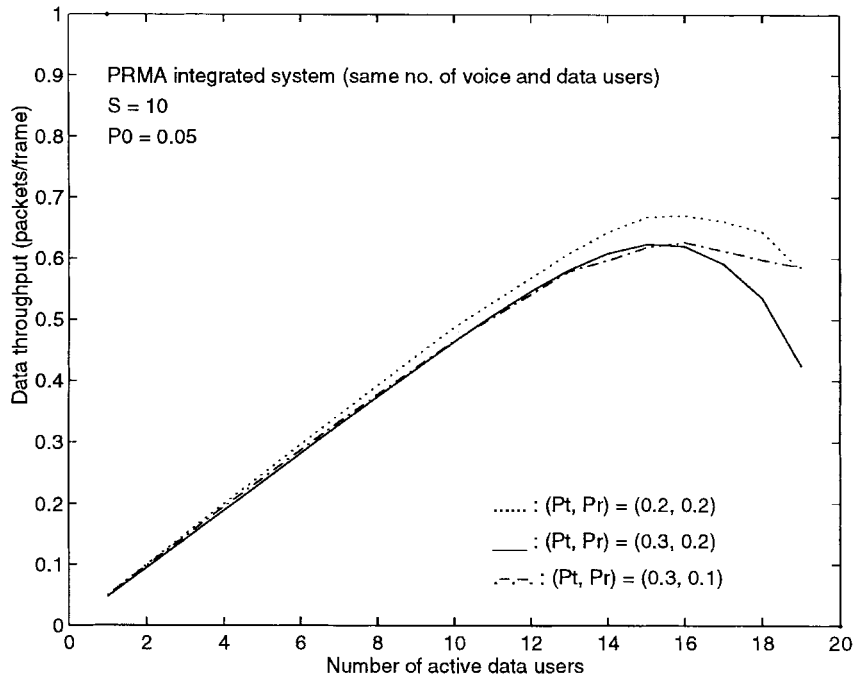
Fig. 15. Impact of (p_t, p_r) on the data throughput performance of integrated PRMA.
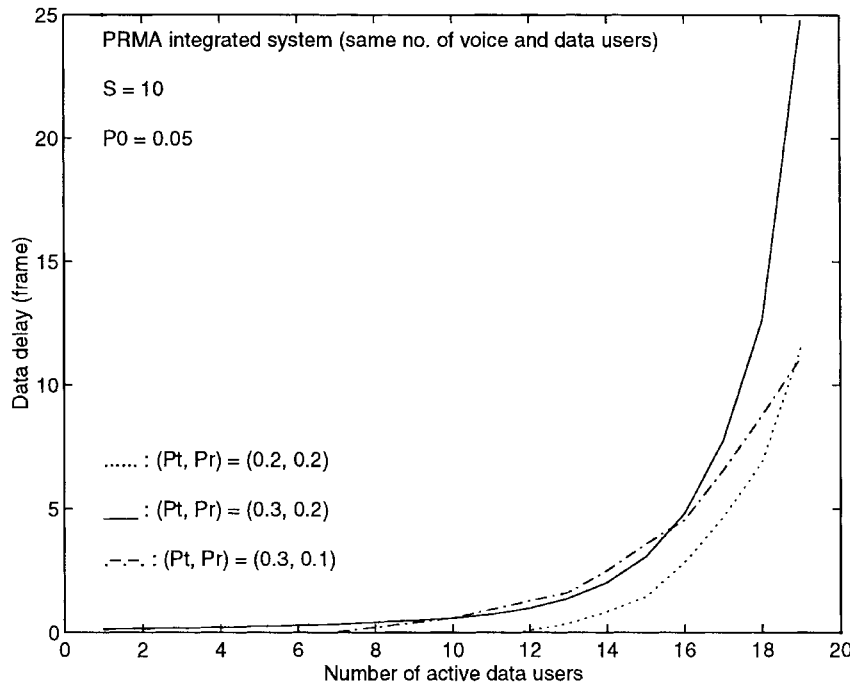

Fig. 16. Impact of (p_t, p_r) on the data-delay performance of integrated PRMA.

maximum number of voice slots per frame, etc.) on the system performance will be investigated for D-TDMA, PRMA, and RAMA.

