PAPER Special Section on Multi-dimensional Mobile Information Networks

Performance Analysis of Dynamic Resource Allocation with Finite Buffers in Cellular Networks

Wei-Yeh CHEN[†a)], ***Student Member***, **Jean-Lien C. WU**[††], ***and*** **Hung-Huan LIU**[†††], ***Nonmembers***

SUMMARY In this paper, we analyzed the performance of dynamic resource allocation with channel de-allocation and buffering in cellular networks. Buffers are applied for data traffic to reduce the packet loss probability while channel de-allocation is exploited to reduce the voice blocking probability. The results show that while buffering data traffic can reduce the packet loss probability, it has negative impact on the voice performance even if channel de-allocation is exploited. Although the voice blocking probability can be reduced with large slot capacity, the improvement decreases as the slot capacity increases. On the contrary, the packet loss probability increases as the slot capacity increases. In addition to the mean value analysis, the delay distribution and the 95% delay of data packets are provided.

key words: *dynamic resource allocation, channel de-allocation, delay distribution, 95% delay, GPRS*

1. Introduction

The provision of integrated voice/data services over wireless links requires that radio resources be allocated to users efficiently to meet the different quality of service (QoS). The bandwidth of wireless links is inherently limited and is generally much less than its wired counterpart. Thus, resource allocation schemes for cellular networks are exploited to utilize the channel utilization efficiently [1]–[4]. It is quite well known that dynamic resource allocation allows communication systems to utilize their resources more efficiently than the traditional fixed allocation schemes.

Cellular networks such as General Packet Radio Service (GPRS) introduced with GSM and Universal Mobile Telecommunications System (UMTS) nowadays can provide both circuit-switched and packet-switched services [5], [6]. In general, real-time applications, such as voice services, are best provided via circuit-switched service, while most data applications are more efficiently provided by packet-switched service. Both GPRS and UMTS allow multiple channels to be allocated to a user served in either circuit-switched mode or packet-switched mode to fulfill its QoS requirement. For instance, channel allocation in GPRS is flexible, where one to eight channels can be allocated to a user or one channel can be shared by several users. In UMTS Terrestrial Radio Access based on TD-CDMA (UTRA-TDD) [7], 240 resource units (15 timeslots in a frame and 16 different code sequences in each timeslot) can be dynamically distributed among users.

Ni and Haggman showed that employing the multi-slot service in GPRS will result in higher blocking probability and longer delay than using the single-slot service, and these effects can be alleviated by implementing the resource allocation scheme with flexible, or dynamic, multi-slot service [8]. Thus, the design of effective and dynamic bandwidth allocation to satisfy different service demands is important in wireless networks. Several bandwidth allocation schemes [9], [10] have been proposed to improve the system performance in integrated voice/data wireless cellular networks. By adopting dynamic resource allocation with possible channel de-allocation, both the blocking probability of voice calls and the wastage probability of radio resource can be effectively reduced. However, none of the above studies considered channel de-allocation mechanism in their analyses.

The analysis and comparison of the performance of dynamic resource allocation with/without channel de-allocation in GSM/GPRS networks can be found in a previous work [11]. In this paper, we will focus on the performance analysis of the dynamic resource allocation with finite buffers. Buffers are applied for data traffic to reduce the packet loss probability while channel de-allocation is exploited to reduce the voice blocking probability. In addition to the mean system delay, the knowledge of delay distribution of packet transfer is essential for real-time applications. In this work, we developed an approach to derive the delay distribution of data packets. We also evaluated the 95% delay which is defined as the maximum delay guaranteed in 95% of all data transfers

The rest of this paper is organized as follows. Section 2 describes the radio resource allocation schemes adopted in this paper. Analytical models of dynamic resource allocation and delay distribution of data traffic are described in Sect. 3. In Sect. 4, comparisons of numerical and simulation results are provided. The effects of different resource allocation capabilities are then investigated. Section 5 concludes this work.

Manuscript received October 24, 2003.
Manuscript revised February 5, 2004.
Final manuscript received March 15, 2004.

[†]The author is with the Department of Information Management, National Taiwan University of Science and Technology, Taipei, Taiwan, R.O.C.

[††]The author is with the Department of Electronic Engineering, National Taiwan University of Science and Technology, Taipei, Taiwan, R.O.C.

[†††]The author is with the Department of Electronic Engineering, Chung Yuan Christian University, Chungli, Taiwan, R.O.C.

a) E-mail: wychen@kwit.edu.tw

2. Resource Allocation

The complete sharing strategy [12] is adopted that the radio resources are completely shared by voice calls and data users, i.e., no guard channels are reserved for either voice or data service. Buffers are exploited to queue data packets which find no channels available upon arrivals.

For data traffic, dynamic resource allocation is applied. To a data request of n channels, the network allocates at most n channels. Assume there are N available channels in the system upon a data request arrival. If $N \geq n$, n channels are allocated to the request. If $0 < N < n$, N channels are allocated to the request. If $N = 0$, the request is queued in the buffer or blocked by the system depending on if there are buffers available. Whenever there are service completions, the queued data packets are allocated with available channels, and are served using the first come first served (FCFS) discipline. Throughout the paper, "slot" and "channel" are used to have the same meaning.

