Analysis of a Hierarchical Cellular System with Reneging and Dropping for Waiting New and Handoff Calls

Cooper Chang, *Student Member, IEEE,* Chung-Ju Chang, *Senior Member, IEEE,* and Kuen-Rong Lo

***Abstract*— In this paper, we analyze a hierarchical cellular system with finite queues for new and handoff calls. Both the effect of the reneging of waiting new calls because of the callers' impatience and the effect of the dropping of queued handoff calls as the callers move out of the handoff area are considered, besides the effect of guard channel scheme. We successfully solve the system by adopting the multidimensional Markovian chain and using the transition-probability matrix and the signal-flow graph to obtain the average new-call blocking probability, the forced termination probability, and the average waiting time of queued new and handoff calls. We further investigate how the design parameters of buffer sizes and guard channel numbers in macrocell and microcells affect the performance of the hierarchical cellular system. The results show that provision of buffering scheme and guard channel scheme can effectively reduce the new-call blocking probability and the forced termination probability in the hierarchical cellular system, and the effectiveness is more significant in the macrocell than in the microcells.**

***Index Terms*— Buffering scheme, dropping, guard channel scheme, hierarchical cellular system, reneging.**

I. Introduction

ONE OF THE important engineering issues in cellular communication systems is to improve spectrum efficiency because teletraffic demands for wireless communications services are increasing. Microcell systems can be given more channels per unit coverage area than macrocell systems so that the spectrum efficiency of microcell systems is better than that of macrocell systems. However, microcell systems are not cost effective in areas with low-user population density due to base-station building cost; they are also not suitable for high-mobility users with large handoff rate. Therefore, cellular systems with hierarchical structure were proposed to take advantages of both microcell and macrocell systems [1]–[5].

Rappaport *et al.* proposed an early personal communication network with a radio link architecture that combines both centralized and distributed control to provide a low-cost mobile radiotelephone service [1]. In [2], Steele and Nofal proposed a priority handoff scheme in a hierarchical personal communication system to reduce the forced termination probability of calls. I *et al.* studied the spectrum sharing strategies in terms of maximizing network capacity in a hierarchical architecture [3]. Yeung and Nanda proposed a macrocell/microcell selection strategy, which can dynamically adjusted the velocity threshold, for a two-tier microcell/macrocell cellular system to increase the traffic load of the system [4]. Rappaport and Hu proposed an overflow scheme for hierarchically communication systems to reduce the blocking probabilities of both new and handoff calls [5]. All the above studies did not consider the buffer provision for the hierarchical cellular system.

Blocking probabilities of new and handoff calls are important performance indexes in designing cellular communication systems. Buffer provisioning for new and/or handoff calls can reduce the blocking probability of new and/or handoff call attempts [6]–[11]. Queueing for handoff calls is necessary since the terminal spends time, named *handoff-dwell time*, in the handoff area. Queueing for new calls is possible due to the *patience* of users. Guérin considered a system with infinite new-call buffer size, but neglected the reneging of the queued new call [8]. Hong and Rappaport proposed an appropriate analytical model and derived performance measures for a cellular mobile telephone system with infinite queueing of handoff calls [7], [9]. We studied a system with finite queues for both new and handoff calls and took reneging and dropping processes into consideration [10]. Because blocking of new-call attempts is more tolerable than forced termination of ongoing calls, guard channel scheme for handoff attempts was also proposed to minimize the blocking probability of handoff calls. Zeng *et al.* considered a system with finite queues for both new and handoff calls, but neglected reneging of waiting new calls [11].

In this paper, we analyze a hierarchical cellular system with finite queues for both new and handoff calls. Overlaid microcells cover high-teletraffic areas to enhance system capacity. Overlaying macrocells cover all of the territory to provide general service in low-teletraffic areas and to provide channels for calls overflowing from the overlaid microcells. Guard channels and waiting queues are provided for handoff calls to minimize the forced termination probabilities; waiting queues are also supported for new calls to reduce the new call blocking probability. Buffered handoff calls are given service priority higher than buffered new calls because interruption of ongoing calls upset customers more than blocking of new calls. Reneging process of waiting new calls and dropping process of waiting handoff calls are further considered in this paper.

Manuscript received May 7, 1997; revised October 9, 1997. This work was supported in part by the National Science Council, Taiwan, R.O.C., under NSC Contract 85-2213-E-009-006.

The authors are with the Department of Communication Engineering, National Chiao Tung University, Hsinchu, Taiwan, R.O.C.

Publisher Item Identifier S 0018-9545(99)05746-1.

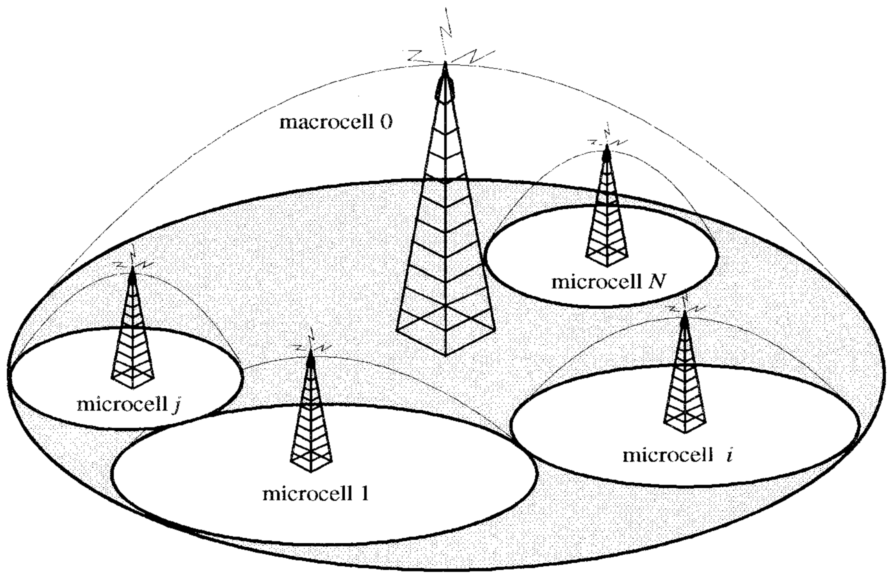


Fig. 1. A typical *macrocell*: N microcells and an overlaying macrocell.

Analysis is via a multidimensional Markov chain approach. The state probabilities are obtained by using state-transition equations since the system possesses a quasi-birth–death Markovian property [12]. We successfully derive the reneging probability of waiting new calls by using a *transition-probability matrix* approach. The transition-probability matrix is composed of one-step state-transition probability of the system, which can be used to find multistep state-transition probability by matrix multiplication. Via some substitution and derivation, the transition-probability matrix approach can further be used to obtain the average waiting time of queued new calls, and the performance measures of handoff calls are also obtained by signal-flow-graph approach. Moreover, we heuristically define a cost function to investigate the optimal guard channel pattern of the system and the suitable queue sizes for both new and handoff calls in the hierarchical cellular system.

The rest of the paper is organized as follows. The channel assignment strategy and basic assumptions for a hierarchical cellular system are presented in Section II. In Section III, we derive the new-call blocking probability and waiting time, the forced termination probability, and the handoff waiting time by the transition-probability matrix approach and the signal-flow graph approach. In Section IV, an example of hierarchically overlaid cellular system with new and handoff queues is illustrated and discussed. Finally, concluding remarks are given in Section V.

II. System Model

The hierarchical cellular system is assumed to consist of macrocells, in which a typical macrocell has an overlaying macrocell and its overlaid microcells. As Fig. 1 shows, the overlaying macrocell, denoted by macrocell 0, has the coverage area bounded by the outermost closed contour. The overlaid microcell, denoted by microcell $i, 1 \leq i \leq N$, has its own coverage area. The area outside microcells, named as the macrocell-only region, is served only by macrocell 0.

The system model of a typical macrocell is shown in Fig. 2. The ith cell, $0 \leq i \leq N$, supports C_i channels and reserves C_{hi} of C_i exclusively for handoff requests of ongoing calls from other cells. Cell i also provides two waiting queues with capacities N_{ni} and N_{hi} for new calls and handoff calls, respectively.

The *channel assignment scheme* of the *macrocell* is described as follows.

