Wireless Networks 9, 393–402, 2003

On Handoff Performance for an Integrated Voice/Data Cellular System

BIN LI
Department of Electrical and Electronic Engineering, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong

LIZHONG LI and BO LI *
Department of Computer Science, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong

XI-REN CAO
Department of Electrical and Electronic Engineering, Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong

Abstract. One of the key challenges in the design of bandwidth allocation policies for a multi-services mobile cellular network is to guarantee the potentially different Quality of Service (QoS) requirement from diverse applications, while at the same to ensure that the scarce bandwidth be utilized efficiently. *Complete Sharing* (CS) and *Dynamic Partition* (DP) schemes have been shown as viable techniques for managing the bandwidth. However, there has been no study that compares their respective performance, which is the focus of this paper. Specifically, in this paper, through both analysis and simulation, we demonstrate that both schemes can achieve comparable performance by proper manipulation of control parameters. The tradeoff is that DP scheme can more easily achieve the target QoS requirement, at the expense of some over-provisioning, thus can potentially lead to less channel efficiency when comparing to a CS based scheme.

Keywords: bandwidth allocation, guarded channel policy, handoff, mobile cellular networks

1. Introduction

We have recently witnessed phenomenal growth in the development and deployment of wireless services, evident from the proliferation of cellular data services and the emerging wireless multimedia applications. This opens up a new research avenue and calls for the re-examination of some of the fundamental issues in wireless cellular networks. Various issues such as QoS guarantee have to be carefully examined. One of the key elements in providing QoS guarantee is an effective *bandwidth allocation policy*, which not only has to ensure that the system guarantee potentially different QoS requirements and provide the necessary service differentiation from diverse applications, but at the same time has to fully utilize the scarce wireless bandwidth available in the wireless cellular networks. Moving from a voice-centric macro-cell wireless network to a multi-services micro-cell or/and pico-cell wireless network brings several new challenges: (1) the characteristics of data and multimedia applications are typically different from that of voice, e.g., data traffic usually consumes more bandwidths than that of voice, but can be adaptive, thus has less stringent service requirement; (2) the limited wireless resources have to be utilized effectively, and fairly allocated to different types of traffic with potentially different service requirements and service differentiation; (3) smaller cell potentially results in more handoffs, making it more difficult to provide the necessary QoS guarantee.

Bandwidth allocation has been extensively studied in single-service wireless cellular networks. The *Guarded Channel* (GC) schemes have been shown to be effective for providing the necessary QoS guarantee in terms of both call termination and call blocking probabilities [5,10,14,16]. One of the challenges in moving to a multi-service system is that the limited bandwidth has to be shared among multiple traffics. In [2] a *Complete Sharing* (CS) and *Complete Partition* (CP) schemes were investigated for two types of traffic, namely *narrow-band* and *wideband*. It assumed that wideband traffic does not have handoff. Huang et al. proposed a *movable boundary* allocation scheme for voice and data traffic [6], in which bandwidth is divided into two sub-pools by two thresholds that can be dynamically adjusted. This facilitates the bandwidth provisioning for different QoS requirements and is adaptive to the changing traffic. The limitations are that both voice and data require the same bandwidth, and there is no service differentiation between voice new calls and handoff.

To the best of our knowledge, there has been no comparative study that investigates the performance of different bandwidth allocation schemes for multi-service wireless cellular networks. This is the focus of this paper. Specifically, we are interested in obtaining the quantitative performance measures for different bandwidth allocation schemes. Although our analysis in this paper focuses on two traffic types, but the derivation and conclusions also provide insight for arbitrary number of traffic types.

A variety of bandwidth allocation schemes have been proposed to support multiple traffics, which can be classified as *Complete Partitioning* (CP), *Complete Sharing* (CS) or hy-

---

* Corresponding author.
E-mail: bli@cs.ust.hk

brid schemes, depending on how the bandwidth are allocated among diverse traffic. In this paper, we first extend the *movable boundary* scheme [6] to consider different bandwidth consumption by the two traffic types, referred as *Dynamic Partition* (DP). We compare this with a CS scheme we proposed in [15,17], called *Dual Threshold Bandwidth Reservation* (DTBR). We compare their respective performance using both analysis and simulation. The study reveals a number of interesting observations and provides insight in searching more suitable bandwidth allocation schemes.

The rest of the paper is organized as follow. We describe the DP and DTBR schemes in section 2. We present the analytical models in section 3, following by the numerical studies in section 4. We conclude the paper in section 5 with discussions on possible avenue for further study.

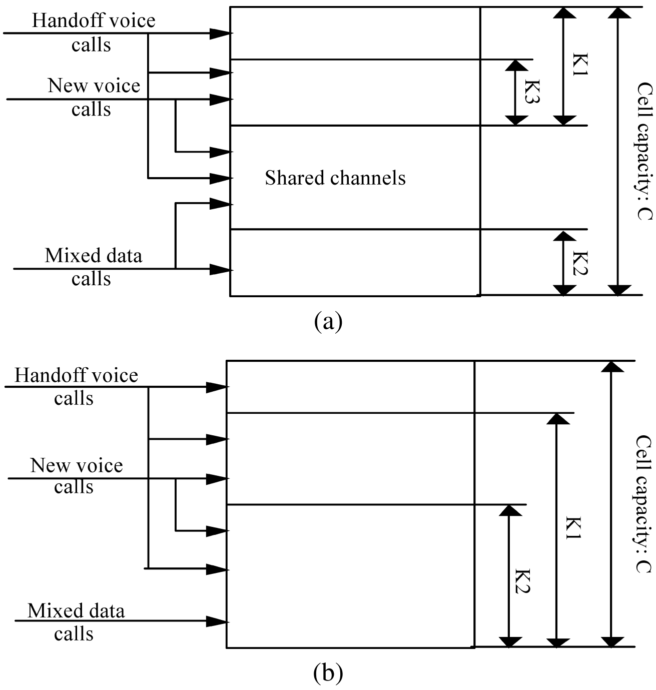
2. Bandwidth allocation strategies

We assume that each cell have a total of C channels. The bandwidth requirements for voice calls and data calls are fixed as 1 and B units, respectively.

2.1. Dynamic partition (DP) scheme

In [6], Haung et al. proposed a voice and data integrated system with finite buffers, in which bandwidths assigned to voice and data are separated (i.e., complete partition). The unique feature of this scheme is that the boundary for the partition is "*movable*", thus can effectively deal with the traffic changes in the system. In this scheme, however, handoff calls were not differentiated form new calls, thus the strict requirement of handoff dropping probability cannot be met. Furthermore it was assumed that the data requires *one unit* of bandwidth. We extend the *movable boundary* scheme to differentiate handoff calls and further take into account the different bandwidth requirement. In the new scheme, called *Dynamic Partition* (DP) scheme (see figure 1(a)), among C channels, K_1 channels (voice-only-area) are reserved exclusively for new/handoff voice calls and K_2 channels (data-only-area) are reserved exclusively for new/handoff data calls. The other $(C - K_1 - K_2)$ channels (shared-area) are fairly shared by both voice and data calls. In order to maintain a low handoff dropping probability for voice calls, we further restrict that new voice calls can only use the K_3 out of K_1 voice channels (i.e., Guarded Channel policy). But handoff voice calls can use all the K_1 channels. The admission control of DP scheme is described as below.

Voice call will be firstly arranged into voice-only-area. If there is no available channel there, it should then be directed into shared-area. On the other hand, data call will be firstly arranged into data-only-area. If the available channels there cannot accommodate it, it should then be directed into shared-area. When a handoff voice call arrive, if there is no channel available in both the voice only area and the shared area, it will be dropped. When a new voice call arrive, if the number of channel occupancy exceeds the threshold K_3 in the voice only area and there is no idle channel in the shared area, it will be rejected. When a data call (new or handoff) arrive, if the number of idle channels in the data only area or in the shared area is less than B, it will be blocked.



(a)
(b)

Figure 1. (a) DP control schema. (b) DTBR control schema.

