Performance evaluation of different resource management strategies in mobile cellular networks *

Enrico Del Re, Romano Fantacci and Giovanni Giambene

Dipartimento di Ingegneria Elettronica, Università degli Studi di Firenze, Via S. Marta 3, 50139 Firenze, Italy

Received October 1997; in final form October 1998

The foreseen mass diffusion of mobile communication services will require the identification of suitable resource management strategies to utilize efficiently the available spectrum. This paper refers to high-mobility cellular systems and carries out a performance evaluation for different channel assignment techniques that belong to the following classes: Fixed Channel Allocation (FCA), or Dynamic Channel Allocation (DCA). Suitable handoff prioritization techniques have been considered to obtain a high quality of service; in particular, the queueing of handoff requests and the use of guard channels have been investigated. The resource management techniques have been compared in terms of the following parameters: the call blocking probability, the call dropping probability, the probability of unsuccessful call and the average number of channel rearrangements per call. The joint use of DCA, guard channels, queueing of handoff requests and channel rearrangements has shown promising results for the management of both new call attempts and handoff requests.

1. Introduction

Future cellular networks will provide mobile users with multimedia services, at anytime, anywhere and in any-form (e.g., voice, data, fax, e-mail, video, etc.) [3]. Mobile networks will be characterized by the coexistence of several interoperating cellular layers: picocells in buildings, microcells in urban areas, terrestrial macrocells that overlay pico/microcells in highly populated areas and cover sub-urban and rural areas, satellite macrocells for the global coverage of the earth (figure 1) [4].

Each user will be identified by a universal personal address regardless of the terminal and the network he/she presently uses. International roaming will be on a global basis. The integration between terrestrial and satellite systems will extend the coverage far beyond the boundaries of terrestrial cellular networks. Services up to 2 Mbit/s will be provided to mobile users. These characteristics will be implemented by third generation cellular systems which are scheduled to start their operations at the beginning of next Century. They will use the bandwidth assigned at the "World Administrative Radio Conference '92" (WARC'92): 1885–2025 MHz (uplink) and

* Work supported by the Italian Space Agency and MURST.

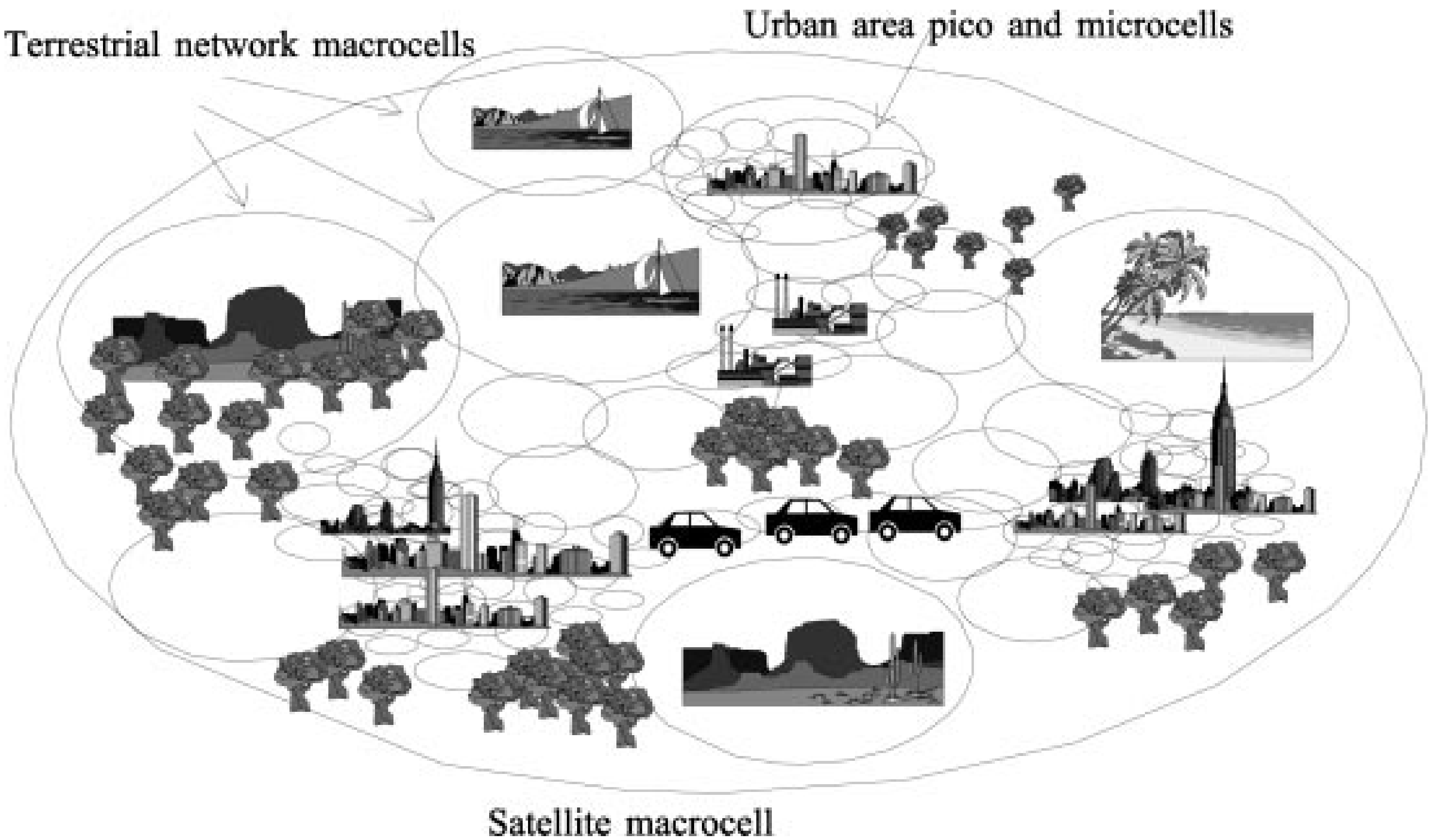

Figure 1. The hierarchical cellular architecture for future global-coverage cellular systems.

2110–2220 MHz (downlink). Both the International Telecommunication Union (ITU) and the European Telecommunications Standards Institute (ETSI) are involved in the standardization process of third generation cellular systems.

Within ITU, the Task Group 8/1 of the ITU-Radiocommunications sector is carrying out the specification of future cellular systems, identified as "Future Public Land Mobile Telecommunications Systems" (FPLMTS) and recently renamed as "International Mobile Telecommunications after the year 2000" (IMT-2000). Key features of IMT-2000 are:

(i) high degree of design commonality worldwide;

(ii) compatibility of IMT-2000 services with fixed networks;

(iii) high quality;

(iv) use of a small pocket communicator with worldwide roaming capability.

The Task Group 8/1 has already produced ITU-R series M Recommendations (some of them have been referenced in [11,13,14,28]) that deal with the characterization of services, network architectures, the satellite segment within IMT-2000, the air interface and so on.

A parallel specification process is carried out within ETSI on the "Universal Mobile Telecommunications System" (UMTS), a compatible standard with IMT-2000. ETSI envisages an evolutionary approach from the "Global System for Mobile communications" (GSM) towards UMTS, since GSM is a worldwide success and it may become a useful platform for the implementation of advanced characteristics. The UMTS standardization is carried out within ETSI by the Special Mobile Group (SMG) with the following subgroups: SMG1 for service aspects, SMG2 for radio aspects, SMG3 for network aspects, and SMG5 which coordinates the work on UMTS and defines system requirements. The first phase of UMTS standardization will be completed within 1999 [2].

The diffusion of mobile communications has grown during the past decade and this trend is expected to continue for the next years. Consequently, a network capacity increase is required. Future mobile networks will manage high traffic densities: a traffic density of hundreds of erlangs per square kilometer is expected in highly-populated urban areas. A possible way for increasing the traffic density managed by the network is to reduce cell sizes; this technique is called *cell-splitting* [21]. Let R be the side of a hexagonal cell; by assuming a fixed number of radio channels available for a given cellular network, if cell sizes reduce (i.e., cells become microcells and picocells), the capacity of the system increases proportionally to $1/R^2$. However, if cell sizes decrease, more base stations are required to cover a given area; then, it is difficult to place properly them in the territory and implementation costs increase as well. Hence, there is a lower bound for the practical values of R and, then, a limit for the application of the cell-splitting technique.

In order to face the expected traffic increase in future mobile communication networks, the viable solution addressed in this paper is represented by the use of resource management strategies that will be able to share the scarcely available radio channels among the greatest number of mobile users while maintaining an acceptable service quality. This paper addresses the study of resource management strategies for circuit-switched voice traffic; this service will be still dominant with respect to data services for some years in mobile cellular networks. The performance evaluation of techniques suitable for packet-switched traffic is left to a further study.

This paper is organized as follows: section 2 deals with the mobility characteristics of different scenarios in future mobile communications systems. An overview of channel allocation techniques and handoff management strategies is given in section 3. Moreover, section 4 presents the new resource management technique proposed in this work. Section 5 introduces the allocation techniques compared. Section 6 shows a theoretical method to evaluate the performance of the fixed channel allocation technique with both guard channels and the queueing of handoff requests. Finally, simulation results and the comparison with analytical predictions are shown in section 7.

2. Mobility scenarios

In this paper, a land mobile cellular network is envisaged and user mobility is described on the basis of the model proposed by Guérin in [15]. Accordingly, the following assumptions are considered:

(i) when a Mobile Station (MS) starts its call in a cell, a motion direction is chosen from four possible orthogonal directions, for the duration of the call;

(ii) an MS moves with a constant speed V equal to an average value suitable for each mobility scenario;

(iii) the cellular network is regular with hexagonal cells;

(iv) an MS is equally likely to start a call in any point of a cell.