1) For a new call originating in the macrocell-only region, it will be served immediately by macrocell 0 if at its arrival the number of idle channels in macrocell 0 is larger than the number of guard channels C_{h0}. If the number of idle channels is not greater than C_{h0} and the new-call queue in cell 0 is not full, the call will be put in the queue. Otherwise, it will be blocked.
2) For a new call originating in the overlaid microcell $i, 1 \leq i \leq N$, it will be served immediately by microcell i if at its arrival the number of idle channels in microcell i is greater than the number of guard channels C_{hi}. If the new call cannot be served by microcell i, but the number of idle channels in macrocell 0 is greater than C_{h0}, it will overflow to and be served by macrocell 0. If neither microcell i nor macrocell 0 can serve the new call, but the new-call queue of microcell i still has waiting rooms, the call will be buffered in the queue. Otherwise, the call will be blocked.
3) For a handoff call coming from neighboring *macrocell* because of its high-mobility behavior, it will be directed to macrocell 0 no matter in which cell the call is in the neighboring *macrocell*. The call will be served immediately by macrocell 0 if there are free channels in macrocell 0. If macrocell 0 has no idle channels, but has free waiting room, the call will be buffered in the queue; otherwise, the call will be blocked.
4) For a handoff call moving from microcell i to macrocell-only region, it will be directed to macrocell 0. The

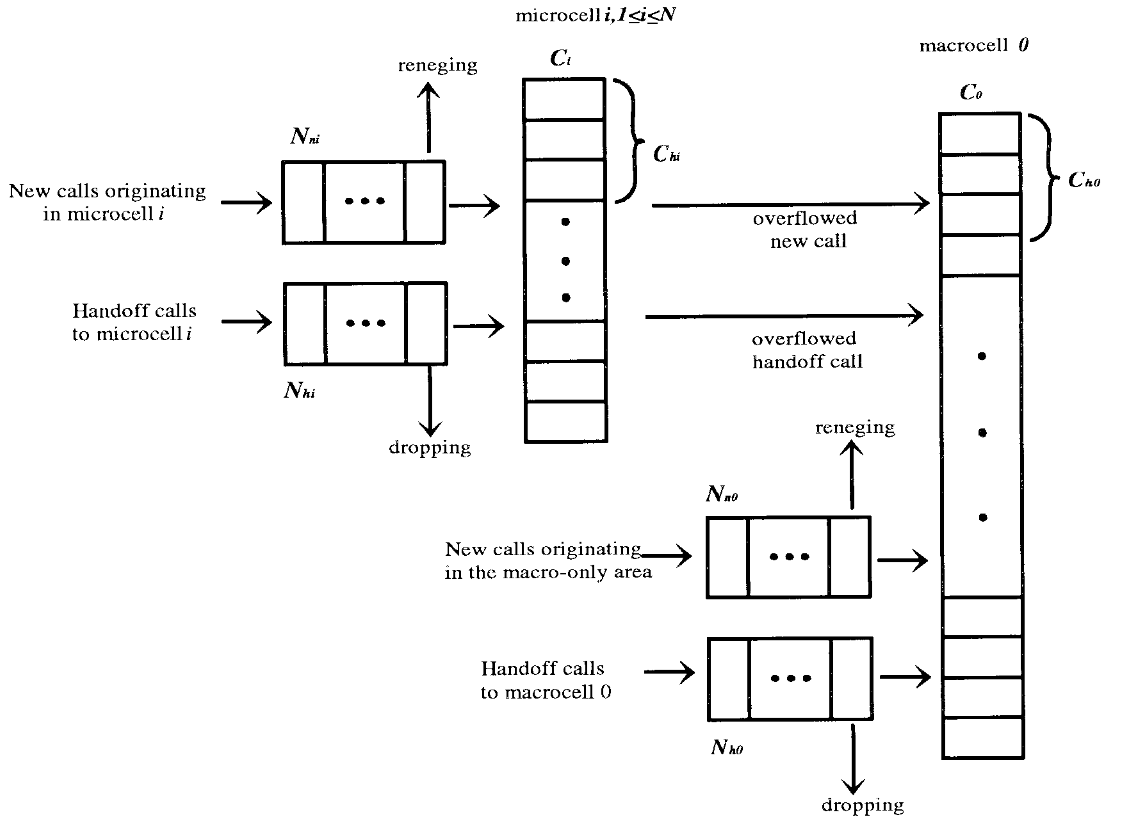


Fig. 2. The system model for a macrocell.

channel assignment scheme is the same as that for the handoff call in 3).

5) For a handoff call moving from microcell i to its neighboring microcell j, it will be served immediately by microcell j if at its arrival there are at least one free channels in microcell j. If microcell j has no idle channel, but the overlaying macrocell 0 has some free ones, the handoff call will overflow to and be served by macrocell 0; otherwise, it will be put in the queue if the queue is not full or be blocked if the queue is full.
6) In each cell, the service priority for buffered handoff calls is higher than that for buffered new calls. That is, whenever a channel is released and becomes available in cell i, $0 \leq i \leq N$, the handoff call buffered in the head of the line in cell i has the first priority to use the idle channel; the buffered new call would not be served until the number of idle channels in cell i is greater than the number of guard channels.
7) For system operation simplicity, the overflow scheme is not applied for buffered new and handoff calls in microcells.
8) The waiting new call in the queue may renege from the system if it cannot access a free channel within its patience time.
9) The waiting handoff call in the queue may be dropped by the system if it cannot access a free channel within its handoff-dwell time.
10) When a mobile platform holding a channel of overlaying macrocell 0 moves across any one of the microcells, no handoff action is needed to take.

Other basic assumptions involved in the model are stated below. The arrival process of new calls originated in the macrocell-only region or microcell i, $1 \leq i \leq N$, is a Poisson process with mean rate λ_0 or λ_i. The arrival process of handoff calls from neighboring macrocell is a Poisson process with mean rate λ_h. The unencumbered session duration of a call, denoted by T, is exponentially distributed with mean $1/\mu$. The time spent in a cell by a mobile is named by cell-dwell time. The cell-dwell time associated with new and handoff calls is assumed to be the same here. The cell-dwell time in cell i, denoted by T_i, is exponentially distributed with mean $1/\mu_i$, $0 \leq i \leq N$. The patience time of waiting new calls in cell i, denoted by T_{ni}, is exponentially distributed with mean $1/\mu_{ni}$. The handoff-dwell time of waiting handoff call to cell i, denoted by T_{hi}, has an exponential distribution with mean $1/\mu_{hi}$. The average fraction of handoff departure from cell i to cell j is denoted by α_{ij}, where $i, j = 0$ denotes the macrocell-only region, $1 \leq i, j \leq N$ denotes the microcell, and $i, j = D$ denotes the neighboring macrocell. Clearly, $\alpha_{ij} = 0$ if $i = j$ and $\Sigma_{j=0}^{N} \alpha_{ij} + \alpha_{iD} = 1$. The system is assumed to be homogeneous with the same assumption listed above for all macrocells.

III. Analysis

A. The System-State Probabilities

We define the system state for the *macrocell* as $s = [u_0, v_0, \cdots, u_i, v_i, \cdots, u_N, v_N]$, where u_i denotes the sum of the number of communicating platforms and waiting handoff calls, and v_i denotes the number of waiting new calls, $0 \le i \le N$. The states form a $(2N+2)$-dimensional sample space S, which is given by

$$S = \{s\colon s = [u_0, v_0, \cdots, u_i, v_i, \cdots, u_N, v_N] \\ 0 \le u_i \le C_i + N_{hi} \\ v_i = 0 \text{ when } 0 \le u_i \le C_i - C_{hi} - 1 \\ 0 \le v_i \le N_{ni} \text{ when } C_i - C_{hi} \le u_i \le C_i + N_{hi} \\ \text{for } 0 \le i \le N\}. \tag{1}$$

We can obtain the limiting probability of state s, denoted by $\pi(s)$, by solving the following stationary state-transition equations:

$$\sum_{t \in S} \pi(t) q(t,s) = 0, \qquad \text{for all } s \in S \tag{2}$$

and the probability conservation condition

$$\sum_{s \in S} \pi(s) = 1 \tag{3}$$

where $q(t,s), t \ne s$ denotes the transition rate from state t to state s and $q(s,s)$ denotes the transition rate out of state s. $q(s,s)$ can be obtained by

$$q(s,s) = -\sum_{t \in S, t \ne s} q(s,t). \tag{4}$$

In the following, we determine the transition rate $q(t,s)$ from state t to state $s = [u_0, v_0, \cdots, u_i, v_i, \cdots, u_N, v_N]$, based on the assumptions and statements 1)–10) of the channel assignment scheme adopted in the previous section.

New Call Arrival:

1) Denote $q_1(t,s)$ to be the transition rate from state t to state s as a new call originates in the macrocell-only region. If macrocell 0 has free channels for the new call, i.e., $t = [u_0 - 1, v_0, \cdots, u_N, v_N], 1 \le u_0 \le C_0 - C_{h0}, v_0 = 0$, the call would be served immediately by macrocell 0 and the transition rate is $q_1(t,s) = \lambda_0$. If macrocell 0 has no available channel for the new call, but has waiting rooms in the new-call queue, i.e., $t = [u_0, v_0 - 1, \cdots, u_N, v_N], C_0 - C_{h0} \le u_0 \le C_0 + N_{h0}, 1 \le v_0 \le N_{n0}$, the new call would be buffered in the new-call queue and $q_1(t,s) = \lambda_0$. Otherwise, $q_1(t,s) = 0$.
2) Denote $q_2(t,s)$ to be the transition rate from state t to state s as a new call originates in microcell $i, 1 \le i \le N$. If $t = [u_0, v_0, \cdots, u_i - 1, v_i, \cdots, u_N, v_N], 1 \le u_i \le C_i - C_{hi}, v_i = 0$, the call would be served immediately by cell i and $q_2(t,s) = \lambda_i$. If $t = [u_0 - 1, v_0, \cdots, u_i, v_i, \cdots, u_N, v_N], C_i - C_{hi} \le u_i \le C_i + N_{hi}, 1 \le u_0 \le C_0 - C_{h0}, v_0 = 0$, the call would overflow to and be served immediately by macocell 0 and $q_2(t,s) = \lambda_i$. If $t = [u_0, v_0, \cdots, u_i, v_i - 1, \cdots, u_N, v_N], C_i - C_{hi} \le u_i \le C_i + N_{hi}, 1 \le v_i \le N_{ni}, C_0 - C_{h0} \le u_0 \le C_0 + N_{h0}$, the call would be buffered in the new-call queue of cell i and $q_2(t,s) = \lambda_i$. Otherwise, $q_2(t,s) = 0$.