2.2. Dual-threshold bandwidth reservation (DTBR)

The rationale for adopting a CS approach in the proposed *Dual-Threshold Bandwidth Reservation* (DTBR) scheme [15,17] (see figure 1(b)) is that it can achieve the maximum channel efficiency, while at the same time still being able to provide service guarantee and differentiation from diverse traffic. In the DTBR scheme, C channels of each cell are divided into *three* regions by two thresholds K_1 and K_2. When the number of channels occupied is less than the threshold K_2, then both data and voice traffic can be admitted into the system; when the number of channels occupied is over the threshold K_2, no data traffic is allowed; when the number of channels occupied is more than the threshold K_1, then only handoff voice calls can be allowed. The handoff voice call will be dropped only if there is no channel available. Under this basic control model, the handoff voice gets highest priority, while data receives lowest service. The reason is that the data traffic can tolerate certain degree of service degradation, while voice cannot. There can be a number of variations using this basic model, for example, whether or not to use a queue to further buffer the handoff request when the requested channels are not available [15]. This has been shown to reduce the handoff dropping probability [5,10]. In this study, for simplicity and for fair comparison with the DP scheme, we only consider the DTBR scheme without queuing.

3. Performance analysis

We consider a homogeneous wireless network where all cells have the same number of channels and experience the same new and handoff call arrival rates. In each cell, the arrivals of new voice calls, new data calls, handoff voice calls and handoff data calls are Poisson distributed with arrival rate λ_{vn}, λ_{dn}, λ_{vh} and λ_{dh}, respectively. Thus the total voice call arrival rate and data call arrival rate are $\lambda_{\mathrm{v}} = \lambda_{\mathrm{vn}} + \lambda_{\mathrm{vh}}$ and $\lambda_{\mathrm{d}} = \lambda_{\mathrm{dn}} + \lambda_{\mathrm{dh}}$, respectively. Since data can usually tolerate some degree of service degradation, new data calls and handoff data calls are not distinguished. Call duration times or call holding times of voice and data are exponentially distributed with the average call duration time $1/\mu_{\mathrm{vr}}$ and $1/\mu_{\mathrm{dr}}$. In addition, the handoff voice and the handoff data are exponentially distributed with mean μ_{vh} and μ_{dh}, respectively. This set of assumptions have been found reasonable as long as the number of mobiles is much larger than the number of channels in a cell, and have been widely used in literature [2,5,6,10,14–17]. The exponential call holding time has also been shown to be valid for a wide range of systems [3,4,12]. In this paper, for simplifying notations, we assume K_2 is a multiple of B.

3.1. Dynamic partition scheme

Scheme DP can be modeled as a three-dimensional Markov chain. It is ergodic [8] if $\rho_{\mathrm{d}} < (C - K_1)/B$. Let P_{ijk} be the steady probability that there are i new voice calls, j handoff voice calls and k data calls in the system. Let $\lceil x \rceil$ denote the greatest integer less than or equal to x. The steady-state balance equations of DP scheme are shown as below.

Case 1. If $i = j = k = 0$, then

$$(\lambda_{\mathrm{v}} + \lambda_{\mathrm{d}})P_{000} = \mu_{\mathrm{v}} P_{100} + \mu_{\mathrm{v}} P_{010} + B\mu_{\mathrm{d}} P_{001}. \tag{1}$$

Case 2. If $0 \leqslant i + j \leqslant K_1$ and $0 \leqslant k < \lceil (C - K_1)/B \rceil$, or $K_1 < i + j < C - K_2$ and $0 < k < \lceil (C - i - j)/B \rceil$, all voice calls (new and handoff) and data calls will be accepted, thus we have

$$\begin{aligned}(\lambda_{\mathrm{vn}} &+ \lambda_{\mathrm{vh}} + \lambda_{\mathrm{d}} + i\mu_{\mathrm{v}} + j\mu_{\mathrm{v}} + kB\mu_{\mathrm{d}})P_{ijk} \\ &= (i+1)\mu_{\mathrm{v}} P_{i+1,jk} + (j+1)\mu_{\mathrm{v}} P_{i,j+1,k} \\ &\quad + (k+1)B\mu_{\mathrm{d}} P_{ij,k+1} + \lambda_{\mathrm{vn}} P_{i-1,jk} + \lambda_{\mathrm{vh}} P_{i,j-1,k} \\ &\quad + \lambda_{\mathrm{d}} P_{ij,k-1}.\end{aligned} \tag{2}$$

Case 3. If $0 < i + j < K_3$ and $k = \lceil (C - K_1)/B \rceil$, voice calls (new and handoff) will be accepted, but data calls will be rejected, thus we have

$$\begin{aligned}(\lambda_{\mathrm{vn}} &+ \lambda_{\mathrm{vh}} + i\mu_{\mathrm{v}} + j\mu_{\mathrm{v}} + kB\mu_{\mathrm{d}})P_{ijk} \\ &= (i+1)\mu_{\mathrm{v}} P_{i+1,jk} + (j+1)\mu_{\mathrm{v}} P_{i,j+1,k} + \lambda_{\mathrm{vn}} P_{i-1,jk} \\ &\quad + \lambda_{\mathrm{vh}} P_{i,j-1,k} + \lambda_{\mathrm{d}} P_{ij,k-1}.\end{aligned} \tag{3}$$

Case 4. If $K_3 \leqslant i + j < K_1$ and $k = \lceil (C - K_1)/B \rceil$, only handoff voice calls will be accepted, new voice calls and data calls will be rejected, thus we have

$$\begin{aligned}(\lambda_{\mathrm{vh}} &+ i\mu_{\mathrm{v}} + j\mu_{\mathrm{v}} + kB\mu_{\mathrm{d}})P_{ijk} \\ &= (j+1)\mu_{\mathrm{v}} P_{i,j+1,k} + \lambda_{\mathrm{vn}} P_{i-1,jk} + \lambda_{\mathrm{vh}} P_{i,j-1,k} \\ &\quad + \lambda_{\mathrm{d}} P_{ij,k-1}.\end{aligned} \tag{4}$$

Case 5. If $i + j = K_1$ and $k = \lceil (C - K_1)/B \rceil$, or $K_1 < i + j < C - K_2$ and $k = \lceil (C - i - j)/B \rceil$, or $i + j = C - K_2$ and $k = K_2/B$, all new/handoff voice calls and data calls will be rejected, thus we have

$$\begin{aligned}(i\mu_{\mathrm{v}} &+ j\mu_{\mathrm{v}} + kB\mu_{\mathrm{d}})P_{ijk} \\ &= \lambda_{\mathrm{vn}} P_{i-1,jk} + \lambda_{\mathrm{vh}} P_{i,j-1,k} + \lambda_{\mathrm{d}} P_{ij,k-1}.\end{aligned} \tag{5}$$

Case 6. If $i + j = C - K_2$ and $k < K_2/B$, data calls will be accepted, but new and handoff voice calls will be rejected, thus we have

$$\begin{aligned}(\lambda_{\mathrm{d}} &+ i\mu_{\mathrm{v}} + j\mu_{\mathrm{v}} + kB\mu_{\mathrm{d}})P_{ijk} \\ &= (k+1)B\mu_{\mathrm{d}} P_{ij,k+1} + \lambda_{\mathrm{vn}} P_{i-1,jk} + \lambda_{\mathrm{vh}} P_{i,j-1,k} \\ &\quad + \lambda_{\mathrm{d}} P_{ij,k-1}.\end{aligned} \tag{6}$$

Above balance equations can be solved by a recursive technique developed by Herzog et al. [7], which is based on the typical feature of Chapman–Kolmogoroff equations that there exist a subset of the state probabilities, called them boundaries, and all other states can be expressed as a linear combination of boundary states. The basic idea of this technique is to choose the boundaries first and to derive the expressions for all remaining state probabilities as functions of the boundary values, and then solve a reduced system of equations for these boundaries. After that, determine all state probabilities by means of the boundaries. This has been shown to be suitable for solving a wide class of queuing problems. Compared with traditional matrix inversion technique, computer time and/or memory requirement can be reduced significantly in this technique.