The user mobility in different scenarios is characterized by the following positive dimensionless parameter:

$$\alpha = \frac{R}{VT_{\mathrm{m}}}, \tag{1}$$

where T_{m} is the average unencumbered call duration time and V is the average MS speed. Note that a decrease (increase) in α corresponds to an increase (decrease) in the user mobility level.

We will only refer to the voice service; then, we will assume $T_{\mathrm{m}} = 3$ min (this is the standard value for the classical fixed telephony). Whereas, R and V depend on the mobility scenario.

When an MS with a call in progress leaves a cell, a new channel must be assigned to it in the destination cell in order to avoid call dropping. This procedure is called (*inter-cell*) *handoff*. When a call is switched to another channel within the same cell for quality or network management needs, the procedure is denoted as *intra-cell handoff* or *channel rearrangement*.

The unencumbered call duration is considered exponentially distributed with average value $T_\mathrm{m} = 1/\mu_\mathrm{d}$. Therefore, due to the memoryless property of this distribution, each time a call originates a handoff towards a cell x (destination cell), the residual call lifetime from the call arrival instant in cell x has the same distribution as the total call duration. According to [15], an MS is within the communication range of a cell for a time (i.e., the MS sojourn time in a cell) which is exponentially distributed with average value $1/\mu_\mathrm{s}$; the rate μ_s is proportional to the rate μ_d as follows:

$$\mu_\mathrm{s} = n_\mathrm{h0}\mu_\mathrm{d} \tag{2}$$

where n_h0 represents the average number of handoff requests per call (without blocking for new call attempts and handoff requests); on the basis of the mobility hypotheses, n_h0 is given by (see [15])

$$n_\mathrm{h0} = \frac{2\sqrt{3}+3}{9\alpha}\ \frac{\text{handoffs}}{\text{call}}. \tag{3}$$

Guérin in [15] validated the use of the exponential distribution for the MS sojourn time in a cell and the previous formulas by simulating the MS mobility according to the above assumptions (Guérin verified a good agreement by assuming also more general hypotheses; e.g., an MS may change its motion direction while its call is in progress).

Parameter n_h0 only depends on the assumptions made on both the cellular topology and the user motion. As α decreases to 0, the average number of handoff requests per call, n_h0, increases; then, handoff requests are more frequent during call lifetime (i.e., MS mobility increases).

In future cellular systems the cell-splitting technique will lead towards small cells, i.e., picocells and microcells. Unfortunately, the reduction of cell sizes will cause more frequent handoff requests during call lifetime. Various mobility scenarios for future cellular systems have been shown in table 1. Typical values of α will be (assuming $T_\mathrm{m} = 3$ min): $0.1 \leqslant \alpha < 0.15$ for picocells, $0.15 \leqslant \alpha < 0.35$ for microcells, $\alpha \geqslant 0.35$ for terrestrial macrocells. Hence, on the basis of (3), typical values of n_h0 are: 6 handoffs/call for picocellular systems, 4 handoffs/call in a microcellular environment and less than 2 handoffs/call for terrestrial macrocells [24]. High values of n_h0 are expected for those Low Earth Orbit Mobile Satellite Systems (LEO-MSSs) where cells are illuminated by fixed spot-beams on the satellite (i.e., spots are not steered to compensate for the satellite motion [26]): in these conditions, cells move very fast on the earth[1] and we have very frequent handoffs between spot-beams during call lifetime [5]. However, the occurrence of these handoffs is predictable and *ad hoc* strategies can exploit this "deterministic" behavior to reserve capacity in advance so as to guarantee successful inter-beam handoffs [27].

[1] The satellite ground-track speed is of the order of 24,000 km/h for a typical LEO altitude of 780 km.

Table 1
Mobility scenarios for the next decade (mobile phone service).

Type of coverage	Mobile environment	Cell side (R, km)	Type of users	Average MS speed (V, km/h)	Average number of handoffs per call (n_{h0}) for an average call duration of 3 min	Managed traffic density (erl/km)^2
Picocells	indoor/ buildings	10^{-1}–$2 \cdot 10^{-1}$	pedestrians	2	$n_{h0} < 7$	>100
Microcells	urban area	$2 \cdot 10^{-1}$–1	cars, buses	30	$n_{h0} < 5$	10–100
Terrestrial macrocells	overlay suburban area rural area	1–35	trains, intercity buses	30–40 40–50 50–60	$n_{h0} < 2$	<10
Satellite macrocells in an integrated system	overlay of the terrestrial cellular coverage		overflow traffic from lower layers	MS speed for GEO systems	for GEO systems: $0.2 < n_{h0} < 0.02$	for GEO systems $<10^{-5}$
	complement of the coverage of the terrestrial network	>200	planes, ships, scarcely populated areas users	satellite ground-track speed for non-GEO systems	for non-GEO systems: $n_{h0} < 5$	for non-GEO systems $<10^{-3}$

In the mobility model considered in this paper there is no distinction between the cell where the call starts (i.e., *the source cell*) and any subsequent cell reached by the MS with the call in progress (i.e., *a transit cell*). This is due to the memoryless property of the exponential distributions which are used to model both the MS sojourn time in a cell and the total call duration.

The handoff probability P_h is the probability that a call in a cell originates a handoff towards an adjacent cell. We have a handoff whenever the (residual) call lifetime is greater than the MS sojourn time in a cell; both these times are exponentially distributed with rates μ_d and μ_s, respectively. Therefore, we obtain the following expression for the handoff probability P_h:

$$P_h = \text{Prob}\{\text{MS sojourn time in a cell} < \text{call lifetime}\} = \frac{\mu_s}{\mu_s + \mu_d} = \frac{n_{h0}}{1 + n_{h0}}. \quad (4)$$

According to (3), probability P_h is a function of parameter α.

The channel holding time in a cell (both source cell and transit cell) is the minimum between two exponentially distributed times: the MS sojourn time in a cell (rate μ_s) and the (residual) call lifetime (rate μ_d). Then, the channel holding time is exponentially distributed with rate μ_h obtained as follows:

$$\mu_h = \mu_d + \mu_s = \mu_d(1 + n_{h0}). \quad (5)$$

Let us assume uniform traffic: parameter λ denotes the mean arrival rate of new call attempts in a cell and parameter λ_h denotes the mean arrival rate of calls in a cell due to handoffs from adjacent cells. We can consider that, in any time interval, an equilibrium exists between the expected number of handoff requests that go into a cell and the expected number of handoff requests which leave that cell towards adjacent cells (*flow conservation condition*). This consideration leads to the following equality between the mean rate of handoff requests that leave a cell and the mean rate of handoff requests which go into this cell:

$$\lambda_h(1 - P_{b2})P_h + \lambda(1 - P_{b1})P_h = \lambda_h, \tag{6}$$

where P_{b1} denotes the blocking probability of new call attempts and P_{b2} denotes the blocking probability of handoff requests (i.e., handoff failure probability).

The left-hand side of (6) represents the average number of handoff requests that leave a cell (due to both new call attempts, $\lambda(1-P_{b1})P_h$, and handoff requests arrived in the cell, $\lambda_h(1-P_{b2})P_h$); the right-hand side of (6) gives the mean number of handoff requests which go into the cell from adjacent ones.

From (6), we obtain λ_h/λ as follows:

$$\frac{\lambda_h}{\lambda} = \frac{(1 - P_{b1})P_h}{1 - (1 - P_{b2})P_h}. \tag{7}$$

According to (7), the average rate of handoff requests towards a cell, λ_h, depends on the mean rate of new call attempts in a cell, λ, the handoff probability P_h and the blocking probabilities P_{b1} and P_{b2}.

3. Overview of radio resource management strategies

Let us assume a cellular network with a hexagonal regular layout. Two different cells may reuse the same channel[2] provided that they are at a suitable distance, called *reuse distance* D, that allows tolerable levels for the co-channel interference. The smaller D is, the greater the degree of resource reuse is. Parameter D depends on transmissions techniques, voice quality, cellular environment, and cell sectorization. In this paper, we will assume a reuse distance $D = \sqrt{3K}R$ with $K = 7$ (the *reuse factor*) [21].

Let us refer to a voice service: circuit switching is used. When a phone call arrives at a cell, if no available channel can be found in this cell, the call is blocked and lost (i.e., *Blocked Calls are Cleared*, BCC). The blocking of calls in cellular systems is a crucial problem that directly affects the quality of service perceived by users. Due to the growth of the cellular market and the scarcely available spectrum, the identification of resource management techniques able to reduce the risk of blocking becomes a

[2] In this paper, we do not consider the physical nature of the radio-communication channel (e.g., a time slot with TDMA, a code with CDMA, a frequency bandwidth with FDMA), that is considered as a "*resource*" of the system.

pressing need. Then, the two following aspects have to be carefully considered when planning future cellular systems:

- The definition of algorithms that assign channels to incoming calls in the cells (due to either new call attempts or handoff requests) so as to pack as much as possible the use of system resources and to increase their utilization (i.e., *channel allocation algorithms*).
- The identification of suitable handoff management techniques that will be able to reduce the risk of handoff failure due to a lack of available resources in the destination cell of the MS. Since call dropping (due to an unsuccessful handoff) is more undesirable for a user than the initial blocking of a new call attempt, the service of handoff requests will be prioritized with respect to the service of new call attempts (i.e., *handoff prioritization strategy*). The unavoidable drawback of any prioritization technique is that it causes an increase in the blocking of new call attempts. Therefore, it is of paramount importance to find a good trade-off between the service quality (user's needs) and the traffic quantity managed by the network (operator's needs). A useful parameter that may summarize both aspects is the probability of unsuccessful call that will be introduced later in this section.