Handoff Call Arrival:

3) Denote $q_3(t,s)$ to be the transition rate from state t to state s as a handoff call comes from its neighboring macrocell. If macrocell 0 has free channels or has no available channel, but has waiting rooms for the handoff calls, i.e., $t = [u_0 - 1, v_0, \cdots, u_N, v_N]$, $1 \le u_0 \le C_0 - C_{h0}, v_0 = 0$, or $C_0 - C_{h0} + 1 \le u_0 \le C_0 + N_{h0}$, the handoff call will be immediately served or queued by macrocell 0 and $q_3(t,s) = \lambda_h$. Otherwise, $q_3(t,s) = 0$.
4) Denote $q_4(t,s)$ to be the transition rate from state t to state s as a handoff call arrives at the macrocell-only region from microcell i, $1 \le i \le N$. If case A $t = [u_0 - 1, v_0, \cdots, u_i + 1, v_i, \cdots, u_N, v_N], 1 \le u_0 \le C_0 - C_{h0}, v_0 = 0$, or $C_0 - C_{h0} + 1 \le u_0 \le C_0 + N_{h0}$, or case B $t = [u_0, v_0, \cdots, u_i + 1, v_i, \cdots, u_N, v_N], u_0 = C_0 + N_{h0}$, the handoff call would be immediately served or queued by macrocell 0 in case A, or be blocked by the system in case B. The transition rate in case A or case B would be $q_4(t,s) = (u_i + 1)\mu_i \alpha_{i0}$ for $0 \le u_i \le C_i - C_{hi} - 1$, $v_i = 0$, or $C_i - C_{hi} \le u_i \le C_i - 1$, or $q_4(t,s) = C_i \mu_i \alpha_{i0}$ for $C_i \le u_i \le C_i + N_{hi} - 1$.
If case C $t = [u_0 - 1, v_0, \cdots, u_i, v_i + 1, \cdots, u_N, v_N]$ and $1 \le u_0 \le C_0 - C_{h0}, v_0 = 0$, or $C_0 - C_{h0} + 1 \le u_0 \le C_0 + N_{h0}$, or case D $t = [u_0, v_0, \cdots, u_i, v_i + 1, \cdots, u_N, v_N]$, $u_0 = C_0 + N_{h0}$, the handoff call would be served or queued by macrocell 0 in case C and be blocked by the system in case D. The transition rate in cases C or D would be $q_4(t,s) = (C_i - C_{hi})\mu_i \alpha_{i0}$ for $u_i = C_i - C_{hi}, 0 \le v_i \le N_{ni} - 1$. Otherwise, $q_4(t,s) = 0$.
5) Denote $q_5(t,s)$ to be the transition rate from state t to state s as a handoff call arrives at microcell j from microcell i, $1 \le i,j \le N$. If case A $t = [u_0, v_0, \cdots, u_i + 1, v_i, \cdots, u_j - 1, v_j, \cdots, u_N, v_N]$, $1 \le u_j \le C_j - C_{hj}, v_j = 0$, or $C_j - C_{hj} + 1 \le u_j \le C_j$, case B $t = [u_0 - 1, v_0, \cdots, u_i + 1, v_i, \cdots, u_j, v_j, \cdots, u_N, v_N]$, $1 \le u_0 \le C_0 - C_{h0}, v_0 = 0$, or $C_0 - C_{h0} + 1 \le u_0 \le C_0$, $C_j \le u_j \le C_j + N_{hj}$, case C $t = [u_0, v_0, \cdots, u_i + 1, v_i, \cdots, u_j - 1, v_j, \cdots, u_N, v_N]$, $C_0 \le u_0 \le C_0 + N_{h0}$, $C_j + 1 \le u_j \le C_j + N_{hj}$, or case D $t = [u_0, v_0, \cdots, u_i + 1, v_i, \cdots, u_j, v_j, \cdots, u_N, v_N]$, $C_0 \le u_0 \le C_0 + N_{h0}$, $u_j = C_j + N_{hj}$, the handoff call would be served immediately in microcell j in case A, be overflowed to and be served by macrocell 0 in case B, be queued in microcell j in case C, and be blocked by the system in case D. The transition rate in all these cases is $q_5(t,s) = (u_i + 1)\mu_i \alpha_{ij}$ for $0 \le u_i \le C_i - C_{hi} - 1, v_i = 0$, or $C_i - C_{hi} \le u_i \le C_i - 1$, $q_5(t,s) = C_i \mu_i \alpha_{ij}$ for $C_i \le u_i \le C_i + N_{hi} - 1$. If case E, $t = [u_0, v_0, \cdots, u_i, v_i + 1, \cdots, u_j - 1, v_j, \cdots, u_N, v_N]$, $1 \le u_j \le C_j - C_{hj}, v_j = 0$, or $C_j - C_{hj} + 1 \le u_j \le C_j$, case F $t = [u_0 - 1, v_0, \cdots, u_i, v_i + 1, \cdots, u_j, v_j, \cdots, u_N, v_N]$, $1 \le u_0 \le C_0 - C_{h0}, v_0 = 0$, or $C_0 - C_{h0} + 1 \le u_0 \le C_0$, $C_j \le u_j \le C_j + N_{hj}$, case G $t = [u_0, v_0, \cdots, u_i, v_i + 1, \cdots, u_j - 1, v_j, \cdots, u_N, v_N]$, $C_0 \le u_0 \le C_0 + N_{h0}$, $C_j + 1 \le u_j \le C_j + N_{hj}$, or case H $t = [u_0, v_0, \cdots, u_i, v_i + 1, \cdots, u_j, v_j, \cdots, u_N, v_N]$,

$C_0 \leq u_0 \leq C_0 + N_{h0}$, $u_j = C_j + N_{hj}$, the call would be served immediately by microcell j in case E, be overflowed and be served by macrocell 0 in case F, be queued in microcell j in case G, or be blocked by the system in case H. For all these cases, $q_5(t,s) = (C_i - C_{hi})\mu_i\alpha_{ij}$ for $u_i = C_i - C_{hi}, 0 \leq v_i \leq N_{ni} - 1$. Otherwise, $q_5(t,s) = 0$.

Reneging (Dropping) of New (Handoff) Calls:

6) Denote $q_6(t,s)$ to be the transition rate from state t to state s as a new call reneges from the new-call queue of cell $k, 0 \leq k \leq N$. If $t = [u_0, v_0, \cdots, u_k, v_k + 1, \cdots, u_N, v_N]$, $C_k - C_{hk} \leq u_k \leq C_k + N_{hk}, 0 \leq v_k \leq N_{nk} - 1$, denoting that there are waiting new calls in the new-call queue of cell k, the transition rate would be $q_6(t,s) = (v_k + 1)\mu_{nk}$. Otherwise, $q_6(t,s) = 0$.
7) Denote $q_7(t,s)$ to be the transition rate from state t to state s as a handoff call is dropped from the handoff-call queue of cell $k, 0 \leq k \leq N$. If $t = [u_0, v_0, \cdots, u_k + 1, v_k, \cdots, u_N, v_N], C_k \leq u_k \leq C_k + N_{hi} - 1$, the transition rate would be $q_7(t,s) = (u_k - C_k + 1)\mu_{hk}$. Otherwise, $q_7(t,s) = 0$.

Handoff Call Departure:

8) Denote $q_8(t,s)$ to be the transition rate from state t to state s as a handoff call departs from cell k, $0 \leq k \leq N$, to its neighboring macrocell. If $t = [u_0, v_0, \cdots, u_k + 1, v_k, \cdots, u_N, v_N]$, $0 \leq u_k \leq C_k - 1$, denoting that the handoff queue of cell k has no waiting handoff calls as the handoff call departure from cell k occurs, the transition rate would be $q_8(t,s) = (u_k + 1)\mu_k\alpha_{kD}$. If $t = [u_0, v_0, \cdots, u_k + 1, v_k, \cdots, u_N, v_N]$, $C_k \leq u_k \leq C_k + N_{hk} - 1$, denoting that cell k has waiting handoff calls as the handoff call departs from cell k, the transition rate would be $q_8(t,s) = C_k\mu_k\alpha_{kD}$. If $t = [u_0, v_0, \cdots, u_k, v_k + 1, \cdots, u_N, v_N], u_k = C_k - C_{hk}, 0 \leq v_k \leq N_{nk} - 1$, denoting that the number of available channels equals to the number of guard channels in cell k as the handoff call departs from cell k, the transition rate would be $q_8(t,s) = (C_k - C_{hk})\mu_k\alpha_{kD}$. Otherwise, $q_8(t,s) = 0$.