Detailed calculation is given below. Choose state probabilities P_{ij0} ($i = 0, 1, \ldots, C - K_1 - K_2 + K_3$ and $j = 0, 1, \ldots, C - K_2$) as the boundaries. We introduce following substitution as in [7]

$$P_{ijk} = \sum_{\beta=0}^{C-K_1-K_2+K_3} \sum_{\gamma=0}^{C-K_2} C_{ijk}^{\beta\gamma} P_{\beta\gamma 0} \tag{7}$$

with

$$C_{ij0}^{\beta\gamma} = \begin{cases} 1 & \text{if } i = \beta \text{ and } j = \gamma, \\ 0 & \text{if } i \neq \beta \text{ or } j \neq \gamma. \end{cases} \tag{8}$$

The coefficients $C_{ijk}^{\beta\gamma}$ can be solved recursively. First, rewrite the balance equation in a general format:

$$\begin{aligned}p_{ijk} = \frac{1}{B_{ijk}}\bigl(&A_{ijk}^{1} P_{i-1,jk} + A_{ijk}^{2} P_{i+1,jk} + A_{ijk}^{3} P_{i,j-1,k} \\ &+ A_{ijk}^{4} P_{i,j+1,k} + A_{ijk}^{5} P_{ij,k-1} + A_{ijk}^{6} P_{ij,k+1}\bigr)\end{aligned} \tag{9}$$

where B_{ijk}, A^{γ}_{ijk} $(\gamma = 1, \ldots, 6)$ can be obtained from equations (1)–(6) accordingly. After some manipulation, P_{ijk} can be expressed as linear combinations of P_{rst} where $t < k$, thus we have

$$\begin{aligned} p_{ijk} = {} & \frac{1}{A^{6}_{ij,k-1}} \big(B_{ij,k-1} P_{ij,k-1} - A^{1}_{ij,k-1} P_{i-1,j,k-1} \\ & \qquad - A^{2}_{ij,k-1} P_{i+1,j,k-1} \big) \\ & - \frac{1}{A^{6}_{ij,k-1}} \big(A^{3}_{ij,k-1} P_{i,j-1,k-1} + A^{4}_{ij,k-1} P_{i,j+1,k-1} \\ & \qquad + A^{5}_{ij,k-1} P_{ij,k-2} \big). \end{aligned} \tag{10}$$

Substituting all state probabilities in equation (10) according to equation (7), we have

$$\begin{aligned} C^{\beta\gamma}_{ijk} = {} & \frac{1}{A^{6}_{ij,k-1}} \big(B_{ij,k-1} C^{\beta\gamma}_{ij,k-1} - A^{1}_{ij,k-1} C^{\beta\gamma}_{i-1,j,k-1} \\ & \qquad - A^{2}_{ij,k-1} C^{\beta\gamma}_{i+1,j,k-1} \big) \\ & - \frac{1}{A^{6}_{ij,k-1}} \big(A^{3}_{ij,k-1} C^{\beta\gamma}_{i,j-1,k-1} + A^{4}_{ij,k-1} C^{\beta\gamma}_{i,j+1,k-1} \\ & \qquad + A^{5}_{ij,k-1} C^{\beta\gamma}_{ij,k-2} \big). \end{aligned} \tag{11}$$

Then, for every fixed pair of (β, γ), where $\beta = 0, 1, \ldots, C - K_1 - K_2 + K_3$ and $\gamma = 0, 1, \ldots, C - K_2$, $C^{\beta\gamma}_{ijk}$ can be determined by solving linear equations (10) recursively by assuming $P_{\beta\gamma 0} = 1$ and $P_{ij0} = 0$ for $i \neq \beta$ or $j \neq \gamma$.

After obtaining all coefficients $C^{\beta\gamma}_{ijk}$, the boundaries state probabilities can be determined by solving reduced system of $(C - K_1 - K_2 + K_3) \times (C - K_2)$ independent equations along with the normalizing condition,

$$\sum_i \sum_j \sum_k P_{ijk} = 1.$$

Having solved the boundaries, all steady state probabilities P_{ijk} can be determined from equation (7). Thus, the voice call blocking probability P_{vb}, the handoff voice call dropping probability P_{vd}, the data call blocking probability P_{db} and the total channel utilization η_{DP} can be derived as below:

$$P_{\mathrm{vb}} = \sum_{\substack{K_3 \leqslant i+j \leqslant K_1, \\ k = \lceil (C-K_1)/B \rceil}} P_{ijk} + \sum_{\substack{K_1 < i+j < C-K_2 \\ k = \lceil (C-i-j)/B \rceil}} P_{ijk} + \sum_{\substack{i+j = C-K_2, \\ k \leqslant K_2/B}} P_{ijk}, \tag{12}$$

$$P_{\mathrm{vd}} = \sum_{\substack{i+j = K_1, \\ k = \lceil (C-K_1)/B \rceil}} P_{ijk} + \sum_{\substack{K_1 < i+j < C-K_2 \\ k = \lceil (C-i-j)/B \rceil}} P_{ijk} + \sum_{\substack{i+j = C-K_2, \\ k \leqslant K_2/B}} P_{ijk}, \tag{13}$$

$$P_{\mathrm{db}} = \sum_{\substack{0 \leqslant i+j \leqslant K_1, \\ k = \lceil (C-K_1)/B \rceil}} P_{ijk} + \sum_{\substack{K_1 < i \leqslant C-K_2, \\ k = \lceil (C-i-j)/B \rceil}} P_{ijk}, \tag{14}$$

$$\eta_{\mathrm{DP}} = \frac{1}{C} \sum_{i=0}^{C-K_1-K_2+K_3} \sum_{j=0}^{C-K_2} \sum_{k=0}^{\lceil (C-K_1)/B \rceil} (i + j + kB) P_{ijk}. \tag{15}$$

3.2. Complete sharing scheme DTBR

Scheme DTBR can be modeled as a two-dimensional Markov chain. It is ergodic [8] if $\rho_{\mathrm{d}} < K_2/B$. Let P_{ij} be the steady probability that there are i voice (new and handoff) calls and j data calls in the system. The steady-state balance equations of DTBR scheme are shown as below.

If $j \neq K_2/B$:

$$P_{ij} = \begin{cases} 0 \quad \text{if } i + jB > C, \\ \dfrac{\lambda_{\mathrm{vh}} P_{i-1,j}}{i\mu_{\mathrm{v}} + jB\mu_{\mathrm{d}}} \quad \text{if } i + jB = C, \\ \dfrac{(i+1)\mu_{\mathrm{v}} P_{i+1,j} + \lambda_{\mathrm{vh}} P_{i-1,j} + (j+1)B\mu_{\mathrm{d}} P_{i,j+1}}{\lambda_{\mathrm{vh}} + i\mu_{\mathrm{v}} + jB\mu_{\mathrm{d}}} \\ \quad \text{if } K_1 < i + jB < C, \\ \dfrac{(i+1)\mu_{\mathrm{v}} P_{i+1,j} + \lambda_{\mathrm{v}} P_{i-1,j} + (j+1)B\mu_{\mathrm{d}} P_{i,j+1}}{\lambda_{\mathrm{vh}} + i\mu_{\mathrm{v}} + jB\mu_{\mathrm{d}}} \\ \quad \text{if } i + jB = K_1, \\ \dfrac{(i+1)\mu_{\mathrm{v}} P_{i+1,j} + \lambda_{\mathrm{v}} P_{i-1,j} + (j+1)B\mu_{\mathrm{d}} P_{i,j+1}}{\lambda_{\mathrm{v}} + i\mu_{\mathrm{v}} + jB\mu_{\mathrm{d}}} \\ \quad \text{if } K_2 < i + jB < K_1, \\ \big((i+1)\mu_{\mathrm{v}} P_{i+1,j} + \lambda_{\mathrm{v}} P_{i-1,j} + (j+1)B\mu_{\mathrm{d}} P_{i,j+1} \\ \quad + \lambda_{\mathrm{d}} P_{i,j-1}\big)(\lambda_{\mathrm{v}} + i\mu_{\mathrm{v}} + jB\mu_{\mathrm{d}})^{-1} \\ \quad \text{if } i + jB = K_2, \\ \big((i+1)\mu_{\mathrm{v}} P_{i+1,j} + \lambda_{\mathrm{v}} P_{i-1,j} + (j+1)B\mu_{\mathrm{d}} P_{i,j+1} \\ \quad + \lambda_{\mathrm{d}} P_{i,j-1}\big)(\lambda_{\mathrm{v}} + i\mu_{\mathrm{v}} + jB\mu_{\mathrm{d}} + \lambda_{\mathrm{d}})^{-1} \\ \quad \text{if } 0 < i + jB < K_2, \\ \dfrac{(i+1)\mu_{\mathrm{v}} P_{1,0} + B\mu_{\mathrm{d}} P_{i,1}}{\lambda_{\mathrm{v}} + \lambda_{\mathrm{d}}} \quad \text{if } i = 0 \text{ and } j = 0, \\ 0 \quad \text{if } i < 0 \text{ or } j < 0. \end{cases} \tag{16}$$