Channel de-allocation is applied for voice arrivals when there are no free channels. One slot of an existing multi-slot data user is de-allocated for a voice arrival. The de-allocated user uses the remaining allocated channels to finish its transmission. When there are no free channels and no multi-slot users in service, the voice call request will be blocked. Channel de-allocation is not applied for data traffic since it will result in higher voice blocking probability, longer data transmission time, and lower channel utilization, with the only benefit of accommodating more data users in the system.

3. Analytical Model

In this section, we first use the continuous-time Markov chain to analyze the performance of dynamic resource allocation in terms of the voice blocking probability, data packet loss probability, and the mean system delay for integrated voice/data services. The phase-type renewal process [13] is exploited to derive the delay distribution and the 95% delay of data packets.

3.1 Mean Value Analysis

To analyze the performance of dynamic n-slot allocation with buffering for integrated voice/data services, it requires n+2 dimensions to formulate the system, i.e., one dimension for voice traffic, n dimensions for data packets in service, and one dimension for data packets queued in the buffer. With a total number of C channels and B buffers, the number of states will be $(C+1)^2(\lfloor C/2 \rfloor +1)\cdots(\lfloor C/n \rfloor +1)(B+1)$, where $\lfloor x \rfloor$ is the floor function of x. Due to the large state space, the queueing analysis will be too difficult to be tractable for numerical solution. To simplify the analysis, we consider the case of $n = 2$, and validate the analysis by simulation experiments. For $n > 2$, the performance results are obtained via simulation.

The analysis focuses on a single cell in isolation and assumes that the network is symmetric and the traffic is homogenous. Let the arrivals of voice call requests form a Poisson process with a mean rate of λ_v. The service time of voice calls is assumed to be exponentially distributed with a mean of $1/\mu_v$. The arrival process of data packets is assumed to be Poisson with a mean rate of λ_d. The service time of each data packet is exponentially distributed with a mean of $1/\mu_d$. The total number of channels in the system is C, and the buffer size is B.

Let the state (i, j, k, l) denote that there are i voice calls, j 1-slot data users in service, and k 2-slot data users in service, and l queued data packets in the system. $\pi_{i,j,k,l}$ denotes the state probability of the system in state (i, j, k, l). Figure 1 shows the state transition diagram of the system. The case of $l > 0$ represents that there are no free channels in the system, therefore, one existing 2-slot data user will be de-allocated upon a voice arrival, and a data arrival will be queued in the buffer if there are buffer space. The dash line in the figure represents the state transition caused by de-allocating a 2-slot user upon a voice arrival. Let S be the set of feasible states,

$$\begin{aligned} S = &\{(i, j, k, 0) \mid 0 \leq i + j + 2k \leq C, 0 \leq i \leq C, \\ &\quad 0 \leq j \leq C, \text{and } 0 \leq k \leq \lfloor C/2 \rfloor\} \\ &\cup \{(i, j, k, l) \mid i + j + 2k = C, 0 \leq i \leq C, \\ &\quad 0 \leq j \leq C, 0 \leq k \leq \lfloor C/2 \rfloor, \text{and } 1 \leq l \leq B\} \qquad (1) \end{aligned}$$

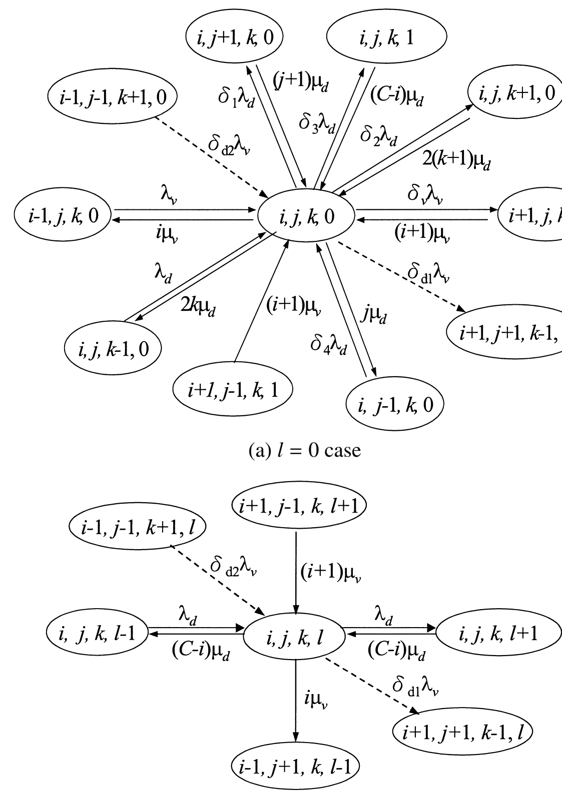


Fig. 1 The state transition diagram.

An indicator function $\varphi(i, j, k, l)$ is used to indicate whether the state (i, j, k, l) is feasible or not, i.e., $\varphi(i, j, k, l) = 1$ if $(i, j, k, l) \in S$. For all $(i, j, k, l) \in S$, the balance equations can be expressed in the following.