Call Completion:

9) Denote $q_9(t,s)$ to be the transition rate from state t to state s as a call is released in cell $k, 0 \leq k \leq N$. If $t = [u_0, v_0, \cdots, u_k + 1, v_k, \cdots, u_N, v_N]$, the transition rate would be $q_9(t,s) = (u_k + 1)\mu$, for $< /t >$ $0 \leq u_k \leq C_k - C_{hk} - 1, v_k = 0$, or $C_k - C_{hk} \leq u_k \leq C_k - 1, 0 \leq v_k \leq N_{nk}$, or $q_9(t,s) = C_k\mu$, for $C_k \leq u_k \leq C_k + N_{hk} - 1, 0 \leq v_k \leq N_{nk}$. If $t = [u_0, v_0, \cdots, u_k, v_k + 1, \cdots, u_N, v_N], u_k = C_k - C_{hk}, 0 \leq v_k \leq N_{nk} - 1$, the transition rate is given by $q_9(t,s) = (C_k - C_{hk})\mu$. Otherwise, $q_9(t,s) = 0$.

Then $q(t,s)$ can be obtained by summing up all possible transition rates from state t to state s, which can be expressed as

$$q(t,s) = \sum_{l=1}^{9} q_l(t,s). \tag{5}$$

B. The Performance Measures

1) Average New-Call Blocking Probability: Here, blocking of new calls in the system occurs in two situations. One is that, at the arrival instant of the new call, the system has no available channel and waiting buffer; the other is that although the new call is temporarily accepted and put in the new-call queue, it finally reneges from the queue due to impatience. For a new call originating in the macrocell-only region or microcell i, the set of system states that belongs to the first situation, denoted by B_0^n or B_i^n, $1 \leq i \leq N$, can be expressed as

$$B_0^n = \{s\colon s \in S, u_0 \geq C_0 - C_{h0}, v_0 = N_{n0}\} \tag{6}$$

$$B_i^n = \{s\colon s \in S, u_i \geq C_i - C_{hi}, u_0 \geq C_0 - C_{h0}, v_i = N_{ni}\}. \tag{7}$$

And the set of system states at which an originating new call joins the new-call queue of macrocell 0 or microcell i, denoted by Q_0^n or Q_i^n, $1 \leq i \leq N$, can be expressed as

$$Q_0^n = \{s\colon s \in S, u_0 \geq C_0 - C_{h0}, 0 \leq v_0 \leq N_{n0} - 1\} \tag{8}$$

$$Q_i^n = \{s\colon s \in S, u_i \geq C_i - C_{hi}, 0 \leq v_i \leq N_{ni} - 1, u_0 \geq C_0 - C_{h0}, \}. \tag{9}$$

Thus, the blocking probability of a new call originated in macrocell-only region or microcell i, denoted by $P_{B_0}^n$ or $P_{B_i}^n$, respectively, can be expressed as

$$P_{B_i}^n = \sum_{s \in B_i^n} \pi(s) + \sum_{s \in Q_i^n} \pi(s) R_i^n(s), \qquad 0 \leq i \leq N \tag{10}$$

where $R_i^n(s)$ is the reneging probability of a waiting new call in cell i given that the system is at state s as it just arrives. We use a quasi-system state to describe the transition process of a waiting new call and a transition-probability matrix method to obtain $R_i^n(s)$. The derivation of $R_i^n(s)$ is given in the Appendix A. Therefore, the average new-call blocking probability of the system, denoted by P_B^n, can be intuitively obtained by

$$P_B^n = \frac{\sum_{i=0}^{N} \lambda_i \cdot P_{B_i}^n}{\sum_{i=0}^{N} \lambda_i}. \tag{11}$$

2) Average Waiting Time of Queued New Calls: Here, we only consider the average waiting time of queued new calls which can successfully access a free channel before reneging. Let $\overline{W}_i^n(s)$ denote the average waiting time of a queued new call which joins in cell i as the system state is at s, $0 \leq i \leq N$. The average waiting time of a queued new call in cell i, denoted by $\overline{W}_i^n$, can be obtained by

$$\overline{W}_i^n = \frac{\sum_{s\in Q_i^n} \pi(s)(1-R_i^n(s))\overline{W}_i^n(s)}{\sum_{s\in Q_i^n} \pi(s)(1-R_i^n(s))}, \qquad 0 \le i \le N. \quad (12)$$

And the average waiting time of the system, denoted by $\overline{W}^n$, can be obtained by

$$\overline{W}^n = \frac{\sum_{i=0}^{N}\left\{\lambda_i \cdot \left[\sum_{s\in Q_i^n} \pi(s)(1-R_i^n(s))\right]\cdot \overline{W}_i^n\right\}}{\sum_{i=0}^{N}\left\{\lambda_i \cdot \left[\sum_{s\in Q_i^n} \pi(s)(1-R_i^n(s))\right]\right\}}. \quad (13)$$

In the derivation of $\overline{W}_i^n(s)$, as in the Appendix A, we use quasi-system state for cell i to describe state transitions of the queued new call. Denote $\overline{\omega}(x,y)$ to be the average holding time at quasi-system state x under the condition that the quasi-system transits to quasi-system state y. Clearly

$$\overline{\omega}(x,y) = \frac{1}{\sum_{z\in S_i^N} q_n^*(x,z)} \quad (14)$$

where S_i^N is the set of quasi-system states and $q_n^*(x,z)$ is the transition rate from quasi-system state x to quasi-system state z. The definitions of S_i^N and $q_n^*(x,z)$ are given in (A.3) and (A.13), respectively. For a queued new call which is initiated at quasi-system state s $(s = z_0)$ and transits k steps, via transient states $z_j, 1 \le j \le k-1$, to service state z_k (the service state z_k is a state at which the call will be served by the system and $z_k \in A_i - \psi$, which is defined in the Appendix A), the average waiting time of this transition path would be $\Sigma_{j=1}^k \overline{\omega}(z_{j-1}, z_j)$ and the transition probability would be $\Pi_{j=1}^k p_{z_{j-1}z_j}$. Therefore, $\overline{W}_i^n(s)$ can be obtained by (15), given at the bottom of the page. Note that $\Sigma_{z_1,\cdots,z_{k-1}\in A_i^c}$ disappears as $k = 1$. The denominator in (15) is equal to $1 - R_i^n(s)$ because it denotes the probability that the queued new call can be successfully served. As for the numerator, since

$$\begin{aligned}&\sum_{k=1}^{n} \sum_{z_1,\cdots,z_{k-1}\in A_i^c} \sum_{z_k\in(A_i-\{\psi\})} \cdot \left\{\left[\prod_{j=1}^{k} p_{z_{j-1}z_j}\right]\cdot\left[\sum_{j=1}^{k}\overline{\omega}(z_{j-1},z_j)\right]\right\}\\ &= \frac{d}{d\delta}\left[\sum_{k=1}^{n} \sum_{z_1,\cdots,z_{k-1}\in A_i^c} \sum_{z_k\in(A_i-\{\psi\})} \cdot\left\{\left(\prod_{j=1}^{k} p_{z_{j-1}z_j}\right) e^{\delta\Sigma_{j=1}^k \overline{\omega}(z_{j-1},z_j)}\right\}\right]_{\delta=0}\\ &= \frac{d}{d\delta}\left[\sum_{k=1}^{n} \sum_{z_1,\cdots,z_{k-1}\in A_i^c} \sum_{z_k\in(A_i-\{\psi\})} \cdot\prod_{j=1}^{k}\left(p_{z_{j-1}z_j} e^{\delta\overline{\omega}(z_{j-1},z_j)}\right)\right]_{\delta=0}\end{aligned} \quad (16)$$

and the term within the square bracket of (16) is given by substituting the matrix element p_{xy} in (A.12) with $p_{xy}e^{\delta\overline{\omega}(x,y)}$, it can be obtained. Therefore, the average waiting time in (15) is yielded.

3) Forced Termination Probability: There are three kinds of handoff attempts in the system. The first is the handoff arriving at the overlaying macrocell 0 from adjacent macrocells, the second is the handoff arriving at the macrocell-only region from microcell i, and the third is the handoff arriving at the microcell j from microcell i, $1 \le i, j \le N$.

Failure of a handoff call coming from its neighboring macrocell occurs in two situations. One is that the handoff call is blocked at its arrival, the other is that the handoff call is temporarily buffered in the waiting queue, but is dropped by the system because the mobile moves out of the handoff area. The sets of system states belonging to these two situations are denoted by B_0^h and Q_0^h, respectively, which can be given by

$$B_0^h = \{s\colon s\in S, u_0 = C_0 + N_{h0}\} \quad (17)$$

$$Q_0^h = \{s\colon s\in S, C_0 \le u_0 \le C_0 + N_{h0} - 1\}. \quad (18)$$

Because the handoff arrival rate from neighboring macrocells is assumed to be uniform, the handoff failure probability, denoted by P_D^h, can be obtained by

$$P_D^h = \sum_{s\in B_0^h}\pi(s) + \sum_{s\in Q_0^h}\pi(s)R_0^h(s) \quad (19)$$

where $R_0^h(s)$ is the dropping probability of a queued handoff call in macrocell 0 given that the system is at state s as it just arrives. We also use quasi-system state to describe the transition process of a queued handoff call and the signal-flow-graph method to obtain $R_0^h(s)$. The derivation of $R_0^h(s)$ is given in the Appendix B.