If $j = K_2/B$:

$$P_{ij} = \begin{cases} 0 & \text{if } i > C - K_2, \\ \dfrac{\lambda_{\mathrm{vh}} P_{i-1,j}}{i\mu_{\mathrm{v}} + jB\mu_{\mathrm{d}}} & \text{if } i = C - K_2, \\ \dfrac{(i+1)\mu_{\mathrm{v}} P_{i+1,j} + \lambda_{\mathrm{vh}} P_{i-1,j}}{\lambda_{\mathrm{vh}} + i\mu_{\mathrm{v}} + jB\mu_{\mathrm{d}}} & \\ \quad \text{if } K_1 - K_2 < i < C - K_2, & \\ \dfrac{(i+1)\mu_{\mathrm{v}} P_{i+1,j} + \lambda_{\mathrm{v}} P_{i-1,j}}{\lambda_{\mathrm{vh}} + i\mu_{\mathrm{v}} + jB\mu_{\mathrm{d}}} & \text{if } i = K_1 - K_2, \\ \dfrac{(i+1)\mu_{\mathrm{v}} P_{i+1,j} + \lambda_{\mathrm{v}} P_{i-1,j}}{\lambda_{\mathrm{v}} + i\mu_{\mathrm{v}} + jB\mu_{\mathrm{d}}} & \text{if } 0 < i < K_1 - K_2, \\ \dfrac{(i+1)\mu_{\mathrm{v}} P_{i+1,j} + \lambda_{\mathrm{d}} P_{i,j-1}}{\lambda_{\mathrm{v}} + jB\mu_{\mathrm{d}}} & \text{if } i = 0, \\ 0 & \text{if } i < 0. \end{cases} \tag{17}$$

Similarly, these balance equations can be solved recursively as in section 3.1. After obtaining all the steady state probabilities P_{ij}, the voice call blocking probability P_{vb}, the handoff voice call dropping probability P_{vd}, the data call blocking probability P_{db} and the total channel utilization

η_{DTBR} can be derived as below:

$$P_{\mathrm{vb}} = \sum_{i+jB \geqslant K_1}^{i+jB=C} P_{ij}, \tag{18}$$

$$P_{\mathrm{vd}} = \sum_{i+jB=C} P_{ij}, \tag{19}$$

$$P_{\mathrm{db}} = \sum_{i+jB \geqslant K_2}^{i+jB=C} P_{ij}, \tag{20}$$

$$\eta_{\mathrm{DTBR}} = \frac{\sum_{i=0}^{C} \sum_{j=0}^{K_2/B} (i + jB) P_{ij}}{C}. \tag{21}$$

4. Numerical results and remarks

In this section, we present numerical results and compare the above two schemes. In order to validate the accuracy of the analysis, we also develop an event-driven simulation. The system performance parameters considered in this paper are handoff voice call dropping probability, voice call blocking probability, data call blocking probability and overall channel utilization. We consider the following system configuration. The total channel number C of each cell is set to be 30 and the bandwidth requirement of data calls B is set to be 2. The data call intensity ρ_{d} is set to be 7, with the average arrival rate of $\lambda_{\mathrm{d}} = 0.007\ \mathrm{s}^{-1}$ and the average service rate of $\mu_{\mathrm{d}} = 0.001\ \mathrm{s}^{-1}$. This is reasonable for typical Internet type of traffic [1]. For voice call, the average service rate is assumed to be $\mu_{\mathrm{v}} = 0.0083\ \mathrm{s}^{-1}$, while the voice call intensity can vary from 6 to 16. To alleviate the transient effect, the simulation was run for a long duration in order to reach the steady state, and the system performance measures were obtained by averaging over the results of 10 independent rounds of simulations.

One of the key QoS measures in wireless cellular networks is the *handoff voice call dropping probability*. As dropping a call in progress generally have more negative impact from user's perception than blocking a newly requested call. Thus, in order to compare the performance of DP and DTBR scheme, we set the target of the handoff voice call dropping probability to be 10^{-3} [16], and study the system performance under a wide range of call intensity. Another important performance measure is the total channel utilization of the system. This is because in wireless cellular networks, radio resources are scarce compared with its wire-line counterparts. So, a good bandwidth allocation scheme should provide QoS guarantee while at the same time has to fully utilize the scarce wireless bandwidth available in the wireless cellular networks.

We first present the performances of DP allocation scheme with different sets of thresholds (K_1, K_2, K_3) that can guarantee target handoff voice call dropping probability (10^{-3}) at $\rho_{\mathrm{v}} = 16$. Then we choose the set of (K_1, K_2, K_3) that can achieve the highest channel utilization as the representative to compare the performance with DTBR schemes having different (K_1, K_2) values.

4.1. Performances of DP scheme

For DP allocation scheme to be ergodic under given system configuration, threshold K_1 should be less than $K_1^{\max}$, where $K_1^{\max} = C - 2\rho_{\mathrm{d}} = 16$. Figures 2–5 illustrate the performances of DP scheme when $\lambda_{\mathrm{vh}} = 0.2\lambda_{\mathrm{v}}$ (i.e., lower user mobility), while figures 6–9 plot the performances of DP scheme when $\lambda_{\mathrm{vh}} = 0.4\lambda_{\mathrm{v}}$ (i.e., higher user mobility).

Figure 2 indicates that, in order to maintain the target handoff voice call dropping probability of 10^{-3} at $\rho_{\mathrm{v}} = 16$, certain amount of channels (K_1) needs to be designated exclusively for voice calls. But due to the nature of complete partition, these voice only channels cannot be shared by data calls, thus when the voice call intensity is low, much lower handoff voice call dropping probability can be achieved.

Figure 3 presents the average channel utilization of DP scheme. It shows that, among all the sets of (K_1, K_2, K_3) parameters, those sets with $K_2 = 0$ (that is, no channels are reserved for data calls) always have higher channel utilization than other sets of (K_1, K_2, K_3) parameters. The reason is

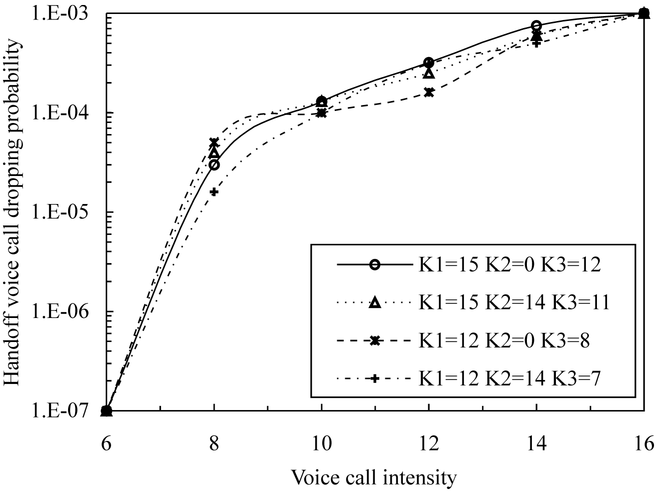


Figure 2. Handoff voice call dropping probability of DP scheme when $\lambda_{\mathrm{vh}} = 0.2\lambda_{\mathrm{v}}$.

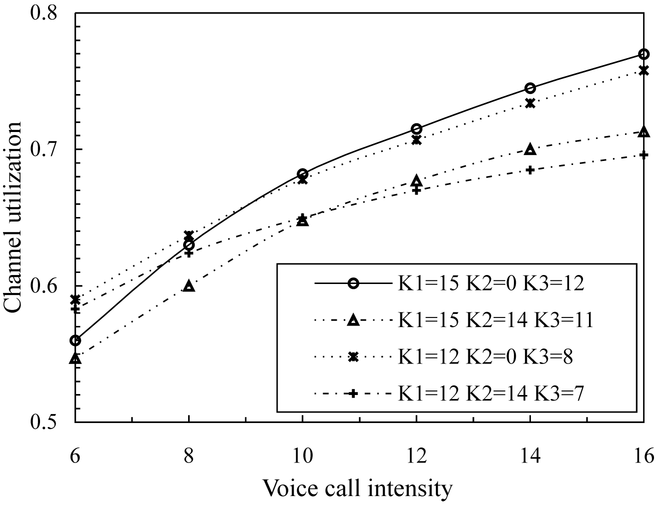


Figure 3. Channel utilization of DP scheme when $\lambda_{\mathrm{vh}} = 0.2\lambda_{\mathrm{v}}$.