In order to cope with the two above mentioned aspects for the achievement of future high-traffic and high-mobility cellular systems, we have considered the techniques described below.

- The channel allocation strategies considered in this paperbelong to the following classes:
  - *Fixed Channel Allocation* (FCA) [21]: channels are permanently assigned to cells. A call in a cell can only be served by an available channel belonging to the set of the cell (if any).
  - *Dynamic Channel Allocation* (DCA) [1,6,22,25,29]: when a call occurs in a cell x, an available resource in x (if any) is selected from a central pool so as to fulfill the reuse distance constraint. The assignment is made on request and a channel is only temporarily allocated to x, i.e., for the duration of the call in x. Various DCA strategies differ on the basis of the criterion used to select a channel among those available in cell x (i.e., those channels that fulfill the reuse distance constraint for cell x). These criteria are mainly based on heuristic conditions of maximum channel packing.
- The handoff prioritization schemes compared in this paper belong to the following classes:
  - *Queuing of Handoff requests* (QH) [8,17,20,31]: this strategy allows the queueing of handoff requests that do not immediately find an available resource in the destination cell.
  - *Guard Channels* (GC), i.e., a group of system resources is reserved for the exclusive service of handoff requests [31]. We have considered that system channels

are divided between *nominal channels* (to serve both new call attempts and handoff requests) and *guard channels* (to serve only handoff requests); both channel sets may be shared among cells by either FCA or DCA. In particular, if we refer to guard channels, we may classify their use according to the following three cases:

(1) guard channels permanently assigned cells by FCA [16,17,31];

(2) guard channels provided by umbrella cells that are used as backup resources to serve the overflow traffic from underlying cells [19,9];

(3) guard channels shared among cells by DCA (this is the new solution proposed in this paper).

This paper only deals with the first guard channel scheme and the last one, since they represent two extreme solutions.

The performance comparison among different resource management techniques has been carried out in terms of the following quality of service parameters:

- the blocking probability of new call attempts P_{b1},
- the handoff failure probability P_{b2},
- the call dropping probability P_{drop} (i.e., the probability that a call will experience a handoff failure before it is over),
- the probability of unsuccessful call, P_{ns} (that is, the probability that a call is either initially blocked or dropped due to the failure of a handoff request).

From [6,17], we obtain the following expression for P_{ns} (case of uniform traffic):

$$P_{ns} = P_{b1} + (1 - P_{b1})P_{drop}. \tag{8}$$

The requirement on P_{drop} has to be more severe than that on P_{b1} to take into account that call dropping affects a call in progress. According ITU-T E.771 Recommendation [23], the values of P_{drop} and P_{b1} should not exceed $5 \cdot 10^{-4}$ and 10^{-2}, respectively. Present cellular systems do not meet these severe requirements that can be considered as target values for high-quality cellular systems.

4. An efficient technique for the management of both new call attempts and handoff requests

This section describes a new resource management technique based on DCA, guard channels and handoff queueing.

4.1. A dynamic channel allocation algorithm with guard channels and channel rearrangements

The DCA solution proposed here includes a handoff prioritization technique based on both guard channels dynamically shared among all the cells and a channel rearrangement scheme.

Conventional guard channel schemes are based on the FCA technique: channels are permanently assigned to cells and within the set of resources of each cell, a number of guard channels is reserved for the exclusive service of handoffs. This solution actually reduces the call dropping probability, but it also significantly increases the blocking probability for new call attempts. A technique is proposed here in order to overcome this drawback: guard channels are shared among cells by DCA. It is expected that this strategy allows a good utilization of guard channels and, then, a smaller number of guard channels are globally required to attain a given level of handoff prioritization. In the allocation strategy proposed in this section, the DCA algorithm is also used to share the nominal channels among the cells of the network.

Cells at a distance D may reuse the same radio channel and are visually marked by the same color. Then, only $K = D^2/(3R^2)$ different colors are necessary in order to cover all the cells as in a mosaic [21]. We define the following function that assigns a color i to each cell x, according to a regular pattern due to the reuse distance D:

$$i = \text{pattern}_D(x), \quad i \in [1, 2, \ldots, K]. \tag{9}$$

Channels are assigned to cells according to DCA; each cell x has a priority order for the use of channels which is graphically represented by the color of cell x. Let us mathematically explain how this ordering is obtained for each color.

Each channel is identified by a number $\eta \in \{0, 1, 2, \ldots, M-1\}$. Within the pool of M channels, we consider G guard channels dynamically shared among all the cells. We assume that both $M-G$ and G are divisible by K. The priority order for the use of channels in a cell x, characterized by a generic color $i = \text{pattern}_D(x)$, is given by the following mapping function:

$$\eta_x(j) = \begin{cases} \left(\dfrac{M-G}{K}(i-1) + j\right) \bmod (M-G), & 0 \leqslant j \leqslant M-G-1 \\ \text{(nominal channels)}, & \\ \left(\dfrac{G}{K}(i-1) + j - (M-G)\right) \bmod (G) + M - G, & M-G \leqslant j \leqslant M-1 \\ \text{(guard channels)}, & \end{cases} \tag{10}$$

where $(a) \bmod (b)$ denotes the remainder of a divided by b.

In (10), index j denotes the priority level (in the allocation phase, $j = 0$ corresponds to the highest priority channel to be used in a cell), whereas $\eta_x(j)$ denotes a channel with jth priority for cells with color $\text{pattern}_D(x)$. In (10), the case $0 \leqslant j \leqslant M-G-1$ is distinguished from the case $M-G \leqslant j \leqslant M-1$, so that the set of guard channels and the set of nominal channels are the same for all the cells, regardless of their color; of course, the priority orders within these sets depend on the color of the cell x. The use of a different channel ordering for different colors allows a packed use of system resources.

Let us describe the management of the channel ordering given by (10) both at the call arrival (i.e., allocation phase) and at the call termination (i.e., deallocation phase) in a cell x:

- When a channel demand arrives at a cell x (due to either a new call attempt or a handoff), the algorithm starts the search for an available channel[3] in cell x (if any) from the channel with number $\eta_x(j = 0)$ given by (10). If this channel is not available in x, the next attempt is done with the channel with number $\eta_x(j = 1)$ given by (10); and so on, until an available channel is found or, after checking all channels (i.e., after $j = M - G - 1$, for a new call attempt, or after $j = M - 1$ for a handoff request), the call is blocked and lost. Note that all M channels are in principle available for handoffs, even if nominal channels are preferred with respect to guard channels.
  Differently from [25], the search does not stop if the first unused channel in cell x is not available in x (note that in this case the call is blocked in [25]). The exhaustive search made in this DCA algorithm privileges the reduction of short-term blocking conditions.
- When a call terminates in a cell x on a channel $h = \eta_x(j = k)$ due to its physical end or a handoff, the lowest priority channel allocated to this cell (i.e., a channel $\eta_x(j)$ with $j \geqslant k$) is released according to the following procedure. Let us define the set of numbers $\Pi_x = \{k, \ldots, M - 1\}$ and its subset $\Psi_x = \{j \in \Pi_x$: channel $\eta_x(j)$ is assigned to cell $x\}$. Note that $\Psi_x \neq \emptyset$, because at least $k \in \Psi_x$ (i.e., before the channel de-allocation in cell x, we have that channel h is still assigned to x). Then, channel $\eta_x(j^*)$ where $j^* = \max\{\Psi_x\}$ is released in cell x. If $j^* > h$ the call in progress on channel $\eta_x(j^*)$ must be switched on channel h (i.e., *channel rearrangement*). This channel rearrangement policy is operated by the network controller in order to pack the use of system resources and in order to release guard channels as soon as possible: such an approach privileges handoff requests against new call attempts. This technique is particularly efficient when the number of guard channels is small with respect to the number of nominal channels and when the traffic is relatively high. Of course, the use of channel rearrangements entails a greater signaling load to be supported by the network per served call. An evaluation of the average number of channel rearrangements per call for this DCA algorithm is shown in section 7 together with a comparison with other allocation strategies presented in this paper.

In a practical implementation, this DCA algorithm should use a distributed database: each cell x has a table where the state of channels in x and in its interfering

[3] A channel is available in cell x if it is not used in x or in the interfering belt of x (i.e., those cells that lie at a distance less than D from x).

cells is recorded. The state of a channel in a cell y represents whether this channel is assigned to y or not. The database is updated whenever a call starts/ends in a cell by sending updating messages towards interfering cells. The distributed database allows that when a channel request occurs in a cell an available channel (if any) can be quickly selected. More details on implementation aspects are given in [7].

4.2. The queueing of handoff requests

Since future cellular systems with pico/microcells will be characterized by very frequent handoff requests during call lifetime, we consider that the guard channel scheme by itself can not give a sufficient prioritization to handoff requests. Therefore, the queueing of handoff requests must also be used: any handoff request towards a cell y, where no channel is available, can be queued (waiting for an available channel in cell y) for the time spent by the related MS to cross the *overlap area*[4] between adjacent cells. Elapsed this maximum queueing time, if no channel becomes available in cell y, and the call is still in progress, the handoff procedure fails and the associated call is forced into termination (i.e., there is a time-out mechanism for handoff requests). According to other papers appeared in the literature [5,17,20], the time spent by an MS to cross the overlap area has been assumed exponentially distributed with mean value equal to $1/\mu_\mathrm{o}$. Parameter μ_o is related to μ_s through the degree of overlap among adjacent cells, S, defined as follows:

$$S = \frac{\mu_\mathrm{s}}{\mu_\mathrm{o}}. \qquad (11)$$

Note that S is a positive dimensionless parameter that depends on the antenna characteristics, the user mobility, the handoff criterion, the transmission techniques and the cellular layout. Obviously, the greater S is, the better the performance of the queueing strategy is. By considering a circular coverage for each cell and a hexagonal regular cellular layout [5], the lower bound for S can be derived on the basis of the mobility assumptions. Under the motion hypotheses considered in section 2, $S \approx 0.1$ in case of minimum overlap. Reasonable values of S range from 0.1 to 0.3. Obviously, future pico/microcellular systems will be characterized by higher overlap degrees than presently designed systems.