Failure of a handoff call coming from microcell i to macrocell-only region occurs also due to being blocked immediately or being dropped from the queue. Because the handoff arrival rate is proportional to the number of communicating users in microcell i, we define $\pi_i(s) \equiv$ Prob{system is in state $s|$ a handoff attempt from microcell i occurs}. $\pi_i(s)$ can

$$\overline{W}_i^n(s) = \lim_{n\to\infty}\frac{\sum_{k=1}^{n} \sum_{z_1,\cdots,z_{k-1}\in A_i^c} \sum_{z_k\in(A_i-\{\psi\})}\left\{\left[\prod_{j=1}^{k} p_{z_{j-1}z_j}\right]\cdot\left[\sum_{j=1}^{k}\overline{\omega}(z_{j-1},z_j)\right]\right\}}{\sum_{k=1}^{n} \sum_{z_1,\cdots,z_{k-1}\in A_i^c} \sum_{z_k\in(A_i-\{\psi\})}\left\{\prod_{j=1}^{k} p_{z_{j-1}z_j}\right\}} \quad (15)$$

be expressed as

$$\pi_i(s) = \frac{h_i(u_i) \cdot \pi(s)}{\sum_{t \in S} h_i(u_i) \cdot \pi(t)} \tag{20}$$

where

$$h_i(u_i) = \begin{cases} u_i, & \text{if } u_i < C_i \\ C_i, & \text{if } u_i \geq C_i. \end{cases} \tag{21}$$

Thus, the handoff failure probability from microcell i to macrocell-only region, denoted by P_{i0}^h, can be obtained by

$$P_{i0}^h = \sum_{s \in B_0^h} \pi_i(s) + \sum_{s \in Q_0^h} \pi_i(s) R_0^h(s). \tag{22}$$

Handoff failure of a call coming from microcell i to microcell j occurs also in two situations. They are being blocked immediately and being dropped from the queue. The sets of system states belonging these cases are denoted by B_j^h and Q_j^h, respectively, which can be given by

$$B_j^h = \{s: s \in S, u_j = C_j + N_{hj}, u_0 \geq C_0\} \tag{23}$$

$$Q_j^h = \{s: s \in S, C_j \leq u_j \leq C_j + N_{hj} - 1, u_0 \geq C_0\}. \tag{24}$$

Because the handoff arrival rate is proportional to the number of communicating users in microcell i, the handoff failure probability, denoted by P_{ij}^h, can be expressed as

$$P_{ij}^h = \sum_{s \in B_j^h} \pi_i(s) + \sum_{s \in Q_j^h} \pi_i(s) R_j^h(s). \tag{25}$$

Once a call is served by macrocell 0, the call will handoff only to neighboring macrocells with a handoff requirement probability, denoted by Θ_0. Θ_0 is defined as and given by

$$\Theta_0 \equiv \text{Prob}\{T > T_0\} = \frac{\mu_0}{\mu_0 + \mu}. \tag{26}$$

And the forced termination of an arbitrarily selected call originated in macrocell-only region, denoted by P_F^0, can be obtained by

$$P_F^0 = \Theta_0 \cdot [P_D^h + (1 - P_D^h) \cdot P_F^0] = \frac{\Theta_0 \cdot P_D^h}{1 - \Theta_0 \cdot (1 - P_D^h)}. \tag{27}$$

Similarly, the handoff requirement probability of the call served in microcell i, denoted by Θ_i, $1 \leq i \leq N$, is defined as and given by

$$\Theta_i \equiv \text{Prob}\{T > T_i\} = \frac{\mu_i}{\mu_i + \mu}. \tag{28}$$

And the forced termination of an arbitrarily selected call originated in microcell i, denoted by P_F^i, can be obtained by

$$P_F^i = \Theta_i \cdot \Bigg\{ \alpha_{iD} \cdot [P_D^h + (1 - P_D^h) \cdot P_F^0] + \alpha_{i0} \cdot [P_{i0}^h + (1 - P_{i0}^h) \cdot P_F^0] + \sum_{j=1, j \neq i}^{N} \alpha_{ij} \cdot [P_{ij}^h + \Upsilon_{ij}^h \cdot P_F^0 + (1 - P_{ij}^h - \Upsilon_{ij}^h) \cdot P_F^j] \Bigg\} \tag{29}$$

where Υ_{ij}^h is the overflowing probability for a handoff attempt from microcell i to microcell j. The overflowing probability Υ_{ij}^h can be obtained by

$$\Upsilon_{ij}^h = \sum_{s \in V_j^h} \pi_i(s) \tag{30}$$

where V_j^h is defined as the set of system states at which the handoff attempts from microcell i to microcell j will overflow to macrocell 0. V_j^h can be expressed as

$$V_j^h = \{s: s \in S, u_j \geq C_j, u_0 \leq C_0 - 1\}. \tag{31}$$

P_F^i can therefore be obtained by solving the set of linear equations shown in (29).

The forced termination probability of the system, denoted by P_F, can be obtained by averaging over the forced termination probability of a call which may originate in the macrocell-only region or microcell i of the system. P_F is given by (32) at the bottom of the page, where Υ_i^n is the overflowing probability of a new call originated in microcell i. Υ_i^n can be expressed as

$$\Upsilon_i^n = \sum_{s \in V_i^n} \pi(s) \tag{33}$$

where V_i^n is defined as the set of the system states at which the new call attempts originated in microcell i can overflow to and be successfully served by macrocell 0. V_i^n can be expressed as

$$V_i^n = \{s: s \in S, u_i \geq C_i - C_{hi}, u_0 \leq C_0 - C_{h0} - 1\}. \tag{34}$$

4) Average Waiting Time of Queued Handoff Calls: Similar to the case of new calls, we only consider the average waiting time of queued handoff calls which can successfully access a free channel before moving away from the handoff area. Let $\overline{W}_i^h(s)$ denote the average waiting time of a queued handoff

$$P_F = \frac{(1 - P_{B_0}^n)\lambda_0 \cdot P_F^0 + \sum_{i=1}^{N} [\Upsilon_i^n \lambda_i \cdot P_F^0 + (1 - \Upsilon_i^n - P_{B_i}^n)\lambda_i \cdot P_F^i]}{\sum_{i=0}^{N} (1 - P_{B_i}^n)\lambda_i} \tag{32}$$

call which joins in cell i, $0 \le i \le N$, at the instant when the system state is at s. $\overline{W}_i^h(s)$ can be obtained by

$$\overline{W}_i^h(s) = \sum_{m_i=C_i}^{u_i} \frac{1}{C_i(\mu+\mu_i)+(m_i-C_i+1)\mu_{hi}}. \quad (35)$$

Let $\Lambda_i^h(s)$ denote the handoff arrival rate impinging on cell i at the instant when the system state is at s. $\Lambda_i^h(s)$ is given by

$$\Lambda_i^h(s) = \begin{cases} \lambda_h + \sum_{j=1}^{N} h_j(u_j)\mu_j\alpha_{j0}, & \text{if } i=0 \\ \sum_{j=1, j\neq i}^{N} h_j(u_j)\mu_j\alpha_{ji}, & \text{if } 1 \le i \le N. \end{cases} \quad (36)$$

The average waiting time for handoff calls of the system, denoted by $\overline{W}^h$, can therefore be obtained by

$$\overline{W}^h = \frac{\sum_{i=0}^{N}\sum_{s\in Q_i^h} \pi(s)\Lambda_i^h(s)(1-R_i^h(s))\cdot \overline{W}_i^h(s)}{\sum_{i=0}^{N}\sum_{s\in Q_i^h} \pi(s)\Lambda_i^h(s)(1-R_i^h(s))}. \quad (37)$$

IV. Numerical Examples and Discussions

In the following examples, a typical macrocell in a hierarchical cellular system is considered to contain one overlaying macrocell and two overlaid microcells. We assume the following system parameters: $(C_0, C_1, C_2) = (32, 14, 14)$, $(1/\mu) = 100$ s, $((1/\mu_0), (1/\mu_1), (1/\mu_2)) = (225, 150, 150)$ s, $((1/\mu_{h0}), (1/\mu_{h1}), (1/\mu_{h2})) = (5, 5, 5)$ s, $((1/\mu_{n0}), (1/\mu_{n1}), (1/\mu_{n2})) = (10, 10, 10)$ s, $(\alpha_{10}, \alpha_{12}, \alpha_{1D}) = (0.2, 0.5, 0.3)$, and $(\alpha_{20}, \alpha_{21}, \alpha_{2D}) = (0.2, 0.5, 0.3)$. The new-call arrival rate in each cell is set to be the same: $\lambda_0 = \lambda_1 = \lambda_2$. We use an iterative method to numerically compute the state probability of the system and show the effects of guard channels and queue capacities of new and handoff calls on the system performance measures.