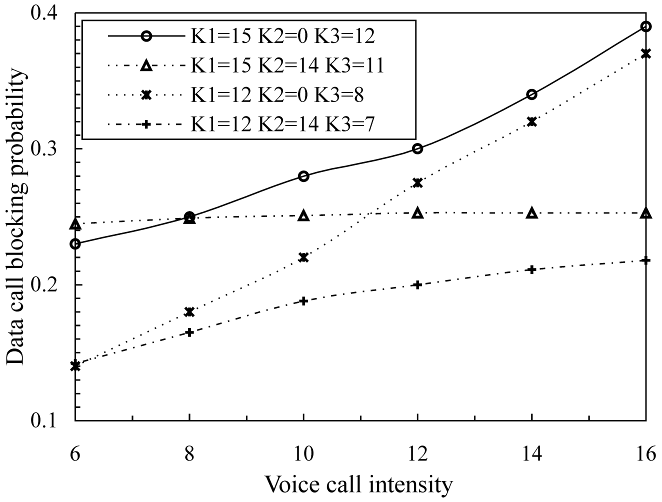


Figure 4. Data call blocking probability of DP scheme when $\lambda_{vh} = 0.2\lambda_v$.

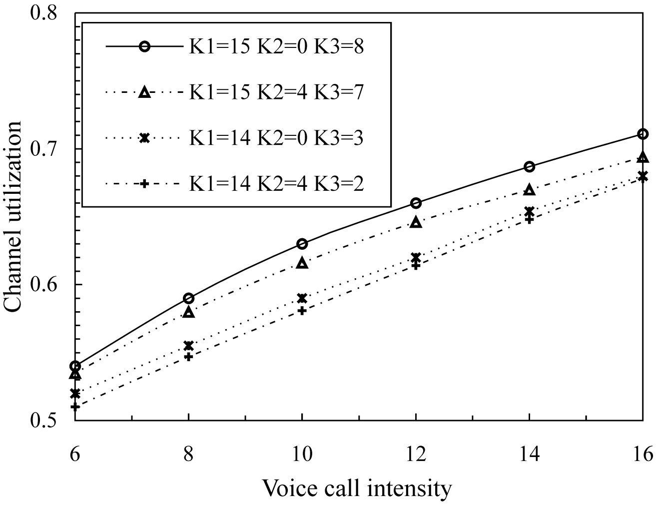


Figure 7. Channel utilization of DP scheme when $\lambda_{vh} = 0.4\lambda_v$.

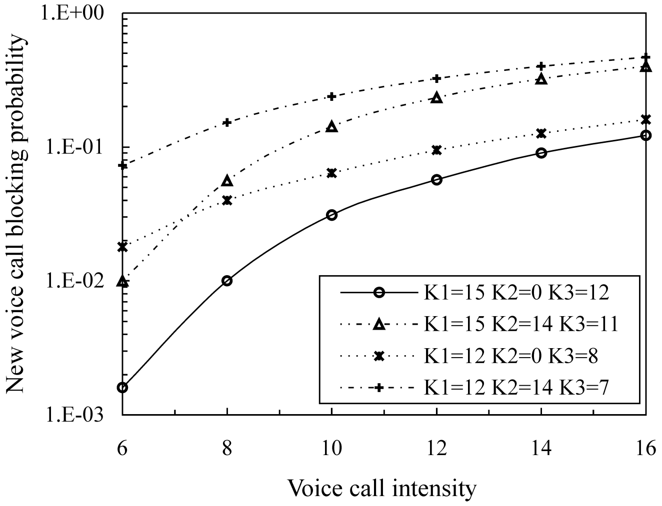


Figure 5. Voice call blocking probability of DP scheme when $\lambda_{vh} = 0.2\lambda_v$.

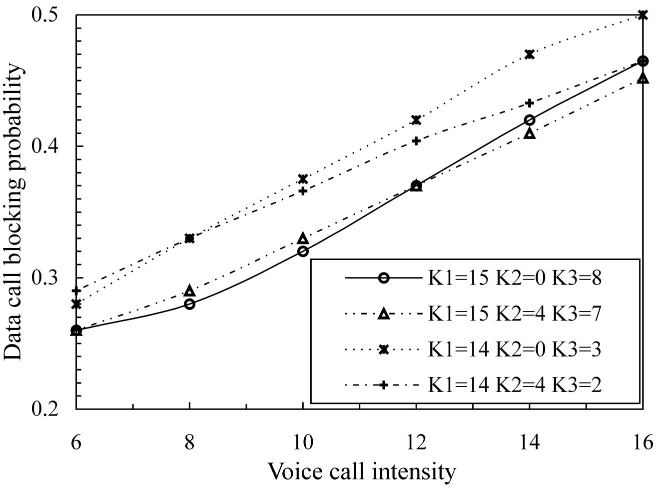


Figure 8. Data call blocking probability of DP scheme when $\lambda_{vh} = 0.4\lambda_v$.

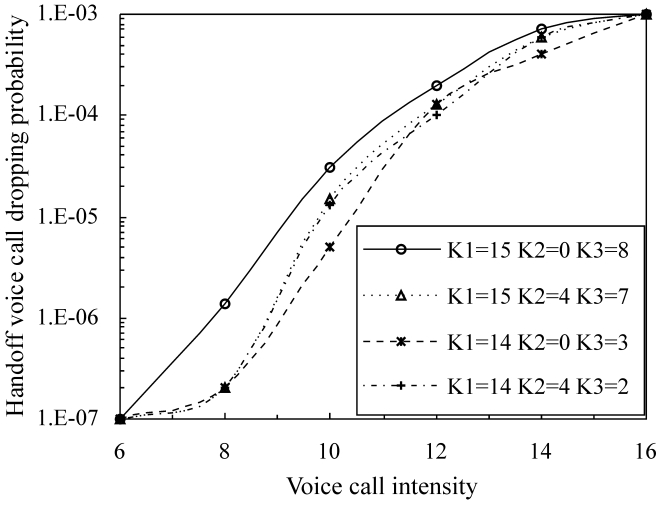


Figure 6. Handoff voice call dropping probability of DP scheme when $\lambda_{vh} = 0.4\lambda_v$.

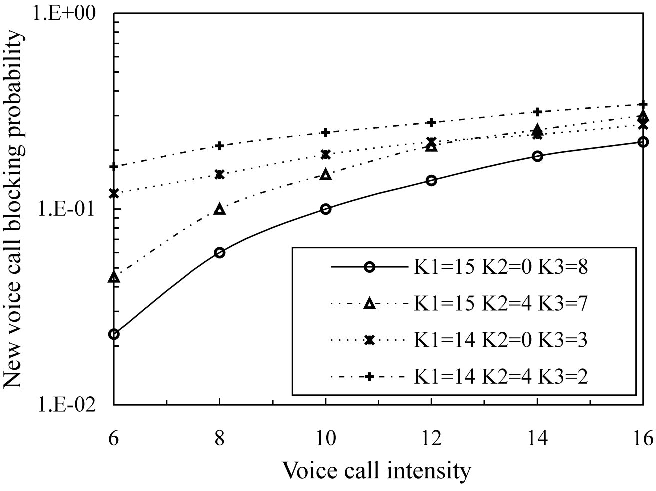


Figure 9. Voice call blocking probability of DP scheme when $\lambda_{vh} = 0.4\lambda_v$.

that, when $K_2 = 0$, except those channels needed to keep the stringent target voice QoS, all other channels can be shared by voice and data calls, thus leads to high channel utilization, especially under heavy traffic loads.

Figure 3 further shows that among the cases of $K_2 = 0$, the case with ($K_1 = 15$, $K_2 = 0$, $K_3 = 12$) can achieve the highest channel utilization when the traffic intensity is high. This is because in this case only $K_1 - K_3 = 3$ channels are reserved exclusively for handoff voice calls; while in other cases, more than 3 channels need to be reserved for handoff voice calls to maintain target QoS. Thus in the case of ($K_1 =$ 15, $K_2 = 0$, $K_3 = 12$), more channels ($C - K_1 - K_2 + K_3 = 27$ in this case) can be used by new voice calls, which leads to relative higher channel utilization. On the other hand, when the voice call intensity is low, the case of ($K_1 = 15$, $K_2 = 0$, $K_3 = 12$) provides much more channels (15 in this case) than handoff voice calls can need, thus leads to relative lower channel utilization compared to other cases. But this phenomenal can not be observed when there are more handoff voice calls ($\lambda_{vh} = 0.4\lambda_v$) as indicated in figure 7. (There, DP scheme with parameters ($K_1 = 15$, $K_2 = 0$, $K_3 = 8$) has the highest channel utilization than other cases under the whole traffic intensities considered.)