1) D-TDMA: First, we investigate the influence of the maximum number of voice slots S_v on the system performance. Keeping the number of information slots per frame $S = 9$ and $(p_t, p_0, p_r) = (0.3, 0.05, 0.2)$ unchanged, we vary the

maximum number of voice slots S_v, $S_v = 7, 8, 9$. The voice-packet-loss probability, the data throughput and delay are illustrated in Figs. 8–10. It is found that although more bandwidth is reserved for data when $S_v = 7$, the data performance is not improved, and the joint performance of voice and data is worse compared to $S_v = 8$. This is caused by a large number of voice contentions in reservation slots. The
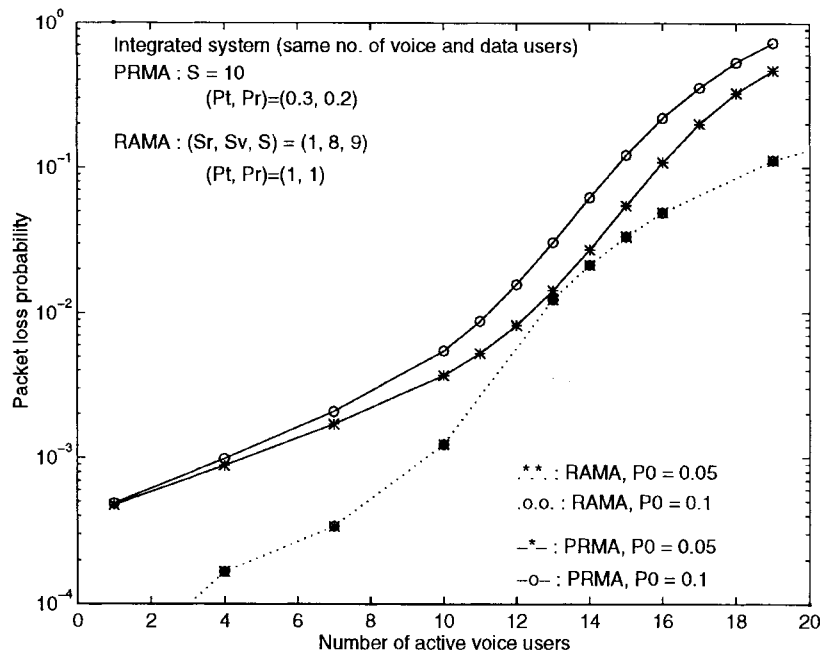
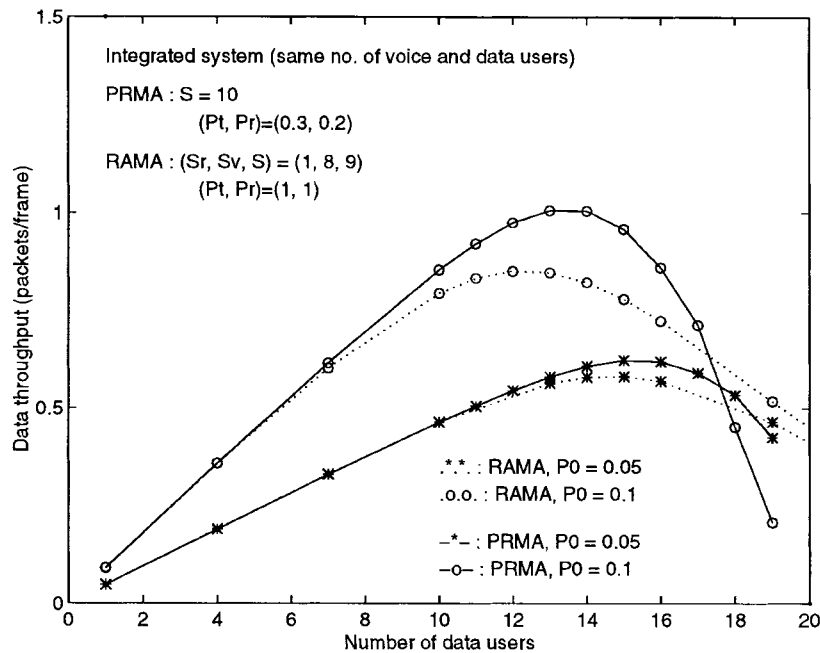Fig. 17. Impact of data traffic (p_0) on the voice-packet-loss probability of integrated PRMA and RAMA.