$l = 0$ case:

$$\begin{aligned}
&\pi_{ijk0}(\delta_v\lambda_v\varphi(i+1, j, k, 0) + \delta_1\lambda_d\varphi(i, j+1, k, 0) \\
&\quad + \delta_2\lambda_d\varphi(i, j, k+1, 0) + \delta_3\lambda_d\varphi(i, j, k, 1) \\
&\quad + i\mu_v\varphi(i-1, j, k, 0) + j\mu_d\varphi(i, j-1, k, 0) \\
&\quad + 2k\mu_d\varphi(i, j, k-1, 0) \\
&\quad + \delta_{d1}\lambda_v\varphi(i+1, j+1, k-1, 0)) \\
&= \lambda_v\pi_{i-1,j,k,0}\varphi(i-1, j, k, 0) \\
&\quad + \delta_4\lambda_d\pi_{i,j-1,k,0}\varphi(i, j-1, k, 0) \\
&\quad + \lambda_d\pi_{i,j,k-1,0}\varphi(i, j, k-1, 0) \\
&\quad + (i+1)\mu_v\pi_{i+1,j,k,0}\varphi(i+1, j, k, 0) \\
&\quad + (i+1)\mu_v\pi_{i+1,j-1,k,1}\varphi(i+1, j-1, k, 1) \\
&\quad + (j+1)\mu_d\pi_{i,j+1,k,0}\varphi(i, j+1, k, 0) \\
&\quad + 2(k+1)\mu_d\pi_{i,j,k+1,0}\varphi(i, j, k+1, 0) \\
&\quad + (C-i)\mu_d\pi_{i,j,k,1}\varphi(i, j, k, 1) \\
&\quad + \delta_{d2}\lambda_v\pi_{i-1,j-1,k+1,0}\varphi(i-1, j-1, k+1, 0)
\end{aligned} \tag{2}$$

$l > 0$ case:

$$\begin{aligned}
&\pi_{ijkl}(\lambda_d\varphi(i, j, k, l+1) \\
&\quad + i\mu_v\varphi(i-1, j+1, k, l-1) \\
&\quad + (C-i)\mu_d\varphi(i, j, k, l-1) \\
&\quad + \delta_{d1}\lambda_v\varphi(i+1, j+1, k-1, l)) \\
&= \lambda_d\pi_{i,j,k,l-1}\varphi(i, j, k, l-1) \\
&\quad + (i+1)\mu_v\pi_{i+1,j-1,k,l+1}\varphi(i+1, j-1, k, l+1) \\
&\quad + (C-i)\mu_d\pi_{i,j,k,l+1}\varphi(i, j, k, l+1) \\
&\quad + \delta_{d2}\lambda_v\pi_{i-1,j-1,k+1,l}\varphi(i-1, j-1, k+1, l)
\end{aligned} \tag{3}$$

where

$$\delta_v = \begin{cases} 1, & \text{if } i + j + 2k < C \\ 0, & \text{otherwise} \end{cases} \tag{4}$$

$$\delta_1 = \begin{cases} 1, & \text{if } i + j + 2k = C - 1 \\ 0, & \text{otherwise} \end{cases} \tag{5}$$

$$\delta_2 = \begin{cases} 1, & \text{if } i + j + 2k \le C - 2 \\ 0, & \text{otherwise} \end{cases} \tag{6}$$

$$\delta_3 = \begin{cases} 1, & \text{if } i + j + 2k = C \\ 0, & \text{otherwise} \end{cases} \tag{7}$$

$$\delta_4 = \begin{cases} 1, & \text{if } i + (j-1) + 2k = C - 1 \\ 0, & \text{otherwise} \end{cases} \tag{8}$$

$$\delta_{d1} = \begin{cases} 1, & \text{if } i + j + 2k = C, \\ & \text{and } 1 \le k \le \lfloor C/2 \rfloor \\ 0, & \text{otherwise} \end{cases} \tag{9}$$

$$\delta_{d2} = \begin{cases} 1, & \text{if } i + j + 2k = C, \\ & \text{and } 0 \le k \le \lfloor C/2 \rfloor - 1 \\ 0, & \text{otherwise} \end{cases} \tag{10}$$

The last term on both sides of Eqs. (2) and (3) is attributed to channel de-allocation caused by voice arrivals. δ_v, δ_1 (or δ_4), and δ_2 are the conditions of transition caused by voice call arrivals, data arrivals using 1 slot, and data arrivals using 2 slots, respectively. δ_3 is the condition of transition caused by data arrivals when there are no free channels. δ_{d1} and δ_{d2} are the conditions for channel de-allocation upon voice arrivals. Applying the constraint $\sum_S \pi_{i,j,k,l} = 1$ to the set of balance equations, we can obtain the steady-state probability $\pi_{i,j,k,l}$ to evaluate the voice blocking probability, packet loss probability, system delay of data packets, mean queue length, and probability of de-allocation.

A voice call arrival will be blocked when there are no free channels and no 2-slot data users in the system. Thus, the blocking probability of voice calls, P_{vb}, can be expressed as

$$P_{vb} = \sum_{l=0}^{B} \sum_{i+j=C} \pi_{i,j,0,l} \tag{11}$$

The packet loss probability, P_{loss}, is the probability that a data packet finds there are no free channels and no buffer space upon arrival, and can be obtained as

$$P_{loss} = \sum_{i+j+2k=C} \pi_{i,j,k,B} \tag{12}$$

The system delay of data packets, W, is defined as the average elapsed time from the instant of data packets admitted into the system to service completion, and can be expressed as

$$W = \frac{1}{\lambda_d(1 - P_{loss})} \cdot \sum_S (j + k + l) \cdot \pi_{i,j,k,l} \tag{13}$$