Fig. 3 shows the average new-call blocking probability P_B^n and the forced termination probability P_F versus the new-call arrival rate of the system λ_n for various guard-channel patterns, where $\lambda_n = \lambda_0 + \lambda_1 + \lambda_2$, $(N_{n0}, N_{n1}, N_{n2}) = (0, 0, 0)$, and $(N_{h0}, N_{h1}, N_{h2}) = (0, 0, 0)$. It can be seen that as C_{h0} increases, P_B^n increases and P_F decreases, while as C_{h1} and C_{h2} increase, both P_B^n and P_F increase. Usually, the guard channel scheme for handoff protection would improve the forced termination probability, but deteriorate the new-call blocking probability. The reason why both P_B^n and P_F deteriorate as C_{h1} and C_{h2} increase is due to the overflow scheme of the system. The overflow scheme provides a new or handoff call impinging on a microcell an alternative of being served by macrocell 0 as the microcell has no free channel to access. The probabilities of new-call blocking and handoff failure in macrocell 0 are consequently much larger than these probabilities in microcells, and therefore the former probabilities play dominant roles in P_B^n and P_F. Also note that the increment of guard channels in microcells would

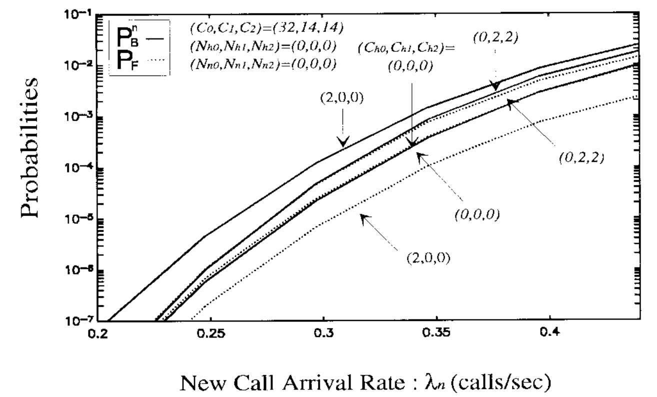


Fig. 3. The probabilities P_B^n and P_F versus new-call arrival rate λ_n for various guard channel pattern (C_{h0}, C_{h1}, C_{h2}).

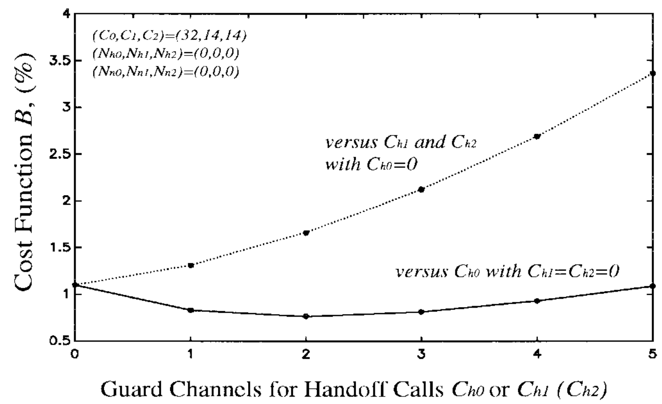


Fig. 4. The cost function B versus the number of guard channel for various C_{h0} or C_{h1} (C_{h2}).

enforce the overflow mechanism and consequently induces much deterioration on P_B^n and P_F. Therefore, in a hierarchical cellular system with overflow scheme, provision of guard channel scheme in microcells will deteriorate both new-call blocking probability and the forced termination probability of the system.

Because interruption of handoff calls upset customers much more than blocking of new calls, we heuristically define a cost function of overall blocking probability, denoted by B, to evaluate the quality-of-service of the system. B is given by

$$B = \zeta \cdot P_B^n + (1-\zeta)\cdot P_F \quad (38)$$

where ζ is a weighting factor to express the stress of P_B^n and P_F laid on the quality-of-service, $0 \le \zeta \le 1$. Fig. 4 shows the cost function B versus C_{h0} as $C_{h1} = C_{h2} = 0$ (depicted by the solid line), and B versus C_{h1} and C_{h2} $(C_{h1} = C_{h2})$ as $C_{h0} = 0$ (depicted by the dotted line), for $\zeta = 0.2$, given that $\lambda_n = 0.4455$ calls per second. We observe that the dotted line B increases monotonously as C_{h1} and C_{h2} grow up, but the solid line B has a minimum value at $C_{h0} = 2$. This result tells us the macrocell needs reservation of guard channels for handoff protection, and the optimal guard-channel pattern for the hierarchical cellular system is (C_{h0}, C_{h1}, C_{h2}) is (2, 0, 0), with $\zeta = 0.2$.

Fig. 5 shows the probabilities P_B^n and P_F versus λ_n for different queue-size patterns of new calls, where

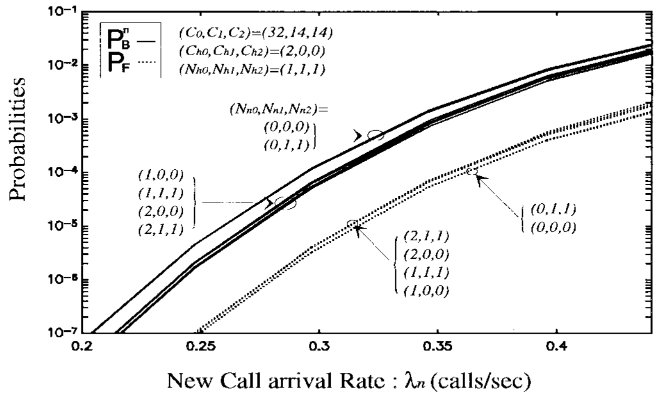


Fig. 5. The probabilities P_B^n and P_F versus new-call arrival rate λ_n for various queue-size patterns of new calls (N_{n0}, N_{n1}, N_{n2}).

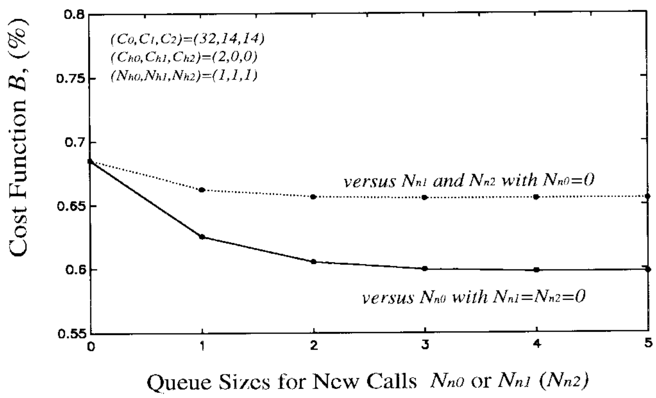


Fig. 6. The cost function B versus queue sizes for new call N_{n0} or N_{n1} (N_{n2}).

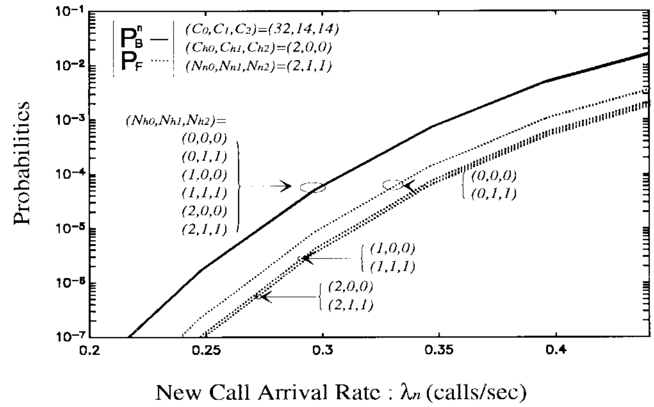


Fig. 7. The probabilities P_B^n and P_F versus new-call arrival rate λ_n for various queue-size patterns of handoff calls (N_{h0}, N_{h1}, N_{h2}).

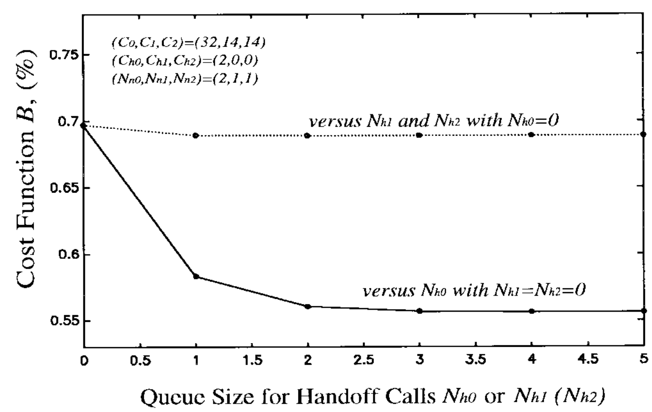


Fig. 8. The cost function B versus queue sizes for handoff call N_{h0} or N_{h1} (N_{h2}).

$(C_{h0}, C_{h1}, C_{h2}) = (2, 0, 0)$, $(N_{h0}, N_{h1}, N_{h2}) = (1, 1, 1)$. We find that the increment of new-call queue sizes in each cell induces improvement on P_B^n, but deterioration on P_F, and the improvement of P_B^n is more significant than the deterioration of P_F, for all traffic loads. Both the improvement on P_B^n and the deterioration on P_F become saturated as the queue capacities become larger because of the effects of the reneging of queued new calls and the dropping of queued handoff calls. The change of P_B^n and P_F caused by N_{n0} is more significant than by N_{n1} and N_{n2}. It is owing to the effect of the overflow scheme used in the system. Fig. 6 shows the cost function B versus N_{n0} as $N_{n1} = N_{n2} = 0$ (depicted by the solid line) and B versus N_{n1} and N_{n2} $(N_{n1} = N_{n2})$ as $N_{n0} = 0$ (depicted by the dotted line) with $\zeta = 0.2$ and $\lambda_n = 0.4455$ calls per second. It can also be found that N_{n0} plays a dominant effect over N_{ni} of microcell i in the hierarchical cellular system with overflow scheme. The saturation of B caused by queue size of macrocell 0 or microcells can be clearly observed from the figure, and of queue-size pattern $(N_{n0}, N_{n1}, N_{n2}) = (2, 1, 1)$ is sufficient in this example.