Finally, we have considered a First Input First Output (FIFO) discipline to serve the queued handoff requests in a cell.

5. Resource management techniques

In this paper the following resource management techniques have been compared under the condition of a total number of M channels, a reuse distance D (and the related cluster size K) and a total number of G guard channels.

[4] The *overlap area* between two adjacent cells is a region where the radio coverages of both cells overlap; it is also called *handoff area*.

- Dynamic Channel Allocation for nominal and Guard Channels and Queueing of Handoff requests (DCA-GC-QH). This is the technique outlined in section 4.
- Fixed Channel Allocation for nominal and Guard Channels and Queueing of Handoff requests (FCA-GC-QH).
- Fixed Channel Allocation for nominal channels, Dynamic Channel Allocation for Guard Channels, Queueing of Handoff requests (FCA&DCA-GC-QH). The DCA scheme used for guard channels is the same of that used for guard channels in the DCA-GC-QH technique.

With DCA-GC-QH, both nominal and guard channels are kept in a central pool and assigned on demand according to the technique described in section 4.

With FCA-GC-QH, the M system channels are divided in groups of $Q = M/K$ channels that are assigned to cells according to FCA; we have H guard channels and $Q - H$ nominal channels within the set of Q channels of each cell (note that $G = HK$). Channel rearrangements are used to reduce the use of guard channels to what is strictly necessary: a channel rearrangement is performed when a call served by a nominal channel z, ends in cell x and there is another call in progress in x on a guard channel g: hence, the call on channel g is switched to channel z and channel g is released in cell x.

Finally, with FCA&DCA-GC-QH the $M - G$ nominal channels are allocated to cells according to FCA: each cell has permanently assigned $(M-G)/K$ channels (i.e., FCA channels). The remaining G resources are guard channels shared among the cells according to DCA. A channel rearrangement is performed in two cases:

(i) a call ends in a cell x on an FCA channel and there is another call served in x by a guard channel of the shared pool;

(ii) a call ends in a cell x on a guard channel and there is another guard channel used in x which is more convenient to deallocate in x, according to the criterion based on (9) and (10).

6. A theoretical study for FCA-GC-QH

In this section, a theoretical approach based on a Markov chain is developed to evaluate the blocking performance of the FCA technique with handoff prioritization based on both queueing and guard channels (FCA-GC-QH) [20]. Let us summarize below the assumptions (explained before) that are used in this section to carry out a performance analysis:

- $Q = M/K$ channels are assigned to each cell by FCA [21];
- H guard channels are reserved in a cell for the exclusive service of handoff requests;

- a uniform traffic distribution is assumed; the arrival processes in a cell for both new call attempts and handoff requests are considered two independent Poisson processes,[5] with mean rates λ and λ_h, where λ_h is related to λ according to (7);
- the channel holding time in a cell (for both new call attempts and handoffs) has an exponential distribution with mean $1/\mu_\mathrm{h}$ obtained from (5);
- the maximum waiting time is a random variable which is considered exponentially distributed, with expected value equal to $1/\mu_\mathrm{o} = S/\mu_\mathrm{s}$, according to (11);
- a FIFO queueing discipline has been assumed for handoff requests.

The above assumption on the exponential distribution for the channel holding time may appear special, but it is widely used in the literature [17,20]: it derives from the exponential distributions for both the MS sojourn time in a cell and the unencumbered call duration, as explained in section 2. Commonly, the call duration is assumed exponentially distributed, whereas the distribution of the MS sojourn time in a cell depends on several factors (i.e., user mobility, cellular layout). If we consider a different distribution for the MS sojourn time in a cell, we can still use (with a good approximation) the blocking results obtained for an exponential distribution on the basis of the results presented in [24], where it is shown that the blocking performance of a channel allocation technique primarily depends on the average MS sojourn time in a cell, but it is insensitive to its variance. As a further validation of these considerations, we refer to the simple case of a loss queueing system of the $M/G/\Theta/\Theta$ type,[6] where the blocking probability for new arrivals is given by the Erlang-B formula (as for the case of an exponentially distributed service time) [18], which only depends on the average arrival rate and the average service time (and the number of servers).

We have considered a finite queue length for handoff requests. The number of places in the queue has been selected according to the following considerations. The maximum number of handoff requests in the queue of a cell is upper-bounded by the maximum number of calls in progress in adjacent cells. Since there are 6 adjacent cells and Q is the maximum number of calls in progress in each cell (for FCA), then, the maximum queue occupancy is $6Q$. Therefore, we have considered $L = 6Q$ waiting places in the queue of handoff requests. In the FCA-GC-QH case with $L = 6Q$, a handoff failure can only be due to a lack of available resources in the destination cell for all the time spent by the related MS to cross the overlap area between adjacent

[5] In a real cellular system we have a finite number of users; therefore, the arrival rate of new call attempts reduces as the number of active users increases. If we assume an infinite population of users we obtain a conservative evaluation of the blocking performance, since the new call arrival process has a constant rate, irrespective of how many calls are already in progress in the cell. Such approximation is quite good in our case, because we consider that the number of users in a cell (see section 7) is much higher than the number of available channels per cell. For more details about the impact of a finite population of users on the performance of a loss queueing system, please refer to [30].

[6] The notation of the $M/G/\Theta/\Theta$ system is explained as follows: M, Poisson arrival process / G, General distribution for the service time / Θ, number of servers / Θ, number of places in the queue. In this case, the number of servers is equal to the number of places in the queue; if a new arrival can not find an available server, it is blocked and lost.

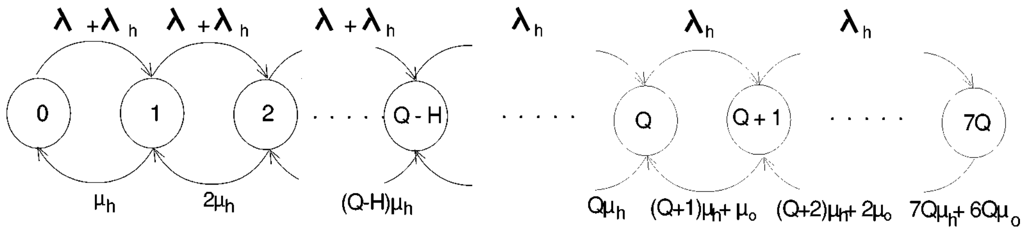

Figure 2. Markov model for FCA-GC-QH.

cells (i.e., there is no handoff failure due to a lack of places in the queue). Practically, the occupancy of the handoff queue is limited by the time-out mechanism for handoff requests due to the times spent by the related MSs to cross the overlap area.

From the above assumptions it follows that with the FCA-GC-QH technique each cell can be modelled as an $M/M/Q/T$ queueing system with nonhomogeneous arrival rates [17] (M: Poisson arrival process/M: service time exponentially distributed/Q: number of channels assigned per cell/T: the maximum number of requests in the system considering both the requests served and those waiting for service), where $T = Q + L$ and $L = 6Q$. The queueing model has been shown in figure 2.

The state of the queueing system under consideration (= a cell of the network) has been defined as sum of the number of calls in service and the number of queued handoff requests. Whenever the system is in a state $n < Q - H$, the gross arrival rate is $\lambda + \lambda_{\mathrm{h}}$; while, if $Q - H \leqslant n \leqslant Q + L$ (i.e., all nominal channels are busy), the gross arrival rate is λ_{h} (figure 2). When the system is in a state $Q + i$, for $i = 1, 2, \ldots, L$, we have the following contributions to the death rate (figure 2):

- $Q\mu_{\mathrm{h}}$, due to the termination of one of the Q calls which are served in the cell (because of either the physical end of a call or a handoff);
- $i\mu_{\mathrm{o}}$, because one of the (previously arrived) i handoff requests in the queue may be unsuccessful;
- $i\mu_{\mathrm{h}}$, because one of the (previously arrived) i handoff requests in the queue may decline, since the related call ends in the overlap area before obtaining service.