Fig. 18. Impact of data traffic (p_0) on the data throughput performance of integrated PRMA and RAMA.

contention success probability is relatively lower. Considering both the voice and data performance, $S_v = 8$ gives the best performance.

There are two reasons for the performance degradation of D-TDMA. One is the shortage of reservation slots and another is the shortage of information slots. We set $(p_0, p_r) = (0.05, 0.2)$ and investigate the influence of (S_r, S_v, S) on the system performance. The voice-packet-loss probability, data throughput, and delay performance are drawn in Figs. 11–13, respectively, for the three cases of $(S_r, S_v, S) = (5, 8, 9), (5, 12, 14)$, and

$(10, 11, 13)$. The performance is improved as the number of information slots is increased from nine in $(S_r, S_v, S) = (5, 8, 9)$ to 14 in $(S_r, S_v, S) = (5, 12, 14)$. But the improvement is not significant, especially for data. With the same bandwidth, $(S_r, S_v, S) = (5, 12, 14)$ does not provide better performance. It can be seen in the figure that when $(S_r, S_v, S) = (10, 11, 13)$, although the number of information slots decreases by one, the system performance improves dramatically from that with $(S_r, S_v, S) = (5, 12, 14)$. This is because when $(S_r, S_v, S) = (5, 12, 14)$, the bottleneck of
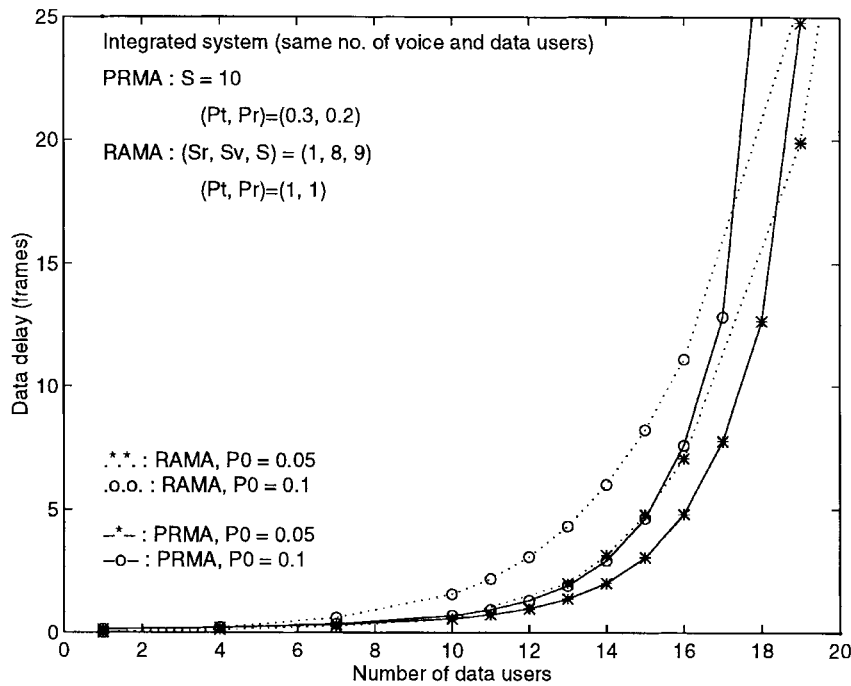
Fig. 19.   Impact of data traffic (p_0) on the data-delay performance of integrated PRMA and RAMA.

the system is in the reservation process. Too many users contending in a few reservation slots causes frequent collisions. Even though there are still some information slots available, no one can get access to them because of contention. Therefore, increasing S_r in this case can significantly improve the system performance.