Fig. 7 shows the probabilities P_B^n and P_F versus λ_n for different queue-size patterns of handoff calls, where $(C_{h0}, C_{h1}, C_{h2}) = (2, 0, 0)$, $(N_{n0}, N_{n1}, N_{n2}) = (2, 1, 1)$. As the size of handoff queue in each cell increases, P_F decreases, but P_B^n remains almost unchanged for all traffic loads. It is because the size of handoff queue has a direct effect on performance of handoff calls, but an indirect effect on that of new calls. The increments of N_{h1} and N_{h2} have almost no effect on the performance measures. It is due to the overflow scheme of the system as mentioned before. The saturation of P_B and P_F as the queue sizes of handoff calls increase can also be observed in the figure. The phenomenon results from the reneging and the dropping of queued new and handoff calls. Fig. 8 shows the cost function B versus N_{h0} as $N_{h1} = N_{h2} = 0$ (depicted by the solid line) and B versus N_{h1} and N_{h2} $(N_{h1} = N_{h2})$ as $N_{h0} = 0$ (depicted by the dotted line), with $\zeta = 0.2$ and $\lambda_n = 0.4455$ calls per second. The saturation of B caused by increasing buffer sizes can also be observed clearly from this figure, and $(N_{h0}, N_{h1}, N_{h2}) = (2, 1, 1)$ is appropriate.

Fig. 9 shows the average waiting time of queued new calls, $\overline{W}^n$, and the waiting time of queued handoff calls, $\overline{W}^h$, versus λ_n, with appropriate design parameters: $(C_{h0}, C_{h1}, C_{h2}) = (2, 0, 0)$, $(N_{n0}, N_{n1}, N_{n2}) = (2, 1, 1)$, and $(N_{h0}, N_{h1}, N_{h2}) = (2, 1, 1)$. We can see from the figure that as the new call arrival rate increases, the increment in $\overline{W}^n$ is more significant than the increment in $\overline{W}^h$. This is because waiting handoff calls have higher priority than waiting new calls.

V. Concluding Remarks

In this paper, we successfully analyze a hierarchical cellular system with overflow scheme, where the system supports finite

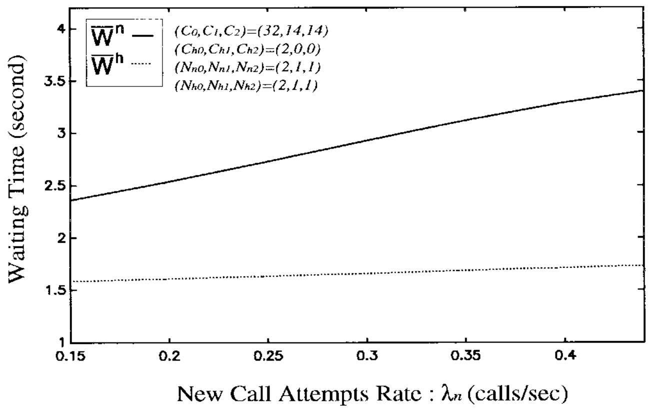


Fig. 9. The average waiting times $\overline{W}^n$ and $\overline{W}^h$ versus new call arrival rate λ_n.

buffers for new and handoff calls. The new call may renege from the system, while the handoff call may be dropped by the system. The guard channel scheme for handoff protection is also considered. We obtain the new-call blocking probability by using transition-probability matrix approach, the forced termination probability via signal-flow graph, and then the average waiting times of new and handoff calls. Interesting system phenomena showed that because of the provision of overflow scheme, design parameters of guard channel and queue size in microcells could be negligible, while these parameters in overlaying macrocell are significant. In other words, in a hierarchical cellular system with overflow scheme, it seems more significant to support guard channel for handoff protection and buffers for new and handoff calls in overlaying macrocell than to provide them in the microcells. We further heuristically propose a cost function not only to justify the above statement, but also to determine the optimal guard channel patterns and the appropriate queue-size patterns for the hierarchical cellular system.

In the study of this paper, all of the channels in the overlaying macrocell are used by the overflowing calls from microcells; the overflow scheme would induce higher traffic load in the overlaying macrocell and thus deteriorate the performance in the macrocell. Therefore, it seems appropriate to set a threshold to confine the maximum number of overflowing new/handoff calls for balancing traffic loads in macrocell and microcells. The topic is now under study.

APPENDIX A
DERIVATION OF $R_i^n(s)$

In order to obtain $R_i^n(s)$, we use quasi-system states for cell i to describe state transitions of the waiting new call in cell i. The quasi-system states for cell i are composed of two kinds of states: absorbing states and transient states. The absorbing states are either the states at which the waiting new call in cell i will renege from the system (the state is called the reneging state) or the states at which the call will be served by the system (the state is called the service state). The transient states are the states denoting that the waiting new call is still in the queue. Except for the reneging state, the quasi-system states for cell i are defined as the system states excluding those new-call arrivals in cell i queueing behind the waiting new call. The reneging state is a virtual state; we here denote it by ψ. Then the set of absorbing states, denoted by A_i, can be expressed as

$$\begin{aligned} A_i = \{x: x = [u_0^*, v_0^*, \cdots, u_i^*, v_i^*, \cdots, u_N^*, v_N^*], x \in S, \\ \text{with constraint of } u_i^* = C_i - C_{hi} - 1, v_i^* = 0\}, \\ \cup \{\psi\}. \end{aligned} \quad \text{(A.1)}$$

The set of transient states, denoted by A_i^c, can be obtained by

$$\begin{aligned} A_i^c = \{x: x = [u_0^*, v_0^*, \cdots, u_i^*, v_i^*, \cdots, u_N^*, v_N^*], x \in S, \\ \text{with constraint of } C_i - C_{hi} \le u_i^* \le C_i + N_{hi}, \\ 0 \le v_i^* \le N_{ni} - 1\}. \end{aligned} \quad \text{(A.2)}$$

And the set of the quasi-system states, denoted by S_i^N, is given by

$$S_i^N = A_i \cup A_i^c. \quad \text{(A.3)}$$

$R_i^n(s)$ can be obtained by summing up the transition probabilities of all possible transitions from state s to reneging state ψ, where state s is a quasi-system state in A_i^c. Denote $p_{xy}^{(m)}$ to be the m-step transition probability from quasi-system state x to quasi-system state y, where $x, y \in S_i^N$. Based on the memoryless assumptions of the system, the m-step quasi-system state-transition probability can be obtained by

$$p_{xy}^{(m)} = \sum_{z \in S_i^N} p_{xz}^{(m-1)} p_{zy} \quad \text{(A.4)}$$

where p_{xz} denotes one-step transition probability from quasi-system state x to quasi-system state z. In matrix form, we have

$$\begin{aligned} P^{(m)} &= P^{(m-1)} P \\ &= P^m \end{aligned} \quad \text{(A.5)}$$

where P is the transition-probability matrix $P = [p_{xy}]$, $P^{(m)} = [p_{xy}^{(m)}]$, and P^m is the mth power of P.

Without loss of generality, P can be expressed as

$$P = \begin{bmatrix} P(A_i, A_i) & P(A_i, A_i^c) \\ P(A_i^c, A_i) & P(A_i^c, A_i^c) \end{bmatrix} \quad \text{(A.6)}$$

where $P(A_1, A_2)$ denotes the transition-probability matrix from states in A_1 to another states in A_2. Because all states in A_i are absorbing states and the transition from an absorbing state to any other states is impossible except to itself, $P(A_i, A_i)$ must be an identity matrix I and $P(A_i, A_i^c)$ must be a zero matrix. So P becomes

$$P = \begin{bmatrix} I & 0 \\ P(A_i^c, A_i) & P(A_i^c, A_i^c) \end{bmatrix}. \quad \text{(A.7)}$$