Let us analyze the state probabilities of the Markov chain in figure 2; by following the same approach proposed in [20], the probability of state n, P_n, can be derived as

$$P_n = \begin{cases} \dfrac{(\lambda + \lambda_{\mathrm{h}})^n}{n!\mu_{\mathrm{h}}^n} P_0, & 1 \leqslant n \leqslant Q - H, \\[2ex] \dfrac{(\lambda + \lambda_{\mathrm{h}})^{Q-H} \lambda_{\mathrm{h}}^{n-Q+H}}{n!\mu_{\mathrm{h}}^n} P_0, & Q - H < n \leqslant Q, \\[2ex] \dfrac{(\lambda + \lambda_{\mathrm{h}})^{Q-H} \lambda_{\mathrm{h}}^{n-Q+H}}{Q!\mu_{\mathrm{h}}^Q \prod_{j=1}^{n-Q} ((Q+j)\mu_{\mathrm{h}} + j\mu_{\mathrm{o}})} P_0, & Q < n \leqslant Q + L, \end{cases} \tag{12}$$

where the idle system probability P_0 is given by

$$P_0 = \left\{ \sum_{n=0}^{Q-H} \left[\frac{(\lambda + \lambda_{\mathrm{h}})^n}{n!\mu_{\mathrm{h}}^n} \right] + \sum_{n=Q-H+1}^{Q} \left[\frac{(\lambda + \lambda_{\mathrm{h}})^{Q-H} \lambda_{\mathrm{h}}^{n-Q+H}}{n!\mu_{\mathrm{h}}^n} \right] + \sum_{n=Q+1}^{Q+L} \left[\frac{(\lambda + \lambda_{\mathrm{h}})^{Q-H} \lambda_{\mathrm{h}}^{n-Q+H}}{Q!\mu_{\mathrm{h}}^Q \prod_{j=1}^{n-Q}((Q+j)\mu_{\mathrm{h}} + j\mu_{\mathrm{o}})} \right] \right\}^{-1}. \tag{13}$$

New arrivals are blocked when all the nominal channels are in use in the cell, i.e., when the queueing system is in states $Q - H \leqslant n \leqslant Q + L$. According to the PASTA (*Poisson Arrivals – See Time Averages)* property [32], P_{b1} results in

$$P_{\mathrm{b1}} = \sum_{n=Q-H}^{Q+L} P_n. \tag{14}$$

Moreover, probability P_{b2} can be obtained according to the following formula:

$$P_{\mathrm{b2}} = P_{Q+L} + \sum_{i=0}^{L-1} P_{Q+i} P_{\mathrm{b2}|i}, \tag{15}$$

where:

$P_{\mathrm{b2}|i} =$ failure probability for a handoff request initially queued at the position $i + 1$ (i.e., we consider the failure of a handoff request that initially arrived at the queue when other i handoff requests was waiting for service and, according to the FIFO policy, it was placed at the position $i + 1$),

$P_{Q+i} =$ probability of the state $n = Q+i$ for the queueing system, derived according to (12) and (13).

We obtain the following formula to express $P_{\mathrm{b2}|i}$:

$$P_{\mathrm{b2}|i} = (1 - P_{\mathrm{hn}})(1 - P_{S|i}), \tag{16}$$

where:

$P_{\mathrm{hn}} =$ probability that the call related to the queued handoff request ends in the handoff area; this is the probability that the time spent in the overlap area is greater than the channel holding time (in the cell):

$$P_{\mathrm{hn}} = \frac{\mu_{\mathrm{h}}}{\mu_{\mathrm{o}} + \mu_{\mathrm{h}}}, \tag{17}$$

$P_{S|i} =$ probability that a handoff request initially queued at the position $i+1$ is served by a channel before the related MS leaves the handoff area.

Then, $P_{S|i}$ can be expressed as the product of two quantities:

$P_{\mathrm{head}|i} =$ probability that a handoff request initially queued at position $i + 1$ reaches the head of the queue,

$P_{S|\text{head}}$ = probability that a handoff request at the head of the queue is served before the related MS leaves the handoff area,

$$P_{S|i} = P_{S|\text{head}} P_{\text{head}|i}. \tag{18}$$

According to the assumptions made in this paper, we obtain the following expressions for $P_{\text{head}|i}$ and $P_{S|\text{head}}$:

$$P_{\text{head}|i} = \prod_{k=1}^{i} \frac{(Q+k)\mu_\text{h} + k\mu_\text{o}}{(Q+k)\mu_\text{h} + (k+1)\mu_\text{o}}, \tag{19}$$

$$P_{S|\text{head}} = \frac{Q\mu_\text{h}}{Q\mu_\text{h} + \mu_\text{o}}. \tag{20}$$

Note that a recursive method is necessary to compute P_{b1} and P_{b2}, because λ_h (i.e., an input parameter) is related to P_{b1} and P_{b2}, through (7). The iterative method is based on parameter λ_h/λ.

We start the iterations with λ_h/λ for $P_{\text{b1}} = P_{\text{b2}} = 0$ (this is the maximum value of λ_h/λ, that, according to (4) and (7), is equal to n_{h0}). With this value of λ_h/λ, probabilities P_n for $n = 0, 1, \ldots, Q + L$ are computed according to formulas (12) and (13). These values are used to compute P_{b1} and P_{b2} and, then, the new value of λ_h/λ. This value is averaged with that obtained at the previous step. A new iteration starts with this average value of λ_h/λ. The iterative method is stopped when the relative difference between the λ_h/λ values computed in two subsequent steps is below a given threshold (i.e., 10^{-3}). Then, according to the mobility assumptions and the uniform traffic, the call dropping probability can be expressed as follows [6,17]:

$$P_{\text{drop}} = \frac{P_{\text{b2}} P_\text{h}}{1 - (1 - P_{\text{b2}}) P_\text{h}}. \tag{21}$$

Finally, probability P_{ns} can be obtained from (8), where P_{drop} is given by (21). Theoretical results have been shown in figure 11 and compared with simulation results as explained in the next section.

7. Simulation results

The following system parameters have been assumed for the simulations:

- The arrival process for new call attempts is Poisson cell-to-cell independent with average call arrival rate equal to $\lambda \in [1, 2.6]$ calls/min/cell (i.e., uniform traffic[7]).

[7] It is well known that the superior performance of DCA-like schemes with respect to FCA-like techniques is more evident in the presence of non-uniform traffic rather than in the presence of uniform traffic [10]. Therefore, the assumption made here of uniform traffic permits to compare the techniques in the conditions less favorable to highlight the advantages of DCA-like schemes. Nevertheless, we will show that our DCA-GC-QH strategy outperforms the other examined techniques (i.e., FCA-GC-QH and FCA&DCA-GC-QH) in the uniform traffic case.

Table 2
Traffic and capacity comparison ($M = 70,\ S = 0.2,\ \alpha = 0.3$, 7 guard channels).

Resource management technique	Maximum traffic per cell, erl (ITU-T requirements)	Maximum number of users per cell
FCA-GC-QH	3.2	128
FCA&DCA-GC-QH	3.6	144
DCA-GC-QH	6	240

- The unencumbered call duration is exponentially distributed with mean value T_{m} equal to 3 min; the traffic intensity per cell due to new call attempts is given by $\lambda T_{\mathrm{m}} \in [3, 8]$ erl/cell.
- The reuse distance is $D = \sqrt{21}R$ (the FCA cluster is formed by $K = 7$ cells).
- The simulated cellular network is parallelogram shaped with 7 cells per side; this network is folded onto itself (both horizontally and vertically), as described in [8].
- The number of channels available for the system, M, is equal to 70 (unless different values are explicitly considered).
- The mobility parameter α varies from 0.2 to 0.5; we have selected $\alpha = 0.3$ for a typical microcellular environment (correspondingly, $n_{\mathrm{h0}} \approx 2.39$ handoffs/call).
- The overlap degree S varies from 0 (i.e., no queueing for handoff requests) to 0.3.
- There are $L = 6M/K$ places in a cell for handoffs which wait to be served.[8]

Figures 3 and 4 show the behaviors in terms of P_{b1} and P_{drop} for the following techniques: FCA-GC-QH, FCA&DCA-GC-QH, DCA-GC-QH with $G = 7$ guard channels within the pool of $M = 70$ channels,[9] $S = 0.2,\ \alpha = 0.3$. The ITU-T target levels for both P_{b1} and P_{drop} [23] are shown by dash-dot lines in these figures (see section 3). The FCA&DCA-GC-QH technique reduces P_{drop} with respect to FCA-GC-QH, but this result is obtained at the expenses of an increase in P_{b1}. DCA-GC-QH gives the best performance in terms of both P_{b1} and P_{drop}. The maximum traffic loads per cell that permit to fulfill ITU-T E.771 requirements on both P_{b1} and P_{drop} for each technique (i.e., FCA-GC-QH, FCA&DCA-GC-QH and DCA-GC-QH) are shown in table 2, where these traffic values have been also translated in terms of the number of simultaneous users per cell by assuming 25 merl/user (we have considered that, on average, a user spends 30–40 min a day at the telephone [32]).

[8] In the previous section, we have shown that for FCA-GC-QH a queue length $L = 6M/K$ assures that no handoff request fails due to a lack of available places in the waiting list of a cell.We have verified by simulations that this result is still valid for both DCA-GC-QH and FCA&DCA-GC-QH, under the conditions detailed at the beginning of this section.

[9] In particular, FCA-GC-QH has 9 nominal channels per cell and 1 guard channel per cell (i.e., $H = 1$); FCA&DCA-GC-QH has 9 nominal channels per cell and 7 guard channels dynamically shared among all the cells; DCA-GC-QH has 63 nominal channels and 7 guard channels and these channels are managed according to (9) and (10). Then, the total number of system channels is 70 for all these techniques.

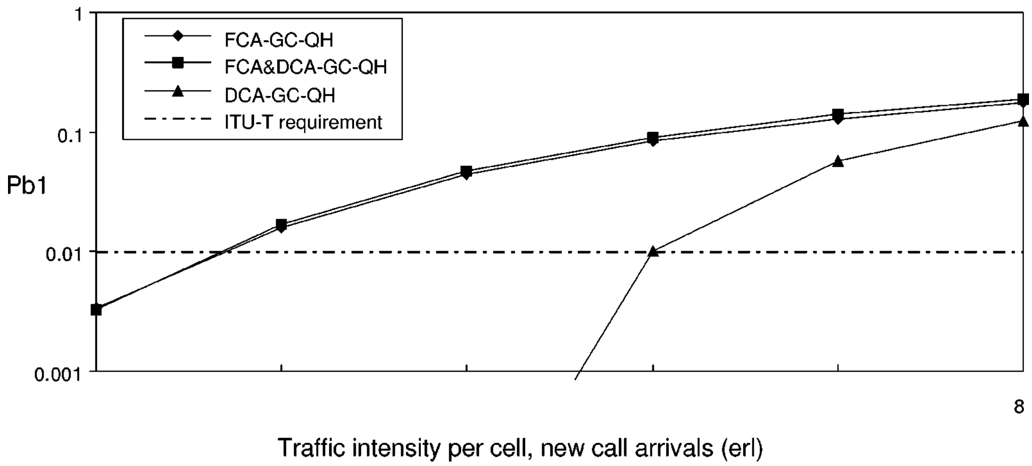

Figure 3. Comparison between FCA-GC-QH, FCA&DCA-GC-QH and DCA-GC-QH in terms of the call blocking probability, P_{b1} ($M = 70$, $S = 0.2$, $\alpha = 0.3$, 7 guard channels).