The mean queue length, L, is

$$L = \sum_{l=1}^{B} \sum_{i+j+2k=C} l \cdot \pi_{i,j,k,l} \tag{14}$$

The probability of de-allocation, $P_{de-alloc}$, is the probability that an existing 2-slot data user will be de-allocated upon a voice arrival, and can be obtained as

$$P_{de-alloc} = \frac{\lambda_v}{\lambda_d(1 - P_{loss})} \cdot \sum_{l=0}^{B} \sum_{i+j+2k=C \& k \ge 1} \pi_{i,j,k,l} \tag{15}$$

3.2 Delay Analysis of Data Traffic

When the mean service time of one traffic class is much larger than that of the other class, the decomposition approximation approach [14] can be exploited and the delay distribution of the concerned traffic class can be analyzed. We are interested in the time distribution of a packet entering the queue until it is completely transmitted. Consider the case that there are m channels available for data packets and l queued packets waiting to be transmitted upon a packet arrival. This data packet arrival has to queue for a time of $(l+1)$ packets to obtain service. The service process of this

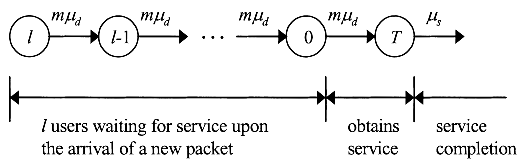


Fig. 2 The phase-type renewal process for a data packet arrival finding l packets already waiting in queue.

packet can be described by the phase-type renewal process and is shown in Fig. 2.

Given that there are m channels for data packets, the departure rate of a packet from the queue, and the service rate of a packet are respectively $m\mu_d$ and μ_s. The mean service rate μ_s can be derived from Eqs. (13) and (14),

$$\frac{1}{\mu_s} = W - \frac{L}{\lambda_d(1 - P_{loss})} \tag{16}$$

In this work, we assume the data service time is exponentially distributed with a mean of $1/\mu_s$. The service rate matrix, $R_{m,l}$, is defined as

$$R_{m,l} = \begin{bmatrix} -m\mu_d & m\mu_d & 0 & \cdots & 0 \\ 0 & -m\mu_d & m\mu_d & \cdots & 0 \\ 0 & 0 & -m\mu_d & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & -\mu_s \end{bmatrix}_{l+2,l+2} \tag{17}$$

The mean system delay, $\bar{t}_{m,l}$, and the system delay distribution, $F_{m,l}(t)$, are given respectively as:

$$\bar{t}_{m,l} = \frac{1}{\mu_s} + \frac{l+1}{m\mu_d}, \quad \mathrm{m} \neq 0 \tag{18}$$

$$F_{m,l}(t) = 1 - \alpha e^{R_{m,l}t}\varepsilon, \quad \mathrm{m} \neq 0 \tag{19}$$

where $\alpha = (1, 0, \cdots, 0)_{l+2}$ is the entrance vector, and ε is a column vector with all entries being one. An arriving data packet will find the system ready to serve, having l packets waiting in queue, or being blocked with respective probabilities denoted as $P_{m,S}$, $P_{m,l}$, and $P_{m,b}$, and are given by

$$P_{m,S} = \sum_{j+2k<m} \pi_{C-m,j,k,0} \tag{20}$$

$$P_{m,l} = \sum_{j+2k=m} \pi_{C-m,j,k,l}, \quad 0 \leq l < B \tag{21}$$

$$P_{m,b} = \sum_{j+2k=m} \pi_{C-m,j,k,B} \tag{22}$$

Thus, given m channels available, the mean system delay is

$$\bar{t}_m = P_{m,S}\frac{1}{\mu_s} + \sum_{l=0}^{B-1} P_{m,l} \cdot \bar{t}_{m,l} \tag{23}$$

and the mean system delay for data packets is

$$\bar{t} = \frac{1}{(1 - P_{loss})} \sum_{m=1}^{C} \bar{t}_m \tag{24}$$

Table 1 The probability that an admitted data user finds all channels being occupied by voice calls for B=10 and mean packet length being 2.

Data traffic load (Erlang)	Probability
1	0.00432
5	0.00000059
10	$< 10^{-8}$

The denominator, 1-P_{loss}, indicates the portion of data packets that are admitted into the system. Similarly, the delay distribution, given m channels available, is

$$F_m(t) = P_{m,S} \cdot F_{m,S}(t) + \sum_{l=0}^{B-1} P_{m,l} \cdot F_{m,l}(t) \tag{25}$$

where $F_{m,S}(t)$ is the service time distribution. Finally, the system delay distribution for data packets is

$$F(t) = \frac{1}{(1 - P_{loss})} \sum_{m=1}^{C} F_m(t) \tag{26}$$

Thereafter, according to the definition, the 95% delay can be obtained with the following expression

$$D = \max\{t \mid F(t) \leq 0.95\} \tag{27}$$

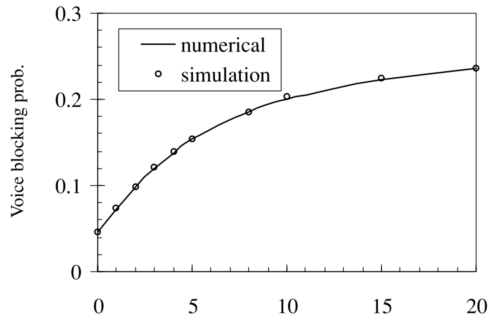
In the analysis, we do not consider the delay contributed by the condition that all channels are occupied by voice calls upon a data arrival, i.e., m=0. In fact, this delay can be neglected because the probability that all channels are occupied by voice calls is very small, which can be verified by evaluating the state probability. Table 1 depicts the probability that an admitted data user finds all channels being occupied by voice calls for different data traffic load with buffer size being 10 and mean packet length being 2. For simplicity, we do not consider to de-allocate existing data packets with more than one channel upon voice arrivals.