Figure 4 illustrates that, by keeping constant of K_1 value, data call blocking probability decrease as the increase of K_2 value. This is because as the increase of K_2 value, more channels can be used exclusively by data calls. On the contrary, figure 5 illustrates that, by keeping constant of K_1 value, new voice call blocking probability increase as the increase of K_2 value. This is because as the increase of K_2 value, fewer channels can be shared by voice calls. Figure 5 also indicates that, DP scheme with ($K_1 = 15$, $K_2 = 0$, $K_3 = 12$) has the lowest new voice call blocking probability. The reason is that, in this case, $C - K_1 - K_2 + K_3 = 27$ channels can be used by new voice calls. While in all other cases, fewer channels can be used by new voice calls.

For DP scheme with higher user mobility (i.e., $\lambda_{vh} = 0.4\lambda_v$), similar results can also be obtained. These are illustrated in figures 6–9. The only difference is the choice of (K_1, K_2, K_3) values. We can see from these figures that, by keeping the target handoff voice call dropping probability, DP scheme with ($K_1 = 15$, $K_2 = 0$, $K_3 = 8$) can achieve the highest channel utilization when $\lambda_{vh} = 0.4\lambda_v$. Thus in following performance comparison between DP and DTBR allocation schemes, only DP scheme with ($K_1 = 15$, $K_2 = 0$, $K_3 = 12$) (when $\lambda_{vh} = 0.2\lambda_v$) and DP scheme with ($K_1 = 15$, $K_2 = 0$, $K_3 = 8$) (when $\lambda_{vh} = 0.4\lambda_v$) are selected as representative.

4.2. Performances of DTBR scheme and its comparison with DP scheme

For DTBR allocation scheme to be ergodic under given system configuration, K_2 should be great than $K_2^{\min}$, where $K_2^{\min} = 2\rho_d = 14$. Thus in the determination of (K_1, K_2) values for DTBR scheme, we first set the values for K_1, then change the value of K_2 from high (K_1) to low ($K_2^{\min}$), till the handoff voice call dropping probability equals to 10^{-3} at $\rho_v = 16$.

Figures 10–13 depict the system performances of DTBR control scheme (including DP as a comparison) when $\lambda_{vh} = 0.2\lambda_v$, while figures 14–17 plot the results when $\lambda_{vh} = 0.4\lambda_v$.

Figure 10 presents the handoff voice call dropping probability for DP and DTBR control schemes. It demonstrates that, compared with DTBR scheme, DP scheme usually provides much lower handoff voice call dropping probability when ρ_v is low. This is because in complete partition DP scheme, those voice only channels needed to maintain target handoff voice call dropping probability at $\rho_v = 16$ cannot be shared by data calls, thus results in an over-provisioning of the bandwidth when ρ_v is low. On the contrary, in DTBR scheme, no channels are dedicated to voice calls. Except those ($C - K_2$) channels needed to keep target handoff voice call dropping probability at $\rho_v = 16$, all other channels can be shared statistically by voice calls and data calls. Thus when voice call intensity becomes low, because there still exist im-

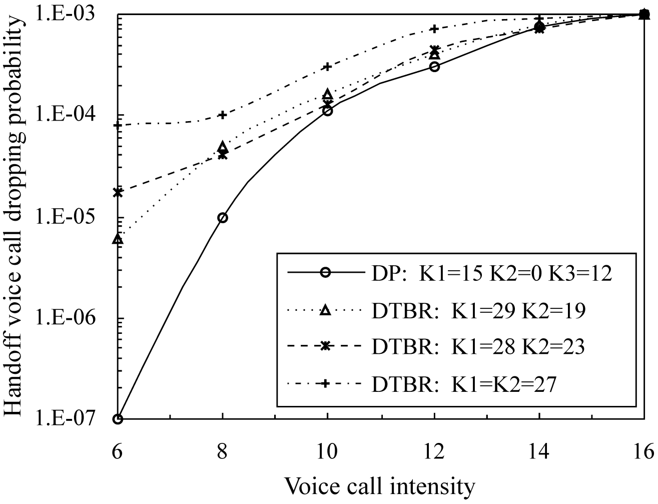


Figure 10. Handoff voice call dropping probability of DP and DTBR schemes when $\lambda_{vh} = 0.2\lambda_v$.

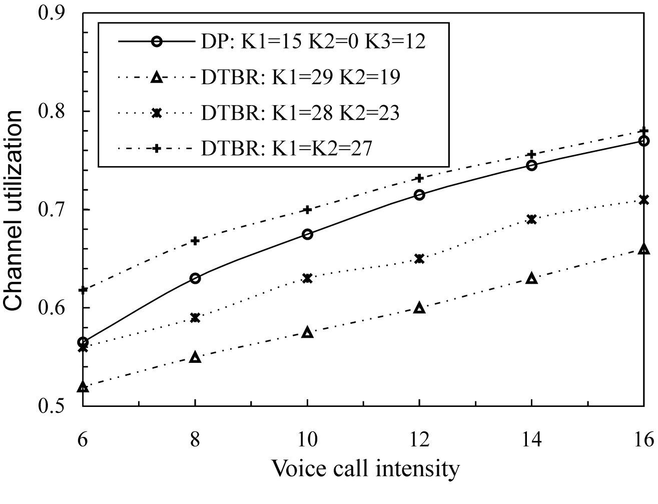


Figure 11. Channel utilization of DP and DTBR schemes when $\lambda_{vh} = 0.2\lambda_v$.

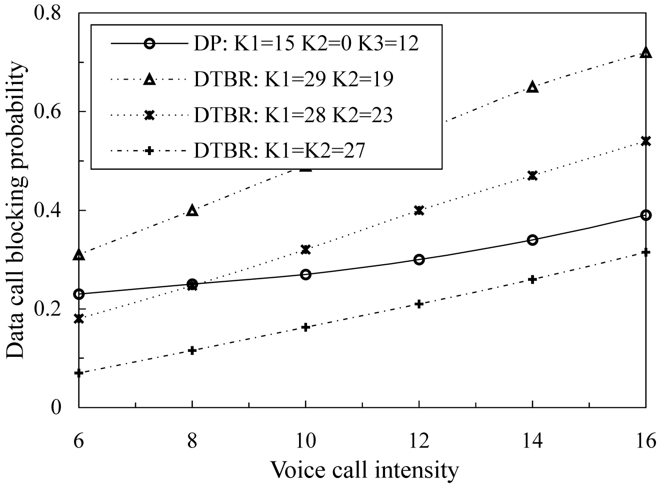


Figure 12. Data call blocking probability of DP and DTBR schemes when $\lambda_{vh} = 0.2\lambda_v$.

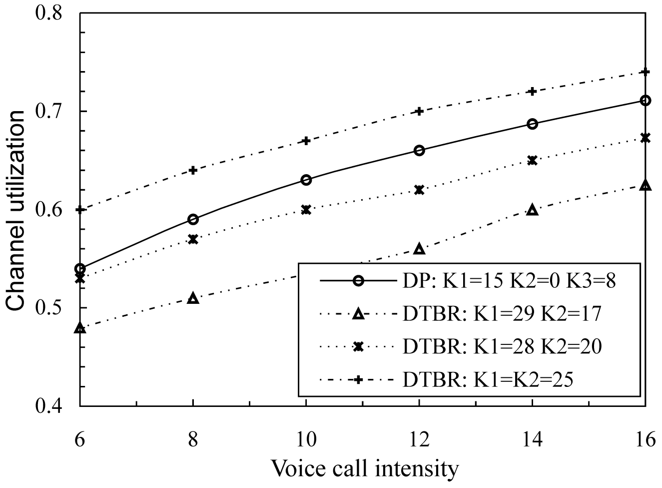


Figure 15. Channel utilization of DP and DTBR schemes when $\lambda_{vh} = 0.4\lambda_v$.