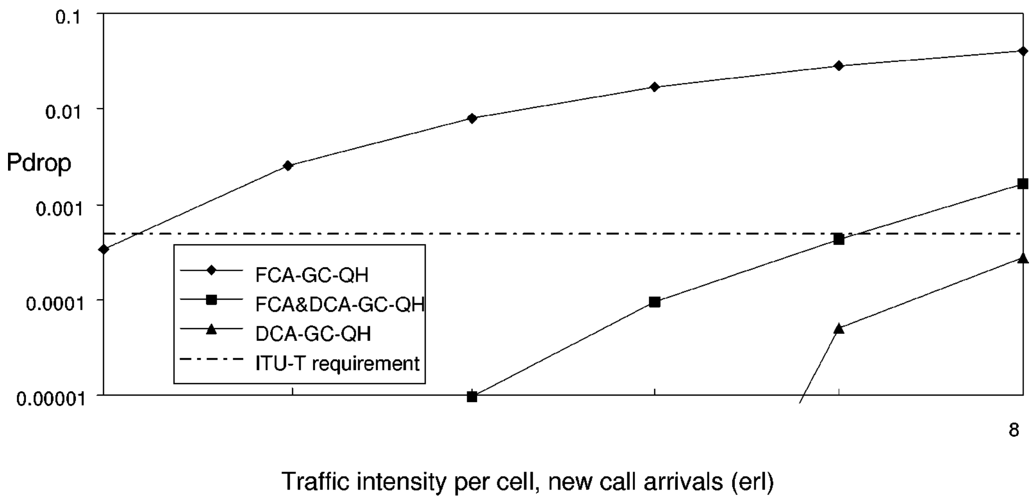

Figure 4. Comparison between FCA-GC-QH, FCA&DCA-GC-QH and DCA-GC-QH in terms of the call dropping probability, P_{drop} ($M = 70$, $S = 0.2$, $\alpha = 0.3$, 7 guard channels).

Under the same assumptions used for figures 3 and 4, figure 5 compares the performance of the proposed techniques in terms of P_{ns}. The evident result is that DCA-GC-QH outperforms FCA&DCA-GC-QH and this technique, in turn, outperforms FCA-GC-QH.

The proposed techniques have been also compared in terms of the average number of channel rearrangements per call accepted into the network, N_r; in figure 6 we have shown the results of this comparison carried out for $M = 70$ channels, $G = 7$ guard channels, $S = 0.2$ and $\alpha = 0.3$. Note that the corresponding blocking performance is shown in figures 3–5. For all these techniques, we have that N_r increases as the traffic

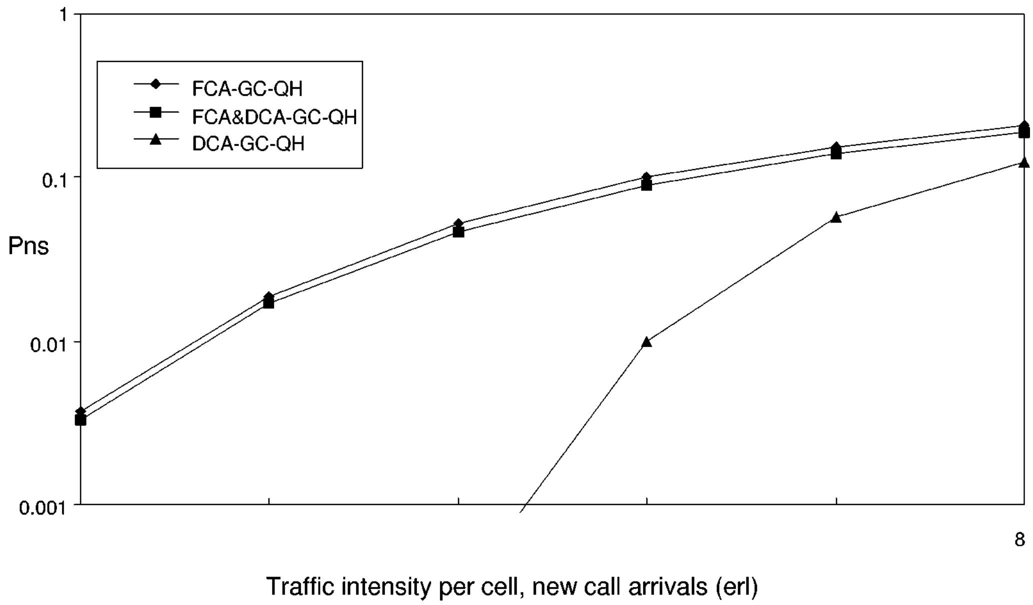

Figure 5. Comparison between FCA-GC-QH, FCA&DCA-GC-QH and DCA-GC-QH in terms of the probability of unsuccessful call, P_{ns} ($M = 70$, $S = 0.2$, $\alpha = 0.3$, 7 guard channels).

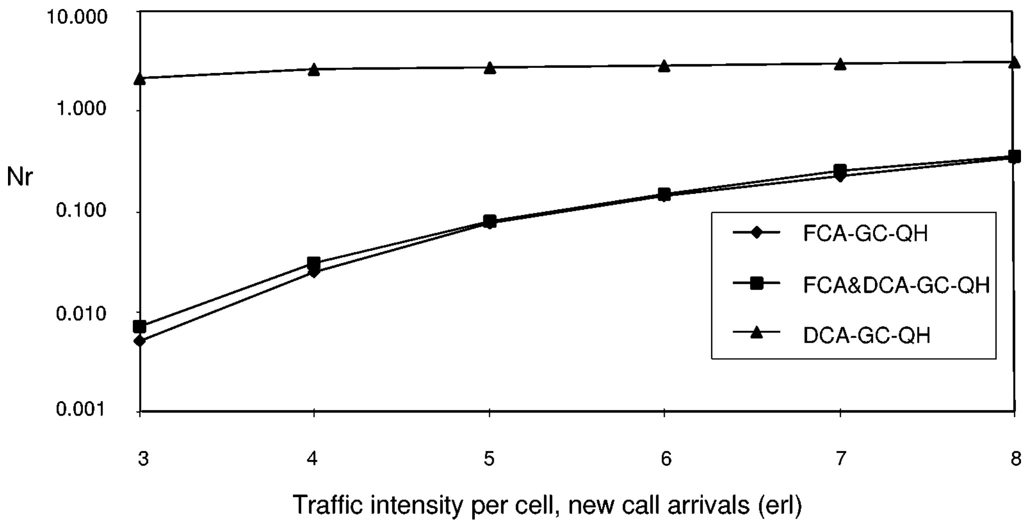

Figure 6. Comparison between FCA-GC-QH, FCA&DCA-GC-QH and DCA-GC-QH in terms of the average number of channel rearrangements per call, N_r ($M = 70$, $S = 0.2$, $\alpha = 0.3$, 7 guard channels).

intensity per cell increases. FCA-GC-QH and FCA&DCA-GC-QH entail about the same value of N_r, but FCA&DCA-GC-QH allows a better handoff management (see figure 4). Finally, DCA-GC-QH requires the highest value of N_r; this is the price that has to be paid in order to attain the best blocking performance with respect to both

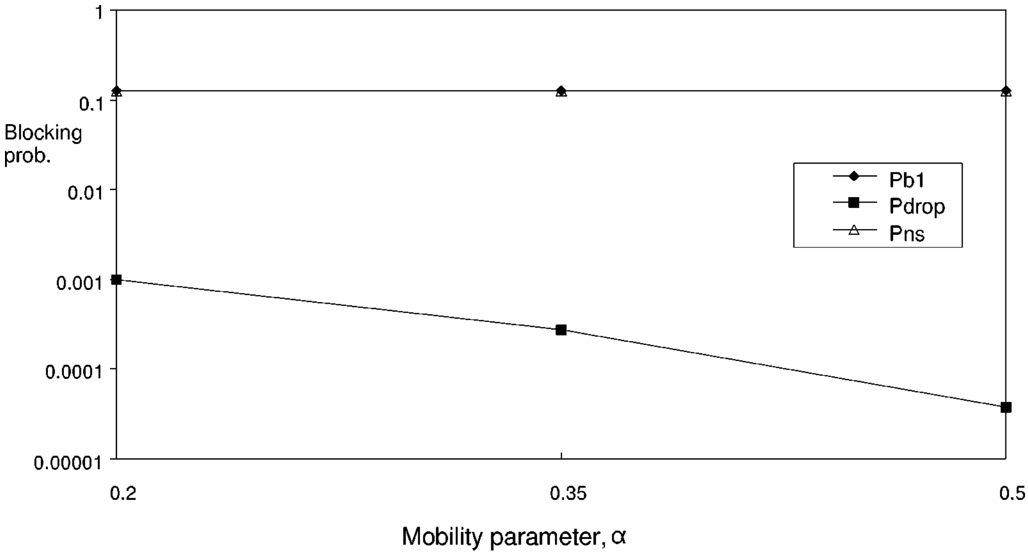

Figure 7. Behavior of P_{b1}, P_{drop} and P_{ns} for the DCA-GC-QH technique as a function of the mobility parameter α ($M = 70$, $S = 0.2$, 8 erl/cell).