4. Numerical and Simulation Results

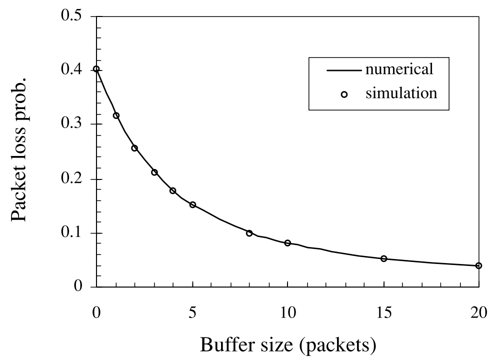
4.1 Validation of Analysis

To validate the numerical results, an event-driven simulator is developed. We will consider the GPRS environment because it supports both circuit-switched and packet-switched services and it also supports channel de-allocation. The total number of channels in a cell is set to be 24. The mean arrival rate of voice calls is taken to be 0.1386 calls/sec, and the mean service time of voice calls is 120 seconds. The voice traffic load is chosen to be 16.63 Erlang corresponding to a 2% blocking probability for 24 channels. The mean arrival rate of data packets is a system parameter and is chosen to be in the range of 0.5 to 6 packets/sec. The mean service time of data packets is taken to be 2 seconds if one channel is allocated. The buffer size ranges from 0 to 20.

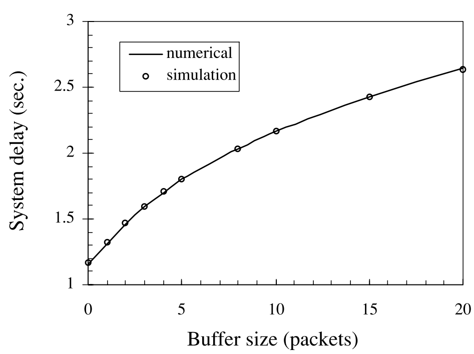
Figure 3 shows the performance comparisons of numerical results and simulation results for different buffer size with data arrival rate being 5 packets/sec. It can be seen



(a) Voice blocking probability

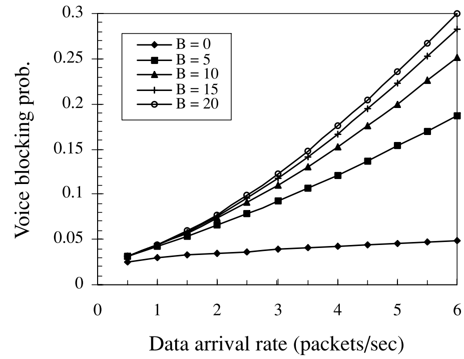


(b) Data packet loss probability

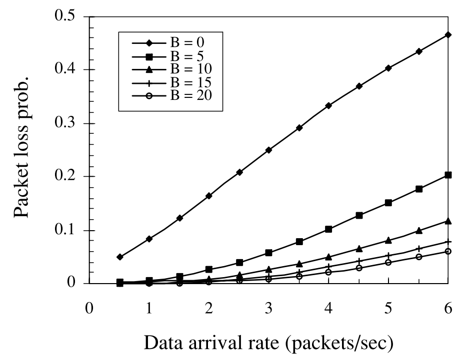


(c) Mean system delay

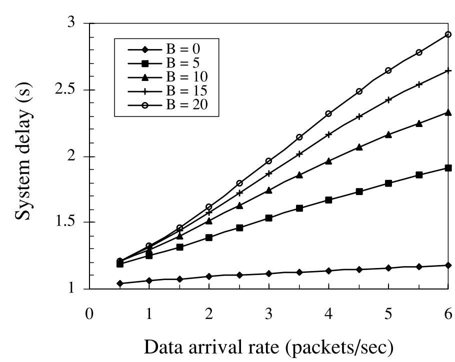
Fig. 3 Performance comparisons of numerical results and simulation results for λ_d=5 packets/sec.



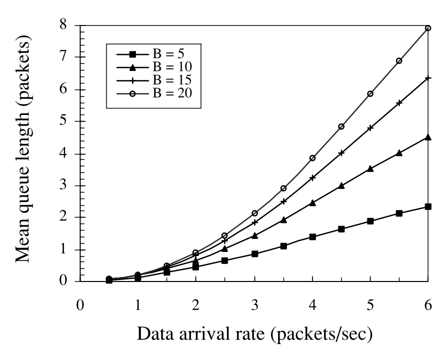
(a) Voice blocking probability



(b) Data packet loss probability



(c) Mean system delay



(d) Mean queue length

Fig. 4 System performance versus buffer size.

that numerical results match very well with simulation results for all performance metrics. Therefore, we can use the results obtained from simulation to investigate the impact of dynamic n-slot ($n > 2$) allocation in the latter subsection.