Then, $P^{(m)}$ can also be expressed as

$$P^{(m)} = \begin{bmatrix} I & 0 \\ P^{(m)}(A_i^c, A_i) & P^{(m)}(A_i^c, A_i^c) \end{bmatrix} \quad \text{(A.8)}$$

where $P^{(m)}(A_i^c, A_i)$ and $P^{(m)}(A_i^c, A_i^c)$ can be obtained by

$$P^{(m)}(A_i^c, A_i) = \sum_{k=1}^{m} P^{k-1}(A_i^c, A_i^c) \cdot P(A_i^c, A_i) \quad \text{(A.9)}$$

and

$$P^{(m)}(A_i^c, A_i^c) = P^m(A_i^c, A_i^c). \tag{A.10}$$

Equation (A.9) shows that the m-step transition probabilities from states in A_i^c to states in A_i are composed of all possible transition probabilities via exactly$k$ steps: $1 \le k \le m$. A k-step transition consists of exactly $(k-1)$ steps of transition among the transient states and a one-step transition from the transient state to the absorbing state. Therefore, $R_i^n(s)$ can be obtained from the transition-probability matrix by

$$R_i^n(s) = \lim_{m\to\infty} p_{s\psi}^{(m)} \tag{A.11}$$

where $\lim_{m\to\infty} p_{s\psi}^{(m)}$ is an element in the transition matrix $P^{(\infty)}(A_i^c, A_i)$, which is given by

$$\begin{aligned} P^{(\infty)}(A_i^c, A_i) &= \lim_{m\to\infty} \sum_{k=1}^{m} P^{k-1}(A_i^c, A_i^c) \cdot P(A_i^c, A_i) \\ &= (I - P(A_i^c, A_i^c))^{-1} \cdot P(A_i^c, A_i). \end{aligned} \tag{A.12}$$

Let $q_n^*(x,y)$ denote the transition rate from quasi-system state $x = [u_0^*, v_0^*, \cdots, u_i^*, v_i^*, \cdots, u_N^*, v_N^*]$, to quasi-system state y. Based on the assumption of the system and the definition of the quasi-system, we have

$$q_n^*(x,y) = \begin{cases} \mu_{ni}, & \text{if } y = \psi \\ 0, & \text{if } y = x \\ 0, & \text{if } y = [u_0^*, v_0^*, \cdots, u_i^*, v_i^*+1, \cdots, u_N^*, v_N^*] \\ q(x,y), & \text{elsewhere} \end{cases} \tag{A.13}$$

where $q(x,y)$ is the transition rate from the system state x to system state y, as defined in Section III. Because the driving processes of the system are assumed to be memoryless, p_{xy} can be obtained by

$$p_{xy} = \frac{q_n^*(x,y)}{\sum_{z\in S_i^N} q_n^*(x,z)}. \tag{A.14}$$

Appendix B
Derivation of $R_i^h(s)$

In order to obtain $R_i^h(s)$, we also use quasi-system states for cell i to describe state transition of the waiting handoff call which joins the queue at system state s. The quasi-system states for the handoff call are also composed of transient states and absorbing states. The transient states denote that the waiting handoff call is still in the queue. The absorbing states, composed of dropping state and service states, denote that the waiting handoff call of interest leaves the waiting queue. The dropping state, which is a virtual state and is denoted by ϕ, is the state that the waiting handoff call is dropped from the queue as it moves away from the handoff area. The service states are the states that the call can be served by the system. For the waiting handoff call, the quasi-system states, except the dropping state ϕ, can be simply identified by u_i^*, which is the sum of the number of busy channels in cell i and waiting handoff calls in cell i coming before the call of interest.

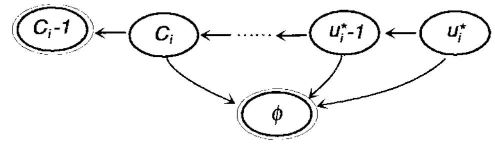


Fig. 10. The transitions of quasi-system states for handoff attempts impinge on cell i.

Fig. 10 shows a signal flow graph that portrays the transition of quasi-system states from the input state u_i^* to the output states, which are the service state $C_i - 1$ or the dropping state ϕ. The one-step transition rate from quasi-system state $x = u_i^*$ to quasi-system state y, denoted by $q_h^*(x,y)$, can be expressed as

$$q_h^*(x,y) = \begin{cases} \mu_{hi}, & \text{if } y = \phi \\ C_i(\mu+\mu_i) + (u_i^* - C_i)\mu_{hi}, & \text{if } y = u_i^* - 1 \\ 0, & \text{elsewhere.} \end{cases} \tag{B.1}$$

The probability that the call of interest can access a free channel is $1 - R_i^h(s)$ because the call of interest must be either served or dropped by the system. Obviously, $1 - R_i^h(s)$ can be obtained by

$$1 - R_i^h(s) = \prod_{u_i^*=C_i}^{u_i} \frac{C_i(\mu+\mu_i) + (u_i^* - C_i)\mu_{hi}}{C_i(\mu+\mu_i) + (u_i^* - C_i + 1)\mu_{hi}} \tag{B.2}$$

where the initial system state $s = [u_0, v_0, \cdots, u_i, v_i, \cdots, u_N, v_N]$.

References

[1] V. R. Kolavennu, S. S. Rappaport, R. R. Duersch, H. L. Lester, and C. M. Puckette, "Traffic performance characterization of a personal radiocommunication system," *Proc. Inst. Elect. Eng.*, vol. 133, pt. F, pp. 550–561, 1986.

[2] R. Steele and M. Nofal, "Teletraffic performance of microcellular personal communication networks," *Proc. Inst. Elect. Eng.*, vol. 139, pt. I, pp. 448–461, 1992.

[3] C. L. I, L. J. Greenstein, and R. D. Gitlin, "A microcell/macrocell cellular architecture for low- and high-mobility wireless users," *IEEE J. Select. Areas Commun.*, vol. 11, pp. 885–891, 1993.

[4] K. L. Yeung and S. Nanda, " Channel management in micro-cell/macrocell cellular radio systems," *IEEE Trans. Veh. Technol.*, vol. 45, no. 4, pp. 601–612, 1996.

[5] S. S. Rappaport and L. R. Hu, "Microcellular communication systems with hierarchical macrocell overlays: Traffic performance models and analysis," *Proc. IEEE*, vol. 82, no. 9, pp. 1383–1397, 1994.

[6] W. C. Y. Lee, *Mobile Cellular Telecommunication Systems.* New York: McGraw-Hill, 1990.

[7] D. Hong and S. Rappaport, "Traffic model and performance analysis for cellular mobile radio telephone system with prioritized and nonprioritized handoff procedures," *IEEE Trans. Veh. Technol.*, vol. VT-35, no. 3, pp. 77–92, 1986.

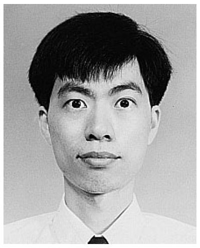
[8] R. Guérin, "Queueing-blocking system with two arrival streams and guard channels," *IEEE Trans. Commun.*, vol. 36, pp. 153–163, Feb. 1988.

[9] D. Hong and S. Rappaport, "Priority oriented channel access for cellular systems serving vehicular and portable radio telephones," *Proc. Inst. Elect. Eng.*, vol. 136, pt. I, no. 5, pp. 339–346, 1989.

[10] C. J. Chang, T. T. Su, and Y. Y. Chiang, "Analysis of a cutoff priority cellular radio system with finite queueing and reneging/dropping," *IEEE/ACM Trans. Networking*, vol. 2, no. 2, pp. 166–175, 1994.

[11] Q. A. Zeng, K. Mukumoto, and A. Fukuda, "Performance analysis of mobile cellular radio system with priority reservation handoff procedures," in *IEEE VTC'94*, pp. 1829–1833.

[12] J. Ye and S. Q. Li, "Analysis of multi-media traffic queues with finite buffer and overload control—Part I: Algorithm," in *IEEE INFOCOM'91*, pp. 1464–1474.



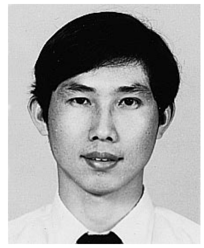
Cooper Chang was born in Keelung, Taiwan, R.O.C. He received the B.S. degree from National Chiao Tung University, Hsinchu, Taiwan, in 1990 and the M.E. degree in electrical engineering from National Tsing Hua University, Taiwan, in 1992. He is currently working towards the Ph.D. degree in communication engineering at National Chiao Tung University.

His current research interests include performance analysis of wireless communications systems and personal communications services.

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

From 1976 to 1988, he was with the Telecommunication Laboratories, Directorate General of Telecommunications, Ministry of Communications, Taiwan, as a Design Engineer, Supervisor, Project Manager, and then Division Director. There, he was involved in designing a digital switching system, ISDN user-network interface, and ISDN service and technology trials. He also acted as a Science and Technical Advisor for the Minister of the Ministry of Communications from 1987 to 1989. In August 1988, he joined the faculty of the Department of Communication Engineering and Center for Telecommunications Research, College of Electrical Engineering and Computer Science, National Chiao Tung University, where he is currently a Professor. He was Director of the Institute of Communication Engineering from August 1993 to July 1995.

Dr. Chang is currently serving as the Chairman of the IEEE Vehicular Technology Society, Taiwan Chapter. His research interests include performance evaluation, PCS, and broad-band networks.



Kuen-Rong Lo was born in Hsinchu, Taiwan, in 1957. He received the B.S. degree in electrical engineering from National Taiwan Institute of Technology, Taipei, Taiwan, in 1984, and the M.S. degree in computer science from National Tsing Hua University, Hsinchu, Taiwan in 1989.

He is currently working toward the Ph.D. degree at National Chiao Tung University, Taiwan. He is also a Project Leader in the Telecommunication Laboratories Chunghwa Telecom Co., Ltd.. His research interests lie in the broad areas of mobile communication systems and integrated services digital networks. Presently, his research is focused on channel assignment schemes, traffic performance modeling, and analysis for mobile cellular communication systems.