FCA-GC-QH and FCA&DCA-GC-QH. The high number of channel rearrangements with DCA-GC-QH is due to the fact that they are employed not only to reduce the use of guard channels to just a minimum, but also to follow a compact allocation pattern.

The impact of various mobility conditions on the performance of the DCA-GC-QH technique (with 7 guard channels within the pool of $M = 70$ system channels and $S = 0.2$) has been evaluated in terms of P_{b1}, P_{drop} and P_{ns} in figure 7; in this case, parameter α ranges from 0.2 (microcellular systems) to 0.5 (macrocellular systems). When mobility increases (i.e., α decreases), P_{b1} and P_{b2} reduce, because the average channel holding time in a cell diminishes. The slight variation of P_{b1} is due its very high value which denotes congestion for new call attempts. This behavior is also emphasized by the adopted handoff prioritization scheme (i.e., guard channels and queueing). Whereas, P_{drop} increases with mobility, since P_{drop} can be roughly considered as the product between the average number of handoffs per call and P_{b2} [5,6]. Since there is a difference of two orders of magnitude between P_{drop} and P_{b1}, the behavior of P_{ns} is analogous to that of P_{b1}.

Moreover, the dependence of the DCA-GC-QH performance on the degree of overlap among adjacent cells, S, has been shown in figure 8. As expected, the call dropping probability significantly decreases as the overlap degree increases. Since the system is saturated for new call attempts (i.e., P_{b1} is quite high), we have practically a quasi-constant behavior of P_{b1}. Note that the case $S = 0$ entails no handoff queueing, because there is no overlap among adjacent cells: this is the worst case.

Figures 9 and 10 present the behavior of FCA-GC-QH ($S = 0.2$, $\alpha = 0.3$, 8 erl/cell) as a function of the number of nominal channels per cell, $Q - H$, where the number of guard channels per cell, H, is a parameter. In these graphs, the dashed lines

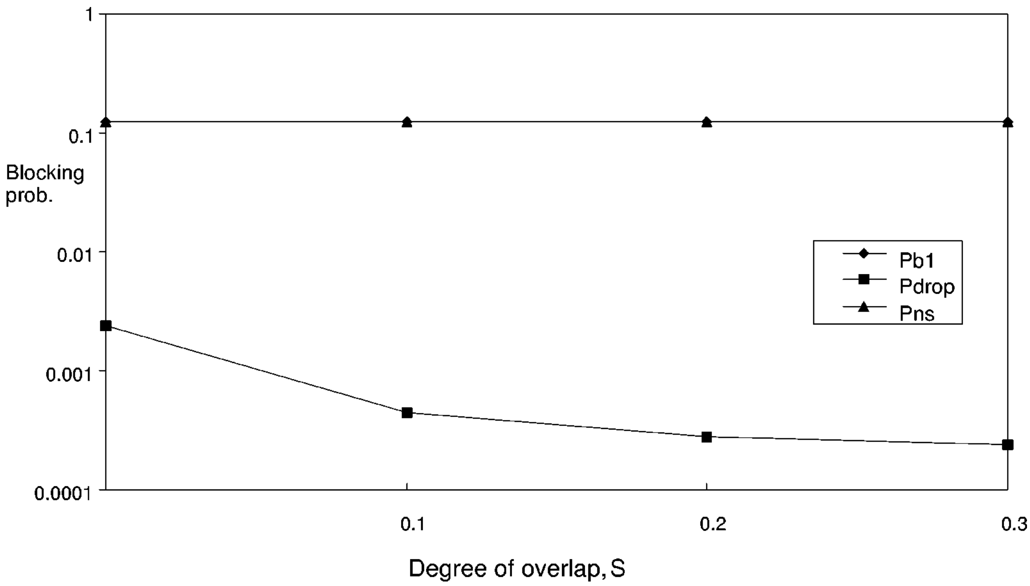

Figure 8. Behavior of P_{b1}, P_{drop} and P_{ns} for the DCA-GC-QH technique as a function of the overlap degree S ($M = 70,\ \alpha = 0.3$, 8 erl/cell).

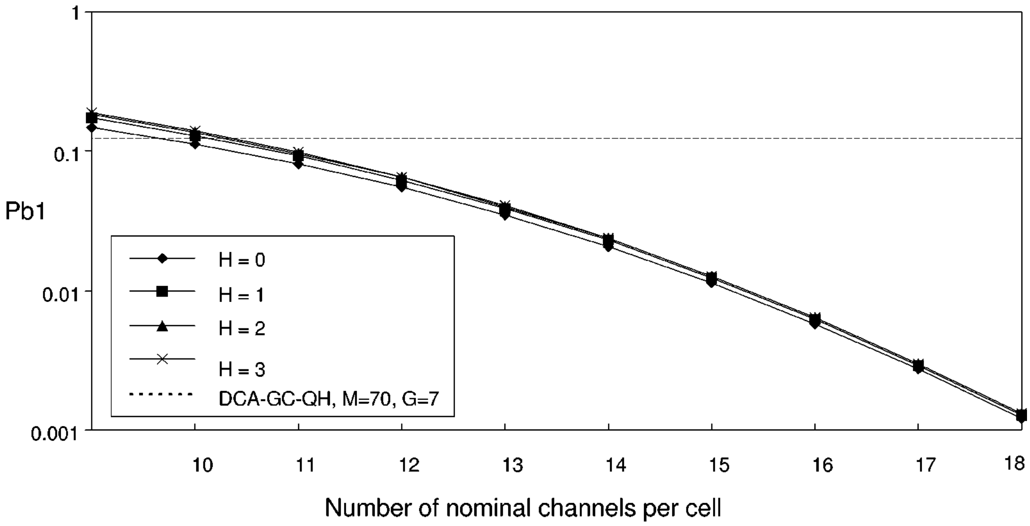

Figure 9. Probability P_{b1} as a function of the number of nominal channels per cell for FCA-GC-QH for different values of the number of guard channels per cell ($S = 0.2,\ \alpha = 0.3$, 8 erl/cell). The dotted line represents the performance for DCA-GC-QH with $M = 70$ and 7 guard channels.

represent the corresponding performance of DCA-GC-QH with 70 system channels within which 7 guard channels are considered. We can note that FCA-GC-QH obtains the same performance of DCA-GC-QH with the following combinations of values of

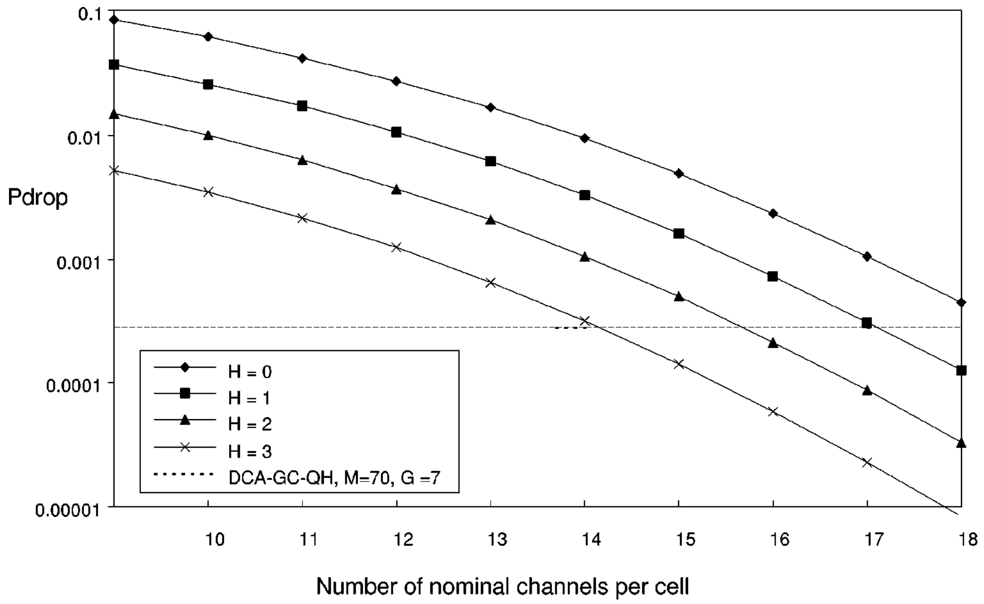

Figure 10. Probability P_{drop} as a function of the number of nominal channels per cell for FCA-GC-QH for different values of the number of guard channels per cell ($S = 0.2$, $\alpha = 0.3$, 8 erl/cell). The dotted line represents the performance for DCA-GC-QH with $M = 70$ and 7 guard channels.