Figure 4 plots the system performance for different buffer sizes as a function of data arrival rate. From Fig. 4(a), it is clear that buffering data traffic has negative impact on the voice performance even channel de-allocation is exploited. The reason is that whenever there are service completions (voice or data), the queued packets will obtain services immediately, causing voice call arrivals to be blocked when there are no channels available and no 2-slot data users in service. Therefore, the voice blocking probability will increase with the increase of data buffer size. It is worth noting that at high data load, voice blocking probability is still below 5% for $B = 0$ due to the contribution of channel de-allocation.

On the contrary, the packet loss probability decreases as the buffer size increases as shown in Fig. 4(b). In addition, the system delay of data packet and mean queue length

Table 2 Comparisons of the mean system delay for ρ_d =10 and B=10.

Mean packet length	Mean system delay (s)		Difference
	Markov chain	Phase-type	
0.5	0.576	0.578	0.35%
2	2.164	2.187	1.06%

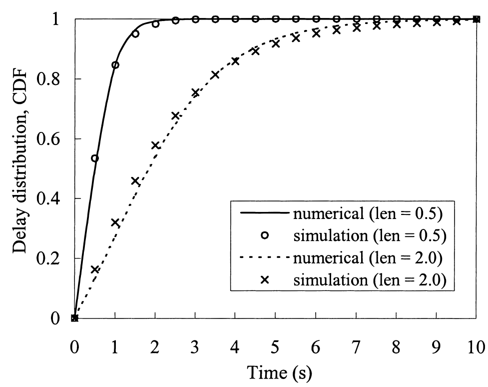


Fig. 5 Delay distribution of data packets for ρ_d=10 and B=10.

increases with data traffic load and buffer size as shown in Figs. 4(c) and 4(d). It can be seen that although buffering has large impact on the system performance, the impact of buffer size to the system performance decreases as buffer size increases.

4.2 Delay Distribution of Data Traffic

Before investigating the delay distribution of data traffic, the accuracy of the phase-type renewal process approach is verified by comparing the result obtained from the Markov chain approach. Table 2 compares the mean system delay obtained from different analytical approaches for various mean packet length when data traffic load ρ_d=10 and buffer size B=10. The mean packet length of 0.5 (or 2) implies that the mean service time of data packets is 0.5 (or 2) seconds if one channel is allocated. It can be seen that the results of these two approaches are very close to each other.

Figure 5 shows the system delay distribution for the mean packet length of 0.5 and 2, respectively, when the data traffic load ρ_d=10 and buffer size B=10. It can be seen that the results for shorter packet length matches better. This is because for shorter packet length, the mean service time of data traffic will be much smaller than that of voice traffic, therefore, applying the decomposition approximation approach gives more accurate results. The comparisons of 95% system delay under different traffic load are shown in Fig. 6. The 95% delay increases rapidly with increased data traffic load at light load. The error between the numerical and simulation results arises from the simplification that we do not consider de-allocation of existing data packets with more than one channel upon voice arrivals. Therefore, the delay value obtained from simulation results will be larger than that obtained from numerical results.