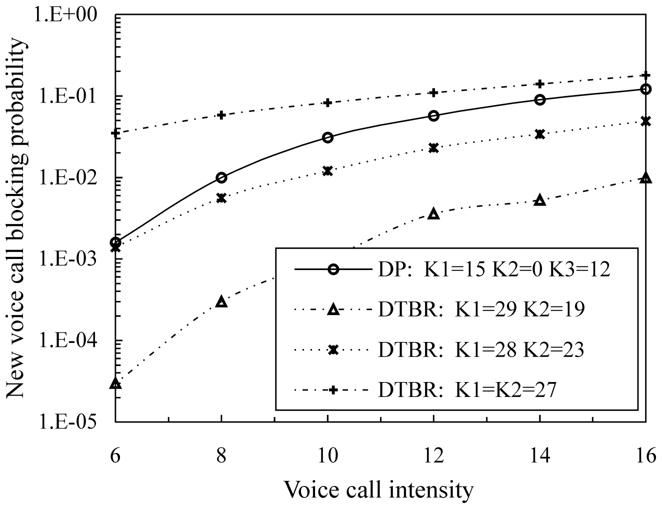


Figure 13. Voice call blocking probability of DP and DTBR schemes when $\lambda_{vh} = 0.2\lambda_v$.

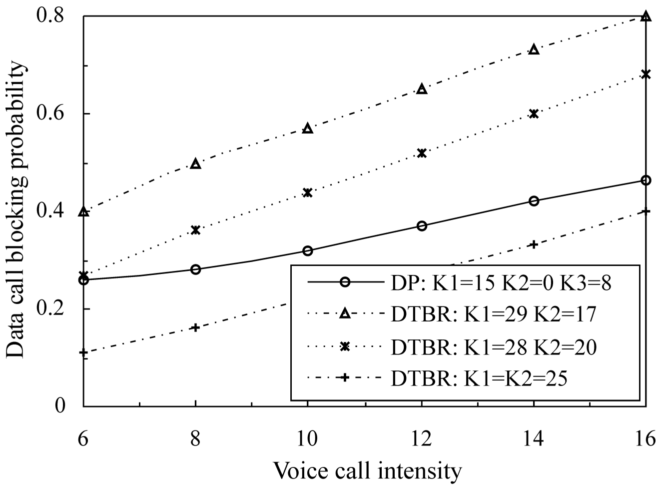


Figure 16. Data call blocking probability of DP and DTBR schemes when $\lambda_{vh} = 0.4\lambda_v$.

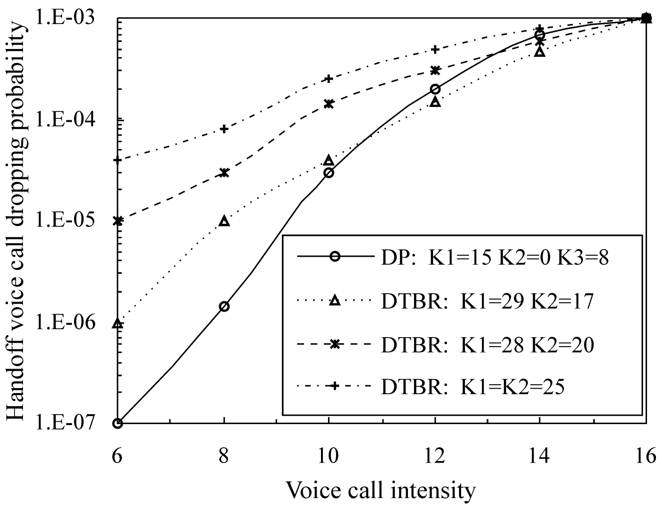


Figure 14. Handoff voice call blocking probability of DP and DTBR schemes when $\lambda_{vh} = 0.4\lambda_v$.

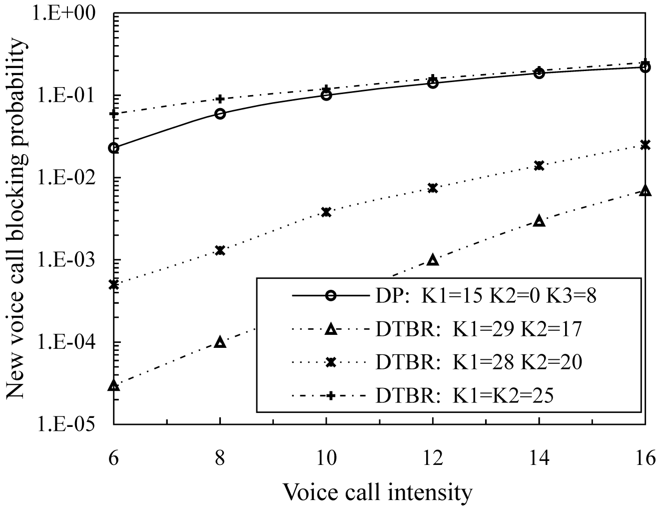


Figure 17. Voice call blocking probability of DP and DTBR schemes when $\lambda_{vh} = 0.4\lambda_v$.

pacts from data calls, handoff voice call dropping probability cannot decrease rapidly.

Figure 11 presents the average channel utilization for DP and DTBR control schemes. It shows that, among all the cases of DTBR scheme (here for clarity we only present three cases), DTBR with one threshold $K_1 = K_2 = 27$ has the highest channel utilization. The reason is that, in this case, no priority is given to new voice calls. Except a small amount of channels ($C - K_2 = 3$ in this case) are reserved for handoff voice calls to keep target QoS, all other available channels can be shared by voice calls and data calls equally. Thus leads to much higher channel utilization.

Figure 11 also indicates that, compared with the best choice of DP scheme, DTBR with one threshold $K_1 = K_2 = 27$ has much higher channel utilization. The reason is that, in DP scheme with ($K_1 = 15$, $K_2 = 0$, $K_3 = 12$), $K_1 = 15$ channels are designated exclusively for voice calls; data calls can share only the remaining $C - K_1 = 15$ channels. While in DTBR with $K_1 = K_2 = 27$, only $C - K_2 = 3$ channels are reserved for handoff voice calls, data calls can share all the other 27 channels, thus leads to higher channel utilization. This can further explain that DTBR with $K_1 = K_2 = 27$ can achieve the lowest data call blocking probability as observed from figure 12.

The high channel utilization and low data call blocking probability for DTBR with one threshold $K_1 = K_2 = 27$ is obtained at the expense of higher new voice call blocking probability, as indicated in figure 13. This is because in this case, no channels are reserved for new voice calls (that is, no priority is given to new voice calls). New voice calls have to contend with handoff voice calls and data calls to get access. On the other hand, because the other two DTBR cases reserve some channels for new voice calls, lower new voice call blocking probability is achieved. Furthermore, because DP with ($K_1 = 15$, $K_2 = 0$, $K_3 = 12$) achieves the lowest voice call blocking probability among all sets of (K_1, K_2, K_3) values (see figure 5), from figure 13 we can draw that DTBR scheme with two thresholds can achieve much lower new voice call blocking probability among all the cases of DP and DTBR schemes.

When we increase λ_{vh} to $0.4\lambda_{\mathrm{v}}$ (i.e., higher user mobility), similar results can be obtained. These are illustrated in figures 14–17. Noticeably the main difference is the selection of the control thresholds in both schemes.

From the above results, in order to maintain the target handoff voice call dropping probability under the given traffic intensities, we can draw the following conclusions:

(1) DP scheme has the lowest handoff voice call dropping probability and relative lower data call blocking probability, while at the same time keeping relative higher channel utilization.

(2) DTBR scheme with one threshold (i.e., there is no channel reserved for new voice call) can achieve the highest channel utilization and the lowest data call dropping probability.

(3) If more stringent handoff voice call dropping probability and voice call blocking probabilities are required, DTBR schemes with two thresholds yield the best performance. This is achieved at the expense of lower channel utilization and potentially higher data call blocking probability.

5. Conclusion

In this paper, we study and compare the performance of two bandwidth allocation schemes for multi-service cellular network, namely Dynamic Partition (DP) and Dual-Threshold Bandwidth Reservation (DTBR) schemes. From the results obtained from both analysis and simulation, we can conclude that, by keeping the stringent handoff voice call dropping probability under given traffic loads: (1) DP scheme always has the lowest handoff voice call dropping probability; (2) DTBR scheme with one threshold can achieve the highest channel utilization and the lowest data call blocking probability; (3) DTBR scheme with two thresholds can achieve the lower new voice call blocking probability.

There are a number of issues that will be addressed in our future research: (1) The bandwidth allocation schemes discussed in this paper are static. However as shown in [13,16], dynamic schemes can always achieve better performance. (2) The handoff rates for both voice and data following convention have been assumed given, as there lacks adequate model that can derive handoff rates based on other parameters such as new call arrival rates and mobility, similar to those in [12]. We are currently working on the dynamic control policy based on the DP scheme, and the online parameters estimation algorithm proposed in [11] seems to be a viable technique. Finally, we will be considering more realistic data arrival using TCP traffic, and exploring the self-similar behavior of the data arrival observed first in wired Ethernet networks [9] and then in wireless cellular data networks [1].