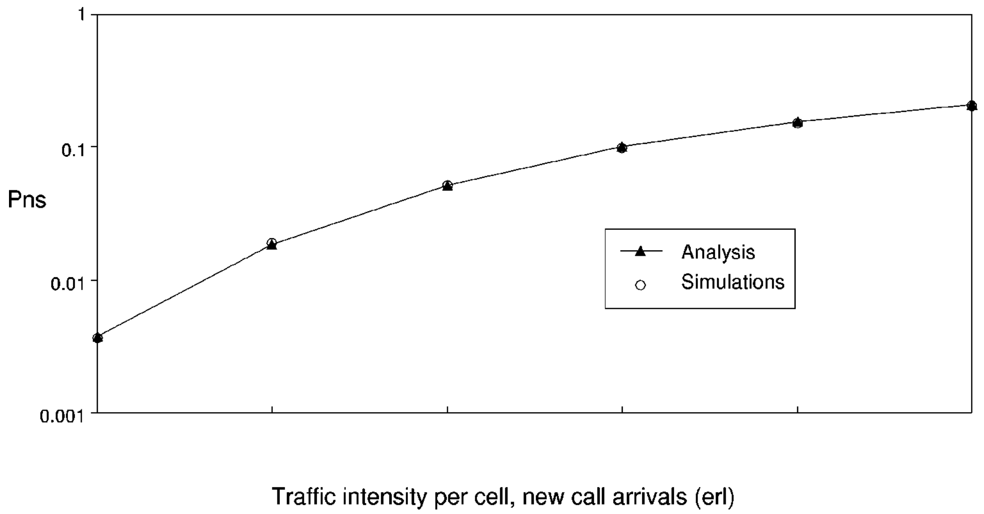

Figure 11. Comparison between theoretical and analytical results for FCA-GC-QH ($M = 70$, $S = 0.2$, $\alpha = 0.3$, 10 channels assigned per cell within which there is 1 guard channel).

$Q - H$ and H: 17/1, 16/2, 14/3; these configurations require a different total number of channels, according to the reuse factor $K = 7$: 126 channels in the cases 17/1 and 16/2, 119 channels in the case 14/3. Since the DCA-GC-QH technique achieves the same results with a total number of 70 channels, we have a significant advantage in sharing by DCA both nominal channels and guard ones among all the cells of the network.

Finally, figure 11 compares simulation and analytical results in terms of P_{ns} for the FCA-GC-QH technique with 7 guard channels within the pool of $M = 70$ system channels, $S = 0.2$ and $\alpha = 0.3$; we can note that there is a very good agreement between simulation results and theoretical predictions.

8. Conclusions

Third generation mobile communication systems will require suitable solutions for both increasing the channel utilization and reducing the risk of handoff failure. We have considered microcells used in a urban environment: handoff requests are frequent during call lifetime. Therefore, we have assumed that a strong handoff prioritization is required as regards new call attempts, if ITU-T E.771 recommendation on blocking probabilities has to be met. The joint use of guard channels and handoff queueing has been presented as a viable solution for increasing network performance.

We have defined a new DCA technique which allocates channels on the basis of a channel ordering defined for each cell according to a regular pattern. System channels are divided in two sets shared by DCA among the cells: the set of nominal channels to serve both new call attempts and handoff requests and the set of guard channels exclusively used to serve handoff requests. A channel rearrangement strategy has been also considered to pack as much as possible the use of channels and to reduce the use of guard channels. Moreover, a FIFO queueing policy has been assumed for handoff requests which can not be immediately served. This resource management strategy, called DCA-GC-QH, has been compared with other allocation techniques that still use guard channels and handoff queueing, but that are different from DCA-GC-QH on the basis of the sharing methods for both nominal channels and guard ones; when both channel sets are shared among cells by FCA, we have the FCA-GC-QH strategy and when nominal channels are assigned by FCA and guard ones are managed according to DCA, we have the FCA&DCA-GC-QH technique.

It has been shown that the proposed DCA-GC-QH technique permits to achieve a better quality of service than the FCA-GC-QH technique and the FCA&DCA-GC-QH strategy. The price that has to be paid for the best performance of DCA-GC-QH is its high average number of channel rearrangements per call accepted into the network. We have demonstrated that DCA-GC-QH attains the same performance of FCA-GC-QH with a significantly reduced number of system channels. Finally, a performance analysis for FCA-GC-QH has been carried out and it has been validated by a good agreement with simulation results.

Acknowledgements

The authors wish to thank Ing. Massimo Pierattini for interesting discussions and the anonymous reviewers for providing useful comments that have enhanced the technical quality of this paper.

References

[1] F. Baiocchi, F. Delli Priscoli, F. Grilli and F. Sestini, The geometric dynamic channel allocation strategy for high traffic FDM/TDMA mobile communication networks, in: *Internat. Zurich Seminar on Digital Communications*, Zurich, Switzerland (March 1994).

[2] E. Berruto, G. Colombo, P. Monogioudis, A. Napolitano and K. Sabatakis, Architectural aspects for the evolution of mobile communications towards UMTS, IEEE Journal on Selected Areas in Communications 15(8) (1997) 1477–1487.

[3] D.C. Cox, Wireless personal communications: What is it?, IEEE Personal Communications (April 1995) 20–35.

[4] E. Del Re, A coordinated European effort for the definition of a satellite integrated environment for future mobile communications, IEEE Communications Magazine 34(2) (1996) 98–104.

[5] E. Del Re, R. Fantacci and G. Giambene, Efficient dynamic channel allocation techniques with handover queueing for mobile satellite networks, IEEE Journal on Selected Areas in Communications 13(2) (1995) 397–405.

[6] E. Del Re, R. Fantacci and G. Giambene, Handoff and dynamic channel allocation techniques in mobile cellular networks, IEEE Transactions on Vehicular Technology 14(2) (1995) 229–237.

[7] E. Del Re, R. Fantacci and G. Giambene, Radio resource management, Deliverable No. 11 of the SAINT Project (RACE II), R2117-FIU-DIE.DR.P211-B1 (September 1995).

[8] E. Del Re, R. Fantacci and G. Giambene, An efficient technique for dynamically allocating channels in satellite cellular networks, in: *Proc. of IEEE GLOBECOM '95*, Singapore (November 13–17, 1995) pp. 1624–1628.

[9] S.A. El-Dolil, W.-C. Wong and R. Steele, Teletraffic performance of highway microcells with overlay macrocells, IEEE Journal on Selected Areas in Communications 7(1) (1989) 71–78.

[10] D. Everitt and D. Manfield, Performance analysis of cellular mobile communication systems with dynamic channel assignment, IEEE Journal on Selected Areas in Communications 7(8) (1989) 1172–1180.

[11] Framework for Services Supported on Future Public Land Mobile Telecommunications Systems (FPLMTS), ITU-R M.816.

[12] Y. Fun Hu and R.E. Sheriff, Traffic model and characteristics, Deliverable No. 10 of the SAINT Project (RACE II), R2117-BRU-EEE.DR.P210-B1 (April 1995).

[13] Future Public Land Mobile Telecommunications Systems (FPLMTS), ITU-R M.687.

[14] Future Public Land Mobile Telecommunications Systems (FPLMTS) Network Architectures, ITU-R M.817.

[15] R. Guérin, Channel occupancy time distribution in cellular radio system, IEEE Transactions on Vehicular Technology 35(3) (1987) 89–99.

[16] R. Guérin, Queuing-blocking system with two arrival streams and guard channels, IEEE Transactions on Communications 36(2) (1988) 153–163.

[17] D. Hong and S.S. Rappaport, Traffic model and performance analysis for cellular mobile radio telephone systems with prioritized and nonprioritized handoff procedures, IEEE Transactions on Vehicular Technology 35(3) (1986) 77–92.

[18] S.S. Lavemberg, *Computer Performance Modeling Handbook* (Academic Press, New York, 1993).

[19] Y.-B. Lin, L.-F. Chang and A. Noerpel, Modeling hierarchical microcell/macrocell PCS architecture, in: *IEEE ICC '95*, Seattle (June 1995) pp. 405–409.
[20] Y.-B. Lin, S. Mohan and A. Noerpel, Queueing channel assignment strategies for PCS hand-off and initial access, in: *Proc. of ICUPC '94*, San Diego, CA (September 27–October 1, 1994).
[21] V.H. MacDonald, The cellular concept, Bell System Technical Journal 58 (January 1979) 15–41.
[22] K. Nakano, K. Kawamura, N. Karasawa, M. Semgoku and S. Shinoda, Traffic characteristics of dynamic channel assignment under non-uniform traffic distribution, in: *Proc. of the 47th IEEE Vehicular Technology Conf.*, Phoenix, USA (1997).
[23] Network grade of service parameters and target values for circuit-switched land mobile services, ITU-T E.771, Blue Book (1995).
[24] S.S. Rappaport, Blocking, hand-off and traffic performance for cellular communication systems with mixed platforms, IEE Proceedings-I 140(5) (1993) 389–401.
[25] P.A. Raymond, Performance analysis of cellular networks, IEEE Transactions on Communications 39(12) (1991) 1787–1793.
[26] J. Restrepo and G. Maral, Coverage concepts for satellite constellations providing communications services to fixed and mobile users, Space Communications 13(2) (1995) 145–157.
[27] J. Restrepo and G. Maral, Providing appropriate grades of service to fixed and mobile users in a non-GEO satellite-fixed cell system, in: *Proc. of the 2nd European Workshop on Mobile/Personal Satcoms*, Rome, Italy (October 9–11, 1996) pp. 79–96.
[28] Satellite Operation within Future Public Land Mobile Telecommunications Systems (FPLMTS), ITU-R M.818.
[29] M. Scheibenbogen, S. Clausen and A. Guntsch, Dynamic channel allocation in hierarchical cellular systems, in: *Proc. of the 46th IEEE Vehicular Technology Conf.*, Atlanta, USA (1996).
[30] M. Schwartz, *Telecommunication Networks: Protocols, Modeling and Analysis* (Addison-Wesley, Reading, MA, 1987).
[31] S. Tekinay and B. Jabbari, Handoff and channel assignments in mobile cellular networks, IEEE Communications Magazine 29(11) (1991) 42–46.
[32] R.W. Wolff, Poisson arrivals see time averages, Operations Research 30 (1982) 223–231.