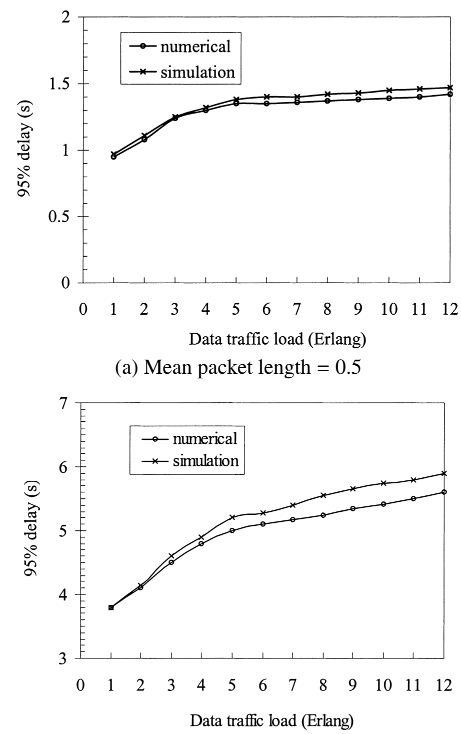


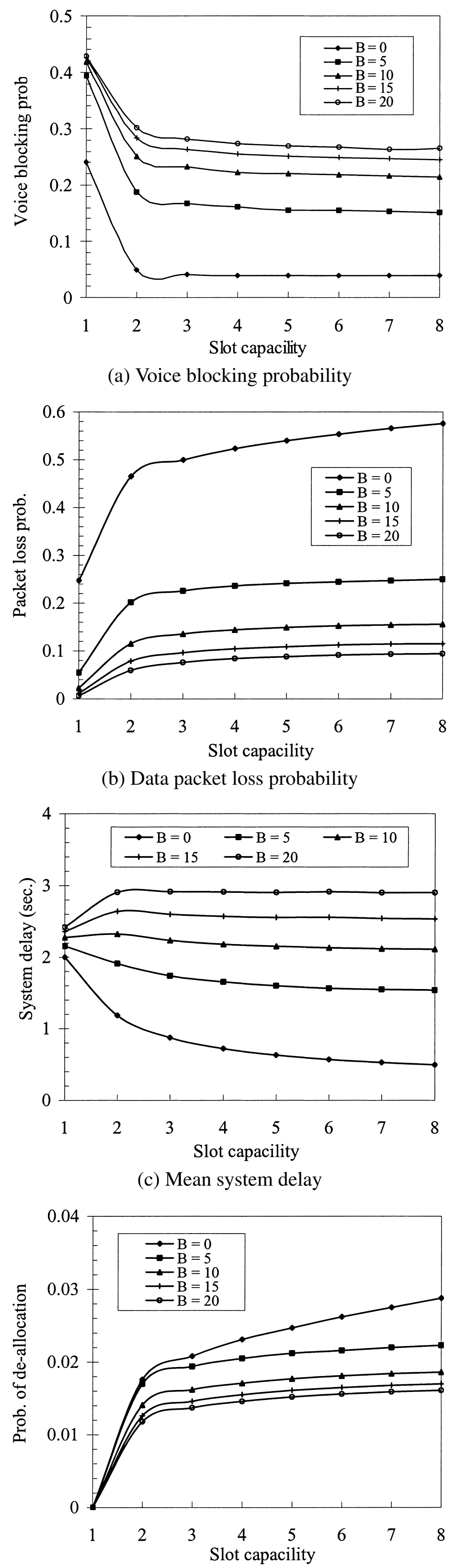
Fig. 6 The 95% delay distribution of data packets for B=10.

4.3 The Effect of Slot Capacity

In this work, the slot capacity means that the maximum number of channels that can be allocated to a data user to transfer its packets. In GPRS, at most 8 channels can be dedicated to a data user, therefore, the maximum slot capacity is 8. For slot capacity greater than 2, when there are no channels available upon voice arrivals, the data users that are allocated with most channels will be de-allocated first. Figure 7 shows the system performance for different slot capacity and buffer sizes at a data packet arrival rate of 6 packets/sec.

As can be seen in Fig. 7(a), the voice blocking probability decreases with increasing slot capacity. It is worth noting that the voice blocking probability can be significantly reduced when the slot capacity increases from one to two. As the slot capacity further increases, the voice blocking probability is almost independent of the number of slot capacity. The reason is that the increment of de-allocation probability is large when the slot capacity increases from one to two, and then the increment becomes less as the slot capacity further increases as can be seen in Fig. 7(d). The impact of multi-slot allocation on voice blocking probability in this work is quite different from that in the work of [9] which says that if many channels are allocated to a packet transmission, voice blocking probability will increase. The main reason is that the channel de-allocation was not considered in their work.

On the contrary, the packet loss probability increases



(a) Voice blocking probability

(b) Data packet loss probability

(c) Mean system delay

(d) Probability of de-allocation

Fig. 7 System performance for different slot capacity.

with increasing slot capacity as shown in Fig. 7(b). This is because data users allocated with more channels will use up channels more quickly, thus the forthcoming data packets will have less chance to obtain services and be dropped when buffer overflows.

It can be seen in Fig. 7(c) that the system delay decreases with increased slot capacity when buffer size is small. However, for large buffer size, e.g., $B > 15$, the system delay increases when the slot capacity increases from one to two. This is because the voice blocking probability decreases significantly when the slot capacity increases from one to two as can be seen in Fig. 7(a). Therefore, more voice calls are admitted into the system, causing more queued packets waiting for service completion. Since the mean service time of voice calls is larger than that of data packets, the system delay increases.

5. Conclusions

In this paper, we developed a simple and effective scheme to investigate the performance of dynamic resource allocation with channel de-allocation and buffering in cellular networks with integrated voice/data services. Using the analytical model, the system performance in terms of voice blocking probability, packet loss probability, mean system delay and delay distribution of data packets can be efficiently obtained. The results are validated by simulation experiments.

The results show that while buffering data traffic can reduce the packet loss probability, it has negative impact on voice performance even if channel de-allocation is exploited. With channel de-allocation, the voice blocking probability can be significantly reduced when the slot capacity increases from one to two. As the slot capacity increases, the voice blocking probability is almost independent of the number of slot capacity. On the contrary, the packet loss probability increases as the slot capacity increases. The operator should thus think deliberately if he/she would provide quick data transmission service at the expense of high packet loss probability and little improvement on voice blocking. To our opinion, the slot capacity ranging from 2 to 4 will be a good choice.

In addition to the mean value analysis, in this work the delay distribution and the 95% system delay of data packets are also provided. For real-time data applications, it is crucial that the delay of packet transfer is below a certain threshold. Therefore, the delay distribution gives an important information in designing the network to provide data services with delay-bounds.

Although we focus on the performance of GPRS networks, the analysis developed in this work can also be applied to the next generation cellular communication systems, such as TD-CDMA, which use codes as the radio resources.

Acknowledgement

This work was supported in part by the National Science of

Council under Contract No. NSC 92-2213-E-011-013.

References

[1] L. Chen, U. Yoshida, H. Murata, and S. Hirose, "Dynamic timeslot allocation algorithms suitable for asymmetric traffic in multimedia TDMA/TDD cellular radio," Proc. VTC 1998 IEEE, vol.2, pp.1424–1428, Ottawa, Canada, May 1998.

[2] C. Mihailescu, X. Lagrange, and P. Godlewski, "Performance evaluation of a dynamic resource allocation algorithm for UMTS-TDD systems," Proc. VTC 2000-Spring IEEE, vol.3, pp.2339–2342, Tokyo, Japan, May 2000.