2) PRMA: The influence of p_t and p_r on the integrated PRMA system performance is illustrated in Figs. 14–16. Three system scenarios are considered, namely, $(p_t, p_r) = (0.3, 0.2), (0.3, 0.1)$, and $(0.2, 0.2)$. It can be found that these two parameters have some influence on the system performance and have to be selected carefully to maintain acceptable voice and data performance. Among these three cases, the system with $(p_t, p_r) = (0.3, 0.1)$ exhibits the best performance when both voice and data performance are considered.

3) RAMA: We will next compare the different impacts of data traffic on the performance of integrated PRMA and RAMA systems. For PRMA, we choose $S = 10$ and $(p_t, p_r) = (0.3, 0.2)$. For RAMA, (S_r, S_v, S) is set to $(1, 8, 9)$. Varying the data arrival rate from 0.05 packets/frame to 0.1 packets/frame, its influence on the performance of these two systems is illustrated in Figs. 17–19, for the voice-packet-loss probability, data throughput, and delay, respectively. It can be found that unlike PRMA (note that D-TDMA has the same behavior as that of PRMA), the amount of data traffic has no impact on the voice performance of RAMA. This is because in RAMA, voice is given higher priority. If a voice terminal is transmitting in the auction slot, data terminals always drop out. Therefore, in RAMA, the data traffic is transparent to voice. But it has to be noted that this superior voice performance is achieved at the expense of slightly degraded data performance.

It can be seen that the data performance of RAMA under these two traffic scenarios is worse than that of PRMA.

VI. CONCLUSION

We observe that although different reservation-type protocols, such as D-TDMA, PRMA, RAMA, etc., are different at a first glance, all of them can be described by several common state variables whose evolvement follows a Markov process. With this observation, a unified performance model has been developed in this paper for the integrated voice and data system with reservation-type MA schemes. The Markovian model is developed, the stationary distribution of system states is evaluated, and three performance measures are defined based on the known stationary distribution. Using PRMA, D-TDMA, and RAMA as examples, the performance of both voice-only and integrated systems have been studied analytically. The analytical results are verified by simulation. It is found that the analytical model is quite accurate, especially in the region of interest. Comparing these three systems, it is observed that RAMA provides the best voice performance at the expense of slightly degraded data performance. The impact of system parameters (such as the voice-permission probability, data-retransmission probability, etc.) on the system performance has also been investigated using the analytical model. One phenomenon which is worth noting is that in RAMA, the data traffic load has no influence on the voice performance, i.e., the data traffic is transparent to voice. In addition, with the increase in the number of users in the system, the performance of RAMA degrades much more slowly than that of D-TDMA or PRMA. This is because in RAMA, an available information slot will always be assigned to a requesting user, independent of the traffic load. On the other hand, the frequent unsuccessful

contentions in D-TDMA and PRMA degrade the system performance significantly when the traffic is heavy. This deterministic assignment feature makes RAMA very attractive for resource sharing in the integrated system.

REFERENCES

[1] N. Amitay, "Distributed switching and control with fast resource assignment/handoff for personal communications systems," *IEEE J. Select. Areas Commun.*, vol. 11, no. 6, pp. 842–849, 1993.

[2] ——, "Resource auction multiple access (RAMA): Efficient method for fast resource assignment in decentralised wireless PCS," *Electron. Lett.*, vol. 28, no. 8, pp. 799–801, 1992.

[3] N. Amitay and L. J. Greenstein, "Resource Auction Multiple Access (RAMA) in the cellular environment," *IEEE Trans. Veh. Technol.*, vol. 43, no. 4, pp. 1101–1111, 1994.