Acknowledgements

The work was supported in part by grants from Research Grants Council (RGC) under contracts CRC98/01.EG03, AOE-E01/99, HKUST6163/00E, HKUST6196/0E and N-HKUST502/02.

References

[1] M. Cheng and L.-F. Chang, Wireless dynamic channel assignment performance under packet data traffic, IEEE Journal on Selected Areas in Communications 17(7) (1999).

[2] B. Epstein and M. Schwartz, Reservation strategies for multimedia traffic in a wireless environment, in: *Proceedings of the IEEE VTC* (July 1995) pp. 165–169.

[3] Y. Fang and I. Chlamtac, Teletraffic analysis and mobility modeling of PCS networks, IEEE Journal on Selected Areas in Communications 17(7) (1999).

[4] Y. Fang, Y.-B. Lin and I. Chlamtac, Channel occupancy times and handoff rate for mobile computing and PCS networks, IEEE Transactions on Computers 47(6) (1998).

[5] R. Guerin, Queuing-blocking systems with two arrival streams and guarded channels, IEEE Transactions on Communications 36(2) (1988).

[6] Y.-R. Haung, Y.-B. Lin and J.M. Ho, Performance analysis for voice/data integration on a finite mobile systems, IEEE Transactions on Vehicular Technology 49(2) (2000).

[7] U. Herzog, L. Woo and K. Chandy, Solution of queuing problems by a recursive technique IBM Journal of Research Development 19(3) (1975).

[8] L. Kleinrock, *Queueing Systems*, Vol. I: *Theory* (Wiley, New York, 1975).

[9] W.E. Leland, M.S. Taqqu, W. Willinger and D.V. Wilson, On the self-similar nature of Ethernet traffic (extended version), IEEE/ACM Transactions on Networking 2(1) (1994).

[10] B. Li, C. Lin and S. Chanson, Analysis of a hybrid cutoff priority scheme for multiple classes of traffic in multimedia wireless networks, Wireless Networks 4(4) (1998).

[11] B. Li, L. Yin, K.Y. Michael Wong and S. Wu, An efficient and adaptive bandwidth allocation scheme for mobile wireless networks based on on-line local parameter estimations, Wireless Networks 7(2) (2001).

[12] Y.-B. Lin, Performance modeling for mobile telephone networks, IEEE Networks (November/December 1997).

[13] X. Luo, B. Li, I. Thng, Y.-B. Lin and I. Chlamtac, A measurement-based pre-assignment scheme with connection-level QoS support for multiservice mobile networks, IEEE Transactions on Wireless Communications 1(3) (2002).

[14] R. Ramjee, R. Nagarajan and D. Towsley, On optimal call admission control in cellular networks, in: *Proceedings of the IEEE Infocom* (March 1996) pp. 43–50.

[15] H.-L. Wu, L.-Z. Li, B. Li, L. Yin, I. Chlamtac and B. Li, On handoff performance for an voice/data integrated cellular system, Part II: Data buffer case, in: *Proceedings of the IEEE PIMRC* (September 2002).

[16] S. Wu, K.Y.M. Wong and B. Li, A dynamic call admission policy with precision QoS guarantee using stochastic control for mobile wireless networks, IEEE/ACM Transactions on Networking 10(2) (2002).

[17] L. Yin, B. Li, Z. Zhang and Y.-B. Lin, An optimal bandwidth allocation scheme for voice/data integrated wireless cellular networks, Part I: No-buffer case, in: *Proceedings of the IEEE WCNC* (September 2000) pp. 258–262.

Bin Li received his Bachelor Degree in industrial electric automation in Huazhong University of Science and Technology, China in 1991. After that, he became a network system engineer in China Telecom GuangDong Branch. He began his postgraduate study in the department of Electrical and Electronic Engineering of Hong Kong University of Science and Technology (HKUST) in 1994, and obtained his Mphil. Degree in 1996. He was the full time Ph.D. student and research assistant in the Department of Electrical and Electronic Engineering of HKUST from 1996 to 1998. He transferred as a part-time student in 1999, and is expected to receive his Ph.D. in 2003. He is now holding the position of Chief Operation Officer of ChinaMotion NetCom Ltd. Mr. Bin Li is a Member of IEEE since 1996. His current research areas include computer and communications systems, wireless technology and network, IP technology and network.

Lizhong Li received his B.S., M.S. and Ph.D degrees in the telecommunications and information systems from University of Electronic Science and Technology of China, Chengdu, P.R. China, in 1987, 1990 and 2000, respectively. Between 1990 and 1996, he worked on SDH, ATM switches and Wireless ATM MAC protocol in Southwestern Institute of Telecommunications Technology, Chengdu, P.R. China. Between 2000 and 2001, he worked as a Research Assistant with the Department of Electronic Engineering, City University of Hong Kong. He is currently a Research Associate with the Department of Computer Science, Hong Kong University of Science and Technology. His current research interests include PCS and wireless mobile networking supporting multimedia, especially in the area of dynamic bandwidth allocation and media access control. He has published over 10 papers in above areas.

Bo Li received the B.S. (summa cum laude) and M.S. degrees in the computer science from Tsinghua University, Beijing, P.R. China, in 1987 and 1989, respectively, and the Ph.D. degree in the computer engineering from University of Massachusetts at Amherst in 1993. Between 1993 and 1996, he worked on high performance routers and ATM switches in IBM Networking System Division, Research Triangle Park, North Carolina. Since then, he has been with the Computer Science Department, the Hong Kong University of Science and Technology. His current research interests include wireless mobile networking supporting multimedia, video multicast and all optical networks using WDM. He has published over 50 journal papers in above areas, primarily in IEEE and ACM. He serves, or has served on editorial board for ACM Mobile Computing and Communications Review (MC2R), ACM/Kluwer Journal of Wireless Networks (WINET), IEEE Journal of Selected Areas in Communications (J-SAC) – Wireless Communication Series, IEEE Transactions on Wireless Communications, IEEE Transactions on Vehicular Technology, SPIE/Kluwer Optical Networking Magazine (ONM), KICS/IEEE Journal of Communications and Networks (JCN). He served as a guest editor for IEEE Communications Magazine Special Issue on Active, Programmable, and Mobile Code Networking (April 2000), IEEE Journal on Selected Areas in Communications Special Issue on Protocols for Next Generation Optical WDM Networks (October 2000), an ACM Performance Evaluation Review Special Issue on Mobile Computing (December 2000), and SPIE/Kluwer Optical Network Magazine Special Issue on Wavelength Routed Networks: Architecture, Protocols and Experiments (January/February 2002). He is currently serving as a guest editor for IEEE Journal on Selected Areas in Communications Special Issue on Recent Advances in Service Overlay Networks, and Special Issue on QoS Delivery in Variable Topology Networks. In addition, he has been involved in organizing over 30 conferences, esp. IEEE Infocom since 1996, and will be the Co-TPC Chair for Infocom'2004. He is a member of ACM and a senior member of IEEE.

Xi-Ren Cao received the M.S. and Ph.D. degrees from Harvard University, in 1981 and 1984, respectively, where he was a research fellow from 1984 to 1986. He then worked as a principal and consultant engineer/engineering manager at Digital Equipment Corporation, U.S.A, until October 1993. Since then, he is a Professor of the Hong Kong University of Science and Technology (HKUST). He is the director of the Center for Networking at HKUST. He held visiting positions at Harvard University, University of Massachusetts at Amherst, AT&T Labs, University of Maryland at College Park, University of Notre Dame, Shanghai Jiaotong University, Nankei University, Tsinghua University, University of Science and Technology of China, and Tongji University. Xi-Ren owns two patents in data communications and published two books in the area of discrete event dynamic systems. He received the Outstanding Transactions Paper Award from the IEEE Control System Society in 1987 and the Outstanding Publication Award from the Institution of Management Science in 1990. He is a Fellow of IEEE, Associate Editor at Large of IEEE Transactions of Automatic Control, and he is/was Board of Governors of IEEE Control Systems Society, associate editor of a number of international journals and chairman of a few technical committees of international professional societies. His current research areas include discrete event dynamic systems, communication systems, signal processing, stochastic processes, and system optimization.