[3] M. Cheung and J.W. Mark, "Resource allocation in wireless networks based on joint packet/call levels QoS constraints," Proc. GLOBECOM'00 IEEE, vol.1, pp.271–275, SanFrancisco, CA, USA, Nov. 2000.

[4] J.B. Kim, M.L. Honig, and S. Jordan, "Dynamic resource allocation for integrated voice and data traffic in DS-CDMA," Proc. VTC 2001 IEEE, vol.1, pp.42–46, Atlantic City, USA, Oct. 2001.

[5] "Digital cellular telecommunications system (Phase 2+); General Packet Radio Service (GPRS); Service description: Stage 1 (GSM 02.60 version 7.2.0 release 1999)," ETSI/TC, Tech. Rep. Rec. GSM 02.60, 1999.

[6] "Selection procedure for the choice of radio transmission technologies of the Universal Mobile Telecommunication System UMTS (UMTS 30.03)," ETSL TR 101 112, April 1998.

[7] "TS 25.22x - Physical Layer, General Description (TDD). Technical report," 3rd Generation Partnership Project, 3GPP, Oct. 1999.

[8] S. Ni and S.G. Haggman, "GPRS performance estimation in GSM circuit switched services and GPRS shared resource systems," Proc. IEEE WCNC'99, vol.3, pp.1417–1421, New Orleans, USA, Sept. 1999.

[9] P. Lin and Y.B. Lin, "Channel allocation for GPRS," IEEE Trans. Veh. Technol., vol.50, no.2, pp.375–387, March 2001.

[10] H.H. Liu, J.L.C. Wu, and W.C. Hsieh, "Delay analysis of integrated voice and data service for GPRS," IEEE Commun. Lett., vol.6, no.8, pp.319–321, Aug. 2002.

[11] W.Y. Chen, J.L.C. Wu, and L.L. Lu, "Performance comparisons of dynamic resource allocation with/without channel de-allocation in GSM/GPRS networks," IEEE Commun. Lett., vol.7, no.1, pp.10–12, Jan. 2003.

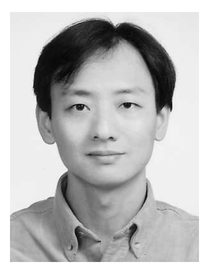
[12] M. Ermel, K. Begain, T. Muller, J. Schuler, and M. Schweigel, "Analytical comparison of different GPRS introduction strategies," Proc. 3rd ACM Int. Workshop on Modeling Analysis and Simulation of Wireless and Mobile Systems, pp.3–10, Boston, USA, Aug. 2000.

[13] L.R. Lipsky, Queueing Theory: A linear algebraic approach, Macmillan Publishing Company, 1992.

[14] S. Ghani and M. Schwartz, "A decomposition approximation for the analysis of voice/data integration," IEEE Trans. Commun., vol.42, no.7, pp.2441–2452, July 1994.

Wei-Yeh Chen received the B.S. degree from National Cheng Kung University in 1986, and M.S. degree from National Chiao Tung University in 1988. He is working toward the Ph.D. degree in the Department of Information Management at the National Taiwan University of Science and Technology. His current research is focused on performance analysis for mobile communication systems.

Jean-Lien C. Wu received the B.S. in electrical engineering from National Taiwan University, Taipei, Taiwan, in 1970, and the Ph.D. degree in electrical engineering from Cornell University, Ithaca, NY, in 1976. Her current research interests are in wireless networks, broadband Internet, and mobile multimedia networks. From 1976 to 1977, she was a lecture in the Dept. of EE in Cornell Univ. From 1978 to 1984, she was Member of the Technical Staff with the Business Switching Systems Dept., Bell Laboratories, Holmdel, NJ. Since 1984, she has been a professor in the Electronic Engineering Department, National Taiwan Univ. of Science and Technology (NTUST), Taipei, Taiwan. She was Head of the department from 1987 to 1990 and Dean of the College of Electrical and Computer Engineering from 1998 to 2001. She is currently Dean of the Academic Affair of the university. Dr. Wu was a member of the editorial advisory board of Elsevier Computer Communications from 1994 to 1999. She was an editor of Journal of Information Science and Engineering (JISE) from 1996 to 2002. She was the Editor-in-Chief of Journal of Chinese Institute of Electrical Engineering (JCIEE) from 2000 to 2003. Since Dec. 2002, Dr. Wu has been the coordinator of the Communication Engineering Program of the Engineering and Applied Sciences Dept. of the National Science Council (NSC), Taiwan. She is also a member of the advisory board of the Science and Technology Advisory Office of the Ministry of Education, Taiwan. Dr. Wu is a member of the IEEE communication society and computer society. She is also member of the Chinese Institute of Engineers society and the Chinese Institute of Electrical Engineering society.



Hung-Huan Liu received the B.S., M.S. and Ph.D. degree in the Department of Electronic Engineering at National Taiwan University of Science and Technology, Taipei, Taiwan in 1994, 1997 and 2003, respectively. He was with Department of Computer Science and Information Engineering at National Penghu Institute of Technology from 1999 to 2004, as a lecturer and later as an assistant professor. Since 2004 spring, he has been at Chung Yuan Christian University, Taiwan, where he is an assistant professor in Department of Electronic Engineering. His research interests include resource allocation, quality of service, wireless LANs, and ad hoc networks.