[4] N. Amitay and S. Nanda, "Speech multiplexing in wireless PCS employing resource auction multiple access (RAMA)," *Electron. Lett.*, vol. 28, no. 25, pp. 2294–2295, 3, 1992.

[5] ——, "Resource auction multiple access (RAMA) for statistical speech multiplexing in wireless PCS," *IEEE Trans. Veh. Technol.*, vol. 43, no. 3, pp. 584–596, 1994.

[6] G. Falk *et al.*, "Integration of voice and data in the wideband packet satellite network," *IEEE J. Select. Areas Commun.*, vol. SAC-1, pp. 1076–1083, Dec. 1983.

[7] D. J. Goodman, "Trends in cellular and cordless communications," *IEEE Commun. Mag.*, vol. 29, pp. 31–40, June 1991.

[8] D. J. Goodman and S. X. Wei, "Efficiency of packet reservation multiple access," *IEEE Trans. Veh. Technol.*, vol. 40, no. 1, pp. 170–180, 1991.

[9] J. G. Gruber and N. Le, "Performance requirements for integrated voice/data networks," *IEEE J. Select. Areas Commun.*, vol. SAC-1, pp. 981–1005, Dec. 1983.

[10] S. S. Lam and L. Kleinrock, "Packet switching in a multiple access broadcast channel: Dynamic control procedures," *IEEE Trans. Commun.*, vol. COM-23, pp. 891–904, Sept. 1975.

[11] V. O. K. Li and X. X. Qiu, "Personal communication systems (PCS)," *Proc. IEEE*, vol. 83, no. 9, pp. 1210–1243, 1995.

[12] N. M. Mitrou, TH. D. Orinos, and E. N. Protonotarios, "A reservation multiple access protocol for microcellular mobile-communication systems," *IEEE Trans. Veh. Technol.*, vol. 39, no. 4, pp. 340–351, 1990.

[13] S. Nanda, "Analysis of packet reservation multiple access: Voice and data integration for wireless networks," in *IEEE GLOBECOM'90*, pp. 1984–1988.

[14] S. Nanda, D. J. Goodman, and U. Timor, "Performance of PRMA: A packet voice protocol for cellular systems," *IEEE Trans. Veh. Technol.*, vol. 40, no. 3, pp. 584–598, 1991.

[15] H. Qi and R. Wyrwas, "Markov analysis for PRMA performance study," in *IEEE VTC'94,* pp. 1184–1188.

[16] X. X. Qiu and V. O. K. Li, "On the capacity of packet reservation multiple access with capture in personal communication systems," *IEEE Trans. Veh. Technol.*, vol. 45, no. 4, pp. 666–675, 1996.

[17] L. Roberts, "Dynamic allocation of satellite capacity through packet reservation," in *AFIP Conf. Proc.*, 1973, vol. 42, pp. 711–716.

[18] I. Rubin and S. Shambayati, "Performance evaluation of a reservation random access scheme for packetized wireless systems with call control," in *IEEE GLOBECOM'92*, pp. 16–20.

[19] N. D. Wilson, R. Ganesh, K. Joseph, and D. Raychaudhuri, "Packet CDMA versus dynamic TDMA for multiple access in an integrated voice/data PCN," *IEEE J. Select. Areas Commun.*, vol. 11, no. 6, pp. 870–884, 1993.

[20] G. Wu, K. Mukumoto, and A. Fukuda, "Analysis of an integrated voice and data transmission system using packet reservation multiple access," *IEEE Trans. Veh. Technol.*, vol. 43, no. 2, pp. 289–297, 1994.

Xiaoxin Qiu (S'93–M'96) received the B.E. and M.E. degrees from Tsinghua University, China, in 1990 and 1991, respectively, and the Ph.D. degree from University of Southern California, Los Angeles, in 1996.

She is with the Broadband Wireless Systems Research Department, AT&T Laboratories–Research. Her research interests are in the areas of wireless communications networks, personal communication systems, and multimedia communications.

Victor O. K. Li (S'80–M'81–SM'86–F'92), for a photograph and biography, see this issue, p. 172.